PART A

PROCEEDINGS OF THE FIFTEENTH

# MARCEL GROSSMANN MEETING ON GENERAL RELATIVITY

Editors

**Elia S. Battistelli**

**Robert T. Jantzen**

**Remo Ruffini**

World Scientific

# THE FIFTEENTH
# MARCEL GROSSMANN MEETING

**On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories**

This page intentionally left blank

PART A

# THE FIFTEENTH
# MARCEL GROSSMANN MEETING

## On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories

Proceedings of the MG15 Meeting
on General Relativity
University of Rome "La Sapienza", Italy 1–7 July 2018


Editors

**Elia S. Battistelli**
University of Rome "La Sapienza"
Rome, Italy

**Robert T. Jantzen**
Villanova University
Villanova, PA, USA

Series Editor

**Remo Ruffini**
International Center for Relativistic Astrophysics (ICRA)
University of Rome "La Sapienza"
Rome, Italy
International Center for Relativistic Astrophysics
Network (ICRANet)
Pescara, Italy



World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

*Published by*

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

*USA office:* 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

*UK office:* 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Control Number: 2022906140

**British Library Cataloguing-in-Publication Data**
A catalogue record for this book is available from the British Library.

**THE FIFTEENTH MARCEL GROSSMANN MEETING**
**On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories**
**(In 3 Volumes)**



ISBN 978-981-125-824-4 (set_hardcover)
ISBN 978-981-125-825-1 (set_ebook for institutions)
ISBN 978-981-125-826-8 (set_ebook for individuals)

ISBN 978-981-125-661-5 (vol. 1_hardcover)

ISBN 978-981-125-662-2 (vol. 2_hardcover)

ISBN 978-981-125-663-9 (vol. 3_hardcover)

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/12843#t=suppl

Desk Editor: Ng Kah Fee

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

**THE MARCEL GROSSMANN MEETINGS**

**Series Editor:** REMO RUFFINI

**Publications in the Series of Proceedings**

Proceedings of the Fifteenth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2018)
Edited by E.S. Battistelli, R.T. Jantzen, R. Ruffini
World Scientific, 2021

Proceedings of the Fourteenth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2015)
Edited by M. Bianchi, R.T. Jantzen, R. Ruffini
World Scientific, 2017

Proceedings of the Thirteenth Marcel Grossmann Meeting on General Relativity
(Stockholm, Sweden, 2012)
Edited by K. Rosquist, R.T. Jantzen, R. Ruffini
World Scientific, 2015

Proceedings of the Twelfth Marcel Grossmann Meeting on General Relativity
(Paris, France, 2009)
Edited by T. Damour, R.T. Jantzen, R. Ruffini
World Scientific, 2012

Proceedings of the Eleventh Marcel Grossmann Meeting on General Relativity
(Berlin, Germany, 2006)
Edited by H. Kleinert, R.T. Jantzen, R. Ruffini
World Scientific, 2007

Proceedings of the Tenth Marcel Grossmann Meeting on General Relativity
(Rio de Janiero, Brazil, 2003)
Edited by M. Novello, S. Perez-Bergliaffa, R. Ruffini
World Scientific, 2005

Proceedings of the Ninth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2000)
Edited by V.G. Gurzadyan, R.T. Jantzen, R. Ruffini
World Scientific, 2002

Proceedings of the Eighth Marcel Grossmann Meeting on General Relativity
(Jerusalem, Israel, 1997)
Edited by T. Piran
World Scientific, 1998

Proceedings of the Seventh Marcel Grossmann Meeting on General Relativity
(Stanford, USA, 1994)
Edited by R.T. Jantzen and G.M. Keiser
World Scientific, 1996

Proceedings of the Sixth Marcel Grossmann Meeting on General Relativity
(Kyoto, Japan, 1991)
Edited by H. Sato and T. Nakamura
World Scientific, 1992

Proceedings of the Fifth Marcel Grossmann Meeting on General Relativity
(Perth, Australia, 1988)
Edited by D.G. Blair and M.J. Buckingham
World Scientific, 1989

Proceedings of the Fourth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 1985)
Edited by R. Ruffini
World Scientific, 1986

Proceedings of the Third Marcel Grossmann Meeting on General Relativity
(Shanghai, People's Republic of China, 1982)
Edited by Hu Ning
Science Press – Beijing and North-Holland Publishing Company, 1983

Proceedings of the Second Marcel Grossmann Meeting on General Relativity
(Trieste, Italy, 1979)
Edited by R. Ruffini
North-Holland Publishing Company, 1982

Proceedings of the First Marcel Grossmann Meeting on General Relativity
(Trieste, Italy, 1975)
Edited by R. Ruffini
North-Holland Publishing Company, 1977

## SPONSORS

International Center for Relativistic Astrophysics Network (ICRANet)
International Center for Relativistic Astrophysics (ICRA)
International Centre for Theoretical Physics (ICTP)
Istituto Nazionale di Astrofisica (INAF)
International Union of Pure and Applied Physics (IUPAP)

## FREEDOM OF MOVEMENT FOR SCIENTISTS

The Marcel Grossman Meetings were founded with the premise that scientists of all nations have a right to meet to exchange knowledge independent of national borders. As such we affirm the IUPAP declaration: "To secure IUPAP sponsorship, the organisers have provided assurance that MG15 will be conducted in accordance with IUPAP principles as stated in the IUPAP resolution passed by the General Assembly in 2008. In particular, no bona fide scientist will be excluded from participation on the grounds of national origin, nationality, or political considerations unrelated to science."

## ACKNOWLEDGEMENTS

We acknowledge the outstanding job done before, during and after the meeting by the ICRANet/ICRA administrative and secretarial staff: Cristina Adamo, Silvia Latorre, Elisabetta Natale, and Cinzia di Niccolo and the conference agency support of Sistema Congressi (Marisa Santori, Ilaria Coccato, Lorena Moretti). Finally this meeting and its proceedings could not have functioned without the dedicated IT support of the ICRA system manager Vittorio Vannini for every MG meeting this century and the ICRANet system manager Gabriele Brandolini, with some temporary assistance from Valerio Antonucci and Damiano Verzulli.

## ORGANIZING BODIES OF THE FIFTEENTH MARCEL GROSSMANN MEETING:

### INTERNATIONAL ORGANIZING COMMITTEE

David Blair, Yvonne Choquet-Bruhat, Thibault Damour, Paolo De Bernardis, Francis Everitt, Riccardo Giacconi, Theodor Haensch, Stephen Hawking, Marc Henneaux, Christine Jones Forman, Roy Kerr, Hagen Kleinert, Jutta Kunz, Claus Laemmerzahl, Malcolm Longair, Tsvi Piran, Jorge Rueda, Remo Ruffini (Chair), Misao Sasaki, Humitaka Sato, Rashid Sunayev, Gerard 't Hooft, Steven Weinberg, Shing-Tung Yau

### LOCAL ORGANIZING COMMITTEE

Amati, L., Angelantonj, C., Barbiellini, G., Bassan, M., Battistelli, E. S. (chair), Belinski, V.A., Belli, L., Benedetti, R., Bernabei, R., Bianchi, M., Bianco, C.L. (co-chair), Bini, D., Buchert, T., Burgio, F., Capozziello, S., Chakrabarti, S.K., Chardonnet, P., Dall'Agata, G., De Angelis, A., De Bernardis, P., Della Valle, M., Di Virgilio, A., Frasca S., Fraschetti, F., Fré, P., Frontera, F., Giavalisco, M., Giommi, P., Gionti, G., Ingrosso, G., Jantzen, R.T., Jetzer, P., Lee, H.-W., Lerda, A., Liberati, S., Longo, R., Luzzi, G., Mandolesi, N., Marmo, G., Menotti, P., Merafina, M., Morselli, A., Pani, P., Pelster, A., Piacentini, F., Pian, E., Procesi, C., Quevedo, H., Ricci, F., Riccioni, F., Rosati, P., Ruffo, S., Scarpetta, E.V., Tavani, M., Titarchuk, L.G., Vereshchagin, G.V., Xue, S.-S., Zen Vasconcellos, C.

### INTERNATIONAL COORDINATING COMMITTEE

ALBANIA: Hafizi, M. ARGENTINA: Arguelles, C., Ghezzi, C.R., Mirabel, F., Romero, G.E. ARMENIA: Sahakyan, N. AUSTRALIA: Blair, D., Ju, L., Lun, A., Manchester, D., Melatos, A., Quinn, P., Scott, S.M., Steele, J.D., AUSTRIA: Aichelburg, P.C., Schindler, S. BELARUS: Kilin, S., Minkevich, A.V., Siutsou, I.A. BELGIUM: Henneaux, M. BOLIVIA: Aguirre, C.B. BOSNIA AND HERZEGOVINA: Pasic, V. BRAZIL: Aguiar, O., Barres de Almeida, U., Berkovits, N.J., Carneiro da Cunha, B., de Gouveia Dal Pino, E.M., de Souza Oliveira Filho, K., Guzzo, M., Hadjimichef, D., Klippert, R., Malheiro, M., Negreiros, R., Peres Menezes, D., Romero Filho, C.A, Shellard, R., Villela, T., Wuensche, C.A., Zen Vasconcellos, C. BULGARIA: Yazadjiev, S. CANADA: Singh, D., Smolin, L., Turok, N. CHILE: Bunster Weitzman, C., Giacomini, A., Reisenegger, A. CHINA: Cai, R.-G., Cao, Z., Chang, J., Chen, J.-S., Chen, X.-L, Dai, Z.-G, Feng, L.-L., Han, W.-B., Jing, Y.-P., Li, T.-P., Lin, W.-B., Lou, Y.-Q., Luo, J., Mei, J.-W., Wang, Y., Wu, X.-P., Wu, Y.-L., Yuan, Y.-F., Zhang, B., Zhang, S.-N., Zhao, G. CHINA (TAIWAN): Chen, C.-M., Chen, P., Lee, D.-S., Lee, W.-L., Ni, W.-T. COLOMBIA: Bargueño de Retes, P., Gonzalez, G., Granda Velasquez, L.N., Núñez, L., Pachón Contreras, L.A., Portilla, B., Valenzuela Toledo, C.A. CROATIA: Dominis Prester,

D., Milekovic, M., Smolčič, V., Smolić, I., Surić, T. CUBA: Pérez Martínez, A., Pérez Rojas, H., Quiros, I. CZECH REPUBLIC:Bicak, J. DENMARK: Naselsky, P. EGYPT: Tawfik, A.N., Wanas, M.I. ESTONIA: Einasto , J., Saar, E. FINLAND: Poutanen, J., Volovik, G. FRANCE: Brillet, A., Buchert, T., Chardonnet, P., Coullet, P., de Freitas Pacheco, J.A., Deruelle, N., Iliopoulos, J., Mignard, F., GEORGIA: Lavrelashvili, G., Machabeli, G. GERMANY: Biermann, P., Blumlein, J., Di Piazza, A., Fritzsch, H., Genzel, R., Gilmozzi, R., Hehl, F., Keitel, C., Kiefer, C., Mirzoyan, R., Neugebauer, G., Nicolai, H., Renn, J., Ringwald, A., Ruediger, A. GREECE: Batakis, N.A., Cotsakis, S., Vagenas, E.C. HUNGARY: Fodor, G., Levai, P. ICELAND: Bjornsson, G., Jakobsson, P. INDIA: Chakrabarti, S.K.,Iyer, B., Padmanabhan, T., Souradeep, T. IRAN: Mansouri, R., Mashhoon, B., Sobouti, Y. IRELAND: O'Murchada, N., Pe'er, A. ISRAEL: Milgrom, M., Nakar, E., Piran, T. ITALY: Belinski, V.A., Bianchi, M., Bianco, C.L., Cherubini, C., Della Valle, M., Falciano, S., Filippi, S., Menotti, P., Merafina, M., Pani, P., Ricci, F., Treves, A., Vereshchagin, G.V., Vitale, S., Xue, S.-S. JAPAN: Fujimoto, M.-K., Makishima, K., Nakamura, T., Sato, K., Shibata, M. KAZAKHSTAN: Abishev, M., Aimuratov, Y., Boshkayev, K., Mychelkin, E.G., Spitaleri, C. KOREA (PYEONGYANG): Kim, J.S. KOREA (SEOUL): Kim, S.P., Kim, S.-W., Lee, H.K., Lee, H.-W., van Putten, M.KYRGYZSTAN: Gurovich, V.Ts. LIBYA: Gadri, M. MEXICO: Breton, N., García-Diaz, A.A., Macías Alvarez, A., Mielke, E.W., Quevedo, H., Rodriguez, L.F. NEW ZEALAND: Visser, M., Wiltshire, D. NORWAY: Elgarøy, Ø., Fonseca Mota, D., Knutsen, H. POLAND: Belczynski, K., Demianski, M., Lewandowski, J., Nurowski, P., Sokolowski, L. PORTUGAL: Costa, M., da Silva, A., Lemos, J.P.S., Lobo, F., Moniz, P., Silva, L.O. ROMANIA: Visinescu, M. RUSSIA: Aksenov, A.G., Arkhangelskaja, I., Bisnovatyi-Kogan, G.S., Blinnikov, S.I., Chechetikin, V.M., Cherepaschuk, A.M., Khriplovich, I.B., Lipunov, V.M., Lukash, V.N., Novikov, I.D., Rudenko, V.N., Starobinsky, A.A. SERBIA: Djordjevic, G., Jovanovic, P., Knežević, Z., Pankov-Hzvojevic, M., Popovic, L., Prodanovic, T., M., Sijacki, D., Simic, S. SLOVAKIA: Balek, V. SLOVENIA: Cadez, A., Gomboc, A., Zavrtanik, D. SOUTH AFRICA: Colafrancesco, S., Larena, J., Maharaj, S., SPAIN: Elizalde, E., Ibanez, J., Perez Mercader, J., Verdaguer, E. SWEDEN: Abramowicz, M.A., Marklund, M., Ryde, F. SWITZERLAND: Durrer, R., Jetzer, P. TURKEY: Aliev, A., Gurses, M. UKRAINE: Novosyadlyj, B., Zaslavski, O.B., Zhuk, A. UK: Cruise, A.M., Frenk, C.S., Green, M., Mavromatos, N., Perry, M., Willingale, R. USA: Abel, T., Ashtekar, A., Bardeen, J.,Carlstrom, J., Cornish, N., Dermer, C., Fan, X., Flanagan, E., Fraschetti, F., Fryer, C., Incera, V., Jantzen, R.T. (Chairperson), Kolb, R., Laguna, P., Longair, M., Lousto, C., Madau, P., Mathews, G., Matzner, R., Melia, F., Mester, J., Michelson, P., Nordtvedt, K., Parker, L., Pretorius, F., Pullin, J., Shapiro, I., Shapiro, S., Shoemaker, D., Smoot, G., Stiavelli, M., Teukolsky, S., van Nieuwenhuizen, P., Zhang, B. UZBEKISTAN: Ahmedov, B., Zalaletdinov, R.M. VATICAN CITY: Gionti, G. VENEZUELA: Herrera, L. VIETNAM: Long, H.N.

## MARCEL GROSSMANN AWARDS

## FIFTEENTH MARCEL GROSSMANN MEETING

### Institutional Award

Planck Scientific Collaboration (ESA) (ESA)

*"for obtaining important constraints on the models of inflationary stage of the Universe and level of primordial non-Gaussianity; measuring with unprecedented sensitivity gravitational lensing of Cosmic Microwave Background fluctuations by large-scale structure of the Universe and corresponding Bpolarization of CMB, the imprint on the CMB of hot gas in galaxy clusters; getting unique information about the time of reionization of our Universe and distribution and properties of the dust and magnetic fields in our Galaxy."*
— presented to its Director General Johann-Dietrich Woerner

### Institutional Award

Hansen Experimental Physics Laboratory (HEPL) at Stanford University

*"for having developed interdepartmental activities at Stanford University at the frontier of fundamental physics, astrophysics and technology."*

### Individual Awards

Lyman Page
*"for his collaboration with David Wilkinson in realizing the NASA Explorer WMAP mission and who now leads the Atacama Cosmology Telescope as its project scientist."*

Rashid Alievich Sunyaev
*"for the development of theoretical tools in the scrutinising, through the CMB, of the first observable electromagnetic appearance of our Universe."*

Shing-Tung Yau
*"for the proof of the positivity of total mass in the theory of general relativity and perfecting as well the concept of quasi-local mass, for his proof of the Calabi conjecture, for his continuous inspiring role in the study of black holes physics ."*

Each recipient is presented with a silver casting of the TEST sculpture by the artist A. Pierelli. The original casting was presented to His Holiness Pope John Paul II on the first occasion of the Marcel Grossmann Awards.

**FOURTEENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

European Space Agency (ESA)

**Individual Awards**

Frank Chen Ny Yang, Tsung Dao Lee, Ken'ichi Nomoto, Martin Rees, Yakov G. Sinai, Sachiko Tsuruta

**THIRTEENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

ALBANOVA UNIVERSITY CENTER, STOCKHOLM

**Individual Awards**

David Arnett, Vladimir Belinski and Isaak M. Khalatnikov, Filippo Frontera

**TWELFTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Institut des Hautes Études Scientifique (IHES)

**Individual Awards**

Jaan Einasto, Christine Jones, Michael Kramer

**ELEVENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Freie Universität Berlin

**Individual Awards**

Roy Kerr, George Coyne, Joachim Trümper

**TENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

CBPF (Brazilian Center for Research in Physics)

**Individual Awards**

Yvonne Choquet-Bruhat, James W. York, Jr., Yval Ne'eman

**NINTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Solvay Institutes

**Individual Awards**

Riccardo Giacconi, Roger Penrose, Cecile and Bryce DeWitt

**EIGHTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Hebrew University of Jerusalem

**Individual Awards**

Tullio Regge, Francis Everitt

**SEVENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Hubble Space Telescope Institute

**Individual Awards**

SUBRAHMANYAN CHANDRASEKHAR, JIM WILSON

**SIXTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Research Institute for Theoretical Physics (Hiroshima)

**Individual Awards**

Minora Oda, Stephen Hawking

**FIFTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The University of Western Australia

**Individual Awards**

Satio Hayakawa, John Archibald Wheeler

**FOURTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Vatican Observatory

**Individual Awards**

William Fairbank, Abdus Salam

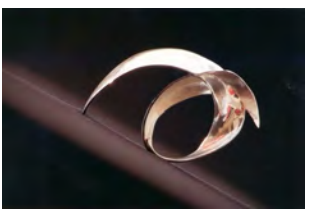


Fig. 1. TEST: sculpture by Attilio Pierelli

**PLANCK SCIENTIFIC COLLABORATION (ESA)**
presented to **Jean-Loup Puget**, the Principal Investigator of the High Frequency Instrument (HFI).

*"for obtaining important constraints on the models of inflationary stage of the Universe and level of primordial non-Gaussianity; measuring with unprecedented sensitivity gravitational lensing of Cosmic Microwave Background fluctuations by large-scale structure of the Universe and corresponding B-polarization of CMB, the imprint on the CMB of hot gas in galaxy clusters; getting unique information about the time of reionization of our Universe and distribution and properties of the dust and magnetic fields in our Galaxy"*

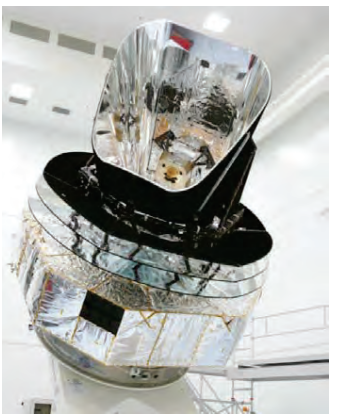
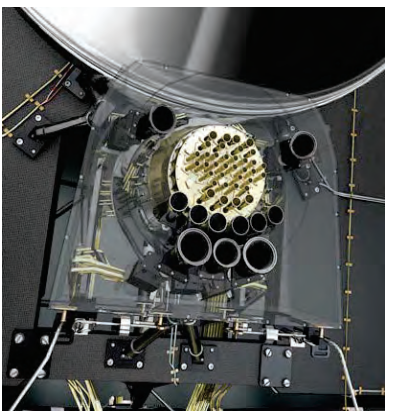


*Photo of the Planck satellite (Courtesy of ESA).*



*Planck focal plane (Courtesy of ESA)*

Planck ESA's mission, was designed to image the temperature and polarization anisotropies of the CMB over the whole sky, with unrivalled angular resolution and sensitivity, pushing the technology to unprecedent limits. In the framework of the highly precision experimental cosmology the legacy Planck results on testing theories of the early universe and the origin of cosmic structure, has provided a major source of information crucial to many cosmological and astrophysical issues. Planck carried out two instruments:

- the High Frequency Instrument (HFI), Principal Investigator: Jean Loup Puget;
- the Low Frequency Instrument (LFI), Principal Investigator: Nazzareno Mandolesi.

The instruments were complementary and using different technology to cross check independently final results and systematic errors. They worked together to produce the overall mission results. The Planck space mission (ESA) has been a wonderful example of Team effort in a large international collaboration, involving scientific, technical and managerial aspects. The unprecedented accuracy of the Planck measurements have established new standards in the determination of fundamental cosmological parameters, as well as new insight in Galactic and extragalactic astrophysics. The Planck full-sky maps in temperature and polarization will remain a lasting legacy for at least dozen years to come. More than 100 papers signed by Planck collaboration have already 30 000 citations in scientific literature. The success of Planck HFI and LFI would not have been possible without the contribution of a large number of talented and dedicated scientists and engineers from many countries of Europe, USA and Canada. HFI was designed to produce high-sensitivity, multi-frequency measurements of the diffuse radiation permeating the sky in all directions in the frequency range of 84 GHz to 1 THz cooled at 100 mK. The instrument consisted of an array of 52 bolometric detectors placed in the focal plane of the telescope. LFI, a microwave instrument, was designed to produce high-sensitivity, multi-frequency measurements of the microwave sky in the frequency range of 27 to 77 GHz. The instrument consisted of an array of 22 tuned radio receivers located in the focal plane of the telescope, cooled at 20 K.

*Jean Loup Puget - PI of the HFI.*

*Nazzareno Mandolesi - PI of the LFI.*

## HANSEN EXPERIMENTAL PHYSICS LABORATORY AT STANFORD UNIVERSITY

presented to **Leo Hollberg**, HEPL Assistant Director

*"to HEPL for having developed interdepartmental activities at Stanford University at the frontier of fundamental physics, astrophysics and technology"*

### Brief History of Stanford's HEPL and Ginzton Laboratories

1947: WW Hansen (right) and team with the Mark I linear accelerator

In 1947, working in the Stanford Physics Department's Microwave Lab, Physics Professor, William W. Hansen and his research team, along with Assistant Professor of Physics and microwave expert, Edward L. Ginzton, completed development on the world's first traveling wave electron linear accelerator. Dubbed the Mark I (see photo) it generated a 1.5 million electron volt (MeV) beam. Hansen's entire report to the U.S. Office of Naval Research (ONR) that funded the project was just one sentence: "We have accelerated electrons."

This successful first step in linear electron acceleration spawned the birth of Stanford's High Energy Physics Lab (HEPL) and Ginzton labs. In 1990, HEPL was renamed the WW Hansen Experimental Physics Lab (also HEPL). HEPL and Ginzton were setup as Stanford's first independent labs. They were organized to facilitate cross-disciplinary research, enabling scientists, engineers, staff and students to work towards common research goals using cutting edge lab equipment and technologies on medium-scale projects. For the past 70 years, the HEPL and Ginzton Labs have spearheaded Stanford's leadership in cross-disciplinary physics and become nurturing homes to a variety of physics-based, research projects: including the following examples:

### Robert Hofstadter's Nobel Prize & Later Work

In 1961, Stanford Professor Robert Hofstadter was awarded the Nobel Prize for his HEPL Mark III Linear Accelerator work on nuclear form factors (nucleons). In the 1980s, Hofstadter became interested in astrophysics and helped design the EGRET telescope in the NASA Compton Gamma Ray Observatory (CGRO).

### Gravity Probe B (GP-B)

2003: Francis Everitt (left) and Brad Parkinson holding a GP-B gyro

In 1959, Physics Department Chair, Leonard Schiff, became interested in using gyroscopes in a satellite to measure the Earth's geodetic effect and the miniscule frame-dragging effect predicted by Albert Einstein's general theory of relativity. Schiff discussed this project with Stanford cryogenic physicist, William Fairbank, and gyroscope expert, Robert Cannon (Aero-Astro department).

In 1962, Fairbank invited post-doc, Francis Everitt, to join the research effort. The team sent a proposal to NASA's Office of Space Sciences requesting funding to develop gyroscopes and a satellite to carry out this unprecedented test. It took 40 years of R&D at Stanford and other places to create and ready the cryogenic satellite and all of its cutting-edge technologies for launch. In 1975, Leonard Schiff moved the GP-B program to HEPL, breathing new life into the lab. In 1981, Francis Everitt became Principal Investigator, a position he still holds. In 1984, Brad Parkinson became Project Manager and a Co-PI, along with Co-PI's John Turneaure and Daniel DeBra.

On 20 April 2004, GP-B launched from Vandenberg AFB into a polar orbit. Data collection began on 28 August 2004 and lasted 50 weeks. Data analysis took five years in order to remove confounding factors in the data. The final results, published in PRL on 31 May 2011, yielded highly accurate geodetic and frame dragging measurements, with 0.28% and 19% margins of error, respectively.

### GPS Spinoffs from GP-B

In the 1990s, Brad Parkinson's research on GPS solutions for positioning the GP-B satellite led to two revolutionary spin-off projects: 1) Wide Area Augmentation System (WAAS) provides highly precise positioning accuracy and itegrity for navigation and the automatic landing of airplanes and 2) Precision Farming adding GPS technology to tractors has enabled the automation of many aspects of farming and has spawned a $1 billion/year industry.

### Fermi Gamma Ray Space Telescope (GLAST)

Peter Michelson, Fermi LAT Principal Investigator

Stanford Physics Professor, Peter Michelson, is a former HEPL Director and the Principal Investigator for the Large Area Telescope (LAT) on board NASA's Fermi Gamma Ray Space Telescope, the successor to CGRO/EGRET. Launched on 11 June 2008, Fermi has been highly successful mapping the gamma-ray sky. Under Michelson's guidance, HEPL's collaborations with Italy are noteworthy. The development of cryogenic bar detectors of gravitational waves, in collaboration with Edoardo Amaldi and his colleagues, established new stringent upper limits to the gravitational waves incident on the Earth. Likewise, the Fermi LAT was developed by a collaboration between Italian INFN and ASI, NASA, and international partners in France, Japan, and Sweden, and used tracking detectors developed, integrated, tested, and qualified for the mission by Italy. GP-B provided the first evidence of frame-dragging on a spinning, superconducting gyroscope. The Fermi detector offers the potential of seeing, through the GeV emission in the Binary Driven Hypernova subclass of long GRBs, the emission from a newly born Black Hole, originating in the induced gravitational collapse of a supernova hypercritically accreting on a binary neutron star companion.

### Robert Byer's LIGO and ACHIP Projects

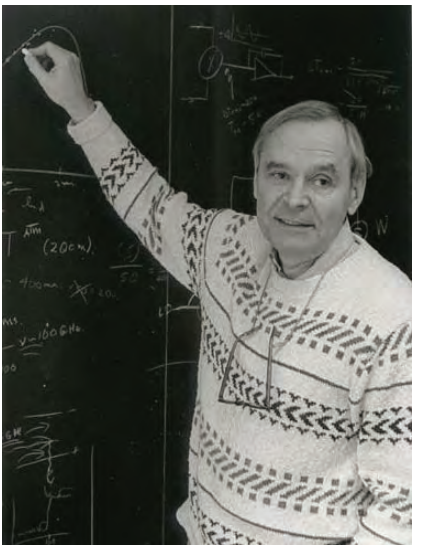
Professor Robert Byer

Robert Byer, former Stanford Dean of Research and former HEPL Director, nurtured the GP-B, GPS and Fermi programs to success during his tenure. He is currently an Applied Physics Professor specializing in lasers and optics. His LIGO Group provided seismic isolation, coatings and materials for the LIGO observatories. His ACHIP project is developing a particle accelerator on a microchip—bringing the HEPL/Ginzton Labs full circle to Hansen's 1947 research.

Professor **LYMAN PAGE**

*"for his collaboration with David Wilkinson in realizing the NASA Explorer WMAP mission and as founding director of the Atacama Cosmology Telescope"*

*Lyman Page*



*David Wilkinson*

This award is given in recognition of Lyman Page's pivotal role in transforming cosmic microwave background observations into a high-precision experimental science over the past two decades. In particular Page provided major contributions to the success of the Wilkinson Microwave Anisotropy Probe (WMAP) space mission, which delivered outstanding measurements of the CMB anisotropy and polarization pattern. He is now continuing his effort by promoting a new generation of experiments like the Atacama Cosmology Telescope to study CMB polarization to greater precision.

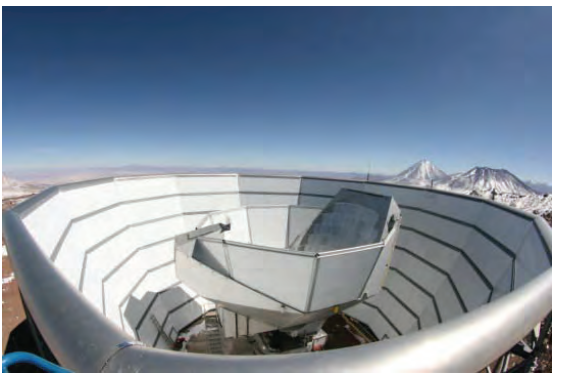


*Photo of the Atacama Cosmology Telescope*

The CMB, the faint afterglow of the Big Bang, is the most powerful probe of the early universe. From its study, we have learned the age of the universe, its major constituents, and have characterized the fundamental fluctuations in gravity that gave rise to cosmic structure. The desire to measure the CMB ever more precisely has driven the development of extraordinary detectors and techniques which will be reviewed in the Lectio Magistralis by Lyman Page. He will describe what we might hope to learn from the CMB in the next decade, including detecting gravitational waves from the birth of the universe if they exist at sufficient amplitude.

Professor **RASHID ALIEVICH SUNYAEV**
*"for the development of theoretical tools in the scrutinising, through the CMB, of the first observable electromagnetic appearance of our Universe"*

*Rashid Sunyaev*

*Rashid Sunyaev and Yakov Borisovich Zeldovich*

Rashid Sunyaev gave extraordinary contributions to the understanding of physical processes in the universe which identified new and uniquely informative targets for observational cosmology. In particular, the Sunyaev-Zeldovich effect, now observed in thousands of clusters of galaxies over the entire sky, has become a cornerstone of cosmology and extragalactic astrophysics, so much so that it is now considered a research field in its own right. Furthermore, Sunyaev's studies of processes in the early universe responsible for angular anisotropy and frequency distortions of the cosmic microwave background have left a profound and lasting legacy for cosmology. In particular, Sunyaev and Zeldovich predicted the presence of acoustic peaks in the CMB angular fluctuation power spectrum and the existence of baryonic acoustic oscillations.
He is currently the project scientist leading the scientific team of the international high-energy astrophysics observatory Spektr-RG being built under the direction of the Russian Space Research Institute.

*Yakov Zeldovich and Remo Ruffini at the audience with Pope John Paul II*

Professor **SHING-TUNG YAU**
*"for the proof of the positivity of total mass in the theory of general relativity and perfecting as well the concept of quasi-local mass, for his proof of the Calabi conjecture, for his continuous inspiring role in the study of black holes physics"*

*Shing-Tung Yau*

Shing-Tung Yau has made fundamental contributions to differential geometry which have influenced a wide range of scientific disciplines, including astronomy and theoretical physics. With Richard Schoen, Yau solved a longstanding question in general relativity by proving that the combined total energy of matter and gravitational field in an asymptotically flat universe is positive. In 1982 Yau was awarded the Fields Medal, the highest award in mathematics, and in 1994 he shared with Simon Donaldson the Crafoord Prize of the Royal Swedish Society in recognition of his development of nonlinear techniques in differential geometry leading to the solution of several outstanding problems.

Another outstanding achievement of Yau is his proof of the Calabi conjecture which allowed physicists to show that string theory is a viable candidate for a unified theory of nature. Furthermore in 2008 Yau (with M.T. Wang) introduced the concept of "quasi-local mass" in general relativity which can be of help to get around the old conundrum — the non-locality of the energy density in relativistic gravity.

During his scientific carrier Yau had more than 50 successful PhD students. At present he is a professor of mathematics at Harvard University where along with research he continues many pedagogical activities. For example, he has created the "Black Hole Initiative", an interdisciplinary center at Harvard University involving a collaboration between principal investigators from the fields of astronomy (Sheperd Doeleman, Avi Loeb and Ramesh Narayan), physics (Andrew Strominger), mathematics (Shing-Tung Yau) and philosophy (Peter Galison). This "Black Hole Initiative" is the first center worldwide to focus on the study of the many facets of black holes.

MG16 in 2021 will mark the $50^{th}$ anniversary of the mass-energy formula for black holes based on the Kerr metric. This timing is an omen that Yau and his school will soon further enlarge our knowledge of this formula with their powerful mathematical analysis.

# PREFACE

The Marcel Grossmann Meetings on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories have always had the goal of bringing together scientists from all over the world enabling them to share recent developments in general relativity and cosmology, paying attention to the interplay between physical predictions and mathematical foundations.

More than 800 scientists met at the University of Rome "la Sapienza" for the Fifteenth Marcel Grossmann Meeting (MG15) during the week of July 1–7, 2018 to discuss theoretical topics and the status of the experimental testing and observations of Einstein's theory of gravitation together with the broad spectrum of gravitational physics related phenomena. The topics discussed ranged from classical and quantum aspects of gravity, relativistic astrophysics, branes and strings, inflation theory, the thermal history of the Universe, to more concrete astrophysics experiments, observations, and modeling, reviewing the current state of the art in theory, observations, and experiments related to relativistic gravitation.

The meeting featured 39 plenary talks during the six mornings of the meeting, plus 6 public lectures given by experts in gravitation and cosmology. 73 parallel sessions, some of which were split over multiple days, kept participants busy with the crucial support of 33 students from Sapienza who managed the coordination of each parallel session and offered technical help.

Three scientists and two collaborations were presented with Marcel Grossmann awards. These were: Lyman Page, Rashid Alievich Sunyaev, Shing-Tung Yau, the Planck Scientific Collaboration (ESA) represented by Jean-Loup Puget, and the Hansen Experimental Physics Laboratory (HEPL) at Stanford University, represented by Leo Hollberg,

The detailed program of plenary and public talks is as follows:

**Shing-Tung Yau** (Harvard University): Quasi-local mass at null infinity
**Malcolm J. Perry** (University of Cambridge): Black hole entropy and soft hair
**Thomas Hertog** (KU Leuven): A smooth exit form eternal inflation
**Jean-Luc Lehners** (Max Planck Institute for Gravitational Physics): No smooth beginning for spacetime
**Ivan Agullo** (Louisiana State University): Loop quantum cosmology and the cosmic microwave background
**Elena Pian** (IASF Bologna):Kilonovae: the cosmic foundries of heavy elements
**Nial Tanvir** (University of Leicester): A new era of gravitational-wave / electromagnetic multi-messenger astronomy
**Tsvi Piran** (Hebrew University of Jerusalem): Mergers and GRBs: past present and future
**Stephan Rosswog** (Stockholm University): Neutron star mergers as heavy element production site

**David Shoemaker** (MIT LIGO Laboratory): LIGO's past and future observations of black hole and neutron star binaries
**Yu Wang** (ICRANet): On the role of binary systems in GW170817/ GRB170817A/AT2017gfo
**Hao Liu** (University of Copenhagen): An independent investigation of gravitational wave data
**Stefano Vitale** (University of Trento): Gravitation wave astronomy in ESA science programme
**Takaaki Kajita** (University of Tokyo): Status of KAGRA and its scientific goals
**Masaki Ando** (University of Tokyo): DECIGO : Gravitational-Wave observations from space
**Jun Luo** (Sun Yat-Sen University): TianQin: a space-borne gravitational wave detector
**Jo Van Den Brand** (Dutch National Institute for Subatomic Physics Nikhef, and VU University Amsterdam): Gravitational wave science and Virgo
**Ernst Maria Rasel** (Leibniz Universität Hannover): Science fiction turns into reality: Interferometry with Bose-Einstein condensates on ground and in space
**Manuel Rodrigues** (Université Paris Saclay): The first results of the MICROSCOPE test of the equivalence principle in space
**Victoria Kaspi** (McGill University): Fast radio bursts
**Bing Zhang** (University of Nevada): From gamma-ray bursts to fast radio bursts: unveiling the mystery of cosmic bursting sources
**Jean-Loup Puget** (CNRS): The Planck mission
**Jorge Armando Rueda Hernandez** (ICRANet): Binary-driven hypernovae and the understanding of gamma-ray bursts
**Remo Ruffini** (ICRANet): The essential role of the nature of the binary progenitors for understanding gamma ray bursts
**Heino Falcke** (Radboud University Nijmegen): Imaging black holes now and in the future
**Luc Blanchet** (Institut d'Astrophysique de Paris): Post-Newtonian theory and gravitational waves
**Jean-Loup Puget** (Université Paris): Frome Planck to Atacama Cosmology Telescope
**Razmik Mirzoyan** (Max Planck Institute for Physics): Gamma-ray and multimessenger highlights with MAGIC
**Elisa Resconi** (Technical University Munich): Neutrino astronomy in the multimessenger era
**Francis Halzen** (University of Wisconsin-Madison): IceCube: opening a new window on the universe from the South Pole
**James Lattimer** (Stony Brook University): The history of R-process
**Ralph Engel** (Karlsruhe Institute of Technology): What have we learned about ultra-high-energy cosmic rays from the Pierre Auger Observatory?

**Paolo De Bernardis** (Sapienza - University of Rome):
**Fabio Gargano** (INFN Bari): DAMPE and its latest results
**Markus Arndt** (University of Vienna): Experiments to probe quantum linearity at the interface to gravity & complexity
**Tobias Westphal** (University of Vienna): Micro-mechanical measurements of weak gravitational forces
**Shu Zhang** (Institute of High Energy Physics, Chinese Academy of Sciences):
**Lorenzo Amati** (INAF - OAS Bologna): Cosmology and multi-messenger astrophysics with Gamma-Ray Bursts
**Elisabetta Cavazzuti** (ASI): Gev LAT observations from GRBs and active galactic nuclei
**Remo Ruffini** (ICRANet): Concluding Remarks
**Roy Kerr** (University of Canterbury): Towards MG16

**Public Lectures**

**Jeremiah Ostriker** (Columbia University): Ultra-light scalars as cosmological dark matter
**Malcolm Longair** (University of Cambridge): Ryle and Hewish: 50 and 100 year anniversaries [Radio astrophysics and the rise of high energy astrophysics]
**Lyman Page** (Princeton University): Measuring the Cosmic Microwave Background
**Marc Henneaux** (Université Libre de Bruxelles): The cosmological singularity
**Anne Archibald** (Newcastle University): Does extreme gravity affect how objects fall?

# Concluding remarks by Prof. Remo Ruffini

For over more than four decades the Marcel Grossmann Meetings have been fostering the interaction between mathematics, relativistic field theories and observations in physics and astrophysics and has witnessed the birth and exponential growth of various new subfields within astrophysics. This has also occurred in 2018 during the MG15 (see Fig. 1), which has seen the participation of approximately one thousand participants from 71 different countries (see Fig. 2). In MG15, as is clear from the above program, we have seen considerable progress in theory on topics ranging from models of quantum gravity to the mathematical structure of Einstein's equations.


Fig. 1. MG15 official Awards Ceremony on Monday, July 2, 2018.

Fig. 2. Group photo of MG15 participants in the Aula Magna, Sapienza University of Rome.

In particular, an MG15 individual award was given to Prof. Rashid Sunyaev, as a member of the relativistic astrophysics school of Zel'dovich-Sunyaev (see Fig. 3) and to Prof. Shing-Tung Yau, as a representative of the Chinese-American school of differential geometry and general relativity (see Fig. 4).

Fig. 3. Rashid Sunyaev receiving the MG15 individual award from Roy Kerr.

Fig. 4. Shing-Tung Yau receiving the MG15 individual award from Roy Kerr.

This has been accompanied by extensive developments in the exploration of the cosmic microwave background with announcements of the final results of the Planck satellite that followed earlier predecessors, like WMAP. In this sense, particularly meaningful have been the MG15

Awards presented to Prof. Jean-Loup Puget, who has successfully accomplished the ESA Plank mission (see Fig. 5) and to Prof. Lyman Page, one of the main participants of the glorious NASA WMAP mission, successfully lead by David Wilkinson (see Fig. 6). I was very fortunate to be in Princeton as an assistant professor at the time WMAP was initially conceived.

Fig. 5. Jean-Loup Puget accepting the MG15 institutional award on behalf of the Planck Scientific Collaboration (ESA) from Roy Kerr.

Fig. 6. Lyman Page accepting the MG15 individual award from Roy Kerr.

X-ray astrophysics gradually successfully evolved from the first rockets of Riccardo Giacconi discovering Scorpius X-1 using well tested Geiger counters, moving on to larger missions using X-ray mirrors in the Einstein Observatory, in XMM and on to the NASA Chandra mission, each new mission introducing new technologies based on previous successes. The

Chandra data on GRB 170817A were amply discussed in the MG15 meeting. Analogously, starting from the earlier gamma ray detectors on the Vela satellite, gamma ray astrophysics reached maturity with the Compton Observatory with the BATSE instruments on board. These two fields joined their separate expertise together in the hybrid Beppo SAX satellite. Equally impressive was the transition from Beppo SAX to the Neil Gehrels Swift Observatory and the Fermi Gamma-Ray Space Telescope that followed EGRET and opened up the field of high energy gamma-ray astrophysics, adding further successes to this ongoing story. One of the most significant contributions to the success of the Fermi mission has been the LAT detector, jointly led by the Hansen Experimental Physics Laboratory (HEPL) at Stanford University, which received the MG15 institutional award, presented to Prof. Leo Hollberg (see Fig. 7).

Fig. 7. Leo Hollberg accepting the MG15 institutional award from Roy Kerr on behalf of the Hansen Experimental Physics Laboratory at Stanford University.

More recently we are witnessing the birth of TeV astrophysics springing from the ground-based MAGIC, HESS and Whipple telescopes, as well as neutrino astrophysics associated with underground (ICE Cube) and underwater laboratories. New tantalizing results concerning the spectrum and composition of cosmic rays have been reported by the Alpha Magnetic Spectrometer experiment on the international space station. All of these

together mark the gradual development of the largest observational effort in history, well recorded in these MG meetings (see Fig. 8). Today relativistic astrophysics is offering us the possibility of utilizing all of these multi-wavelength observations simultaneously in the study of GRBs and active galactic nuclei. Indeed the description of the distinct temporally discrete episodes for the GRB 130427A, one of the most complex astrophysical objects ever observed, requires the simultaneous knowledge of observations from all of these various wavelengths. This exponential growth in all these areas of astrophysics has been accompanied by attempts at data acquisition of gravitational wave signals arriving on the Earth. After the failure of the initial attempt by Misner and Weber in 1972, the new attempt by LIGO-VIRGO was discussed extensively in this MG15 meeting.

Fig. 8. From the left to right: Leo Hollberg, Rashid Sunyaev, Shing-Tung Yau and Remo Ruffini during the MG15 official awards ceremony.

Indeed in addition to the many topics discussed in the plenary lectures, there was a special session dedicated to kilonova and gravitational waves chaired by Enrico Costa on July 3 and a session on gravitational waves chaired by Claus Lämmerzahl on July 4. Many intense discussions took place at the meeting, both in the plenary and parallel sessions, bringing together different points of view, at times controversial, which stimulated a large number of articles published in leading scientific journals in the months following the meetings. The impossibility of finding these discussions in these proceedings is simply explained: it is due to the incomprehensible decision by Sapienza University not to renew the collaboration agreement with ICRANet. What is incomprehensible is that both the research and teaching activities of ICRA and ICRANet are

recognized internationally at a worldwide level, in more than 70 nations. Locks were put on the ICRA Offices (see Fig. 9).

Fig. 9. Remo Ruffini shows to Barry Barish the locked offices of ICRA at Sapienza University.

The preparation of the proceedings was halted. It was not possible to follow up with post-conference interactions and the impossibility of access to our offices and interact with our students. The invited lecturers, clearly indicated in the conference program and some of the world leaders in their fields who had just presented their extremely interesting lectures stimulating widespread interest, mostly chose to publish their results in international journals giving rise to some of the most highly cited articles in recent times. Only a minority were published here. This was also the case for some of the ICRA-ICRANet lectures, which appeared in:

- J.F. Rodriguez, J.A. Rueda, R. Ruffini, *On the final gravitational wave burst from binary black hole mergers*, Astron. Rep. 62, 940 (2018).
- J.A. Rueda, R. Ruffini, Y. Wang, Y. Aimuratov, U. Barres de Almeida, C.L. Bianco, Y.C. Chen, R.V. Lobato, C. Maia, D. Primorac, R. Moradia, and J.F. Rodriguez, *GRB 170817A-GW170817-AT 2017gfo and the observations of NS-NS, NS-WD and WD-WD mergers*, JCAP 10, 006 (2018); and J.A. Rueda, R. Ruffini, Y. Wang, C.L. Bianco, J.M. Blanco-Iglesias, M. Karlica, P. Lorén-Aguilar, R. Moradi and N. Sahakyan, *Electromagnetic emission of white dwarf binary mergers*, JCAP 03, 044 (2019).

- R. Ruffini, R. Moradi, J. A. Rueda, L. Becerra, C. L. Bianco, C. Cherubini, S. Filippi, Y. C. Chen, M. Karlica, N. Sahakyan, Y. Wang, and S. S. Xue, *On the GeV Emission of the Type I BdHN GRB 130427A*, The Astrophysical Journal, 886, 82 (2019).

The dialogue we had started at MG15 undoubtedly had a very stimulating effect on the community which have been followed up and expanded upon in the subsequent MG16 meeting in 2021. Paradoxically, the moments of greatest difficulty for ICRA members have coincided with the flow of new ideas which promoted the steps forward towards the final identification of some fundamental laws of black hole electrodynamics. The very night ICRA offices were locked up and its electronic communications cut, we had submitted a paper to The Astrophysical Journal for publication containing the first fundamental steps for identifying the "inner engine" of GRBs as discussed in talks in this conference. After one year and four months a judge ordered Sapienza to return the ICRA offices.

I am looking forward to a renewed successful collaboration with Sapienza University, enjoyed for more than three decades, to pursue a new common "path" for the rapidly expanding knowledge of the largest distant objects in our Universe. There is still room for improvement: now that the differences have been identified in published articles, a joint effort on an attentive scientific analysis on both sides can lead to the understanding of the nature of these differences and to jointly converge to common solutions. For this to occur, dialogue and not obstruction of office space is needed. As soon as these clarifications will be achieved, we will be able to proceed in a broad scientific effort of common topics of interest in Einstein's theory, in relativistic astrophysics, ranging from the astrophysics of black holes, neutron stars, dark matter and primordial cosmology, observing the rules of basic academic relations.

I am grateful to all members of the IOC and to Carlo Luciano Bianco, Nathalie Deruelle, Rahim Moradi, Tsvi Piran, Jorge Rueda and Narek Sahakyan for advice and discussion.

Remo Ruffini
Chair of the MG International Organizing Committee
Director of ICRANet
President of ICRA

January 10, 2022

This page intentionally left blank

# CONTENTS



## PART A
## PLENARY AND REVIEW TALKS

## PARALLEL SESSIONS

## PART B

## PART C

# PART A

This page intentionally left blank

# Radio Astronomy and the rise of high energy astrophysics: Two anniversaries

Malcolm Longair

*Cavendish Laboratory,*
*University of Cambridge,*
*Cambridge, UK*
** E-mail: msl1000@cam.ac.uk*

This essay celebrates the 100th anniversary of the birth of Martin Ryle and the 50th anniversary of the discovery of pulsars by Jocelyn Bell and Antony Hewish. Ryle and Hewish received the 1974 Nobel Prize in Physics, the first in the area of astrophysics. Their interests strongly overlapped, one of the key papers on the practical implementation of the technique of aperture synthesis being co-authored by Ryle and Hewish. The discovery of pulsars and the roles played by Hewish and Bell are described. These key advances were at the heart of the dramatic rise of high energy astrophysics in the 1960s and led to the realisation that general relativity is central to the understanding of high energy astrophysical phenomena.

*Keywords*: Martin Ryle, Antony Hewish, Jocelyn Bell-Burnell, earth-rotation aperture synthesis, high energy astrophysics, radio sources, neutron stars, supermassive black holes

## 1. Two Anniversaries

2018 is a cause for celebration in the high energy astrophysical community. The two anniversaries are the centenary of the birth of Martin Ryle (1918-1984) and the 50th anniversary of the announcement of the discovery of pulsars in 1968, associated with the names of Antony Hewish and Jocelyn Bell-Burnell. They were all members of the Radio Astronomy Group in the Cavendish Laboratory. I was present as a graduate student and research fellow through the exciting period from 1963 to 1970.

Martin Ryle and Antony Hewish were awarded the Nobel Prize in Physics in 1974, the first to be awarded in astrophysics. The citation reads:

> 'for their pioneering research in radio astrophysics: Ryle for his observations and inventions, in particular of the aperture synthesis technique, and Hewish for his decisive role in the discovery of pulsars.'

Their experimental work was central to the realisation that high energy particles and strong magnetic fields play a major role in modern astrophysics and that general relativity is essential in order to understand high energy astrophysical phenomena.

## 2. The Origins of Radio Astronomy and the Impact of the Second World War

The story begins in 1933 with the discovery of radio waves from our Galaxy by Karl Jansky who was working at the Bell Telephone Laboratories.[1] The discovery arose

from the need to understand sources of radio interference in radio communications. Grote Reber followed up Jansky's discovery with his own home-made radio telescope. By 1940, he had succeeded in making the first map of the radio emission from the Galaxy, the results being published in the *Astrophysical Journal*.[2] Little attention was paid to this work by professional astronomers. The only paper was a discussion which showed that the spectrum of the radiation found by comparing Jansky's and Reber's intensity measurements could not be the thermal emission of hot gas clouds.[3]

Immediately after the Second World War, astronomy was about to be changed out of all recognition as compared with the pre-War era. The contributory causes can be summarised as follows:

- The opening up of the whole of the electromagnetic spectrum for astronomical observation was facilitated by huge advances in radio and electronic techniques as well as by the availability of rockets from which observations could be made from above the Earth's atmosphere.
- Investment in pure science increased dramatically as the benefits of basic research for defence purposes and for the benefit of society were appreciated.
- Scientific electronic computation began in earnest and would provide the means for advancing all scientific disciplines dramatically.
- Astronomy became one of the 'Big Sciences'.

These all contributed to many of the great and unexpected astronomical discoveries of the succeeding years.[a]

After the Second World War, a number of University Groups began to investigate the nature of the cosmic radio emission discovered by Jansky. The principal groups involved were at Cambridge, Manchester and Sydney, all of them led by scientists who came from a background in radar. The science of radio communication and detection developed at a great pace during the War under the combined pressures of defending the UK from incoming enemy aircraft and rockets and developing air-borne radar and radio location techniques. The Cambridge efforts were led by Martin Ryle who assembled a brilliant team of young physicists to attack these problems. The Radio Astronomy Group was remarkably tight-knit and everyone contributed to the various technical challenges.

Two of these were of particular importance. The first was the need to achieve higher angular resolution and sensitivity of the antennae and receiver systems. The second was the need to understand the origin and nature of the 'twinkling' or 'scintillation' of the radio sources. Ryle and Hewish worked on both problems, as can be appreciated from a list of some of their joint papers.

[a]For more details, see my book *The Cosmic Century: A History of Astrophysics and Cosmology*. Cambridge: Cambridge University Press (2006). For the high energy astrophysical aspects of the story, see also my book *High Energy Astrophysics: third edition*. Cambridge: Cambridge University Press (2011).

Fig. 1. The Cambridge Radio Astronomy Group in the early 1950s. Those seated in the middle row are (left to right) Francis Graham Smith, Martin Ryle and Antony Hewish. (Courtesy and copyright the Cavendish Laboratory, University of Cambridge.

- Ryle, M. and Hewish, A. (1950), The Effects of the Terrestrial Ionosphere on the Radio Waves from Discrete Sources in the Galaxy. [4]
- Ryle, M. and Hewish, A. (1955), The Cambridge Radio Telescope. [5]
- Ryle, M. and Hewish, A. (1960), The Synthesis of Large Radio Telescopes. [6]
- Scott, P. F., Ryle, M. and Hewish, A. (1961), The First Results of Radio Star Observations using the Method of Aperture Synthesis. [7]

Immediately after the Second World War, there was very little money, but Ryle and his colleagues were able to make very good use of surplus war equipment, including high quality radio antennae, a large amount of coaxial cable and other items brought back to the UK as German war booty. It was quickly understood that the way to achieving the goals of higher angular resolution and greater sensitivity was to use radio interferometry and, in particular, to implement the techniques of aperture synthesis in which both the amplitude and phase of the interferometric observations are preserved. Martin Ryle's contribution of genius was the practical implementation of Earth-rotation aperture synthesis which resulted in both high angular resolution and high sensitivity images of the radio sky.

Optical telescopes reflect the light of a distant object from a parabolic mirror which has the property that the signals from a distant object reflected from all parts of the mirror surface travel the same distance to the focus. The radio astronomers realised that the reflecting surfaces do not need to be part of the same surface. If

the path lengths to the focus from the source are the same, the interferometric data provide the necessary amplitude and phase information to begin the reconstruction of the image on the sky. To ensure that the waves travel the same distance, delay lines needed to be introduced so that the signals from the two telescopes were combined in phase. Increasing the number of antennae increases the number of possible pairings of antennae, the short baselines providing the large-scale structure and the long baselines the fine detail.

These techniques were exploited in a series of radio interferometers constructed, first of all, at the Rifle Range site just behind the Cambridge University Rugby Ground and then at the Lord's Bridge Observatory which was opened in 1957 once the full significance of radio astronomy for astrophysics and cosmology had become apparent. Radio astronomy hit the headlines in 1955 with the first results of the Second Cambridge (2C) Survey. The dramatic result was that there is a large excess of extragalactic radio sources at large distances, implying that these objects had evolved strongly with cosmic epoch. This was initially a controversial result but it led to the need for deeper surveys and the continued development of radio interferometric techniques. The first large interferometer on the Lord's Bridge site was the 4C radio telescope completed in about 1960. The surveys of the Northern sky carried out by the telescope showed convincingly the evolutionary nature of the radio source population. These radio surveys were carried out with fixed telescopes which mapped the sky by allowing the Earth's rotation to provide a scan of the sky above the telescope.

To sample the two-dimensional structure of the sources, Ryle and his colleagues pioneered the concept that it is simplest to build a one-dimensional interferometer and then use the Earth's rotation to carry one telescope about another as viewed from a point on the sky. In this way information is obtained corresponding to the annulus of a large telescope with diameter equal to the maximum separation of the elements of the interferometer. By adding together a number of baselines with different spacings, the equivalent of a single large telescope with diameter equal to the longest baseline separation can be synthesised with much improved sensitivity. Ryle and Ann Neville used the 4C telescope system in 1962 to create the first fully two-dimensional map of a region about the North Celestial Pole using the Earth-rotation synthesis technique.[8] Every available receiver in the Observatory was needed to make the observations.

The implementation of fully-steerable aperture-synthesis radio telescopes was realised with the construction of the Cambridge One-Mile Telescope (OMT). It required a great deal of innovation in electronics, path compensation and computation. The new generation of electronic computers, the Cambridge EDSAC-1 and 2 machines, was essential to carry out the Fourier transforms to convert the interferometric data into two-dimensional maps. The Fast Fourier transform was implemented to make these computations feasible in a reasonable time.

Fig. 2. The Cambridge One-Mile Telescope, the world's first fully-steerable, general purpose, Earth-rotation aperture synthesis radio telescope.

In 1965, the first radio images from the One-Mile Telescope, the world's first fully-steerable, general purpose, Earth-rotation aperture synthesis radio telescope (Fig. 2), were made of the radio galaxy Cygnus A and the supernova remnant Cassiopaeia A.[9] I was there in the EDSAC control room when the first maps came out of the computer printer. The next step was to extend these techniques to higher frequencies with larger numbers of telescopes and this was achieved with the 5-kilometre (Ryle) in the early 1970s. This resulted in much higher angular resolution and sensitivity. The success of these telescopes led to the construction of even more powerful instruments such as the Very Large Array in the USA.

It is remarkable that over the 25 year period from the end of the Second World War, the sensitivity of radio astronomical observations increased by a factor of about one million and the imaging capability of the telescope system improved from several degrees to a few arcseconds, comparable to that of ground-based optical telescopes. This was Martin Ryle's legacy to radio astronomy. After 1972, his health declined and his interest changed to wind power, sustainability and opposition to nuclear power.

The major impact of radio astrophysics upon astrophysics and cosmology in general cannot be overstated. The discovery of Galactic and extragalactic radio sources revealed the importance of relativistic astrophysics for astronomy in general. To summarise the change of perspective:

- Enormous energies in relativistic particles and magnetic fields were needed to account for the synchrotron radio emission of the radio sources and involved the conversion of $10^6\, M_\odot c^2$ of mass into these forms of energy, at the same time ejecting them far beyond the confines of the host galaxy.
- The role of relativistic jets in powering the huge energies in relativistic particles and magnetic fields became apparent.
- The discovery of the quasars and the BL-Lac objects opened up quite new challenges for the astrophysics of these objects in all wavebands.
- The extreme variability of some of the quasars and BL-Lac objects led to the realisation that supermassive black holes had to be involved in the most extreme active galactic nuclei.
- Evidence for the cosmological evolution of extragalactic radio sources, both radio galaxies and quasars, showed that major changes had taken place in the properties of these objects over the last 75% of the age of the Universe.

These discoveries were first reviewed internationally at the first Texas Symposium on *Relativistic Astrophysics* held in Dallas, Texas in 1963. At the closing dinner, Thomas Gold remarked:

> 'Everyone is pleased: the relativists who feel they are being appreciated, who are suddenly experts in a field which they hardly knew existed; the astrophysicists for having enlarged their domain, their empire by the annexation of another subject - general relativity.'

This was the beginning of high energy astrophysics in its modern guise.

## 3. The Discovery of Pulsars

The discovery of pulsars in 1967 is associated with the names of Antony Hewish and Jocelyn Bell-Burnell, but the seeds of their achievement were sown long before during the immediate post-War years. During that period, part of Hewish's research involved understanding the nature of the scintillations of the intensities of radio sources caused by intervening moving plasma clouds. Just as stars twinkle even on the clearest nights, so point sources of radio emission are observed to scintillate, particularly at long radio wavelengths. Their cause is the deflections of radio rays when they pass through irregularities in the ionospheric plasma.

The theory of the process of scintillation was worked out in detail by Hewish in 1951 in a paper entitled 'The diffraction of radio waves in passing through a phase-changing ionosphere'.[10] The paper set out the theoretical background needed to understand the short-term fluctuations in the intensities of radio sources due to irregularities in an ionised plasma. The same concepts could be used to understand the physics of fluctuations due to ionospheric, interplanetary and interstellar electron density fluctuations. This theoretical paper was followed in 1952 by another entitled 'The Diffraction of Galactic Radio Waves as a Method of Investigating the

Irregular Structure of the Ionosphere'.[11] Applying these concepts to observations of the fluctuating radio signals, Hewish showed that the scale of the irregularities ranged from 2 to 10 km, that the variation of the electron content was about $5 \times 10^9$ electrons cm$^{-2}$ and that the irregularities are at a height of about 400 km. These irregularities moved with a steady wind-like motion at a velocity of the order 100 to 300 m s$^{-1}$.

The same technique could be used to study the solar corona, the region of hot plasma surrounding the Sun. The radio source Taurus A (the Crab Nebula) was observed at varying angular distances from the Sun and the variability of the signal could be accounted for by scattering because of the presence of fluctuations of the electron density in the solar corona. In his paper of 1955 'The Irregular Structure of the Outer Regions of the Solar Corona', Hewish derived the sizes and electron densities of coronal irregularities in the distance range 5 to 15 solar radii.[12]

In 1954, Hewish had remarked in his notebooks that, if the angular sizes of the extragalactic radio sources were small enough, they would illuminate the solar corona with a coherent radio signal and so give rise to rapid time variations in their intensities. This idea was forgotten until about 1962 when Margaret Clarke showed that two of the compact 3CR radio sources ($\theta \lesssim 2$ arcsec) varied very rapidly in intensity. Hewish realised that his old idea was the answer.

By 1964, a number of radio quasars were known and some of these radio sources had small angular sizes. With Paul Scott and Derek Wills, Hewish showed that the radio scintillations were due to scattering of the radio waves by inhomogeneities in the ionised plasma flowing out from the Sun, the Solar Wind. This wind had been predicted by Eugene Parker in 1958 and observed by the Soviet Luna satellites in 1959 and by the US Mariner-2 satellite in 1962. The paper by Hewish, Scott and Wills showed how radio source scintillations could be used to map the outflowing Solar Wind.[13]

Hewish realised that a large, low-frequency array dedicated to the measurement of the scintillations of compact radio sources would provide a new approach to the study of three important astronomical areas:

- it would enable many more quasars to be discovered,
- their angular sizes could be estimated,
- the structure and velocity of the Solar Wind could be determined.

In 1965, he designed a large array to undertake these studies and was awarded a grant of £17,286 by the UK Department of Scientific and Industrial Research to construct it, as well as outstations for measuring the velocity of the Solar Wind. To obtain adequate sensitivity at the low observing frequency of 81.5 MHz (3.7 m wavelength), the array had to be large, 4.5 acres (1.8 hectares) in area, in order to record the rapidly fluctuating intensities of bright radio sources on time-scales as short as one tenth of a second.

Jocelyn Bell joined the 4.5 acre array project as a graduate student in October 1965. She was involved in the construction of the telescope, including knocking the posts into the ground, and then became responsible for the network of cables connecting the dipoles. The telescope was commissioned during July 1967 with the objective of mapping the whole sky once a week so that the variation of the scintillation of the sources with solar elongation could be studied. The array consisted of 2,048 full-wave dipoles arranged in 16 rows of 128 elements. Each row was 470 m long and the north-south extent of the array was 45 m.

A key aspect of the array was that it had to measure the fractional scintillations of the radio sources in real time. Before the days of high speed digital computers, this was achieved by electronic processing of the incoming signals. On a strip chart, the top trace showed the intensity of the source as it passed through the beam of the telescope. This signal was then passed through a high-pass filter so that only the fluctuating component was registered in the middle trace, from which the noise power in the fluctuating component could be displayed in the bottom trace.

While the array was being constructed, Leslie Little and Hewish carried out a theoretical investigation of the strength of the scintillations as a function of heliocentric coordinates. They demonstrated how the angular sizes of the sources could be estimated from measurements of the amplitudes of the scintillations when sources were observed at different solar elongations.[14] A key point was that the scintillations decrease to very small amplitudes when observed at large angles from the Sun.

The commissioning of the 4.5 acre array proceeded through the summer of 1967. Hewish suggested that Bell create sky charts for each strip of the sky each day, noting all the scintillating sources. If the scintillating sources were present on successive weeks at the same astronomical coordinates, they were likely to be real sources, whereas if they were simply interference, for example caused by a nearby unsuppressed tractor or motorcycle, they would not recur at the same astronomical coordinates. This was a very demanding task requiring great persistence, patience and attention to detail on Bell's part since she had to keep up with the very high rate at which the charts were being produced by the telescope, over 200 metres per week.

The discovery of the pulsar CP 1919 was made by Bell on 6 August 1967, the story of the discovery being contained in Appendix 1 of her PhD dissertation. The remarkable feature of CP1919 was that the source scintillated at roughly the 100% level in the anti-solar direction, quite contrary to the expectations of the scintillation models of Little and Hewish. Furthermore, the source was highly variable and not always present. It was not observed again until 28 November 1968, this time with a much short time-constant in the receiver system – the pulses were detected separately for the first time. To everyone's astonishment, the signal consisted entirely of a sequence of pulses with repetition period 1.33 sec, the period being stable to better than one part in $10^6$.

The following two months were what Hewish described as the most exciting of

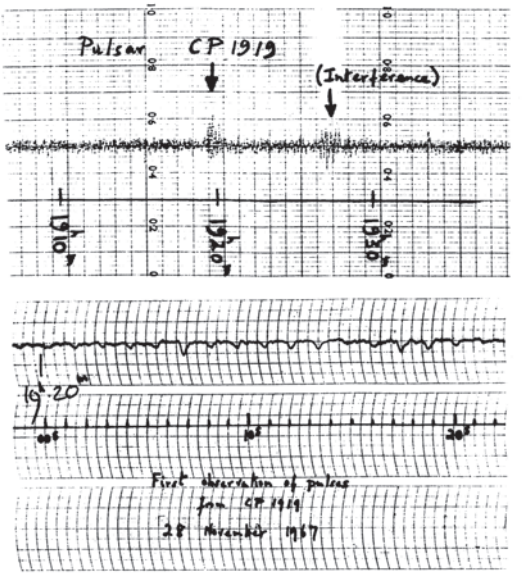



Fig. 3. The discovery record of CP 1919 taken on 6 August 1967. (Courtesy and copyright the Cavendish Laboratory, University of Cambridge and Churchill College Archives.)

his scientific career. Nothing like this had been observed in astronomy before and Hewish and his colleagues had to be absolutely certain of the correctness of the observations. It was essential to carry out follow-up observations and experiments:

- All sources of terrestrial inference had to be excluded.
- If the source was associated with extraterrestrial emissions, including the notorious 'Little Green Men (LGM)', the motion of a planet about the parent star would be easily detectable. The motion of the Earth about the Sun was observed, but no orbital motion of the source.
- The low frequency signals displayed dispersion, the high frequency signals arriving earlier than the low frequencies. This enabled a rough distance of 65 parsec (about 200 light years) to be estimated for the source.
- Three other similar sources were discovered by Bell including one with a period of only 0.25 seconds.

The discovery was kept under tight wraps until Hewish and his colleagues were absolutely convinced that they had discovered a new type of astronomical phenomenon. I was in the next door office to Hewish at the time and I knew nothing about what was going on until he gave a lecture about the discovery in the week before the *Nature* paper was published. The paper 'Observation of a Rapidly Pulsating Radio Source' was submitted for publication in *Nature* on 9th February 1968 and published on 24th February 1968.[15]

Within a few months, Thomas Gold convincingly associated the pulsars with magnetised, rotating neutron stars.[16] The radio pulses are caused by beams of very high energy particles emitting coherent radiation of extraordinarily high brightness temperature escaping from the poles of a magnetised rotating neutron star. When the beam passes across the line of sight to the observer, an intense burst of radio emission is observed.

Very soon after the discovery, large numbers of pulsars were discovered. By now, well over 2000 radio pulsars are known and they are of the greatest astrophysical importance as the last stable stars before collapse to a black hole ensues. The neutron stars represent matter in bulk at nuclear densities and offer many challenges for physicists and astrophysicists. Perhaps most significant was the fact that relativistic stars really exist in nature — general relativity is essential in working out their stability.

In 1972, neutron stars were discovered as the compact X-ray emitting sources in X-ray binary systems by Riccardo Giacconi and his colleagues from observations with the UHURU X-ray observatory.[17] In these sources, the energy source is the accretion of matter from the normal primary star onto the poles of the neutron star.

In 1975, Russell Hulse and Joseph Taylor discovered that the pulsar PSR 1913+16 is a member of a binary neutron star system.[18] This was a fabulous gift to relativists since it can be considered to be a perfect clock in a rotating frame of reference. The binary neutron star system loses energy by the radiation of gravitational waves and one of the great discoveries was the measurement of the speeding up of the binary due to this process. The remarkable agreement between theory and experiment shows that general relativity is the best theory of relativistic gravity we possess.

The discovery of the pulsars resulted in the award of the Nobel prize to Hewish in 1974. Hewish continued his research on the use of the scintillation technique to chart 'interplanetary weather', work which is of the considerable importance because of its impact upon the GPS system. Bell went on to become a distinguished member of the UK scientific community and has received many awards recognising her role in the discovery of pulsars, most recently the $3 M Special Breakthrough Prize in Fundamental Physics in 2018 — she has generously donated this remarkable prize to the Institute of Physics to support research studentships. In June 2007, she was created Dame Jocelyn Bell-Burnell in the UK honours list. She has been President of the UK Institute of Physics and of the Royal Society of Edinburgh. She has recently been appointed Chancellor of the University of Dundee.

## References

1. K. G. Jansky, Electrical Disturbances Apparently of Extraterrestrial Origin, *Proceedings of the Institution of Radio Engineers*, **21**, 1387–1398 (1933).
2. G. Reber, Cosmic Static, *Astrophysical Journal*, **91**, 621–624 (1940).
3. L. G. Henyey and P. C. Keenen, *Astrophysical Journal*, **91**, 625–630 (1940).

4. M. Ryle and A. Hewish, The Effects of the Terrestrial Ionosphere on the Radio Waves from Discrete Sources in the Galaxy. *Monthly Notices of the Royal Astronomical Society*, **110**, 381–394 (1950).
5. M. Ryle and A. Hewish, The Cambridge Radio Telescope. *Memoirs of the Royal Astronomical Society*, **67**, 97–105 (1955).
6. M. Ryle and A. Hewish, The Synthesis of Large Radio Telescopes. *Monthly Notices of the Royal Astronomical Society*, **120**, 220–230 (1960).
7. P. F. Scott, M. Ryle and A. Hewish, First Results of Radio Star Observations Using the Method of Aperture Synthesis. *Monthly Notices of the Royal Astronomical Society*, **122**, 95–111 (1961).
8. M. Ryle and A. C. Neville, A Radio Survey of the North Polar Region with a 4.5 Minutes of Arc Pencil-beam System. *Monthly Notices of the Royal Astronomical Society*, **125**, 39–56 (1962).
9. M. Ryle, B. Elsmore and A. C. Neville, High Resolution Observations of the Radio Sources in Cygnus and Cassiopeia. *Nature*, **205**, 1259–1262 (1965).
10. A. Hewish, The diffraction of radio waves in passing through a phase-changing ionosphere. *Proceedings of the Royal Society of London*, **209**, 81–96, (1951).
11. A. Hewish, The Diffraction of Galactic Radio Waves as a Method of Investigating the Irregular Structure of the Ionosphere. *Proceedings of the Royal Society of London*, **214**, 494–514 (1952).
12. A. Hewish, The Irregular Structure of the Outer Regions of the Solar Corona. *Proceedings of the Royal Society of London*, **228**, 238–251, (1955).
13. A. Hewish, P. F. Scott and D. Wills, Interplanetary Scintillation of Small Diameter Radio Sources. *Nature*, **203**, 1214-1217 (1964).
14. L. T. Little and A. Hewish, Interplanetary scintillation and its relation to the angular structure of radio sources. *Monthly Notices of the Royal Astronomical Society*, **134**, 221-237 (1966).
15. A. Hewish, S. J. Bell, J. D. H. Pilkington, P. F. Scott and R. A. Collins, Observation of a Rapidly Pulsating Radio Source. *Nature*, **217**, 709-713 (1968).
16. T. Gold. Rotating Neutron Stars as the Origin of Pulsating Radio Sources *Nature*, **218**, 731–732 (1968).
17. H. Tananbaum, H., H. Gursky, E. M. Kellogg, R. Levinson, E. Schreier and R. Giacconi. Discovery of a Periodic Binary X-ray Source in Hercules from UHURU, *Astrophysical Journal*, **174**, L144–L149 (1972).
18. R. A. Hulse and J. H. Taylor, Discovery of a Pulsar in a Binary System, *Astrophysical Journal Letters*, **195**, L51–L53 (1975).

# The CMB then and now with snapshots of Dave Wilkinson

Lyman A. Page Jr.

*Department of Physics, Princeton University, Princeton, NJ 08544-0708, USA*
**E-mail: page@princeton.edu*

We comment on the advance in measurements of the anisotropy of the CMB spanning from when Remo Ruffini was starting his academic career and watching the field's development to the present.

*Keywords*: Cosmic Microwave Background, Cosmology, David T. Wilkinson.

## 1. Ruffini and Wilkinson at Princeton

Remo Ruffini and Dave Wilkinson and were faculty colleagues at Princeton in the early 1970s. Remo was working in the group established by John Wheeler on theoretical aspects of general relativity: gravitational radiation, black holes, neutron stars, and the coupling of electromagnetic and gravitational radiation process to name a few topics. Dave on the other hand was working with colleagues in Bob Dicke's "gravity research group" on experimental aspects of gravity: tests of the equivalence principle, lunar laser ranging, pulsar timing, measurements of the CMB, and the search for primeval galaxies. Just this simple listing gives a sense for what the intellectual atmosphere must have been like. So many possibilities were just about within reach but yet beyond the limits of the empirical tests of the times. Peebles describes the environment at Princeton in these days, and what led Dicke and Wheeler, quite independently, to pursue general relativity [1]. The topics with which Remo and Dave were engaged are, of course, at the forefront of modern physics and as exciting as ever. In the following I'll focus on just one element of the activity in the 1970s, the cosmic microwave background (CMB), and give a brief update on the status of observations.

## 2. The CMB anisotropy in the 1970s

Figure 1 shows the state of measurements of the anisotropy as of 1979. It is taken from "Finding the Big Bang" [2]. It was known that the anisotropy should be present at some level because there are galaxies and clusters of galaxies. They formed via gravitational instability and the same instability affects the CMB. But the level was uncertain and it was not known whether foregrounds or, for example, an early epoch of reionization would mask the anisotropy. The number of groups who were measuring the anisotropy was small. Those who had published results between 1965 and 1979 are shown in the figure. The year 1967, not long after the initial discovery, was particularly active, with limits coming from the "Isotropometer,"

the first instrument designed explicitly to measure the anisotropy [3,4], "Stanford" [5,6], "Crawford Hill" [7], and "Aerospace" [8].

"White Mountain" [9] was also run by the Stanford group. Conklin was searching for evidence of the CMB dipole. It was not appreciated until many years later that his was probably the first experiment to detect anomalous microwave emission [10]. With "RATAN" [11–13] the Soviets were limiting the anisotropy with a highly asymmetric beam while Steve Boughn published his senior thesis at Princeton, "KaDip" [14], on a search for the quadrupole. "XBal" [15] was the first CMB balloon experiment for measuring the anisotropy. Additional limits came from "NRAO-P" [16], "Goldstone" [17], and "Parkes" [18]. The "U2" [19] experiment made a big splash because of its novel platform, a modified U2 spy plane, and a clear detection of the CMB dipole, although at least one other measurement of the dipole had already been presented at a conference [20]. In addition, the U2 was a testing ground for the DMR experiment aboard COBE that discovered the anisotropy [21]. "Testa-Griga" [22] is of special note. It was the first experiment to use bolometers on the ground to search for the anisotropy. Today, all active CMB anisotropy experiments use bolometers. Towards the end of the decade came "GBank-R" [23] and then "MIT" [24,25], a balloon-borne bolometric radiometer.

These were small experiments (or small observing programs on existing telescopes) by today's standards. The total number of separate authors in all the sixteen different measurements spanning this fifteen-year period is about thirty. By comparison, the Planck mission began around 1996 and the number of authors on the most recent parameters paper [26] is just over 175.

## 3. Wilkinson in the 1990s

Before the MAP collaboration formed, Dave had been thinking about a satellite to follow on from COBE. Not much more than a decade had passed since the results in Figure 1 but the field was advancing rapidly. Figure 2 shows what Dave had on his mind. One day when I was walking by his office he called out and asked if I was interested in working on a satellite project. I said "sure!" Dave formed a small group of Ed Wollack, one of his graduate students (affectionally called "Waveguide Wollack" by Ken Ganga, Michael Joyce, and José Gonzalez), Norm Jarosik, a research scientist in the group, and me. When we weren't working on other experiments, we'd iterate on the design shown in the figure. Telling of the times, Dave called the satellite PIE, the Princeton Isotropy Experiment. The anisotropy had yet to be discovered. His vision, though, was far from "PIE in the sky." After pursuing a few different paths [2], we ended up collaborating with the group at NASA Goddard Space Flight Center with Chuck Bennett as the PI. The Goddard group had independently been thinking about a space mission. The proposal, submitted in 1995, was a partnership between Princeton and Goddard. Although Dave passed away in 2002, he saw the first maps from the mission he

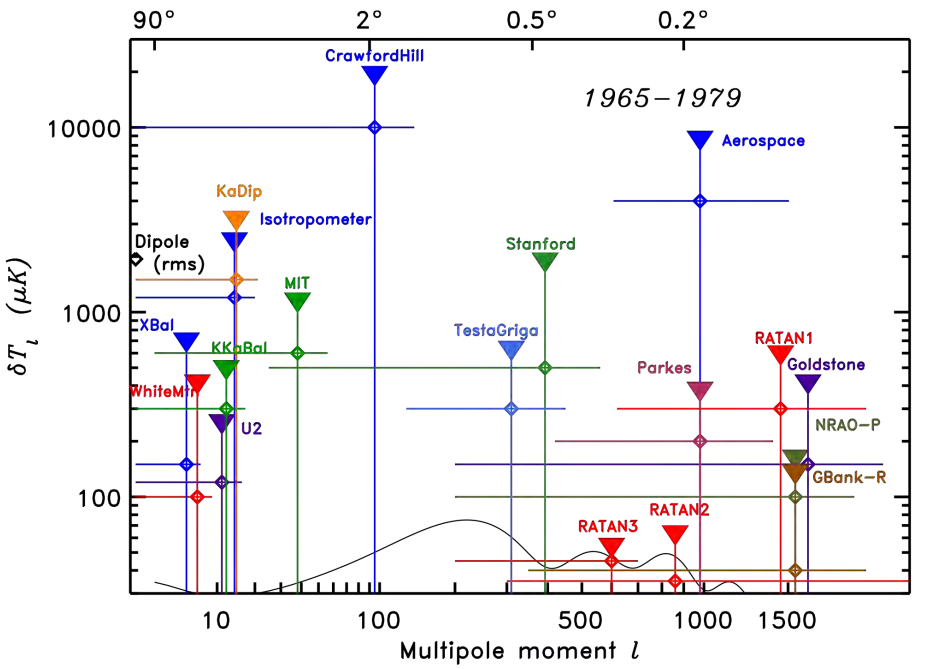



Fig. 1. Status of anisotropy measurements as of 1979. The quantity on the y-axis is $\delta T_\ell = (\ell(\ell+1)C_\ell/2\pi)^{1/2}$ where $C_\ell$ is the angular power spectrum of the CMB. At the time of the measurements, different conventions were used. The original data were converted to the modern standard format as discussed in "Finding the Big Bang" [2]. The triangles show $1\sigma$ upper limits; the horizontal lines show the range of $\ell$ corresponding to the limit. The smooth curve near the bottom of the plot shows the LCDM model in the same units.

had dreamed about just a decade earlier. The satellite was renamed the Wilkinson Microwave Anisotropy Probe (WMAP) in his honor.

## 4. The state of the art

Figure 3 shows the state of the art in anisotropy measurements from groups that have published in the past few years. Since the discovery of the anisotropy by COBE in 1992 a whole field has been born and matured. We have learned more than we ever thought, or even knew, possible in the early 1990s. Although I don't think Dave would have been surprised by the advance in instrumentation, I think he might have been surprised at how kind Nature has been in revealing her secrets through the CMB.

There is a lot captured in Figure 3. For the temperature anisotropy (TT), Planck has the smallest error bars up to about $\ell = 2000$. The higher resolution ground-based ACT and SPT data fill in from $\ell = 500$ up to about $\ell = 8000$. For those in the field, there are important differences between the measurements, but if one takes a step back the agreement is a major success for the field. It is important

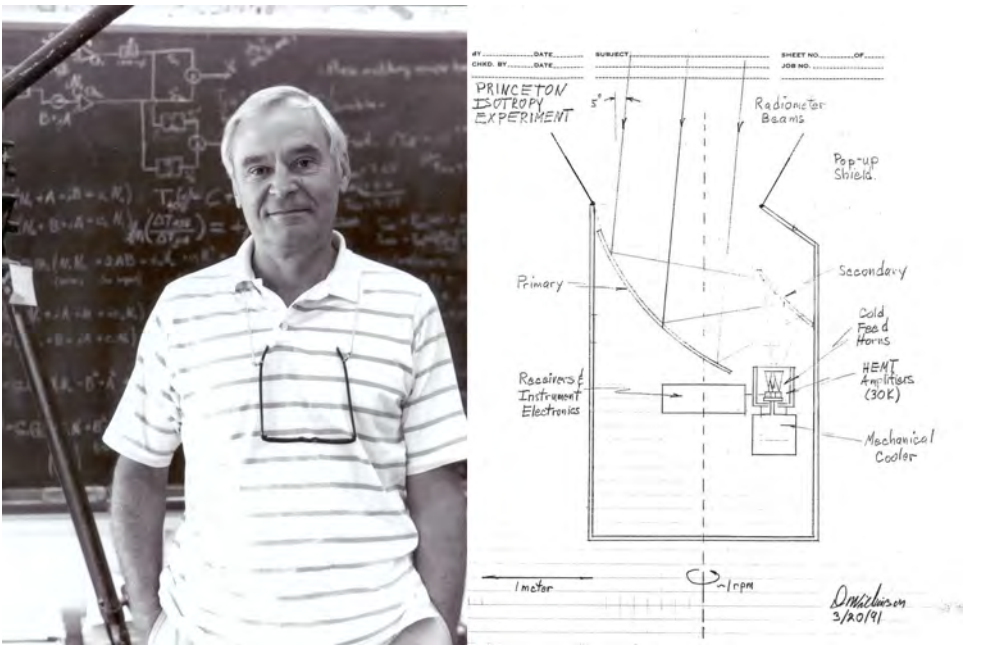



Fig. 2. Dave Wilkinson in front of the blackboard and his 1991 vision for a dedicated CMB satellite.

to note that the agreement is not just in the power spectrum but is in the maps. Different competing teams with different methods of analysis measure the same thing when compared. Both ACT and SPT calibrate off of the Planck maps in a range of $\ell$s near $\ell \sim 1000$.

The two lines through the TT spectrum are the WMAP and Planck basic six-parameter LCDM models. The first thing to note is their similarity. Two completely different satellites, based on different technologies, and analyzed with different methods give the same results.

Sometimes the LCDM model is characterized as just the six parameter fit to cosmological data but it is, of course, more than that. To the limits of measurement, the anisotropy is Gaussian and the fluctuations are adiabatic. In addition, the fluctuations are super horizon. As the universe expands, we sample ever more of a pre-existing sea of primordial fluctuations outside our current Hubble volume.

The Planck/WMAP agreement is especially notable because the data sets are considerably different. Although Planck was the union of two different technologies so that it could cover, importantly, from 30-800 GHz, the following will just address the HFI. On the one hand, Planck is much more sensitive than WMAP. For example the Planck 143 GHz channels, their most sensitive, have a combined instantaneous sensitivity[a] of 14 $\mu$Ks$^{1/2}$ relative to the CMB whereas for WMAP, the instantaneous

[a]All quoted sensitivities are relative to the CMB and for the temperature anisotropy, as opposed to the polarization.

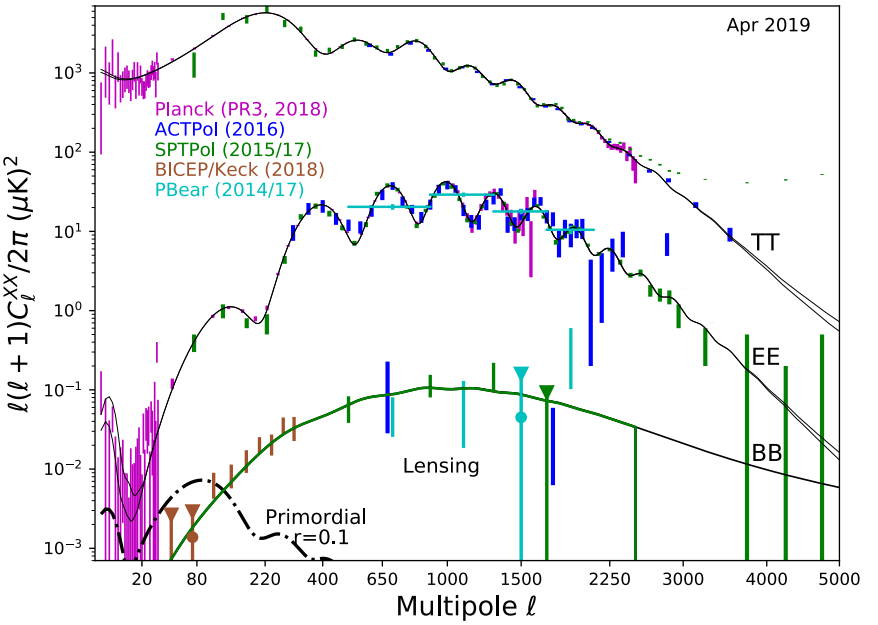



Fig. 3. Recent CMB anisotropy measurements. The y-axis in this plot is the square of the quantity in Figure 1. The TE spectrum is not shown. The data come from Planck, ACTPol, SPTPol, BICEP2/Keck, and Polarbear/Simons Array. Maps from ACT, SPT (and of course WMAP) are publicly available through NASA's LAMBDA site (https://lambda.gsfc.nasa.gov) and Planck's are available through https://wiki.cosmos.esa.int/planck-legacy-archive/index.php/CMB_maps. The *rms* fluctuations for the full range of $\ell$s are TT: 113 $\mu$K, EE 6.4 $\mu$K, BB lensing: 0.43 $\mu$K, and BB primordial: 0.12 $\mu$K for the full range for $r = 0.1$ but 0.09 $\mu$K for $30 < \ell < 160$.

sensitivity in any band is a factor of 30 or more higher[27]. On the other hand, the WMAP data are much cleaner. There are no A/D non-linearities, no long detector time constants, and barely any cosmic rays or glitches to excise. The distribution of WMAP's raw data is Gaussian[28] so that the noise in the sky pixels can be characterized simply by the number of observations.

The HFI sensitivity is notable. Planck still holds the record for the single most sensitive detector that has observed the CMB[29], 50 $\mu$Ks$^{1/2}$. These bolometers came through the efforts of Andrew Lange and Paul Richards, and their power was dramatically demonstrated with the Boomerang experiment[30], led by Lange and La Sapienza's Paolo deBernardis. WMAP was based on the much different broad-band coherent amplifiers developed by NRAO's Marian Pospieszalski.

The next line down on the plot shows the EE spectrum. To minimize "clutter," the TE spectrum is not plotted. The largest discrepancy between WMAP and Planck is on the optical depth, $\tau$. This is a difficult parameter to pin down because it is manifest at large scales, where the foreground emission is large and instrumental systematic errors can have an outsized effect. In Figure 3 the discrepancy is seen as

the difference between the two thin solid curves at $\ell < 20$ for the EE polarization. Planck simply does not see as much low-$\ell$ power as WMAP did. The Planck value is $\tau = 0.054 \pm 0.0076$ [26]. Weiland *et al.*, [31] taking into account WMAP and a range of systematic errors recommends $\tau = 0.07 \pm 0.02$. Reionization is process that is begging to be measured better. Among other things our understanding of it impacts how well one can determine the sum of neutrino masses from the CMB [32].

The LCDM model has a number of built-in cross checks. The same model that fits TT predicts the TE spectrum, the EE spectrum, and the lensing BB spectrum. The physics of these spectra are different. For example, the polarization is produced in a relatively short time right around decoupling as opposed to the longer time over which the temperature anisotropies are produced. As can be seen in the plot, the EE pattern is in excellent agreement with the model that fits TT. The adjustable parameters between them are the polarization efficiency and detector angles.

In Figure 3 one can also see that the Planck TT spectrum runs out of statistical power near $\ell \sim 2300$ and the EE spectrum does so at $\ell \sim 1500$. Both ACT and SPT are more sensitive in this range. In particular, this means that with increased sensitivity the cosmological parameters can be redone using only CMB polarization. This program is in the works.

The bottom curve shows the lensing BB spectrum. The signal is produced by gravitational lensing of the CMB E-modes [33] as the CMB propagates to us from the decoupling surface through the matter fluctuations in the volume of the universe. The fluctuations in matter density arise from the same primordial power spectrum that gives rise to the TT/TE/EE CMB spectra. Thus the lensing ties together two manifestations of the primordial power spectrum. Lensing also affects TT and TE and, to within a few $\sigma$, the lensing effects agree with the those of the primary anisotropy.

The much-discussed primordial B-modes are shown on the bottom left. If they exist at measurable levels, they will give us new insights into the physics of the birth of the universe and to how gravitational waves couple to fundamental fields. The plotted data are from the BICEP2/Keck experiment [34].

## 5. Challenges

Between the COBE discovery in 1992 and 1995, there were half dozen groups publishing measurements of the anisotropy. Most of the results were in the form of power spectra and only a few could be confirmed through cross correlation with another experiment or through a repeated independent measurement in, say, a subsequent year. This prompted Dave Wilkinson, who often admonished the community to "show our warts," to write "A warning label for cosmic microwave background anisotropy experiments." [35]. The abstract is still relevant:

*It is demonstrated that there may be undetected systematic errors or foreground sources lurking in measurements of anisotropy in the cosmic microwave background radiation (CMBR). The elaborate statistical analysis techniques used to identify tiny*

*signals in hours of noisy data are discussed. The important errors in the results are probably not generated by Gaussian noise, but by spurious signals and unknown foreground sources that are hiding in that noise, along with whatever CMBR signal is being detected. It is likely that among the many reports of statistically significant sky signals there are some real detections of CMBR anisotropy. However, given the number of ways that results can be contaminated, and the modest level of candor and quantitative assessment of possible systematic effects in the experimental papers, theorists should treat most current "detections" with skepticism.*

The field has advanced enormously since then. We routinely produce *maps* of the temperature anisotropy and *maps* of the polarization that can be compared between experiments. In contrast to reporting power spectra, maps allow one to check both the amplitude and phase of any purported signal. This gives faith in the results, but also assumes that the maps are public. There are still claims of different, beyond LCDM, phenomena at the 2-3$\sigma$ level and time will tell if these turn into something real. The statistical methods are now much more sophisticated but at the same time Dave's "tiny signals" have been redefined. In 1995, a typical error bar was $10 - 20$ $\mu$K for $\delta\ell \sim 50$; today it is an order of magnitude or more smaller. The warning label still obtains: one must be skeptical of new unconfirmed claims based on coadding months and months of data.

The overall challenge of the measurement can be appreciated from a few numbers. On the ground, the best sites are at the South Pole and Chile. At 150 GHz, the physical temperature of the surroundings is 250 K and the atmosphere is 10 K. The *rms* of the E-mode polarization is 6 $\mu$K. To measure this to an accuracy of 1% requires control/understanding of the environment to better than a part in $10^6$. The advances over the years have not only been in detector sensitivity but in technique. In the mid-1990s we did not know how best to extract such small signals.

The field is entering a new era. For example, when one measures B-mode polarization at large angular scales, one is measuring foregrounds. Thus understanding, measuring, and modeling polarized foregrounds has become an active area of research. So far, the only measured polarized foregrounds are from synchrotron and dust emission. In temperature these components show spatial variations in their spectra indices and they are non-Gaussian. The polarized emission comes from different regions of the interstellar medium and a definitive departure from power law behavior has not been published. Though we might get lucky in modeling them, the field is prepared for the challenge of dealing with a complicated sky. To measure primordial B-modes with say $r = 0.001$ with a S/N of a few entails measuring a 10 nK level signal with a few nK precision. The lensing produces a corresponding *rms* of roughly 65 nK. In a clean part of the sky, the *rms* level of dust emission of the sweet spot for ground observing at $30 < \ell < 160$ is roughly 220 nK.

At smaller angular scales, gains can still be made with only modest cleaning. Figure 4 shows the dust foregrounds at 150 GHz, one of the prime observing frequencies. The top grey line is for TT. The bottom two swaths are for the polarization

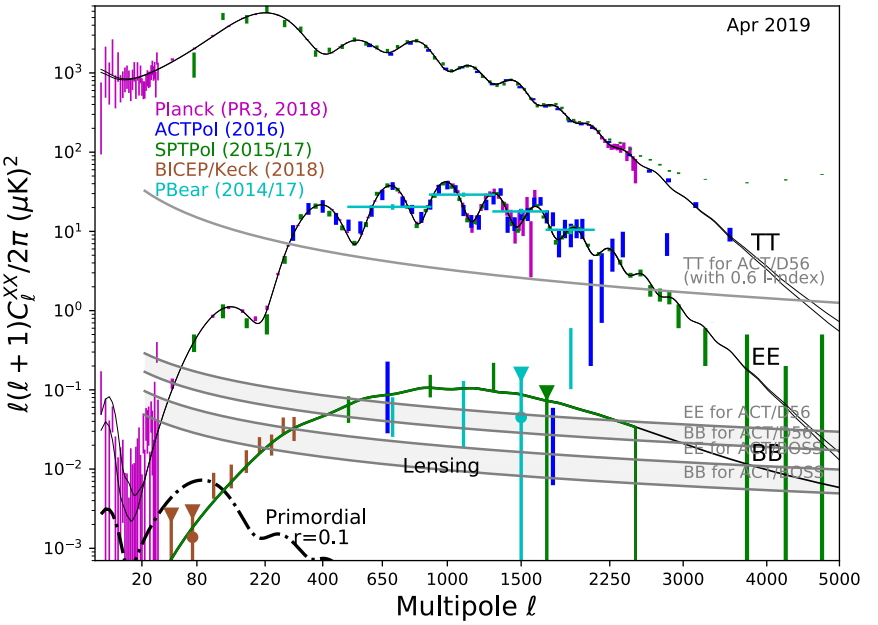



Fig. 4. Foreground emission from dust at 150 GHz for different regions of sky. The top grey line is for TT and the bottom two swaths are for polarized emission in a typical off-plane (top) and a relatively clean (bottom) region. Cleaner parts of the sky exist.

in different regions. The BB foreground levels are typically a factor of two smaller than the EE levels as indicated by the width of the various bands.

## 6. The future

The CMB is the gift that keeps on giving. We are far from learning all we can from it. Largely driven by the different demands of measuring large and small angular scales, there are two broad classes of instruments: "small telescopes" for large angular scales, particularly primordial B-modes, and "large telescopes" for smaller angular scales corresponding to $\ell > 300$. The Polarbear/Simons Array instrument[b] straddles these two broad classes.

### 6.1. *Large angular scales*

On the ground, the leading large-angular scale experiment is BICEP2/Keck which operates from the South Pole. It's advantages are the relatively stable atmospheric conditions at the South Pole, a large array of detectors, and the nearby SPT to help

[b]The Polarbear/Simons Array project is distinct from the Simons Observatory project.

delens the signal. The limitation is access to large clean, perhaps even the cleanest, areas of sky. The current results are from $400\,\mathrm{deg}^2$. As can be seen in Figure 3, delensing will be needed before too long.

The CLASS experiment is operating in Chile with the most sensitive array yet at 40 GHz: $32\,\mu\mathrm{Ks}^{1/2}$. (The previous record holder was the QUIET experiment[36] at $69\,\mu\mathrm{Ks}^{1/2}$.) Higher frequencies are fielded and observing. Their strategy is to cover a large area of sky, currently 75%, to search for primordial B-modes and measure the optical depth[37].

The QUBIC[38] team is building a novel bolometric interferometer to search for B-modes. It will be situated in northern Argentina, not too far from the ACT/SO/PB-SA/CLASS site. This is the most novel of the current generation of experiments. It is the sort of thing I think Dave would have liked to try. Other examples of him taking a direction different from the rest of the field are his dual-horn correlation receiver experiment done with Peter Timbie[39] and his development of a mm-wave detector based on Rydberg atoms which was never published. The first of these presaged the successful compact interferometric arrays of CBI and DASI.

The Simons Observatory collaboration is in the process of building three "Small Aperture Telescopes" (SATs)[40] that will be situated in Chile. As of this writing, the first is scheduled to be fielded in 2020. In contrast to BICEP2/Keck, they will observe roughly 4000 $\mathrm{deg}^2$. For this area, delensing is not needed to reach a sensitivity of $\sigma_r = 0.002$. (See Table 1 ). The CMB-S4 collaboration envisions a collection of 14 half-meter aperture receivers.

There are a number of balloon-borne experiments in various stages of development. Currently the most advanced is SPIDER[41]. It has completed one balloon flight mapping 4000 $\mathrm{deg}^2$ of the sky at 90 and 150 GHz. The preliminary in-flight net sensitivity is $7.1\,\mu\mathrm{Ks}^{1/2}$ at 90 GHz and $5.3\,\mu\mathrm{Ks}^{1/2}$ at 150 GHz. The LSPE[42], EBEX-IDS[43] (next generation EBEX[44]), and PIPER[45] experiments (with first flight in Oct. 2017) are in various stages of testing. There are also some recent proposals for measuring the optical depth from balloons. During the conference, La Sapienza's OLIMPO experiment was launched. While its science goal is the characterization of SZ clusters, it has advanced detector technology relevant to B-mode measurements. In particular, it gave the first demonstration of KID detectors in a space environment[46].

Space is the ultimate environment for CMB measurements. For a measurement as important B-modes, a new satellite mission is a natural and arguably necessary next step. There have been a number of proposals that have reached various stages of review. They include COrE, COrE+, PRISM, PIXIE, EPIC, EPIC-IM, CMB-Bharat, PICO, and LiteBIRD. The number of proposals indicates the general realization of the importance of such a mission. One clear advantage of space is the ability to look at a wide range of frequencies to clean foreground emission as assessed, for example, by Remazeilles et al.[47]. The most developed of the satellite missions is LiteBIRD, which has already passed a number of review milestones.

Table 1. Comparison of some projections for future large experiments

| | Existing | Simons Observatory | CMB-S4 | LiteBIRD | PICO |
|---|---|---|---|---|---|
| Estimated first light | Current | 2021 | 2027 | 2027 | TBD |
| Angular scale | | $\ell > 30$ | $\ell > 30$ | $\ell < 200$ | $\ell < 4000$ |
| $\sigma_r$ | 0.03 | $2 \times 10^{-3}$ | $0.5 \times 10^{-3}$ | $1 \times 10^{-3}$ | $1 \times 10^{-4}$ |
| $\sigma_{N_{eff}}$ | 0.2 | 0.05 | 0.03 | | 0.06 (95% cl) |
| $\sigma_\tau$ | 0.0076 | | | | 0.002 |
| $\sigma_{\Sigma m_\nu}$ (meV) | 100 | 30 | 26 | | 15 |

*Note*: The current values of $\sigma$ are from BICEP2/Keck [34] for the tensor to scalar ratio, $r$, and from Planck [26] for the other parameters. The Simons Observatory values [48] are the "goals" and include foreground cleaning. The CMB-S4 projections [49] for $\sigma_{\Sigma m_\nu}$ were originally 15 meV but are here modified according to Allison et al. [32]. The other projection are from the LiteBIRD [50] and PICO [51] projects. See references for constraints on additional parameters.

### 6.2. *Small angular scales*

At angular scales where one can measure multiple SZ clusters as well as the high $\ell$ part of the CMB ($200 < \ell < 8000$) the two major telescopes are ACT and SPT. There are both in their third generation of receiver and continue to produce great science. ACT, of which I am a part, has mapped half the sky with receivers that have a combined sensitivity, for just two of three arrays, of about $8\,\mu\mathrm{Ks}^{1/2}$ at 90 GHz and $11\,\mu\mathrm{Ks}^{1/2}$ at 150 GHz [52] in nominally good (pwv=1.3 mm) observing conditions. In other words, they exceed the sensitivity of Planck even though they are observing through the atmosphere.

The next generation of large telescopes are already in the works. The Simons Foundation has made a significant award for a new effort, called the Simons Observatory [c], to be built in Chile. Not only are there the three SATs mentioned above, but there will be a new 6-m telescope based on the CCAT-prime design [53]. The telescope and the associated receiver are currently under construction. The currently scheduled delivery date for the telescope is late 2020. There are yet further plans for additional large telescopes under the CMB-S4 umbrella. Table 1 shows projections for what these new efforts aim to achieve.

The CMB field is alive and growing. There is steady progress on new techniques and on improving the detector sensitivity. The more we learn, the more we know what we can learn. Through the CMB, and the CMB combined with other observations, we are piecing together an interlocking picture of the universe that combines high energy theory, particle physics, general relativity, and a host of exciting astronomical phenomena. The potential for discovery is high and the future prospects exciting.

[c]See https://simonsobservatory.org for more details.

**References**


1. P. J. E. Peebles, Robert Dicke and the naissance of experimental gravity physics, 1957-1967, *European Physical Journal H* **42** (June 2017).
2. P. J. E. Peebles, L. A. Page, Jr. and R. B. Partridge, *Finding the Big Bang* March 2009.
3. D. T. Wilkinson and R. B. Partridge, Large Scale Density Inhomogeneities in the Universe, *Nature* **215**, p. 719 (August 1967).
4. R. B. Partridge and D. T. Wilkinson, Isotropy and Homogeneity of the Universe from Measurements of the Cosmic Microwave Background, *Physical Review Letters* **18**, 557 (April 1967).
5. E. K. Conklin and R. N. Bracewell, Isotropy of the Cosmic Background Radiation at 10690 MHz, *Physical Review Letters* **18**, p. 614 (April 1967).
6. E. K. Conklin and R. N. Bracewell, Limits on Small Scale Variations in the Cosmic Background Radiation, *Nature* **216**, 777 (November 1967).
7. R. W. Wilson and A. A. Penzias, Isotropy of Cosmic Background Radiation at 4080 Megahertz, *Science* **156**, 1100 (May 1967).
8. E. E. Epstein, On the Small-Scale Distribution at 3.4-MM Wavelength of the Reported 3 deg K Background Radiation, *Ap. J. Lett.* **148**, p. L157 (June 1967).
9. E. K. Conklin, Velocity of the Earth with Respect to the Cosmic Background Radiation, *Nature* **222**, 971 (June 1969).
10. M. Lu, J. Dunkley and L. Page, Evidence for Anomalous Dust-correlated Emission at 8 GHz, *Ap. J.* **749**, p. 165 (April 2012).
11. Y. N. Pariiskii and T. B. Pyatunina, An Attempt to Detect Protogalaxies from Fluctuations in the Primordial Background Brightness., *Soviet Astron Lett.* **14**, p. 1067 (June 1971).
12. Y. N. Pariiskii, New limit of fluctuations in the primordial background radiation of the universe, *Soviet Astron Lett.* **17**, p. 291 (December 1973).
13. I. N. Pariiskii, Z. E. Petrov and L. N. Cherkov, A search for primordial perturbations of the universe - Observations with the RATAN-600 radio telescope, *Soviet Astronomy Letters* **3**, 263 (December 1977).
14. S. P. Boughn, D. M. Fram and R. B. Patridge, Isotropy of the Microwave Background at 8-MILLIMETER Wavelength, *Ap. J.* **165**, p. 439 (May 1971).
15. P. S. Henry, Isotropy of the 3 K Background, *Nature* **231**, 516 (June 1971).
16. Y. N. Parijskij, New Limit on Small-Scale Irregularities of "blackbody" Radiation, *Ap. J. Lett.* **180**, p. L47 (March 1973).
17. R. L. Carpenter, S. Gulkis and T. Sato, Search for Small-Scale Anisotropy in the 2.7 deg K Cosmic Background Radiation at a Wavelength of 3.56 Centimeters, *Ap. J. Lett.* **182**, p. L61 (June 1973).
18. K. S. Stankevich, Fine-scale temperature variations in the 11.1-cm cosmic background radiation, *Soviet Astronomy Letters* **18**, p. 126 (August 1974).
19. G. F. Smoot, M. V. Gorenstein and R. A. Muller, Detection of anisotropy in the cosmic blackbody radiation, *Physical Review Letters* **39**, 898 (October 1977).
20. B. E. Corey and D. T. Wilkinson, A Measurement of the Cosmic Microwave Background Anisotropy at 19 GHz, *Proc. BAAS* **8**, p. 351 (March 1976).
21. G. F. Smoot, C. L. Bennett, A. Kogut, E. L. Wright, J. Aymon, N. W. Boggess, E. S. Cheng, G. de Amici, S. Gulkis, M. G. Hauser, G. Hinshaw, P. D. Jackson, M. Janssen, E. Kaita, T. Kelsall, P. Keegstra, C. Lineweaver, K. Loewenstein, P. Lubin, J. Mather, S. S. Meyer, S. H. Moseley, T. Murdock, L. Rokke, R. F. Silverberg, L. Tenorio, R. Weiss and D. T. Wilkinson, Structure in the COBE differential microwave radiometer first-year maps, *Ap. J. Lett* **396**, L1 (September 1992).

22. N. Caderni, V. de Cosmo, R. Fabbri, B. Melchiorri, F. Melchiorri and V. Natale, Infrared limit on the fine-scale anisotropy of the cosmic background radiation, *Phys. Rev. D.* **16**, 2424 (October 1977).
23. L. Rudnick, Ionized gas in X-ray clusters of galaxies - Radio limits, *Ap. J.* **223**, 37 (July 1978).
24. D. Muehlner, Spectrum and isotropy of the submillimeter background radiation, *Proc. Inf. and Submm. Ast.* **63**, 143 (1977).
25. D. K. Owens, D. J. Muehlner and R. Weiss, A large-beam sky survey at millimeter and submillimeter wavelengths made from balloon altitudes, *Ap. J.* **231**, 702 (August 1979).
26. Planck Collaboration, N. Aghanim, Y. Akrami, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday, R. B. Barreiro, N. Bartolo, S. Basak, R. Battye, K. Benabed, J.-P. Bernard, M. Bersanelli, P. Bielewicz, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, M. Bucher, C. Burigana, R. C. Butler, E. Calabrese, J.-F. Cardoso, J. Carron, A. Challinor, H. C. Chiang, J. Chluba, L. P. L. Colombo, C. Combet, D. Contreras, B. P. Crill, F. Cuttaia, P. de Bernardis, G. de Zotti, J. Delabrouille, J.-M. Delouis, E. Di Valentino, J. M. Diego, O. Doré, M. Douspis, A. Ducout, X. Dupac, S. Dusini, G. Efstathiou, F. Elsner, T. A. Enßlin, H. K. Eriksen, Y. Fantaye, M. Farhang, J. Fergusson, R. Fernandez-Cobos, F. Finelli, F. Forastieri, M. Frailis, E. Franceschi, A. Frolov, S. Galeotta, S. Galli, K. Ganga, R. T. Génova-Santos, M. Gerbino, T. Ghosh, J. González-Nuevo, K. M. Górski, S. Gratton, A. Gruppuso, J. E. Gudmundsson, J. Hamann, W. Handley, D. Herranz, E. Hivon, Z. Huang, A. H. Jaffe, W. C. Jones, A. Karakci, E. Keihänen, R. Keskitalo, K. Kiiveri, J. Kim, T. S. Kisner, L. Knox, N. Krachmalnicoff, M. Kunz, H. Kurki-Suonio, G. Lagache, J.-M. Lamarre, A. Lasenby, M. Lattanzi, C. R. Lawrence, M. Le Jeune, P. Lemos, J. Lesgourgues, F. Levrier, A. Lewis, M. Liguori, P. B. Lilje, M. Lilley, V. Lindholm, M. López-Caniego, P. M. Lubin, Y.-Z. Ma, J. F. Macías-Pérez, G. Maggio, D. Maino, N. Mandolesi, A. Mangilli, A. Marcos-Caballero, M. Maris, P. G. Martin, M. Martinelli, E. Martínez-González, S. Matarrese, N. Mauri, J. D. McEwen, P. R. Meinhold, A. Melchiorri, A. Mennella, M. Migliaccio, M. Millea, S. Mitra, M.-A. Miville-Deschênes, D. Molinari, L. Montier, G. Morgante, A. Moss, P. Natoli, H. U. Nørgaard-Nielsen, L. Pagano, D. Paoletti, B. Partridge, G. Patanchon, H. V. Peiris, F. Perrotta, V. Pettorino, F. Piacentini, L. Polastri, G. Polenta, J.-L. Puget, J. P. Rachen, M. Reinecke, M. Remazeilles, A. Renzi, G. Rocha, C. Rosset, G. Roudier, J. A. Rubiño-Martín, B. Ruiz-Granados, L. Salvati, M. Sandri, M. Savelainen, D. Scott, E. P. S. Shellard, C. Sirignano, G. Sirri, L. D. Spencer, R. Sunyaev, A.-S. Suur-Uski, J. A. Tauber, D. Tavagnacco, M. Tenti, L. Toffolatti, M. Tomasi, T. Trombetti, L. Valenziano, J. Valiviita, B. Van Tent, L. Vibert, P. Vielva, F. Villa, N. Vittorio, B. D. Wandelt, I. K. Wehus, M. White, S. D. M. White, A. Zacchei and A. Zonca, Planck 2018 results. VI. Cosmological parameters, *arXiv e-prints* (July 2018).
27. S. Hanany, M. D. Niemack and L. Page, *CMB Telescopes and Optical Systems*, in *Planets, Stars and Stellar Systems. Volume 1: Telescopes and Instrumentation*, eds. T. D. Oswalt and I. S. McLean 2013, p. 431.
28. N. Jarosik, C. Barnes, C. L. Bennett, M. Halpern, G. Hinshaw, A. Kogut, M. Limon, S. S. Meyer, L. Page, D. N. Spergel, G. S. Tucker, J. L. Weiland , E. Wollack and E. L. Wright, First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: On-Orbit Radiometer Characterization, *Ap.J.S.* **148**, 29 (Sep 2003).
29. Planck HFI Core Team, P. A. R. Ade, N. Aghanim, R. Ansari, M. Arnaud, M. Ashdown, J. Aumont, A. J. Banday, M. Bartelmann, J. G. Bartlett, E. Battaner,

K. Benabed, A. Benoît, J.-P. Bernard, M. Bersanelli, R. Bhatia, J. J. Bock, J. R. Bond, J. Borrill, F. R. Bouchet, F. Boulanger, T. Bradshaw, E. Bréelle, M. Bucher, P. Camus, J.-F. Cardoso, A. Catalano, A. Challinor, A. Chamballu, J. Charra, M. Charra, R.-R. Chary, C. Chiang, S. Church, D. L. Clements, S. Colombi, F. Couchot, A. Coulais, C. Cressiot, B. P. Crill, M. Crook, P. de Bernardis, J. Delabrouille, J.-M. Delouis, F.-X. Désert, K. Dolag, H. Dole, O. Doré, M. Douspis, G. Efstathiou, P. Eng, C. Filliard, O. Forni, P. Fosalba, J.-J. Fourmond, K. Ganga, M. Giard, D. Girard, Y. Giraud-Héraud, R. Gispert, K. M. Górski, S. Gratton, M. Griffin, G. Guyot, J. Haissinski, D. Harrison, G. Helou, S. Henrot-Versillé, C. Hernández-Monteagudo, S. R. Hildebrandt, R. Hills, E. Hivon, M. Hobson, W. A. Holmes, K. M. Huffenberger, A. H. Jaffe, W. C. Jones, J. Kaplan, R. Kneissl, L. Knox, G. Lagache, J.-M. Lamarre, P. Lami, A. E. Lange, A. Lasenby, A. Lavabre, C. R. Lawrence, B. Leriche, C. Leroy, Y. Longval, J. F. Macías-Pérez, T. Maciaszek, C. J. MacTavish, B. Maffei, N. Mandolesi, R. Mann, B. Mansoux, S. Masi, T. Matsumura, P. McGehee, J.-B. Melin, C. Mercier, M.-A. Miville-Deschênes, A. Moneti, L. Montier, D. Mortlock, A. Murphy, F. Nati, C. B. Netterfield, H. U. Nørgaard-Nielsen, C. North, F. Noviello, D. Novikov, S. Osborne, C. Paine, F. Pajot, G. Patanchon, T. Peacocke, T. J. Pearson, O. Perdereau, L. Perotto, F. Piacentini, M. Piat, S. Plaszczynski, E. Pointecouteau, R. Pons, N. Ponthieu, G. Prézeau, S. Prunet, J.-L. Puget, W. T. Reach, C. Renault, I. Ristorcelli, G. Rocha, C. Rosset, G. Roudier, M. Rowan-Robinson, B. Rusholme, D. Santos, G. Savini, B. M. Schaefer, P. Shellard, L. Spencer, J.-L. Starck, P. Stassi, V. Stolyarov, R. Stompor, R. Sudiwala, R. Sunyaev, J.-F. Sygnet, J. A. Tauber, C. Thum, J.-P. Torre, F. Touze, M. Tristram, F. van Leeuwen, L. Vibert, D. Vibert, L. A. Wade, B. D. Wandelt, S. D. M. White, H. Wiesemeyer, A. Woodcraft, V. Yurchenko, D. Yvon and A. Zacchei, Planck early results. IV. First assessment of the High Frequency Instrument in-flight performance, *A&Ast* **536**, p. A4 (December 2011).

30. P. de Bernardis, P. A. R. Ade, J. J. Bock, J. R. Bond, J. Borrill, A. Boscaleri, K. Coble, B. P. Crill, G. De Gasperis, P. C. Farese, P. G. Ferreira, K. Ganga, M. Giacometti, E. Hivon, V. V. Hristov, A. Iacoangeli, A. H. Jaffe, A. E. Lange, L. Martinis, S. Masi, P. V. Mason, P. D. Mauskopf, A. Melchiorri, L. Miglio, T. Montroy, C. B. Netterfield, E. Pascale, F. Piacentini, D. Pogosyan, S. Prunet, S. Rao, G. Romeo, J. E. Ruhl, F. Scaramuzzi, D. Sforna and N. Vittorio, A flat Universe from high-resolution maps of the cosmic microwave background radiation, *Nature* **404**, 955 (April 2000).
31. J. L. Weiland, K. Osumi, G. E. Addison, C. L. Bennett, D. J. Watts, M. Halpern and G. Hinshaw, Effect of Template Uncertainties on the WMAP and Planck Measures of the Optical Depth Due to Reionization, *Ap.J.* **863**, p. 161 (August 2018).
32. R. Allison, P. Caucal, E. Calabrese, J. Dunkley and T. Louis, Towards a cosmological neutrino mass detection, *PRD* **92**, p. 123535 (Dec 2015).
33. M. Zaldarriaga and U. Seljak, Gravitational lensing effect on cosmic microwave background polarization, *PRD* **58**, p. 023003 (July 1998).
34. BICEP2 Collaboration, Keck Array Collaboration, P. A. R. Ade, Z. Ahmed, R. W. Aikin, K. D. Alexander, D. Barkats, S. J. Benton, C. A. Bischoff, J. J. Bock, R. Bowens-Rubin, J. A. Brevik, I. Buder, E. Bullock, V. Buza, J. Connors, J. Cornelison, B. P. Crill, M. Crumrine, M. Dierickx, L. Duband, C. Dvorkin, J. P. Filippini, S. Fliescher, J. Grayson, G. Hall, M. Halpern, S. Harrison, S. R. Hildebrandt, G. C. Hilton, H. Hui, K. D. Irwin, J. Kang, K. S. Karkare, E. Karpel, J. P. Kaufman, B. G. Keating, S. Kefeli, S. A. Kernasovskiy, J. M. Kovac, C. L. Kuo, N. A. Larsen, K. Lau, E. M. Leitch, M. Lueker, K. G. Megerian, L. Moncelsi, T. Namikawa, C. B. Netterfield, H. T. Nguyen, R. O'Brient, R. W. Ogburn, S. Palladino, C. Pryke, B. Racine, S. Richter, A. Schillaci, R. Schwarz, C. D. Sheehy, A. Soliman, T. St. Germaine, Z. K. Staniszewski, B. Steinbach, R. V. Sudiwala, G. P. Teply, K. L. Thompson, J. E. Tolan,

C. Tucker, A. D. Turner, C. Umiltà, A. G. Vieregg, A. Wandui, A. C. Weber, D. V. Wiebe, J. Willmert, C. L. Wong, W. L. K. Wu, H. Yang, K. W. Yoon and C. Zhang, Constraints on Primordial Gravitational Waves Using Planck, WMAP, and New BICEP2/Keck Observations through the 2015 Season, *Physical Review Letters* **121**, p. 221301 (November 2018).

35. D. Wilkinson, A warning label for cosmic microwave background anisotropy experiments., *Proc. Part. Phys. and Cos.* (1995).
36. C. Bischoff, A. Brizius, I. Buder, Y. Chinone, K. Cleary, R. N. Dumoulin, A. Kusaka, R. Monsalve, S. K. Næss, L. B. Newburgh, G. Nixon, R. Reeves, K. M. Smith, K. Vanderlinde, I. K. Wehus, M. Bogdan, R. Bustos, S. E. Church, R. Davis, C. Dickinson, H. K. Eriksen, T. Gaier, J. O. Gundersen, M. Hasegawa, M. Hazumi, C. Holler, K. M. Huffenberger, W. A. Imbriale, K. Ishidoshiro, M. E. Jones, P. Kangaslahti, D. J. Kapner, C. R. Lawrence, E. M. Leitch, M. Limon, J. J. McMahon, A. D. Miller, M. Nagai, H. Nguyen, T. J. Pearson, L. Piccirillo, S. J. E. Radford, A. C. S. Readhead, J. L. Richards, D. Samtleben, M. Seiffert, M. C. Shepherd, S. T. Staggs, O. Tajima, K. L. Thompson, R. Williamson, B. Winstein, E. J. Wollack and J. T. L. Zwart, The Q/U Imaging ExperimenT Instrument, *Ap.J.* **768**, p. 9 (May 2013).
37. D. J. Watts, B. Wang, A. Ali, J. W. Appel, C. L. Bennett, D. T. Chuss, S. Dahal, J. R. Eimer, T. Essinger-Hileman, K. Harrington, G. Hinshaw, J. Iuliano, T. A. Marriage, N. J. Miller, I. L. Padilla, L. Parker, M. Petroff, K. Rostem, E. J. Wollack and Z. Xu, A Projected Estimate of the Reionization Optical Depth Using the CLASS Experiment's Sample Variance Limited E-mode Measurement, *Ap.J.* **863**, p. 121 (Aug 2018).
38. J. Aumont, S. Banfi, P. Battaglia, E. S. Battistelli, A. Baù, B. Bélier, D. Bennett, L. Bergé, J. P. Bernard, M. Bersanelli, M. A. Bigot-Sazy, N. Bleurvacq, G. Bordier, J. Brossard, E. F. Bunn, D. Buzi, A. Buzzelli, D. Cammilleri, F. Cavaliere, P. Chanial, C. Chapron, G. Coppi, A. Coppolecchia, F. Couchot, R. D'Agostino, G. D'Alessandro, P. de Bernardis, G. De Gasperis, M. De Petris, T. Decourcelle, F. Del Torto, L. Dumoulin, A. Etchegoyen, C. Franceschet, B. Garcia, A. Gault, D. Gayer, M. Gervasi, A. Ghribi, M. Giard, Y. Giraud-Héraud, M. Gradziel, L. Grand sire, J. C. Hamilton, D. Harari, V. Haynes, S. Henrot-Versillé, N. Holtzer, J. Kaplan, A. Korotkov, L. Lamagna, J. Lande, S. Loucatos, A. Lowitz, V. Lukovic, B. Maffei, S. Marnieros, J. Martino, S. Masi, A. May, M. McCulloch, M. C. Medina, S. Melhuish, A. Mennella, L. Montier, A. Murphy, D. Néel, M. W. Ng, C. O'Sullivan, A. Paiella, F. Pajot, A. Passerini, A. Pelosi, C. Perbost, O. Perdereau, F. Piacentini, M. Piat, L. Piccirillo, G. Pisano, D. Prêle, R. Puddu, D. Rambaud, O. Rigaut, G. E. Romero, M. Salatino, A. Schillaci, S. Scully, M. Stolpovskiy, F. Suarez, A. Tartari, P. Timbie, M. Tristram, G. Tucker, D. Viganò, N. Vittori, F. Voisin, B. Watson, M. Zannoni and A. Zullo, QUBIC Technical Design Report, *arXiv e-prints* , p. arXiv:1609.04372 (Sep 2016).
39. P. T. Timbie and D. T. Wilkinson, Low-noise interferometer for microwave radiometry, *Review of Scientific Instruments* **59**, 914 (Jun 1988).
40. N. Galitzki, A. Ali, K. S. Arnold, P. C. Ashton, J. E. Austermann, C. Baccigalupi, T. Baildon, D. Barron, J. A. Beall, S. Beckman, S. M. M. Bruno, S. Bryan, P. G. Calisse, G. E. Chesmore, Y. Chinone, S. K. Choi, G. Coppi, K. D. Crowley, K. T. Crowley, A. Cukierman, M. J. Devlin, S. Dicker, B. Dober, S. M. Duff, J. Dunkley, G. Fabbian, P. A. Gallardo, M. Gerbino, N. Goeckner-Wald, J. E. Golec, J. E. Gudmundsson, E. E. Healy, S. Henderson, C. A. Hill, G. C. Hilton, S.-P. P. Ho, L. A. Howe, J. Hubmayr, O. Jeong, B. Keating, B. J. Koopman, K. Kiuchi, A. Kusaka, J. Lashner, A. T. Lee, Y. Li, M. Limon, M. Lungu, F. Matsuda, P. D. Mauskopf, A. J. May, N. McCallum, J. McMahon, F. Nati, M. D. Niemack, J. L. Orlowski-Scherer, S. C. Parshley, L. Piccirillo, M. Sathyanarayana Rao, C. Raum, M. Salatino,

J. S. Seibert, C. Sierra, M. Silva-Feaver, S. M. Simon, S. T. Staggs, J. R. Stevens, A. Suzuki, G. Teply, R. Thornton, C. Tsai, J. N. Ullom, E. M. Vavagiakis, M. R. Vissers, B. Westbrook, E. J. Wollack, Z. Xu and N. Zhu, The Simons Observatory: instrument overview, **10708**, p. 1070804 (Jul 2018).

41. R. Gualtieri, J. P. Filippini, P. A. R. Ade, M. Amiri, S. J. Benton, A. S. Bergman, R. Bihary, J. J. Bock, J. R. Bond, S. A. Bryan, H. C. Chiang, C. R. Contaldi, O. Doré, A. J. Duivenvoorden, H. K. Eriksen, M. Farhang, L. M. Fissel, A. A. Fraisse, K. Freese, M. Galloway, A. E. Gambrel, N. N. Gandilo, K. Ganga, R. V. Gramillano, J. E. Gudmundsson, M. Halpern, J. Hartley, M. Hasselfield, G. Hilton, W. Holmes, V. V. Hristov, Z. Huang, K. D. Irwin, W. C. Jones, C. L. Kuo, Z. D. Kermish, S. Li, P. V. Mason, K. Megerian, L. Moncelsi, T. A. Morford, J. M. Nagy, C. B. Netterfield, M. Nolta, B. Osherson, I. L. Padilla, B. Racine, A. S. Rahlin, C. Reintsema, J. E. Ruhl, M. C. Runyan, T. M. Ruud, J. A. Shariff, J. D. Soler, X. Song, A. Trangsrud, C. Tucker, R. S. Tucker, A. D. Turner, J. F. v. d. List, A. C. Weber, I. K. Wehus, D. V. Wiebe and E. Y. Young, SPIDER: CMB Polarimetry from the Edge of Space, *Journal of Low Temperature Physics* **193**, 1112 (Dec 2018).
42. S. Aiola, G. Amico, P. Battaglia, E. Battistelli, A. Baó, P. de Bernardis, M. Bersanelli, A. Boscaleri, F. Cavaliere, A. Coppolecchia, A. Cruciani, F. Cuttaia, A. D'Addabbo, G. D'Alessand ro, S. De Gregori, F. Del Torto, M. De Petris, L. Fiorineschi, C. Franceschet, E. Franceschi, M. Gervasi, D. Goldie, A. Gregorio, V. Haynes, N. Krachmalnicoff, L. Lamagna, B. Maffei, D. Maino, S. Masi, A. Mennella, G. Morgante, F. Nati, M. W. Ng, L. Pagano, A. Passerini, O. Peverini, F. Piacentini, L. Piccirillo, G. Pisano, S. Ricciardi, P. Rissone, G. Romeo, M. Salatino, M. Sandri, A. Schillaci, L. Stringhetti, A. Tartari, R. Tascone, L. Terenzi, M. Tomasi, E. Tommasi, F. Villa, G. Virone, S. Withington, A. Zacchei and M. Zannoni, The Large-Scale Polarization Explorer (LSPE), **8446**, p. 84467A (Sep 2012).
43. F. Aubin, S. Hanany, B. R. Johnson, A. Lee, A. Suzuki, B. Westbrook and K. Young, Developments of Highly Multiplexed, Multi-chroic Pixels for Balloon-Borne Platforms, *Journal of Low Temperature Physics* **193**, 298 (Nov 2018).
44. EBEX Collaboration, M. Abitbol, A. M. Aboobaker, P. Ade, D. Araujo, F. Aubin, C. Baccigalupi, C. Bao, D. Chapman, J. Didier, M. Dobbs, S. M. Feeney, C. Geach, W. Grainger, S. Hanany, K. Helson, S. Hillbrand, G. Hilton, J. Hubmayr, K. Irwin, A. Jaffe, B. Johnson, T. Jones, J. Klein, A. Korotkov, A. Lee, L. Levinson, M. Limon, K. MacDermid, A. D. Miller, M. Milligan, K. Raach, B. Reichborn-Kjennerud, C. Reintsema, I. Sagiv, G. Smecher, G. S. Tucker, B. Westbrook, K. Young and K. Zilic, The EBEX Balloon-borne Experiment—Detectors and Readout, *Ap.J. S.* **239**, p. 8 (Nov 2018).
45. S. Pawlyk, P. A. R. Ade, D. Benford, C. L. Bennett, D. T. Chuss, R. Datta, J. L. Dotson, J. R. Eimer, D. J. Fixsen, N. N. Gand ilo, T. M. Essinger-Hileman, M. Halpern, G. Hilton, G. F. Hinshaw, K. Irwin, C. Jhabvala, M. Kimball, A. Kogut, L. Lowe, J. J. McMahon, T. M. Miller, P. Mirel, S. H. Moseley, S. Rodriguez, E. Sharp, P. Shirron, J. G. Staguhn, D. F. Sullivan, E. R. Switzer, P. Taraschi, C. E. Tucker, A. Walts and E. J. Wollack, The primordial inflation polarization explorer (PIPER): current status and performance of the first flight, **10708**, p. 1070806 (Jul 2018).
46. S. Masi, P. de Bernardis, A. Paiella, F. Piacentini, L. Lamagna, A. Coppolecchia, P. A. R. Ade, E. S. Battistelli, M. G. Castellano, I. Colantoni, F. Columbro, G. D'Alessandro, M. De Petris, S. Gordon, C. Magneville, P. Mauskopf, G. Pettinari, G. Pisano, G. Polenta, G. Presta, E. Tommasi, C. Tucker, V. Vdovin, A. Volpe and D. Yvon, Kinetic Inductance Detectors for the OLIMPO experiment: in–flight operation and performance, *arXiv e-prints* , p. arXiv:1902.08993 (Feb 2019).

47. M. Remazeilles, C. Dickinson, H. K. K. Eriksen and I. K. Wehus, Sensitivity and foreground modelling for large-scale cosmic microwave background B-mode polarization satellite missions, *MNRAS* **458**, 2032 (May 2016).
48. P. Ade, J. Aguirre, Z. Ahmed, S. Aiola, A. Ali, D. Alonso, M. A. Alvarez, K. Arnold, P. Ashton, J. Austermann and et al., The Simons Observatory: science goals and forecasts, *JCAP* **2**, p. 056 (February 2019).
49. K. N. Abazajian, P. Adshead, Z. Ahmed, S. W. Allen, D. Alonso, K. S. Arnold, C. Baccigalupi, J. G. Bartlett, N. Battaglia, B. A. Benson, C. A. Bischoff, J. Borrill, V. Buza, E. Calabrese, R. Caldwell, J. E. Carlstrom, C. L. Chang, T. M. Crawford, F.-Y. Cyr-Racine, F. De Bernardis, T. de Haan, S. di Serego Alighieri, J. Dunkley, C. Dvorkin, J. Errard, G. Fabbian, S. Feeney, S. Ferraro, J. P. Filippini, R. Flauger, G. M. Fuller, V. Gluscevic, D. Green, D. Grin, E. Grohs, J. W. Henning, J. C. Hill, R. Hlozek, G. Holder, W. Holzapfel, W. Hu, K. M. Huffenberger, R. Keskitalo, L. Knox, A. Kosowsky, J. Kovac, E. D. Kovetz, C.-L. Kuo, A. Kusaka, M. Le Jeune, A. T. Lee, M. Lilley, M. Loverde, M. S. Madhavacheril, A. Mantz, D. J. E. Marsh, J. McMahon, P. D. Meerburg, J. Meyers, A. D. Miller, J. B. Munoz, H. N. Nguyen, M. D. Niemack, M. Peloso, J. Peloton, L. Pogosian, C. Pryke, M. Raveri, C. L. Reichardt, G. Rocha, A. Rotti, E. Schaan, M. M. Schmittfull, D. Scott, N. Sehgal, S. Shandera, B. D. Sherwin, T. L. Smith, L. Sorbo, G. D. Starkman, K. T. Story, A. van Engelen, J. D. Vieira, S. Watson, N. Whitehorn and W. L. Kimmy Wu, CMB-S4 Science Book, First Edition, *arXiv e-prints* (October 2016).
50. Y. Sekimoto, P. Ade, K. Arnold, J. Aumont, J. Austermann, C. Baccigalupi, A. Banday, R. Banerji, S. Basak, S. Beckman, M. Bersanelli, J. Borrill, F. Boulanger, M. L. Brown, M. Bucher, E. Calabrese, A. Challinor, Y. Chinone, F. Columbro, A. Cukierman, D. Curtis, P. de Bernardis, M. de Petris, M. Dobbs, T. Dotani, L. Duband, A. Ducout, K. Ebisawa, T. Elleflot, H. Eriksen, J. Errard, R. Flauger, C. Franceschet, U. Fuskeland, K. Ganga, J. R. Gao, T. Ghigna, J. Grain, A. Gruppuso, N. Halverson, P. Hargrave, T. Hasebe, M. Hasegawa, M. Hattori, M. Hazumi, S. Henrot-Versille, C. Hill, Y. Hirota, E. Hivon, D. T. Hoang, J. Hubmayr, K. Ichiki, H. Imada, H. Ishino, G. Jaehnig, H. Kanai, S. Kashima, Y. Kataoka, N. Katayama, T. Kawasaki, R. Keskitalo, A. Kibayashi, T. Kikuchi, K. Kimura, T. Kisner, Y. Kobayashi, N. Kogiso, K. Kohri, E. Komatsu, K. Komatsu, K. Konishi, N. Krachmalnicoff, C. L. Kuo, N. Kurinsky, A. Kushino, L. Lamagna, A. T. Lee, E. Linder, B. Maffei, M. Maki, A. Mangilli, E. Martinez-Gonzalez, S. Masi, T. Matsumura, A. Mennella, Y. Minami, K. Mistuda, D. Molinari, L. Montier, G. Morgante, B. Mot, Y. Murata, A. Murphy, M. Nagai, R. Nagata, S. Nakamura, T. Namikawa, P. Natoli, T. Nishibori, H. Nishino, F. Noviello, C. O'Sullivan, H. Ochi, H. Ogawa, H. Ogawa, H. Ohsaki, I. Ohta, N. Okada, G. Patanchon, F. Piacentini, G. Pisano, G. Polenta, D. Poletti, G. Puglisi, C. Raum, S. Realini, M. Remazeilles, H. Sakurai, Y. Sakurai, G. Savini, B. Sherwin, K. Shinozaki, M. Shiraishi, G. Signorelli, G. Smecher, R. Stompor, H. Sugai, S. Sugiyama, A. Suzuki, J. Suzuki, R. Takaku, H. Takakura, S. Takakura, E. Taylor, Y. Terao, K. L. Thompson, B. Thorne, M. Tomasi, H. Tomida, N. Trappe, M. Tristram, M. Tsuji, M. Tsujimoto, S. Uozumi, S. Utsunomiya, N. Vittorio, N. Watanabe, I. Wehus, B. Westbrook, B. Winter, R. Yamamoto, N. Y. Yamasaki, M. Yanagisawa, T. Yoshida, J. Yumoto, M. Zannoni and A. Zonca, Concept design of the LiteBIRD satellite for CMB B-mode polarization, **10698**, p. 106981Y (Aug 2018).
51. S. Hanany, M. Alvarez, E. Artis, P. Ashton, J. Aumont, R. Aurlien, R. Banerji, R. B. Barreiro, J. G. Bartlett, S. Basak, N. Battaglia, J. Bock, K. K. Boddy, M. Bonato, J. Borrill, F. Bouchet, F. Boulanger, B. Burkhart, J. Chluba, D. Chuss, S. E. Clark, J. Cooperrider, B. P. Crill, G. De Zotti, J. Delabrouille, E. Di Valentino,

J. Didier, O. Doré, H. K. Eriksen, J. Errard, T. Essinger-Hileman, S. Feeney, J. Filippini, L. Fissel, R. Flauger, U. Fuskeland, V. Gluscevic, K. M. Gorski, D. Green, B. Hensley, D. Herranz, J. C. Hill, E. Hivon, R. Hložek, J. Hubmayr, B. R. Johnson, W. Jones, T. Jones, L. Knox, A. Kogut, M. López-Caniego, C. Lawrence, A. Lazarian, Z. Li, M. Madhavacheril, J.-B. Melin, J. Meyers, C. Murray, M. Negrello, G. Novak, R. O'Brient, C. Paine, T. Pearson, L. Pogosian, C. Pryke, G. Puglisi, M. Remazeilles, G. Rocha, M. Schmittfull, D. Scott, P. Shirron, I. Stephens, B. Sutin, M. Tomasi, A. Trangsrud, A. van Engelen, F. Vansyngel, I. K. Wehus, Q. Wen, S. Xu, K. Young and A. Zonca, PICO: Probe of Inflation and Cosmic Origins, *arXiv e-prints* , p. arXiv:1902.10541 (Feb 2019).

52. S. K. Choi, J. Austermann, J. A. Beall, K. T. Crowley, R. Datta, S. M. Duff, P. A. Gallardo, S. P. Ho, J. Hubmayr, B. J. Koopman, Y. Li, F. Nati, M. D. Niemack, L. A. Page, M. Salatino, S. M. Simon, S. T. Staggs, J. Stevens, J. Ullom and E. J. Wollack, Characterization of the Mid-Frequency Arrays for Advanced ACTPol, *Journal of Low Temperature Physics* **193**, 267 (Nov 2018).
53. M. D. Niemack, Designs for a large-aperture telescope to map the CMB 10x faster, *Appl. Opt.* **55**, p. 1686 (Mar 2016).

# Quasi-local mass at axially symmetric null infinity*

Po-Ning Chen

*Department of Mathematics, University of California, Riverside, USA*
*E-mail: poningc@ucr.edu*

Mu-Tao Wang

*Department of Mathematics, Columbia University, New York, NY 10027, USA*
*E-mail: mtwang@math.columbia.edu*

Ye-Kai Wang

*Department of Mathematics, National Cheng Kung University, Taiwan*
*E-mail: ykwang@mail.ncku.edu.tw*

Shing-Tung Yau

*Department of Mathematics, Harvard University, USA*
*E-mail: yau@math.harvard.edu*

We give a brief review of the definition of the Wang-Yau quasilocal mass and discuss the evaluation of which on surfaces of unit size at null infinity of an axi-symmetric spacetime in Bondi-van der Burg-Metzner coordinates.

## 1. Introduction

As is well known, it is not possible to find the mass density of gravity in general relativity. The expression for the mass density would have to consist of first derivatives of the metric tensor which are zero in suitable chosen coordinates at a point. But we still desire to measure the total mass in a spacelike region bounded by a closed surface. The mass due to gravity should be computable from the intrinsic and the extrinsic geometry of the surface. It has been considered to be one of the important questions to find the right definition. Penrose gave a talk [24] at the Institute for Advanced Study in 1979 and listed it as the first one in his list of major open problems. The quantity is called quasilocal mass.

Many people including Penrose, [25] Hawking, [16] Brown-York, [5] Hawking-Horowitz, [17] Bartnik [2] and others worked on this problem and various definitions have been given. Several important contributions have also been made by Bartnik, [1] Shi-Tam, [29] and Liu-Yau. [20,21] This article discusses the Wang-Yau quasilocal mass definition discovered in 2009 [33,34].

*P.-N. Chen is supported by NSF grant DMS-1308164 and Simons Foundation collaboration grant #584785, Y.-K. Wang is supported by MOST Taiwan grant 105-2115-M-006-016-MY2, 107-2115-M-006-001-MY2, and S.-T. Yau is supported by NSF grants PHY-0714648 and DMS-1308244. This material is based upon work supported by the National Science Foundation under Grants No. DMS-1405152 and No. DMS-1810856 (Mu-Tao Wang). Part of this work was carried out when the authors were visiting the National Center of Theoretical Sciences at National Taiwan University in Taipei, Taiwan.

## 2. The definition of Wang-Yau quasilocal mass

Quasilocal mass is attached to a 2-dimensional spacetime surface which is a topological 2-sphere (the boundary of a spacelike region). But with different intrinsic geometry and extrinsic geometry, we expect to read off the effect of gravitation in the spacetime vicinity of the surface. Suppose the surface is spacelike, i.e., the induced metric $\sigma$ is Riemannian. An essential part of the extrinsic geometry is measured by the mean curvature vector field $\mathbf{H}$, which is a normal vector field of the surface such that the null expansion along any null normal direction $\ell$ is given by the paring of $\mathbf{H}$ and $\ell$.

To evaluate the quasilocal mass of a 2-surface $\Sigma$ with the physical data $(\sigma, \mathbf{H})$, one solves the *optimal isometric embedding system* (see (1) in the next paragraph), which gives an embedding of $\Sigma$ into the Minkowski spacetime with the image surface $\Sigma_0$ that has the same induced metric $\sigma$ as $\Sigma$. We then compare the extrinsic geometries of $\Sigma$ and $\Sigma_0$ and evaluate the quasilocal mass from $\sigma, \mathbf{H}$, and $\mathbf{H_0}$.

The physical surface $\Sigma$ with physical data $(\sigma, \mathbf{H})$ gives $(\sigma, |\mathbf{H}|, \alpha_{\mathbf{H}})$, where $|\mathbf{H}|$ is the norm of the mean curvature vector field and $\alpha_{\mathbf{H}}$ is the connection one-form determined by the mean curvature gauge. As long as the mean curvature vector field $\mathbf{H}$ is spacelike, $|\mathbf{H}|$ is positive and $\alpha_{\mathbf{H}}$ is well-defined. Given an isometric embedding $X : \Sigma \to \mathbb{R}^{3,1}$ of $\sigma$, let $\Sigma_0 = X(\Sigma)$ be the image and $(\sigma, |\mathbf{H_0}|, \alpha_{\mathbf{H_0}})$ be the corresponding data of $\Sigma_0$. Let $T$ be a future timelike unit Killing field of $\mathbb{R}^{3,1}$ and define $\tau = -\langle X, T\rangle$ as a function defined on the surface $\Sigma$ where $\langle \cdot, \cdot \rangle$ is the Minkowski metric on $\mathbb{R}^{3,1}$.

The optimal isometric embedding system and the quasilocal mass can be expressed in terms of a function $\rho$ and a 1-form $j_a$ on $\Sigma$ given by

$$\rho = \frac{\sqrt{|\mathbf{H_0}|^2 + \frac{(\Delta\tau)^2}{1+|\nabla\tau|^2}} - \sqrt{|\mathbf{H}|^2 + \frac{(\Delta\tau)^2}{1+|\nabla\tau|^2}}}{\sqrt{1+|\nabla\tau|^2}}$$

$$j_a = \rho\nabla_a\tau - \nabla_a\left(\sinh^{-1}(\frac{\rho\Delta\tau}{|\mathbf{H_0}||\mathbf{H}|})\right) - (\alpha_{\mathbf{H_0}})_a + (\alpha_{\mathbf{H}})_a,$$

where $\nabla_a$ is the covariant derivative with respect to the metric $\sigma$, $|\nabla\tau|^2 = \nabla^a\tau\nabla_a\tau$ and $\Delta\tau = \nabla^a\nabla_a\tau$.

The optimal isometric embedding system seeks for a solution $(X, T)$ that satisfies

$$\begin{cases} \langle dX, dX\rangle & = \sigma \\ \nabla^a j_a & = 0. \end{cases} \tag{1}$$

We note that the first equation is the isometric embedding equation into the Minkowski spacetime $\mathbb{R}^{3,1}$. The quasilocal mass is then defined to be

$$E(\Sigma, X, T) = \frac{1}{8\pi}\int_\Sigma \rho. \tag{2}$$

Several remarks are in order:

(1) $\Sigma_0$ is the "unique" surface in the Minkowski spacetime that best matches the physical surface $\Sigma$. If $\Sigma$ happens to be a surface in the Minkowski spacetime, the above procedure identifies $\Sigma_0 = \Sigma$ up to a global isometry.

(2) A prototype form of the quasilocal mass which corresponds to the special case $\tau = 0$ in (2) (due to Brown-York,[5] Liu-Yau,[20] Booth-Mann,[4] Kijowski[19]) is

$$\frac{1}{8\pi}\int_{\Sigma}\left(|\mathbf{H_0}| - |\mathbf{H}|\right).$$

The positivity is proved by Shi-Tam[29] and Liu-Yau.[21] However, for a surface in the Minkowski spacetime, the above expression may not be zero.[22]

The optimal isometric embedding system gives the necessary correction, so that the Wang-Yau definition is positive in general and zero for surfaces in the Minkowski spacetime.[33,34]

The derivation of the Wang-Yau definition (2) relies on both physical theory and mathematical theory. From the Hamilton-Jacobi analysis of the Einstein-Hilbert action (Brown-York,[5] Horowitz-Hawking[17]) a surface Hamiltonian $\mathfrak{H}(\Sigma)$ is obtained and the quasilocal energy should be $\mathfrak{H}(\Sigma) - \mathfrak{H}(\Sigma_0)$ for a reference surface $\Sigma_0$ in the reference Minkowski spacetime. The precise definitions of the surface Hamiltonians still depend on the choices of a normal gauge and a timelike vector field along the surface $\Sigma$ as an observer.

On the other hand, the mathematical theory of isometric embeddings (Nirenberg,[23] Pogorelov[26]) is used to find the reference surface $\Sigma_0$ and a variational approach leads to a canonical gauge that anchors the choices of the gauge of the surface Hamiltonian.

In general, the optimal isometric embedding system is difficult to solve. Suppose $(X, T)$ is a solution and suppose the corresponding $\rho$ is positive, then $E(\Sigma, X, T)$ is a local minimum[8] and the nearby optimal isometric embedding system is solvable by an inverse function theorem argument. In a perturbative configuration, when a family of surfaces limit to a surface in the Minkowski spacetime, then the optimal isometric embedding system is solvable, again subject to the positivity of the limiting mass. This applies to the case of large sphere limits[7] and small sphere limits.[12]

## 3. Quasilocal mass at null infinity

### 3.1. *Large sphere limit at null infinity*

Consider an isolated system surrounding a source. In terms of the Bondi-Sachs coordinate system $(u, r, x^2 = \theta, x^3 = \phi)$,[3,27,31,32] near future null infinity $\mathfrak{I}^+$ the spacetime metric takes the form:

$$-Vdu^2 - 2Ududr + \sigma_{ab}(dx^a + W^a du)(dx^b + W^b du), a, b = 2, 3$$

such that each $V$, $U$, $W^a$ admit expansions in terms of integral powers of $r$ with

$$\begin{aligned} V &= 1 - \frac{2m(u,\theta,\phi)}{r} + O(r^{-2}) \\ U &= 1 + O(r^{-2}) \\ \sigma_{ab}dx^a dx^b &= r^2(d\theta^2 + \sin^2\theta d\phi^2) + 2r(cd\theta^2 - c\sin^2\theta d\phi^2 + 2\underline{c}\sin\theta d\theta d\phi) + O(1) \\ W^a &= O(r^{-2}) \end{aligned}$$

The Bondi-Trautman mass is

$$M_{BT}(u) = \frac{1}{4\pi}\int_{S^2} m(u,\theta,\phi), \tag{3}$$

where $m(u,\theta,\phi)$ is the mass aspect function defined at $\mathfrak{I}^+$. As a result of the vacuum Einstein equation, the mass loss formula states

$$\frac{d}{du}M_{BT}(u) = -\frac{1}{4\pi}\int_{S^2}\left(c_u^2 + (\underline{c}_u)^2\right) \leq 0. \tag{4}$$

This important formula represents the first theoretical verification of gravitational wave/radiation in the nonlinear setting.

In Ref. 7, we evaluate the large sphere limit of quasilocal mass which recovers the Bondi-Trautman mass. At a retarded time $u = u_0$, we consider the family of large spheres $\Sigma_r$ parametrized by an affine parameter $r$. The positivity of the Bondi-Trautman mass guarantees the unique solvability of the optimal isometric embedding system (1) with a solution $(X_r, T_r)$. Suppose $X_r$ and $T_r$ admit expansions:

$$T_r = T^{(0)} + \sum_{k=1}^{\infty} T^{(-k)} r^{-k}$$

$$X_r = rX^{(1)} + X^{(0)} + \sum_{k=1}^{\infty} X^{(-k)} r^{-k},$$

then $T^{(0)}$ is shown to be proportional to the Bondi-Sachs energy-momentum and $T^{(0)}$ being future timelike makes $T^{(-k)}$ and $X^{(-k+1)}$ solvable inductively for $k = 1, 2\cdots$.

### 3.2. *Unit sphere limits at null infinity*

Both the Bondi-Trautman mass and the mass loss formula are global statements about $\mathfrak{I}^+$, i.e. they require the knowledge of all directions of $(\theta,\phi)$. The limit of quasilocal mass introduced in the following provides a quasilocal quantity along a single direction $(\theta,\phi)$ at $\mathfrak{I}^+$.

Consider a null geodesic $\gamma(d)$ that approaches $\mathfrak{I}^+$. Around each point on $\gamma(d)$, consider a geodesic 2-sphere $\Sigma_d$ of unit radius. The geometry of $\Sigma_d$ approaches the geometry of a standard unit round sphere of $\mathbb{R}^3$.

In the limit $d \to \infty$, we obtain two quantities. The first one is $\lim_{d\to\infty} E(\Sigma_d)$ which is of the order of $\frac{1}{d^2}$ with $E(\Sigma_d) \geq 0$. The second one is obtained by exploiting the vanishing of the $\frac{1}{d}$ term and appears as a loop integral on the limiting surface that is of the order of $\frac{1}{d}$.

Several cases have been computed:

(1) Linear gravitational perturbation of the Schwarzschild black hole à la Chandrasekhar. The linearized vacuum Einstein equation is solved by separation of variables and solutions of linearized waves are obtained. The optimal isometric embedding system can be solved and the quasilocal mass can be evaluated by solving

$$\begin{aligned} \Delta(\Delta+2)\tau &= \text{physical data}, \\ (\Delta+2)N &= \text{physical data}, \end{aligned} \tag{5}$$

where $\tau$ and $N$ are functions on the standard 2-sphere and $\Delta$ is the Laplace operator. All distinctive features of the waves such as frequency and mode parameters are recovered.

(2) The Vaidya spacetime [a]

$$-(1-\frac{2m(u)}{r})du^2 - 2dudr + r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

The quasilocal mass of a unit sphere approaching null infinity is

$$E(\Sigma_d) = -\frac{1}{8\pi d^2}\int_{S^2} (\partial_u m)\sin^2(\hat{\theta}) + l.o.t. \geq 0 \tag{6}$$

The positivity of quasilocal mass corresponds to the mass loss formula in the Vaidya case.

One may expect that the limit of quasilocal mass in the direction of $(\theta_0, \phi_0)$ will recover the value of the mass aspect function $m(u, \theta_0, \phi_0)$. But notice that the mass aspect function is not pointwise positive, only the integrated Bondi-Trautman mass is positive by Schoen-Yau [28] and Horowitz-Perry. [18] In the Vaidya case, the mass aspect is recovered from the loop integral.

## 4. Unit sphere limit in BVM coordinate near null infinity

The Vaidya spacetime is a model for the Einstein-null dust system. In particular, the condition $-\partial_u m \geq 0$ in (6) can be interpreted as local energy condition for matters. In order to study the contribution of quasilocal mass at the purely gravitational level, we study the null infinity of a vacuum spacetime. We discuss the case of an axi-symmetric spacetime in the Bondi- van der Burg-Metzner coordinates near null infinity. The leading terms of the spacetime metric are of the form:

$$-(1-\frac{2\mathbf{M}}{r})du^2 - 2dudr - 2\mathbf{U}dud\theta + (r^2 + 2r\mathbf{C})d\theta^2 + (r^2 - 2r\mathbf{C})\sin^2\theta d\phi^2,$$

where $\mathbf{M}, \mathbf{U}, \mathbf{C}$ are functions of $u$ and $\theta$.

[a] The mass aspect function used here differs to that of Ref. 13 by a factor of 2.

The $\frac{1}{d^2}$ term of the quasilocal mass of a unit sphere approaching null infinity is (up to a constant factor)

$$\int_{B^3}[(\partial_u \mathbf{C})^2 + \det(h_0^{(-1)} - h^{(-1)})] + \frac{1}{4}\int_{S^2}[(tr_\Sigma k^{(-1)})^2 - \tau^{(-1)}\tilde{\Delta}(\tilde{\Delta}+2)\tau^{(-1)}], \quad (7)$$

in which $h^{(-1)}$ and $k^{(-1)}$ depend on the physical data and $h_0^{(-1)}$ and $\tau^{(-1)}$ depend on the solution of the optimal isometric embedding system. A priori, the expression may depend on all $\mathbf{M}, \mathbf{U}$, and $\mathbf{C}$. However all occurrences of $\mathbf{M}$ and $\mathbf{U}$ are cancelled and the final answer only involves the function $\mathbf{C}$ and is completely independent of the mass aspect function $\mathbf{M}$. The expression (7) is manifestly positive by Wang-Yau's theorem.

### 4.1. *Final remark*

Given a spacetime surface $(\Sigma, \sigma, \mathbf{H})$, we find the reference surface $(\Sigma_0, \sigma, \mathbf{H_0})$ through the optimal isometric embedding equation and evaluate the quasilocal mass. This is a nonlinear and coordinate independent theory. The procedure is canonical and is accompanied by a uniqueness statement. In particular, the definition does not involve any ad hoc referencing or normalization. The calculation of the quasilocal mass does not assume any a priori knowledge of null infinity.

The positivity of the unit sphere limit should correspond to a "quasilocal" mass loss formula, or "quasilocal gravitational radiation" at null infinity. The evaluation in the full generality of Bondi-Sachs coordinates (without assuming axi-symmetry) will appear in an upcoming paper.

## References


1. R. Bartnik, *Quasi-spherical metrics and prescribed scalar curvature*, J. Differential Geom. **37** (1993), no. 1, 31–71.
2. R. Bartnik, *New definition of quasi-local mass*, Phys. Rev. Lett. **62** (1989), no. 20, 2346–2348.
3. H. Bondi, M. G. J. van der Burg, and A. W. K. Metzner, *Gravitational waves in general relativity. VII. Waves from axi-symmetric isolated systems*, Proc. Roy. Soc. Ser. A 269 (1962) 21–52.
4. I. S. Booth and R. B. Mann, *Moving observers, nonorthogonal boundaries, and quasilocal energy*, Phys. Rev. D (3) **59** (1999), 064021.
5. J. D. Brown and J. W. York, *Quasi-local energy and conserved charges derived from the gravitational action,* Phys. Rev. D (3) **47** (1993), no. 4, 1407–1419.
6. S. Chandrasekhar, *The mathematical theory of black holes*, reprint of the 1992 edition, Oxford Classic Texts in the Physical Sciences, Oxford Univ. Press, New York.
7. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Evaluating quasi-local energy and solving optimal embedding equation at null infinity*, Comm. Math. Phys. **308** (2011), no.3, 845–863.

8. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Minimizing properties of critical points of quasi-local energy*, Comm. Math. Phys. **329** (2014), no.3, 919–935
9. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Conserved quantities in general relativity: from the quasi-local level to spatial infinity*, Comm. Math. Phys. **338** (2015), no.1, 31–80.
10. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Quasi-local energy in presence of gravitational radiation*, Int. J. Mod. Phys. D 25, 164501 (2016).
11. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Quasi-local mass in the gravitational perturbations of black holes,* in preparation.
12. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Evaluating small sphere limit of the Wang-Yau quasi-local energy*, Comm. Math. Phys. **357** (2018), no. 2, 731–774
13. P.-N. Chen, M.-T. Wang, and S.-T. Yau, *Quasi-local mass at the null infinity of the Vaidya spacetime*, Nonlinear analysis in geometry and applied mathematics, 33–48, Harv. Univ. Cent. Math. Sci. Appl. Ser. Math., 1, Int. Press, Somerville, MA, 2017
14. D. Christodoulou, *Nonlinear nature of gravitation and gravitational-wave experiments*, Phys. Rev. Lett. **67** (1991), no. 12, 1486–1489.
15. A. J. Dougan and L. J. Mason, *Quasilocal mass constructions with positive energy*, Phys. Rev. Lett., **67** (1991), no. 16, 2119–2122.
16. S. W. Hawking, *Gravitational radiation in an expanding universe*, J. Math. Phys. **9**, 598 (1968).
17. S. W. Hawking and G. T. Horowitz, *The gravitational Hamiltonian, action, entropy and surface terms*, Classical Quantum Gravity **13** (1996), no. 6, 1487–1498.
18. G. T. Horowitz and M. J. Perry, *Gravitational energy cannot become negative*, Phys. Rev. Lett. **48** (1982), no. 6, 371–374.
19. J. Kijowski, *A simple derivation of canonical structure and quasi-local Hamiltonians in general relativity,* Gen. Relativity Gravitation **29** (1997), no. 3, 307–343.
20. C.C. M. Liu and S.T. Yau, *Positivity of quasilocal mass,* Phys. Rev. Lett. **90** (2003), no. 23, 231102.
21. C.-C. M. Liu and S.-T. Yau, *Positivity of quasi-local mass II,* J. Amer. Math. Soc. **19** (2006), no. 1, 181–204.
22. N. Ó Murchadha, L. B. Szabados and K. P. Tod, Comment on: "Positivity of quasi-local mass" Phys. Rev. Lett. **92** (2004), no. 25, 259001, 1 p.
23. L. Nirenberg, *The Weyl and Minkowski problems in differential geometry in the large*, Comm. Pure Appl. Math. 6 (1953), 337–394.
24. R. Penrose, *Some unsolved problems in classical general relativity*, Seminar on Differential Geometry, pp. 631–668, Ann. of Math. Stud., 102, Princeton Univ. Press, Princeton, N.J., 1982.
25. R. Penrose, *Quasi-local mass and angular momentum in general relativity*, Proc. Roy. Soc. London Ser. A **381** (1982), no. 1780.

26. A. V. Pogorelov, *Regularity of a convex surface with given Gaussian curvature,* (Russian) Mat. Sbornik N.S. **31** (73), (1952), 88–103.
27. R. K. Sachs, *Gravitational waves in general relativity, VIII. Waves in asymptotically flat space-time.* Proc. Roy. Soc. Ser. A 270 1962 103–126.
28. R. Schoen and S.-T. Yau, *Proof that the Bondi mass is positive*, Phys. Rev. Lett. **48** (1982), no. 6, 369–371.
29. Y. Shi and L.-F. Tam, *Positive mass theorem and the boundary behaviors of compact manifolds with nonnegative scalar curvature*, J. Differential Geom. **62** (2002), no. 1, 79–125.
30. K. P. Tod, Penrose's quasi-local mass, in *Twistors in mathematics and physics*, 164–188, London Math. Soc. Lecture Note Ser., 156, Cambridge Univ. Press, Cambridge.
31. A. Trautman, *Boundary conditions at infinity for physical theories*, Bull. Acad. Polon. Sci. 6 (1958), 403–406; reprinted as arXiv:1604.03144.
32. A. Trautman, *Radiation and boundary conditions in the theory of gravitation*, Bull. Acad. Polon. Sci., 6 (1958), 407–412; reprinted as arXiv:1604.03145.
33. M.-T. Wang, and S.-T. Yau, *Quasi-local mass in general relativity*, Phys. Rev. Lett. **102** (2009), no. 2, no. 021101.
34. M.-T. Wang, and S.-T. Yau, *Isometric embeddings into the Minkowski space and new quasi-local mass*, Comm. Math. Phys. **288** (2009), no. 3, 919–942.

# Black hole entropy from soft hair

Malcolm J. Perry

*Department of Applied Mathematics and Theoretical Physics,*
*Centre for Mathematical Sciences,*
*University of Cambridge,*
*Wilberforce Road,*
*Cambridge CB3 0WA,*
*UK.*
**E-mail: malcolm@damtp.cam.ac.uk*

We start by looking at why we believe that black holes have entropy. According to Boltzmann, the entropy is a measure of the number of microstates of a system. We suggest here that the entropy arises from a holographic conformal field theory on the black hole horizon. Finally, we discuss some of the implications for the information paradox.

*Keywords*: Black holes. Thermodynamics, Holography. Information Paradox.

## 1. Introduction

Black holes were first thought about by John Michell back in 1784.[1] He reasoned that if the escape velocity from an object like a star exceeded the speed of light, then it would give rise to an object that cannot be seen optically but whose gravitational field would betray its existence. These ideas were given substance by Einstein's general theory of relativity and the subsequent discovery of solutions of the field equations that represented stationary black holes.[2–6] One puzzling feature of stationary black holes is that they are completely characterised by just a few parameters; their mass $M$, angular momentum $\mathbf{J}$ and electric charge $Q$.[7–11] It is this observation that is the basis for the information paradox.[12]

The thermodynamics of black holes in general relativity has a history starting in 1972. The first relevant discovery, the area theorem, was made by Hawking.[13] He found that the area of a black hole horizon could never decrease provided the null convergence condition $R_{ab}k^ak^b \geq 0$ holds for every null vector $k^a$. Shortly after this, Jacob Bekenstein suggested[14] that the entropy of a black hole must be proportional to the area of its event horizon. His reasoning was based on three observations. The first was that if a Kerr-Newman black hole increased its mass by an amount $dM$, then

$$dM = \frac{\kappa dA}{8\pi} + \Phi dQ + \mathbf{\Omega}\cdot d\mathbf{J} \tag{1}$$

where $\kappa$ is the black hole surface gravity, $A$ the area of the event horizon, $\Phi$ the electrostatic potential of the black hole and $\mathbf{\Omega}$ its angular velocity. $\kappa$, $A$, $\Phi$ and $\mathbf{\Omega}$ are all determined in terms of $M$, $\mathbf{J}$ and $Q$. The second piece of evidence was that he reasoned that a black hole must have some kind of internal structure that

resulted from it method of formation. That would give rise to an entropy

$$S = -\sum_n p_n \ln p_n \tag{2}$$

arising from the probability of the occupation of the $n^{th}$-state being $p_n$. Finally, he showed that it was necessary for this black hole entropy to be added to the thermodynamic entropy of the rest of the universe in order to have a consistent theory of thermodynamics. This came about because otherwise dropping a box of radiation into a black hole would cause the entropy of the universe to decrease, in contradiction to the second law of thermodynamics.

His ideas were met with a certain amount of scepticism because black holes were thought to have vanishing temperature. Despite that, Bardeen, Carter and Hawking[15] pointed out the similarities between the first law of thermodynamics and (1) and also the second law of thermodynamics and the area theorem.

In 1974, Hawking[16,17] showed that black holes had a temperature $T_H$ of $\hbar\kappa/(2\pi)$. Unlike previous work, his calculation was quantum mechanical in nature. Black holes would emit particles with a thermal spectrum at a temperature given by $T_H$. By identifying (1) with the first law of thermodynamics, one can immediately infer that the entropy must be given by $A/(4\hbar)$. The area theorem is thereby identified with the second law of thermodynamics. A somewhat different view of entropy was taken in Ref. 18. The idea here was to use the path integral for gravity to derive black hole entropy. Although gravity is unrenormalizable, there is no obstacle to using the path integral to lowest order as the uncontrollable divergences only occur at one loop or beyond. The action for pure gravity is, including the Gibbons-Hawking-York boundary terms[18,19]

$$I[g,h] = \frac{1}{16\pi}\int_{\mathcal{M}} R(g)\sqrt{\| g \|}d^4x + \frac{1}{8\pi}\int_{\partial\mathcal{M}} K\sqrt{\| h \|}d^3x + C[h] \tag{3}$$

where now $\mathcal{M}$ is the spacetime manifold with metric $g$ and Ricci scalar $R(g)$. The boundary of $\mathcal{M}$ is $\partial\mathcal{M}$ with metric $h$ and second fundamental form $K$. $C[h]$ is any functional of $h$ and is designed to make the action of flat spacetime vanish. Suppose one wants to find the partition function for a black hole spacetime. Then one wants to compute $Z = \mathrm{tr}(e^{-\beta\mathcal{H}})$ where $\beta$ is the inverse temperature and $\mathcal{H}$ the Hamiltonian. This can be done by realising that $e^{i\mathcal{H}t}$ is the time evolution operator and so if $t$ is identified with $t + i\beta$ then $Z$ is given by

$$Z = \int D[g]e^{-I[g,h]/\hbar}, \tag{4}$$

where now the integral is over all metrics $g$ of positive definite signature and that approach flat space at infinity and are periodic in imaginary time $t$ with period $\beta = T_H^{-1}$.

The Schwarzschild metric is

$$ds^2 = -(1 - \frac{2M}{r})dt^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2 d\Omega^2\,, \tag{5}$$

where $M$ is the mass of the static black hole with horizon at $r = 2M$ and $d\Omega^2$ is the metric on the unit 2-sphere. Taking $t = i\tau$ so that the geometry is as described above gives the Euclidean metric

$$ds^2 = (1 - \frac{2M}{r})d\tau^2 + (1 - \frac{2M}{r})^{-1}dr^2 + r^2 d\Omega^2 \tag{6}$$

Now, $r = 2M$ is a conical singularity that is resolved provided that $\tau$ is identified with period $8\pi M$.[20] This periodicity is precisely the same periodicity expected from the Hawking calculation of the black hole temperature. Interpreting the exponential of the action as the partition function, reproduces the black hole entropy. In this calculation, the region $r < 2M$ has been removed from consideration. Implicitly this means that the internal degrees of freedom have been traced over. One is thereby led to believe that the black hole does have some kind of internal structure that cannot be probed by external observers who just look at the classical geometry. The same kind of reasoning can be applied to the Kerr-Newman metric too.

Black holes evaporate. The black hole uniqueness theorems suggest that the only properties that a stationary black hole has are just the mass, charge and spin. As a consequence, there is a tension with the ideas of quantum mechanics. If a black hole completely disappears, then the final state should be unitarily equivalent to the initial state. Obviously, there are enormous number of ways in which the black hole could form. The black hole, once it has settled down to a more or less equilibrium state, is described by just those three parameters. The Hawking radiation is thermal and characterised by the Hawking temperature. Such a final state consisting of Hawking radiation will not be unitarily related to the initial state that gave rise to the black hole. This is the information paradox. It might be that quantum mechanical information really is lost in gravitational collapse. But then, the whole edifice of quantum mechanics would need to be rethought. The incredible success of quantum mechanics would seem to discourage such a viewpoint. Alternatively, there might be something wrong with the uniqueness theorems. It is this latter possibility that we will investigate here.

In what follows, we will use covariant phase space methods[21–27] to understand the nature of charges in general relativity and the consequences for the physics of black holes. The reason for using the covariant phase space method is to preserve as much as possible of the covariance of theory. Had we picked the more conventional canonical methods, we would be forced to pick a particular time coordinate which would obscure matters. Furthermore, it would it impossible to understand what happens on null surfaces such as the event horizon. Our aim now is to try to understand something about the microscopic origin of black hole entropy. The hope is that this will aid a resolution of the information paradox.

In pure general relativity, one can start with the Einstein-Hilbert action $I$ given by

$$I = \frac{1}{16\pi}\int_{\mathcal{M}} R(g)\sqrt{\| g \|}d^4x. \tag{7}$$

This action omits the boundary terms, but these are not germane to the discussion that follows. One finds the Einstein equation by performing a variation of the action induced by a variation of the metric $g_{ab} \to g_{ab} + h_{ab}$. This results in the variation $\delta I$ given by

$$\delta I = \int_{\mathcal{M}} \; (R_{ab} - \frac{1}{2} R \, g_{ab}) \; h^{ab} \, \sqrt{\| \; g \; \|} d^4x + \int_{\partial \mathcal{M}} \theta \tag{8}$$

The three-form on the boundary $\theta(g, h)$ is known as the presymplectic potential and has components

$$(*\theta)_a = \frac{1}{16\pi} \left( \nabla_b h^b_a - \nabla_a h \right) \tag{9}$$

where $h = h_{ab} g^{ab}$. In canonical general relativity, the boundary term would be thought of

$$\int_{\partial \mathcal{M}} \; \sum_i \; p_i \delta q^i \tag{10}$$

where $q^i$ are the generalised coordinates, $p_i$ are the generalised momenta and $i$ represent the tensor indices of these fields.

The presymplectic density $\omega(g; h, h')$ is defined by a second variation $g_{ab} \to g_{ab} + h'_{ab}$

$$\omega(g; h, h') = \delta\theta(g, h') - \delta'\theta(g, h) \tag{11}$$

Finally, the symplectic form for general relativity is

$$\Omega_\Sigma = \int_\Sigma \omega \tag{12}$$

where $\Sigma$ is any partial Cauchy surface in the spacetime. In the language of the canonical theory, $\Omega_\Sigma$ would be

$$\int_\Sigma \; \sum_i \; \delta p_i \wedge \delta q^i. \tag{13}$$

One property that $\omega$ has is that if the background metric $g_{ab}$ obeys the Einstein equation and both $h_{ab}$ and $h'_{ab}$ obey the linearised Einstein equations, then $\omega$ is closed. Thus $\Omega_\Sigma$ is constant under variations of $\Sigma$ as long as the boundaries of $\Sigma$ are fixed.

In general relativity, the symmetry group is the group of diffeomorphisms. An infinitesimal coordinate transformation is specified by a vector field $\zeta^a$. This induces a variation in any tensor field given by the Lie derivative of that field. Thus, for example, the variation of the metric is given by

$$\delta g_{ab} = \mathcal{L}_\zeta g_{ab} = g_{ab} + \nabla_a \zeta_b + \nabla_b \zeta_a. \tag{14}$$

The bulk term in the variation of the action $\delta I$ is invariant under such a transformation but the boundary term is not. The infinitesimal co-ordinate transformations obey an algebra whose composition law is

$$\mathcal{L}_\zeta \mathcal{L}_\eta - \mathcal{L}_\eta \mathcal{L}_\zeta = \mathcal{L}_{[\zeta, \eta]} \tag{15}$$

where $\zeta$ and $\eta$ are two (smooth) vector fields and

$$[\zeta, \eta] = \mathcal{L}_\zeta \eta = -\mathcal{L}_\eta \zeta \tag{16}$$

Suppose now that in $\Omega$ one makes $h'_{ab}$ a gauge transformation given by the vector field $\zeta$. Then $\Omega$ can be written as a boundary integral. Explicitly,

$$Q_\zeta = \frac{1}{16\pi} \int_{\partial\Sigma} F_{ab}\ dS^{ab} \tag{17}$$

with

$$F_{ab} = -2\zeta_{[a}\nabla_{b]}h + 2\zeta_{[a}\nabla^c h_{b]c} - 2\zeta^c \nabla_{[a} h_{b]c} - h\nabla_{[a}\zeta_{b]} + 2h_{c[a}\nabla^c \zeta_{b]}. \tag{18}$$

Let $\partial\Sigma$ is a closed 2-surface $S$, for example the celestial sphere or a black hole event horizon. One would like to interpret $Q_\zeta$ as the variation in the Noether charge conjugate to $\zeta$ that is enclosed in the interior of $S$ as one moves between the metric $g_{ab}$ and $g_{ab} + h_{ab}$. There is a complication with this idea because in such a change, there might be a flux of charge crossing $S$. To take account of this possibility, one needs to examine $Q_\zeta$ and identify such terms and subtract them out. In more mathematical language, one we want $Q_\zeta$ to be a function of state. As such it must be a 1-form on the infinite-dimensional phase space of the theory. This 1-form needs to be exact so that if one goes along a path $\Gamma$ between $g_{ab}$ and $g_{ab} + h_{ab}$, then $Q_\zeta$ is independent of the path $\Gamma$, and therefore dependent only the end-points of that path. The definition on $Q_\zeta$ thus needs to be modified by the ddition of a suitable counterterm $Q_\zeta \to Q_\zeta + Q^{ct}_\zeta$. Finding $Q^{ct}_\zeta$ needs to be done on a case by case basis as has been elegantly explained in detail by Wald and Zoupas.[26]

In the case that $\zeta$ were a time translation, then $Q_\zeta$ would be the quasi-local mass enclosed in $S$.[28] If it were a spatial translation then the momentum. If $\zeta$ were a Killing vector, then $Q_\zeta$ would be the same as the Komar integral.[29] If $\zeta$ were a supertranslation or super-rotation at null infinity, then $Q_\zeta$ would be the corresponding supertranslation or super-rotation charge. Equally, one can define charges on the black hole horizon and these are the soft charges or soft black hole hair.[30]

Diffeomorphism invariance of general relativity means that the charges $Q_\zeta$ lie in some representation of group of coordinate transformations. Thus

$$\delta_\zeta Q_\eta - \delta_\eta Q_\zeta = Q_{[\zeta,\eta]} \tag{19}$$

Were this relation not to hold, general coordinate invariance would be violated, in gross contradiction to our expectations of what should be true in physics. However, what we find is that this relationship does not hold for charges on black hole event horizons. Instead, we find

$$\delta_\zeta Q_\eta - \delta_\eta Q_\zeta = Q_{[\zeta,\eta]} + K(\zeta, \eta) \tag{20}$$

where $K(\zeta, \eta)$ is a central extension of this algebra.[31] We will now explore a particular example and move on to its interpretation.

We start from the Kerr metric in Boyer-Lindquist coordinates.

$$ds^2 = \rho^2(\frac{dr^2}{\Delta} + d\theta^2) + (r^2 + a^2)\sin^2\theta d\phi^2 - dt^2 + \frac{2Mr}{\rho^2}(a\sin^2\theta d\phi - dt)^2 \quad (21)$$

where

$$\Delta = r^2 - 2Mr + a^2 \quad (22)$$

and

$$\rho^2 = r^2 + a^2\cos^2\theta. \quad (23)$$

$M$ is the mass of the black hole and $J = Ma$ is its angular momentum. $\Delta = 0$ at $r_\pm$ with $r_+$ being the location of the outer horizon, $r_-$ the location of the inner horizon and

$$r_\pm = M \pm \sqrt{M^2 - a^2} \quad (24)$$

Now we will define "conformal" coordinates[32] and assume that the black hole is not extreme so that $m^2 > a^2$.

$$w^+ = \sqrt{\frac{r - r_+}{r - r_-}}\, e^{2\pi T_R \phi} \quad (25)$$

$$w^- = \sqrt{\frac{r - r_+}{r - r_-}}\, e^{2\pi T_L \phi - t/2M} \quad (26)$$

$$y = \sqrt{\frac{r_+ - r_-}{r - r_-}}\, e^{\pi(T_L + T_R)\phi - t/4M} \quad (27)$$

where

$$T_L = \frac{r_+ + r_-}{4\pi a} \quad \text{and} \quad T_R = \frac{r_+ - r_-}{4\pi a} \quad (28)$$

The future outer horizon is $w^- = 0$ and the past outer horizon is $w^+ = 0$. The azimuthal coordinate $\phi$ is identified with period $2\pi$ and this induces an identification on $w^+, w^-$ and $y$ as

$$w^+ \to e^{4\pi^2 T_R} w^+, \quad w^- \to e^{4\pi^2 I_L} w^-, \quad y \to e^{2\pi^2(T_L + T_R)} y. \quad (29)$$

The line element close to the horizon bifurcation surface $w^+ = w^- = 0$ is[33]

$$ds^2 = \frac{4\rho_+^2}{y^2} dw^+ dw^- + \frac{16M^2a^2\sin^2\theta}{\rho_+^2 y^2} dy^2 + \rho_+^2 d\theta^2 + O(w^+, w^-) \quad (30)$$

where $\rho_+^2 = r_+^2 + a^2\cos^2\theta$. If one looks at the $w^+, w^-, y-$ plane by setting $\theta$ to be a constant, then this line element is that of $AdS_3/\Gamma$ with $\Gamma$ being some discrete group. So close to the horizon bifurcation surface, the geometry of spacetime is some kind of warped product of the line segment $\theta \in [0, \pi]$ and a deformed portion of three-dimensional anti-de Sitter space. The classic work of Brown and Henneaux[31] shows

that in spacetimes that are asymptotic to anti-de Sitter spacetime, the diffeomorphism algebra has anomalies. One might therefore suspect that something similar happens in the case of the Kerr black hole.

Consider the diffeomorphism given by the vector field $\zeta_n$

$$\zeta_n = \epsilon_n(w^+)\partial_+ + \frac{1}{2}\epsilon'(w^+)y\partial_y \tag{31}$$

with

$$\epsilon_n(w^+) = 2\pi T_R(w^+)^{(1+\frac{in}{2\pi T_R})} \tag{32}$$

and $n$ being any integer. It should be noted that under the identifications of either $w^+ \to w^+ e^{4\pi^2 T_R}$ or $y \to y e^{2\pi^2(T_L+T_R)}$ that $\zeta_n$ is invariant. This vector field is well-defined on the future horizon. These vector fields obey the Witt (or centerless Virasoro) algebra with the commutator

$$[\zeta_n, \zeta_m] = i(m-n)\zeta_{n+m} \tag{33}$$

Similarly, one can find a second vector field $\tilde{\zeta}_n$ given by

$$\tilde{\zeta}_n = \tilde{\epsilon}_n(w^-)\partial_- + \tfrac{1}{2}\tilde{\epsilon}'(w^-)y\partial_y \tag{34}$$

and $\tilde{\epsilon}$ being given

$$\tilde{\epsilon}_n(w^-) = 2\pi T_L(w^-)^{(1+\frac{in}{2\pi T_L})} \tag{35}$$

again with $n$ being any integer. This vector field is well-defined on the past horizon. Again, it is invariant under the identifications $w^- \to w^- e^{4\pi^2 T_L}$ or $y \to y e^{2\pi^2(T_L+T_R)}$. It too obeys the Virasoro algebra

$$[\tilde{\zeta}_n, \tilde{\zeta}_m] = i(m-n)\tilde{\zeta}_{n+m}. \tag{36}$$

Both of these vector fields are well-defined on the bifurcation surface and commute with each other there

$$[\zeta_n, \tilde{\zeta}_m] = 0. \tag{37}$$

These vector fields can be used to generate charges on the bifurcation surface. To do this we need to introduce an appropriate counterterm. This is given by

$$-\frac{1}{8\pi}\int dS^{ab}\ \nabla_a(\zeta^c h^d_b)N_{cd} \tag{38}$$

where $N_{ab}$ are the components of the volume form on the normal bundle to the horizon. There is a precisely similar expression for the fields $\tilde{\zeta}_n$. One then finds that the charges on the bifurcation surface obey the algebra

$$[Q_n, Q_m] = i(n-m)Q_{n+m} + in^3 J\delta_{n,-m} \tag{39}$$

for the right-handed algebra and

$$[\tilde{Q}_n, \tilde{Q}_m] = i(n-m)\tilde{Q}_{n+m} + in^3 J\delta_{n,-m} \tag{40}$$

for the left-handed algebra. Finally, the left and right algebras commute with each other

$$[Q_n, \tilde{Q}_m] = 0. \tag{41}$$

In both cases, the central terms shown here correspond to the conventionally normalised Virasoro algebra with central charges given by $c_L = c_R = 12J$. Thus the diffeomorphism algebra has an anomaly.

We postulate that this anomaly is cancelled by holographic degrees of freedom on the horizon expressed in terms of a two-dimensional conformal field theory. Consider for a moment the expressions for the absorption probabilities for particles incident on a Kerr black hole. Suppose we look at a particle with energy $\delta E$ and angular momentum parallel to the black hole spin $\delta J$. Then we observe that the absorption probability obtains a suggestive factor of

$$|\Gamma(1 + \frac{i\omega_L}{2\pi T_L})|^2 \ |\Gamma(1 + \frac{i\omega_R}{2\pi T_R})|^2 \tag{42}$$

where

$$\omega_L = \frac{2M^3}{J}\delta E \quad \omega_R = \frac{2M^3}{J}\delta E - \delta J. \tag{43}$$

This is precisely what is to be expected for a conformal field theory where the left-handed degrees of freedom are at a temperature of $T_L$ and the right-handed degrees of freedom are at a temperature of $T_R$ and one is asking for the absorption probability for particles of energy $\omega_L$ in the left-handed sector and energy $\omega_R$ in the right-handed sector. We take it that there are no coincidences in nature and therefore we really can attribute our observations to the existence of holographic degrees of freedom on the horizon described by a two-dimensional conformal field theory.

A general property of conformal field theories, provided the central charge is sufficiently large, was first described by Cardy.[34] The entropy for a system with central charges $c_L$ and $c_R$ for the two sectors at temperatures $T_L$ and $T_R$ is given by

$$S = \frac{\pi^2}{3}(c_L T_L + c_R T_R). \tag{44}$$

Plugging in our expressions for $c_L, c_R, T_l$ and $T_R$ gives

$$S = \frac{1}{4}A. \tag{45}$$

It is hard to believe that this is a coincidence. It appears therefore we have identified the degrees of freedom responsible for black hole entropy.[33]

Subsequent to the conference, it has been shown that the same methods reproduce the entropy for the Kerr-Newman family of black holes[35] and for uncharged black holes in anti-de Sitter spacetime.[36]

A key question is to ask how this affects our view of the information paradox. We have shown how to account for black hole entropy in terms of a holographic two-dimensional conformal field theory living on the black hole horizon. It is however far from clear that the states of such a theory can record all of the quantum mechanical information that is pertinent to black hole formation from ordinary matter. We are therefore left with a collection of problems that need exploration and solution before there can be any claim of solving the information paradox. We conclude this essay with a summary of outstanding issues. Does the horizon conformal field theory contain a complete description of the black hole formation process? How does the Hawking radiation encode this information so as to preserve unitary time evolution? Why is it that the black hole entropy is independent of the spectrum of elementary particles when the number of ways a black hole can be formed is highly dependent on that spectrum. For example, if there were a million different species of electron, the number of ways a black hole could form would be vastly higher than if there a single type of electron. Nevertheless, the Hawking entropy would be same.

Suppose a particle falls into a black hole. Classically, a co-moviong observer sees it pass through the horizon without anything obvious happening. In the case of a Schwarzschild black hole, it will reach the singularity in a finite amount of proper time. The singularity is a boundary of spacetime and so we believe the particle to have disappeared. In the case of rotating black holes, it seems plausible that it will also inevitably reach a singularity as the inner horizon of a Kerr black hole is unstable and is presumed to become singular once any energy-momentum arrives there. However, if the particle is to leave an imprint on the state of the horizon conformal field theory, it appears to have violated the quantum no-cloning theorem. Roughly speaking, the no-cloning theorem says that you cannot duplicate the state of a particle by unitary time evolution. A number of technical assumptions go into this amongst which is a notion of locality, a dubious assumption in the case of gravitation.

Then there are some more challenging issues. What happens to the singularity? It is a classical concept and shows that classical general relativity is an incomplete theory. What happens quantum mechanically? There is no satisfactory answer at present. What are the final stages of black hole evaporation? The picture presented seems to suggest that all symmetries in nature are gauge symmetries and not global symmetries. For example in the standard model, baryon number is a global symmetry, but it is hard to see how this could be encoded in the picture presented here. There is one ambitious theory that predicts that all symmetries are gauge symmetries and that is string theory. Although string theory is successful in resolving the divergence problems of quantum gravity, and potentially geometrizing the spectrum of elementary particles, it is far from being a theory of spacetime. Hopefully, the picture here will provide a guide to the true nature of quantum gravity, but there are immense and exciting challenges to the construction of such a theory. Eventually, we hope that the construction of such a theory will lead to deep insights into the nature of our Universe.

**Acknowledgements**

The work presented here was joint with the late Stephen Hawking, Andy Strominger and Sasha Haco. We hope that it is a fitting memory to Stephen who was a dear friend for almost forty years.

We are grateful to Sangmin Choi, Geoffrey Comp‘ere, Peter Galison, Monica Guica, Dan Harlow, Roy Kerr, Alex Lupsasca, Juan Maldacena, Alex Maloney, Suvrat Raju and Maria Rodriguez for useful conversations. This paper was supported in part by DOE de-sc0007870, the John Templeton Foundation, the Black Hole Initiative at Harvard, the Radcliffe Institute for Advanced Study and the UK STFC. We would also like to thank the George and Cynthia Mitchell Foundation, whose generous support allowed for over a decade of scientific workshops, spanning Cooks' Branch in Texas and Great Brampton House and Brinsop Court in England, without which this collaboration would not have been possible. The work began at a Mitchell Foundation retreat and was completed at another. We are profoundly grateful to the Foundation and personally to Sheridan Lorenz and George Mitchell for their contribution to the development of fundamental science.

**References**

1. J. Michell, "On the Means of Discovering the Distance, Magnitude, of the Fixed Stars, in Consequence of the Diminution of the Velocity of Their Light, in Case Such a Diminution Should be Found to Take Place in any of Them, and Such Other Data Should be Procured from Observations, as Would be Farther Necessary for That Purpose.," Phil. Trans. Roy. Soc. Lond. **74** (1784), 35-57. doi:10.1098/rstl.1784.0008
2. K. Schwarzschild, "On the gravitational field of a mass point according to Einstein's theory," Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.) (1916), 189-196. [arXiv:physics/9905030 [physics]].
3. H. Reissner, "Über die Eigengravitation des elektrischen Feldes nach der Einsteinschen Theorie". Annalen der Physik (1916), 50 (9): 106–120. doi:10.1002/andp.19163550905.
4. G. Nordström, "On the Energy of the Gravitational Field in Einstein's Theory". Verhandl. Koninkl. Ned. Akad. Wetenschap., Afdel. Natuurk., Amsterdam. (1918) 26: 1201–1208.
5. R. P. Kerr, "Gravitational field of a spinning mass as an example of algebraically special metrics," Phys. Rev. Lett. **11** (1963), 237-238. doi:10.1103/PhysRevLett.11.237
6. E. T. Newman, R. Couch, K. Chinnapared, A. Exton, A. Prakash and R. Torrence, J. Math. Phys. **6** (1965), 918-919 doi:10.1063/1.1704351
7. D. C. Robinson, "Uniqueness of the Kerr black hole," Phys. Rev. Lett. **34** (1975), 905-906 doi:10.1103/PhysRevLett.34.905
8. P. O. Mazur, "Proof of Uniqueness of the Kerr-Newman Black Hole Solution," J. Phys. A **15** (1982), 3173-3180 doi:10.1088/0305-4470/15/10/021

9. B. K. Harrison, K. S. Thorne, M. Wakano, and J. A. Wheeler, "Gravitation Theory and Gravitational Collapse (Univ. of Chicago Press, Chicago, 1965).
10. J. D. Bekenstein, "Nonexistence of baryon number for static black holes," Phys. Rev. D **5** (1972), 1239-1246 doi:10.1103/PhysRevD.5.1239
11. J. D. Bekenstein, "Nonexistence of baryon number for black holes. ii," Phys. Rev. D **5** (1972), 2403-2412 doi:10.1103/PhysRevD.5.2403
12. S. W. Hawking, "Breakdown of Predictability in Gravitational Collapse," Phys. Rev. D **14** (1976), 2460-2473 doi:10.1103/PhysRevD.14.2460
13. S.W. Hawking and J.B. Hartle, "Energy and angular momentum flow into a black hole," Commun. Math. Phys. **27** (1972), 283-290 doi:10.1007/BF01645515.
14. J.D. Bekenstein,"Black holes and entropy," Phys. Rev. D **7** (1973), 2333-2346 doi:10.1103/PhysRevD.7.2333.
15. J. M. Bardeen, B. Carter and S. Hawking, "The Four laws of black hole mechanics," Commun. Math. Phys. **31** (1973), 161-170 doi:10.1007/BF01645742
16. S.W. Hawking, "Particle Creation by Black Holes," Commun. Math. Phys. **43** (1975), 199-220 doi:10.1007/BF02345020
17. S.W. Hawking, "Black hole explosions," Nature **248** (1974), 30-31 doi:10.1038/248030a0
18. G.W. Gibbons and S.W. Hawking, "Action Integrals and Partition Functions in Quantum Gravity," Phys. Rev. D **15** (1977), 2752-2756 doi:10.1103/PhysRevD.15.2752
19. J. W. York, Jr., "Role of conformal three geometry in the dynamics of gravitation," Phys. Rev. Lett. **28** (1972), 1082-1085 doi:10.1103/PhysRevLett.28.1082
20. G. W. Gibbons and M. J. Perry, "Black Holes and Thermal Green's Functions," Proc. Roy. Soc. Lond. A **A358** (1978), 467-494 doi:10.1098/rspa.1978.0022
21. R. E. Peierls, "The Commutation laws of relativistic field theory," Proc. Roy. Soc. Lond. A **A214** (1952), 143-157 doi:10.1098/rspa.1952.0158
22. G. J. Zuckerman, "Action Principles and Global geometry," in "Mathematical Aspects of String Theory," ed S. T. Yau, World Scientific, Singapore, 1986. doi:10.1142/9789812798411.0013
23. C. Crnkovic and E. Witten, "Covariant Description of Canonical Formalsim in Geoemtrical Theories," in eds S. W. Hawking and W. Israel, "Three Hundred Years of Gravitation,", Cambridge University Press, 1989.
24. J. Lee and R. M. Wald, "Local symmetries and constraints," J. Math. Phys. **31** (1990), 725-743 doi:10.1063/1.528801
25. R. M. Wald, "Black hole entropy is the Noether charge," Phys. Rev. D **48** (1993) no.8, 3427-3431 doi:10.1103/PhysRevD.48.R3427 [arXiv:gr-qc/9307038 [gr-qc]].
26. R. M. Wald and A. Zoupas, "A General definition of 'conserved quantities' in general relativity and other theories of gravity," Phys. Rev. D **61** (2000), 084027 doi:10.1103/PhysRevD.61.084027 [arXiv:gr-qc/9911095 [gr-qc]].

27. A. Ashtekar and M. Streubel, "Symplectic Geometry of Radiative Modes and Conserved Quantities at Null Infinity," Proc. Roy. Soc. Lond. A **A376** (1981), 585-607 doi:10.1098/rspa.1981.0109
28. J. D. Brown and J. W. York, Jr., "Quasilocal energy and conserved charges derived from the gravitational action," Phys. Rev. D **47** (1993), 1407-1419 doi:10.1103/PhysRevD.47.1407 [arXiv:gr-qc/9209012 [gr-qc]].
29. A. Komar, "Positive-Definite Energy Density and Global Consequences for General Relativity," Phys. Rev. **129** (1963) no.4, 1873 doi:10.1103/PhysRev.129.1873
30. S. W. Hawking, M. J. Perry and A. Strominger, "Soft Hair on Black Holes," Phys. Rev. Lett. **116** (2016) no.23, 231301 doi:10.1103/PhysRevLett.116.231301 [arXiv:1601.00921 [hep-th]].
31. J. D. Brown and M. Henneaux, "Central Charges in the Canonical Realization of Asymptotic Symmetries: An Example from Three-Dimensional Gravity," Commun. Math. Phys. **104** (1986), 207-226 doi:10.1007/BF01211590
32. A. Castro, A. Maloney and A. Strominger, "Hidden Conformal Symmetry of the Kerr Black Hole," Phys. Rev. D **82** (2010), 024008 doi:10.1103/PhysRevD.82.024008
33. S. Haco, S. W. Hawking, M. J. Perry and A. Strominger, "Black Hole Entropy and Soft Hair," JHEP **12** (2018), 098 doi:10.1007/JHEP12(2018)098 [arXiv:1810.01847 [hep-th]].
34. J. L. Cardy, "Operator Content of Two-Dimensional Conformally Invariant Theories," Nucl. Phys. B **270** (1986), 186-204 doi:10.1016/0550-3213(86)90552-3
35. S. Haco, M. J. Perry and A. Strominger, "Kerr-Newman Black Hole Entropy and Soft Hair," [arXiv:1902.02247 [hep-th]].
36. M. Perry and M. J. Rodriguez, "Central Charges for AdS Black Holes," [arXiv:2007.03709 [hep-th]].

# No smooth beginning for spacetime

Jean-Luc Lehners

*Max-Planck-Institute for Gravitational Physics (Albert-Einstein-Institute)*
*Am Mühlenberg 1, Potsdam, Germany*
*E-mail: jlehners@aei.mpg.de*

In these proceedings, I will review an obstruction for theories of the beginning of the universe which can be formulated as semiclassical path integrals. Hartle and Hawking's no boundary proposal and Vilenkin's tunneling proposal are examples of such theories. Each may be formulated as the quantum amplitude for obtaining a final 3-geometry by integrating over 4-geometries. The result is obtained using a new mathematical tool – Picard-Lefschetz theory – for defining the semiclassical path integral for gravity. The Lorentzian path integral for quantum cosmology with a positive cosmological constant is mathematically meaningful in this approach, but the Euclidean version is not. Framed in this way, the resulting framework and predictions are unique. Unfortunately, the outcome is that primordial gravitational wave fluctuations are unsuppressed.

*Keywords*: Quantum cosmology; Big bang; MG15 Proceedings.

## 1. Introduction

There is an old and attractive idea that the universe could have been created out of "nothing" via a quantum creation event. More specifically, since in a closed universe all total charges are zero, there is the intuition that no conservation law can prevent such an event from happening – and since everything that is not forbidden will eventually occur, so goes the logic, such a process must be possible. One could object to this line of reasoning by saying that in the absence of time, it is not at all clear that such a process must "eventually" occur. A more pertinent line of enquiry however is to first see if such an idea can be made concrete. This has indeed been attempted, in particular with the "no-boundary" proposal of Hartle and Hawking[1–3] and the tunneling wavefunction of Vilenkin[4–6]. These proposals have been heuristically formulated within the path integral approach to quantisation, but now with gravity included. The idea then is that the big bang is replaced by a closed and regular geometry corresponding to the Euclidean version of de Sitter space (for an illustration see Fig. 7). Such a geometry can arise as a saddle point of the path integral for gravity with a cosmological constant $\Lambda$, as we shall see in more detail below. What was left unspecified in these proposals is how to properly and precisely define the original path integral itself. Here Hartle and Hawking implicitly had in mind a Euclidean version of the path integral, since they hoped that this would lead to better convergence properties. Meanwhile Vilenkin stated that his proposal was to be viewed in the framework of a Lorentzian path integral. However, neither approach was made very precise, and thus the true consequences of these proposals remained somewhat heuristic.

The object of our work was to make both proposals mathematically precise. The outcome of this clarification was however rather negative: we discovered that both

proposals lead to unstable fluctuations, and that consequently the replacement of the big bang by a smooth geometry remains an untenable proposition at present.

The present review is based on Refs.[7–11]

## 2. Path Integral

In this overview we consider a universe with a positive cosmological constant $\Lambda$, described by the action

$$S = \frac{1}{2}\int_{\mathcal{M}} \mathrm{d}^4x\sqrt{-g}\,(R - 2\Lambda) + \int_{\partial\mathcal{M}} \mathrm{d}^3y\sqrt{g^{(3)}}K\,, \tag{1}$$

where we have set $8\pi G = 1$. The second term, involving the 3-metric $g^{(3)}_{ij}$ and the trace of the second fundamental form $K$ of the boundary $\partial\mathcal{M}$, renders the variational principle well posed. We start by truncating the theory to the simplest cosmologies, with line element

$$\mathrm{d}s^2 = -N(t)^2\mathrm{d}t^2 + a(t)^2\mathrm{d}\Omega_3^2\,, \tag{2}$$

with $\mathrm{d}\Omega_3^2$ the metric of a homogeneous, isotropic 3-sphere. Later on, we will add gravitational wave perturbations.

The Feynman path integral for the reduced theory is

$$G[a_1;a_0] = \int \mathcal{D}N\mathcal{D}\pi\mathcal{D}a\mathcal{D}p\mathcal{D}C\mathcal{D}\bar{P}\, e^{\frac{i}{\hbar}\int_0^1\left[\dot{N}\pi+\dot{a}p+\dot{C}\bar{P}-NH\right]\mathrm{d}t}\,, \tag{3}$$

where, in addition to $a$, $N$ and the fermionic ghost $C$, we have introduced the conjugate momenta $p, \pi$ and $\bar{P}$, and the corresponding Liouville measure. Without loss of generality, we can choose the range of the time coordinate to be $0 \leq t \leq 1$. The Hamiltonian constraint $H[a,p;N,\pi;C,\bar{P}] = H_{EH}[a,p] + H_g[N,\pi;C,\bar{P}]$ consists of the Einstein-Hilbert Hamiltonian $H_{EH}$, in our case a minisuperspace Hamiltonian, and a Batalin, Fradkin and Vilkovisky (BFV) ghost Hamiltonian $H_g$. The ghost term breaks time reparametrization symmetry and fixes the proper-time gauge $\dot{N} = 0$.[12] For minisuperspace models, most of the path integrals can be performed analytically, yielding

$$G[a_1;a_0] = \int_{0^+}^{\infty}\mathrm{d}N\int_{a=a_0}^{a=a_1}\mathcal{D}a\, e^{iS(N,a)/\hbar}\,, \tag{4}$$

which has a very simple interpretation. The path integral $\int \mathcal{D}a e^{iS(N,a)/\hbar}$ represents the quantum mechanical amplitude for the universe to evolve from $a_0$ to $a_1$ in a proper time $N$. The integral over the lapse function indicates that we should consider paths of every proper duration $0 < N < \infty$. Teitelboim[13] showed that this choice of integration domain leads to the causal ordering of the $a_0$ and $a_1$, *i.e.* $a_0$ precedes $a_1$. This allows us to describe both an expanding $a_1 > a_0$ and a contracting $a_1 < a_0$ universe, since the direction of the arrow of time is determined by the Feynman propagator and not by the choice of boundary conditions.

The action in (4) reduces to

$$S = 2\pi^2 \int_0^1 \mathrm{d}t N \left( -3a\frac{\dot{a}^2}{N^2} + 3a - a^3\Lambda \right) . \tag{5}$$

We are faced with a functional integral over $a(t)$, and an ordinary integral over the proper time $N$. In fact, we can simplify the calculation by noticing that redefining the lapse function $N(t) \to N(t)/a(t)$ renders the action (5) quadratic in $q(t) \equiv a(t)^2$, allowing the path integral over $q(t)$ to be performed exactly[14]. In these new variables, the action (5) becomes

$$S = 2\pi^2 \int_0^1 \mathrm{d}t \left( -\frac{3}{4N}\dot{q}^2 + N(3 - \Lambda q) \right) . \tag{6}$$

The equation of motion and the constraint following from this action are

$$\ddot{q} = \frac{2\Lambda}{3}N^2; \quad \frac{3}{4N^2}\dot{q}^2 + 3 = \Lambda q \, . \tag{7}$$

With boundary conditions $q(0) = q_0$ and $q(1) = q_1$, the general solution to the first equation (before imposing the constraint) is

$$\bar{q} = \frac{\Lambda}{3}N^2 t^2 + \left( -\frac{\Lambda}{3}N^2 + q_1 - q_0 \right) t + q_0 \, . \tag{8}$$

Writing the full solution as

$$q(t) = \bar{q}(t) + Q(t) \, , \tag{9}$$

the path integral becomes

$$G[q_1; q_0] = \int_0^\infty \mathrm{d}N e^{2\pi^2 i S_0/\hbar} \int_{Q[0]=0}^{Q[1]=0} \mathcal{D}Q e^{2\pi^2 i S_2/\hbar} \, , \tag{10}$$

with

$$S_0 = \int_0^1 \mathrm{d}t \left( -\frac{3}{4N}\dot{\bar{q}}^2 + 3N - N\Lambda\bar{q} \right) , \quad S_2 = -\frac{3}{4N}\int_0^1 \mathrm{d}t \, \dot{Q}^2 \, . \tag{11}$$

The path integral over $Q$ is Gaussian and can be evaluated exactly:

$$\int_{Q[0]=0}^{Q[1]=0} \mathcal{D}Q e^{2\pi^2 i S_2/\hbar} = \sqrt{\frac{3\pi i}{2N\hbar}} \, . \tag{12}$$

The propagator thus reduces to an ordinary integral

$$G[q_1; q_0] = \sqrt{\frac{3\pi i}{2\hbar}} \int_0^\infty \frac{\mathrm{d}N}{N^{1/2}} e^{2\pi^2 i S_0/\hbar} . \tag{13}$$

with

$$S_0 = N^3 \frac{\Lambda^2}{36} + N\left( -\frac{\Lambda}{2}(q_0 + q_1) + 3 \right) + \frac{1}{N}\left( -\frac{3}{4}(q_1 - q_0)^2 \right) . \tag{14}$$

Equation (13) is an oscillatory integral, implying that it is only conditionally convergent. Below, we will use Picard-Lefschetz theory to show that it is indeed convergent, and to evaluate it. However, it is instructive to see that an elementary

proof of convergence may be provided, without the use of analytic continuation. We will sketch it here.

For many oscillatory integrals, convergence can be demonstrated with the Leibniz convergence test for alternating series. A real alternating series is defined as

$$a = \pm \sum_{i=0}^{\infty} (-1)^i a_i \,, \tag{15}$$

with $a_i$ positive real numbers. The Leibniz convergence test states that the series is convergent when the arguments decrease monotonically, i.e. $a_{i+1} \leq a_i$ for sufficiently large $i$, and the argument goes to zero in the limit of large $i$, i.e. $\lim_{i\to\infty} a_i = 0$. To see the relation to oscillatory integrals, consider the integral

$$I = \int_0^\infty \mathrm{d}x \; e^{if(x)} \,, \tag{16}$$

for a real valued polynomial $f$ (for the Fresnel integrals $f(x) = x^2$). The real and imaginary parts of $I$ are given by

$$\mathrm{Re}[I] = \int_0^\infty \mathrm{d}x \;\; \cos(f(x)) \,, \quad \mathrm{Im}[I] = \int_0^\infty \mathrm{d}x \;\; \sin(f(x)) \,. \tag{17}$$

For simplicity we concentrate on the real part. Let us assume that the leading term of $f$ goes like $x^n$ in the limit $x \to \infty$ for $n \in \mathbb{N}$. A change of coordinates $u = x^n$ gives the integral

$$\mathrm{Re}[I] = \int_0^\infty \frac{\mathrm{d}u}{n u^{1-1/n}} \;\; \cos(f(\sqrt[n]{u})) \,, \tag{18}$$

and ensures that $f(\sqrt[n]{u}) \sim u$ for large $u$. Now let the zero crossings of the argument be given by $z_i$ for $i \in \mathbb{N}$. The real part of the integral can be written as an alternating series

$$\mathrm{Re}[I] = \left[ \int_0^{z_0} + \sum_{i=0}^{\infty} \int_{z_i}^{z_{i+1}} \right] \frac{\mathrm{d}u}{n u^{1-1/n}} \;\; \cos(f(u^{1/n})) \,, \tag{19}$$

$$= c \pm \sum_{i=0}^{\infty} (-1)^i \left| \int_{z_i}^{z_{i+1}} \frac{\mathrm{d}u}{n u^{1-1/n}} \;\; \cos(f(u^{1/n})) \right| \,, \tag{20}$$

$$= c \pm \sum_{i=0}^{\infty} (-1)^i a_i \,, \tag{21}$$

without changing the order of summation, with $c$ the integral over the interval $(0, z_0)$, either the positive or the negative sign for $\pm$ depending on the details of $f$, and the positive real numbers

$$a_i = \left| \int_{z_i}^{z_{i+1}} \frac{\mathrm{d}u}{n u^{1-1/n}} \;\; \cos(f(u^{1/n})) \right| \,. \tag{22}$$

The argument of the alternating series can be dominated with a simple approximation

$$a_i = \left| \int_{z_i}^{z_{i+1}} \frac{\mathrm{d}u}{n u^{1-1/n}} \cos(f(u^{1/n})) \right| \tag{23}$$

$$< \int_{z_i}^{z_{i+1}} \frac{\mathrm{d}u}{n u^{1-1/n}} = \sqrt[n]{z_{i+1}} - \sqrt[n]{z_i} = b_i \,. \tag{24}$$

In the limit of large $u$ the function $f(\sqrt[n]{u})$ asymptotes to a function proportional to $u$. For this reason, in the limit of large $i$, the zero crossings $z_i$ in $u$ will asymptote to a regular spacing, leading to the conclusion that for $n > 1$ and for sufficiently large $i$, the coefficients $b_i$ satisfy the conditions of the Leibniz convergence test. Since $a_i < b_i$ for all $i$ we conclude that $\mathrm{Re}[I]$ converges when $n > 1$. A similar argument can be given for the imaginary part of $I$, making $I$ conditionally convergent. This discussion applies to a more general class of integrals. When the integral function $f(x)$ diverges as $x^{-n}$ in the limit $x \to 0$ with $n \in \mathbb{N}$, the change of coordinates $u = x^{-n}$ leads to convergence for $n > 1$. More generally, when $f$ is not a polynomial but dominates some polynomial $x^n$ with $n > 1$, the oscillatory integral can be shown to converge due to cancellations from oscillations at large $x$. Our integral above, when expressed in terms of the canonically normalised integration variable $M \equiv N^{1/2}$, involves the asymptotic behaviour $\sim \int dM e^{i(\Lambda^2 M^6 - q_1^2/M^2)/\hbar}$, and hence is manifestly convergent. In order to evaluate its value, however, Picard-Lefschetz theory is a preferable tool to which we turn our attention now.

### 2.1. *Picard-Lefschetz theory*

Picard-Lefschetz theory deals with oscillatory integrals like

$$I = \int_D \mathrm{d}x \, e^{iS[x]/\hbar}, \tag{25}$$

where $\hbar$ is a real parameter, the action $S[x]$ is a real-valued function and the integral is taken over a real domain $D$, usually defined by the singularities of the integrand or, in higher dimensional or path integral cases, its partial integrals (see e.g.[15] for a more complete overview). One is typically interested in the behavior of the integral for small values of the parameter $\hbar$: in quantum mechanical applications, taking $\hbar$ to zero is a nice way to study the classical limit.

When faced with an integral in the form of (25), the idea of Picard-Lefshetz theory is to interpret $S[x]$ as a holomorphic function of $x \in \mathbb{C}$, the complex plane. Cauchy's theorem allows us to deform the integration contour from the real domain $D$ on the real $x$-axis into a contour we now call $\mathcal{C}$ in the complex $x$-plane, while keeping its endpoints fixed. In particular, we seek to deform $\mathcal{C}$ into a "steepest descent" contour passing through one or more critical points of $S[x]$, *i.e.* points where $\partial_x S = 0$. By the Cauchy-Riemann equations, the real part of the exponent, $\mathrm{Re}[iS[x]]$, which controls the magnitude of the integrand, has a saddle point in the real two-dimensional (Re[$x$],Im[$x$])-plane there. The steepest descent contour

through the saddle point is defined as the path along which $\mathrm{Re}[iS[x]]$ decreases as rapidly as possible.

A simple example is provided by $S[x] = x^2$, with a critical point at $x = 0$. Writing $x = \mathrm{Re}[x] + i\,\mathrm{Im}[x]$, we have $\mathrm{Re}[iS[x]] = -2\,\mathrm{Re}[x]\mathrm{Im}[x]$. The magnitude of the integrand decreases most rapidly along the contour $\mathrm{Im}[x]=+\mathrm{Re}[x]$ which is the steepest descent contour. Conversely, it increases most rapidly along the contour $\mathrm{Im}[x] = -\mathrm{Re}[x]$, which is the steepest ascent contour. As we shall discuss, steepest descent contours generically lead to convergent integrals, and in this case they are known as Lefschetz thimbles $\mathcal{J}_\sigma$.

In more detail, we write the exponent $\mathcal{I} = iS/\hbar$ and its argument $x$ in terms of their real and imaginary parts, $\mathcal{I} = h + iH$ and $x = u^1 + iu^2$. Downward flow is then defined by

$$\frac{\mathrm{d}u^i}{\mathrm{d}\lambda} = -g^{ij}\frac{\partial h}{\partial u^j}\,, \tag{26}$$

with $\lambda$ a parameter along the flow and $g_{ij}$ a Riemannian metric introduced on the complex plane. The real part of the exponent $h$ (known as the Morse function) decreases on such a flow away from its critical points, because $\frac{\mathrm{d}h}{\mathrm{d}\lambda} = \sum_i \frac{\partial h}{\partial u^i}\frac{\mathrm{d}u^i}{\mathrm{d}\lambda} = -\sum_i \left(\frac{\partial h}{\partial u^i}\right)^2 < 0$, with the fastest rate of decrease occuring in the direction of "steepest descent", which maximises the magnitude of the gradient. Defining the latter requires that we introduce a metric.

For the simple examples we discuss here, the obvious metric $\mathrm{d}s^2 = |\mathrm{d}x|^2$ is sufficient. Defining complex coordinates, $(u, \bar{u}) = \big((\mathrm{Re}([x] + i\mathrm{Im}[x]), (\mathrm{Re}[x] - i\mathrm{Im}[x])\big)$, the metric is $g_{uu} = g_{\bar{u}\bar{u}} = 0$, $g_{u\bar{u}} = g_{\bar{u}u} = 1/2$. Then $h = (\mathcal{I} + \bar{\mathcal{I}})/2$ and (26) becomes

$$\frac{\mathrm{d}u}{\mathrm{d}\lambda} = -\frac{\partial\bar{\mathcal{I}}}{\partial\bar{u}}, \quad \frac{\mathrm{d}\bar{u}}{\mathrm{d}\lambda} = -\frac{\partial\mathcal{I}}{\partial u}\,. \tag{27}$$

The imaginary part of the exponent $H = \mathrm{Im}[iS/\hbar]$ is conserved along these flows, since

$$\frac{\mathrm{d}H}{\mathrm{d}\lambda} = \frac{1}{2i}\frac{\mathrm{d}(\mathcal{I} - \bar{\mathcal{I}})}{\mathrm{d}\lambda} = \frac{1}{2i}\left(\frac{\partial\mathcal{I}}{\partial u}\frac{\mathrm{d}u}{\mathrm{d}\lambda} - \frac{\partial\bar{\mathcal{I}}}{\partial\bar{u}}\frac{\mathrm{d}\bar{u}}{\mathrm{d}\lambda}\right) = 0\,. \tag{28}$$

Thus the integrand $e^{iS[x]/\hbar}$ – which was a purely oscillatory factor in the original integral – does not oscillate at all when evaluated along a downward flow (see Fig. 1). Instead, it decreases monotonically so that the integral converges absolutely and "as rapidly as possible." For a downward flow originating at a saddle, $\lambda$ runs from $-\infty$ at the saddle point to positive values as $h$ decreases. The Lefschetz thimble associated with a given saddle is defined as the set of downward flows leaving the saddle in this way.

Analogously, upward flows are defined via

$$\frac{\mathrm{d}u^i}{\mathrm{d}\lambda} = +g^{ij}\frac{\partial h}{\partial u^j}\,, \tag{29}$$

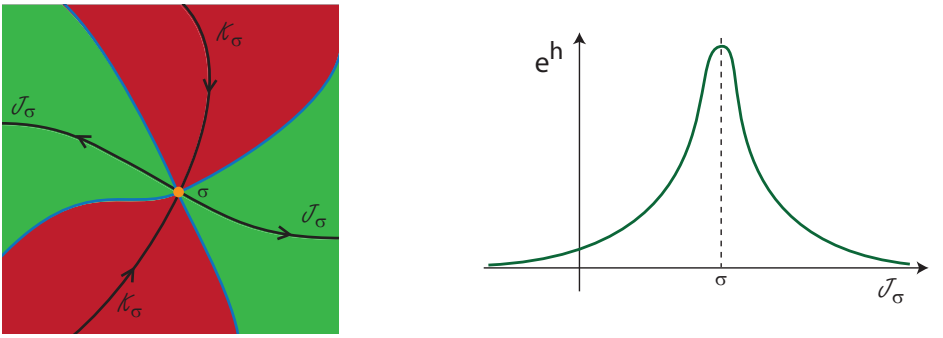



Fig. 1. *Left panel:* From a saddle point $\sigma$ emanate upward ($\mathcal{K}_\sigma$) and downward ($\mathcal{J}_\sigma$) flows, which are located in the wedges $J_\sigma$ (in green) and $K_\sigma$ (in red) respectively, defined as the regions where the Morse function $h$ is lower (higher) than its value at the saddle, respectively. The arrows along the flows indicate the direction of descent, and the downward flow $\mathcal{J}_\sigma$ is known as a Lefschetz thimble. The wedges are separated by blue lines along which $h$ is constant and equal to the value at the saddle point $h(p_\sigma.)$ *Right panel:* Along a Lefschetz thimble the real part $h$ of the exponent decreases as fast as possible, ensuring an absolutely convergent integral.

with $H$ likewise being conserved along these flows. Every critical point has an upward flow which, in analogy to the downward flow, is labelled $\mathcal{K}_\sigma$.

Lefschetz thimbles and upward flows only intersect at a single critical point (in case of a degeneracy, a small perturbation may be added), the one where both are defined. With a suitable choice of orientation, we can write for the intersection number

$$\mathrm{Int}(\mathcal{J}_\sigma, \mathcal{K}_{\sigma'}) = \delta_{\sigma\sigma'}. \tag{30}$$

Our objective is to deform the original integral (25) into one evaluated over a sum of Lefschetz thimbles. That is, we would like to write

$$\mathcal{C} = \sum_\sigma n_\sigma \mathcal{J}_\sigma, \tag{31}$$

in a homological sense, for some integers $n_\sigma$ which may take the values 0 or $\pm 1$ when accounting for the orientation of the contour over each thimble. It follows from these equations that $n_\sigma = \mathrm{Int}(\mathcal{C}, \mathcal{K}_\sigma) = \mathrm{Int}(D, \mathcal{K}_\sigma)$, since the intersection number is topological and will not change if we deform the contour $\mathcal{C}$ back to the original, real domain $D$. Thus a necessary and sufficient condition for a given thimble $\mathcal{J}_\sigma$ to be relevant is that a steepest *ascent* contour from the critical point $p_\sigma$ intersects the original, real integration domain $D$. In this circumstance, intuitively, there is no obstacle to smoothly "sliding" the intersection point from the real axis along $\mathcal{K}_\sigma$ down to $p_\sigma$, and in the process deforming the original integration contour onto the the thimble $\mathcal{J}_\sigma$. This is an argument that shows that if one starts from a real Lorentzian theory, then as long as the action is analytic one *never* obtains semiclassical enhancement factors such as are found in the Euclidean approach.

Once we have deformed the contour from the real axis to run through a set of thimbles associated with the contributing critical points, we have:

$$I = \int_D \mathrm{d}x\, e^{iS[x]/\hbar} = \int_{\mathcal{C}} \mathrm{d}x\, e^{iS[x]/\hbar} = \sum_\sigma n_\sigma \int_{\mathcal{J}_\sigma} \mathrm{d}x\, e^{iS[x]/\hbar}. \tag{32}$$

As (32) indicates, typically more than one Lefschetz thimble contributes to the Lorentzian path integral, with given boundary conditions, even in mini-superspace quantum cosmology.

The integral taken over a thimble is absolutely convergent if

$$\left| \int_{\mathcal{J}_\sigma} \mathrm{d}x e^{iS[x]/\hbar} \right| \leq \int_{\mathcal{J}_\sigma} |\mathrm{d}x| \left| e^{iS[x]/\hbar} \right| = \int_{\mathcal{J}_\sigma} |dx| e^{h(x)} < \infty\,. \tag{33}$$

Defining the length along the curve as $l = \int |dx|$, the integral will converge if $h(x(l)) < -\ln(l) + A$, for some constant $A$, as $l \to \infty$, which is a rather weak requirement.

We have then expressed the original integral as a sum of absolutely convergent steepest descent integrals. In an expansion in $\hbar$, we have

$$I = \int_D \mathrm{d}x\, e^{iS[x]/\hbar} = \sum_\sigma n_\sigma\, e^{i\,H(p_\sigma)} \int_{\mathcal{J}_\sigma} e^h \mathrm{d}x \approx \sum_\sigma n_\sigma\, e^{iS(p_\sigma)/\hbar} \left[A_\sigma + \mathcal{O}(\hbar)\right], \tag{34}$$

where $A_\sigma$ represents the result of the leading-order Gaussian integral about the critical point $p_\sigma$. Sub-leading terms may be evaluated perturbatively in $\hbar$.

### 2.2. *Relevant and irrelevant saddle points*

The action $S_0$ has four saddle points in the complex plane, which are solutions of

$$\partial S_0/\partial N = \Lambda^2 N_s^4 + (-6\Lambda(q_0+q_1)+36)\, N_s^2 + 9(q_1-q_0)^2 = 0\,, \tag{35}$$

given by

$$N_s = c_1 \frac{3}{\Lambda} \left[ \left( \frac{\Lambda}{3} q_0 - 1 \right)^{1/2} + c_2 \left( \frac{\Lambda}{3} q_1 - 1 \right)^{1/2} \right], \tag{36}$$

with $c_1, c_2 \in \{-1, 1\}$. The action evaluated at these saddle points is given by

$$S_0^{saddle} = -c_1 \frac{6}{\Lambda} \left[ \left( \frac{\Lambda}{3} q_0 - 1 \right)^{3/2} + c_2 \left( \frac{\Lambda}{3} q_1 - 1 \right)^{3/2} \right]. \tag{37}$$

Each of these four saddle points corresponds to a Lefschetz thimble $\{\mathcal{J}_\sigma\}$, and a steepest ascent contour $\{\mathcal{K}_\sigma\}$. Each is also associated with wedges $J_\sigma, K_\sigma$ in which the real part of the exponent $iS/\hbar$ is respectively lower and higher than the saddle point value. Writing the original integration contour in terms of the Lefschetz thimbles

$$(0^+, \infty) = \sum_\sigma n_\sigma \mathcal{J}_\sigma\,, \tag{38}$$

we approximate the propagator using the saddle point approximation in the limit $\hbar \to 0$,

$$\begin{aligned} G[q_1;q_0] &= \sum_\sigma n_\sigma \sqrt{\frac{3\pi i}{2\hbar}} \int_{\mathcal{J}_\sigma} \frac{\mathrm{d}N}{N^{1/2}} e^{2\pi^2 i S_0/\hbar} \\ &\approx \sum_\sigma n_\sigma \sqrt{\frac{3\pi i}{2\hbar}} \frac{e^{2\pi^2 i S_0^{saddle}/\hbar}}{N_s^{1/2}} \int_{\mathcal{J}_\sigma} \mathrm{d}N e^{\frac{i\pi^2}{\hbar} S_{0,NN}(N-N_s)^2} \left[1 + \mathcal{O}\left(\hbar^{1/2}\right)\right] \\ &\approx \sum_\sigma n_\sigma \sqrt{\frac{3\pi i}{2\hbar}} \frac{e^{2\pi^2 i S_0^{saddle}/\hbar}}{N_s^{1/2}} e^{i\theta_\sigma} \int_{\mathcal{J}_\sigma} \mathrm{d}n e^{-\frac{\pi^2}{\hbar}|S_{0,NN}|n^2} \left[1 + \mathcal{O}\left(\hbar^{1/2}\right)\right] \\ &\approx \sum_\sigma n_\sigma \sqrt{\frac{3i}{2N_s|S_{0,NN}|}} e^{i\theta_\sigma} e^{2\pi^2 i S_0^{saddle}/\hbar} \left[1 + \mathcal{O}\left(\hbar^{1/2}\right)\right] , \end{aligned} \tag{39}$$

where we defined $N - N_s \equiv n e^{i\theta}$ with $n$ real and $\theta$ being the angle of the Lefschetz thimble with respect to the positive real $N$ axis.

Earlier approaches amount to choosing a particular contour in the complex $N$ plane "by hand," on the basis of some preconceived notions. However, the virtue of the Lorentzian path integral combined with Picard-Lefschetz theory is that the proper combination of saddle points and relative phases between them is completely fixed.

The "no-boundary" conditions were proposed by Hartle and Hawking as a theory of initial conditions for the universe[1–3]. The idea is that in the path integral one should sum only metrics whose only boundary is provided by the final spatial hypersurface (corresponding to the current state of the universe). To implement "no-boundary" conditions, we must take $q_0 = 0$ and find a 4-metric which is regular there. This is possible for positive spatial curvature, as assumed here. The "no-boundary" condition is supplemented with the constraint equation (7) evaluated at $q = 0$,

$$\dot{q}^2 = -4N^2 \qquad (q = 0) \ . \tag{40}$$

We will take the final boundary to correspond to a late time configuration, where the universe has become large, $q_1 > \frac{3}{\Lambda}$. The saddle points of the action are given by

$$N_{s,nb1} = +\frac{3}{\Lambda}\left[i \pm \left(\frac{\Lambda}{3}q_1 - 1\right)^{1/2}\right] , \quad N_{s,nb2} = -\frac{3}{\Lambda}\left[i \pm \left(\frac{\Lambda}{3}q_1 - 1\right)^{1/2}\right] , \tag{41}$$

with corresponding actions

$$S_{0,nb1} = -\frac{6}{\Lambda}\left[-i \pm \left(\frac{\Lambda}{3}q_1 - 1\right)^{3/2}\right] , \quad S_{0,nb2} = +\frac{6}{\Lambda}\left[-i \pm \left(\frac{\Lambda}{3}q_1 - 1\right)^{3/2}\right] . \tag{42}$$

Note that saddle points in the upper half plane lead to a $e^{i2\pi^2 S_0} \sim e^{-12\pi^2/(\hbar\Lambda)}$, while those in the lower half plane lead to $e^{i2\pi^2 S_0} \sim e^{+12\pi^2/(\hbar\Lambda)}$.

Given the saddle points, we can determine the wedges and the curves of steepest descent and ascent emanating from them. We use the fact that curves with $\mathrm{Re}(iS_0)$

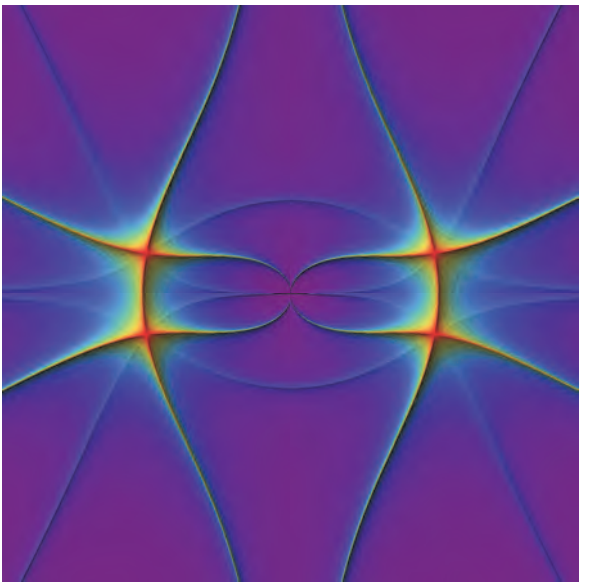


Fig. 2. For this numerical example we have chosen $\Lambda = 3, q_0 = 0, q_1 = 10$. The saddle points then lie at $\pm 3 \pm i$. Shown in the present figure are both the boundaries of wedges (lines of constant real part of the integrand/imaginary part of the action – light blue lines) and the flow lines (lines of constant real part of the action – red/green lines). More specifically, the plot shows both $Abs[\mathrm{Im}(S(N) - S(N_s))]$ and $Abs[\mathrm{Re}(S(N) - S(N_s))]$, where lighter colours correspond to smaller values. The four saddle points are located at the intersections of the flow lines. More details are provided in Fig. 3.

specify the boundaries of the wedges, and that $\mathrm{Im}(iS_0)$ is constant along the flow lines to determine them numerically – see also[16]. For the case of interest to us, the wedge boundaries and flow lines are shown in Fig. 2, while the directions of the flows are sketched in Fig. 3.

Note that the downward flow lines (Lefschetz thimbles) of the upper saddle points can indeed be deformed to the real $N$ line, while the downward flow lines of the lower saddle points cannot. Moreover, only saddle point 1 can be linked to the original integration contour (the positive real half line) via an upward flow, and hence the appropriate integration contour, along which the integral will be manifestly convergent, is given by the Lefschetz thimble $\mathcal{J}_1$ also indicated by the dashed orange line in Fig. 3. Saddle point 1 lies at

$$N^{+}_{s,nb1} = +\frac{3}{\Lambda}\left[i + (\frac{\Lambda}{3}q_1 - 1)^{1/2}\right], \tag{43}$$

and the action evaluated on the saddle point is

$$S^{+}_{0,nb1} = -\frac{6}{\Lambda}\left[-i + (\frac{\Lambda}{3}q_1 - 1)^{3/2}\right]. \tag{44}$$

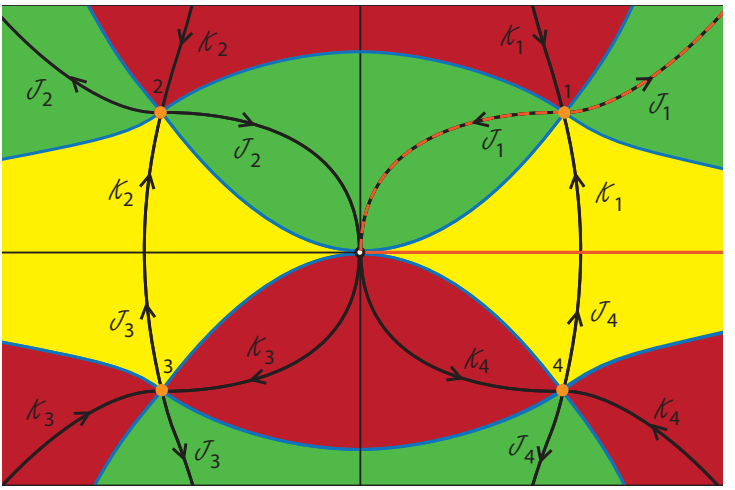



Fig. 3. A sketch of the wedges and flow lines emanating from the saddle points in the complex $N$ plane, for "no-boundary" conditions $q_0 = 0, q_1 > \frac{3}{\Lambda}$. The loci of the steepest ascent/descent flows (in black) and of the boundaries between wedges (in blue) were determined numerically in Fig. 2. Here the arrows indicate the direction of steepest descent. We have coloured the wedges such that regions $J_\sigma$ with a lower value of the magnitude of the integrand than the corresponding saddle point are green, and regions $K_\sigma$ with a higher value are red, with the exception of the yellow regions which have a value intermediate between the two saddle point values. Comparing with the adjacent colours then avoids any ambiguity. The original integration contour along the positive real axis is shown in orange, and the deformed contour which Picard-Lefschetz theory picks out as the preferred integration cycle is marked in dashed orange. Again neither the flow lines, nor the original or final integration contours, include the point at $N = 0$. Only saddle point 1 in the upper right quadrant can be linked to the original integration contour via an upward flow, and this implies that the (orange-dashed) downward flow from this saddle point is the correct Lefschetz thimble along which the path integral should be performed.

For saddle points of the form (36), we have

$$S_{0,NN} = \frac{2c_2}{N_s} \left(\Lambda q_0 - 3\right)^{1/2} \left(\Lambda q_1 - 3\right)^{1/2} , \tag{45}$$

implying that $Arg(N_s) = -\alpha + Arg\left[\left(\Lambda q_0 - 3\right)^{1/2} \left(\Lambda q_1 - 3\right)^{1/2}\right]$. For the "no-boundary" conditions we thus find $Arg(N_s) + \alpha = \frac{\pi}{2}$, and this implies $\theta - \frac{1}{2} Arg(N_s) = 0$. In the saddle point approximation, we thus obtain the wavefunction

$$G_{nb}[q_1; 0] \approx e^{i\frac{\pi}{4}} \frac{3^{1/4}}{2(\Lambda q_1 - 3)^{1/4}} e^{-12\pi^2/(\hbar\Lambda) - i4\pi^2\sqrt{\frac{\Lambda}{3}}(q_1 - \frac{3}{\Lambda})^{3/2}/\hbar} . \tag{46}$$

Note that the real part of the classical action for the dominant saddle point is negative, as expected from the general arguments presented earlier. This concludes the explicit derivation of our result that the relevant saddle point contributes a weighting $e^{-12\pi^2/(\hbar\Lambda)}$, the inverse of the Hartle-Hawking result.

## 3. Unstable Perturbations

We now turn our attention to the perturbations, treated in general relativistic linear perturbation theory. The second order action for a linearized (tensor) perturbation $\phi_l$ with principal quantum number $l$ is given in terms of the background squared scale factor $q(t)$ as

$$S^{(2)}[q,\phi;N] = \frac{1}{2}\int_0^1 \left[q^2\frac{\dot{\phi}_l^2}{N} - Nl(l+2)\phi_l^2\right]\mathrm{d}t$$
$$= \frac{1}{2}\int_0^1 \left[\frac{\dot{\chi}_l^2}{N} + N\left(\frac{\ddot{q}}{q} - \frac{l(l+2)}{q^2}\right)\chi_l^2\right]\mathrm{d}t - \frac{1}{2N}\left[\frac{\dot{q}}{q}\chi_l^2\right]_0^1, \qquad (47)$$

where we have re-expressed the dimensionless tensor metric perturbation $\phi_l$ in terms of the canonically normalized field $\chi_l = q\,\phi_l$. Note that we have orthonormalized the modes on the unit sphere (thus no prefactor of $2\pi^2$ appears in the action). Also, to avoid needless complexity in the equations we only consider a single mode. It is straightforward to amend all the formulae we derive by replacing $l(l+1)(l+2)\phi_l^2$ with $\sum_{lmn} l(l+1)(l+2)\phi_{lmn}^2$ where the $\phi_{1,lmn}$ are the expansion coefficients expressing in the final tensor perturbation in terms of orthonormal tensor spherical harmonics on the three sphere. For ease of notation, where there is no danger of confusion, we will also usually drop the subscript $l$. Note that the perturbation of the lapse $N$ is non-dynamical in the absence of matter and may be set to zero.

If we neglect the backreaction of the linear perturbations on the background, such as is certainly reasonable for small final amplitude $\phi_1$, then we can evaluate the path integral first for $q$ and then for $\phi$, using the classical solution for the background $q$ in the action (47) for $\phi$. To integrate out the perturbations, we again just find the classical solution and use this to evaluate the classical action. The total semiclassical exponent is then given by $S^{(0)}[q_1;N] + S^{(2)}[q_1,\phi_1;N]$ and we then perform the final ordinary integral over $N$ using saddle point methods. We shall not calculate any functional determinants here, although this is perfectly possible. These should not alter any conclusions about the Picard-Lefschetz flow, nor the final semiclassical exponent, in any regime where the semiclassical expansion is valid.

The no boundary path integral on a contour $\mathcal{C}$ is then given, in this leading semiclassical approximation, by

$$G_{\mathcal{C}}[q_1,\phi_1;0] \propto \int_{\mathcal{C}} \frac{\mathrm{d}N}{\sqrt{D(N,q_1,\hbar)}} e^{i\bar{S}^{(0)}[q_1;N]/\hbar + i\bar{S}^{(2)}[q_1,\phi_1;N]/\hbar}\,, \qquad (48)$$

where $\bar{S}^{(2)}[q_1,\phi_1;N]$ is the classical action for the perturbation, in the background $\bar{q}$, satisfying $\phi(1) = \phi_1$ as well as a second condition we shall define shortly. The quantity $D(N,q_1,\hbar)$ is the functional determinant which is in principle calculable in terms of the classical modes and as a series expansion in $\hbar$.

In the leading semiclassical approximation, we can perform the path integral by the saddle point method, *i.e.*, by solving the equations of motion and computing the classical action. The boundary condition on the perturbations at $t = 0$ is delicate

because the background geometry is sufficiently singular for a range of real values of the lapse, that the perturbations obey a singular equation of motion. We shall find that, nevertheless, for generic complex $N$, the criterion of finite classical action selects a unique perturbation mode.

At fixed $N$, the classical equation for $\chi$ following from (47) is

$$\ddot{\chi} = \left( \frac{\ddot{\bar{q}}}{\bar{q}} - \frac{N^2 l(l+2)}{\bar{q}^2} \right) \chi \,. \tag{49}$$

Near $t = 0$, this becomes

$$\ddot{\chi} \approx -\frac{N^2 l(l+2)}{(q_1 - \Lambda N^2/3)^2} \frac{\chi}{t^2} \equiv \frac{\gamma^2 - 1}{4} \frac{\chi}{t^2} \,, \tag{50}$$

from which we see $\chi \sim t^{\frac{1}{2}(1 \pm \gamma)}$, as $t \to 0$. Notice that the equation of motion for $\chi$ is singular and this results in some unusual properties of the perturbations, as we explain below.

For small real $N$, we take $\gamma$ to be real and positive. Provided $N$ is real and smaller in magnitude than a particular value $N_-$, then both solutions for $\chi$ are monotonic in $t$ and both vanish at $t = 0$. However, only one of them has finite action so it is natural to select that one as the saddle point solution. For real $N$ larger in magnitude than $N_-$ but smaller than another, larger value, $N_+$, $\gamma$ is imaginary and the solutions oscillate an infinite number of times as they approach $t = 0$. In fact, both solutions have a finite regularized action, so the finite action criterion becomes ambiguous for $N$ in this range. Increasing the magnitude of $N$ beyond $N_+$, while keeping $N$ real, we see that $\gamma$ becomes real once again. However, as we explain shortly, in this latter regime, there are *no* finite action classical solutions.

The two critical values are given by

$$N_- = \frac{3}{\Lambda} \sqrt{2l(l+2) + q_1 \Lambda/3 - 2\sqrt{l(l+2)(l(l+2) + q_1 \Lambda/3)}} \,, \tag{51}$$

$$N_+ = \frac{3}{\Lambda} \sqrt{2l(l+2) + q_1 \Lambda/3 + 2\sqrt{l(l+2)(l(l+2) + q_1 \Lambda/3)}} \,, \tag{52}$$

with geometric mean $N_\star \equiv \sqrt{N_+ N_-} = \sqrt{3q_1/\Lambda}$. It follows that we can take

$$\gamma = \frac{\sqrt{(N_-^2 - N^2)(N_+^2 - N^2)}}{(N_\star^2 - N^2)} \,, \tag{53}$$

defined to be real and positive for small $N$ and defining $\gamma$ at complex $N$ by analytic continuation. The branch cuts needed to define the square roots are conveniently placed along the real intervals $-N_+ < N < -N_-$ and $N_- < N < N_+$. On the upper side of the cuts, $\gamma$ is negative imaginary and on the lower sides it is positive imaginary. Away from the cuts, as is evident from (47), the action integral converges at $t = 0$ only for the mode behaving as $t^{\frac{1}{2}(1+\gamma)}$ as $t \to 0$. The complete solution of (49) with this small $t$ behavior is

$$\chi(t) = \bar{q}(t)^{\frac{1}{2}} \left( \frac{t}{3q_1 + (t-1)N^2 \Lambda} \right)^{\frac{\gamma}{2}} \left( (3q_1 - \Lambda N^2)(1+\gamma) + 2\Lambda N^2 t \right) \,, \tag{54}$$

and the corresponding, correctly normalized classical solution is

$$\phi(t) = \phi_1 \frac{\chi(t)}{\bar{q}(t)} \frac{q_1}{\chi(1)}. \tag{55}$$

This solution allows us to calculate the classical action from (47). With an integration by parts and using the equations of motion, we find

$$S^{(2)}[q_1, \phi_1; N] = \left[\bar{q}^2 \frac{\bar{\phi}\dot{\phi}}{2N}\right]_0^1 = \frac{l(l+2)q_1\phi_1^2}{4N(3l(l+2)+q_1\Lambda)} \left(-3q_1 - N^2\Lambda + \gamma(N_\star^2 - N^2)\right), \tag{56}$$

which is real where $\gamma$ is real, but gains a negative or positive imaginary part (meaning that the semiclassical exponent $iS/\hbar$ gains a positive or negative real part) as $N$ approaches the real axis from above or below the branch cuts. This behavior is illustrated in Figure 4.

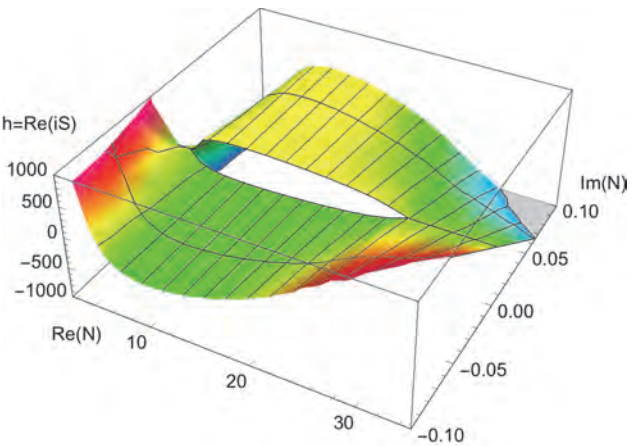



Fig. 4. The Morse function $h = Re(iS/\hbar)$ around a branch cut, in units where $\hbar = 1$ and for the parameters $\Lambda = 3, q_1 = 101, l = 10, \phi_1 = 1$. At the cut, the Morse function reaches its maximum at $N_\star = 10$ coming from the upper half plane, and its minimum also at $N_\star$, though approaching the cut from below.

There is one additional important consideration: for real $N$, the background solution for the scale factor is real and quadratic in $t$. For $N > N_\star$ (or $N < -N_\star$) the background solution starts at $q = 0$, then turns negative before crossing $q = 0$ a second time, at $t_s = 1 - \frac{3q_1}{\Lambda N^2}$, to eventually reach $q_1$ at $t = 1$. Thus there is a second singularity in these real but off-shell-in-$N$ background geometries. It is obvious from (54) that if $\chi$ behaves as $t^{\frac{1}{2}(1\pm\gamma)}$ near $t = 0$, then it behaves as $(t_s - t)^{\frac{1}{2}(1\mp\gamma)}$ near $t = t_s$. Thus, for real $\gamma$ and $N > N_\star$ then if the action integral converges at $t = 0$, it diverges at $t = t_s$, and vice versa. We conclude that for $N > N_+$ or $N < -N_+$ *no* solution of the perturbation equations of motion has finite action. Hence, in performing the integration over $N$ in the last step (48) of our calculation, however

we deform the contour $\mathcal{C}$, we cannot allow it to cross the real $N$-axis for real $N$ beyond the outer ends of the two branch cuts.

Finally, notice that at large $|N|$ in the complex $N$-plane, the background action $\bar{S}^{(0)} \sim N^3$ – dominates over the perturbation action $\bar{S}^{(2)} \sim N$. The same holds in the small $|N|$ limit, where both the background and the perturbation diverge like $1/N$ (the background and the perturbation action have opposite sign). As a consequence the asymptotic regions of convergence are preserved when we add linearized tensor perturbations.

The leading contribution of the perturbations, in the semi-classical approximation, comes from the relevant saddle point of the background. At this saddle points, the parameter $\gamma$ is precisely equal to $l+1$, meaning that the tensor modes $\phi_l$ behave as $t^{l/2}$ near the singularity, which means they are regular there. The value of the classical action at the relevant saddle point is then

$$\bar{S}^{(2)}(N_s^+) = -i\frac{\phi_1^2 q_1}{2}\frac{l(l+2)}{l+1+i\sqrt{q_1\Lambda/3-1}}. \tag{57}$$

There are two simplifying regimes. If the wavelength on the final three-geometry, $\sim \sqrt{q_1}l^{-1}$ is well within the Hubble radius $\sqrt{3/\Lambda}$, we obtain

$$i\frac{\bar{S}^{(2)}(N_s^+)}{\hbar} \approx +\frac{\phi_1^2 q_1}{2\hbar}l, \quad l \gg \sqrt{\frac{\Lambda q_1}{3}}, \tag{58}$$

a result which is independent of $\Lambda$. In the opposite limit, we obtain the result for the "frozen" modes in the expanding de Sitter spacetime, which have passed out of the de Sitter Hubble radius and ceased to evolve. In this case, we obtain

$$i\frac{\bar{S}^{(2)}(N_s^+)}{\hbar} \approx +\frac{3}{2\Lambda}l(l+1)(l+2)\phi_1^2 - i\sqrt{\frac{3q_1}{4\Lambda}}l(l+2)\phi_1^2, \quad \sqrt{\frac{\Lambda q_1}{3}} \gg l. \tag{59}$$

As we can see, the relevant saddle point leads to an inverse Gaussian distribution, meaning that the tensor modes are out of control.

## 4. Comments

In view of the unsettling nature of our result, one may ask whether our approximations have broken down, or whether we have somehow incorrectly implemented the idea of the no-boundary proposal. In this section we will address these issues and show that our result, to the best of our knowledge, really is robust.

### 4.1. *Backreaction*

First one may wonder whether the instability which we have demonstrated only occurs for very small perturbations, and whether it will be cured for larger perturbations. The $l = 2$ modes are particularly interesting in this regard, as a possible non-linear completion of the metric exists in the form of the Bianchi IX line element

$$ds_{IX}^2 = -N_p^2(t)dt_p^2 + \sum_m \left(\frac{l_m(t)}{2}\right)^2 \sigma_m^2\,, \tag{60}$$

where $N_p$ is the physical lapse function and $\sigma_1 = \sin\psi d\theta - \cos\psi\sin\theta d\varphi$, $\sigma_2 = \cos\psi d\theta + \sin\psi\sin\theta d\varphi$, and $\sigma_3 = -(d\psi + \cos\theta d\varphi)$ are differential forms on the three sphere such that $0 \leq \psi \leq 4\pi$, $0 \leq \theta \leq \pi$, and $0 \leq \phi \leq 2\pi$. For ease of notation we will denote a derivative w.r.t. physical time $t_p$ by an overdot in this section (and only in this section). Employing the original definition of Misner, we can re-write the three scale factors as

$$l_{1,2}(t_p) = a(t_p)\exp\left[\frac{1}{2}\left(\beta_+(t_p) + \pm\sqrt{3}\beta_-(t_p)\right)\right], \tag{61}$$

$$l_3(t_p) = a(t_p)\exp\left[-\beta_+(t_p)\right], \tag{62}$$

which makes it clear that $a$ is the average scale factor while the $\beta$s quantify anisotropic perturbations. In these coordinates the action becomes

$$S = 2\pi^2\int dt_p N_p a\left[\frac{1}{N_p^2}\left(-3\dot{a}^2 + \frac{3}{4}a^2\left(\dot{\beta}_+^2 + \dot{\beta}_-^2\right)\right) - U(\beta_+,\beta_-)\right], \tag{63}$$

where the full non-linear potential is given by

$$\begin{aligned} U(\beta_+,\beta_-) = -2\left(e^{2\beta_+} + e^{-\beta_+ - \sqrt{3}\beta_-} + e^{-\beta_+ + \sqrt{3}\beta_-}\right) \\ + \left(e^{-4\beta_+} + e^{2\beta_+ - 2\sqrt{3}\beta_-} + e^{2\beta_+ + 2\sqrt{3}\beta_-}\right) \end{aligned} \tag{64}$$

The equations of motion for $a, \beta_+, \beta_-$ are given by

$$\frac{\ddot{a}}{a} + \frac{1}{2}\frac{\dot{a}^2}{a^2} + \frac{3}{8}\left(\dot{\beta}_+^2 + \dot{\beta}_-^2\right) - \frac{N_p^2}{6a^2}U(\beta_+,\beta_-) = 0\,, \tag{65}$$

$$\ddot{\beta}_\pm + 3\frac{\dot{a}}{a}\dot{\beta}_\pm + \frac{2}{3}\frac{N_p^2}{a^2}U_{,\beta_\pm} = 0\,. \tag{66}$$

Expanding the last equation to linear order we obtain

$$\ddot{\beta}_\pm + 3\frac{\dot{a}}{a}\dot{\beta}_\pm + 8\frac{N_p^2}{a^2}\beta_\pm = 0\,. \tag{67}$$

A comparison with Eq. (49) confirms that the $\beta$s are non-linear versions of the $l = 2$ modes – more specifically, they are non-linear versions of two $l = 2$ modes which are such that they preserve the Bianchi IX symmetry. To match with our earlier normalization conventions, one has to re-scale

$$\beta_\pm = \frac{1}{\sqrt{3}\pi}\phi_\pm\,, \tag{68}$$

where $\phi_\pm$ denote two separate $l = 2$ modes. The structure of the potential $U$ shows that when going beyond linear order, the equations of motion lead to direct couplings between these two $l = 2$ modes.

In the present section we work in a gauge where $N_p = 1$ and where one then has to determine the value of the time coordinate of the final hypersurface on which the boundary conditions $q_0 = 0, q_1 = a_1^2, \phi_\pm = \phi_{1\pm}$ are satisfied. This is done using the shooting method discussed in[17]. In this method, the (generally complex valued) second time derivatives of $\phi_\pm$ at the no boundary point $a = 0$ are adjusted using an

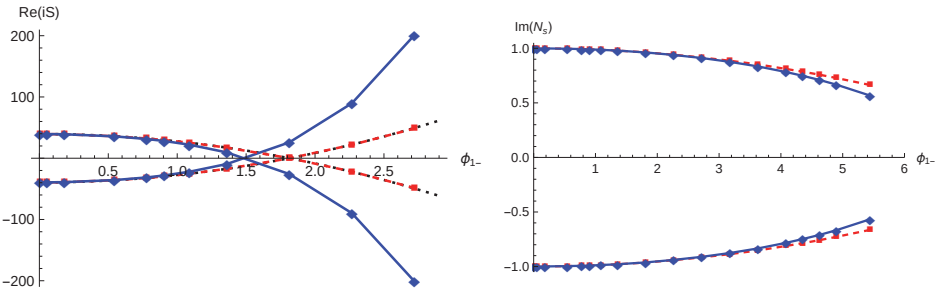



Fig. 5. These graphs show the weighting at the saddle points (left panel) and the imaginary part of the saddle point locations (right panel) as a function of the $l = 2$ anisotropy mode amplitude $\phi_{1-}$, for $\Lambda = 3$. In the plot of the action (left) the line starting at $+12\pi^2/\Lambda = +4\pi^2$ for $\phi_{1-} = 0$ corresponds to the saddle points in the lower half $N$ plane, while the line starting at $-4\pi^2$ corresponds to the saddle points in the upper half plane. In black (mostly hidden behind the red line) are the linear results without backreaction, in red the results including backreaction but still in linear perturbation theory, and in blue the results stemming from solving the fully backreacted Einstein equations. For values of $\phi_{1-}$ below 1 the linear and non-linear results agree to high precision, while one can see that at larger values of the anisotropy the non-linear corrections enhance the instability of the fluctuations, and move the saddle points further towards the real $N$-axis. Note that the weighting of the upper saddle points surpasses that of the lower ones when backreaction is still entirely negligible. Moreover, the non-linear effects of the full Einstein equations imply that the (unstable) upper saddle points come to dominate already for smaller amplitudes of the fluctuations.

optimization algorithm such that at a final time $t_f$ the desired real values $q_1, \phi_{1\pm}$ are simultaneously reached. The total time interval $\int N_p \mathrm{d}t_p = t_f$ can then also be related to the lapse function $N$ using the change of coordinates $N_p \mathrm{d}t_p = N q^{-1/2} \mathrm{d}t$,

$$N = \int_0^1 N \mathrm{d}t = \int_0^{t_f} a(t_p) \mathrm{d}t_p \,. \tag{69}$$

Our results are shown in Fig. 5. For ease of comparison with linear perturbation theory, we only show results for the case where a single $l = 2$ mode (here chosen to be $\phi_{1-}$) takes on a non-trivial value on the final hypersurface. The left panel shows how the weighting of the saddle point solution (for saddles in the upper half plane) increases as the perturbation amplitude is increased. The opposite behavior is seen for the saddle points in the lower half plane. As is evident from the figure, backreaction at second order in perturbation theory is utterly negligible. Even more importantly, the effects of the instability are even stronger when non-linear terms are included, and the dominance of the upper saddle point over the lower ones occurs already for smaller values of $\phi_{1-}$ than in the linear theory. Also, as shown in the right panel, the saddle point moves faster towards the real $N$-axis in the non-linear theory. These results consolidate our analytic results, and indicate that the inclusion of non-linear terms only reinforces the instability that we have identified.

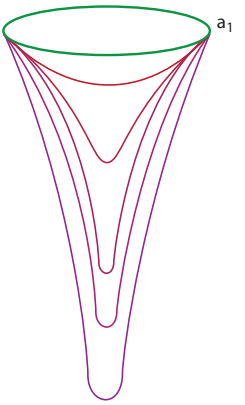



Fig. 6. We may consider a path integral over purely regular geometries, having as their only boundary a large late time universe with scale factor $a_1$. The integral can be pictured as a sum over complexified 4-spheres, bearing in mind that a complexified 4-sphere contains a Lorentzian de Sitter section. One may argue that such a restricted sum is closest in spirit to the original idea of the no-boundary proposal[1,2].

### 4.2. *Sums over regular metrics*

The idea of the no-boundary proposal is to sum over compact and regular metrics. However, as we have seen, in the minisuperspace example discussed so far there were many singular metrics off-shell, although the saddle point geometries were regular. One may wonder if one should restrict the sum in the path integral even furthst, so as to allow only regular metrics off- and on-shell. Could this lead to a different result? Following[18] we can consider the simplest and most symmetric possibility, namely a sum over 4-spheres with given boundaries $a_0 = 0$ and $a_1 > 0$ and arbitrary radius,

$$a(t) = \pm r \sin\left(\frac{N_E\, t}{r}\right) \tag{70}$$

with $a_1 = \pm r \sin\left(\frac{N_E}{r}\right)$ and $\dot{a}_1 = \pm N_E \cos\left(\frac{N_E}{r}\right)$. Accordingly, one should think of $N_E$ as being fixed by the boundary conditions and the sum to be over $r$. Given that $N_E$ will in general be a complex number, we should also expect $r$ to be complex, and that the integral will be over a contour in the complex $r$ plane. As discussed in[11], in fact the relevant saddle point geometry remains the same, and the associated perturbations remain unstable. Moreover, one can also explicitly check that in this calculation backreaction effects are small, so that the instability is confirmed.

## 5. Outlook

Despite the fact that intuitively, the idea of a smooth beginning for spacetime is very attractive, and of clear aesthetic appeal, our results have shown that upon closer inspection it does not work, as it leads to unsuppressed fluctuations (see Fig. 7).

One lesson to be drawn from this result is that it is worthwhile analysing the mathematical setup of a physical theory in some detail, as heuristic arguments can be misleading in unexpected ways. Moreover, quantum gravity appears to be much more restrictive than naively thought. This is in fact an extremely encouraging thought, as it may allow us to make progress regarding the unification of quantum theory and general relativity.

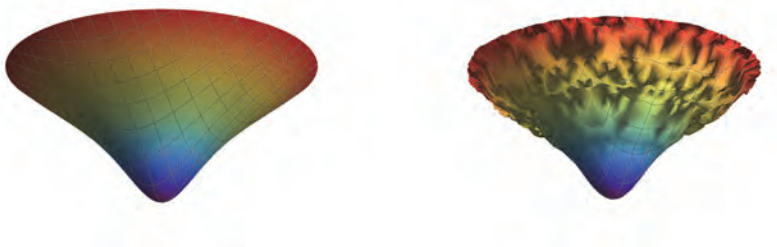


Fig. 7. The beginning of the universe is depicted at the bottom, while our current state of the universe should be thought of as the boundary at the top. *Left panel:* The originally hoped for situation of a smooth beginning. *Right panel:* A detailed investigation of the path integral shows that fluctuations are unstable, and the probability for large perturbations is unsuppressed. Thus the no-boundary and tunneling proposals are untenable in their current forms.

The most important result of our investigations is probably that the Lorentzian path integral is well defined: it gives sensible results for other types of boundary conditions, in particular for classical boundary conditions[7], and it is in fact simply the no-boundary condition that fails. Moreover, our analysis clearly shows that the Euclidean path integral does not exist (see Fig. 3 where one can see that integral along imaginary lapse direction necessarily diverges, either near $N = 0$ or towards imaginary infinity). Other, intrinsically complex integration contours for the lapse were proposed in reaction to our work[19,20], but these ultimately lead to inconsistent results[9,10]. Thus one may take the Lorenztian path integral as a solid basis for further investigations into quantum cosmology, allowing us to explore questions such as: what are the conditions for quantum field theory in curved spacetime to be a good approximation? Are the complex big-crunch-avoiding solutions that were found heuristically in[21] of physical relevance or not? How severe is the trans-Planckian problem of inflation? Which assumptions must be modified in order to obtain a consistent theory of initial conditions? The search remains wide open!

**Acknowledgments**

My sincere thanks to Alice Di Tucci, Job Feldbrugge and Neil Turok for enjoyable and enlightening collaborations on these subjects. I would also like to thank Sebastian Bramberger, Juan Diaz Dorronsoro, Jonathan Halliwell, Thomas Hertog, Oliver Janssen, Alex Vilenkin and Yannick Vreys for stimulating discussions. The

author gratefully acknowledges the support of the European Research Council in the form of the ERC Consolidator Grant CoG 772295 “Qosmology”.

## References

1. S. W. Hawking, *Pontif. Acad. Sci. Scr. Varia* **48** (1982) 563
2. J. B. Hartle and S. W. Hawking, *Phys. Rev. D* **28** (1983) 2960
3. S. W. Hawking, *Nucl. Phys. B* **239** (1984) 257
4. A. Vilenkin, *Phys. Lett.* **117B** (1982) 25.
5. A. Vilenkin, *Phys. Rev. D* **27** (1983) 2848.
6. A. Vilenkin, *Phys. Rev. D* **30** (1984) 509.
7. J. Feldbrugge, J. L. Lehners and N. Turok, *Phys. Rev. D* **95** (2017) no.10, 103508 [arXiv:1703.02076 [hep-th]].
8. J. Feldbrugge, J. L. Lehners and N. Turok, *Phys. Rev. Lett.* **119** (2017) no.17, 171301 [arXiv:1705.00192 [hep-th]].
9. J. Feldbrugge, J. L. Lehners and N. Turok, *Phys. Rev. D* **97** (2018) no.2, 023509 [arXiv:1708.05104 [hep-th]].
10. J. Feldbrugge, J. L. Lehners and N. Turok, arXiv:1805.01609 [hep-th].
11. A. Di Tucci and J. L. Lehners, arXiv:1806.07134 [gr-qc].
12. C. Teitelboim, *Phys. Rev. D* **28** (1983) 297.
13. C. Teitelboim, *Phys. Rev. Lett.* **50** (1983) 705.
14. J. J. Halliwell, *Phys. Rev. D* **38** (1988) 2468.
15. E. Witten, *AMS/IP Stud. Adv. Math.* **50** (2011) 347 [arXiv:1001.2933 [hep-th]].
16. J. J. Halliwell and J. Louko, *Phys. Rev. D* **39** (1989) 2206.
17. S. F. Bramberger, S. Farnsworth and J. L. Lehners, *Phys. Rev. D* **95** (2017) no.8, 083513 [arXiv:1701.05753 [hep-th]].
18. J. J. Halliwell and J. Louko, *Phys. Rev. D* **40** (1989) 1868.
19. J. Diaz Dorronsoro, J. J. Halliwell, J. B. Hartle, T. Hertog and O. Janssen, *Phys. Rev. D* **96** (2017) no.4, 043505 [arXiv:1705.05340 [gr-qc]].
20. J. Diaz Dorronsoro, J. J. Halliwell, J. B. Hartle, T. Hertog, O. Janssen and Y. Vreys, *Phys. Rev. Lett.* **121** (2018) no.8, 081302 [arXiv:1804.01102 [gr-qc]].
21. S. F. Bramberger, T. Hertog, J. L. Lehners and Y. Vreys, *JCAP* **1707** (2017) no.07, 007 [arXiv:1701.05399 [hep-th]].

# The role of redundancy in blind signal estimation for multiple gravitational wave detectors

Hao Liu*

*The Niels Bohr Institute & Discovery Center, Blegdamsvej 17, Denmark*
*Key laboratory of Particle and Astrophysics, Institute of High Energy Physics, CAS, 19B YuQuan Road, Beijing, China.*
*E-mail: liuhao@nbi.dk*

James Creswell, Sebastian von Hausegger, and Pavel Naselsky

*The Niels Bohr Institute & Discovery Center, Blegdamsvej 17, Denmark*
*E-mail: james.creswell@nbi.ku.dk, s.vonhausegger@nbi.dk and naselsky@nbi.dk*

Andrew D. Jackson

*The Niels Bohr International Academy, Blegdamsvej 17, DK-2100 Copenhagen, Denmark*
*E-mail: jackson@nbi.dk*

In this work, we extend our previous blind GW signal estimation method [1] to the case of multiple detectors and show that, with a full use of redundancy, it gives promising results, e.g. a faster decay of fluctuations than that expected from the central limit theorem. This method, whose design explicitly accounts for redundancy in multiple measurements, considerably improves the efficiency of signal extraction in a multi-detector network.

## 1. Introduction

Given that detector noise is several orders of magnitude larger than expected gravitational wave (GW) signals from black hole mergers, the analysis of LIGO's data is challenging. This is one reason why LIGO consists of two independent detectors separated by 3000 km. It is reasonable to divide the task into three relatively distinct parts. These include event detection, waveform extraction, and the identification of its physical origin. Here, we will be concerned with the two final steps in this process, focusing on black hole mergers like GW150914. Since thermal and seismic effects lead to substantial low frequency noise and since quantum noise is dominant at high frequencies, it is possible to reduce the effects of noise by band-pass filtering (or equivalent operations) to a restricted frequency range e.g., 30 – 300 Hz. LIGO extracts waveforms by comparing the data from the Livingston and Hanford detectors with (suitably time-shifted) elements of a bank of templates describing the merger of two black holes with various masses, spins and initial conditions. A satisfactory level of agreement would then lead to the acceptable conclusion that the data is *not inconsistent* with a black hole merger. An equally satisfactory conclusion would be that, *if* the event is due to a black hole merger, it has the parameters corresponding to the best-fit template. However, in our view, the far stronger conclusion that such agreement proves that the event actually *is* due to a black hole merger with the corresponding parameters is unwarranted.

Moreover, there are disturbing features in the noise associated with LIGO's events. Although this noise is complex, its two essential characteristics are clear and unambiguous. First, LIGO noise is neither stationary nor Gaussian.[2] In fact, both stationarity and Gaussianity are implicitly assumed in the Bayesian likelihood analysis as a diagonal approximation for the noise covariance matrices for both the Hanford and Livingston detectors.[3] Second, the LIGO "residuals" (defined as the cleaned data minus the best-fit template) for the two detectors should not be correlated. In fact, the residuals associated with GW150914 are strongly correlated when shifted by the same time lag as the template itself and considerably larger than the expected level of "accidental" correlations for LIGO data.[4,5] Since the Hanford (H) and Livingston (L) detectors are assumed to be completely independent except for a possible GW signal, this correlation in the residuals tells us that a template-based analysis does not provide a reliable description of the common signal seen by the H and L detectors.

Template-based analyses have a strong tendency to be self-fulfilling, this alone should be sufficient to emphasize the importance of maintaining a clean separation between the extraction of a signal and attempts to divine its physical origin.

In order to make optimal use of the independence of the H and L detectors and to quantify the uncertainties in previously extracted waveforms, we have constructed a template-free method that includes minimization of residual correlations in order to determine a "best common signal" for GW150914.[1] Since the residual for an individual detector is a complicated function of time (or frequency), there are many ways to realize the desired absence of correlations. This means that our algorithm yields a family of best common signals that enable us to estimate the probability, $p$, that the best common signal is a black hole merger template. For GW150914, we find $p = 0.008$. It is unclear whether this small probability is indicative of inadequacies in the gravitational wave templates, imperfect knowledge of the acceptance of the instruments, or of a completely different physical explanation — either astrophysical or terrestrial — of this event. As the amplitude of the signal becomes smaller with respect to the noise level, the relative width of the envelope of best common solutions obtained by this method grows and eventually covers the line of zero signal. The signal to noise ratio (SNR) for other possible black hole merger events is smaller than that found for GW150914, and we find that all of these subsequent events are consistent with a best common signal of precisely zero.

The fact that LIGO does not make use of the only relevant information available about noise properties (i.e., the fact that L and H residuals must be uncorrelated) also suggests that their analysis does not optimally exploit the benefits of redundancy that should accrue from two independent measurements of the same signal. Since the power of redundancy becomes clearer as the degree of redundancy is increased, our primary concern in this paper is to extend the results of Ref. 1 to the case of an arbitrary number of GW detectors. We emphasize that this approach is not merely a pedagogic exercise. There is considerable current discussion about adding additional detectors to the GW network. Since the effects an additional

detector can have on the accuracy of signal attraction depends sensitively on the degree to which redundancy is exploited, we believe that the present work can be of practical relevance in deciding the extent to which such an investment is justified.

We begin in section 2 by presenting simple schematic models that illustrate the dramatic differences found when analyzing multi-detector events with or without consideration of the effects of redundancy. These models will provide a gauge of the extent to which redundancy is realized in practice. The extension of our earlier model for the blind estimation of the common signal to the case of many detectors will be described in section 3 along with the results of realistic simulations. Finally, section 4 contains a discussion based on these results.

## 2. The value of redundancy and its price

We all have an intuitive understanding of the important role that redundancy can play in the accurate determination of a signal transmitted in the presence of noise. When told an important telephone number in a noisy environment, our immediate reaction is to ask for it to be repeated. Hearing the same number twice greatly increases our confidence that it has been transmitted correctly. Clearly, redundancy is equally important in a scientific context, and it was surely one of the primary reasons that LIGO wished to have multiple GW detectors. In the analysis of GW170814, LIGO makes use of the Virgo detector as a consistency check for the results from Hanford and Livingston.[6] The additional information from Virgo improves the false-alarm rate compared to the two detector case. In general, the SNR of the LIGO/Virgo network is defined as a sum in quadrature of the individual detector SNRs, i.e.[7]

$$\rho^2_{\text{network}} = \sum_{i=1}^{M} \rho_i^2. \tag{1}$$

The precondition for this definition is ideal detector noise uncorrelated between detectors, with off-diagonal terms of the noise covariance matrix neglected. Furthermore, the addition of detector SNRs in this way fails to properly exploit redundancy, because the network SNR can be dominated by the SNR of a single detector. A high network SNR does not necessarily imply agreement between detectors.

Our aim in this section is to provide a better understanding of both the power of redundancy and the price that must be paid to obtain it.

We first consider a schematic but instructive example of redundancy in which there are two independent measurements of a common signal consisting of $N$ pieces of data. In the absence of noise, the two measured signals will be identical to one another and to the true signal. Now, simulate noise by assuming that there is a probability, $p$, that any given piece of genuine data, $d_i$, has been replaced by noise $n_i$ (with $n_i \neq d_i$). The probability that both of the measured signals are free of errors and therefore correct is evidently $(1-p)^{2N}$, which vanishes exponentially as the signal becomes more complex (i.e., $N \to \infty$). It is necessary, however, to

consider the possibility of false positives, for which the signals are not correct in spite of being identical. To investigate this question, assume that the probability that two randomly drawn pieces of noise are identical is $q$. (E.g., If the $n_i$ are randomly drawn digits between 0 and 9, $q = 1/9$.) The probability that the two signals will be identical (either genuinely or accidentally) is

$$\sum_{m=0}^{N} (1-p)^{2(N-m)} p^{2m} q^m \frac{N!}{(N-m)!m!} = \left[(1-p)^2 + p^2 q\right]^N \quad . \tag{2}$$

Thus, the probability that identical results of the two measurements will actually be correct is

$$\left[\frac{(1-p)^2}{(1-p)^2 + p^2 q}\right]^N \quad .$$

The extension of this problem to the case where there are $M$ detectors is straightforward. In this case, all $M$ detectors see the same event. The probability that all of the measured signals are free of errors (and therefore measured correctly) is evidently $(1-p)^{MN}$, which clearly decreases exponentially with increasing $M$. The generalization of Eq. 2 describing the probability that these $M$ signals will be identical (but not necessarily correct) then becomes

$$P = \left[(1-p)^M + p^M q^{M-1}\right]^N \quad , \tag{3}$$

and thus the probability that identical results for all $M$ measurements will actually be the correct signal is

$$P = \left[\frac{(1-p)^M}{(1-p)^M + p^M q^{M-1}}\right]^N \quad . \tag{4}$$

To illustrate this result, consider the transmission of an 8-digit telephone number in a noisy environment. For this case, assume $N = 8$, $p = 1/4$, and $q = 1/9$ and consider the cases of $M = 2$, 3, and 4 detectors. The probability that identical (but not necessarily correct) results will be obtained for all $M$ measurements is approximately 0.01, 0.001, and 0.0001, respectively. The probabilities that these identical results will be false are 0.0935, 0.00365, and 0.000135, respectively. The exponential decrease of this error is evident and emphasizes the dramatic improvement in accuracy that results from $M$ identical measurements of the same signal. Unfortunately, it also reminds us that the probability of actually obtaining $M$ identical signals also decreases exponentially with $M$.

The preceding example might seem to indicate that redundant measurement of a given signal will lead to signal detection and extraction with a confidence that grows to 1 exponentially with the redundancy $M$. We now wish to consider a second example to illustrate that this is not necessarily the case. To this end, consider $M$ independent measurements each of which for simplicity consists of a common signal, $s$, and an independent realization of $N$ pieces of random noise.

The data obtained at detector $k$ is thus $d^{(k)} = s + n^{(k)}$.[a] We wish to extract $s$ from this data without making unwarranted assumptions about the noise such as stationarity and/or Gaussianity. Given this strong constraint, the only assumption that can be made is that the measurements in the various detectors are genuinely independent and that there are therefore no correlations in their noise realizations To be concrete, we will imagine that cross-correlations are given as the Pearson cross-correlation to be adopted below.[b]

In these circumstances, it might seem natural to approximate the best common signal as the average record

$$\mathcal{S} = \frac{1}{M}\sum_{k=1}^{M} d^{(k)} = s + \frac{1}{M}\sum_{k=1}^{M} r^{(k)} \ . \tag{5}$$

Given this guess, we can re-express the individual data strings as $d^{(k)} = \mathcal{S} + \rho^{(k)}$ where the residuals are given as

$$\rho^{(k)} = \frac{1}{M}\left[(M-1)r^{(k)} - \sum_{j\neq k} r^{(j)}\right] \ . \tag{6}$$

For sufficiently large $N$, it is reasonable to make the approximation that the cross-correlators $C(s, r^{(k)}) = 0$ for all $k$ and $C(r^{(j)}, r^{(k)}) = 0$ for all $j \neq k$. This leads to the result that

$$C(s,\mathcal{S}) = \frac{1}{\sqrt{1+1/M}} \approx 1 - \frac{1}{2M} \ \text{ and } \ C(\rho^{(j)},\rho^{(k)}) = -\frac{1}{M-1} \ . \tag{7}$$

It is true that $\mathcal{S}$ converges to the exact result $s$ and that the correlations between the residuals, $\rho^{(k)}$, vanish as expected in the limit of large $M$. Unfortunately, these convergence rates, which are an elementary consequence of the central limit theorem, are far too slow to be useful. The fact that all of the $M(M-1)/2$ correlators between the residuals have the same value of $-1/(M-1)$ is also unphysical. Thus, the assumption that the signal can be approximated by Eq. 5 is unjustified. For the case $M = 2$, the cross-correlator $C(s, \mathcal{S})$ has the unsatisfactorily small value of $\sqrt{2/3}$, and there is a perfect anti-correlation between the residuals with $C(\rho_1, \rho_2) = -1$. As noted in our earlier work[4,8], similarly large and unphysical correlations in the residuals determined by LIGO for GW150914[9] suggest the existence of problems with the corresponding GW signal.[c]

[a] Note that this is the model used for Bayesian analysis of the LIGO events.[3]

[b] For two vector records of length $N$, we first shift the records so that each has average value zero and rescale them so that the scalar product of each vector with itself is 1. The Pearson cross-correlation is then the scalar product of these shifted and rescaled vectors and will have a value between $-1$ and $+1$.

[c] The presence of correlations in the Hanford and Livingston residuals determined in original template-based analysis of GW150914 raises questions about this analysis. It is important for others to confirm the existence of these correlations. This can be done using the publicly available data, as described on our webpage:
`http://www.nbi.ku.dk/gravitational-waves/residual_correlations_notebook.html`

The two examples presented in this section can serve as a measure of the extent to which the benefits of redundancy have been realized by a given method of signal extraction. When redundancy is not exploited, we find that the extracted signal converges to the true signal with an error that vanishes slowly (i.e., like $1/\sqrt{M}$) as the degree of redundancy increases. In contrast, a maximal implementation of redundancy leads to an exponentially decreasing error rate as a function of $M$. This test will be applied in practice in the following sections of this paper. It should be noted, however, that this increased confidence level has a relatively high price. As we have seen, the probability that an event will pass the redundancy test also vanishes exponentially with $M$. Thus, if event rates are too low, it may be impossible to realize fully the benefits of redundancy. We stress that the examples here are highly schematic and are intended to illustrate the general fact that the benefit of multiple independent measurements depends sensitively on the way these measurements are analyzed. They do not tell us to how the benefits of redundancy can be maximally realized in the case of GW data. We consider one such approach in the following section.

## 3. Application of the blind estimation method to multiple GW detectors

### 3.1. *Basis of the blind estimation method: cross-correlation and Fisher transformation*

We briefly review our previous work on blind estimation. [1] The strain signal detected by LIGO in the $i$-th detector is assumed to be

$$X_i(t) = a_i \cdot h(t, \Delta\tau_i, \Delta\theta_i) + N_i(t), \tag{8}$$

where $X_i(t)$ is the total strain data, $N_i(t)$ is the noise, and $a_i \cdot h(t, \Delta\tau_i, \Delta\theta_i)$ is the gravitational wave signal with given amplitude $a_i$, time lag $\Delta\tau_i$ and phase shift $\Delta\theta_i$, which contain both contributions from projection and detector acceptance. As mentioned in Ref. 1, we pre-match detector data to roughly remove the contributions of $\Delta\tau_i$ and $\Delta\theta_i$ (this can be done precisely with, e.g., an EM-counterpart), and then the equation becomes

$$X_i(t) = a_i \cdot h(t) + N_i(t). \tag{9}$$

For a blind estimation, we also need to further derive the residual noise, for which we consider two data sets $X_1$ and $X_2$ of length $N$ (e.g. cleaned strain data from two independent detectors) which contain a common signal $A$. The amplitude of $A$ could potentially be different in each detector, either due to projection or detector acceptances. For convenience, we assume that $X_1$, $X_2$, and $A$ have been shifted to have zero average values and normalized to have variance unity. Then the Pearson cross-correlation coefficient of two such vectors, $C_{X_1 X_2}$, is simply the inner product $S_{X_1 X_2}$ given by

$$S_{X_1 X_2} = \frac{1}{N-1} \sum_{k=1}^{N} X_1(k) \cdot X_2(k). \tag{10}$$

The residuals are defined as

$$R_i = X_i - A \cdot \frac{S_{AX_i}}{S_{AA}} = X_i - A \cdot S_{AX_i}. \tag{11}$$

As mentioned above, the amplitude of the term $A$ (i.e. $S_{AX}/S_{AA}$) can be different for two dececctors. By construction, the correlations of both $R_1$ and $R_2$ with $A$ are zero.

The criterion for determining the blind estimate of $A$ is to maximize the $C_{AX_i}$ while simultaneously minimizing the cross-correlation between the residuals $C_{R_1 R_2}$. Note that the residuals $R_1$ and $R_2$ are not automatically normalized.

For familiarity and simplicity, we obtain approximate Gaussianity of the resulting correlations by using the Fisher transformation:[10]

$$Z_{XY} = \frac{1}{2} \log \left( \frac{1 + C_{XY}}{1 - C_{XY}} \right). \tag{12}$$

### 3.2. *Extension of the likelihood approach*

In our previous work,[1] a blind GW-template estimation was done by considering the likelihood that a given initial guess, $A$, is the common signal observed by two detectors ($X_1$ and $X_2$) as

$$\log(L) = Z^2_{AX_1} + Z^2_{AX_2} - k Z^2_{R_1 R_2}, \tag{13}$$

where $Z_{AX_1}$ and $Z_{AX_2}$ represent the similarity between $A$ and the detector data as measured by the Pearson cross correlation, $Z_{R_1 R_2}$ represents the similarity between the residuals from the two detectors, and $k$ is a constant factor determining the relative weight of the two contributions. The likelihood function is designed to be maximized at higher $Z^2_{AX_1}$ and $Z^2_{AX_2}$ and lower $Z^2_{R_1 R_2}$. The initial guess $A$ is then improved by a random walk approach until the likelihood reaches an oscillatory region (see Fig. 2 of Ref. 1). The oscillatory region is used to estimate the range of fluctuation for each pixel as shown in Figs. 4 and 5 of Ref. 1.

The above method was initially designed for the GW150914 event for which there were only two detectors in the network. A natural extension of the method is to apply it to multiple detectors, so the likelihood function becomes:

$$\log(L) = \sum_{i=1}^{M} (Z_{AX_i})^2 - k \sum_{i=1}^{M-1} \sum_{j=i+1}^{M} Z_{R_i R_j}, \tag{14}$$

where $M$ is the number of detectors, $X_i$ is the data from the $i$-th detector, $A$ is the blind estimate of the signal, and $R_i$ is the residual after removing this estimate from the $i$-th detector. With this modified likelihood function, the blind estimation method presented in Ref. 1 can be extended to the case of multiple detectors.

Also note that this kind of likelihood approach is different from a likelihood approach that assumes either a known covariance matrix or a known theoretical model, or even both. Here, neither the covariance matrix nor the theoretical model is assumed, and one starts only from basic ideas about "correlated signals and uncorrelated noise". Thus the likelihood approach here is totally blind, i.e. it makes minimal assumptions. This is unlike template-free methods currently in use by LIGO, such as BayesWave and oLIB, which begin with the assumption of a stationary Gaussian noise model.[11,12] Finally, we note that if there is any reliable additional information, such as a known correlation between two of the detectors, or an especially low SNR in one of the detectors, then such information can also be added to Eq. 14 by changing the weights of the corresponding terms.

### 3.3. *Test and results*

To test the performance of the blind estimation method for the case of multiple detectors, we run a simulation as follows: We select the GW150914 waveform template as the input "real signal", and inject it into genuine strain data taken $2, 3, \ldots, M+1$ seconds after the GW150914 event to simulate the data from multiple detectors. Here, we are mainly interested in the trend of how the error of estimation decreases with increasing number of detectors. Thus, for convenience, we assume identical projections and similar noise levels for all detectors (see also Sec 3.1 of Ref. 1). In practice, when multiple detectors have different projections and signal-to-noise ratios, the overall performance will become worse. For simplicity, we will neglect these concerns here.

For comparison, we adopt a reference estimator of the common signal that is simply the average of the data from the individual detectors and further assume that the detector noise is Gaussian. The estimated error will then scale like $1/\sqrt{M}$. We shall compare the performance of our method with this reference.

As mentioned above, the range of fluctuations is an immediate result from the oscillations of the likelihood function caused by chance correlations. In the simulation here, we select the 10th and 90th percentiles of the fluctuation range and use their difference as a measure of the range of fluctuation for each pixel. This quantity is averaged over the entire time range of the event as the final estimator of the uncertainty of blind estimation:

$$E_M = \langle S_{90}(t) - S_{10}(t) \rangle_M \,. \tag{15}$$

The result of the simulation described above is given in Fig. 1 where it is apparent that the blind estimation (black) performs better than the $1/\sqrt{M}$ reference (red). This result is not surprising because the $1/\sqrt{M}$ reference is obtained by simple averaging without consideration of the residual correlations. In other words, an estimation method that considers both the correlations between signal and data (first term in Eq. 14) and the correlations between residuals (second term in Eq. 14) will certainly give better results than simple averaging.

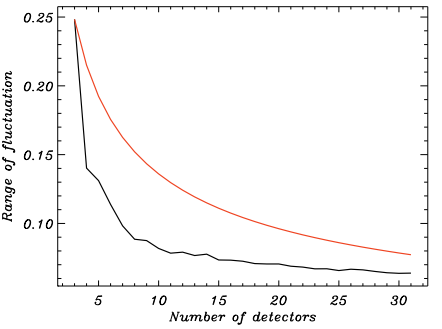



Fig. 1. The range of fluctuation defined by Eq. 15. The horizontal axis shows the number of detectors, and the vertical axis shows the estimator of uncertainty, $E_M$, calculated using the convention that strain data is shifted to have zero mean and normalized to have variance unity. The black line is the performance of the multi-detector blind estimation method, and the red line is the expectation of the central limit theorem that scales as $1/\sqrt{M}$ given by simple averaging, which has been normalized to the black line at $M = 3$.

We also show the input signal and the range of error obtained from blind estimations for the extreme example of 32 detectors in Fig. 2. We include, for test purporses, a slight variation of the above method in which the first term in Eq. 14, $\propto Z^2_{AX_i}$, is dropped. One can see that even without this seemingly essential term, one still gets unbiased estimations of the "real signal". Only the error of estimation is larger.

The error of the estimated common signal can also be evaluated by comparison to the real signal $h(t)$. Three cases are considered: the simple average of multiple detectors, $B_1(t)$; the average of the blind estimations, $B_2(t)$; and the average of blind estimations calculated only using the residual correlation terms in Eq. 14, $B_3(t)$. For each of the three cases, we determine the deviation from the real signal $h(t)$ as

$$\begin{aligned} \delta_1 &= B_1(t) - h(t) \\ \delta_2 &= B_2(t) - h(t) \\ \delta_3 &= B_3(t) - h(t), \end{aligned} \tag{16}$$

and calculate the standard deviations as $\sigma_1$, $\sigma_2$ and $\sigma_3$ respectively. For a given number of detectors, we calculate two ratios

$$\begin{aligned} r_{12} &= \sigma_1/\sigma_2 \\ r_{13} &= \sigma_1/\sigma_3, \end{aligned} \tag{17}$$

which are defined such that larger values correspond to better performance than simply averaging. Our blind estimation method is expected to give smaller uncertainties than simple averaging, thus we should see $r_{12} > 1$. Larger values of $r_{12}$ indicate better results given by the blind estimation. On the other hand, we expect

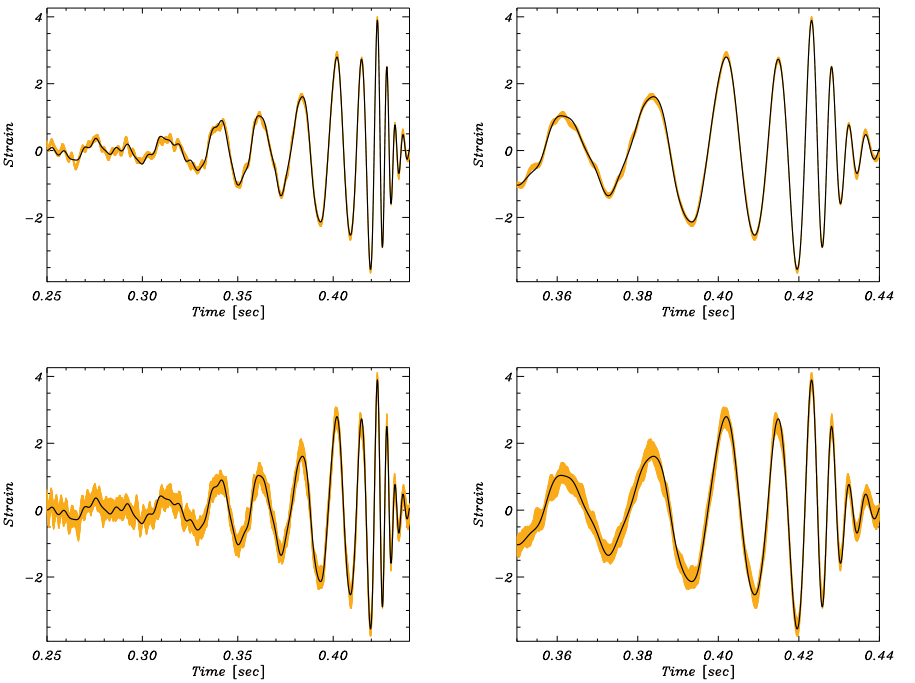



Fig. 2. The real input template (black) and the 10%–90% range of uncertainty (see Eq. 15) given by the blind estimation for the case of 32 detectors. Upper: using the whole likelihood, as in Eq. 14. Lower: same as the upper panels but only the residual terms are used. Left: the full time range. Right: only the second half.

that a blind estimation only using the correlations of the residuals is worse than simple averaging, but it should still be a reasonable estimation, thus $r_{13}$ should be less than 1 but not much lower.

In Fig. 3, we show $r_{12}$ and $r_{13}$ as functions of the number of detectors. The ratio $r_{12}$ lies around 1.2–1.4, indicating that simple averaging gives a 20%–40% larger error than blind estimation. Also, $r_{13}$ is about 0.8–0.9, indicating that, even from only regarding the residual–residual terms of the likelihood, one can still get a reasonably good estimation of the real input signal. Therefore, Figs. 2 and 3 suggest that to only compare template to data and ignore residuals is an inefficient use of experimental resources. We also see from Fig. 3 that, if the number of detectors is less than 10, then the blind estimation method is significantly better than simple averaging. With increasing $M$, however, the improvement of the method slows. This is possibly due to the finite record length. Since in the near future, the number of GW detectors will not exceed 10, a blind estimation method such as this one is especially important.

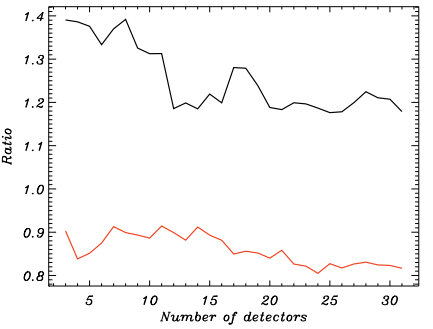



Fig. 3. The ratios $r_{12}$ (black) and $r_{13}$ (red) as functions of the number of detectors. $r_{12} > 1$ means our blind estimation gives lower estimation error than a simple averaging, while $r_{13}$ close to 1 means even without using the template–data term, one can still get a reasonably good estimation of the potential common signal using the same blind estimation method.

## 4. Discussions

The aim of the present paper is two-fold. On the one hand, an exploration of the role of redundancy in the analysis of LIGO data can help better to understand the reliability of events already observed. On the other, it can offer some guidance regarding the most fruitful way to analyze data in the coming multi-detector era of gravitational wave science. It is our firm conviction that template-based analysis alone is in principle circular and thus fundamentally flawed. Clearly, a bank that only contains templates for black hole mergers can never detect anything other than black hole mergers. The best that one can hope for is to claim that an event is *not inconsistent* with black hole merger and then make the best case possible that other possible origins — terrestrial as well as astrophysical — can be excluded. Rather, we are convinced that data analysis should begin with the template-free extraction of a best common signal that can later be compared with specific physical models. This conviction led to the development of the blind signal estimation initially presented in Ref. 1 and extended here. This method enabled us to determine the probability that the best common signal could be described by a gravitational wave template. Unfortunately, this probability is remarkably low. The probability that the common signal is LIGO's original published template was found to be $4\times10^{-6}$. The best GW template was found to have substantially higher masses (38 and 48 solar masses) and high spins (0.96 and $-0.85$, respectively). While better than the published template, the probability that it was the best common signal still had the unacceptably small value of 0.008. It should be noted that GW150914 is by far the strongest event seen. A similar template-free analysis of all other putative gravitational wave events is consistent with a common signal of zero.

Gravitational wave signals are characteristically much smaller than measured data indicating that the noise is much stronger than the signal. Thus, a reliable

blind analysis of GW data requires detailed knowledge of the origin and nature of detector noise. Unfortunately, it is generally acknowledged that this noise is neither Gaussian nor stationary. In order to make a reliable detection of a GW signal, it is essential that there be no correlation between the residuals observed at individual detectors. This obvious requirement lies at the heart of the justification for incurring the expense of performing redundant measurements of a given signal with two or more independent detectors. We have thus provided simple schematic examples to show that the accuracy of detection can be improved exponentially with an increasing degree of redundancy. This stands in sharp contrast to the far slower convergence expected from a simple average of the measured signals. These results are supported by simulations based on a realistic gravitational wave form and real LIGO noise data. These results indicate the important role played by the residual correlations in the data analysis. Indeed, it is possible to obtain a reasonably good estimation of the injected signal by using only these terms. In summary, these results provide a strong reminder of the importance of exploiting redundancy in the analysis of both present and future gravitational wave data and suggest that this can be accomplished by including the suppresion of residual correlations in the construction of a satisfactory likelihood function.

## Acknowledgments

The authors thank the organizers of the MG15 conference. This work has made use of the LIGO software package and data. Our research was funded in part by the Danish National Research Foundation (DNRF) and by Villum Fonden through the Deep Space project. Hao Liu is supported by the Youth Innovation Promotion Association, CAS.

## References

1. H. Liu, J. Creswell, S. von Hausegger, A. D. Jackson and P. Naselsky, A blind search for a common signal in gravitational wave detectors, *JCAP* **2**, p. 013 (February 2018).
2. B. P. Abbott *et al.*, Sensitivity of the Advanced LIGO detectors at the beginning of gravitational wave astronomy, *Phys. Rev.* **D93**, p. 112004 (2016), [Addendum: Phys. Rev.D97,no.5,059901(2018)].
3. J. Veitch *et al.*, Parameter estimation for compact binaries with ground-based gravitational-wave observations using the LALInference software library, *Phys. Rev.* **D91**, p. 042003 (2015).
4. H. Liu and A. D. Jackson, Possible associated signal with GW150914 in the LIGO data, *JCAP* **10**, p. 014 (October 2016).
5. J. Creswell, S. von Hausegger, A. D. Jackson, H. Liu and P. Naselsky, On the time lags of the LIGO signals, *JCAP* **8**, p. 013 (August 2017).

6. B. P. Abbott *et al.*, GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence, *Phys. Rev. Lett.* **119**, p. 141101 (2017).
7. C. Cutler and E. E. Flanagan, Gravitational waves from merging compact binaries: How accurately can one extract the binary's parameters from the inspiral waveform?, *Phys. Rev. D* **49**, 2658 (Mar 1994).
8. P. Naselsky, A. D. Jackson and H. Liu, Understanding the LIGO GW150914 event, *JCAP* **8**, p. 029 (August 2016).
9. LIGO Scientific Collaboration and Virgo Collaboration, Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, p. 061102 (Feb 2016).
10. R. A. Fisher, Frequency distribution of the values of the correlation coeffients in samples from an indefinitely large population, *Biometrika* **10**, 507 (1915).
11. N. J. Cornish and T. B. Littenberg, BayesWave: Bayesian Inference for Gravitational Wave Bursts and Instrument Glitches, *Class. Quant. Grav.* **32**, p. 135012 (2015).
12. R. Lynch, S. Vitale, R. Essick, E. Katsavounidis and F. Robinet, Information-theoretic approach to the gravitational-wave burst detection problem, *Phys. Rev. D* **95**, p. 104046 (May 2017).

# Status of KAGRA and its science goals

Takaaki Kajita, for the KAGRA collaboration

*Institute for Cosmic Ray Research, The University of Tokyo, Kashia, 277-8582, Japan*
†*E-mail: kajita@icrr.u-tokyo.ac.jp*

KAGRA is an interferometric gravitational-wave detector with 3-km arms constructed at Kamioka, Gifu, Japan. One of the key features of KAGRA is the cryogenic mirrors for the 3 km arm cavities. KAGRA plans to begin the operation before the end of 2019. KAGRA plans to join the network observation of global gravitational wave interferometers.

*Keywords*: Gravitational waves, interferometer, cryogenic, underground.

## 1. Introduction

The gravitational wave astronomy has begun with the detection of gravitational waves (GWs) in the past three years. LIGO Scientific Collaboration and Virgo Collaboration have reported mergers of binary black holes[1,2] and a merger of binary neutron stars[3]. These observations have already proved that the gravitational waves are very important for the understanding of the Universe.

## 2. Overview of KAGRA

KAGRA[4] is a laser interferometer with 3 km arms, constructed in Kamioka, Gifu, Japan, and was in its final installation phase as of August 2018. After the conference, in April 2019, the construction has been completed. KAGRA is located underground for smaller seismic noise. KAGRA will use 4 cryogenic mirrors for the 3 km arm cavities to reduce thermal noise around the detector's most sensitive band at around 100 Hz. The KAGRA collaboration is an international collaboration with more than 200 collaborators.

The optical configuration of KAGRA is similar to LIGO and Virgo, with 3 km arm length and various optical cavities in the arms and recycling systems. Each mirror is suspended by vibration isolation system. Depending on the requirement on the vibration isolation of the mirrors, there are essentially 4 types of vibration isolation systems. For example, the arm cavity mirrors are suspended by a 13.5-meter-long vibration isolation system. Figures 1 and 2 show the interferometer configuration and 4 types of vibration isolation systems.

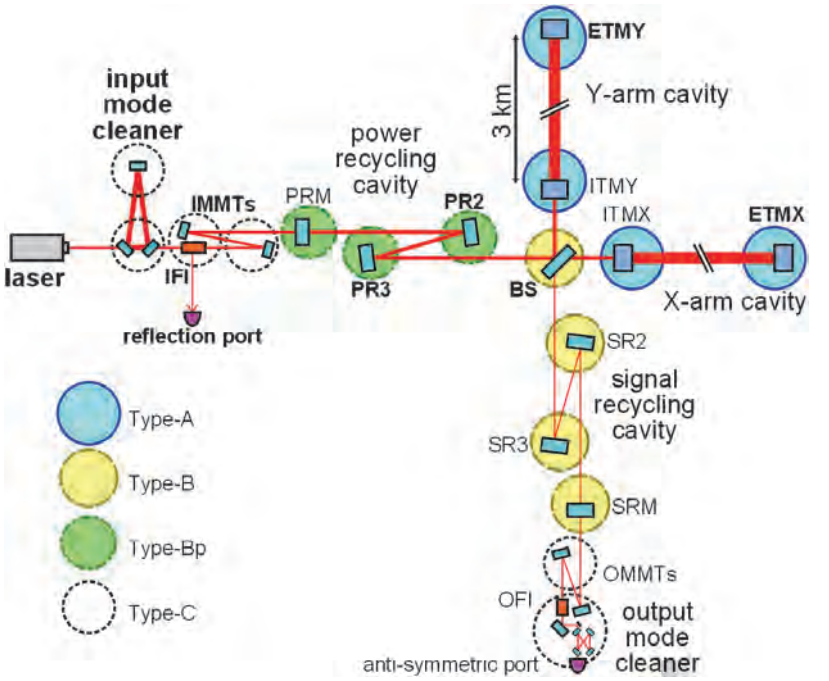



Fig. 1. Schematic of the KAGRA interferometer. All the mirrors shown are suspended inside the vacuum tanks with four types of vibration isolation systems.

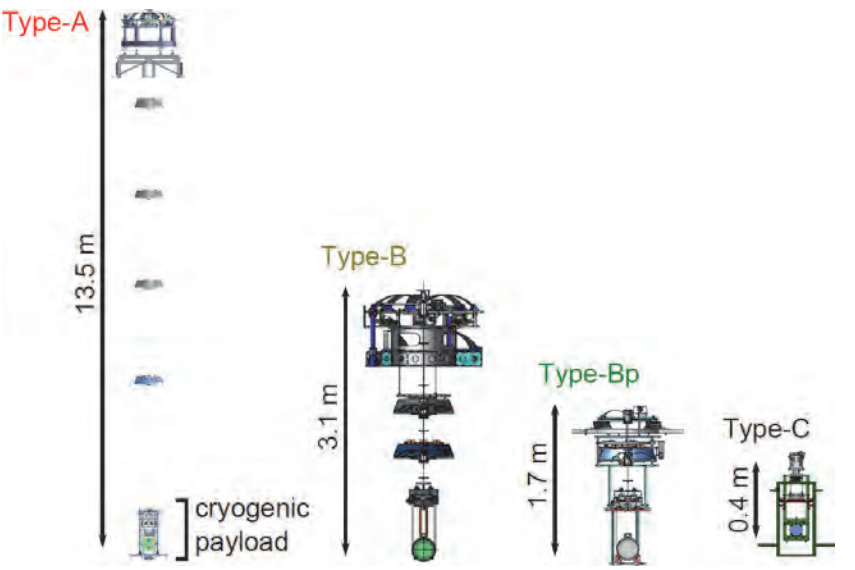



Fig. 2. Vibration isolation system for mirrors. The location of these systems are shown in Fig. 1.

Four cryogenic mirrors are suspended by Type-A system (see Fig. 2). The top part is a room temperature system and the bottom part is cryogenic. Figure 3 shows the details of the cryogenic payload.

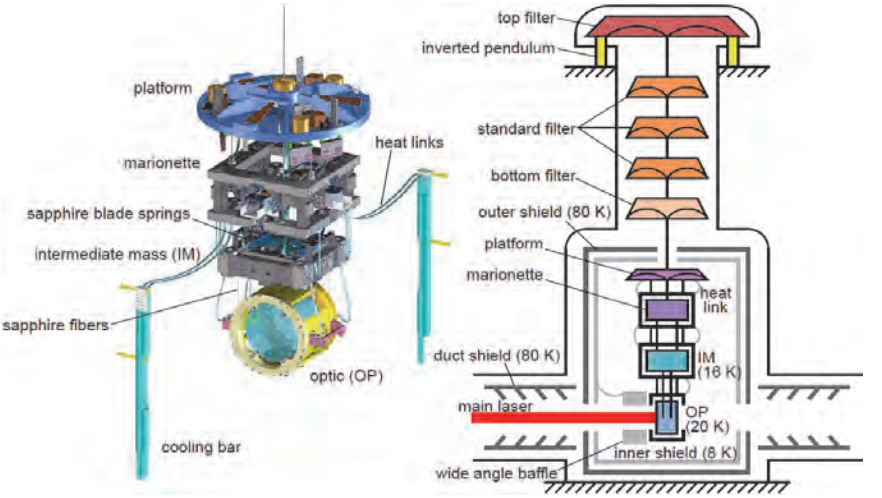



Fig. 3. Left: The drawing of the cryogenic payload under Type-A. Right: The schematic of the cryogenic suspension system of sapphire test masses. Upper part of the suspension system outside of the thermal shield are at room temperature.

## 3. Status of KAGRA

The KAGRA project was approved by the Japanese government in 2010. The excavation of the underground facility including the 3 km × 3 km arm tunnels was finished in March 2014. The installation of vacuum tubes for the 3 km × 3 km arms was completed in February 2015. KAGRA had the initial 3 km interferometer operation in the spring of 2016 with the simplest configuration. Then in the spring of 2018, interferometer operation with one cryogenic mirror was carried out. As of summer, 2018, the construction of the KAGRA interferometer was in the final stage. The construction was finished in April 2019. KAGRA is in the middle of the interferometer commissioning as of May 2019.

## 4. Plan of KAGRA

LIGO and Virgo began the third observation run (O3) in April 2019. O3 is expected to continue until the end of March 2020. Therefore, KAGRA plans to join O3 before the end of 2019. After joining O3, KAGRA would like to fully join the later observation runs (O4, O5 and later).

Figure 4 shows the sensitivity of KAGRA as a function of time[5]. It is estimated that the sensitivity of KAGRA during O3 will be limited (8–25 Mpc) because of the short time for the commissioning and noise hunting before joining O3.

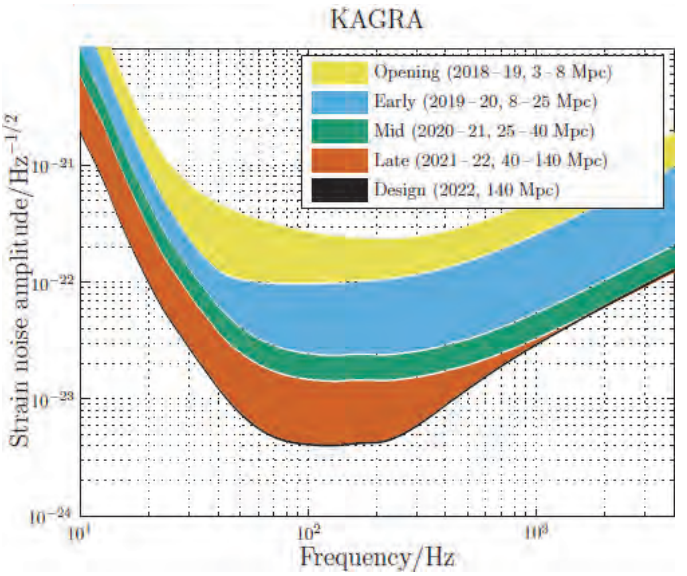



Fig. 4. The expected sensitivity of KAGRA as a function of time[5].

## 5. Scientific merits

It I well known that the scientific merit should be maximized if the operations of the interferometers are coordinated. Therefore, KAGRA plans to join the global network of gravitational wave interferometers, starting in O3.

Let us assume that we need at least 3 interferometers in operation to get scientific results, such as the determination of the source direction. Assuming that each interferometer has the duty cycle of 80%, the probability of the 3-detector coincidence in operation is 51% for 3 interferometer configurations. If there are 4 interferometers including KAGRA, the 3-detector coincidence in operation should be improved to 81%. Clearly adding the 4$^{\text{th}}$ interferometer (namely, adding KAGRA) has a significant impact on the effective observation time of the interferometers.

Initially there were only 2 LIGO interferometers. During that period the accuracy of the source localization was limited. Adding Virgo had a significant impact on the source localization. Since KAGRA is located far from both LIGO and Virgo, KAGRA's operation is expected to make the source localization and waveform reconstruction more precise.

Finally, it should be mentioned that KAGRA is located in the same mountain as Super-Kamiokande. Therefore, if a supernova explodes in our Galaxy, KAGRA and Super-Kamiokande should observe gravitational wave and neutrinos signals, respectively. These data can be analyzed without worrying the relative timing, which is a merit of KAGRA.

## 6. Summary

KAGRA is a unique GW interferometer with the underground site and the cryogenic technology. KAGRA has finished the installation in April 2019, and plans join O3 in late 2019. We would like to contribute to the global network of gravitational wave interferometers and contribute to the science of gravitational wave astronomy.

## Acknowledgments

We would like to thank the Advanced Technology Center (ATC) of NAOJ, the Mechanical Engineering Center of KEK, and the machine shop of the Institute of Solid State Physics (ISSP) of the University of Tokyo for technical support.

Also, we would like to thank LIGO Scientific Collaboration and Virgo Collaboration for various supports.

This work is supported by MEXT, JSPS Leading-edge Research Infrastructure Program, JSPS Grant-in-Aid for Specially Promoted Research 26000005, JSPS Grant-in-Aid for Scientific Research on Innovative Areas 2905: JP17H06358, JP17H06361 and JP17H06364, and JSPS Core-to-Core Program A. Advanced Research Networks.

## References

1. B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116** (2016) 061102.
2. B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), arXiv:1811.12907.
3. B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **119** (2017) 161101.
4. T. Akutsu *et al.* (KAGRA collaboration), arXiv:1901.03569.
5. B. P. Abbott *et al.* (KAGRA Collaboration, LIGO Scientific Collaboration and Virgo Collaboration), Living Rev Relativ (2018) 21:3.

# The TianQin project

Jun Luo

*TianQin Research Center for Gravitational Physics, Sun Yat-sen University (Zhuhai Campus), 2 Daxue Rd., Zhuhai 519082, P. R. China.*
*MOE Key Laboratory of Fundamental Physical Quantities Measurement & Hubei Key Laboratory of Gravitation and Quantum Physics, PGMF and School of Physics, Huazhong University of Science and Technology, 1037 Luoyu Rd., Wuhan 430074, P. R. China.*
*E-mail: junluo@mail.sysu.edu.cn*

In this talk, I give a brief introduction to the TianQin project, which aims to start space-based gravitational wave detection in the 2030s. My main focus will be on the background, the preliminary concept, the scientific objectives, the development of key technologies, the current progress and the international collaboration of the project.

*Keywords*: TianQin; Gravitational waves; Space-based gravitational wave detection.

## 1. Introduction

Discussions leading to the TianQin project started in 2013 when a team of scientists from the Sternberg Astronomical Institute, Moscow State University, joined our team at the Huazhong University of Science and Technology (HUST) to discuss ideas for future gravitational experiments in space. The focus of the discussion gradually shifted to a feasible gravitational wave mission in space from China. The name "TianQin" was proposed during a meeting in March 2014, followed by the first international workshop on the TianQin science mission in December 2014. A systematic development of the TianQin project was initiated in Sun Yat-sen University (SYSU) in early 2015. The paper summarizing the preliminary mission concept of TianQin was submitted to *Class. Quant. Grav.* on September 1, 2015 [1], 13 days before LIGO detected its first gravitational wave signal.

The name of the planned detector (which will be consisted of three satellites), TianQin, is the phonetic spelling of two Chinese characters that, when put together, mean a harp in space. By choosing this name, the detector is metaphorically seen as a musical instrument in space to be played by nature with gravitational waves.

In this talk, I shall give a brief introduction to the TianQin project, including the background (section 2), the preliminary concept (section 3), the scientific objectives (section 4), the development of key technologies (section 5), the current progress (section 6) and the international collaboration (section 7) of the project.

## 2. Background

For more than 100 years, General Relativity has passed numerous non-trivial experimental tests, including [2]:

- orbital precession of mercury,
- deflection of light by sun,

- gravitational redshift,
- Shapiro delay,
- frame dragging,
- gravitational lensing,

and so on. The first detection of gravitational waves by LIGO in 2015[3] opened a new era when General Relativity can be tested under extreme conditions where even black holes can be radically deformed.

Gravitational waves are extremely weak. In the first event detected by LIGO, GW150914, two black holes with masses $29M_\odot$ and $36M_\odot$ merged at about 1.3 billion light years away[3], producing gravitational waves with strength at the order $10^{-21}$ when reaching Earth. The effect of such gravitational wave is comparable to deforming the distance between the Sun and the Earth by the size of an atom! For this reason, it has taken people a whole century to detect gravitational waves after its prediction by Einstein in 1916.

Gravitational waves provide a new method to study the universe, providing crucial information on the origin and growth of stars, galaxies and the Universe itself, and on the nature of gravity and black holes. New discoveries have already been made with the few gravitational wave events detected since 2015, including:

- Showing that massive stellar mass black holes are more abundant than expected[4];
- Demonstrating the feasibility of multi-messenger astronomy[5,6];
- Showing that binary neutron star mergers are cosmic factories of heavy elements and are central engines of short gamma ray bursts[7].

Close to the frequency band of about 10Hz$\sim 10^4$Hz that has already been opened up with ground based detectors, the millihertz (mHz) frequency band (which typically corresponds to the frequency range $10^{-4}$Hz$\sim$ 1Hz) also has many types of important astronomical and cosmological sources:

- Galactic compact binaries can produce gravitational waves with periods in the order of a few minutes;
- Systems involving massive black holes can produce gravitational waves with periods from minutes to years;
- The birth and the initial expansion of the Universe may leave detectable gravitational waves at all frequencies.

To detect gravitational waves in the mHz frequency band, we need a laser interferometer with arm lengths at the order $10^5$km or greater and to stay away from the seismic noise that has become a limiting factor below about 10Hz for ground based detectors. So the only feasible way is to put the detector in space.

The idea of using a laser interferometer in space to detect gravitational waves can be traced back to the 1970s, and the first such mission concept LAGOS was proposed in the 1980s[9]. By far the most studied mission concept for space-based

gravitational wave detection is LISA, which envisages three spacecraft forming a regular triangle with each side measuring about 2.5 million kilometers[10]. LISA has been selected as the L3 mission in the Cosmic Vision 2015-2025 programme of ESA and is expected to launch in 2034[11].

When TianQin was first proposed in 2014, it was more intended to be an experiment rather than an observatory, with the main goal to verify the prediction of gravitational waves by General Relativity[1]. After the first detection of gravitational waves by LIGO, this primitive goal of TianQin has to be updated. Thanks to the fact that all laser interferometer-based gravitational wave detectors are wide band detectors, TianQin has the natural capability of being a space-based gravitational wave observatory. However, the initial goal of TianQin has allowed it to take some special features that are not shared by any other mission concept proposed. One of our task is to investigate the consequence of such special features when TianQin is to be treated as a gravitational wave observatory.

## 3. The preliminary concept of TianQin

A description of the mission concept of TianQin has been presented in[1]. Here I only summarize some of the key features:

- TianQin will be consisted of three satellites, forming a regular triangle constellation;
- The TianQin satellites will be on nearly identical geocentric orbits with radii at the order $10^5$ kilometers;
- The plane of the TianQin constellation is nearly perpendicular to the ecliptic (the original reason was to let the plane face the ultra-compact binary system RX J0806.3+1527, so as to maximize the response of TianQin to this particular source).

The adoption of geocentric orbits brings some advantages for TianQin: the transfer time for the TianQin satellites to enter the scientific operation orbits is at the order of dozens of days and TianQin foresees little difficulty with communicating with Earth.

However, the same geocentric orbits also bring some extra challenges.

Firstly, TianQin is facing a complicated celestial dynamical environment for being close to Earth and Moon, which directly leads to the question that if TianQin can even find such orbits that are suitable for gravitational wave detection. For TianQin, a candidate orbit need to satisfy the following constraints in order not to interfere with gravitational wave detection:

- The distance between any two satellites need to be very stable, e.g. varying less than 1% throughout the mission lifetime;
- The relative velocity between any two satellites need to be small, e.g. being less than 10m/s during scientific observation time;

- The angle between any pair of arms of the constellation need to be stable, e.g. varying no more than 0.1 degree in the short term (several months) and no more than 0.2 degree in the long term (years).

It has been shown that an orbit satisfying all the above requirement does exist[1].

Recent study has produced more candidate orbits with different orientations, for which all the above constraints are satisfied[12]. Knowledge of the orbit is important for data analysis purpose. An analytical approximation has also been obtained for the orbits of the TianQin satellites, based on which the response of TianQin to gravitational waves have been calculated[13]. The response allows us to expedite the procedure of simulating the stain data output of the detector with decent accuracy and conduct subsequent investigations on the data analysis techniques for various sources.

Secondly, all the orbits known for geocentric gravitational wave mission have their orientation nearly fixed in space. This can be seen in the many examples studied in[14]. As a result, the telescopes used for inter-satellite laser ranging will periodically point toward the Sun. Varying solar radiation on the telescopes can lead to temperature fluctuation and temperature gradient in the satellites, causing problem for the detection of gravitational waves.

With TianQin, a solution to this problem is made possible by a particular feature of the mission: the "standing" orbital plane. The plane of the TianQin constellation is facing J0806, and the location of the latter is about 4.7° below the ecliptic. As a result, the plane of the TianQin constellation is nearly perpendicular to the ecliptic and it will sweep through Sun only twice a year.

When the plane of the TianQin constellation comes too close to the Sun, there will be times that the telescopes point too close to the direction of the Sun, causing problems for the observation. TianQin adopts a "3-month on + 3-month off" detection scheme (to be further optimized) to cope with the problem. The orbit of the Earth can be partitioned into four sections, each has about 3 months:

- From early June to early September and from early December to early March, sunlight is at large angles with respect to the plane of the TianQin constellation. During such times, the telescopes will be well protected from the Sunlight.
- From early March to early June and from early September to early December, sunlight is at small angles with respect to the plane of the TianQin constellation. During such times, there can be direct sunshine on the telescopes and TianQin will suspend observation.

As such, TianQin solves the problem but pays the price of having shortened observation time.

## 4. The scientific objectives of TianQin

The main gravitational wave sources for TianQin include Galactic compact binaries, massive black hole binary coalescence, extreme mass ratio inspirals, stellar mass black hole inspirals, possible first order phase transition in the early Universe, and possibly some unforeseen sources[8]:

- With Galactic compact binaries, TianQin seeks to study the formation and evolution of compact Galactic binaries, to combine GW+EM observation to obtain comprehensive understanding of the Galactic binary systems.
- With massive black hole binary coalescence, TianQin seeks to discover seed black holes in the early universe, to depict massive black hole growth process, to study the surrounding environment of massive black hole merger, to test the Kerr-ness of the post-merger object, and to test deviation from General Relativity.
- With extreme mass ratio inspirals, TianQin seeks to study the dynamic environment around black holes in the nearby universe, to explore the fundamental nature of gravity and black holes, including: the multipolar structure and Kerr-ness of the central massive object, the beyond-general relativity emission channels, the propagation properties of gravitational waves, and the presence of massive fields around massive black holes.
- With stellar mass black hole inspirals, TianQin seeks to facilitate multi-band and multi-messenger observation, and to enhance parameter estimation accuracy, to provide better understanding of the system as well as the nature of gravity.
- With the waveform of a binary system, TianQin seeks to constrain the parameters that characterize the deviation of modified theories of gravity from general relativity.
- TianQin seeks to detect stochastic gravitational waves background originated from stellar mass black hole mergers or even binary neutron star mergers, and to measure or set limit on cosmic origin stochastic background (e.g., first order phase transitions).

We also expect enhanced science output if there is enough overlap in the operation times of TianQin and LISA[15,16].

## 5. The development of key technologies and research teams

TianQin will rely on high precision intersatellite laser interferometry to detect gravitational waves. Two test masses will be placed inside each of the three TianQin satellites. These test masses will be used as the end points for laser interferometry between the satellites. The ideal situation is that the test masses exactly follow the geodesics determined by the ambient gravitational field. In reality, however, there is environmental effect on the test masses due to electromagnetic force, particle collision and so on. So the variation of distance between the test masses and the satellite

is closely monitored and the information is used to control the satellite to follow the motion of the test masses. For this process to work, one will need high precision inertial sensors, micro-Newton thrusters and a dragfree control mechanism.

For the inertial sensor, a preliminary conceptual design for TianQin has been presented in[1]. TianQin requires that the resolution of the inertial sensor is at the order $10^{-15}\mathrm{m/s^2/Hz^{1/2}}$ in the mHz frequency band.

Our team has started working on inertial sensors since 2000. A space electrostatic accelerometer with a resolution of $4\times10^{-8}\mathrm{m/s^2/Hz^{1/2}}$ and a dynamic range of $10^{-2}$ $\mathrm{m/s^2}$ is being tested and functions well in flight from Nov. 2013 up to now[18]. A second space electrostatic accelerometer with a resolution of $3\times10^{-10}$ $\mathrm{m/s^2/Hz^{1/2}}$ and a dynamic range of $10^{-5}\mathrm{m/s^2}$ has been put to use in space from April to Sep. 2017[19].

For intersatellite laser interferometry, TianQin requires that the displacement measurement noise is at the order $10^{-12}\mathrm{m/Hz^{1/2}}$ in the mHz frequency band. In order to achieve this level of accuracy, we need technologies with laser interferometer (including ultra-stable optical bench, laser, telescopes, and clocks and so on), and ultra-stable temperature control of the satellite platform. A preliminary conceptual design for the space laser interferometry for TianQin has been given in[1].

Our team has started working on intersatellite laser interferometry since 2002. We have built a 10-m prototype of intersatellite laser ranging system in 2010 and a resolution of 3.2 nm has been achieved[20]. An ultra-precise phasemeter has been developed in 2012 and a noise level of $1.2\mu\mathrm{rad/Hz^{1/2}}$ at 1Hz has been achieved[21]. Recently, a novel scheme of intersatellite laser beam acquisition has been developed. The averaged acquisition time is 10 s for a scanning radius of 1 mrad with a success rate of 99%[22].

Apart from that for the key payloads, satellite technology is needed to provide ultra-stable and clean environment for the scientific payload and to form and maintain a highly coordinated constellation throughout the scientific observation period. A team responsible for the satellite/system technology has been assembled in SYSU to study the problems in this direction.

A team responsible for theoretical and data analysis has also been assembled in SYSU.

Apart from the teams at HUST and SYSU, there are many other groups in China that have technology background related to space-based gravitational wave detection. There is an effort to engage all these teams in the work of TianQin.

## 6. Recent progress

In order to support the development of key technologies, we are constructing a few dedicated research facilities and have started flight experiments on TianQin key technologies.

### 6.1. *Dedicated research facilities*

As of present we are constructing three dedicated TianQin research facilities:

- The Payload Research Base, which is responsible for all key technology research and development for the project.
- A laser ranging station, which will be used to develop laser ranging capability to TianQin satellites.
- The Ground Simulation Facility, which is responsible for integrated test and research on the TianQin technologies and prototypes.

The Payload Research Base is consisted of the TianQin Research Building and a cave lab. The TianQin Research Building has about 30 thousand square meters in total area and the cave lab has about 10 thousand square meters in total area. The construction of the Payload Research Base and the laser ranging station has started on the SYSU Zhuhai campus in the end of 2017. The TianQin Research Building will be delivered by the end of 2019. The tunnel of the cave lab will be finished by the end of 2020.

The laser ranging station will be equipped with a 1.2 meter telescope, plus an education and outreach facility. The laser ranging station will be ready in the early part of 2019.

The Ground Simulation Facility is a big effort in the TianQin project. The idea is to have a facility that can simulate as close as possible the various aspects of a space-based gravitational wave mission. The facility will have the capability to simulate space environment, inertial reference, inter-satellite laser interferometry, the formation of TianQin constellation and the process of space-based gravitational wave observation. The facility will also aid in signal abstraction and data analysis. The facility will be located on the SYSU Shenzhen Campus.

### 6.2. *Flight experiments*

There are two space experiment projets going on at the moment:

- Laser ranging to the Chang'E 4 (CE4) relay satellite;
- The TQ-1 experimental satellite.

Laser ranging technology will be used to help tracking the TianQin satellites. In order to bring the needed technology to mature, the TianQin project has planned a lunar laser ranging program, which involves (1) creating new generation corner cube retro-reflectors to be deployed on the surface of Moon or to be carried by high Earth orbit satellites, and (2) upgrading/constructing laser ranging stations on the ground. The CE4 relay satellite (QueQiao) has been launched on May 21, 2018 and has successfully entered a Lissajous orbit around the Earth-Moon L2 point. Our team have created a single large aperture hollow corner cube retro-reflector (CCR) and have installed it on the QueQiao satellite. Our next step is to do laser ranging

experiment to the CCR onboard the QueQiao satellite. As part of the project, the Yunnan Observatory (located in Kunmin, China) has successfully carried out the first lunar laser rang experiment in China early 2018. A new laser ranging station is also being constructed on the SYSU Zhuhai campus and is expected to become available in the spring of 2019.

We are also preparing for the first experimental satellite, TQ-1, on TianQin key technologies. The satellite will be equipped with an inertial sensor reaching the resolution level $10^{-12}\mathrm{m/s^2/Hz^{1/2}}$ at 0.1Hz, a laser interferometer reaching the resolution level $0.1\mathrm{nm/Hz^{1/2}}$ at 0.1Hz. The mission has been approved by the China National Space Administration and is scheduled for launch in late 2019.

## 7. International collaboration

International collaboration is an important aspect of the TianQin project. The effort on the TianQin project is expected to span some 15 years. Due to the long duration of the effort, a core team is necessary to make sure that the project evolves as expected. Currently the core team of TianQin is consisted of two teams, located in two universities in China, Huazhong University of Science and Technology (HUST) and Sun Yat-sen University (SYSU). The team at HUST was formed in 1983 and has grown to more than 300 researchers and students by now. SYSU has established in 2016 a new center dedicated to the TianQin project, the TianQin Research Center for Gravitational Physics. The center has grown to more than 100 researchers and students by now.

The TianQin collaboration has been formally established in the end of 2018 during the fifth international workshop on the TianQin science mission.

## Acknowledgement

This work was supported in part by the National Natural Science Foundation of China (Grants No. 91636111, 11690022, 11703098, 11475064).

## References

1. J. Luo *et al.* [TianQin Collaboration], *TianQin: a space-borne gravitational wave detector*, Class. Quant. Grav. **33**, no. 3, 035010 (2016) doi:10.1088/0264-9381/33/3/035010 [arXiv:1512.02076 [astro-ph.IM]].
2. C. M. Will, *The Confrontation between General Relativity and Experiment*, Living Rev. Rel. **17**, 4 (2014) doi:10.12942/lrr-2014-4 [arXiv:1403.7377 [gr-qc]].
3. B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, no. 6, 061102 (2016) doi:10.1103/PhysRevLett.116.061102 [arXiv:1602.03837 [gr-qc]].
4. B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], *Astrophysical Implications of the Binary Black-Hole Merger GW150914*, Astrophys. J. **818**,

no. 2, L22 (2016) doi:10.3847/2041-8205/818/2/L22 [arXiv:1602.03846 [astro-ph.HE]].

5. B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*, Phys. Rev. Lett. **119**, no. 16, 161101 (2017) doi:10.1103/PhysRevLett.119.161101 [arXiv:1710.05832 [gr-qc]].
6. B. P. Abbott *et al.* [LIGO Scientific and Virgo and Fermi GBM and INTEGRAL and IceCube and IPN and Insight-Hxmt and ANTARES and Swift and Dark Energy Camera GW-EM and DES and DLT40 and GRAWITA and Fermi-LAT and ATCA and ASKAP and OzGrav and DWF (Deeper Wider Faster Program) and AST3 and CAASTRO and VINROUGE and MASTER and J-GEM and GROWTH and JAGWAR and CaltechNRAO and TTU-NRAO and NuSTAR and Pan-STARRS and KU and Nordic Optical Telescope and ePESSTO and GROND and Texas Tech University and TOROS and BOOTES and MWA and CALET and IKI-GW Follow-up and H.E.S.S. and LOFAR and LWA and HAWC and Pierre Auger and ALMA and Pi of Sky and DFN and ATLAS Telescopes and High Time Resolution Universe Survey and RIMAS and RATIR and SKA South Africa/MeerKAT Collaborations and AstroSat Cadmium Zinc Telluride Imager Team and AGILE Team and 1M2H Team and Las Cumbres Observatory Group and MAXI Team and TZAC Consortium and SALT Group and Euro VLBI Team and Chandra Team at McGill University], *Multi-messenger Observations of a Binary Neutron Star Merger*, Astrophys. J. **848**, no. 2, L12 (2017) doi:10.3847/2041-8213/aa91c9 [arXiv:1710.05833 [astro-ph.HE]].
7. B. P. Abbott *et al.* [LIGO Scientific and Virgo and Fermi-GBM and INTEGRAL Collaborations], *Gravitational Waves and Gamma-rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*, Astrophys. J. **848**, no. 2, L13 (2017) doi:10.3847/2041-8213/aa920c [arXiv:1710.05834 [astro-ph.HE]].
8. Y. M. Hu, J. Mei and J. Luo, *Science prospects for space-borne gravitational-wave missions*, Natl. Sci. Rev. **4**, no. 5, 683 (2017). Hu:2017yoc
9. R. T. Stebbins, P. L. Bender, J. E. Faller, J. L. Hall, D. Hils and M. A. Vincent, *A Laser interferometer for gravitational wave astronomy in space*,
10. H. Audley *et al.* [LISA Collaboration], *Laser Interferometer Space Antenna*, arXiv:1702.00786 [astro-ph.IM].
11. `http://www.esa.int/Our_Activities/Space_Science/Gravitational_wave_mission_selected_planet-hunting_mission_moves_forward`
12. B. B. Ye *et al.*, *Optimizing orbits for TianQin*, in preparation.
13. X. C. Hu *et al.*, *Fundamentals of the orbit and response for TianQin*, Class. Quant. Grav. **35**, no. 9, 095008 (2018) doi:10.1088/1361-6382/aab52f [arXiv:1803.03368 [gr-qc]].
14. `https://pcos.gsfc.nasa.gov/physpag/GW_Study_Rev3_Aug2012-Final.pdf`

15. H. T. Wang *et al.*, *Science with TianQin: Preliminary Results on Massive Black Hole Binaries,* arXiv:1902.04423 [astro-ph.HE].
16. C. Shi *et al.*, *Science with TianQin: Preliminary Results on Testing the No-hair Theorem with Ringdown Signals,* arXiv:1902.08922 [gr-qc].
17. T. Robson, N. Cornish and C. Liu, *The construction and use of LISA sensitivity curves*, arXiv:1803.01944 [astro-ph.HE].
18. S. B. Qu *et al.*, *Self-calibration method of the bias of a space electrostatic accelerometer*, Review of Scientific Instruments 87, 11 (2016).
19. Y. Bai *et al.*, *Research and Development of Electrostatic Accelerometers for Space Science Missions at HUST*, Sensors. 2017; 17(9):1943.
20. H.-C. Yeh *et al*, *Intersatellite laser ranging with homodyne optical phase locking for Space Advanced Gravity Measurements mission*, Review of Scientific Instruments 82, 4 (2011).
21. Y.-R. Liang *et al*, *Fundamental limits on the digital phase measurement method based on cross-correlation analysis*, Review of Scientific Instruments 83, 095110 (2012).
22. J.-Y. Zhang *et al*, *Inter-satellite laser link acquisition with dual-way scanning for Space Advanced Gravity Measurements mission featured*, Review of Scientific Instruments 89, 064501 (2018).

# Analytic approximations in GR and gravitational waves

Luc Blanchet**

*$\mathcal{GR}\varepsilon\mathbb{C}\mathcal{O}$, Institut d'Astrophysique de Paris,*
*UMR 7095, CNRS, Sorbonne Universités & UPMC Univ Paris 6,*
*$98^{bis}$ boulevard Arago, 75014 Paris, France*
** *E-mail: luc.blanchet@iap.fr*

Analytic approximation methods in general relativity play a very important role when analyzing the gravitational wave signals recently discovered by the LIGO & Virgo detectors. In this contribution, we present the state-of-the-art and some recent developments in the famous post-Newtonian (PN) or slow-motion approximation, which has successfully computed the equations of motion and the early inspiral phase of compact binary systems. We discuss also some interesting interfaces between the PN and the gravitational self-force (GSF) approach based on black-hole perturbation theory, and between PN and the post-Minkowskian (PM) approximation, namely a non-linearity expansion valid for weak field and possibly fast-moving sources.

*Keywords*: gravitational waves, compact binary systems, post-Newtonian approximation, post-Minkowskian approximation, perturbation theory

## 1. Methods to generate gravitational wave templates

The LIGO & Virgo detectors have opened up a fantastic new avenue in Astronomy with the discovery of gravitational waves (GWs) generated by the orbital motion and merger of binary black hole and neutron star systems.[1,2] This also highlights the crucial role played by analytic approximation methods in general relativity (GR), since they permit an accurate description of the two-body problem in GR, which is of direct use in the data analysis of the detectors.[44]

The most important method in this respect is the post-Newtonian (PN) approximation, which is an expansion when the slowness parameter $\epsilon_{\rm PN} = v/c$ of the compact binary system tends to zero, where $v$ is the relative orbital velocity and $c$ the speed of light. For gravitationally bound systems such as compact binaries on quasi-circular orbits, the PN approximation comes along with the post-Minkowskian (PM) one, namely a non-linearity expansion around the Minkowski background, with small expansion parameter $\gamma_{\rm PM} = Gm/(rc^2)$, where $r$ is the size of the orbit and $m$ the total mass of the source. Indeed, in the bounded case we have $\gamma_{\rm PM} \sim \epsilon^2_{\rm PN}$. However, the most important physical application of the PM approximation is for unbound orbits, when $\gamma_{\rm PN}$ and $\epsilon_{\rm PM}$ are unrelated, *i.e.*, the problem of scattering of ultra-relativistic particles ($\epsilon_{\rm PN} \lesssim 1$) and small deviation angle. The PM approximation is sometimes called the weak-field fast-moving approximation.

Black hole perturbation theory constitutes another large body of analytic approximations in GR. In the context of compact binary systems, this approximation is important, first, for analyzing the post-merger waveform of two black holes (BHs) during the so-called ringdown phase, when the newly formed BH emits quasi-normal mode radiation, and, secondly, for describing the dynamics and

GWs of asymmetric compact binaries, *i.e.*, endowed with an extreme mass ratio, $\nu = m_1 m_2/(m_1+m_2)^2 \ll 1$. In the latter case the perturbation method takes the more suggestive name of gravitational self-force (GSF), since it is concerned with the modifications of the background geometry of the larger BH and of the geodesic motion of the particle, due to the self field generated by the particle itself.

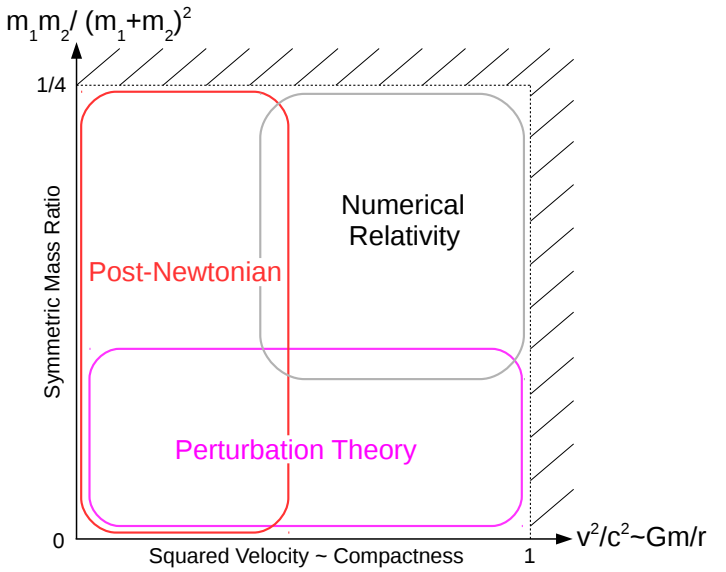



Fig. 1. Analytic approximation and numerical techniques to build GW templates for the compact binary inspiral and merger, depending on the symmetric mass ratio $\nu = m_1 m_2/m^2$ ($m = m_1+m_2$) and the slowness-weak-field parameter $\epsilon_{\rm PN} = v/c \sim \sqrt{Gm/rc^2}$. PN theory and perturbative GSF analysis can be compared in the slow motion weak field regime, $\epsilon_{\rm PN} \ll 1$ thus $r \gg Gm/c^2$, of an extreme mass ratio compact binary, $\nu \ll 1$.

The domains of validity of these approximation methods, together with interesting mutual overlaps, are illustrated in Figs. 1 and 2. Also shown in Fig. 1 is some comparison with numerical relativity (NR), which succeeded at solving the long standing binary black hole challenge.[4,45,93] At first sight it could seem that NR would be able to tackle and solve the complete problem of the inspiral, merger and ringdown for two compact objects. However, in order to monitor the early inspiral of two neutron stars, thousands of orbital cycles have to be computed with high precision. Then the computing times of NR become prohibitively long, and the precision of the NR simulation will never be competitive with that provided by the PN approximation. A fact that is of uttermost importance for building GW templates (and was not *a priori* obvious several years ago[42]), is that the overlap between PN and NR exists and is quite significant. On the other hand, when the

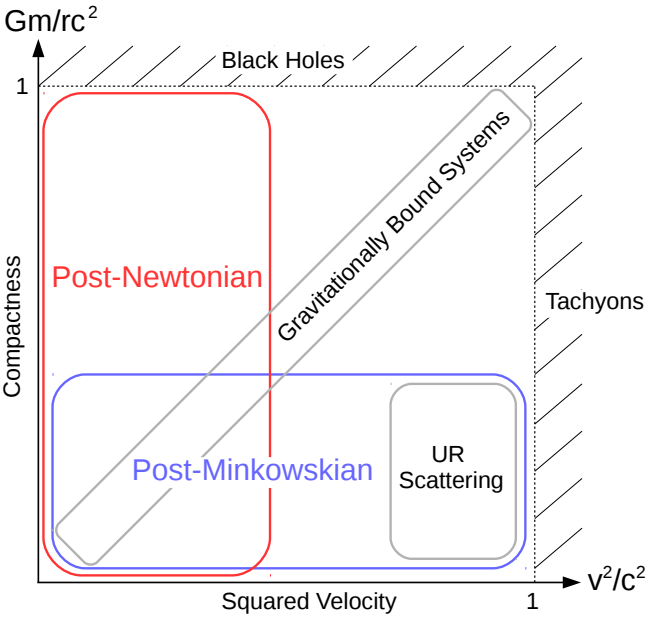



Fig. 2. Comparison between the PN and PM approximations. Gravitationally bound systems such as compact binaries on quasi-circular orbits, stand roughly on the diagonal. For such systems the PM approximation actually reduces to the PN approximation. The PM expansion is a weak-field expansion, defined with no restriction on the slowness parameter $\epsilon_{\rm PN}$, and is mostly relevant in the case of unbound orbits, such as the ultra-relativistic (UR) scattering of two BHs.

mass ratio between the two bodies is extreme, the full NR approach is unfeasible, due to the different length scales corresponding to the very different physical sizes of the compact bodies. We shall discuss in this article the important intersections between PN and GSF as shown in Fig. 1, and between PN and PM, see Fig. 2.

The GW templates are defined as the theoretical prediction from GR, and weighted in the Fourier domain by the detector's spectral density of noise. The templates are cross-correlated with the detector's output, and the correlation builds up when a good match occurs between a particular template and the real signal.[44] This technique is highly sensitive to the phase evolution of the signal, which, in PN templates of compact binary coalescence, is computed from the energy balance between the decay of the binary's energy and (minus) the GW flux.

In principle, as there is a significant overlap between the PN and NR regimes, the templates are obtained by matching together the best PN waveform for the inspiral phase (currently known to order 3.5PN) to a highly accurate numerical waveform for the merger and ringdown phases.[67] For low mass compact binaries, such as double neutron star systems, the detectors are mostly sensitive to the inspiral phase prior to the final coalescence, and the currently known analytical PN templates are accurate

enough for detection, at least when the compact bodies have moderate spins. Thus, the data analysis of neutron star binaries events such as GW170817, is essentially based on the 3.5PN templates.

For larger masses, like BH binary events such as GW150914, the merger occurs at lower frequency, right in the middle of the detector's frequency band. Since only a few orbital cycles before the merger are seen, the match between NR and the PN is not very good and the GW templates are essentially based on the NR results. Nevertheless it is worth pointing out that the mere zeroth-order Newtonian waveform, *i.e.*, based on the Einstein quadrupole formula, gives a reasonably satisfying physical interpretation of the signal even for GW150914!

In practice, for the data analysis of large mass BH binary events, it is important to dispose of analytic rather than numerical templates, since the cross correlations must be performed with many templates associated by many trial parameters (masses and spins). In this case the templates are based on effective analytical methods that interpolate somehow between the initial PN and final NR phases. Two such techniques play a key role in the data analysis of the BH events. One is called the hybrid inspiral-merger-ringdown (IMR or IMR-Phenom) waveform and is constructed by matching together the PN and NR waveforms in an overlapping time interval described phenomenologically.[3] The other technique consists of recasting the real two-body dynamics, as given by PN theory, into a simpler one-body dynamics described in a non perturbative analytic way. The so-called effective-one-body (EOB) waveform obtained in that way extends the domain of validity of the PN approximation (because it is non perturbative) and can therefore be compared and matched to the NR waveform.[43] The IMR and EOB waveforms are extensively used in the LIGO & Virgo data analysis of the recent binary BH events.

## 2. State-of-the-art on equations of motion

The equations of motion (EOM) of a self-gravitating $N$-body system are written in PN like form, up to 4PN order, using one's favorite coordinate system in GR, as

$$\frac{\mathrm{d}\boldsymbol{v}_a}{\mathrm{d}t} = \boldsymbol{A}_a^{\mathrm{N}} + \frac{1}{c^2}\boldsymbol{A}_a^{\mathrm{1PN}} + \frac{1}{c^4}\boldsymbol{A}_a^{\mathrm{2PN}} + \frac{1}{c^5}\boldsymbol{A}_a^{\mathrm{2.5PN}} + \frac{1}{c^6}\boldsymbol{A}_a^{\mathrm{3PN}} + \frac{1}{c^7}\boldsymbol{A}_a^{\mathrm{3.5PN}} + \frac{1}{c^8}\boldsymbol{A}_a^{\mathrm{4PN}} + \mathcal{O}\left(\frac{1}{c^9}\right). \tag{1}$$

The first term is of course, the usual Newtonian acceleration of $N$ "planets",

$$\boldsymbol{A}_a^{\mathrm{N}} = -\sum_{b\neq a}\frac{Gm_b}{r_{ab}^2}\boldsymbol{n}_{ab}\,. \tag{2}$$

The historical works in GR started in the early days of this theory, and solved the problem of the EOM at the 1PN level beyond the Newtonian term.[58,79] This famous Lorentz-Droste-Einstein-Infeld-Hoffmann 1PN correction is fully given as

$$\boldsymbol{A}_a^{\mathrm{1PN}} = -\sum_{b\neq a}\frac{Gm_b}{r_{ab}^2}\boldsymbol{n}_{ab}\Big[v_a^2 + 2v_b^2 - 4(v_a v_b) - \frac{3}{2}(n_{ab}v_b)^2$$

$$- 4\sum_{c\neq a}\frac{Gm_c}{r_{ac}} - \sum_{c\neq b}\frac{Gm_c}{r_{bc}}\left(1-\frac{r_{ab}}{2r_{bc}}(n_{ab}n_{bc})\right)\Big]$$
$$+\sum_{b\neq a}\frac{Gm_b}{r_{ab}^2}\boldsymbol{v}_{ab}\big[4(n_{ab}v_a)-3(n_{ab}v_b)\big] - \frac{7}{2}\sum_{b\neq a}\sum_{c\neq b}\frac{G^2m_bm_c}{r_{ab}r_{bc}^2}\boldsymbol{n}_{bc}\,, \qquad (3)$$

where we denote $\boldsymbol{v}_a = \mathrm{d}\boldsymbol{y}_a/\mathrm{d}t$, $\boldsymbol{v}_{ab} = \boldsymbol{v}_a - \boldsymbol{v}_b$, $r_{ab} = |\boldsymbol{y}_a - \boldsymbol{y}_b|$, $\boldsymbol{n}_{ab} = (\boldsymbol{y}_a - \boldsymbol{y}_b)/r_{ab}$ and the parenthesis indicate the usual Euclidean scalar product, *e.g.* $(n_{ab}v_b) = \boldsymbol{n}_{ab}\cdot\boldsymbol{v}_b$.

Up to the 2PN level the system is *conservative*, *i.e.*, admits the ten invariants associated with the symmetries of the Poincaré group. The first *dissipative* effect appears at the 2.5PN order and features the radiation reaction damping of the system by GW emission. The 2PN and 2.5PN approximations were motivated by the Hulse-Taylor binary pulsar and worked out at the time of its discovery.[33,47,50,72,87,98] Later the motivation for the 3PN EOM came from the development of the LIGO & Virgo detectors and the need of accurate GW templates for inspiralling compact binaries. The 3PN EOM took some time to be fully understood and completed,[23,30,37,51,54,60,69,71,73] together with the relatively easier dissipative 3.5PN term.[70,76,86,88]

Three techniques have been undertaken to obtain the 4PN EOM. One is based on the Arnowitt-Deser-Misner (ADM) Hamiltonian formalism of GR in ADM coordinates,[52,53,74,75] and has led to complete results but for the appearance of one "*ambiguity*" parameter. The second technique is based on the Fokker action of GR in harmonic coordinates,[7–10,83] and has obtained complete results, *i.e.*, free of any ambiguity parameter. The third one is the effective field theory (EFT),[59,61,62,65] which yielded partial results up to now (the terms $\propto G^4$ still being in progress), but is expected to also be free of any ambiguity parameter.[90]

In this section we describe the approach based on the Fokker action in harmonic coordinates. We start with the gravitation-plus-matter action of GR, in which the gravitational piece includes the usual harmonic gauge-fixing term, and the matter term is that of $N$ particles without spins, and with negligible internal structure:

$$S = \frac{c^3}{16\pi G}\int \mathrm{d}^4x\,\sqrt{-g}\Big[R - \frac{1}{2}g_{\mu\nu}\Gamma^\mu\Gamma^\nu\Big] - \sum_{a=1}^{N} m_a c^2\int \mathrm{d}t\,\sqrt{-(g_{\mu\nu})_a\,v_a^\mu v_a^\nu/c^2}\,. \quad (4)$$

Here $\Gamma^\mu = g^{\rho\sigma}\Gamma^\mu_{\rho\sigma}$ and we use, for practical calculations, the Landau-Lifshitz form of the action (*i.e.*, modulo a total divergence).

The Fokker action is obtained when we insert back into (4) an explicit PN solution of the corresponding gauge-fixed Einstein field equations, and given by an explicit functional of the particle's trajectories, *i.e.*,

$$\bar{g}_{\mu\nu}(\mathbf{x},t) = g_{\mu\nu}[\mathbf{x};\boldsymbol{y}_a(t),\boldsymbol{v}_a(t),\cdots]\,. \quad (5)$$

The ellipsis indicate that the metric also depends on accelerations $\boldsymbol{a}_a = \mathrm{d}\boldsymbol{v}_a/\mathrm{d}t$, derivatives of accelerations $\boldsymbol{b}_a = \mathrm{d}\boldsymbol{a}_a/\mathrm{d}t$, *etc.*, since we do not perform any replacements of accelerations when iterating the Einstein field equations, the EOM being

considered off-shell at this stage. Substituting $\bar{g}_{\mu\nu}$ into Eq. (4) defines the Fokker action $S_{\rm F}[\boldsymbol{y}_a, \boldsymbol{v}_a, \cdots]$, and the EOM of the self-gravitating system of particles are obtained as the (generalized Euler-Lagrange) equations

$$\frac{\delta S_{\rm F}}{\delta \boldsymbol{y}_a} = 0\,. \tag{6}$$

Once they have been constructed, the EOM can be order reduced by replacing all the higher-order accelerations by their expressions coming from the lowest-order PN equations. The Fokker action describes only the conservative dynamics, and the dissipative effects have to be added separately in the EOM. Note that the Fokker action is equivalent to the EFT action in the "tree-level" approximation, in which we neglect quantum loops.

An interesting feature of the local (near zone) 4PN dynamics is that there is an imprint of GW tails propagating at infinity. The tails are secondary non-linear waves caused by backscattering of linear waves onto the space-time curvature generated by the total mass $M$ of the source. Part of the effect can be seen as a tail-induced modification of the leading 2.5PN radiation reaction force at the relative 1.5PN order.[16,18,21] However, associated with this dissipative piece, there exists also a conservative effect which thus enters into the 4PN conservative dynamics. Its contribution to the Fokker action reads as[7,52,62,65]

$$S_{\rm F}^{\rm tail} = \frac{G^2 M}{5c^8} \int_{-\infty}^{+\infty} {\rm d}t\, I_{ij}^{(3)}(t) \int_0^{+\infty} {\rm d}\tau \ln\left(\frac{\tau}{\tau_0}\right) \left[I_{ij}^{(4)}(t-\tau) - I_{ij}^{(4)}(t+\tau)\right]. \tag{7}$$

Here $I_{ij} = \sum_a m_a y_a^{\langle i} y_a^{j\rangle}$ is the Newtonian quadrupole moment of the system (the angular brackets refer to the symmetric-trace-free projection), the superscript $(n)$ denotes multiple time derivatives, and $\tau_0$ is an arbitrary constant. To leading order the total mass $M$ reduces to $\sum_a m_a$, but at higher order it should involve the contribution of the gravitational binding energy of the particles. An elegant rewriting of Eq. (7) is with the Hadamard "*Partie finie*" (Pf) integral,

$$S_{\rm F}^{\rm tail} = \frac{G^2 M}{5c^8} \operatorname*{Pf}_{\tau_0} \iint \frac{{\rm d}t\, {\rm d}t'}{|t-t'|} I_{ij}^{(3)}(t)\, I_{ij}^{(3)}(t')\,. \tag{8}$$

Due to this conservative tail contribution, the 4PN dynamics is non-local in time, and this entails subtleties in the derivation of the invariants of motion, which have been recently fully elucidated.[9,41,52,53]

The calculation crucially relies on the systematic use of dimensional regularization (DR), to cure both ultra-violet (UV) divergences due to the model of point particles adopted to describe the compact objects, and infra-red (IR) divergences that start appearing precisely at the 4PN order and are associated with GW tails. We are here borrowing DR from EFT and quantum field theory, and we use it in the classical $N$-body problem as a mean to preserve the diffeomorphism invariance of GR. For this reason we conjecture that DR is the only known regularization technique able to successfully solve the problem of EOM in high PN approximations.

In an initial calculation (valid for two particles, $N = 2$), we used DR for UV divergences but a variant of the Hadamard regularization (HR) for IR divergences.[7,9] Based on some trial calculations, using various types of regularizations, we conjectured that the results of different IR regularizations will physically differ by at most two parameters called ambiguities and denoted $\delta_1$ and $\delta_2$.[9] This finding was in agreement with an earlier suggestion.[53] Modulo unphysical shifts of the trajectories, the two offending ambiguity terms in the Fokker Lagrangian ($S_{\rm F} = \int {\rm d}t L_{\rm F}$) turn out to appear at the difficult $G^4$ level and be of the form

$$\delta L_{\rm F} = \frac{G^4(m_1+m_2)\,m_1^2 m_2^2}{c^8 r_{12}^4}\Big(\delta_1 (n_{12}v_{12})^2 + \delta_2 v_{12}^2\Big)\,. \tag{9}$$

To determine what the values of these ambiguities are we embarked on the DR treatment of the IR divergences. Consider a typical term in the Fokker Lagrangian with non-compact support and generic function $F$, which diverges at infinity. With HR such term is treated as

$$L_{\rm F}^{\rm HR} = \underset{B=0}{\rm FP}\int_{r>\mathcal{R}} {\rm d}^3\mathbf{x}\left(\frac{r}{r_0}\right)^B F(\mathbf{x})\,. \tag{10}$$

Since we focus on IR divergences we consider only the far zone contribution $r > \mathcal{R}$, where $\mathcal{R}$ denotes an arbitrary large radius, typically the inner radius of the wave zone. The Finite Part (FP) operation is closely related to the Hadamard partie finie Pf, and consists of applying analytical continuation in $B \in \mathbb{C}$, expanding the integral when $B$ tends to zero, and keeping only the coefficient of the zero-th power of $B$ in that expansion (discarding any strictly positive or negative power of $B$). On the other hand, with DR the same term is treated as

$$L_{\rm F}^{\rm DR} = \int_{r>\mathcal{R}} \frac{{\rm d}^d\mathbf{x}}{\ell_0^{d-3}}\,F^{(d)}(\mathbf{x})\,, \tag{11}$$

where $F^{(d)}$ is the $d$-dimensional analogue of the generic function $F$ in (10), and where $\ell_0$ is the characteristic length scale associated with DR. We find that the difference between the two prescriptions is given by[8]

$$L_{\rm F}^{\rm DR} - L_{\rm F}^{\rm HR} = \sum_q \left[\frac{1}{(q-1)\varepsilon} - \ln\left(\frac{r_0}{\ell_0}\right)\right]\int {\rm d}\Omega_{2+\varepsilon}\,\varphi_{3,q}^{(\varepsilon)}(\mathbf{n}) + \mathcal{O}(\varepsilon)\,. \tag{12}$$

The functions $\varphi_{p,q}^{(\varepsilon)}$ represent the coefficients of $r^{-p-q\varepsilon}$ (with $p, q \in \mathbb{Z}$) in the expansion of $F^{(d)}$ when $r \to +\infty$ along the direction $\mathbf{n}$. We pose $\varepsilon = d - 3$ and neglect the terms dying with $\varepsilon \to 0$. The angular integration in (12) is over the sphere in $d - 1 = 2 + \varepsilon$ dimensions. Notice that the result (12) depends only on the singular coefficients $\varphi_{3,q}^{(\varepsilon)}$ of the expansion of $F^{(d)}$ at infinity, and that the arbitrary scale $\mathcal{R}$ has disappeared from it.

The formula (12) contains an IR pole $\propto 1/\varepsilon$. In the language of the EFT, this pole comes from the "potential mode" contribution. However, we have also to take into account the 4PN tail effect given by Eq. (8) but in $d$ dimensions. In the EFT language this will correspond to the "radiation" contribution and should

also contain a pole $\propto 1/\varepsilon$, but this time of UV type. Our explicit calculation has shown that the two IR and UV poles exactly cancel out (modulo unphysical shifts of the trajectories),[8,83] in complete agreement with general arguments within the EFT.[90] The 4PN tail term in $d$ dimensions takes the same form as (8) but with the arbitrary scale $\tau_0$ determined by the DR scale $\ell_0$ as

$$\tau_0^{\rm DR} = \frac{2\ell_0}{c\sqrt{4\pi}} \exp\Big[\frac{1}{2\varepsilon} - \frac{1}{2}\gamma_{\rm E} - \frac{41}{60}\Big]\,, \tag{13}$$

with $\gamma_{\rm E}$ denoting the Euler constant, and the UV pole $\propto 1/\varepsilon$ cancelling the IR one in (12). The result (13) has been obtained thanks to a "matching" equation relating the near zone which is the domain of validity of the PN approximation, to the far zone where GW tails propagate. Finally we find that the rational fraction $-\frac{41}{60}$ in (13), is just the one necessary and sufficient to determine the values of the two ambiguity parameters in (9) as

$$\delta_1 = -\frac{2179}{315}\,, \qquad \delta_2 = \frac{192}{35}\,, \tag{14}$$

therefore resolving the problem of ambiguities. The values (14) are consistent with numerical and analytical GSF calculations of the energy and periastron advance for circular orbits in the small mass ratio limit.[8,52,53] Remarkably, the result (13) agrees with that of Galley *et al.*,[65] obtained by means of a diagrammatic evaluation of the tail term in $d$ dimensions with EFT methods. On the other hand, the lack of a consistent matching between the near and far zones in the ADM Hamiltonian formalism,[52,53,74,75] and therefore a complete control of the tail term (8) including the final determination of Eq. (13), forces this formalism to be still plagued by one ambiguity parameter, denoted $C$ in[52].

## 3. State-of-the-art on GW generation

The two basic ingredients in the theoretical PN analysis correspond to the two sides of the energy balance equation obeyed by the binary's orbital frequency and phase. Since the orbit will have circularized by radiation reaction at the time when the signal enters the detectors' bandwidth there is no need to invoke the balance equation for the orbital angular momentum. Thus we just impose

$$\frac{{\rm d}E}{{\rm d}t} = -\mathcal{F}\,. \tag{15}$$

The energy $E$ is nothing but the Noetherian conserved energy $E$ associated with the Fokker Lagrangian computed in the previous section. On the other hand, the GW energy flux $\mathcal{F}$ on the right-hand side is obtained from a GW generation formalism. From Eq. (15) one deduces the time evolution of the binary's orbital frequency $\omega$ and orbital phase $\phi$ by solving

$$\phi = \int \omega\,{\rm d}t = -\int \frac{\omega}{\mathcal{F}}\frac{{\rm d}E}{{\rm d}\omega}\,{\rm d}\omega\,. \tag{16}$$

At the 4.5PN order, for circular orbits, the conserved energy function is given by

$$E = -\frac{m\,\nu c^2 x}{2}\left\{1 + \left(-\frac{3}{4} - \frac{\nu}{12}\right)x + \left(-\frac{27}{8} + \frac{19}{8}\nu - \frac{\nu^2}{24}\right)x^2 \right.$$
$$+ \left(-\frac{675}{64} + \left[\frac{34445}{576} - \frac{205}{96}\pi^2\right]\nu - \frac{155}{96}\nu^2 - \frac{35}{5184}\nu^3\right)x^3$$
$$+ \left(-\frac{3969}{128} + \left[-\frac{123671}{5760} + \frac{9037}{1536}\pi^2 + \frac{896}{15}\gamma_{\rm E} + \frac{448}{15}\ln(16x)\right]\nu\right.$$
$$\left.\left. + \left[-\frac{498449}{3456} + \frac{3157}{576}\pi^2\right]\nu^2 + \frac{301}{1728}\nu^3 + \frac{77}{31104}\nu^4\right)x^4\right\}, \tag{17}$$

where $m = m_1 + m_2$ is the total mass, $\nu = m_1 m_2/m^2$ is the symmetric mass ratio, $\gamma_{\rm E}$ is Euler's constant, and we employ for convenience the PN ordering parameter $x = (\frac{Gm\omega}{c^3})^{2/3}$ defined from the orbital frequency $\omega$ of the circular orbit, and which constitutes an invariant in a large class of coordinate systems. Notice that Eq. (17) is valid up to the 4.5PN order *included*, as there is no term at the 4.5PN order in the conserved energy for circular orbits.

The most complete formula for the GW flux is valid at the 3.5PN order beyond the Einstein quadrupole formula. However, the 4.5PN coefficient is also known,[84] while the 4PN coefficient is in progress. This formula has been obtained by application of a GW generation formalism based on a Multipolar-Post-Minkowskian (MPM) expansion for the external field of an isolated source,[15,20–22] and followed by a matching to the inner (near zone) PN field of that source.[17,19,32,91] The first important step in this computation is the obtention of the multipole moments of the source, $I_{i_1\cdots i_\ell}$ (mass type) and $J_{i_1\cdots i_\ell}$ (current type). The most difficult of these moments (because it necessitates the highest PN precision) is the mass quadrupole moment $I_{i_1 i_2}$, given at 3.5PN order for quasi-circular orbits as

$$I_{i_1 i_2} = m\nu\left(A\,x_{\langle i_1 i_2\rangle} + B\,\frac{r^2}{c^2}v_{\langle i_1 i_2\rangle} + \frac{G^2m^2\nu}{c^5 r}\,C\,x_{\langle i_1}v_{i_2\rangle}\right), \tag{18}$$

where the terms are explicitly given by[38,39]

$$A = 1 + \gamma\left(-\frac{1}{42} - \frac{13}{14}\nu\right) + \gamma^2\left(-\frac{461}{1512} - \frac{18395}{1512}\nu - \frac{241}{1512}\nu^2\right)$$
$$+ \gamma^3\left(\frac{395899}{13200} - \frac{428}{105}\ln\left(\frac{r}{r_0}\right) + \left[\frac{3304319}{166320} - \frac{44}{3}\ln\left(\frac{r}{r'_0}\right)\right]\nu\right.$$
$$\left. + \frac{162539}{16632}\nu^2 + \frac{2351}{33264}\nu^3\right), \tag{19a}$$

$$B = \frac{11}{21} - \frac{11}{7}\nu + \gamma\left(\frac{1607}{378} - \frac{1681}{378}\nu + \frac{229}{378}\nu^2\right)$$
$$+ \gamma^2\left(-\frac{357761}{19800} + \frac{428}{105}\ln\left(\frac{r}{r_0}\right) - \frac{92339}{5544}\nu + \frac{35759}{924}\nu^2 + \frac{457}{5544}\nu^3\right), \tag{19b}$$

$$C = \frac{48}{7} + \gamma\left(-\frac{4096}{315} - \frac{24512}{945}\nu\right). \tag{19c}$$

Here the PN ordering parameter is $\gamma = \frac{Gm}{rc^2}$ where $r$ is the separation distance in harmonic coordinates. Note the two constant scales entering the logarithmic terms at the 3PN order, one being the length scale $r_0$ coming from the MPM algorithm,[20] while the other one $r'_0$ comes from the 3PN EOM in harmonic coordinates.[30]

The second step is the relationship between the multipole moments of the source, and the so-called "radiative" multipole moments parametrizing the observable GW at future null infinity. Such relationship involves in particular the well-known tail effects and their iterations. At the 4.5PN order the radiative mass quadrupole moment $U_{i_1 i_2}$ is related to the mass quadrupole moment of the source $I_{i_1 i_2}$ by

$$\begin{aligned} U_{i_1 i_1}(t) = I^{(2)}_{i_1 i_2}(t) &+ \frac{GM}{c^3}\int_0^{+\infty} \mathrm{d}\tau\, I^{(4)}_{i_1 i_2}(t-\tau)\left[2\ln\left(\frac{c\tau}{2b_0}\right) + \frac{11}{6}\right] \\ &+ \frac{G^2M^2}{c^6}\int_0^{+\infty} \mathrm{d}\tau\, I^{(5)}_{i_1 i_2}(t-\tau)\left[2\ln^2\left(\frac{c\tau}{2b_0}\right) + \frac{11}{3}\ln\left(\frac{c\tau}{2b_0}\right)\right. \\ &\qquad \left. - \frac{214}{105}\ln\left(\frac{c\tau}{2r_0}\right) + \frac{124627}{22050}\right] \\ &+ \frac{G^3M^3}{c^9}\int_0^{+\infty} \mathrm{d}\tau\, I^{(6)}_{i_1 i_2}(t-\tau)\left[\frac{4}{3}\ln^3\left(\frac{c\tau}{2b_0}\right) + \frac{11}{3}\ln^2\left(\frac{c\tau}{2b_0}\right)\right. \\ &\qquad + \frac{124627}{11025}\ln\left(\frac{c\tau}{2b_0}\right) - \frac{428}{105}\ln\left(\frac{c\tau}{2b_0}\right)\ln\left(\frac{c\tau}{2r_0}\right) \\ &\qquad \left. - \frac{1177}{315}\ln\left(\frac{c\tau}{2r_0}\right) + \frac{129268}{33075} + \frac{428}{315}\pi^2\right]. \end{aligned} \tag{20}$$

For simplicity, we have not included here the non-linear memory effect which arises at 2.5PN order,[22,46,102,106] as well as many instantaneous (non-tails) terms, that are relatively easy to compute. The terms at 1.5PN, 3PN and 4.5PN orders shown in (20) correspond to what can rightly be called the "tail", the "tail-of-tail", and the "tail-of-tail-of-tail", respectively.[84] The expression (20) contains still another arbitrary scale $b_0$, parametrizing the coordinate transformation between harmonic coordinates and radiative coordinates. We find that the scale $b_0$ as well as the two previous scales $r_0$ and $r'_0$ in Eq. (18) cleanly cancel out in the GW flux, expressed in terms of the invariant PN parameter $x$, which is finally given by[24–27,31,38–40]

$$\begin{aligned} \mathcal{F} = \frac{32c^5}{5G}\nu^2 x^5 \Bigg\{ &1 + \left(-\frac{1247}{336} - \frac{35}{12}\nu\right)x + 4\pi x^{3/2} \\ &+ \left(-\frac{44711}{9072} + \frac{9271}{504}\nu + \frac{65}{18}\nu^2\right)x^2 + \left(-\frac{8191}{672} - \frac{583}{24}\nu\right)\pi x^{5/2} \\ &+ \left(\frac{6643739519}{69854400} + \frac{16}{3}\pi^2 - \frac{1712}{105}\gamma_{\mathrm{E}} - \frac{856}{105}\ln(16\,x)\right. \\ &\qquad \left. + \left[-\frac{134543}{7776} + \frac{41}{48}\pi^2\right]\nu - \frac{94403}{3024}\nu^2 - \frac{775}{324}\nu^3\right)x^3 \\ &+ \left(-\frac{16285}{504} + \frac{214745}{1728}\nu + \frac{193385}{3024}\nu^2\right)\pi x^{7/2} + F_{4\mathrm{PN}}\,x^4 \end{aligned}$$

$$+\left(\frac{265978667519}{745113600}-\frac{6848}{105}\gamma_{\rm E}-\frac{3424}{105}\ln(16x)+\left[\frac{2062241}{22176}+\frac{41}{12}\pi^2\right]\nu\right.$$
$$\left.-\frac{133112905}{290304}\nu^2-\frac{3719141}{38016}\nu^3\right)\pi x^{9/2}\Bigg\}. \tag{21}$$

This is valid up to 4.5PN order, with the notable exception that the 4PN coefficient, denoted $F_{\rm 4PN}$ in (21), is not yet known. However, from BH perturbation theory we know already the test mass limit of this coefficient, *i.e.*, in the small mass ratio limit $\nu \to 0$:[63,64,96,100,101]

$$F_{\rm 4PN} = -\frac{323105549467}{3178375200}+\frac{232597}{4410}\gamma_{\rm E}-\frac{1369}{126}\pi^2$$
$$+\frac{39931}{294}\ln 2-\frac{47385}{1568}\ln 3+\frac{232597}{8820}\ln x+\mathcal{O}(\nu)\,. \tag{22}$$

Of course, this nice result from BH perturbation theory will have to be confirmed by PN theory, which will also be able to provide the mass ratio corrections $\mathcal{O}(\nu)$.

## 4. PN theory versus GSF theory

The conservative dynamics and GWs of compact binary systems in the extreme mass ratio limit, is the realm of the perturbative gravitational self force (GSF) theory.[55,57,66,85,92,94] For the conservative dynamics, a comparison between GSF computations and traditional PN calculations was initiated some years ago,[56] applying to systems that are at once slowy moving and with extreme mass ratio, see the overlap region between PN and perturbation theory in Fig. 1. In recent years the possibility for this comparison has been dramatically extended. Such progress is due in large part due to high precision numerical and analytical computations from a self force perspective,[12–14,80–82,95,99] and to extensive analytical computations within the PN approximation.[28,29,34,35,77]

For a particle moving on an exact circular orbit around a Schwarzschild BH (neglecting radiation reaction), one disposes of a very interesting quantity, which is the invariant associated with the helical Killing symmetry, appropriate for exact circular orbits.[56] This invariant, denoted $u_1^t$, is defined by

$$u_1^\mu = u_1^t\, K_1^\mu\,, \tag{23}$$

where $u_1^\mu$ is the normalized four-velocity of the particle 1 (with mass $m_1 \ll m_2$), $K^\mu$ is the helical Killing vector (HKV) and $K_1^\mu$ the HKV at the location of the particle. Adopting a coordinate system in which the HKV reads $K^\mu\partial_\mu = \partial_t + \omega\,\partial_\varphi$, where $\omega$ is the orbital frequency of the circular orbit, the invariant reduces to the time component of the four-velocity hence its name $u_1^t$, and we have

$$u_1^t = \frac{1}{z_1} = \left[-(g_{\mu\nu})_1\frac{v_1^\mu v_1^\nu}{c^2}\right]^{-1/2}, \tag{24}$$

where $(g_{\mu\nu})_1$ is the metric evaluated at the particle's location, following a certain regularization (here $v_1^\mu = {\rm d}y_1^\mu/{\rm d}t$ denotes the coordinate velocity, *i.e.*, $y_1^0 = ct$ and

$v_1^0 = c$). The inverse of $u_1^t$ appears to be a redshift $z_1$, and sometimes $u_1^t$ itself is called the redshift. [For a generalization of the notion of redshift to eccentric orbits, see[5].] In the exact test mass limit $\nu = 0$, the invariant reduces to the one appropriate to a Schwarzschild BH,

$$u^t_{\rm Schw} = \frac{1}{\sqrt{1-3y}}\,, \tag{25}$$

where $y = (\frac{Gm_2\omega}{c^3})^{2/3}$ is the frequency-related PN parameter associated with the larger BH mass $m_2$. The GSF part is then defined as the coefficient of the mass ratio correction beyond Eq. (25),

$$u^t_1 = u^t_{\rm Schw} + \nu\, u^t_{\rm GSF} + \mathcal{O}(\nu^2)\,. \tag{26}$$

It is clear that with this approximation the symmetric mass ratio $\nu = \frac{m_1 m_2}{m^2}$ can be replaced by the ordinary mass ratio $q = \frac{m_1}{m_2}$. In the PN approximation, the GSF part of the redshift factor appears to be an infinite PN series of the type

$$u^t_{\rm GSF} = \sum_{j=0}^{+\infty} \big(\alpha_j + \beta_j \ln y\big)\, y^{j+1}\,. \tag{27}$$

We have included terms linear in the logarithm of $y$, but we neglect (just for this discussion) the higher powers of $\ln y$, that occur at large PN orders. Recall that the most general structure of the PN expansion involves any (integer) powers of the logarithm, $\sim (\ln y)^k y^{j+1}$.[20]

Here we report the results that have been obtained so far using the "traditional" PN approach. Recall that the PN method heavily relies on dimensional regularization (DR) to treat both UV and IR divergences, see Sec. 2. Another feature of the PN calculation is that it requires a machinery of tails and related non-linear effects, see Eq. 20. In the PN approach one computes $u_1^t$ as a redshift in harmonic coordinates using Eq. (24), and evaluates the metric at the particle's location with DR. In that way the GSF redshift has been obtained up to 4PN order as[12,28,29,77]

$$\begin{aligned} u^t_{\rm GSF} =& -y - 2y^2 - 5y^3 + \left(-\frac{121}{3} + \frac{41}{32}\pi^2\right) y^4 \\ &+ \left(-\frac{1157}{15} + \frac{677}{512}\pi^2 - \frac{128}{5}\gamma_{\rm E} - \frac{64}{5}\ln(16y)\right) y^5 + \mathcal{O}(y^6)\,. \end{aligned} \tag{28}$$

In addition, PN theory has been able to fix the logarithmic term at the 5PN order, which is due to a subdominant tail effect, namely

$$\beta_5 = \frac{956}{105}\,, \tag{29}$$

while the coefficient $\alpha_5$ is known only from GSF methods but has not yet been checked with PN theory. The results (28)–(29) are in full agreement with direct GSF computations. This constitutes a strong confirmation of the adequation of DR for traditional PN calculations, as well as of the procedure of subtraction of the singular field which is employed by GSF theory.

A feature of the PN expansion of the redshift factor at high orders is the appearance of half-integral PN approximations, say $\frac{n}{2}$PN. At first sight these terms sound surprising because the dynamics is purely conservative (exactly circular orbits with a HKV), and we are used to the fact that half-integral PN approximations like 2.5PN are associated with dissipative radiation reaction. The existence of such terms in the conservative redshift factor, starting at 5.5PN order, has been pointed out with numerical GSF methods,[99] and later it was proved that these terms originate from iterated non-linear tail effects, called "tails-of-tails".[34,35] The leading 5.5PN, next-to-leading 6.5PN and next-to-next-to-leading 7.5PN coefficients in this category have been found to be

$$\alpha_{\frac{11}{2}} = -\frac{13696}{525}\pi\,, \qquad \alpha_{\frac{13}{2}} = \frac{81077}{3675}\pi\,, \qquad \alpha_{\frac{15}{2}} = \frac{82561159}{467775}\pi\,, \tag{30}$$

while the corresponding $\beta_j$'s are zero. Notice that 7.5PN is arguably the highest order ever reached by traditional PN methods. Again the PN results (30) are in full agreement with numerical and analytical results derived by GSF methods.

## 5. PN versus PM

The post-Minkowskian approximation has been developed in many pionneering works.[6,11,68,89,104,105] Notably the gravitational scattering angle of two relativistic particles has been obtained up to 2PM order (quadratic in $G$).[103] Recently there has been a renewal of interest in the PM approximation. Ledvinka, Schäfer and Bičák[78] obtained a closed-form expression for the Hamiltonian of $N$ particles in the 1PM approximation, and new works appeared on the gravitational scattering angle and the link between the PM expansion and the EOB formalism.[48,49] Here we outline our own contribution,[36] which concerns the comparison between 1PM and the recent 4PN calculation of the EOM of compact binaries. The Fig. 2 showed the domain of validity of the PM approximation *versus* that of the PN expansion.

At the 1PM approximation the field equations for a system of $N$ particles in harmonic coordinates, deduced from the gauge-fixed action (4), read

$$\Box h^{\mu\nu} = \frac{16\pi G}{c^2}\sum_{a=1}^{N} m_a \int_{-\infty}^{+\infty} \mathrm{d}\tau_a\, u_a^\mu u_a^\nu \delta^{(4)}(x - y_a)\,, \tag{31}$$

where $\Box$ denotes the flat space-time d'Alembertian operator, $\delta^{(4)}$ is the four-dimensional Dirac function, $y_a^\mu$ are the particle's worldlines and $u_a^\mu = \mathrm{d}y_a^\mu/(c\mathrm{d}\tau_a)$ their normalized four-velocities, with the special-relativistic proper time $\mathrm{d}\tau_a = \sqrt{-\eta_{\mu\nu}\mathrm{d}y_a^\mu \mathrm{d}y_a^\nu/c^2}$. We solve Eqs. (31) by means of the standard Lienard-Wiechert procedure. Adopting a parametrization by the coordinate time $t$, *i.e.*, such that $y_a^\mu = (ct, \boldsymbol{y}_a)$, the retarded time $t_a^{\mathrm{ret}}$ on the trajectory $a$ associated with the propagation from $a$ to the field point $x^\mu = (ct, \mathbf{x})$, is given by the implicit retardation equation $t_a^{\mathrm{ret}} = t - r_a^{\mathrm{ret}}/c$ with $r_a^{\mathrm{ret}} = |\mathbf{x} - \boldsymbol{y}_a(t_a^{\mathrm{ret}})|$. The solution of (31) is then

$$h^{\mu\nu}(x) = -\frac{4G}{c^2}\sum_a \frac{m_a\, u_a^\mu u_a^\nu}{r_a^{\mathrm{ret}}\,(ku)_a^{\mathrm{ret}}}\,, \tag{32}$$

where $k_a^\mu = [x^\mu - y_a^\mu(t_a^{\rm ret})]/r_a^{\rm ret} = (1, \boldsymbol{n}_a^{\rm ret})$ is the Minkowski null vector between $a$ and the field point, $(ku)_a^{\rm ret} = k_\mu^a u_a^\mu = \gamma_a(-1 + \boldsymbol{n}_a^{\rm ret}\cdot\boldsymbol{v}_a/c)$ is the usual redshift factor, with $\gamma_a = u_a^0$ the Lorentz factor and $v_a^\mu = cu_a^\mu/\gamma_a = (c, \boldsymbol{v}_a)$, the velocities being computed at the retarded time $t_a^{\rm ret}$.

However, we repeatedly use the fact that the accelerations are of order $G$ and therefore their contributions in (32) will be of order $G^2$, hence negligible with the 1PM approximation. Thus, we are allowed to assume that the four velocities $u_a^\mu$ and Lorentz factors $\gamma_a$ are constant. Furthermore, neglecting terms of order $G^2$, we can solve the retardation equation to get the retarded time $t_a^{\rm ret}$, the distance $r_a^{\rm ret}$, and the direction $\boldsymbol{n}_a^{\rm ret}$, in terms of their current values at time $t$, *i.e.*, the "instantaneous" distance $r_a = |\mathbf{x} - \boldsymbol{y}_a(t)|$ and direction $\boldsymbol{n}_a = [\mathbf{x} - \boldsymbol{y}_a(t)]/r_a$. In this way, Eq. (32) becomes equivalent to

$$h^{\mu\nu} = -\frac{4G}{c^2}\sum_a \frac{m_a\, u_a^\mu u_a^\nu}{r_a\sqrt{1+(n_a u_a)^2}}\,, \tag{33}$$

which is valid at any field point except at the singular locations of the particles. Nevertheless, we can easily extend its validity to the particles by using a self-field regularization. For this purpose, it is sufficient to discard the self-field contribution from the sum of particles. An explicit self-field regularization process yielding the same result was implemented in [6]. Therefore, at the location of particle $a$, we have

$$(h^{\mu\nu})_a = -\frac{4G}{c^2}\sum_{b\neq a} \frac{m_b\, u_b^\mu u_b^\nu}{r_{ab}^2\big[1+(n_{ab}u_b)^2\big]^{1/2}}\,, \tag{34}$$

where the sum runs over all particles except $a$, we pose $r_{ab} = |\boldsymbol{y}_a - \boldsymbol{y}_b|$, and denote $n_{ab}^0 = 0$ and $n_{ab}^i = [\boldsymbol{y}_a - \boldsymbol{y}_b]/r_{ab}$.

The EOM of the particles is just the geodesic equation, computed at the linearized order consistent with our approximation, and we obtain

$$\begin{aligned}\frac{{\rm d}u_a^\mu}{{\rm d}\tau_a} = -\frac{1}{c}\sum_{b\neq a}\frac{Gm_b}{r_{ab}^2\big[1+(n_{ab}u_b)^2\big]^{3/2}}\bigg[&(2\epsilon_{ab}^2-1)n_{ab}^\mu \\ &+(2\epsilon_{ab}^2+1)\Big(-(n_{ab}u_a)+\epsilon_{ab}(n_{ab}u_b)\Big)u_a^\mu + \Big(4\epsilon_{ab}(n_{ab}u_a)-(2\epsilon_{ab}^2+1)(n_{ab}u_b)\Big)u_b^\mu\bigg]\,.\end{aligned} \tag{35}$$

We use $\epsilon_{ab} = -(u_a u_b)$ as a shorthand notation. Equivalently we have also the non-covariant form of the EOM (*i.e.*, PN like form), in which we introduce the ordinary velocities and accelerations and the relevant Lorentz factors (with $\boldsymbol{v}_{ab} = \boldsymbol{v}_a - \boldsymbol{v}_b$),

$$\begin{aligned}\frac{{\rm d}\boldsymbol{v}_a}{{\rm d}t} = -\gamma_a^{-2}\sum_{b\neq a}\frac{Gm_b}{r_{ab}^2\big[1+\gamma_b^2(n_{ab}v_b)^2/c^2\big]^{3/2}}\bigg[&(2\epsilon_{ab}^2-1)\boldsymbol{n}_{ab} \\ &+\gamma_b\Big(-4\epsilon_{ab}\gamma_a(n_{ab}v_a)+(2\epsilon_{ab}^2+1)\gamma_b(n_{ab}v_b)\Big)\frac{\boldsymbol{v}_{ab}}{c^2}\bigg]\,.\end{aligned} \tag{36}$$

At the 1PM order the EOM are conservative, thus admit conserved integrals of energy, angular momentum and linear momentum. Indeed, the radiation reaction

dissipative effects are at least 2PM, *i.e.*, second order in $G$. The closed-form expression of the energy through 1PM reads $E = \sum_a m_a c^2 \gamma_a + V$, with the first term being the usual special-relativistic expression and

$$V = G \sum_a \sum_{b \neq a} \frac{m_a m_b}{r_{ab} \left[1 + \gamma_b^2 (n_{ab} v_b)^2 / c^2\right]^{1/2}} \left\{ \gamma_a \left( 2\epsilon_{ab}^2 + 1 - 4 \frac{\gamma_b}{\gamma_a} \epsilon_{ab} \right) \right. \tag{37}$$
$$\left. + \frac{\gamma_b^2}{\gamma_a} \left(2\epsilon_{ab}^2 - 1\right) \frac{\dot{r}_{ab}(n_{ab} v_b) - (v_{ab} v_b)}{\left(v_{ab}^2 - \dot{r}_{ab}^2\right) \left[1 + \gamma_b^2 (n_{ab} v_b)^2 / c^2\right] + \frac{\gamma_b^2}{c^2} \left(\dot{r}_{ab}(n_{ab} v_b) - (v_{ab} v_b)\right)^2} \right\}.$$

We have verified[36] that Eqs. (36) and (37) perfectly reproduce the PN results in harmonic coordinates, in the case of two particles ($N = 2$) at the linear order in $G$ and up to the 4PN order.[8,10]

Next we consider the Lagrangian associated with the EOM (35)–(36), for any $N$, in harmonic coordinates. The Lagrangian will be given by the special-relativistic expression plus terms of order $G$, and again, we neglect higher-order terms in $G$. In PN theory, it is known that the Lagrangian in harmonic coordinates is a *generalized* one, depending not only on positions and velocities $\boldsymbol{y}_a, \boldsymbol{v}_a$ but also on accelerations $\boldsymbol{a}_a = \mathrm{d}\boldsymbol{v}_a/\mathrm{d}t$.[50] Such accelerations are contained in terms at least linear in $G$, so that, replacing the accelerations by the EOM would yield negligible terms of order $G^2$ at least. However, it is not allowed to replace accelerations in a Lagrangian while remaining in the same coordinate system. Such replacement is equivalent to a shift in the particles' trajectories (or so-called "contact" transformation), *i.e.*, the new Lagrangian is physically equivalent to the original one but written in a different coordinate system.[97] Furthermore, by employing the technique of double-zero (or multiple-zero) terms, it is sufficient to consider a Lagrangian that is linear in accelerations (since the procedure can work for any PN order, and is thus formally valid at the 1PM order). Therefore, we look for a Lagrangian of the form

$$L\big[y, v, a\big] = -\sum_a \frac{m_a c^2}{\gamma_a} + \lambda + \sum_a q_a^i a_a^i \,. \tag{38}$$

We symbolize the functional dependence by $L[y, v, a] \equiv L[\{\boldsymbol{y}_a, \boldsymbol{v}_a, \boldsymbol{a}_a\}]$. The terms $\lambda$ and $q_a^i$ are of order $G$ and depend only on positions and velocities, *i.e.*, $\lambda[y, v]$ and $q_a^i[y, v]$. Denoting by $p_a^i$ and $q_a^i$ the conjugate momenta associated with the positions $y_a^i$ and velocities $v_a^i$, *i.e.*,

$$p_a^i = \frac{\delta L}{\delta v_a^i} = \frac{\partial L}{\partial v_a^i} - \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial a_a^i}\right), \tag{39a}$$

$$q_a^i = \frac{\delta L}{\delta a_a^i} = \frac{\partial L}{\partial a_a^i}, \tag{39b}$$

the EOM take the ordinary Euler-Lagrange form

$$\frac{\mathrm{d}p_a^i}{\mathrm{d}t} = \frac{\partial L}{\partial y_a^i}, \tag{40}$$

while the conserved energy $E$ is given by the generalized Legendre transformation

$$E = \sum_a \Big( p_a^i v_a^i + q_a^i a_a^i \Big) - L \,. \tag{41}$$

In both (40) and (41) we are allowed to replace the accelerations by the EOM. For instance, the term $q_a^i a_a^i$ in $E$ will be second-order in $G$ and can be neglected at 1PM order. With (38) we obtain the EOM

$$f_a^i = \frac{\delta \lambda}{\delta y_a^i} + \ddot{q}_a^i \,, \tag{42}$$

where $f_a^i = m_a \frac{\mathrm{d}}{\mathrm{d}t}(\gamma_a v_a^i)$ and $\frac{\delta \lambda}{\delta y_a^i} = \frac{\partial \lambda}{\partial y_a^i} - \frac{\mathrm{d}}{\mathrm{d}t}(\frac{\partial \lambda}{\partial v_a^i})$, and the dots refer to time derivatives. The potential $V$ (such that $E = \sum_a m_a c^2 \gamma_a + V$) reads

$$V = \sum_a v_a^i \frac{\partial \lambda}{\partial v_a^i} - \lambda - \sum_a v_a^i \dot{q}_a^i \,. \tag{43}$$

The left-hand sides of (42) and (43) are known from Eqs. (35)–(36) and (37). However, the two equations are not independent, since $f_a^i$ and $V$ satisfy the constraint

$$\frac{\mathrm{d}V}{\mathrm{d}t} + \sum_a v_a^i f_a^i = 0 \,. \tag{44}$$

In order to find $L$, our strategy is to determine first a particular Lagrangian $\hat{L}$, characterized by $(\hat{\lambda}, \hat{q}_a^i)$, such that the conjugate momenta $\hat{q}_a^i$ obey

$$\sum_a v_a^i \hat{q}_a^i = 0 \,. \tag{45}$$

To order $G$, the same equation is also satisfied by the time derivative $\dot{\hat{q}}_a^i$. Therefore, for the particular solution $(\hat{\lambda}, \hat{q}_a^i)$, the equation (43) reduces to an ordinary Legendre transformation,

$$V = \sum_a v_a^i \frac{\partial \hat{\lambda}}{\partial v_a^i} - \hat{\lambda} \,. \tag{46}$$

To determine $\hat{\lambda}$, we note that the potential $V$ given by (41) reduces in the limit $c \to +\infty$ to the Newtonian approximation, namely $V = U + \mathcal{O}(1/c^2)$ where

$$U = -\sum_{a<b} \frac{G m_a m_b}{r_{ab}} \,. \tag{47}$$

If we subtract its Newtonian limit $U$ to $V$, we get a quantity which tends to zero when $c \to +\infty$ like $\mathcal{O}(1/c^2)$. Then, it is straightforward to show that a well-behaved solution of Eq. (46) is

$$\hat{\lambda} = -U + \frac{1}{c} \int_c^{+\infty} \mathrm{d}s \left[ V\Big( \boldsymbol{y}_a, \frac{\boldsymbol{v}_a}{s} \Big) - U(\boldsymbol{y}_a) \right] . \tag{48}$$

Namely, we have to insert into Eq. (37) all the relevant factors $c$ and make the replacement of $c$ by $s$, then integrate over the "speed of light" $s$ from the physical value $c$ up to infinity. The bound $s \to +\infty$ of the integral corresponds to the

Newtonian limit and we see from the definition of the Newtonian potential (47) that the integral is convergent. The first term in Eq. (48) represents the Newtonian approximation with the correct minus sign for a Lagrangian, and the integral represents formally the complete series of PN corrections, but resummed in the PM approximation. The result (48) can be rewritten in a simpler way as the Hadamard "partie finie" (Pf) of the integral, in the same sense as was used in Eq. (8), for taking care of the divergence at infinity:

$$\hat{\lambda} = \mathrm{Pf}\,\frac{1}{c}\int_{c}^{+\infty} \mathrm{d}s\, V\Big(\boldsymbol{y}_a, \frac{\boldsymbol{v}_a}{s}\Big)\,. \tag{49}$$

For this very simple type of divergence $\sim s^0 + \mathcal{O}(s^{-2})$ the Pf does not depend on an arbitrary constant, unlike in (8). The expressions (48)–(49) give a particular solution of the equation (43) but we still have to adjust $\hat{q}_a^i$ in order to satisfy the EOM, see (42). Thus, we look for $\hat{q}_a^i$ satisfying

$$\ddot{\hat{q}}_a^i = f_a^i - \frac{\delta\hat{\lambda}}{\delta y_a^i}\,, \tag{50}$$

where the right-hand side is known. To order $G$ we have been able to integrate twice this relation to determine $\hat{q}_a^i$, and that solution automatically satisfies the constraint (45) by virtue of (44).

Finally we have found a particular Lagrangian $(\hat{\lambda}, \hat{q}_a^i)$. Now the general solution $(\lambda, q_a^i)$ can be obtained by adding an arbitrary total time-derivative $\mathrm{d}F/\mathrm{d}t$, where $F$ is a function of the positions $y_a^i$ and velocities $v_a^i$. Hence the general solution (for the class of harmonic-coordinate Lagrangians that are linear in accelerations) reads

$$\lambda = \hat{\lambda} + \sum_a v_a^i \frac{\partial F}{\partial y_a^i}\,, \tag{51a}$$

$$q_a^i = \hat{q}_a^i + \frac{\partial F}{\partial v_a^i}\,. \tag{51b}$$

At 1PM order the Lagrangian in harmonic coordinates irreducibly depends on accelerations, *i.e.*, it is impossible to determine $F$ such that $q_a^i = 0$. However, we know that the accelerations in a Lagrangian can be eliminated by appropriate shifts of the trajectories. In fact, it can be shown that the particular solution $\hat{\lambda}$ found in (48)–(49) represents an ordinary Lagrangian which is physically equivalent but expressed in some shifted (non harmonic) variables.[36]

Given the complicated structure of $V$ in Eq. (37), we could not find a closed form expression for the 1PM harmonic coordinate Lagrangian in the general case. However, we could easily work out the integral (48) in the PN approximation $c \to +\infty$ to any order. We start from the known 4PN expansion of the potential $V$ following from (37), and explicitly perform the integration (48) term by term, to obtain the corresponding 4PN expansion of $\hat{\lambda}$. Then, we derive the coefficient of accelerations $\hat{q}_a^i$ at 4PN order, see[36] for details. Finally, we find a unique total

time-derivative, with some function $F_{\rm PN}$ given in the form of a PN expansion, so that the Lagrangian satisfyingly agrees up to order $G$ with the published 4PN Lagrangian.[8,10] Furthermore, we have pushed the analysis to the next order and obtained all the terms of order $G$ in the harmonic coordinates Lagrangian up to the 5PN order.[36]

In another application, we worked out the case of equal masses for which it is possible to find a closed form expression for the Lagrangian, and we have verified that the associated Hamiltonian differs from the one obtained by Ledvinka, Schäfer and Bičák[78] in the ADM Hamiltonian by a mere canonical transformation.

## References


1. B. Abbott et al., *Observation of gravitational waves from a binary black hole merger*, Phys. Rev. Lett., 116 (2016), p. 061102.
2. ——, *Gw170817: Observation of gravitational waves from a binary neutron star inspiral*, Phys. Rev. Lett., 119 (2017), p. 161101.
3. P. Ajith, M. Hannam, S. Husa, Y. Chen, B. Bruegmann, N. Dorband, D. Mueller, F. Ohme, D. Pollney, C. Reisswig, L. Santamaria, and J. Seiler, *Inspiral-merger-ringdown waveforms for black-hole binaries with non-precessing spins*, Phys. Rev. Lett, 106 (2011), p. 241101.
4. J. G. Baker, J. Centrella, D.-I. Choi, M. Koppitz, and J. van Meter, *Gravitational wave extraction from an inspiraling configuration of merging black holes*, Phys. Rev. Lett., 96 (2006), p. 111102.
5. L. Barack and N. Sago, *Beyond the geodesic approximation: conservative effects of the gravitational self-force in eccentric orbits around a schwarzschild black hole*, Phys. Rev. D, 83 (2011), p. 084023.
6. L. Bel, T. Damour, N. Deruelle, J. Ibanez, and J. Martin, *Poincaré invariant gravitational field and equations of motion of point-like objects: The post-linear approximation of general relativity*, Gen. Relativ. Gravit., 13 (1981), p. 963.
7. L. Bernard, L. Blanchet, A. Bohé, G. Faye, and S. Marsat, *Fokker action of non-spinning compact binaries at the fourth post-Newtonian approximation*, Phys. Rev. D, 93 (2016), p. 084037.
8. L. Bernard, L. Blanchet, A. Bohé, G. Faye, and S. Marsat, *Dimensional regularization of the ir divergences in the fokker action of point-particle binaries at the fourth post-newtonian order*, Phys. Rev. D, 96 (2017), p. 104043.
9. ——, *Energy and periastron advance of compact binaries on circular orbits at the fourth post-newtonian order*, Phys. Rev. D, 95 (2017), p. 044026.
10. L. Bernard, L. Blanchet, G. Faye, and T. Marchand, *Center-of-mass equations of motion and conserved integrals of compact binary systems at the fourth post-newtonian order*, Phys. Rev. D, 97 (2018), p. 044037.

11. B. Bertotti and J. Plebanski, *Theory of gravitational perturbations in the fast motion approximation*, Ann. Phys. (N. Y.), 11 (1960), p. 169.
12. D. Bini and T. Damour, *Analytical determination of the two-body gravitational interaction potential at the fourth post-newtonian approximation*, Phys. Rev. D, 87 (2013), p. 121501(R).
13. ——, *Analytic determination of the eight-and-a-half post-Newtonian self-force contributions to the two-body gravitational interaction potential*, Phys. Rev. D, 89 (2014), p. 104047.
14. ——, *High-order post-Newtonian contributions to the two-body gravitational interaction potential from analytical gravitational self-force calculations*, Phys. Rev. D, 89 (2014), p. 064063.
15. L. Blanchet, *Radiative gravitational fields in general relativity. ii. asymptotic behaviour at future null infinity*, Proc. Roy. Soc. Lond. A, 409 (1987), pp. 383–399.
16. ——, *Time asymmetric structure of gravitational radiation*, Phys. Rev. D, 47 (1993), pp. 4392–4420.
17. ——, *Second post-newtonian generation of gravitational radiation*, Phys. Rev. D, 51 (1995), pp. 2559–2583.
18. ——, *Gravitational radiation reaction and balance equations to post-newtonian order*, Phys. Rev. D, 55 (1997), pp. 714–732.
19. ——, *On the multipole expansion of the gravitational field*, Class. Quant. Grav., 15 (1998), pp. 1971–1999.
20. L. Blanchet and T. Damour, *Radiative gravitational fields in general relativity. i. general structure of the field outside the source*, Phil. Trans. Roy. Soc. Lond. A, 320 (1986), pp. 379–430.
21. ——, *Tail-transported temporal correlations in the dynamics of a gravitating system*, Phys. Rev. D, 37 (1988), p. 1410.
22. ——, *Hereditary effects in gravitational radiation*, Phys. Rev. D, 46 (1992), pp. 4304–4319.
23. L. Blanchet, T. Damour, and G. Esposito-Farèse, *Dimensional regularization of the third post-newtonian dynamics of point particles in harmonic coordinates*, Phys. Rev. D, 69 (2004), p. 124007.
24. L. Blanchet, T. Damour, G. Esposito-Farèse, and B. R. Iyer, *Gravitational radiation from inspiralling compact binaries completed at the third post-newtonian order*, Phys. Rev. Lett., 93 (2004), p. 091101.
25. ——, *Dimensional regularization of the third post-newtonian gravitational wave generation of two point masses*, Phys. Rev. D, 71 (2005), p. 124004.
26. L. Blanchet, T. Damour, and B. R. Iyer, *Gravitational waves from inspiralling compact binaries: Energy loss and wave form to second post-newtonian order*, Phys. Rev. D, 51 (1995), pp. 5360–5386.
27. L. Blanchet, T. Damour, B. R. Iyer, C. M. Will, and A. G. Wiseman, *Gravitational radiation damping of compact binary systems to second post-newtonian order*, Phys. Rev. Lett., 74 (1995), pp. 3515–3518.

28. L. Blanchet, S. Detweiler, A. Le Tiec, and B. Whiting, *Higher-order post-Newtonian fit of the gravitational self-force for circular orbits in the schwarzschild geometry*, Phys. Rev. D, 81 (2010), p. 084033.
29. ——, *Post-Newtonian and numerical calculations of the gravitational self-force for circular orbits in the schwarzschild geometry*, Phys. Rev. D, 81 (2010), p. 064004.
30. L. Blanchet and G. Faye, *General relativistic dynamics of compact binaries at the third post-newtonian order*, Phys. Rev. D, 63 (2001), p. 062005.
31. L. Blanchet, G. Faye, B. R. Iyer, and B. Joguet, *Gravitational-wave inspiral of compact binary systems to 7/2 post-newtonian order*, Phys. Rev. D, 65 (2002), p. 061501(R). Erratum *Phys. Rev. D*, 71:129902(E), 2005.
32. L. Blanchet, G. Faye, and S. Nissanke, *Structure of the post-newtonian expansion in general relativity*, Phys. Rev. D, 72 (2005), p. 044024.
33. L. Blanchet, G. Faye, and B. Ponsot, *Gravitational field and equations of motion of compact binaries to 5/2 post-newtonian order*, Phys. Rev. D, 58 (1998), p. 124002.
34. L. Blanchet, G. Faye, and B. Whiting, *Half-integral conservative post-newtonian approximations in the redshift factor of black hole binaries*, Phys. Rev. D, 89 (2014), p. 064026.
35. ——, *High-order half-integral conservative post-newtonian coefficients in the redshift factor of black hole binaries*, Phys. Rev. D, 90 (2014), p. 044017.
36. L. Blanchet and A. Fokas, *Equations of motion of self-gravitating n-body systems in the first post-minkowskian approximation*, Phys. Rev. D, 98 (2018), p. 084005.
37. L. Blanchet and B. R. Iyer, *Third post-newtonian dynamics of compact binaries: Equations of motion in the center-of-mass frame*, Class. Quant. Grav., 20 (2003), p. 755.
38. ——, *Hadamard regularization of the third post-newtonian gravitational wave generation of two point masses*, Phys. Rev. D, 71 (2004), p. 024004.
39. L. Blanchet, B. R. Iyer, and B. Joguet, *Gravitational waves from inspiralling compact binaries: Energy flux to third post-newtonian order*, Phys. Rev. D, 65 (2002), p. 064005. Erratum *Phys. Rev. D*, 71:129903(E), 2005.
40. L. Blanchet, B. R. Iyer, C. M. Will, and A. G. Wiseman, *Gravitational wave forms from inspiralling compact binaries to second-post-newtonian order*, Class. Quant. Grav., 13 (1996), pp. 575–584.
41. L. Blanchet and A. Le Tiec, *First law of compact binary mechanics with gravitational-wave tails*, Class. Quant. Grav., 34 (2017), p. 164001.
42. P. Brady, J. Creighton, and K. Thorne, *Computing the merger of black-hole binaries: The ibbh problem*, Phys. Rev. D, 58 (1998), p. 061501.
43. A. Buonanno and T. Damour, *Effective one-body approach to general relativistic two-body dynamics*, Phys. Rev. D, 59 (1999), p. 084006.

44. A. Buonanno and B. Sathyaprakash, *Sources of gravitational waves: Theory and observations*, in General Relativity and Gravitation: A Centennial Perspective, A. Ashtekar, B. Berger, J. Isenberg, and M. MacCallum, eds., 2015, p. 513.
45. M. Campanelli, C. O. Lousto, P. Marronetti, and Y. Zwlochower, *Accurate evolutions of orbiting black-hole binaries without excision*, Phys. Rev. Lett., 96 (2006), p. 111101.
46. D. Christodoulou, *Nonlinear nature of gravitation and gravitational-wave experiments*, Phys. Rev. Lett., 67 (1991), pp. 1486–1489.
47. T. Damour, *Gravitational radiation and the motion of compact bodies*, in Gravitational Radiation, N. Deruelle and T. Piran, eds., Amsterdam, 1983, North-Holland Company, pp. 59–144.
48. ——, *Gravitational scattering, post-minkowskian approximation and effective one-body theory*, Phys. Rev. D, 94 (2016), p. 104015.
49. ——, *High-energy gravitational scattering and the general relativistic two-body problem*, Phys. Rev. D, 97 (2018), p. 044038.
50. T. Damour and N. Deruelle, *Lagrangien généralisé du système de deux masses ponctuelles, à l'approximation post-post-newtonienne de la relativité générale*, C. R. Acad. Sc. Paris, 293 (1981), p. 537.
51. T. Damour, P. Jaranowski, and G. Schäfer, *Dimensional regularization of the gravitational interaction of point masses*, Phys. Lett. B, 513 (2001), pp. 147–155.
52. T. Damour, P. Jaranowski, and G. Schäfer, *Non-local-in-time action for the fourth post-newtonian conservative dynamics of two-body systems*, Phys. Rev. D, 89 (2014), p. 064058.
53. ——, *On the conservative dynamics of two-body systems at the fourth post-newtonian approximation of general relativity*, Phys. Rev. D, 93 (2016), p. 084014.
54. V. de Andrade, L. Blanchet, and G. Faye, *Third post-newtonian dynamics of compact binaries: Noetherian conserved quantities and equivalence between the harmonic-coordinate and adm-hamiltonian formalisms*, Class. Quant. Grav., 18 (2001), pp. 753–778.
55. B. De Witt and R. Brehme, *Radiation damping in a gravitational field*, Ann. Phys. (N.Y.), 9 (1960), pp. 220–259.
56. S. Detweiler, *A consequence of the gravitational self-force for circular orbits of the schwarzschild geometry*, Phys. Rev. D, 77 (2008), p. 124026.
57. S. Detweiler and B. Whiting, *Self-force via a green's function decomposition*, Phys. Rev. D, 67 (2003), p. 024025.
58. A. Einstein, L. Infeld, and B. Hoffmann, *The gravitational equations and the problem of motion*, Ann. Math., 39 (1938), pp. 65–100.
59. S. Foffa, P. Mastrolia, R. Sturani, and C. Sturm, *Effective field theory approach to the gravitational two-body dynamics at fourth post-newtonian order and quintic in the newton constant*, Phys. Rev. D, 95 (2017), p. 104009.

60. S. Foffa and R. Sturani, *Effective field theory calculation of conservative binary dynamics at third post-Newtonian order*, Phys. Rev. D, 84 (2011), p. 044031.
61. ——, *The dynamics of the gravitational two-body problem in the post-Newtonian approximation at quadratic order in the Newton's constant*, Phys. Rev. D, 87 (2013), p. 064011.
62. ——, *Tail terms in gravitational radiation reaction via effective field theory*, Phys. Rev. D, 87 (2013), p. 044056.
63. R. Fujita, *Gravitational radiation for extreme mass ratio inspirals to the 14th post-newtonian order*, Prog. Theor. Phys., 127 (2012), p. 583.
64. ——, *Gravitational waves from a particle in circular orbits around a schwarzschild black hole to the 22nd post-newtonian order*, Prog. Theor. Phys., 128 (2012), p. 971.
65. C. R. Galley, A. K. Leibovich, R. A. Porto, and A. Ross, *Tail effect in gravitational radiation reaction: Time nonlocality and renormalization group evolution*, Phys. Rev. D, 93 (2016), p. 124010.
66. S. Gralla and R. Wald, *A rigorous derivation of gravitational self-force*, Class. Quant. Grav., 25 (2008), p. 205009.
67. M. Hannam, S. Husa, J. A. Gonzalez, U. Sperhake, and B. Brueg-mann, *Where post-newtonian and numerical-relativity waveforms meet*, Phys. Rev. D, 77 (2007), p. 044020.
68. P. Havas and J. Goldberg, *Lorentz-invariant equations of motion of point masses in the general theory of relativity*, Phys. Rev., 128 (1962), p. 398.
69. Y. Itoh, *Equation of motion for relativistic compact binaries with the strong field point particle limit: Third post-newtonian order*, Phys. Rev. D, 69 (2004), p. 064018.
70. ——, *Third-and-a-half order post-newtonian equations of motion for relativistic compact binaries using the strong field point particle limit*, Phys. Rev. D, 80 (2009), p. 024003.
71. Y. Itoh and T. Futamase, *New derivation of a third post-newtonian equation of motion for relativistic compact binaries without ambiguity*, Phys. Rev. D, 68 (2003), p. 121501(R).
72. Y. Itoh, T. Futamase, and H. Asada, *Equation of motion for relativistic compact binaries with the strong field point particle limit: The second and half post-newtonian order*, Phys. Rev. D, 63 (2001), p. 064038.
73. P. Jaranowski and G. Schäfer, *Third post-newtonian higher order adm hamilton dynamics for two-body point-mass systems*, Phys. Rev. D, 57 (1998), pp. 7274–7291.
74. P. Jaranowski and G. Schäfer, *Towards the fourth post-Newtonian Hamiltonian for two-point-mass systems*, Phys. Rev. D, 86 (2012), p. 061503(R).

75. ——, *Dimensional regularization of local singularities in the 4th post-newtonian two-point-mass hamiltonian*, Phys. Rev. D, 87 (2013), p. 081503(R).
76. C. Königsdörffer, G. Faye, and G. Schäfer, *The binary black-hole dynamics at the third-and-a-half post-newtonian order in the adm-formalism*, Phys. Rev. D, 68 (2003), p. 044004.
77. A. Le Tiec, L. Blanchet, and B. Whiting, *The first law of binary black hole mechanics in general relativity and post-newtonian theory*, Phys. Rev. D, 85 (2012), p. 064039.
78. T. Ledvinka, G. Schäfer, and J. Bičák, *Relativistic closed-form hamiltonian for many-body gravitating systems in the post-minkowskian approximation*, Phys. Rev. Lett., 100 (2008), p. 251101.
79. H. Lorentz and J. Droste, *The motion of a system of bodies under the influence of their mutual attraction, according to Einstein's theory*, Nijhoff, The Hague, 1937, p. 330. Versl. K. Akad. Wet. Amsterdam **26**, 392 and 649 (1917).
80. S. Mano, H. Susuki, and E. Takasugi, *Analytic solutions of the regge-wheeler equation and the post-minkowskian expansion*, Prog. Theor. Phys., 96 (1996), p. 549.
81. ——, *Analytic solutions of the teukolsky equation and their low frequency expansions*, Prog. Theor. Phys., 95 (1996), p. 1079.
82. S. Mano and E. Takasugi, *Analytic solutions of the teukolsky equation and their properties*, Prog. Theor. Phys., 97 (1997), p. 213.
83. T. Marchand, L. Bernard, L. Blanchet, and G. Faye, *Ambiguity-free completion of the equations of motion of compact binary systems at the fourth post-newtonian order*, Phys. Rev. D, 97 (2018), p. 044023.
84. T. Marchand, L. Blanchet, and G. Faye, *Gravitational-wave tail effects to quartic non-linear order*, Class. Quant. Grav., 33 (2016), p. 244003.
85. Y. Mino, M. Sasaki, and T. Tanaka, *Gravitational radiation reaction to a particle motion*, Phys. Rev. D, 55 (1997), p. 3457.
86. S. Nissanke and L. Blanchet, *Gravitational radiation reaction in the equations of motion of compact binaries to 3.5 post-newtonian order*, Class. Quant. Grav., 22 (2005), p. 1007.
87. T. Ohta, H. Okamura, T. Kimura, and K. Hiida, *Higher-order gravitational potential for many-body system*, Prog. Theor. Phys., 51 (1974), pp. 1220–1238.
88. M. Pati and C. Will, *Post-newtonian gravitational radiation and equations of motion via direct integration of the relaxed einstein equations. ii. two-body equations of motion to second post-newtonian order, and radiation-reaction to 3.5 post-newtonian order*, Phys. Rev. D, 65 (2002), p. 104008.
89. M. Portilla, *Scattering of two gravitating particles: classical approach*, J. Phys. A, 13 (1980), p. 3677.
90. R. Porto and I. Rothstein, *On the apparent ambiguities in the post-newtonian expansion for binary systems*. eprint arXiv, 2017.

91. O. Poujade and L. Blanchet, *Post-newtonian approximation for isolated systems calculated by matched asymptotic expansions*, Phys. Rev. D, 65 (2002), p. 124020.
92. A. Pound, *Self-consistent gravitational self-force*, Phys. Rev. D, 81 (2010), p. 024023.
93. F. Pretorius, *Evolution of binary black hole spacetimes*, Phys. Rev. Lett., 95 (2005), p. 121101.
94. T. Quinn and R. Wald, *Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime*, Phys. Rev. D, 56 (1997), p. 3381.
95. N. Sago, L. Barack, and S. Detweiler, *Two approaches for the gravitational self force in black hole spacetime: Comparison of numerical results*, Phys. Rev. D, 78 (2008), p. 124024.
96. M. Sasaki, *Post-newtonian expansion of the ingoing-wave regge-wheeler function*, Prog. Theor. Phys., 92 (1994), pp. 17–36.
97. G. Schäfer, *Acceleration-dependent lagrangians in general relativity*, Phys. Lett. A, 100 (1984), p. 128.
98. G. Schäfer, *The gravitational quadrupole radiation-reaction force and the canonical formalism of adm*, Ann. Phys. (N. Y.), 161 (1985), pp. 81–100.
99. A. Shah, J. Friedmann, and B. Whiting, *Finding high-order analytic post-newtonian parameters from a high-precision numerical self-force calculation*, Phys. Rev. D, 89 (2014), p. 064042.
100. H. Tagoshi and M. Sasaki, *Post-newtonian expansion of gravitational-waves from a particle in circular orbit around a schwarzschild black-hole*, Prog. Theor. Phys., 92 (1994), pp. 745–771.
101. T. Tanaka, H. Tagoshi, and M. Sasaki, *Gravitational waves by a particle in circular orbit around a schwarzschild black hole: 5.5 Post-Newtonian formula*, Prog. Theor. Phys., 96 (1996), pp. 1087–1101.
102. K. Thorne, *Gravitational-wave bursts with memory: The christodoulou effect*, Phys. Rev. D, 45 (1992), p. 520.
103. K. Westpfahl, *Highspeed scattering of charged and uncharged particles in general relativity*, Fortschr. Physik, 33 (1985), p. 417.
104. K. Westpfahl and M. Goller, *Gravitational scattering of two relativistic particles in post-linear approximation*, Lett. Nuovo Cim., 26 (1979), p. 573.
105. K. Westpfahl and H. Hoyler, *Gravitational bremsstrahlung in post-linear fast-motion approximation*, Lett. Nuovo Cim., 27 (1980), p. 581.
106. A. Wiseman and C. Will, *Christodoulou's nonlinear gravitational-wave memory: Evaluation in the quadrupole approximation*, Phys. Rev. D, 44 (1991), pp. R2945–R2949.

# DAMPE and its latest results

Fabio Gargano on behalf of the DAMPE Collaboration

*Istituto Nazionale di Fisica Nucleare (INFN) Sezione di Bari, I-70125 Bari, Italy*
*E-mail: fabio.gargano@ba.infn.it*

Dark Matter Particle Explorer (DAMPE), the first Chinese astronomical satellite, was launched into a Sun-synchronous orbit at an altitude of about 500 km on 17 December 2015. DAMPE is a high-energy particle detector optimized for observations of cosmic ray electrons and gamma-rays up to about 10 TeV. The on-orbit performance of the detector, the calibration, and the latest results on cosmic rays and gamma-rays of DAMPE will be presented.

*Keywords*: Astro-particle; Cosmic-rays; Space Mission; Gamma-ray astronomy

## 1. Introduction

The DArk Matter Particle Explorer (DAMPE)[1] is a particle detector launched on 17 December 2015 in a Sun-synchronous orbit at a 500 km altitude. Its main scientific goals are the study of cosmic-rays, namely electrons, photons, protons and nuclei; the observation of gamma-rays from astrophysical sources; the search for dark matter signatures, for electromagnetic counterparts of gravitational waves or neutrinos and for exotic particles. DAMPE has been designed to detect particles with energy from some GeV up to tens of TeV with an energy resolution $\leq 1.5\%$ at 100 GeV for electromagnetic showers, an accurate angular resolution $\leq 0.1^\circ$ at 100 GeV for gamma-rays and a field of view of 1 sr.

## 2. The DAMPE instrument

DAMPE is composed by four sub-detectors (Figure 1): a Plastic Scintillator Detector (PSD), a Silicon-Tungsten Tracker (STK), a BGO Calorimeter (CALO) and a Neutron Detector (NUD). The PSD has two tasks: to work as a veto for charged particles and to measure the charge number $Z$ of incident high-energy particles. It is made by two layers of staggered scintillator bars in orthogonal arrangements, providing information on the $x$ and $y$ coordinates. The STK is devoted to reconstruct the particle tracks and to convert photons. It consists of 12 position-sensitive silicon detector planes (6 for the $x$-, 6 for the $y$-coordinate); 3 layers of tungsten, each 1 mm thick, are between the silicon planes 2, 3, 4 and 5 to enhance the conversion probability of gamma rays into electron-positron pairs. The BGO calorimeter is used to measure the energy deposition of incident particles and to reconstruct the electromagnetic shower profile. It is composed of 308 BGO crystal bars optically isolated from each other and arranged in 14 layers of 22 bars each; the bars of a layer are orthogonal to those of the adjacent one, to reconstruct the shower in the $xz$ and $yz$ views. The total depth of the calorimeter is 32 radiation lengths and 1.6 nuclear interaction lengths. The last sub-detector is the NUD, made of four

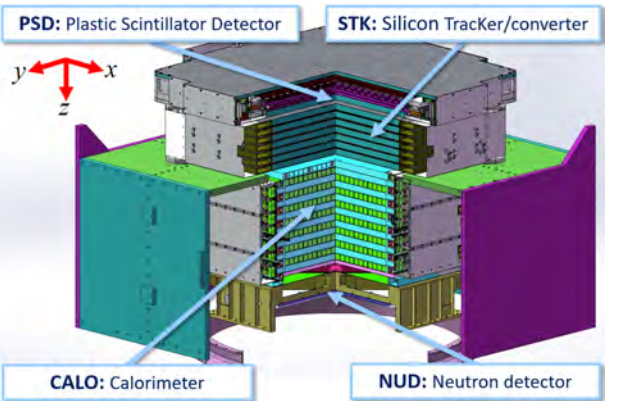



Fig. 1. The DAMPE instrument.

boron-loaded plastic scintillators. It is used to improve overall hadron identification efficiency.

### 2.1. *The Plastic Scintillation array Detector (PSD)*

The main purpose of the PSD is to provide charged-particle background rejection for the gamma-ray detection and to measure the absolute value of the charge (hereafter Z) of incident high-energy particles in a wide range (i.e., $Z \leq 26$). Therefore high detection efficiency, large dynamic range, and good charge resolution are required for charged particle detection of PSD.

The PSD has an active area of $82.5 \times 82.5\text{cm}^2$, which is larger than the on-axis cross-section of other subdetectors of DAMPE. The PSD consists of 82 plastic scintillators bars (EJ-200 produced by Eljen) arranged in two planes, each with a double layer configuration. Each bar is 88.4cm long with a $2.81.0\text{cm}^2$ cross-section; the signals are read out by two Hamamatsu R4443 Photomultiplier Tubes (PMTs) coupled to the ends of each scintillator bar.

The bars in the top plane are perpendicular to those in the bottom plane. The bars of the two layers of a plane are staggered by 0.8cm, allowing full coverage of the detector with the active area of scintillators without any gap. As the efficiency of a single layer is $\geq 0.95$, the PSD provides an overall efficiency $\geq 0.9975$ for charged particles. The segmented structure of the PSD allows suppressing the spurious veto signals due to the "backsplash effect", which can lead to a misidentification of gamma rays as charged particles.

Since the PSD is used to identify cosmic-ray nuclei from helium to iron ($Z = 26$), a wide dynamic range extending up to $\sim 1400$ times the energy deposition of a minimum ionizing particle (MIP) is required. To cover such a broad range with good energy resolution, a double dynode readout scheme for each PMT has been

implemented. Signals from the dynode with high gain cover the range from 0.1 MIPs to 40 MIPs, while those from the dynode with low gain cover the range from 4 MIPs to 1600 MIPs; the overlap region can be used for cross-calibration. The dynode signals are coupled to the VA160 ASIC chip developed by IDEAS. This chip integrates the charge sensitive preamplifier, the shaper and the holding circuit. The performance of the PSD has been tested with the relativistic heavy ion beams at CERN. In this test, the primary Argon beam of 40 GeV/n was sent onto a 40 mm polyethylene target, and the secondary fragments with A/Z = 2 were selected by beam magnets, thus allowing to study the PSD response to all the stable nuclei with Z = 2/18. Fig. 2 shows the reconstructed charge spectra for different ions ($Z \geq 2$) from one PSD module within the beam spot. In this figure, the Helium contribution has been removed for clarity (the He fraction is much higher than that of other ion species). The signals from both sides of each module are used (geometric mean) and the quenching effect has been corrected based on the ion response from the same test.

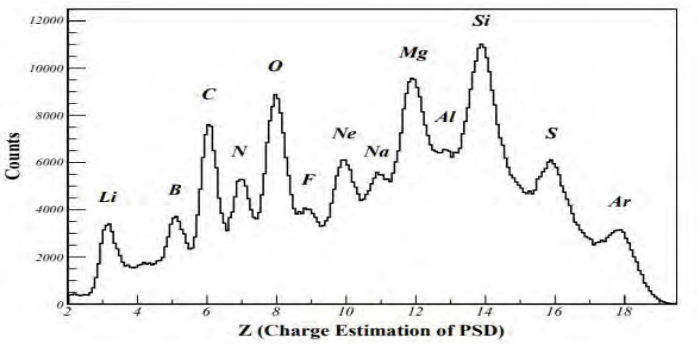



Fig. 2. Reconstructed charge spectra of PSD for nuclei with A/Z = 2, generated by a 40 GeV/n 40Ar beam. The helium peak has been removed for clarity.

### 2.2. *The Silicon-Tungsten tracKer-converter (STK)*

The DAMPE STK is designed to accomplish the following tasks: precise particle track reconstruction with a resolution better than 80 $\mu$m for most of the incident angles, measurement of the electrical charge of incoming cosmic rays, and photon conversion to electron-positron pairs.

It is composed of six position-sensitive double (X and Y) planes of silicon detectors with a total area of about $7m^2$. Multiple thin tungsten layers have been inserted in the tracker structure in order to enhance the photon conversion rate while keeping negligible multiple scattering of electron/positron pairs (above $\sim$ 5 GeV). The total thickness of STK corresponds to about one radiation length, mainly due to

the tungsten layers. The STK detector is equipped with a total of 768 single-sided AC-coupled silicon microstrip detectors (SSD). Four SSDs are assembled together with a wire bonded strip-to-strip connection to form a silicon detector ladder. The total strip length along a ladder is about 37cm. The ladders are glued on the seven support trays to form the 12 STK silicon layers. Each silicon layer consists of 16 ladders. The silicon ladders on the bottom surface of each tray are placed orthogonally with respect to the ones of the top surface of the lower tray, in order to measure the X-Y coordinates of the incident particles. The silicon microstrip sensors produced by Hamamatsu Photonics have a size of $95 \times 95 \times 0.32$ mm$^3$ and each SSD is segmented in 768 strips. The strips are 48 $\mu$m wide and 93.196 mm long with a pitch of 121 $\mu$m.

The signal shaping and amplification is performed by six VA140 ASIC chips (produced by IDEAS)

The ions charge identification power of STK was evaluated with a dedicated test conducted on single ladder unit at CERN with a lead beam. The particle charge can be identified by looking at the mean value of the signal associated with the track. Due to the non-linearity of the VAs above a signal of 200 fC, the identification of ions above Oxygen with the STK becomes non-trivial and on-going work is under preparation to improve the charge identification power (see Fig. 3).

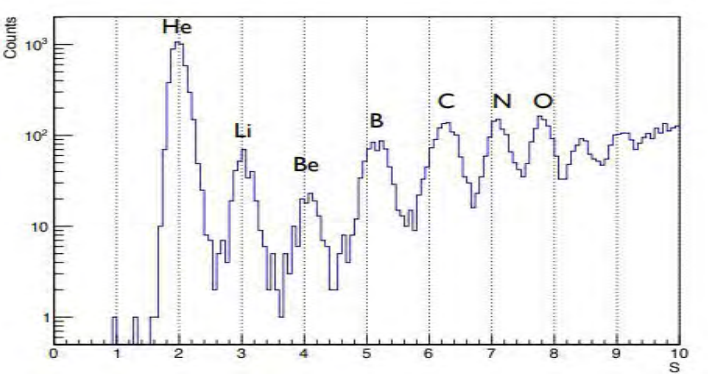



Fig. 3. STK signal mean distribution for nuclei produced by a lead beam on target, after removing Z=1 particles. The signal mean, with current reconstruction procedures, allows for the identification of ions until Oxygen.

### 2.3. *The BGO calorimeter (BGO)*

The BGO calorimeter onboard DAMPE has three primary purposes: (1) measuring the energy deposition of incident particles; (2) imaging the 3D profile (both longitudinal and transverse) of the shower development, and provide electron/hadron discrimination; (3) providing the level 0 trigger for the DAMPE data acquisition

system. Each crystal is readout by two Hamamatsu R5610A-01 PMTs mounted on both ends (named S0 and S1, respectively): the left/right light asymmetry provides a measurement of the position of the energy deposit along the bar. The signals are read out from three different dynodes (dy2, dy5, dy8), thus allowing to cover a very large dynamic range of energy deposition, The PMTs are coupled to the crystals with optical filters, which attenuate the scintillation light produced in the BGO. The filter on S1 has a 5× attenuation factor with respect to the one on S0. The signals are sent to VA160 chip which is composed of a charge sensitive pre-amplifier, a CR-RC shaping amplifier and a sample-hold circuit. The ground calibration of BGO has been performed using both the data collected in a beam test campaign at CERN and cosmic ray data collected from ground. The calibration procedure includes the measurement of the pedestals, the evaluation of the calibration constants from the MIP peaks, the evaluation of the dynode ratios, and the measurement of the bar attenuation lengths.

The data shown in the figure was obtained during the beam test campaigns performed at CERN. The linearity of reconstructed energy is better than 1%, and the energy resolution is better than 1.2% at the energies above 100 GeV (see Fig. 4).

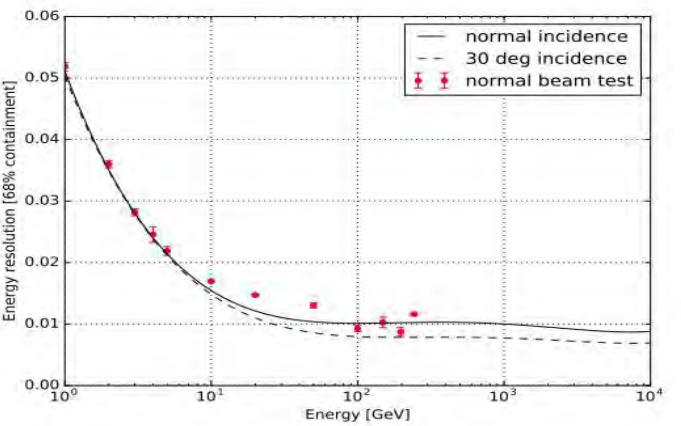



Fig. 4. Energy resolution for gamma rays and electrons/positrons at normal incidence (solid line) and at 30° off-axis angle (dashed line). DAMPE beam test results (with electrons) are over-plotted

### 2.4. *The NeUtron Detector (NUD)*

The main purpose of the NUD is to perform electron/hadron identification using the neutrons produced in hadronic showers initiated in the BGO calorimeter. In fact, for given initial particle energy, the neutron content of a hadronic shower is expected to be one order of magnitude larger than that of an electromagnetic shower. Once

the neutrons are created, they are quickly thermalized in the BGO calorimeter, and the total neutron activity over a few microseconds is measured by NUD.

It consists of four $30 \times 30 \times 1.0\text{cm}^3$ blocks of boron-loaded plastic scintillator (Eljen Technologies EJ-254), with 5% boron concentration by weight which has the natural $^{10}B$ abundance of 20%. Each scintillator is wrapped with a layer of aluminum film for photon reflection, anchored in aluminum alloy framework by silicone rubber.

The scintillators are embedded with wavelength shift fibers for optical transmission in order to reduce the fluorescence attenuation and increase photon collection efficiency, and then the signals are read out by corner-on Hamamatsu R5610A-01 PMTs.

Preliminary estimates, based on on-orbit calibration data, show that a rejection power is $\sim 12.5$ for incoming particles with BGO energy deposit above 800 GeV.

## 3. On-orbit performance

### 3.1. *Trigger rate*

The on-orbit trigger rate depends on the flux of the cosmic rays. It is lower in the equatorial regions but higher in the polar ones due to the Earth's magnetic field bending effect.

Four different triggers have been implemented: Unbiased trigger, MIP trigger, High Energy trigger and Low Energy trigger. They are "OR-ed" to generate the global trigger signal for the detector.

The Unbiased trigger requires signals in the two top BGO layers exceeding a low threshold of ∼0.4MIPs in each hit BGO bar. The MIP trigger aims to select particles crossing all the BGO layers. The High Energy trigger selects events with energy depositions in the top four BGO layers exceeding a high threshold of ∼10MIPs in each hit BGO Bar. The Low Energy trigger is similar to the High Energy one, but with a lower threshold of ∼2MIPs. A periodic signal of 100 Hz is also implemented in the trigger board for pedestal calibration.

The Unbiased, MIP and Low Energy triggers are pre-scaled with the ratios of 512 : 1, 4 : 1, 8 : 1, respectively, when the satellite is in the low latitude region ($\pm 20°$). At high latitudes, the MIP trigger is disabled and the pre-scaler ratios of Unbiased and Low Energy triggers are set to 2048 : 1 and 64 : 1, respectively.

The expected average rate of global triggers is about 70 Hz in flight (the rate of High Energy triggers is 50 Hz, the rate of Unbiased triggers is about 2.5 Hz) (see Fig. 5 and 6)

The DAQ systems work in an "event by event" mode, and a 3 ms time interval is set to acquire each event so that the dead time is fixed to 3 ms as a consequence.

About 15 GB of raw and control data are transmitted to ground each day and, after off-line processing, 85 GB of reconstructed data is produced. The yearly data production is 35 TB.

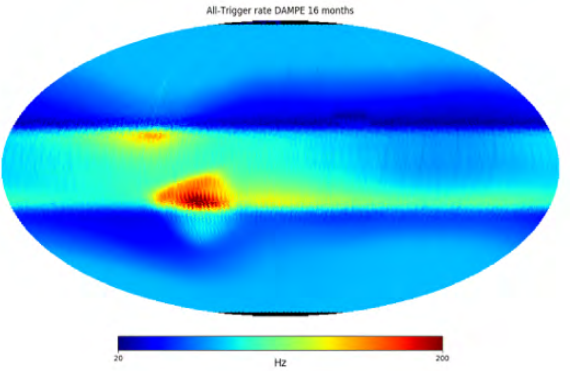



Fig. 5. All trigger rate. It is evident the effect of the pre-scaling at $\pm 20^{\circ}$ and the South Atlantic Anomaly.

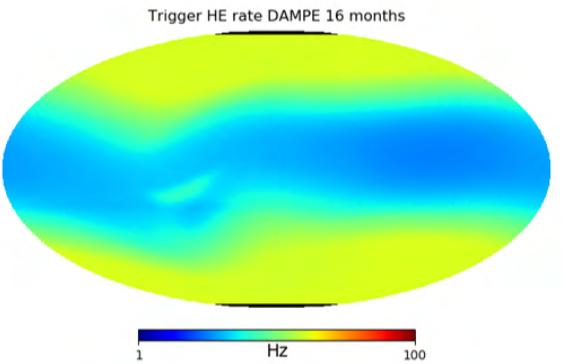



Fig. 6. High Energy trigger rate

### 3.2. *STK Alignment*

The position resolution of the STK silicon sensors is better than 70 $\mu$m. At the same time, the construction precision of the mechanical assembly of STK is about 100 $\mu$m. To maximally profit from the position resolution capabilities of the silicon detectors, alignment parameters are introduced to correct for shifts and rotations of each silicon sensor with respect to its position in the nominal instrument design. Alignment corrections are calculated based on the data and then used in the reconstruction procedure to correct the coordinates of particle hits.

Given the mechanical stability of the tracker structure and the limited temperature variation of the detector, the alignment parameters are updated every two weeks to ensure optimal tracking performance. We estimate the effective position resolution for protons in the x and y internal tracker planes, after the alignment procedure is applied, to be $47\pm2$ $\mu$m for events arriving at normal incidence ($\leq 10^{\circ}$), $41\pm2$ $\mu$m for intermediate inclinations and $45\pm3$ $\mu$m at high incidence angles ($\geq 45^{\circ}$) (see Fig. 7).

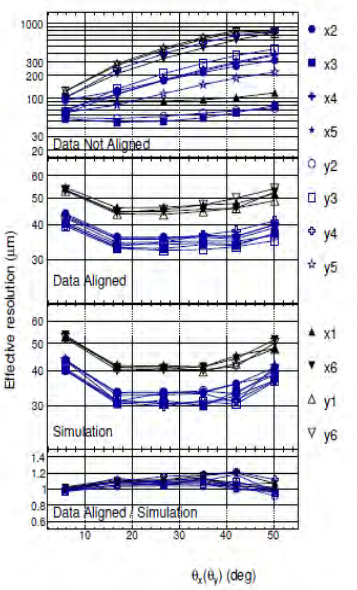



Fig. 7. Alignment results for different STK planes and comparison with simulation.

### 3.3. *Absolute energy calibration*

A method to determine the absolute energy scale of DAMPE has been developed within the Collaboration[2]. In fact, cosmic-ray electrons and positrons below a certain rigidity are bent back to space by the Earth's magnetic field and this causes a clear cutoff in their spectrum. This cutoff was calculated using the geomagnetic field model IGRF-12[3] and back-tracking the particle trajectories in the magnetic field, yielding a value of about 13 GeV, which was compared with the measured cutoff in the interval of McIlwain $L$ parameter between 1.0 and 1.14. A correction factor of 1.2% to be applied to DAMPE data was obtained (Fig. 8).

For a complete review of the on-orbit calibration of all the sub detectors please refer to[12].

## 4. Measurement of the all-electron spectrum

The behavior of the spectrum of the cosmic electrons and positrons (CREs) at very high energies is related to some of the most relevant questions of physics and astrophysics[4,5], and its study is one of the main goals of DAMPE, that can extend its measurements up to about 10 TeV with an acceptance of 0.3 $m^2$sr and an energy resolution of $\sim$1% for energies above 100 GeV.

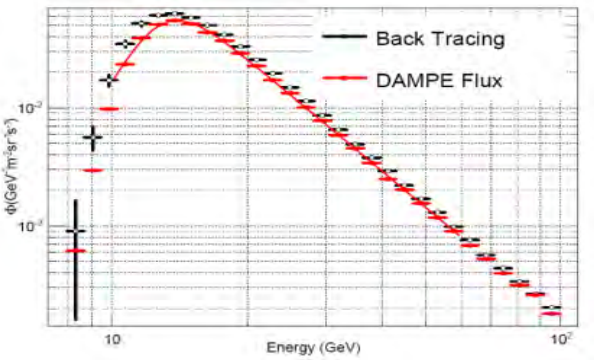



Fig. 8. Energy calibration: the geomagnetic cutoff (left) and its stability over time (right).

To study CREs, the particle identification (PID) and the separation from protons are crucial and can be carried out in several ways, based on the sub-detectors response and on the event topology. In our main analysis, a dedicated parameter, called $\zeta$, was computed from the lateral shower development size and the energy deposition in the last layer of the BGO. Its capability to discriminate electrons from protons was validated with beam tests at CERN: for 90% electron efficiency, the hadron background is ∼2% at 1 TeV and ∼10% at 5 TeV. Other PID strategies have been investigated within the Collaboration: for example, machine learning algorithms and neural networks showed to be very effective and are currently used in new analyses.

The CRE spectrum (Fig. 9) has been measured by DAMPE from about 20 GeV up to about 5 TeV[6]. The main feature is a spectral break at ∼1 TeV, so a smoothly broken power law, with $\gamma$ varying from 3.1 to 3.9, fits DAMPE data. The next step is to improve the statistics and search for spectral structures that may be related to possible sources such as nearby pulsars, or dark matter.

## 5. Measurement of cosmic protons and nuclei spectra

Among DAMPE's tasks, the measurement of the charge of the incident cosmic rays plays a relevant role. The scientific goal is to improve the direct measurement of the cosmic rays spectrum from a few GeV up to 100 TeV — a region that is poorly investigated — thanks to the very good calorimetric and geometric characteristics of the instrument. The capabilities of the PSD and the STK to measure the charge of ions nuclei have been evaluated and tuned with a beam test campaign at the CERN SPS, in which many tests have been performed with Lead and Argon beams that produced ion fragments with a wide range of charge number $Z$. Once in orbit, the PSD has measured the charge from protons to Iron ions with a resolution ranging from 0.2 to 0.4, while the STK has measured the charge from protons to Oxygen ions. Fig. 10 shows the cosmic rays charge spectrum after two years of on-orbit data taking[13].

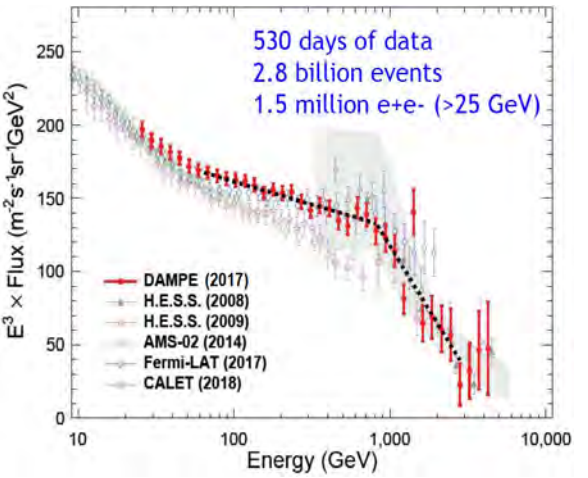



Fig. 9. The all-electron spectrum measured by DAMPE, compared with results by HESS[7,8], AMS-02[9], Fermi[10] and CALET[11].

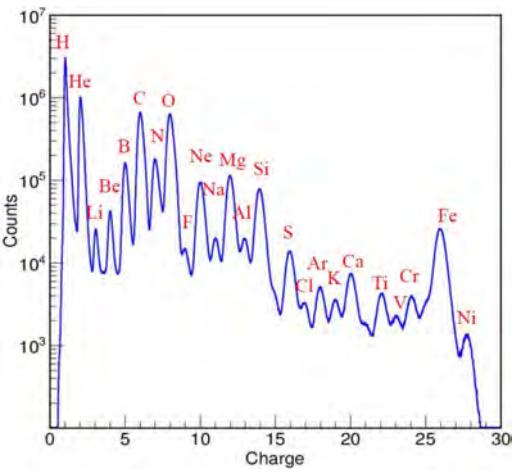



Fig. 10. The cosmic rays charge spectrum measured by DAMPE after two years of data taking.

Also, the fluxes of protons and Helium ions (Fig. 11) as a function of the kinetic energy are being measured. The preliminary results show good agreement with the PAMELA[14] and AMS-02[15] data for $E - k < 200$ GeV. With the accumulation of data, the region from 1 TeV up to 100 TeV is being analyzed by DAMPE with high precision and new results are expected soon.

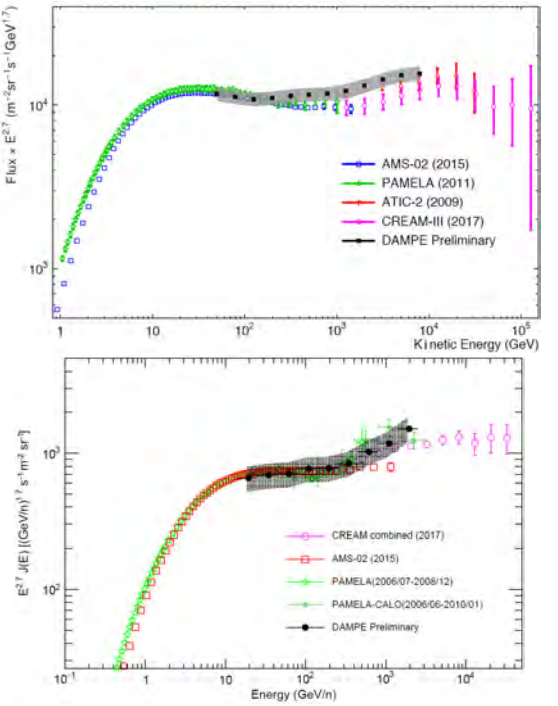



Fig. 11. The preliminary fluxes of CR protons (left) and Helium (right) measured by DAMPE, compared with previous results by PAMELA [14], AMS-02 [15], ATIC-2 [16] and CREAM [17]. The error bars represent statistical errors and the grey bands show the systematic errors.

## 6. Photons

DAMPE is also able to detect high-energy cosmic gamma rays. The photon selection is a challenging task since the background of charged particles has fluxes much higher than the galactic gamma-ray emission: the minimum rejection power required at 100 GeV is $10^5$ for protons and $10^3$ for electrons. The rejection techniques are based on the event topology. Protons are mainly suppressed using the PSD response and the shower profile in the BGO calorimeter, with a contribution from the NUD, while electrons are mainly rejected using the PSD response and the first plane of the STK. Also, innovative procedures, like neural networks and "random forest" classifiers [19], are under test and show excellent particle identification performance. An average number of 150 photons is identified daily by DAMPE. The major sources are resolved with a good angular resolution and their positions agree with those measured by the Fermi-LAT.

DAMPE can also observe the time behavior of variable sources. For instance, Fig. 12 shows the phase profile of the Geminga pulsar while Fig. 13 shows the short-term variability of the CTA 102 blazar.

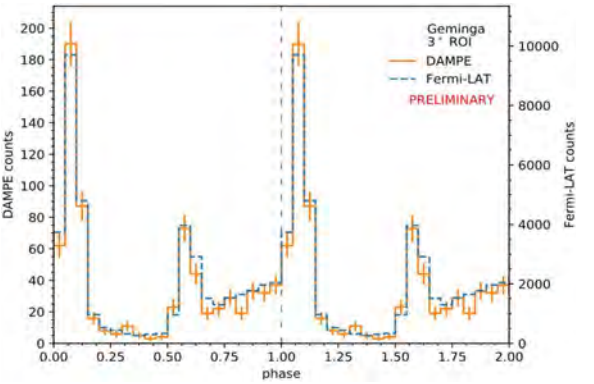



Fig. 12. Phase profile of the Geminga pulsar, compared with Fermi[18].

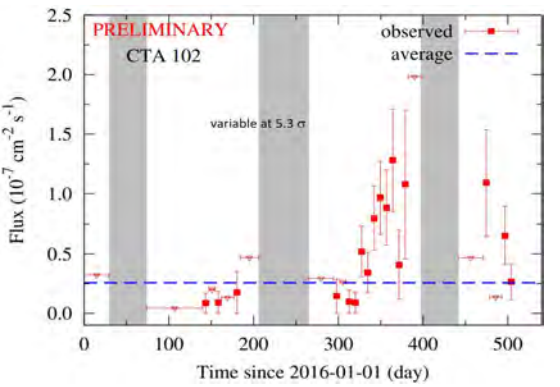



Fig. 13. Short-term variations in the gamma-ray emission of the CTA 102 blazar[20].

Finally, DAMPE participates to the multi-messenger observation of high-energy cosmic phenomena: for example, it has observed the extragalactic source TXS 0506+056 which is the possible origin of the 290 TeV muon neutrino detected by IceCube in September 2017[21].

## 7. Summary

DAMPE is working extremely well since December 2015. Its geometric factor is about 0.3 $m^2$sr for electrons, its Si-W tracker reaches a 40 $\mu$m spatial resolution and a 0.15° angular resolution and its 32 $X-0$ deep BGO calorimeter provides, at energies above 100 GeV, ∼1% energy resolution for cosmic gamma rays and ∼40% for hadrons. Absolute energy calibration is performed using the geomagnetic cutoff.

The all-electron spectrum has been precisely measured up to TeV energies and a spectral break has been directly measured at ∼1 TeV; improved precision of

this spectrum may shed light on nearby sources, anisotropies, dark matter. Proton, Helium and nuclei measurements are ongoing with good preliminary results. Photon detection capability has been assessed; more statistics are being collected to profit the excellent energy resolution at high energy.

## Acknowledgments

The DAMPE mission is founded by the strategic priority science and technology projects in space science of the Chinese Academy of Sciences XDA04040000 and XDA04040400. In China it is supported by the National Key R&D Program 2016YFA0400200, the National Basic Research Program 2013CB837000, the NSFC grants 11525313, 11622327, XDB23040000, 11273070, 11303096, 11303105, 11303106, 11303107, 11673075, U1531126, U1631111 and the 100 Talents Program of Chinese Academy of Sciences. In Europe, the activities are supported by the Italian National Institute for Nuclear Physics (INFN), the Italian University and Research Ministry (MIUR), the Swiss National Science Foundation (SNSF) and the University of Geneva.

## References

1. Chang J *et al.* 2017 *Astropart. Phys.* **95** 6
2. Zang J *et al.* 2017 *Proc. of Sci.* PoS(ICRC2017)197
3. Thebault E *et al.* 2015 *Earth, Planets and Space* **67** 79
4. Fan Y Z, Zhang B and Chang J 2010 *Int. J. Mod. Phys.* D **19** 2011
5. Bertone G, Hooper D and Silk J 2005 *Phys. Rep.* **405** 279
6. Ambrosi G *et al.* 2017 *Nature* **552** 63
7. Aharonian F *et al.* 2008 *Phys. Rev. Lett.* **101** 261104
8. Aharonian F *et al.* 2009 *Astron. Astrophys.* **508** 561
9. Aguilar M *et al.* 2014 *Phys. Rev. Lett.* **113** 221102
10. Abdollahi S *et al.* 2017 *Phys. Rev.* D **95** 082007
11. Adriani O *et al.* 2018 *Phys. Rev. Lett.* **120** 261102
12. Ambrosi G *et al.* 2019 *Astropart. Phys.* **106** 18
13. Dong T *et al.* 2019 *Astropart. Phys.* **105** 31
14. Adriani O *et al.* 2011 *Science* **332** 69
15. Aguilar M *et al.* 2015 *Phys. Rev. Lett.* **114** 171103
16. Panov A D *et al.* 2009 *Bull. Russ. Acad. Sci. Phys.* **73** 564
17. Yoon Y S *et al.* 2017 *Astrophys. J.* **839** 5
18. Abdo A A *et al.* 2010 *Astrophys. J.* **720** 272
19. Garrappa S *et al.* 2017 *Proc. of Sci.* PoS(ICRC2017)603
20. Xu Z L *et al.* 2016 *Astronomer's Telegram* No. 9901
21. Aartsen M G *et al.* 2018 *Science* **361** 6398

# A brief review of binary driven hypernova

Jorge A. Rueda

*ICRA and Dipartimento di Fisica, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, I-00185 Roma, Italy*
*ICRANet, Piazza della Repubblica 10, I-65122 Pescara, Italy*
*ICRANet-Ferrara, Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Saragat 1, I–44122 Ferrara, Italy*
*Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Saragat 1, I–44122 Ferrara, Italy*
*INAF, Istituto di Astrofisica e Planetologia Spaziali, Via Fosso del Cavaliere 100, 00133 Rome, Italy*
*E-mail: jorge.rueda@icra.it*

Remo Ruffini

*ICRA and Dipartimento di Fisica, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, I-00185 Roma, Italy*
*ICRANet, Piazza della Repubblica 10, I-65122 Pescara, Italy*
*INAF, Viale del Parco Mellini 84, 00136 Rome, Italy*
*E-mail: ruffini@icra.it*

Rahim Moradi and Yu Wang*

*ICRA and Dipartimento di Fisica, Università di Roma "La Sapienza", Piazzale Aldo Moro 5, I-00185 Roma, Italy*
*ICRANet, Piazza della Repubblica 10, I-65122 Pescara, Italy*
*INAF – Osservatorio Astronomico d'Abruzzo, Via M. Maggini snc, I-64100, Teramo, Italy*
*E-mail: rahim.moradi@inaf.it*
** E-mail: yu.wang@inaf.it*

Binary driven hypernova (BdHN) models long gamma-ray burst (GRBs) as occurring in the binary systems involving a carbon-oxygen core ($CO_{core}$) and a companion neutron star (NS) or a black hole (BH). This model, first proposed in 2012, succeeds and improves upon the fireshell model and the induced gravitational collapse (IGC) paradigm. After nearly a decade of development, the BdHN model has reached a nearly complete structure, giving explanation to all the observables of long bursts into its theoretical framework, and has given a refined classification of long GRBs according to the original properties of the progenitors. In this article, we present a summary of the BdHN model and the physical processes at work in each of the envisaged Episodes during its occurrence and lifetime, duly contextualized in the framework of GRB observations.

*Keywords*: MG15 Proceedings; High-energy Astrophysics; Gamma-ray Burst; Binary Driven Hypernova.

## 1. Introduction

| Class | Type | Number | *In-state* | *Out-state* | $E_{p,i}$ (MeV) | $E_{iso}$ (erg) | $E_{iso,Gev}$ (erg) |
|---|---|---|---|---|---|---|---|
| Binary Driven | I | 329 | $CO_{core}$-NS | $\nu$NS-BH | $\sim 0.2$–$2$ | $\sim 10^{52}$–$10^{54}$ | $\gtrsim 10^{52}$ |
| Hypernova | II | (30) | $CO_{core}$-NS | $\nu$NS-NS | $\sim 0.01$–$0.2$ | $\sim 10^{50}$–$10^{52}$ | – |
| (BdHN) | III | (19) | $CO_{core}$-NS | $\nu$NS-NS | $\sim 0.01$ | $\sim 10^{48}$–$10^{50}$ | – |
| | IV | 5 | $CO_{core}$-BH | $\nu$NS-BH | $\gtrsim 2$ | $> 10^{54}$ | $\gtrsim 10^{53}$ |

GRBs occur in binary systems of two main classes, binary driven hypernovae (BdHNe) and binary mergers (BMs), observationally corresponding to long and short GRBs. In BdHNe, the long GRB is generated by a type Ic SN explosion of an evolved star occurring in presence of a close-by NS or BH companion. This article is dedicated to the BdHN systems. In BMs, the short GRB is generated from the merger of two compact stars, mostly from the NS-NS systems, and for low energetic short GRBs, the binary white dwarfs could be the progenitors, see Rueda et al. [19,20] and references therein.

The progenitor of a BdHN is a binary system composed of a carbon-oxygen core ($CO_{core}$) and a magnetized neutron star (NS) companion in a tight orbit (period of the order of a few minutes). In some cases, the companion might be a stellar-mass BH (see below). We focus here on the more frequent case of a NS companion. At the end of its thermonuclear evolution, the iron core of the pre-SN star (the $CO_{core}$) undergoes gravitational collapse, forming a new NS (hereafter $\nu$NS) at the SN centre. In the $\nu$NS formation process, a strong shockwave of kinetic energy $\sim 10^{51}$ erg expands outward and when it emerges (SN breakout) expels the $CO_{core}$ outer layers as the SN ejecta. Part of the ejecta is subsequently accreted onto the companion NS and also onto the $\nu$NS by fallback. There are different possible fates for the NS due to the hypercritical accretion process [1,2,4]. For short binary periods ($\lesssim 5$ min), the NS reaches the critical mass for gravitational collapse and forms a BH. We call this subclass BdHN of type I (BdHN I). Thus, a BdHN I leads to a new binary composed of a $\nu$NS originated by the SN, and a BH originated by the collapse of the NS companion. For longer binary periods, the hypercritical accretion onto the NS is not sufficient to bring it to the critical mass, and a more massive NS (MNS) is formed. This subclass is named BdHN of type II (BdHN II). A BdHN II leads to a new binary composed of a $\nu$NS and a massive NS. For very long binary periods, the accretion energy is significantly lower than the above types, and only the hypernova is observed. We call this subclass of sources of type III (BdHN III). In addition, we have BdHN type IV (BdHN IV) for the progenitors of a $CO_{core}$ and a companion BH, and it leads to a new binary of NS and BH.

Having given the physical picture and the classifications of BdHNe, we will present in the following sections the theory and associated observables of the BdHNe. The BdHN starts from the final evolution of the binary stars, including the SN and the accretion of the SN ejecta onto the companion star, to the formation of BH and the particle acceleration mechanisms processing therein, then to the generated relativistic outflow propagates and interacts with the SN ejecta and the interstellar medium (ISM) giving rise to the emissions.

## 2. Binary Accretion

For the binary accretion and the forming of BH, we refer to the theoretical work of Becerra et al. [1–4], Cipolletta et al. [6], Fryer et al. [7,8], Rueda & Ruffini [15] and the observational papers of Izzo et al. [9,10], Ruffini et al. [26,34], Wang et al. [37].

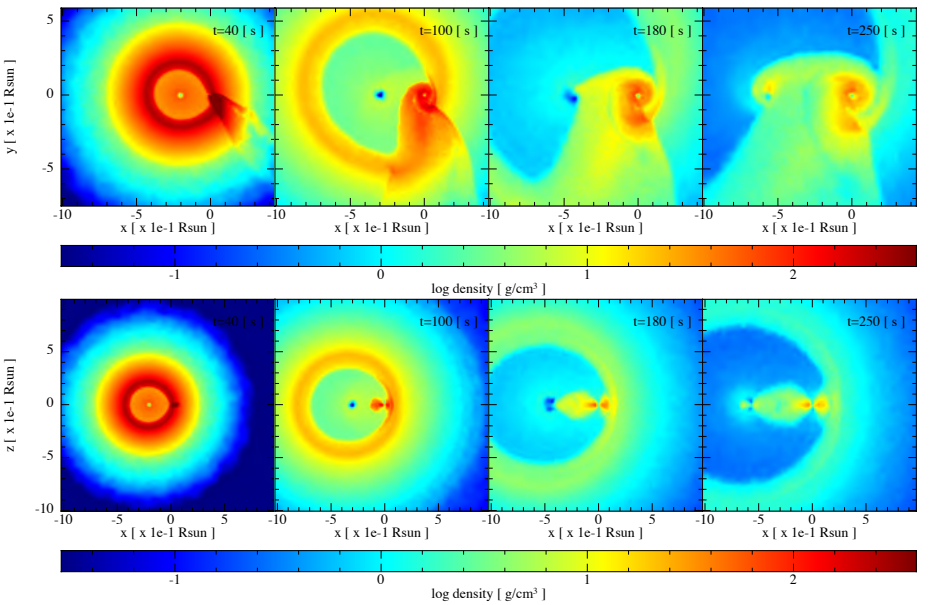



Fig. 1. Snapshot of the SPH simulation of the binary accretion. The initial binary system consists of a $\mathrm{CO}_{rmcore}$ ($M_{rmzams} = 25M_{\odot}$) and an NS ($2M_{\odot}$) with an initial orbital period of about 5 min. The upper panel shows the mass density on the equatorial plane of the binary at different times of the simulation, while the lower panel corresponds to the plane orthogonal to the equator. At $t = 40$ s, particles captured by NS can be seen forming a kind of tail behind them, then these particles form a circle around NS, and at $t = 100$ s a thick disk is observed. At $t = 180$ s, NS starts to accrete the surrounding matters. After about one initial orbital period, at $t = 250$ s, a disk-like structure has formed around the two stars. This figure is cited from Becerra et al. [3].

Rueda & Ruffini [15] has been the first article to consider accretion of SN ejecta onto a very close-by companion star of the binary period of minutes, and it gave the physical picture and the theoretical architectures of a simple one-dimensional model that calculates the Bondi-Hoyle-Lyttleton hypercritical accretion rate. Fryer et al. [8] numerically simulated for the first time the BdHN hypercritical accretion. Following the collapse of the $CO_{\rm core}$ of forming an SN, of which the ejecta falling onto the Bondi-Hoyle surface of the companion with an accretion rate $> 10^{-2}\mathrm{M}_{\odot}\ \mathrm{s}^{-1}$, these one-dimensional numerical simulations give the density and the velocity profiles till the NS reaches the critical mass of BH in tens or hundreds of seconds. Becerra et al. [2] went one step further performing two-dimensional numerical simulations and incorporating angular momentum transport from the SN ejecta to the NS of hypercritical accretion. Those simulation show that under some conditions outflow is necessarily formed because of the excess of angular momentum. Fryer et al. [7] demonstrated that most BdHN with tight orbits (i.e. BdHN I) remain bound after the explosion and accretion, even when a large fraction of mass (over half of the total binary mass) is lost, and a large kick velocity is induced. Becerra et al. [1] performed the first three-dimensional numerical simulations of the BdHNe process,

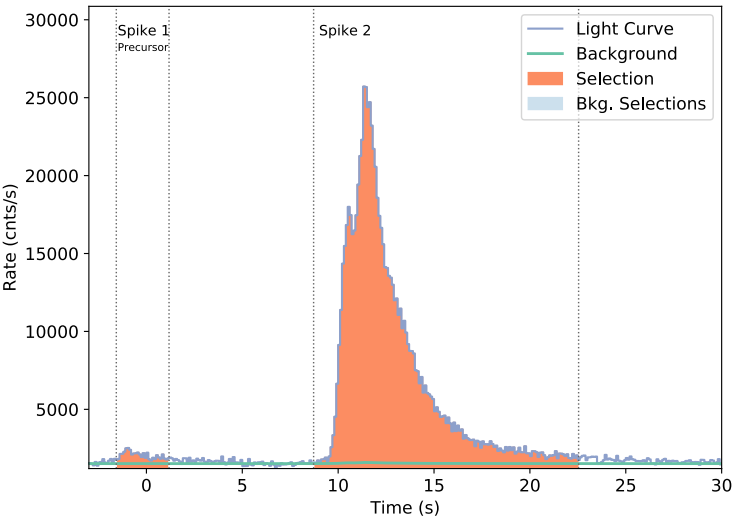



Fig. 2. Light curve of prompt emission of GRB 180728A observed by Fermi-GBM. It contains two pulses. The first pulse ranges from −1.57 s to1.18 s. The second pulse rises at 8.72 s, peaks at 11.50 s, and fades at 22.54 s. This figure is quoted from [37].

which were further upgraded and improved in Becerra et al. [3] via smoothed-particle-hydrodynamics (SPH) simulations, see e.g. in figure 1. A wide selection of initial parameters and several NS equations of state have been there tested. It was there evaluated the outcome of the NS and the $\nu$NS after the hypercritical accretion, namely whether they reach or not the mass-shedding limit, or gravitationally collapses to a BH, or become a more massive and fast-spinning NS. The development of accretion theory and simulations has led to clarifying the physical processes below the Bondi radius: the dominant pressure is supported by the random pressure of the infalling matter, the magnetic pressure is negligible. Such a pressure provides a very high temperature of 1–10 MeV, generating a large abundance of neutrinos and photons. The photons are trapped within the inflowing material, as its diffusion velocity is slower than the inflow velocity. The escape of neutrinos takes away most gravitational energy from the accreted flow, allowing the hypercritical accretion to continue for a given period of time. Becerra et al. [4] further investigated the neutrino flavour oscillations that occur during the propagation of neutrinos emitted from the surface of a neutron star. The final neutrino flow is composed of $\sim 55\%$ ($\sim$62%) of $\sim$ MeV electronic neutrinos for the normal (inverted) neutrino mass hierarchy. In addition, Cipolletta et al. [6] present the numerical calculation and give useful fitting formulas for the location, binding energy and angular momentum of the last stable orbit of test particles around rotating NSs in full general relativity. The results

of this work allows to estimate in full general relativity the amount of energy and angular transferred by the accreting matter to an accreting, rotating NS.

Let us now dive into one specific example. GRB 180728A well demonstrates the binary accretion scenario by its two pulses in the prompt emission[37], see figure 2. At a given time, the $CO_{core}$ collapses, forms a $\nu$NS, and produces an SN explosion. A powerful shockwave is generated and emerges from the SN ejecta. A typical SN shockwave carries $\sim 10^{51}$ erg of kinetic energy, which is partially converted into electromagnetic radiation with an efficiency of $\sim 10\%$. Thus, the energy of $\sim 10^{50}$ erg is consistent with the total energy of the first pulse, which lasts $\sim 2$ s and contains $\sim 8 \times 10^{49}$ erg in keV-MeV photons. The second pulse rises at $\sim 10$ s and subsides at $\sim 10$ s, with a luminosity $\sim 2 \times 51$ erg s$^{-1}$. The distance of the binary separation can be estimated by the delay time between these two pulses, i.e. $\sim 10$ s. Because of the SN ejecta front shell moves at $\sim 0.1c$, we estimate the distance of the binary separation to be about $3 \times 10^{10}$ cm. By given the binary separation and some typical initial parameters, our simulation shows the total mass accreted is $\sim 10^{-2}\ M_\odot$, most of the mass is accreted in $\sim 10$ s with an accretion rate of $\sim 10^{-3}\ M_\odot$ s$^{-1}$. These results are consistent with the second pulse whose total energy is $\sim 10^{51}$ erg, considering an increase in efficiency of $\sim 10\%$, and the luminosity of $\sim 10^{50}$ erg s$^{-1}$ in 10 s duration. The spectrum of the second pulse contains a thermal component which again hints the action of the accretion process. A time-resolved analysis of the thermal component suggests that a mildly relativistic source is expanding and radiating. This radiation is interpreted as an adiabatic expansion heat outflow from the accretion region. The Rayleigh-Taylor convective instability plays a role in the initial accretion phase, driving matter out of the accreting NS with a final velocity of the order of the speed of light. As the matter expands and cools, the temperature evolution from the theory is again consistent with the observations. This kind of thermal emission of BdHN was first found in[9,10]. GRB 180728A offers a good example of BdHN II, and for BdHN I, which has a tighter binary separation, see also the case of GRB 130427A as an example[26,37].

## 3. The Inner Engine and the GeV Emission

We have introduced the inner engine theory for the explanation of the GRB high-energy (GeV) emission observed in some BdHN I after the prompt emission phase. We here summarize the inner engine properties following the calculations presented in Campion et al.[5], Moradi et al.[14], Rueda & Ruffini[16], Ruffini et al.[32]; see also Liang et al.[11], Moradi et al.[12,13], Ruffini et al.[29,33].

Once the NS reaches the critical mass, a fast-rotating BH forms which contains sufficient rotational energy ($> 10^{54}$ erg) to power a GRB. Ruffini et al.[32] proposed an efficient way to extract energy from the newborn, Kerr BH.

The inner engine is composed of this newborn rotating BH, surrounded by the magnetic field inherited from the collapsed NS[17], and the ionized very low density of

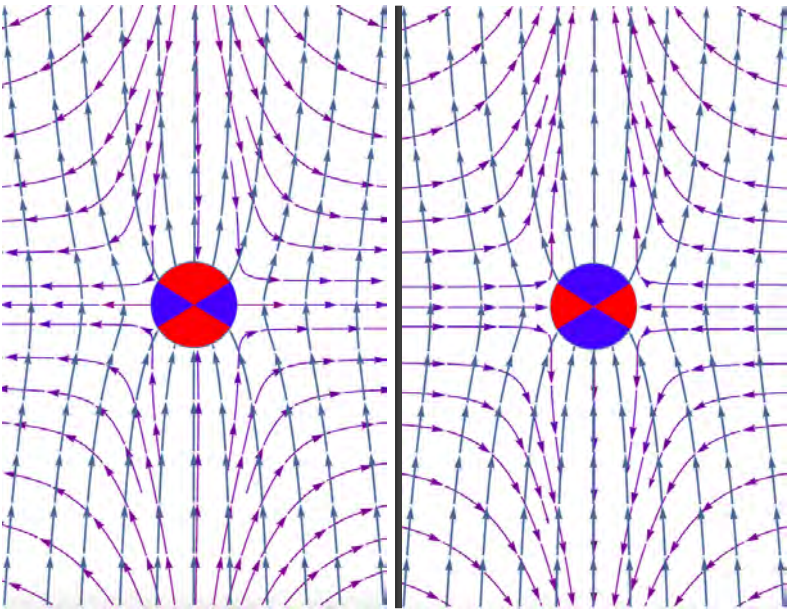


Fig. 3. The electromagnetic field lines of Wald solution. The blue lines indicate the magnetic field lines, and the purple lines indicate the electric field lines. **left:**. The magnetic field is parallel to the spin of the Kerr BH, so parallel to the rotation axis. The electric field lines are inward for the polar angle $\theta <\sim \pi/3$, so the electrons are accelerated away from the BH. For $\theta > \pi/3$, the electric field lines are outward, so the protons are accelerated away from the BH. **right:** The magnetic field is antiparallel to the rotational axis of the Kerr BH. The electric field lines are outward for the polar angle $\theta <\sim \pi/3$, so the protons will be accelerated away from the BH. The electric field lines are inward for $\theta > \pi/3$, so the electrons will be accelerated away from the BH. This figure is quoted from Ruffini et al. [32].

matter ($\sim 10^{-14}$ g cm$^{-3}$) of the SN ejecta [22]. For an aligned magnetic field with the angular momentum of the Kerr BH, an electric field is induced by gravitomagnetic interaction as described by the Wald solution [35]; see figure 3. The medium around the BH provides a sufficient amount of ionized particles that are accelerated to ultra-relativistic energies by the induced electric field, thereby emitting synchrotron and curvature radiation at expenses of the BH rotation energy.

The synchrotron radiation emitted from the accelerated charged particles has been calculated for different polar angles, see figure 4. Along the polar axis, the electric and magnetic fields are aligned, so there are no radiation losses and electrons can reach energies as large as $\sim 10^{18}$ eV, becoming a source of ultrahigh-energy cosmic rays (UHECRs). At larger angles, where electrons propagate across the magnetic lines producing synchrotron photons in the GeV energy domain.

The parameters of the inner engine, namely the BH mass and spin, and the surrounding magnetic field strength have been inferred from the following conditions: 1) the Kerr BH extractable energy accounts for the observed GeV radiation energetics in BdHN I, 2) the synchrotron radiation luminosity explains the observed

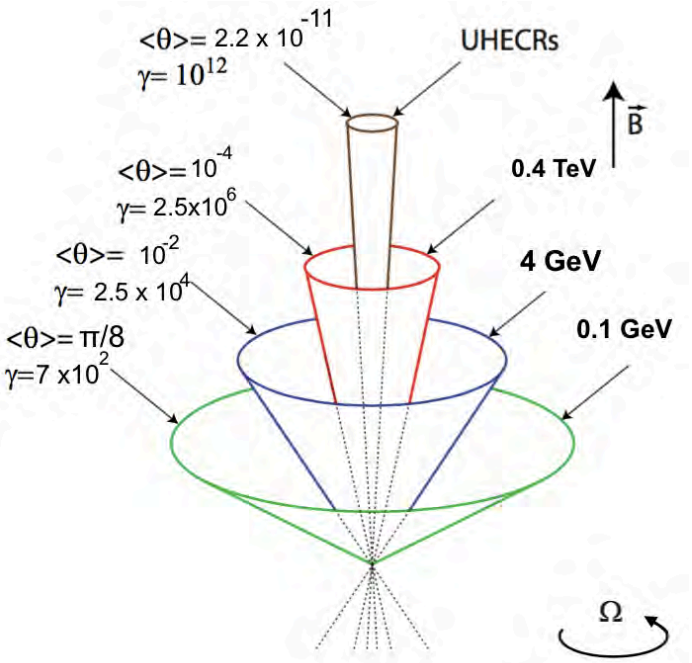



Fig. 4. The radiation emitted by the synchrotron emission of accelerated electrons in the different bands (0.1 GeV to 0.4 TeV) from different angles. And in the narrow polar cone, The UHECRs are produced. Arrows indicate the BH rotational direction and the external magnetic field direction. This figure is quoted from Ruffini et al. [32].

GeV luminosity, and 3) the emitted GeV photons can indeed scape from the system without suffering from magnetic pair production. The case of GRB 130427A has been analyzed in Ruffini et al. [32] and GRB 190114C in Moradi et al. [14]. For instance, in GRB 190114C the accelerated electrons radiate $1.8 \times 10^{53}$ erg in GeV photons via the synchrotron mechanism, and this procedure lasts for years following a power-law decay of light-curve with power-law index $-1.2$.

The mass and spin of the BHs came precisely what expected from the gravitational collapse of the fast rotating NS by accretion, and the strength of the magnetic field surrounding the BH turns out to be a few $10^{10}$ G. Therefore, the intensity of the induced accelerating electric field is undercritical in this BdHN I episode. The above magnetic field strength lower than expected to be inherited from the NS, and that could be the result from a screening process during the GRB prompt emission by electron-positron pairs; see the next section and Campion et al. [5] for details on this interesting physical process. In a recent comprehensive analysis of all the up-to-know identified BdHN I and the GeV emission observed in some of them, it has been inferred that the GeV emission must be emitted within an angle of 60° from the BH rotation axis; see Ruffini et al. [34] for details. This result is in agreement with the theoretical expectation; see Moradi et al. [14].

Another discovery of this inner engine model is that this energetic emission from GRB is not continuous but proceeds in a repetitive sequence of discrete impulse events. Since the inner engine repeats the procedure of charge (BH spin and magnetic field induce electric field) and discharge (electric field accelerates the charged particle that escape from the system), the repetition time grows slowly along with the loss of BH rotational energy. Along the emission of these discrete events the magnetic field keeps constant, but the BH spin decreases after each event by a well defined amount given by the concept of *blackholic quantum* described in Rueda & Ruffini [16]. The blackholic quanta explaining the GeV emission are characterized by an energy $\Delta E_q \sim 10^{38}$ erg, emitted over a timescale $\tau_q \sim 10^{-15}$ s. The fraction of BH angular momentum extracted after each event is $\Delta J_q/J \sim 10^{-16}$, i.e. $\Delta J_q \sim 10^{33}$ g cm$^2$ s$^{-1}$ [14], where $J$ is the Kerr BH angular momentum. This result is indeed unexpected, and it seems to be a general property not only of GRBs but also of the supermassive Kerr BHs in active galactic nuclei; see e.g. Moradi et al. [14] for the analysis of M87*.

## 4. The Prompt and the Afterglow Emission

This section gives the details of the propagation and the radiation of the relativistic outflow, based on the theoretical articles of Ruffini et al. [21,23,24,25] and the observational articles of Rueda et al. [17], Ruffini et al. [26,30,31], Wang et al. [36,37,38].

### 4.1. *The Ultrarelativistic Prompt Emission (UPE) Phase*

Moradi et al. [13] investigated the ultrarelativistic prompt emission (UPE) phase of the BdHN I in which the electric field is overcritical, generating an optically thick electron-positron plasma by vacuum polarization. The plasma expands and self-accelerates to ultra-relativistic by converting its internal energy. Eventually, it reaches the transparency point and releases photons in the MeV energy domain. Campion et al. [5] studied the channel of producing electron-positron pairs via high energy photons interacting with magnetic fields, the motion of these pairs generates a current which induces another dominant magnetic field that screens the original one.

We now focus on the relativistic electron-position plasma that leads to the UPE phase. The hydrodynamics of this plasma have been formulated and simulated in the articles that established the *fireshell* model, and which have been adopted by the BdHN model. Ruffini et al. [23] considered a Reissner-Nordstrom electromagnetic BH generates the electron-positron pairs. The system is expected to be thermalized to a plasma configuration due to the huge pair density and cross-section of the $e^+ + e^- \rightarrow \gamma + \gamma$ process. The evolution of plasma is governed by the hydrodynamic equations including, the conservation of energy-momentum, the conservation of the baryon number, the rate equation for electron-positron annihilation, and the equation of state. By integrating the equations numerically under the Reissner-Nordström metric and compared with the analytical analysis, the temperature drops

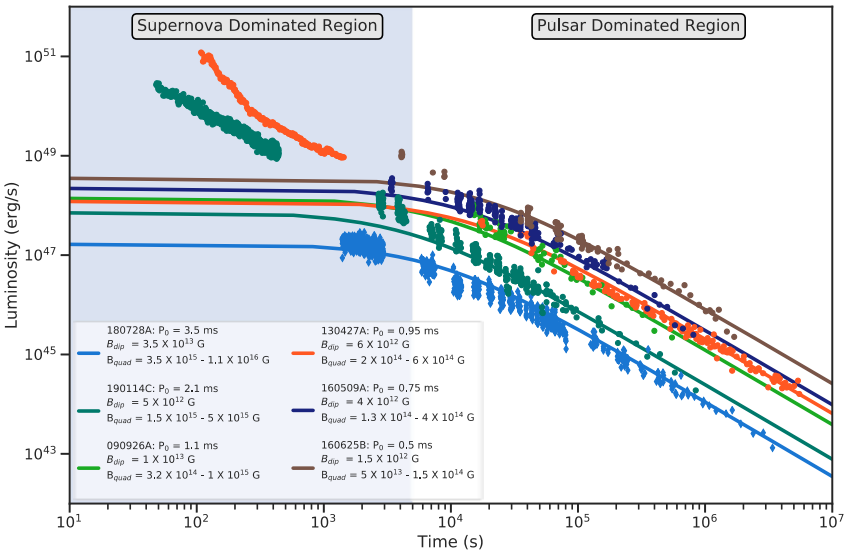



Fig. 5. The brown, dark blue, orange, green, and bright blue dots correspond to the bolometric light curves of GRB 160625B, 160509A, 130427A, 190114C, and 180728A, respectively. The thick lines are energy injection from the NS spin-down, which energizes the late-time afterglow (white background), while in the early time (blue background) the remaining kinetic energy of the outermost shell of the SN ejecta plays a dominant role. This figure is quoted from Rueda et al. [17].

as the internal energy are converted to kinetic energy. The plasma Lorentz factor is accelerated to several hundred, and the radiation releases when reaching the transparency radius. Ruffini et al. [24] adopted a more realistic baryonic environment where the plasma propagates. Baryons are incorporated into the hydrodynamic equations. It is clear by solving the equations that the Lorentz factor of plasma keeps increasing in the beginning until engulfing the baryons, a drop of Lorentz factor occurs, then it goes up again and finally reaches saturation. Ruffini et al. [25] systematically reviewed the creation and annihilation of the electron-positron pairs, thermalization, oscillation and their applications in GRB observations. The plasma penetrates the low-density region of the SN ejecta, see Becerra et al. [1] and Ruffini et al. [30] for details, then propagates in the ISM, and radiates, accounting for the prompt emission. For the observations, the light-curve and spectrum of the prompt emission have been successfully fitted by solving the hydrodynamic equations plus a density profile of the circumburst medium. Izzo et al. [10] offers an example of GRB 090618, of which the system starts by $\sim 2.5 \times 10^{53}$ erg electron-position plasma. The plasma propagates in the circumburst density of 0.6 cm$^{-3}$, and collides with dense clouds of mass $\sim 10^{24}$ g at the distance of $10^{15}$ cm to $10^{16}$ cm. The plasma finally self-accelerates up to transparency reaching a Lorentz factor $\sim 500$, thereby producing the observed emission.

We refer the reader to the most recent analysis of the UPE phase in BdHN I presented in Moradi et al. [13], where the physical origin of the UPE phase has been scrutinized taking as a proxy GRB 190114C. The UPE phase of GRB 190114C is observed in the rest-frame time interval $t_{\rm rf} = 1.9$–$3.99$ s, by the *Fermi*-GBM in 10 keV–10 MeV energy band. Thanks to the high signaltonoise ratio of *Fermi*-GBM data, a time-resolved spectral analysis of the UPE emission has evidenced a sequence of similar blackbody plus cut-off power-law spectra (BB+CPL), on ever decreasing time intervals. In it, the inner engine operates in an overcritical electric field regime. The electron-positron pair electromagnetic plasma in presence of a baryon load, *a PEMB pulse*, is therein originated from a vacuum polarization quantum process. This initially optically thick plasma self-accelerates, giving rise at the transparency radius to the MeV radiation observed by *Fermi*-GBM. For the first time, it has been quantitatively shown how the inner engine, by extracting the rotational energy of the Kerr BH, produces a series of PEMB pulses. Therefore, a quantum vacuum polarization process sequences with decreasing time bins occurs. We compute the Lorentz factors, the baryon loads and the radii at transparency, as well as the value of the magnetic field in each sequence. It has been therefore found there is an underlying fundamental hierarchical structure, linking the quantum electrodynamics regime of the UPE to the classical electrodynamics regime of the GeV emission after the UPE. The PEMB pulses of the UPE have been found to be characterized by the emission of blackholic quanta of energy $\sim 10^{45}$ erg, over a timescale $\sim 10^{-9}$ s.

Let us summarize GRB 190114C. The initial magnetic field left over by the collapse of the accreting NS and rooted in the surrounding material is very strong ($\sim 10^{14}$ G), so it induces a sizeable electric field that surpasses the critical value near the horizon. The overcritical electric field transfers its energy to the electron-positron pairs by the vacuum polarization and is later emitted as the UPE phase of $2.5 \times 10^{53}$ erg. The magnetic field becomes then screened to a few $10^{10}$ G in a few seconds [5], consequently the size of the region above the BH horizon with overcritical electric field (the *dyadoregion*) shrinks, and its energy stored becomes insignificant. This marks the end of the UPE phase and after it, the above inner engine mechanism by which the induced electric field accelerate electrons within a few horizon radii becomes the main channel of taking away the BH rotation energy in form of GeV photons [14].

### 4.2. *The Afterglow Emission Phase*

Another part of the plasma hindered by the SN ejecta accelerates the SN outermost shell to mildly-relativistic velocities. The breakout of the plasma (shockwave) from the outermost shell at $\sim 10^2$ s radiate photons of keV energies which explain the observed X-ray flares [30]. The synchrotron emission in the outermost shell accounts for the early afterglow X-ray emission [26]. Rotational energy from the $\nu$NS rotational is injected into the SN ejecta, then radiated by the synchrotron emission, accounts

for the plateau and late-time ($\sim 10^4$ s) afterglow[17,21,37]. The emission of the $\nu$NS as a pulsar becomes directly observable when the synchrotron luminosity fades off below the pulsar radiation luminosity. About $\sim 15$ days (rest-frame time) after the SN explosion that triggered the BdHN, the optical emission from the nickel decay in the SN ejecta reaches the maximum, there may appear a bump on the optical light-curve[26,37,38].

Ruffini et al.[30], Wang et al.[36] statistically analyzed the X-ray flares observed in the early afterglow. A general pattern of thermal component of temperature $\sim 1$ keV was found, suggesting that the flare is generated from a mildly-relativistic expanding shell of Lorentz factor $< 4$ at a distance $\sim 10^{12}$ cm. The observation of flares is consistent with our simulation of $\sim 10^{53}$ erg of plasma impacts on the SN ejecta of a few solar masses, leading to the formation of a shock propagating inside the SN ejecta until reaching the outermost shell. The density profile and velocity profile of the accelerated ejecta are obtained. Precisely, the shockwave breaks out at $\sim 10^{12}$ cm and the outermost shell is accelerated to Lorentz factor $\sim 2$–$5$. This feature was also extensively studied for GRB 151027A[31]. Along with the conversion of the kinetic energy of the outermost shell into radiation, the early afterglow exhibits a steep decay behaviour. Then, the energy injected from the $\nu$NS dominates the afterglow, the light-curve shows a plateau followed by a normal power-law decay, shown in figure 5. Taking GRB 180728A as an example[37], from the conversion of angular momentum, the $\mathrm{CO_{core}}$ collapses to a fast spinning NS of initial spin period $\sim 3$ ms. Such a newborn NS allows the presence of multipolar magnetic fields, with a quadrupole magnetic field $\sim 10^{15}$ G and a dipole field $\sim 10^{13}$ G, the spin-down of the $\nu$NS injects energy into the outflowing ejecta whose synchrotron emission fits the late-time X-ray afterglow. A comprehensive analysis of the afterglow of a few long GRB afterglows within the above synchrotron mechanism of the BdHN model has been presented in Rueda et al.[17].

## 5. Conclusion

We can draw some general conclusions with the aid of the BdHN evolution shown in figure 6. (a) Our picture starts with a binary system consisting of two main-sequence stars of intermediate mass, say $15M_\odot$ and $12M_\odot$, respectively. (b) At a given time, the more massive star undergoes a core-collapse SN and forms an NS. (c) The system enters the X-ray binary phase. (d) The system has overcome binary interactions and common-envelope phases (not shown in the diagram) which have led to the hydrogen and helium envelopes of the ordinary star to have been stripped off, remaining a star which is rich in carbon and oxygen, referred to as $\mathrm{CO_{core}}$. At this stage, the system is a $\mathrm{CO_{core}}$-NS binary, which is considered as the initial configuration of the BdHN model. (e) At this stage the orbit of the binary has shrink to a period of the order of a few minutes. The $\mathrm{CO_{core}}$ explodes into an SN (of type Ic in view of the absence of hydrogen and helium in its outermost layers), expelling several solar masses. These ejecta begin to expand, and a rapidly rotating $\nu$NS is

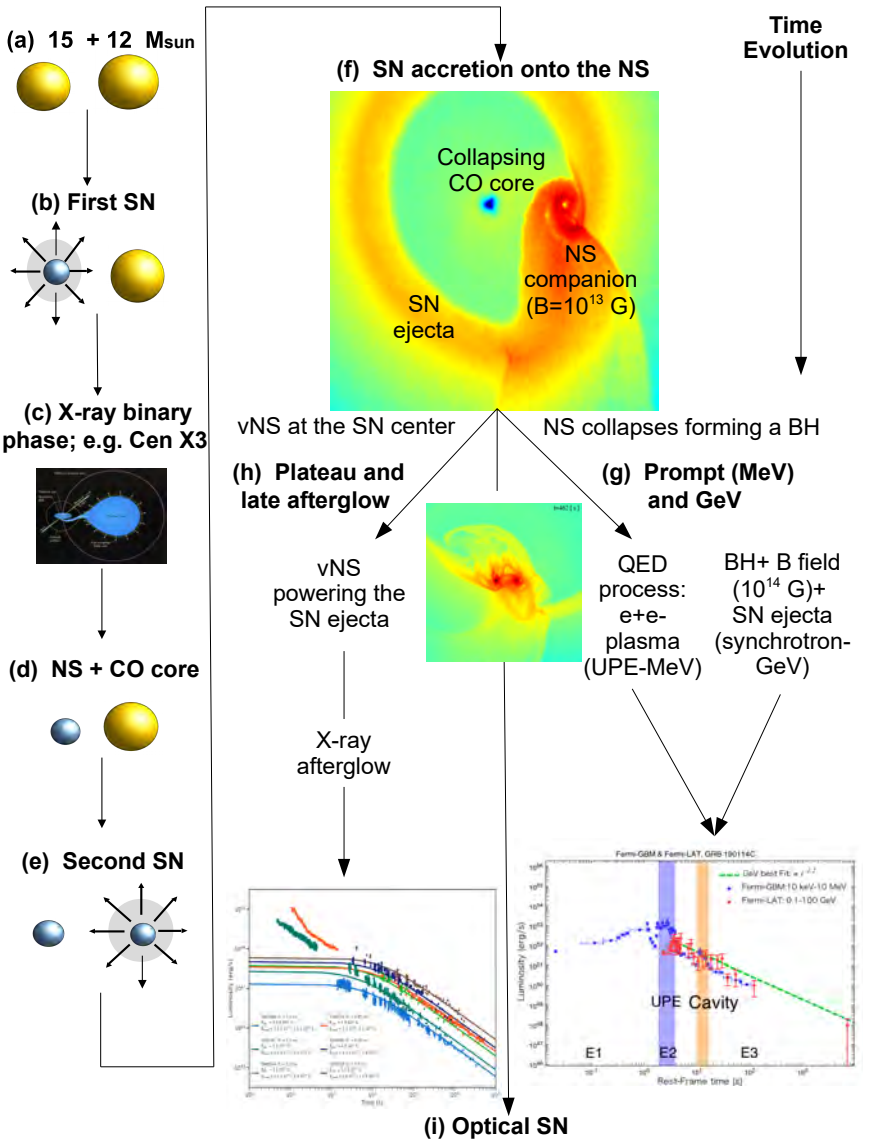



Fig. 6. Diagram of the evolutionary path of BdHN. Including binary evolution, SN explosion, NS accretion, BH formation, GRB prompt and afterglow emissions and SN appearence. This figure is quoted from [17].

left in the centre. (f) Depending on the initial NS mass and binary separation, the SN ejecta accretes onto the NS companion and onto the $\nu$NS, forming a massive NS (BdHN II) or BH (BdHN I; this example). At this stage, the system is a new NS and BH binary surrounded by the expanding ejecta. (g) The inner engine composed by the newborn BH, the surrounding magnetic field and ionized plasma is formed, and

its activity explains the GRB UPE phase and the subsequent GeV emission. The magnetic field in the inner engine at BH formation is overcritical, so it induces (by gravitomagnetic interaction with the BH spin) an overcritical electric field, so the UPE phase operates in an overcritical regime. A quantum electrodynamical process of vacuum polarization takes place leading to an electron-positron pair plasma pulses (PEMB pulses) that expand to ultrarelativistic velocity reaching transparency with a Lorentz factor of up to hundreds, and emitting MeV photons. The magnetic field is then screened to undercritical values by currents produced by the motion of the electron and positrons, and the inner engine classical electrodynamical process of particle acceleration emitting GeV photons by synchrotron radiation becomes the relevant process of emission. (h) The spin-down energy of the $\nu$NS injects energy into the expanding SN ejecta emitting the observed X-ray afterglow by synchrotron radiation. (i) The appearance of the energy release owing to nickel decay in the SN ejecta is observed at optical wavelengths.

## Acknowledgments

We acknowledge the contribution from the ICRANet group and the collaborators of establishing the BdHN model. Remo Ruffini, Jorge Rueda and Yu Wang gave the plenary talks in the MG 15 meeting on behalf of the contributors.

## References

1. Becerra, L., Bianco, C. L., Fryer, C. L., Rueda, J. A., & Ruffini, R. 2016, ApJ, 833, 107
2. Becerra, L., Cipolletta, F., Fryer, C. L., Rueda, J. A., & Ruffini, R. 2015, ApJ, 812, 100
3. Becerra, L., Ellinger, C. L., Fryer, C. L., Rueda, J. A., & Ruffini, R. 2019, ApJ, 871, 14
4. Becerra, L., Guzzo, M. M., Rossi-Torres, F., et al. 2018, ApJ, 852, 120
5. Campion, S., Rueda, J., Ruffini, R., & Xue, S. 2021, Physics Letters B, 820, 136562. https://www.sciencedirect.com/science/article/pii/S0370269321005025
6. Cipolletta, F., Cherubini, C., Filippi, S., Rueda, J. A., & Ruffini, R. 2017, PRD, 96, 024046
7. Fryer, C. L., Oliveira, F. G., Rueda, J. A., & Ruffini, R. 2015, Physical Review Letters, 115, 231102
8. Fryer, C. L., Rueda, J. A., & Ruffini, R. 2014, ApJ, 793, L36
9. Izzo, L., Rueda, J. A., & Ruffini, R. 2012, A&A, 548, L5
10. Izzo, L., Ruffini, R., Penacchioni, A. V., et al. 2012, A&A, 543, A10
11. Liang, L., Ruffini, R., Rueda, J. A., et al. 2019, arXiv e-prints, arXiv:1910.12615
12. Moradi, R., Li, L., Rueda, J. A., et al. 2021, arXiv e-prints, arXiv:2103.09158
13. Moradi, R., Rueda, J. A., Li, L., Ruffini, R., & Wang, Y. 2021, Submitted to ApJ

14. Moradi, R., Rueda, J. A., Ruffini, R., & Wang, Y. 2021, A&A, 649, A75
15. Rueda, J. A., & Ruffini, R. 2012, ApJ, 758, L7
16. —. 2020, European Physical Journal C, 80, 300
17. Rueda, J. A., Ruffini, R., Karlica, M., Moradi, R., & Wang, Y. 2020, ApJ, 893, 148
18. Rueda, J. A., Aimuratov, Y., de Almeida, U. B., et al. 2017, International Journal of Modern Physics D, 26, 1730016
19. Rueda, J. A., Ruffini, R., Wang, Y., et al. 2018, JCAP, 2018, 006
20. —. 2019, JCAP, 2019, 044
21. Ruffini, R., Karlica, M., Sahakyan, N., et al. 2018, ApJ, 869, 101
22. Ruffini, R., Melon Fuksman, J. D., & Vereshchagin, G. V. 2019, ApJ, 883, 191
23. Ruffini, R., Salmonson, J. D., Wilson, J. R., & Xue, S. S. 1999, A&A, 350, 334
24. Ruffini, R., Salmonson, J. D., Wilson, J. R., & Xue, S.-S. 2000, A&A, 359, 855
25. Ruffini, R., Vereshchagin, G., & Xue, S. 2010, Phys. Rep., 487, 1
26. Ruffini, R., Wang, Y., Enderli, M., et al. 2015, ApJ, 798, 10
27. Ruffini, R., Rueda, J. A., Muccino, M., et al. 2016, ApJ, 832, 136
28. Ruffini, R., Rodriguez, J., Muccino, M., et al. 2016, ArXiv:1602.03545, arXiv:1602.03545
29. Ruffini, R., Rueda, J. A., Moradi, R., et al. 2018, arXiv e-prints, arXiv:1811.01839
30. Ruffini, R., Wang, Y., Aimuratov, Y., et al. 2018, ApJ, 852, 53
31. Ruffini, R., Becerra, L., Bianco, C. L., et al. 2018, ApJ, 869, 151
32. Ruffini, R., Moradi, R., Rueda, J. A., et al. 2019, ApJ, 886, 82
33. Ruffini, R., Li, L., Moradi, R., et al. 2019, arXiv e-prints, arXiv:1904.04162
34. Ruffini, R., Moradi, R., Rueda, J. A., et al. 2021, MNRAS, 504, 5301
35. Wald, R. M. 1974, Phys. Rev., D10, 1680
36. Wang, Y., Aimuratov, Y., Moradi, R., et al. 2018, Mem. Soc. Astron. Ital., 89, 293
37. Wang, Y., Rueda, J. A., Ruffini, R., et al. 2019, ApJ, 874, 39
38. Wang, Y., Ruffini, R., Kovacevic, M., et al. 2015, Astronomy Reports, 59, 667

# IceCube: Opening a new window on the universe from the South Pole

Francis Halzen

*Wisconsin IceCube Particle Physics Center*
*Madison, WI 53703 USA*
* *E-mail: halzen@icecube.wisc.edu*

After updating the status of the measurements of the cosmic neutrino flux by the IceCube experiment, we summarize the observations of the first identified source of cosmic rays and speculate on the connection between the two observations.

*Keywords*: multimessenger, neutrinos, cosmic rays, blazars, cosmic accelerators

## 1. Detecting cosmic neutrinos

Cosmic rays have been studied for more than a century. They reach energies in excess of $10^8$ TeV, populating an extreme universe that is opaque to electromagnetic radiation[1,2]. We don't yet know where or how particles are accelerated to these extreme energies, but with the observation of a distant blazar in coincidence with the direction and time of a very high energy muon neutrino, neutrino astronomy has made a breakthrough in resolving this puzzle[3,4]. The rationale for finding cosmic ray sources by observing neutrinos is straightforward: near neutron stars and black holes, gravitational energy released in the accretion of matter can power the acceleration of protons ($p$) or heavier nuclei that subsequently interact with gas ("$pp$") or ambient radiation ("$p\gamma$") to produce neutrinos originating from the decay of pions and other secondary particles. In the case of photoproduction, both neutral and charged pion secondaries are produced in the processes $p + \gamma_{\rm bg} \to p + \pi^0$ and $p + \gamma_{\rm bg} \to n + \pi^+$, for instance. While neutral pions decay as $\pi^0 \to \gamma + \gamma$ and create a flux of high-energy gamma rays, the charged pions decay into three high-energy neutrinos ($\nu$) and anti-neutrinos ($\bar{\nu}$) via the decay chain $\pi^+ \to \mu^+ + \nu_\mu$ followed by $\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e$, and the charged-conjugate process. We refer to these photons as pionic photons. They provide the rational for multimessenger astronomy and should be distinguished from photons radiated by electrons that may be accelerated along with the cosmic rays.

High-energy neutrinos interact predominantly with matter via deep inelastic scattering off nucleons: the neutrino scatters off quarks in the target nucleus by the exchange of a $Z$ or $W$ weak boson, referred to as *neutral current* and *charged current* interactions, respectively. Whereas the neutral current interaction leaves the neutrino state intact, in a charged current interaction a charged lepton is produced that shares the initial neutrino flavor. The average relative energy fraction transferred from the neutrino to the lepton is at the level of 80% at high energies. The struck nucleus does not remain intact and its high-energy fragments typically initiate hadronic showers in the target medium.

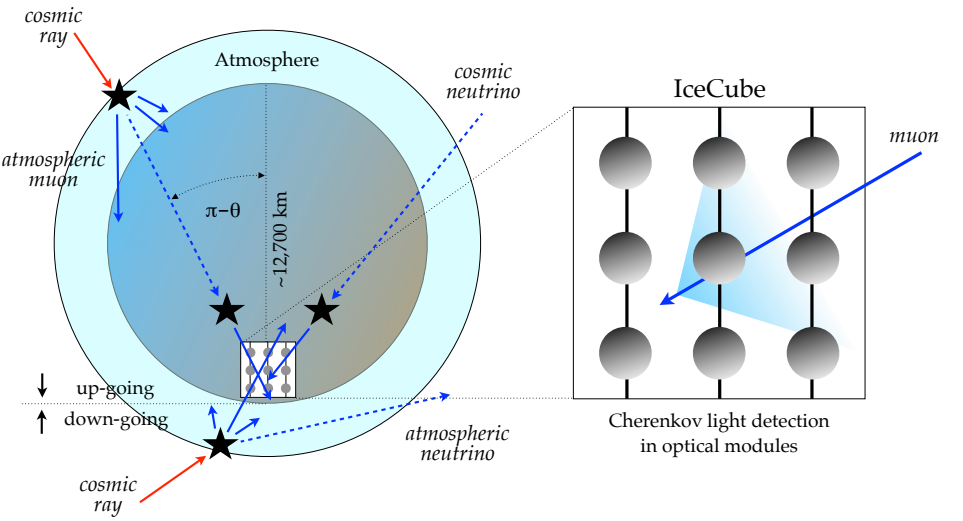



Fig. 1. The principal idea of neutrino telescopes from the point of view of IceCube located at the South Pole. Neutrinos dominantly interact with a nucleus in a transparent medium like water or ice and produce a muon that is detected by the wake of Cherenkov photons it leaves inside the detector. The background of high-energy muons (solid blue arrows) produced in the atmosphere can be reduced by placing the detector underground. The surviving fraction of muons is further reduced by looking for upgoing muon tracks that originate from muon neutrinos (dashed blue arrows) interacting close to the detector. This still leaves the contribution of muons generated by atmospheric muon neutrino interactions. This contribution can be separated from the diffuse cosmic neutrino emission by an analysis of the combined neutrino spectrum.

Immense particle detectors are required to collect cosmic neutrinos in statistically significant numbers. Already by the 1970s, it had been understood[5] that a kilometer-scale detector was needed to observe the cosmogenic neutrinos produced in the interactions of cosmic rays with background microwave photons[6]. The IceCube project has transformed one cubic kilometer of natural Antarctic ice into a Cherenkov detector. Photomultipliers embedded in the ice transform the Cherenkov light radiated by secondary particles produced in neutrino interactions into electrical signals using the photoelectric effect; see Figs. 1 and 2. Computers at the surface use this information to reconstruct the light patterns produced to infer the arrival directions, energies and flavor of the neutrinos.

Two patterns of Cherenkov radiation are of special interest, “tracks” and “cascades.” The term “track” refers to the Cherenkov emission of a long-lived muon passing through the detector after production in a charged current interaction of a muon neutrino inside or in the vicinity of the detector. Because of the large background of muons produced by cosmic ray interactions in the atmosphere, the observation of muon neutrinos is limited to upgoing muon tracks that are produced

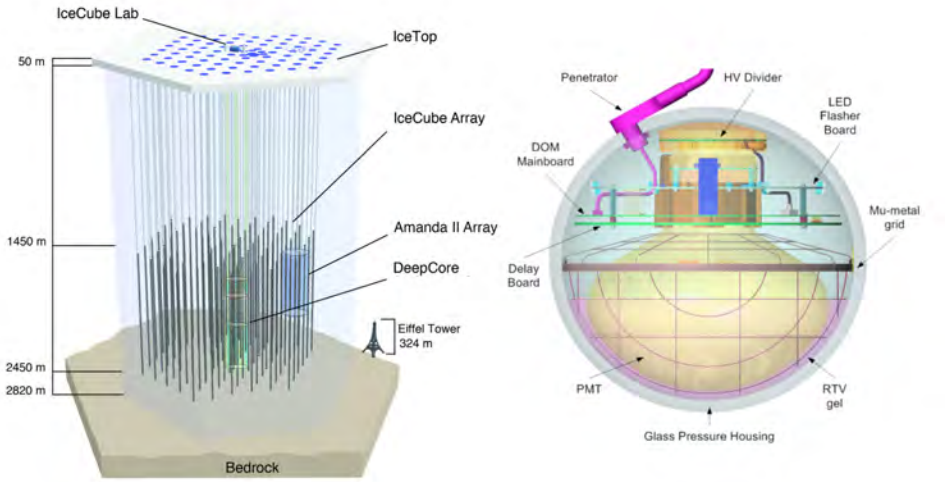



Fig. 2. Architecture of the IceCube observatory (left) and the schematics of a digital optical module (right) (see Ref. 7 for details).

by neutrinos that have passed through the Earth that acts as a neutrino filter. The remaining background consists of atmospheric neutrinos, which are indistinguishable from cosmic neutrinos on an event-by-event basis. However, the steeply falling spectrum ($\propto E^{-3.7}$) of atmospheric neutrinos allows identifying diffuse astrophysical neutrino emission above a few hundred TeV by a spectral analysis. The atmospheric background is also reduced for muon neutrino observation from point-like sources, in particular transient neutrino sources.

Energetic electrons and taus produced in interactions of electron and tau neutrinos, respectively, will initiate an electromagnetic shower that develops over less than 10 meters. This shower as well as the hadronic particle shower generated by the target struck by a neutrino in the ice radiate Cherenkov photons. The light pattern is mostly spherical and referred to as a "cascade." The direction of the initial neutrino can only be reconstructed from the Cherenkov emission of secondary particles produced close to the neutrino interaction point, and the angular resolution is worse than for track events. On the other hand, the energy of the initial neutrino can be constructed with a better resolution than for tracks.

Two methods are used to identify cosmic neutrinos. Traditionally, neutrino searches have focused on the observation of muon neutrinos that interact primarily outside the detector to produce kilometer-long muon tracks passing through the instrumented volume. Although this allows the identification of neutrinos that interact outside the detector, it is necessary to use the Earth as a filter in order to remove the background of cosmic-ray muons. This limits the neutrino view

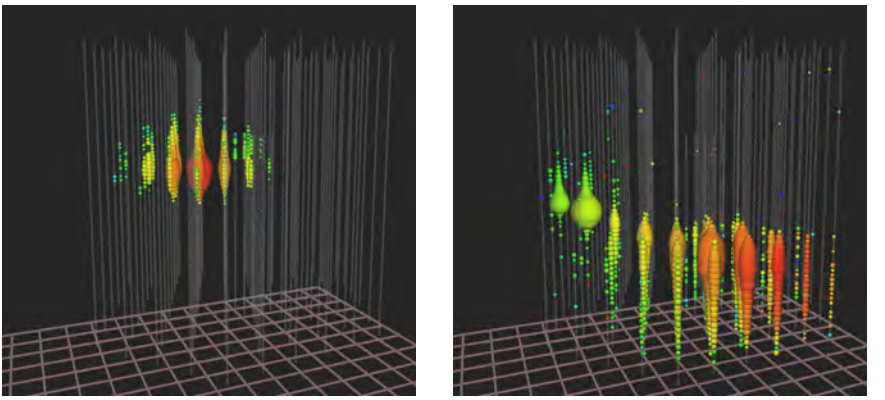


Fig. 3. **Left Panel:** Light pool produced in IceCube by a shower initiated by an electron or tau neutrino. The measured energy is 1.14 PeV, which represents a lower limit on the energy of the neutrino that initiated the shower. White dots represent sensors with no signal. For the colored dots, the color indicates arrival time, from red (early) to purple (late) following the rainbow, and size reflects the number of photons detected. **Right Panel:** An upgoing muon track traverses the detector at an angle of 11° below the horizon. The deposited energy, i.e., the energy equivalent of the total Cherenkov light of all charged secondary particles inside the detector, is 2.6 PeV.

to a single flavor and half the sky. An alternative method exclusively identifies high-energy neutrinos interacting inside the detector, so-called high-energy starting events (HESE). It divides the instrumented volume of ice into an outer veto shield and a $\sim$ 420-megaton inner fiducial volume. The advantage of focusing on neutrinos interacting inside the instrumented volume of ice is that the detector functions as a total absorption calorimeter, measuring the neutrino energy of cascades with a 10-15 % resolution[8]. Furthermore, with this method, neutrinos from all directions in the sky can be identified, including both muon tracks as well as secondary showers, produced by charged-current interactions of electron and tau neutrinos, and neutral current interactions of neutrinos of all flavors. For illustration, the Cherenkov patterns initiated by an electron (or tau) neutrino of about 1 PeV energy and a muon neutrino losing 2.6 PeV energy in the form of Cherenkov photons while traversing the detector are contrasted in Fig. 3.

In general, the arrival times of photons at the optical sensors, whose positions are known, determine the particle's trajectory, while the number of photons is a proxy for the deposited energy, i.e., the energy of all charged secondary particles from the interactions. For instance, for the cascade event shown in the left panel of Fig. 3, more than 300 digital optical modules (DOMs) report a total of more than 100,000 photoelectrons. The two abovementioned methods of separating neutrinos from the cosmic-ray muon background have complementary advantages. The long tracks produced by muon neutrinos can be pointed back to their sources with a $\leq 0.4°$

angular resolution. In contrast, the reconstruction of the direction of cascades in the HESE analysis, in principle possible to a few degrees, is still in the development stage in IceCube[8]. Their reconstruction, originally limited to within $10^\circ \sim 15^\circ$ of the direction of the incident neutrino, has now achieved resolutions closer to $5^\circ$[9]. Determining the deposited energy from the observed light pool is, however, relatively straightforward, and a resolution of better than 15 % is possible; the same value holds for the reconstruction of the energy deposited by a muon track inside the detector.

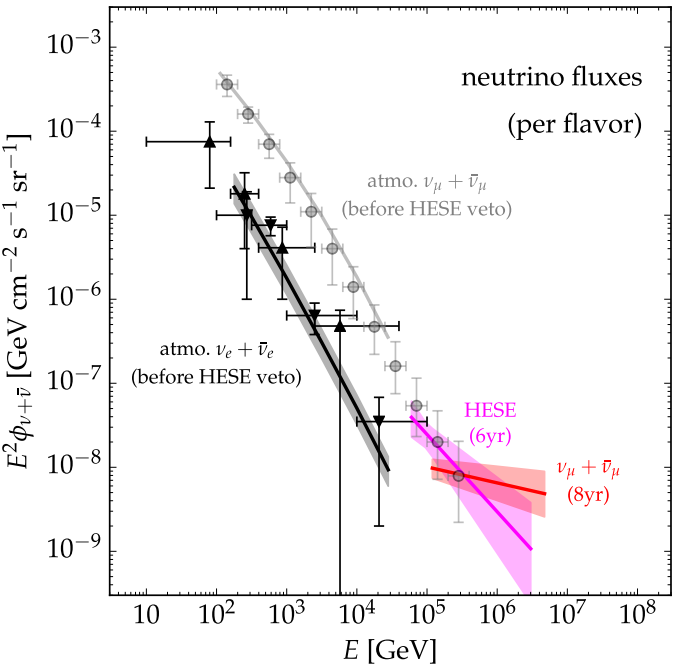



Fig. 4. Summary of diffuse neutrino observations (per flavor) by IceCube. The black and grey data show IceCube's measurement of the atmospheric $\nu_e + \bar{\nu}_e$ [10,11] and $\nu_\mu + \bar{\nu}_\mu$ [12] spectra. The magenta line and magenta-shaded area indicate the best-fit and $1\sigma$ uncertainty range of a power-law fit to the six-year HESE data. Note that the HESE analysis vetoes atmospheric neutrinos and can probe astrophysical neutrinos below the atmospheric neutrino flux. The corresponding fit to the eight-year $\nu_\mu + \bar{\nu}_\mu$ analysis is shown in red. Figure from Ref. 13.

Using the Earth as a filter, a flux of neutrinos has been identified that is predominantly of atmospheric origin. IceCube has measured this flux over three orders of magnitude in energy with a result that is consistent with theoretical calculations. However, with eight years of data, an excess of events is observed at energies beyond 100 TeV[14–16], which cannot be accommodated by the atmospheric flux; see Fig. 4. Allowing for large uncertainties on the extrapolation of the atmospheric

component to higher energy, the statistical significance of the excess astrophysical flux is $6.7\sigma$. While IceCube measures only the energy of the secondary muon inside the detector, from Standard Model physics we can infer the energy spectrum of the parent neutrinos. The cosmic neutrino flux is well described by a power law with a spectral index $\Gamma = 2.19 \pm 0.10$ and a normalization at 100 TeV neutrino energy of $(1.01^{+0.26}_{-0.23}) \times 10^{-18}\,\mathrm{GeV}^{-1}\mathrm{cm}^{-2}\mathrm{sr}^{-1}$ [16]. The neutrino energies contributing to this power-law fit cover the range from 119 TeV to 4.8 PeV.

Using only two years of data, it was the alternative HESE method, which selects neutrinos interacting inside the detector, that revealed the first evidence for cosmic neutrinos [17,18]. The segmentation of the detector into a surrounding veto and active signal region has been optimized to reduce the background of atmospheric muons and neutrinos to a handful of events per year, while keeping most of the cosmic signal. Neutrinos of atmospheric and cosmic origin can be separated not only by using their well-measured energy but also on the basis that background atmospheric neutrinos reaching us from the Southern Hemisphere can be removed because they are accompanied by particles produced in the same air shower where the neutrinos originate. A sample event with a light pool of roughly one hundred thousand photoelectrons extending over more than 500 meters is shown in the left panel of Fig. 3. With PeV energy, and no trace of accompanying muons from an atmospheric shower, these events are highly unlikely to be of atmospheric origin. The six-year data set contains a total of 82 neutrino events with deposited energies ranging from 60 TeV to 10 PeV. The data are consistent with an astrophysical component with a spectrum close to $E^{-2}$ above an energy of $\sim 200$ TeV. In summary, IceCube has observed cosmic neutrinos using both methods for rejecting background. Based on different methods for reconstruction and energy measurement, their results agree, pointing at extragalactic sources whose flux has equilibrated in the three flavors after propagation over cosmic distances [19] with $\nu_e : \nu_\mu : \nu_\tau \sim 1 : 1 : 1$.

An extrapolation of this high-energy flux to lower energy suggests an interesting excess of events in the $30 - 100$ TeV energy range over and above a single power-law fit; see Fig. 5. This conclusion is supported by a subsequent analysis that has lowered the threshold of the starting-event analysis [20] and by a variety of other analyses. The astrophysical flux measured by IceCube is not featureless; either the spectrum of cosmic accelerators cannot be described by a single power law or a second component of cosmic neutrino sources emerges in the spectrum. Because of the self-veto of atmospheric neutrinos in the HESE analysis, i.e., the veto triggered by accompanying atmospheric muons, it is very difficult to accommodate the component below 100 TeV as a feature in the atmospheric background.

In Figure 5 we show the arrival directions of the most energetic events in the eight-year upgoing $\nu_\mu + \bar{\nu}_\mu$ analysis ($\odot$) and the six-year HESE data sets. The HESE data are separated into tracks ($\otimes$) and cascades ($\oplus$). The median angular resolution of the cascade events is indicated by thin circles around the best-fit position.

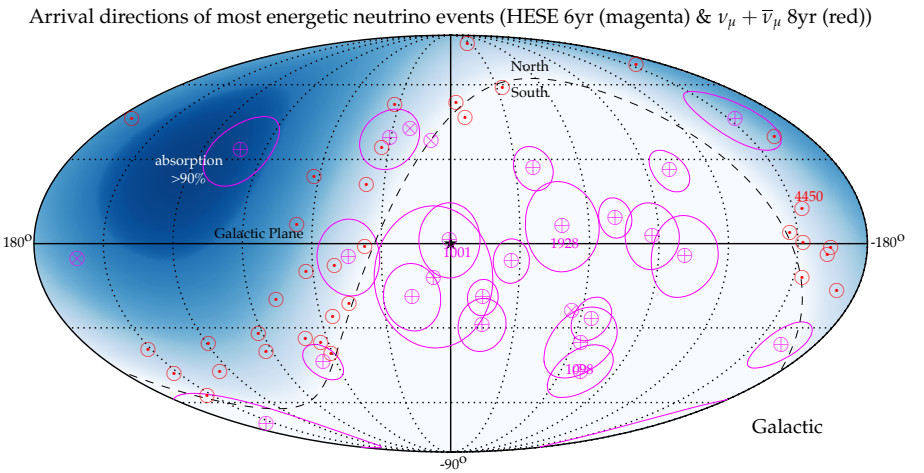



Fig. 5. Mollweide projection in Galactic coordinates of the arrival direction of neutrino events. We show the results of the eight-year upgoing track analysis [16] with reconstructed muon energy $E_\mu \gtrsim 200$ TeV ($\odot$). The events of the six-year high-energy starting event (HESE) analysis with deposited energy larger than 100 TeV (tracks $\otimes$ and cascades $\oplus$) are also shown [16,21,22]. The thin circles indicate the median angular resolution of the cascade events ($\oplus$). The blue-shaded region indicates the zenith-dependent range where Earth absorption of 100 TeV neutrinos becomes important, reaching more than 90% close to the nadir. The dashed line indicates the horizon and the star ($\star$) the Galactic Center. We highlight the four most energetic events in both analyses by their deposited energy (magenta numbers) and reconstructed muon energy (red number). Figure from Ref. 13.

The most energetic muons with energy $E_\mu > 200$ TeV in the upgoing $\nu_\mu + \bar{\nu}_\mu$ data set accumulate near the horizon in the Northern Hemisphere. Elsewhere, muon neutrinos are increasingly absorbed in the Earth causing the apparent anisotropy of the events in the Northern Hemisphere. Also HESE events with deposited energy of $E_{\rm dep} > 100$ TeV suffer from absorption in the Earth and are therefore mostly detected when originating in the Southern Hemisphere. After correcting for absorption, the arrival directions of cosmic neutrinos are isotropic, suggesting extragalactic sources. In fact, no correlation of the arrival directions of the highest energy events, shown in Fig. 5, with potential sources or source classes has reached the level of $3\sigma$ [20].

## 2. IceCube neutrinos and Fermi photons

The most important message emerging from the IceCube measurements of the high-energy cosmic neutrino flux is not apparent yet: the prominent and surprisingly important role of protons relative to electrons in the nonthermal universe. Photons are produced in association with neutrinos when accelerated cosmic rays produce neutral and charged pions in interactions with target photons or nuclei in the vicinity

of the accelerator. Targets include strong radiation fields that may be associated with the accelerator as well as concentrations of matter, such as molecular clouds in their vicinity. Additionally, pions can be produced in the interaction of cosmic rays with the extragalactic background light (EBL) when propagating through the interstellar or intergalactic background. As already discussed in section 1, a high-energy flux of neutrinos is produced in the subsequent decay of charged pions via $\pi^+ \to \mu^+ + \nu_\mu$ followed by $\mu^+ \to e^+ + \nu_e + \bar{\nu}_\mu$ and the charge–conjugate processes. High-energy gamma rays result from the decay of neutral pions, $\pi^0 \to \gamma + \gamma$. Pionic gamma rays and neutrinos carry, on average, 1/2 and 1/4 of the energy of the parent pion, respectively. With these approximations, the neutrino production rate $Q_{\nu_\alpha}$ (units of $\mathrm{GeV}^{-1}\mathrm{s}^{-1}$) can be related to the one for charged pions as

$$\sum_\alpha E_\nu Q_{\nu_\alpha}(E_\nu) \simeq 3\,[E_\pi Q_{\pi^\pm}(E_\pi)]_{E_\pi \simeq 4E_\nu} \ . \tag{1}$$

Similarly, the production rate of pionic gamma-rays is related to the one for neutral pions as

$$E_\gamma Q_\gamma(E_\gamma) \simeq 2\,[E_\pi Q_{\pi^0}(E_\pi)]_{E_\pi \simeq 2E_\gamma} \ . \tag{2}$$

Note, that the relative production rates of pionic gamma rays and neutrinos only depend on the ratio of charged-to-neutral pions produced in cosmic-ray interactions, denoted by $K_\pi = N_{\pi^\pm}/N_{\pi^0}$. Pion production by cosmic rays in interactions with photons can proceed resonantly in the processes $p + \gamma \to \Delta^+ \to \pi^0 + p$ and $p + \gamma \to \Delta^+ \to \pi^+ + n$. These channels produce charged and neutral pions with probabilities 2/3 and 1/3, respectively. However, the additional contribution of nonresonant pion production changes this ratio to approximately 1/2 and 1/2. In contrast, cosmic rays interacting with matter, e.g., hydrogen in the Galactic disk, produce equal numbers of pions of all three charges: $p + p \to N_\pi\,[\,\pi^0 + \pi^+ + \pi^-] + X$, where $N_\pi$ is the pion multiplicity. From above arguments we have $K_\pi \simeq 2$ for cosmic ray interactions with gas ($pp$) and $K_\pi \simeq 1$ for interactions with photons ($p\gamma$) [13].

With this approximation we can combine Eqs. (1) and (2) to derive a powerful relation between the pionic gamma-ray and neutrino production rates:

$$\frac{1}{3}\sum_\alpha E_\nu^2 Q_{\nu_\alpha}(E_\nu) \simeq \frac{K_\pi}{4}\left[E_\gamma^2 Q_\gamma(E_\gamma)\right]_{E_\gamma = 2E_\nu} \ . \tag{3}$$

The prefactor 1/4 accounts for the energy ratio $\langle E_\nu \rangle / \langle E_\gamma \rangle \simeq 1/2$ and the two gamma rays produced in the neutral pion decay.

Note that this relation relates pionic neutrinos and gamma rays without any reference to the cosmic ray beam; it simply reflects the fact that a $\pi^0$ produces two $\gamma$ rays for every charged pion producing a $\nu_\mu + \bar{\nu}_\mu$ pair, which cannot be separated by current experiments.

Before applying this relation to a cosmic accelerator, we have to take into account the fact that, unlike neutrinos, gamma rays interact with photons of the cosmic

microwave background before reaching Earth. The resulting electromagnetic shower subdivides the initial photon energy, resulting in multiple photons in the GeV-TeV energy range by the time the photons reach Earth. Calculating the cascaded gamma-ray flux accompanying IceCube neutrinos is straightforward [23,24].

As an illustration, an example of $\gamma$-ray and neutrino emission is shown as blue lines in Fig. 6 assuming that the underlying $\pi^0$ / $\pi^\pm$ production follows from cosmic-ray interactions with gas in the universe. In this way, the initial emission spectrum of $\gamma$-rays and neutrinos from pion decay is almost identical to the spectrum of cosmic rays (assumed to be a power law, $E^{-2.19}$ as is the case for the diffuse cosmic neutrino flux above and energy of 100 TeV), after accounting for the different normalizations and energy scales. The flux of neutrinos arriving at Earth (blue dashed line) follows this initial CR emission spectrum. However, the observable flux of $\gamma$-rays (blue solid lines) is strongly attenuated above 100 GeV by interactions with extragalactic background photons [13].

The overall normalization of the emission is chosen in a way that the model does not exceed the isotropic $\gamma$-ray background observed by the Fermi satellite (blue data). This implies an upper limit on the neutrino flux shown as the blue dashed line. Interestingly, the neutrino data shown in Fig. 6 saturates this limit above 100 TeV. Moreover, the HESE data that extends to lower energies is only marginally consistent with the upper bound implied by the model (blue dashed line). This example shows that multimessenger studies of $\gamma$-ray and neutrino data are powerful tools to study the neutrino production mechanism and to constrain neutrino source models [25].

The matching energy densities of the extragalactic gamma-ray flux detected by Fermi and the high-energy neutrino flux measured by IceCube suggest that, rather than detecting some exotic sources, it is more likely that IceCube to a large extent observes the same universe conventional astronomy does. Clearly, an extreme universe modeled exclusively on the basis of electromagnetic processes is no longer realistic. The finding implies that a large fraction, possibly most, of the energy in the nonthermal universe originates in hadronic processes, indicating a larger role than previously thought. The high intensity of the neutrino flux below 100 TeV in comparison to the Fermi data might indicate that these sources are even more efficient neutrino than gamma-ray sources [26,27].

Interestingly, the common energy density of photons and neutrinos is also comparable to that of the ultra-high-energy extragalactic cosmic rays (above $10^9$ GeV) observed, for instance, by the Auger observatory [28] (green data). Unless accidental, this indicates a common origin of the signal and illustrates the potential of multi-messenger studies. A scenario where the high-energy neutrinos observed at IceCube could actually originate in the same sources could be realized as follows: the cosmic ray sources can be embedded in environments that act as "storage rooms" for cosmic rays with energies far below the "ankle" ($E_{\rm CR} \ll 1$EeV). This energy-dependent

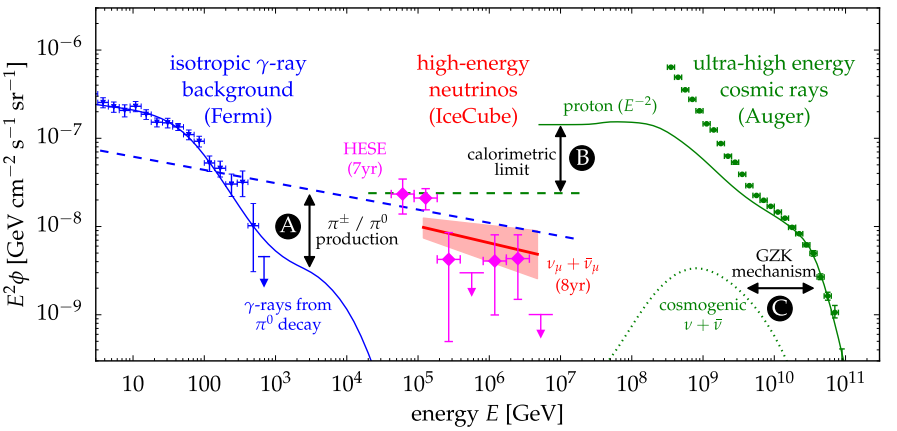



Fig. 6. The spectral flux ($\phi$) of neutrinos inferred from the eight-year upgoing track analysis (red fit) and the seven-year HESE analysis (magenta fit) compared to the flux of unresolved extragalactic γ-ray sources[31] (blue data) and ultra-high-energy cosmic rays[28] (green data). The neutrino spectra are indicated by the best-fit power-law (solid line) and $1\sigma$ uncertainty range (shaded range). We highlight the various multimessenger interfaces: **A:** The joined production of charged pions ($\pi^{\pm}$) and neutral pions ($\pi^0$) in cosmic-ray interactions leads to the emission of neutrinos (dashed blue) and γ-rays (solid blue), respectively. **B:** Cosmic ray emission models (solid green) of the most energetic cosmic rays imply a maximal flux (calorimetric limit) of neutrinos from the same sources (green dashed). **C:** The same cosmic ray model predicts the emission of cosmogenic neutrinos from the collision with cosmic background photons (GZK mechanism). Figure from Ref. 13.

trapping can be achieved via cosmic ray diffusion in magnetic fields. While these cosmic rays are trapped, they can produce γ-rays and neutrinos via collisions with gas. If the conditions are right, this mechanism can be so efficient that the total energy stored in cosmic rays below the ankle is converted to that of γ-rays and neutrinos. These "calorimetric" conditions can be achieved, for instance, in starburst galaxies[29] or galaxy clusters[30].

The extragalactic γ-ray background observed by Fermi[31] has contributions from identified point-like sources on top of an isotropic γ-ray background (IGRB) shown in Fig. 6. This IGRB is expected to consist mostly of emission from the same class of γ-ray sources that are individually below Fermi's point-source detection threshold (see, e.g., Ref. 32). A significant contribution of γ-rays associated with IceCube's neutrino observation would have the somewhat surprising implication that indeed many extragalactic γ-ray sources are also neutrino emitters, while none had been detected so far. This dramatically changed when IceCube developed methods for performing real-time multiwavelength observations in cooperation with some twenty other observatories to identify the sources and build on the discovery of cosmic neutrinos to launch a new era in astronomy[33,34]. This effort led to the identification

of a distant flaring blazar as a cosmic ray accelerator in a multimessenger campaign launched by a 290 TeV energy neutrino detected from the constellation of Orion on September 22, 2017[3].

## 3. The first truly multimessenger campaign

Neutrinos only originate in environments where protons are accelerated to produce pions and other particles that decay into neutrinos. Neutrinos can thus exclusively pinpoint cosmic ray accelerators, and this is exactly what one neutrino did on September 22, 2017.

IceCube detects muon neutrinos, a type of neutrino that leaves a well-reconstructed track in the detector roughly every five minutes. Most of them are low-energy neutrinos produced in the Earth's atmosphere, which are of interest for studying the neutrinos themselves, but are a persistent background when doing neutrino astronomy. In 2016, IceCube installed an online filter that selects from this sample, in real time, very high energy neutrinos that are likely to be of cosmic origin[34]. We reconstruct their energy and celestial coordinates, typically in less than one minute, and distribute the information automatically via the Gamma-ray Coordinate Network to a group of telescopes around the globe and in space for follow-up observations. These telescopes look for electromagnetic radiation from the arrival direction of the neutrino, searching for coincident emission that can reveal its origin.

The tenth such alert[35], IceCube-170922A, on September 22, 2017, reported a well-reconstructed muon neutrino with an energy of 290 TeV and, therefore, with a high probability of originating in an astronomical source. The Fermi telescope detected a flaring blazar aligned with the cosmic neutrino within 0.06 degrees. The source is a known blazar, a supermassive black hole spitting out high-energy particles in twin jets aligned with its rotation axis which is directed at Earth. This blazar, TXS 0506+056, had been relatively poorly studied until now, although it was identified as the highest energy gamma ray source detected by EGRET from any blazar with two photons above 40 GeV[36]. The set of observations triggered by the September 22 neutrino has yielded a treasure trove of multiwavelength data that will allow us to probe the physics of the first cosmic ray accelerator. An optical telescope eventually measured its distance[37], which was found to be 4 billion light-years. Its large distance points to a special galaxy, which sets it apart from the ten-times-closer blazars such as the Markarian sources that dominate the extreme gamma-ray sky observed by NASA's Fermi satellite.

TXS 0506+056 was originally flagged by the Fermi[38] and Swift[39] satellite telescopes. Follow-up observations with the MAGIC air Cherenkov telescope[40] identified it as a rare TeV blazar with the potential to produce the very high energy neutrino detected by IceCube. The source was subsequently scrutinized in X-ray, optical, and radio wavelengths. This is a first, truly multimessenger observation: none of the instruments could have made this breakthrough independently. In total,

more than 20 telescopes observed the flaring blazar as a highly variable source in a high state[3].

It is important to realize that nearby blazars like the Markarian sources are at a redshift that is ten times smaller, and therefore TXS 0506+056, with a similar flux despite the greater distance, is one of the most luminous sources in the Universe. It likely belongs to a special class of blazars that accelerate proton beams as revealed by the neutrino. This is further supported by the fact that a variety of previous attempts to associate the arrival directions of cosmic neutrinos with a variety of Fermi blazar catalogues failed.

Informed by the multimessenger campaign, IceCube searched its archival neutrino data up to and including October 2017 in the direction of IC170922 using the likelihood routinely used in previous searches. This revealed a spectacular burst of over a dozen high-energy neutrinos in 110 days in the 2014-2015 data with a spectral index consistent with the one observed for the diffuse cosmic neutrino spectrum[4].

Interestingly, the AGILE collaboration, which operates an orbiting X-ray and gamma ray telescope, reported a day-long flare in the direction of a previous neutrino alert sent on Juli 31, 2016[41]. The flare occurred more than one day before the time of the alert. In light of the rapid daily variations observed near the peak emission of the TXS 0506+056 flare at the time of IC170922A, this may well be a genuine coincidence.

## 4. Flaring sources and the high-energy neutrino flux

The extraordinary detection of more than a dozen cosmic neutrinos in the 2014 flare despite its 0.34 redshift further suggests that TXS 0506+056 belongs to a special class of sources that produce cosmic rays. The single neutrino flare dominates the flux of the source over the 9.5 years of archival IceCube data, leaving IC170922 as a less luminous second flare in the sample. We will show that a subset of about 5% of all blazars bursting once in 10 years at the level of TXS 0506+056 in 2014, can accommodate the diffuse flux cosmic neutrino flux observed by IceCube. We already pointed out in the previous section that the energy density corresponding to this flux is similar to the one in the highest energy cosmic rays.

In order to calculate the flux of high-energy neutrinos from a population of sources, we follow[42] and relate the diffuse neutrino flux to the injection rate of cosmic rays and their efficiency to produce neutrinos in the source. For a class of sources with density $\rho$ and neutrino luminosity $L_\nu$, the all-sky neutrino flux is

$$\sum_\alpha E_\nu^2 \frac{dN_\nu}{dE_\nu} = \frac{1}{4\pi}\frac{c}{H_0}\xi_z L_\nu \rho, \tag{4}$$

where $\xi_z$ is a factor of order unity that parametrizes the integration over the redshift evolution of the sources. The relation can be adapted to a fraction $\mathcal{F}$ of sources

with episodic emission of flares of duration $\Delta t$ over a total observation time $T$:

$$\sum_\alpha E_\nu^2 \frac{dN_\nu}{dE_\nu} = \frac{1}{4\pi}\frac{c}{H_0}\xi_z L_\nu \rho \mathcal{F}\frac{\Delta t}{T}\,. \tag{5}$$

Applying this relation to the 2014 TXS 0506+056 burst that dominates the flux over the 9.5 years of neutrino observations, yields

$$\begin{aligned} 3\times 10^{-11}\,\mathrm{TeV cm^{-2} s^{-1} sr^{-1}} = & \frac{\mathcal{F}}{4\pi}\left(\frac{R_H}{3\,\mathrm{Gpc}}\right)\left(\frac{\xi_z}{0.7}\right)\left(\frac{L_\nu}{1.2\times 10^{47}\,\mathrm{erg/s}}\right) \\ & \left(\frac{\rho}{1.5\times 10^{-8}\,\mathrm{Mpc^{-3}}}\right)\left(\frac{\Delta t}{110\,\mathrm{d}}\frac{10\,\mathrm{yr}}{T}\right), \end{aligned} \tag{6}$$

a relation which is satisfied for $\mathcal{F}\sim 0.05$. In summary, a special class of BL Lac blazars that undergo $\sim$ 110-day duration flares like TXS 0506+056 once every 10 years accommodates the observed diffuse flux of high-energy cosmic neutrinos. The class of such neutrino-flaring sources represents 5% of the sources. The argument implies the observation of roughly 100 muon neutrinos per year. This is exactly the flux of cosmic neutrinos that corresponds to the $E^{-2.19}$ diffuse flux measured above 100 TeV. (Note that the majority of these neutrinos cannot be separated from the atmospheric background, leaving us with the reduced number of very high energy events discussed in the previous sections).

As previously discussed, he energetics of the cosmic neutrinos is matched by the energy of the highest energy cosmic rays, with [43]

$$\frac{1}{3}\sum_\alpha E_\nu^2 \frac{dN_\nu}{dE_\nu} \simeq \frac{c}{4\pi}\left(\frac{1}{2}(1-e^{-f_\pi})\,\xi_z t_H \frac{dE}{dt}\right). \tag{7}$$

The cosmic rays' injection rate $dE/dt$ above $10^{16}$ eV is $(1-2)\times 10^{44}$ erg Mpc$^{-3}$ yr$^{-1}$ [44,45]. From Eq. 6 it follows that the energy densities match for a pion production efficiency of the neutrino source of $f_\pi \gtrsim 0.4$. This high efficiency requirement is consistent with the premise that a special class of efficient sources is responsible for producing the high-energy cosmic neutrino flux seen by IceCube. The sources must contain sufficient target density in photons, possibly protons, to generate the large value of $f_\pi$. It is clear that the emission of flares producing the large number of cosmic neutrinos detected in the 2014 burst must correspond to major accretion events onto the black hole lasting a few months. The pionic photons will lose energy in the source and the neutrino emission is not accompanied by a flare as was the case for the 2017 event; the Fermi data, consistent with the scenario proposed, reveal photons with energies of tens of GeV, but no flaring activity.

A key question is whether the neutrino and gamma ray spectra for the 2014 neutrino burst from TXS 0506+056 satisfy the multimessenger relationship introduced in section 2. With the low statistics of the very high-energy gamma ray measurements during the burst period, the energetics is a more robust measure for evaluating the connection, especially because the source is opaque to high-energy gamma rays, as indicated by the large value of $f_\pi$, and the pionic gamma rays will

lose energy inside the source before cascading in the microwave photon background; for details see Ref. 43.

It is worth noting that this model for the diffuse neutrino flux clarifies why earlier attempts to associate it with blazars were unsuccessful; see Ref. 46,47. Clearly, the time-integrated studies are not applicable to time-dependent sources. Moreover, with a subclass of more energetic sources with lower density responsible for the diffuse flux, the constraints on blazars obtained from the relation between the point source limits and the diffuse flux are mitigated. Additional issues with this limit arise from the use of point source sensitivity (defined as 90% C.L.) rather than discovery potential (defined as $5\sigma$ C.L.) as a constraint. The former does not provide a robust statistical measure. It has been argued that the limits from the nonobservation of a point source in time-integrated searches would suppress the contribution of blazars to the total neutrino emission because the time-averaged flux of sources is directly correlated with bursts[48]. However, two points are missed in this argument. First, this argument only applies if the duty cycle of the source and the duration of the flare, supplies neutrino bursts shorter than the total time that the detector has been taking data. Second, it neglects the excess buried under the atmospheric background which might not be statistically compelling yet. For instance, in the case of TXS 0506+056, the time-integrated search reports a significance of $2.1\sigma$ only.

We are aware that these speculations are qualitative and that they may be premature and the hope is that multimessenger astronomy will provide us with more clues after the breakthrough event of September 22, 2017 that generated an unmatched data sample over all wavelengths of the electromagnetic spectrum on the first identified cosmic ray accelerator.

## 5. Summary

Getting all the elements of this puzzle to fit together is not easy, but they suggest that the blazar may contain important clues on the origin of cosmic neutrinos and cosmic rays. This breakthrough is just the beginning and raises intriguing questions. What is special about this source? Can a subclass of blazars to which it belongs accommodate the diffuse flux observed by IceCube? Are these also the sources of all high-energy cosmic rays or only of some? The TXS 0506+056 emission over the last 10 years is dominated by the single flare in 2014. If this is characteristic of the subclass of sources that it belongs to, identifying additional sources will be difficult unless more and larger neutrino telescopes yield more frequent and higher statistics neutrino alerts.

Unlike the previous SN1987A and GW170817 multimessenger events, this event could not have been observed with a single instrument. Without the initial coincident observation, IC170922A would be just one more of the few hundred cosmic neutrinos detected by IceCube and the accompanying radiation just one more flaring blazar observed by Fermi-LAT. Neutrino astronomy was born with a supernova

in 1987. Thirty years later, this recent event involves neutrinos that are tens of millions of times more energetic and are from a source a hundred thousand times more distant.

## References


1. K. Kotera and A. V. Olinto, The Astrophysics of Ultrahigh Energy Cosmic Rays, *Ann. Rev. Astron. Astrophys.* **49**, 119 (2011).
2. M. Ahlers and F. Halzen, High-energy cosmic neutrino puzzle: a review, *Rept. Prog. Phys.* **78**, p. 126901 (2015).
3. M. Aartsen *et al.*, Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A, *Science* **361**, p. eaat1378 (2018).
4. M. Aartsen *et al.*, Neutrino emission from the direction of the blazar TXS 0506+056 prior to the IceCube-170922A alert, *Science* **361**, 147 (2018).
5. A. Roberts, The Birth of high-energy neutrino astronomy: A Personal history of the DUMAND project, *Rev.Mod.Phys.* **64**, 259 (1992).
6. V. Berezinsky and G. Zatsepin, Cosmic rays at ultrahigh-energies (neutrino?), *Phys.Lett.* **B28**, 423 (1969).
7. R. Abbasi *et al.*, The IceCube Data Acquisition System: Signal Capture, Digitization, and Timestamping, *Nucl.Instrum.Meth.* **A601**, 294 (2009).
8. M. Aartsen *et al.*, Energy Reconstruction Methods in the IceCube Neutrino Telescope, *JINST* **9**, p. P03009 (2014).
9. T. Yuan, Improving the angular resolution in icecube cascade reconstruction, *TeV Particle Astrophysics 2017 (TeVPA2017), Columbus, Ohio, USA* (2017).
10. M. Aartsen *et al.*, Measurement of the Atmospheric $\nu_e$ flux in IceCube, *Phys.Rev.Lett.* **110**, p. 151105 (2013).
11. M. G. Aartsen *et al.*, Measurement of the Atmospheric $\nu_e$ Spectrum with IceCube, *Phys. Rev.* **D91**, p. 122004 (2015).
12. R. Abbasi *et al.*, Measurement of the atmospheric neutrino energy spectrum from 100 GeV to 400 TeV with IceCube, *Phys.Rev.* **D83**, p. 012001 (2011).
13. M. Ahlers and F. Halzen, Opening a New Window onto the Universe with IceCube, *Prog. Part. Nucl. Phys.* **102**, 73 (2018).
14. M. G. Aartsen *et al.*, Evidence for Astrophysical Muon Neutrinos from the Northern Sky with IceCube, *Phys. Rev. Lett.* **115**, p. 081102 (2015).
15. M. G. Aartsen *et al.*, Observation and Characterization of a Cosmic Muon Neutrino Flux from the Northern Hemisphere using six years of IceCube data, *Astrophys. J.* **833**, p. 3 (2016).
16. M. G. Aartsen *et al.*, The IceCube Neutrino Observatory - Contributions to ICRC 2017 Part II: Properties of the Atmospheric and Astrophysical Neutrino Flux (2017).
17. M. Aartsen *et al.*, First observation of PeV-energy neutrinos with IceCube, *Phys.Rev.Lett.* **111**, p. 021103 (2013).

18. M. G. Aartsen *et al.*, Evidence for High-Energy Extraterrestrial Neutrinos at the IceCube Detector, *Science* **342**, p. 1242856 (2013).
19. M. G. Aartsen *et al.*, Flavor Ratio of Astrophysical Neutrinos above 35 TeV in IceCube, *Phys. Rev. Lett.* **114**, p. 171102 (2015).
20. M. G. Aartsen *et al.*, Lowering IceCube's Energy Threshold for Point Source Searches in the Southern Sky, *Astrophys. J.* **824**, p. L28 (2016).
21. M. Aartsen *et al.*, Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data, *Phys.Rev.Lett.* **113**, p. 101101 (2014).
22. M. G. Aartsen *et al.*, The IceCube Neutrino Observatory - Contributions to ICRC 2015 Part II: Atmospheric and Astrophysical Diffuse Neutrino Searches of All Flavors, in *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015): The Hague, The Netherlands, July 30-August 6, 2015*, 2015.
23. R. J. Protheroe and T. Stanev, Electron–photon cascading of very high-energy gamma-rays in the infrared background, *Monthly Notices of the Royal Astronomical Society* **264**, 191 (1993).
24. M. Ahlers, L. Anchordoqui, M. Gonzalez-Garcia, F. Halzen and S. Sarkar, GZK Neutrinos after the Fermi-LAT Diffuse Photon Flux Measurement, *Astropart.Phys.* **34**, 106 (2010).
25. K. Murase, M. Ahlers and B. C. Lacki, Testing the Hadronuclear Origin of PeV Neutrinos Observed with IceCube, *Phys.Rev.* **D88**, p. 121301 (2013).
26. K. Murase, D. Guetta and M. Ahlers, Hidden Cosmic-Ray Accelerators as an Origin of TeV-PeV Cosmic Neutrinos, *Phys. Rev. Lett.* **116**, p. 071101 (2016).
27. K. Bechtol, M. Ahlers, M. Di Mauro, M. Ajello and J. Vandenbroucke, Evidence against star-forming galaxies as the dominant source of IceCube neutrinos, *Astrophys. J.* **836**, p. 47 (2017).
28. A. Aab *et al.*, The Pierre Auger Observatory: Contributions to the 34th International Cosmic Ray Conference (ICRC 2015), in *Proceedings, 34th International Cosmic Ray Conference (ICRC 2015): The Hague, The Netherlands, July 30-August 6, 2015*, 2015.
29. A. Loeb and E. Waxman, The Cumulative background of high energy neutrinos from starburst galaxies, *JCAP* **0605**, p. 003 (2006).
30. V. Berezinsky, P. Blasi and V. Ptuskin, Clusters of galaxies as a storage room for cosmic rays, *Astrophys J.* **487**, 529 (1997).
31. M. Ackermann *et al.*, The spectrum of isotropic diffuse gamma-ray emission between 100 MeV and 820 GeV, *Astrophys. J.* **799**, p. 86 (2015).
32. M. Di Mauro and F. Donato, Composition of the Fermi-LAT isotropic gamma-ray background intensity: Emission from extragalactic point sources and dark matter annihilations, *Phys. Rev.* **D91**, p. 123001 (2015).
33. M. G. Aartsen *et al.*, Very High-Energy Gamma-Ray Follow-Up Program Using Neutrino Triggers from IceCube, *JINST* **11**, p. P11009 (2016).
34. M. G. Aartsen *et al.*, The IceCube Realtime Alert System, *Astropart. Phys.* **92**, 30 (2017).

35. C. Kopper and E. Blaufuss, IceCube-170922A - IceCube observation of a high-energy neutrino candidate event., *GRB Coordinates Network, Circular Service, No. 21916, #1 (2017)* **21916** (2017).
36. B. L. Dingus and D. L. Bertsch, The highest energy emission detected by egret from blazars, 251 (2001), [AIP Conf. Proc.587,251(2001)].
37. S. Paiano, R. Falomo, A. Treves and R. Scarpa, The redshift of the BL Lac object TXS 0506+056, *Astrophys. J.* **854**, p. L32 (2018).
38. Y. T. Tanaka, S. Buson and D. Kocevski, Fermi-LAT detection of increased gamma-ray activity of TXS 0506+056, located inside the IceCube-170922A error region., *The Astronomer's Telegram* **10791** (September 2017).
39. P. A. E. A. Keivani, J. A. Kennea, D. B. Fox, D. F. Cowen, J. P. Osborne, F. E. Marshall and Swift-IceCube Collaboration, Further Swift-XRT observations of IceCube 170922A, *The Astronomer's Telegram* **10792** (September 2017).
40. R. Mirzoyan, First-time detection of VHE gamma rays by MAGIC from a direction consistent with the recent EHE neutrino event IceCube-170922A, *The Astronomer's Telegram* **10817** (October 2017).
41. F. Lucarelli *et al.*, AGILE detection of a candidate gamma-ray precursor to the ICECUBE-160731 neutrino event, *Astrophys. J.* **846**, p. 121 (2017).
42. F. Halzen and D. Hooper, High-energy neutrino astronomy: The Cosmic ray connection, *Rept.Prog.Phys.* **65**, 1025 (2002).
43. F. Halzen, A. Kheirandish and T. Weisgarber, A neutrino from a distant blazar sets a tipping point in the search for the most powerful cosmic objects, *In preparation* (2018).
44. M. Ahlers and F. Halzen, Minimal Cosmogenic Neutrinos, *Phys. Rev.* **D86**, p. 083010 (2012).
45. B. Katz, E. Waxman, T. Thompson and A. Loeb, The energy production rate density of cosmic rays in the local universe is $\sim 10^{44-45}$erg Mpc$^{-3}$ yr$^{-1}$ at all particle energies (2013).
46. M. Kowalski, Status of High-Energy Neutrino Astronomy, *J. Phys. Conf. Ser.* **632**, p. 012039 (2015).
47. P. Mertsch, M. Rameez and I. Tamborra, Detection prospects for high energy neutrino sources from the anisotropic matter distribution in the local universe, *JCAP* **1703**, p. 011 (2017).
48. K. Murase, F. Oikonomou and M. Petropoulou, Blazar Flares as an Origin of High-Energy Cosmic Neutrinos? (2018).

# Breakdown of the equivalence between gravitational mass and energy due to quantum effects

Andrei G. Lebed

*Department of Physics, University of Arizona, 1118 E. 4th Street, Tucson, Arizona 85721, USA*

*L. D. Landau Institute for Theoretical Physics, RAS, 2 Kosygina Street, Moscow 117334, Russia*
*lebed@physics.arizona.edu*

We review our recent theoretical results about inequivalence between passive and active gravitational masses and energy in the semiclassical variant of general relativity, where the gravitational field is not quantized but matter is quantized. To this end, we consider the simplest quantum body with internal degrees of freedom — a hydrogen atom. We concentrate our attention on the following physical effects, related to electron mass. The first one is the inequivalence between passive gravitational mass and energy at the microscopic level. Indeed, the quantum measurement of gravitational mass can give a result which is different from the expected one, $m \neq m_e + \frac{E_1}{c^2}$, where the electron is initially in its ground state; $m_e$ is the bare electron mass. The second effect is that the expectation values of both the passive and active gravitational masses of stationary quantum states are equivalent to the expectation value of the energy. The most spectacular effects are the inequivalence of the passive and active gravitational masses and the energy at the macroscopic level for an ensemble of coherent superpositions of stationary quantum states. We show that, for such superpositions, the expectation values of passive and active gravitational masses are not related to the expectation value of energy by Einstein's famous equation, $m \neq \frac{E}{c^2}$. In this paper, we also improve several drawbacks of the original pioneering works.

*Keywords*: Equivalence principle; mass–energy equivalence; quantum gravity.

## 1. Introduction

The equivalence principle (EP) between gravitational and inertial masses in combination with the local Lorentz invariance of spacetime is known to be a keystone of classical general relativity[1,2]. In the current scientific literature, there exists widespread discussion of whether it can survive in a possible quantum theory of gravity (see, for example, Refs. 3–5). Since quantum gravitation theory has not been elaborated yet, the EP is often studied in the framework of the so-called semiclassical approach to quantum gravity, where the gravitational field is not quantized, but matter is quantized[3–5]. Note that the EP for a composite body is not a trivial notion even in general relativity in the absence of quantum effects. Indeed, as shown in Refs. 6–8, the external gravitational field is coupled not directly with the energy of a composite body but with the combination $R+3K+2P$, where $R$, $K$ and $P$ are the rest, kinetic and potential energies, respectively. As mentioned in Ref. 8, and considered in detail in Ref. 9, the above mentioned combination can be changed into the expected total energy if we choose the proper local coordinates, where the line interval has the Minkowski form. Therefore, in classical general relativity,

passive gravitational mass is equivalent to the inertial one for a composite body [8,9], as expected.

On the other hand, as shown in Ref. 7, active gravitational mass of a composite classical body is equivalent to its energy only after averaging the gravitational mass over time. Semiclassical analysis [5] of Einstein's field equations has demonstrated that the expectation values of active gravitational mass and energy are equivalent only for stationary quantum states of a composite quantum body. The situation is different for an ensemble of coherent quantum superpositions of stationary quantum states, where the expectation values of active gravitational mass can oscillate in time [5] even for superpositions with constant expectation values of energy. The results of Ref. 5 go against the equivalence of active gravitational mass and energy even at the macroscopic level in quantum gravity, which has to modify the EP. Note that quantum effects also change the status of the EP for the passive gravitational mass of a quantum body with internal degrees of freedom. As discussed in Ref. 10, quantum effects break the equivalence of passive gravitational mass and energy at the microscopic level.

Let us consider this phenomenon in more detail. Suppose there is a hydrogen atom in its ground state, $E_1$, and we switch on the gravitational field to measure the electron mass, which is expected to be equal to $m = m_e + \frac{E_1}{c^2}$, where $m$ is the electron passive gravitational mass and $m_e$ is the bare electron mass. Contrary to the above mentioned common expectation, it has been shown in Ref. 10 that quantum measurements of the mass can give the following result $m = m_e + \frac{E_n}{c^2}$, where $E_n$ is energy of electron $nS$ orbital in a hydrogen atom, although the corresponding probabilities are very small. Note that the influence of quantum effects on the EP is even more dramatic and, as shown in Ref. 11, the equivalence between passive gravitational mass and energy is broken even at the macroscopic level. Indeed, the above-mentioned equivalence exists for the expectation values of the mass and energy only for stationary quantum states. In accordance with results of Ref. 11, the equivalence between the expectation values of passive gravitational mass and energy is broken for an ensemble of coherent quantum superpositions of stationary quantum states in a hydrogen atom. Of course, all these statements are not restricted by this atom but are common properties of any quantum body with internal degrees of freedom.

## 2. Goal

In Sec. 3, we discuss that there is no equivalence between passive gravitational mass and energy of electron in a hydrogen atom at a microscopic level, using the local Lorentz invariance, which defines electron wave functions in a gravitational field (we call it method-1). We start from quantum state with a definite electron energy in the absence of external gravitational field, $E_1$, and show that quantum

measurement of the mass in the field can give values different from expected one, $m \neq m_e + \frac{E_1}{c^2}$, although the corresponding probabilities are small. In Sec. 4, we discuss the same results, using corrections to the Schrödinger equation for electron in a hydrogen atom, which contain the so-called virial term in the gravitational field. We stress the importance of the virial term for the breakdown of the above mentioned equivalence (and, thus, for the breakdown of the EP) at a microscopic level (we call this method-2). In Sec. 5, we discuss the breakdown of the equivalence between passive gravitational mass and energy (and, as a result — breakdown of the EP) at macroscopic level. We show that this equivalence survives for macroscopic ensemble of stationary quantum electron states in a hydrogen atom due to the so-called quantum virial theorem. On the other hand, it is also shown that for coherent ensembles of superpositions of stationary quantum states, the above mentioned equivalence is not survived due to the quantum virial term. During our calculations in Secs. 2–4, we use property of the local Lorentz invariance of a spacetime in general relativity as well as consider passive gravitational mass to be a quantity proportional to weight of a composite body whose center-of-mass is fixed in gravitational field by some forces of nongravitational origin. Finally, in Sec. 6, we discuss the EP between active gravitational mass and energy for electron in a hydrogen atom at macroscopic level. Our results are similar to that for passive gravitational mass. Indeed, we show that the equivalence (and, thus, the EP) survives for macroscopic ensembles of stationary quantum states, whereas for macroscopic coherent ensembles of superpositions of quantum states, the equivalence is broken. In Sec. 7, we come to the conclusion that the EP has to be seriously reformulated in the presence of quantum effects in general relativity.

## 3. Inequivalence of Passive Gravitational Mass and Energy at Microscopic Level (Method-1)

### 3.1. *Electron wave function in a hydrogen atom with a definite energy in the absence of gravitational field*

Suppose that, at $t < 0$, there is no gravitational field and electron is in a ground state of a hydrogen atom, characterizing by the following wave function:

$$\Psi_1(r,t) = \exp\left(\frac{-im_e c^2 t}{\hbar}\right) \exp\left(\frac{-iE_1 t}{\hbar}\right) \Psi_1(r), \tag{1}$$

which is solution of the corresponding Schrödinger equation:

$$i\hbar \frac{\partial \Psi_1(r,t)}{\partial t} = \left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r}\right] \Psi_1(r,t). \tag{2}$$

[Here, $E_1$ is electron ground state energy, $r$ is distance between electron with coordinates $(x, y, z)$ and proton; $\hbar$ is the Planck constant, $c$ is the velocity of light.]

### 3.2. *Electron wave functions in a hydrogen atom in the presence of gravitational field*

At $t = 0$, we perform the following Gedanken experiment. We switch on a weak centrosymmetric (e.g. the Earth's) gravitational field, with position of center-of-mass of the atom (i.e. proton) being fixed in the field by some nongravitational forces. It is known that, in a weak field approximation, curved spacetime is characterized by the following interval[1,2]:

$$ds^2 = -\left(1 + 2\frac{\phi}{c^2}\right)(cdt)^2 + \left(1 - 2\frac{\phi}{c^2}\right)(dx^2 + dy^2 + dz^2), \quad \phi = -\frac{GM}{R}. \tag{3}$$

Below we introduce the proper local coordinates,

$$\tilde{t} = \left(1 + \frac{\phi}{c^2}\right)t, \quad \tilde{x} = \left(1 - \frac{\phi}{c^2}\right)x, \quad \tilde{y} = \left(1 - \frac{\phi}{c^2}\right)y, \quad \tilde{z} = \left(1 - \frac{\phi}{c^2}\right)z, \tag{4}$$

where interval has Minkowski's form[1,2],

$$d\tilde{s}^2 = -(cd\tilde{t})^2 + (d\tilde{x}^2 + d\tilde{y}^2 + d\tilde{z}^2). \tag{5}$$

[Here, we stress that, since we are interested in calculations of some quantum transition amplitudes with the first-order accuracy with respect to the small parameter, $|\frac{\phi}{c^2}| \ll 1$, we disregard in Eqs. (3)–(5) and therein below all terms of the order of $\frac{\phi^2}{c^4}$. We pay attention that near the Earth's surface the above discussed parameter is small and is equal to $|\frac{\phi}{c^2}| \sim 10^{-9}$.]

Due to the local Lorentz invariance of a spacetime in general relativity, if we disregard the so-called tidal terms in the Hamiltonian [i.e. if we don't differentiate the potential $\phi(R)$], then new wave functions, written in the local proper coordinates (4) (with fixed proton's position), satisfy at, $t$, $\tilde{t} > 0$, the similar Schrödinger equation:

$$i\hbar\frac{\partial\tilde{\Psi}(\tilde{r},\tilde{t})}{\partial\tilde{t}} = \left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial\tilde{x}^2} + \frac{\partial^2}{\partial\tilde{y}^2} + \frac{\partial^2}{\partial\tilde{z}^2}\right) - \frac{e^2}{\tilde{r}}\right]\tilde{\Psi}(\tilde{r},\tilde{t}). \tag{6}$$

[Note that it is easy to show that the above disregarded tidal terms have relative order of $\frac{r_0}{R_0}$, where $r_0$ is the Bohr radius and $R_0$ is distance between a hydrogen atom and center of source of gravitational field. Near the Earth's surface they are very small and are of the relative order of $\frac{r_0}{R_0} \sim 10^{-17}$.]

We stress that it is very important that the wave function (1) is not a solution of the Schrödinger equation (6) anymore and, thus, is not characterized by definite energy and weight in the gravitational field (3). Moreover, a general solution of

Eq. (6) can be written in the proper local coordinates in the following way:

$$\tilde{\Psi}(\tilde{r},\tilde{t}) = \exp\left(\frac{-im_e c^2 \tilde{t}}{\hbar}\right) \sum_{n=1}^{\infty} \tilde{a}_n \Psi_n(\tilde{r}) \exp\left(\frac{-iE_n \tilde{t}}{\hbar}\right), \tag{7}$$

where the wave functions $\Psi_n(\tilde{r})$ are solutions[12] for the so-called $nS$ atomic orbitals of a hydrogen atom with energies $E_n$ and are normalized in the proper local space,

$$\int \Psi_n^2(\tilde{r}) d^3\tilde{r} = 1. \tag{8}$$

[It is possible to show that only $1S \to nS$ quantum transitions amplitudes are nonzero in a hydrogen atom in the gravitational field (3), which corresponds only to real wave functions. Therefore, we keep in Eq. (7) only $nS$ atomic orbitals and everywhere below disregard difference between $\Psi_n(r)$ and $\Psi_n^*(r) = \Psi_n(r)$.]

Note that the normalized wave function (1) can be rewritten in the proper local spacetime coordinates (4) in the following way:

$$\Psi_1(\tilde{r},\tilde{t}) = \exp\left[\frac{-im_e c^2 \left(1-\frac{\phi}{c^2}\right)\tilde{t}}{\hbar}\right] \exp\left[\frac{-iE_1 \left(1-\frac{\phi}{c^2}\right)\tilde{t}}{\hbar}\right]$$
$$\times \left(1+\frac{\phi}{c^2}\right)^{3/2} \Psi_1\left[\left(1+\frac{\phi}{c^2}\right)\tilde{r}\right]. \tag{9}$$

It is important that the gravitational field (3) can be considered as a sudden perturbation to the Hamiltonian (2), therefore, at $t = \tilde{t} = 0$, the wave functions (7) and (9) have to be equal to each other:

$$\left(1+\frac{\phi}{c^2}\right)^{3/2} \Psi_1\left[\left(1+\frac{\phi}{c^2}\right)\tilde{r}\right] = \sum_{n=1}^{\infty} \tilde{a}_n \Psi_n(\tilde{r}). \tag{10}$$

From Eq. (10), it directly follows that

$$\tilde{a}_1 = \left(1+\frac{\phi}{c^2}\right)^{3/2} \int_0^{\infty} \Psi_1\left[\left(1+\frac{\phi}{c^2}\right)\tilde{r}\right] \Psi_1(\tilde{r}) d^3\tilde{r}, \tag{11}$$

and

$$\tilde{a}_n = \left(1+\frac{\phi}{c^2}\right)^{3/2} \int_0^{\infty} \Psi_1\left[\left(1+\frac{\phi}{c^2}\right)\tilde{r}\right] \Psi_n(\tilde{r}) d^3\tilde{r}, \quad n > 1. \tag{12}$$

### 3.3. *Probabilities and amplitudes*

Below, we calculate quantum mechanical amplitudes (11) and (12) in a linear approximation with respect to the gravitational potential,

$$\tilde{a}_1 = 1 + O\left(\frac{\phi^2}{c^4}\right), \tag{13}$$

and

$$\tilde{a}_n = \left(\frac{\phi}{c^2}\right)\int_0^\infty \left[\frac{d\Psi_1(\tilde{r})}{d\tilde{r}}\right]\tilde{r}\Psi_n(\tilde{r})d^3\tilde{r}, \quad n>1. \tag{14}$$

We stress that the wave function (7) is a series of wave functions, which have definite weights in the gravitational field (3). This means that they are characterized by the following definite electron passive gravitational masses:

$$m_n = m_e + \frac{E_n}{c^2}. \tag{15}$$

In accordance with the most general properties of quantum mechanics, this means that, if we do a measurement of gravitational mass for wave function (1) and (9), we obtain quantum values (15) with the probabilities: $\tilde{P}_n = |\tilde{a}_n|^2$, where $\tilde{a}_n$ are given by Eqs. (13) and (14).

Let us show that

$$\int_0^\infty \left[\frac{d\Psi_1(\tilde{r})}{d\tilde{r}}\right]\tilde{r}\Psi_n(\tilde{r})d^3\tilde{r} = \frac{V_{1n}}{E_n - E_1}, \quad n>1, \tag{16}$$

where $\hat{V}(\tilde{r})$ is the so-called quantum virial operator [12]:

$$\hat{V}(r) = -2\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial\tilde{x}^2}+\frac{\partial^2}{\partial\tilde{y}^2}+\frac{\partial^2}{\partial\tilde{z}^2}\right) - \frac{e^2}{\tilde{r}}, \tag{17}$$

and

$$V_{1,n} = \int_0^\infty \Psi_1(\tilde{r})\hat{V}(\tilde{r})\Psi_n(\tilde{r})d^3\tilde{r}. \tag{18}$$

To this end, we rewrite the Schrödinger equation in gravitational field (6) in terms of the initial coordinates $(x, y, z)$

$$(m_e c^2 + E_1)\Psi_1\left[\left(1-\frac{\phi}{c^2}\right)r\right] = \left[m_e c^2 - \frac{1}{\left(1-\frac{\phi}{c^2}\right)^2}\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) - \frac{1}{\left(1-\frac{\phi}{c^2}\right)}\frac{e^2}{r}\right]\Psi_1\left[\left(1-\frac{\phi}{c^2}\right)r\right]. \tag{19}$$

Then, keeping as usual only terms of the first order with respect to the small parameter $|\frac{\phi}{c^2}| \ll 1$, we obtain

$$E_1\Psi_1(r) - \frac{\phi}{c^2}E_1 r\left[\frac{d\Psi_1(r)}{dr}\right] = \left[-\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r} + \frac{\phi}{c^2}\hat{V}(r)\right] \times \left[\Psi_1(r) - \frac{\phi}{c^2}r\left[\frac{d\Psi_1(r)}{dr}\right]\right], \tag{20}$$

and as a result

$$-E_1 r\left[\frac{d\Psi_1(r)}{dr}\right] = \left[-\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r}\right]\left[-r\frac{d\Psi_1(r)}{dr}\right] + \hat{V}(r)\Psi_1(r). \tag{21}$$

Let us multiply Eq. (21) on $\Psi_1(r)$ and integrate over space,

$$-E_1 \int_0^\infty \Psi_n(r) r \left[\frac{d\Psi_1(r)}{dr}\right] d^3r = \int_0^\infty \Psi_n(r) \left[-\frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right)-\frac{e^2}{r}\right] \times \left[-r\frac{d\Psi_1(r)}{dr}\right] d^3r + \int_0^\infty \Psi_n(r)\hat{V}(r)\Psi_1(r)d^3r. \tag{22}$$

Taking into account that the Hamiltonian operator is the Hermitian one, we rewrite Eq. (22) as

$$E_1 \int_0^\infty \Psi_n(r) r \left[\frac{d\Psi_1(r)}{dr}\right] d^3r = E_n \int_0^\infty \Psi_n(r) r \left[\frac{d\Psi_1(r)}{dr}\right] d^3r - \int_0^\infty \Psi_1(r)\hat{V}(r)\Psi_n(r)d^3r. \tag{23}$$

Then, Eqs. (16)–(18) directly follow from Eq. (23).

As a result, the calculated amplitudes (14) and the corresponding probabilities for $n \neq 1$ can be rewritten as functions of matrix elements (18) of the virial operator (17),

$$\tilde{a}_n = \left(\frac{\phi}{c^2}\right)\frac{V_{1,n}}{E_n - E_1}, \tag{24}$$

and

$$\tilde{P}_n = |\tilde{a}_n|^2 = \left(\frac{\phi}{c^2}\right)^2 \left(\frac{V_{1,n}}{E_n - E_1}\right)^2. \tag{25}$$

Note that near the Earth's surface, where $\frac{\phi^2}{c^4} \approx 0.49 \times 10^{-18}$, the probability for $n = 2$ in a hydrogen atom can be calculated as

$$\tilde{P}_2 = |\tilde{a}_2|^2 = 1.5 \times 10^{-19}, \tag{26}$$

where

$$\frac{V_{1,2}}{E_2 - E_1} = 0.56. \tag{27}$$

It is important that nonzero matrix elements (18) of the virial operator (17) for $n \neq 1$ are also responsible for breakdown of the equivalence between active gravitational mass and energy for a quantum body with internal degrees of freedom [5].

## 4. Inequivalence of Passive Gravitational Mass and Energy at Microscopic Level (Method-2)

### 4.1. *Schrödinger equation with a definite energy in the absence of gravitational field*

As in the previous section, at $t < 0$, gravitational field is zero and electron occupies ground state in a hydrogen atom, characterizing by the wave function (1). As we

have already discussed, the wave function (1) corresponds to the $1S$ electron orbital and is known to be a ground state solution of Eq. (2).

### 4.2. *Schrödinger equation in the presence of gravitational field*

Let us consider the same Gedanken experiment as in Sec. 3. We switch on the weak gravitational field (3) and obtain Eq. (6) for the wave functions in the proper local spacetime coordinates (4). But, in this section, we rewrite Eq. (6) in the initial spacetime coordinates, $(t, x, y, z)$,

$$i\hbar\frac{\partial\Psi(\mathbf{r},t)}{\partial t} = \left\{\left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r}\right]\right.$$
$$\left. + \left(\frac{\phi}{c^2}\right)\left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}+\frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r} + \hat{V}(\mathbf{r})\right]\right\}\Psi(\mathbf{r},t), \tag{28}$$

where the virial operator [12], $\hat{V}(r)$, is equal to (17). From Eq. (28), it directly follows that the external gravitational field (3) is coupled not only to Hamiltonian (2) but also to the virial operator (17). It is important that the virial term (17) does not commute with the Hamiltonian (2), therefore, it breaks the equivalence of the passive gravitational mass and energy for electron in a hydrogen atom.

#### 4.2.1. *More general Lagrangian*

Here, we derive Hamiltonian (28) from more general Lagrangian. Let us consider the Lagrangian of a three-body system: a hydrogen atom and the Earth in inertial coordinate system, treating gravitation (3) as a small perturbation in Minkowski's spacetime. In this case, we can make use of the results of Ref. 7, where the corresponding $n$-body Lagrangian is calculated as a sum of the following four terms:

$$L = L_{\rm kin} + L_{\rm em} + L_G + L_{e,G}, \tag{29}$$

where $L_{\rm kin}$, $L_{\rm em}$, $L_G$, and $L_{e,G}$ are kinetic, electromagnetic, gravitational and electric-gravitational parts of the Lagrangian, respectively. We recall that, in our approximation, we keep in the Lagrangian and Hamiltonian only terms of the order of $(v/c)^2$ and $|\phi|/c^2$ as well as keep only classical kinetic and the Coulomb electrostatic potential energies couplings with external gravitational field. It is possible to show that, in our case, different contributions to the Lagrangian (29) can be simplified:

$$L_{\rm kin} + L_{\rm em} = -Mc^2 - m_p c^2 - m_e c^2 + m_e\frac{\mathbf{v}^2}{2} + \frac{e^2}{r}, \tag{30}$$

$$L_G = G\frac{m_p M}{R} + G\frac{m_e M}{R} + \frac{3}{2}G\frac{m_e M}{R}\frac{\mathbf{v^2}}{c^2}, \tag{31}$$

$$L_{e,G} = -2G\frac{M}{Rc^2}\frac{e^2}{r}, \tag{32}$$

where, as usual, we use the inequality $m_p \gg m_e$, with $m_p$ being the bare proton mass.

If we keep only those terms in the Lagrangian which are related to electron motion (as usual, proton is supposed to be supported by some nongravitational forces in the gravitational field), then we can write the Lagrangian (30)–(32) in the following form:

$$L = -m_e c^2 + m_e \frac{\mathbf{v}^2}{2} + \frac{e^2}{r} - \frac{\phi(R)}{c^2}\left[m_e + 3m_e\frac{\mathbf{v}^2}{2} - 2\frac{e^2}{r}\right], \quad \phi(R) = -G\frac{M}{R}. \tag{33}$$

It is easy to show that the corresponding electron Hamiltonian is

$$\tilde{H} = \left\{\left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r}\right] + \left(\frac{\phi}{c^2}\right)\left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r} + \hat{V}(\mathbf{r})\right]\right\}. \tag{34}$$

Note that Eq. (34) exactly coincides with electron Hamiltonian (28), obtained by us in the previous section.

4.2.2. *More general Hamiltonian*

Let us derive the Hamiltonian (28) and (34) from more general arguments. The so-called gravitational Stark effect (i.e. the mixing effect between even and odd wave functions in a hydrogen atom in gravitational field) was studied in Ref. 13 in the weak external gravitational field (3). Note that the corresponding Hamiltonian was derived in $1/c^2$ approximation and a possibility of center-of-mass of the atom motion was taken into account. The main peculiarity of the calculations in the above-mentioned paper was the fact that not only terms of the order of $\phi/c^2$ were calculated, as in our case, but also terms of the order of $\phi'/c^2$. Here, we use a symbolic notation $\phi'$ for the first derivatives of gravitational potential. In accordance with the existing tradition, we refer to the latter terms as to the tidal ones. Note that the Hamiltonian (3.24) was obtained in Ref. 13 directly from the Dirac equation in a curved spacetime of general relativity. As shown in Ref. 13, it can be rewritten for the corresponding Schrödinger equation as a sum of the four terms:

$$\hat{H}(\hat{\mathbf{P}}, \hat{\mathbf{p}}, \tilde{\mathbf{R}}, r) = \hat{H}_0(\hat{\mathbf{P}}, \hat{\mathbf{p}}, r) + \hat{H}_1(\hat{\mathbf{P}}, \hat{\mathbf{p}}, \tilde{\mathbf{R}}, r) + \hat{H}_2(\hat{\mathbf{p}}, \mathbf{r}) + \hat{H}_3(\hat{\mathbf{P}}, \hat{\mathbf{p}}, \tilde{\mathbf{R}}, r), \tag{35}$$

$$\hat{H}_0(\hat{\mathbf{P}}, \hat{\mathbf{p}}, r) = m_e c^2 + m_p c^2 + \left[\frac{\hat{\mathbf{P}}^2}{2(m_e + m_p)} + \frac{\hat{\mathbf{p}}^2}{2\mu}\right] - \frac{e^2}{r}, \tag{36}$$

$$\hat{H}_1(\hat{\mathbf{P}}, \hat{\mathbf{p}}, \tilde{\mathbf{R}}, r) = \left\{m_e c^2 + m_p c^2 + \left[3\frac{\hat{\mathbf{P}}^2}{2(m_e + m_p)} + 3\frac{\hat{\mathbf{p}}^2}{2\mu} - 2\frac{e^2}{r}\right]\right\}\left(\frac{\phi - \mathbf{g}\tilde{\mathbf{R}}}{c^2}\right), \tag{37}$$

$$\hat{H}_2(\hat{\mathbf{p}},\mathbf{r}) = \frac{1}{c^2}\left(\frac{1}{m_e}-\frac{1}{m_p}\right)[-(\mathbf{gr})\hat{\mathbf{p}}^2 + i\hbar\mathbf{g}\hat{\mathbf{p}}]$$
$$+\frac{1}{c^2}\mathbf{g}\left(\frac{\hat{\mathbf{s}}_\mathbf{e}}{m_e}-\frac{\hat{\mathbf{s}}_p}{m_p}\right)\times\hat{\mathbf{p}} + \frac{e^2(m_p-m_e)}{2(m_e+m_p)c^2}\frac{\mathbf{gr}}{r}, \tag{38}$$

$$\hat{H}_3(\hat{\mathbf{P}},\hat{\mathbf{p}},\tilde{\mathbf{R}},r) = \frac{3}{2}\frac{i\hbar\mathbf{gP}}{(m_e+m_p)c^2} + \frac{3}{2}\frac{\mathbf{g}(\mathbf{s_e}+\mathbf{s_p})\times\mathbf{P}}{(m_e+m_p)c^2}$$
$$-\frac{(\mathbf{gr})(\mathbf{Pp})+(\mathbf{Pr})(\mathbf{gp})-i\hbar\mathbf{gP}}{(m_e+m_p)c^2}, \tag{39}$$

where $\mathbf{g} = -G\frac{M}{R^3}\mathbf{R}$. Note that we use the following notations in Eqs. (35)–(39): $\tilde{\mathbf{R}}$ and $\mathbf{P}$ stand for coordinate and momentum of a hydrogen atom center-of-mass, respectively; whereas, $\mathbf{r}$ and $\mathbf{p}$ stand for relative electron coordinate and momentum in center-of-mass coordinate system; $\mu = m_e m_p/(m_e+m_p)$ is the so-called reduced electron mass. We point out that $\hat{H}_0(\hat{\mathbf{P}},\hat{\mathbf{p}},r)$ is the Hamiltonian of a hydrogen atom in the absence of the field. It is important that the Hamiltonian $\hat{H}_1(\hat{\mathbf{P}},\hat{\mathbf{p}},\tilde{\mathbf{R}},r)$ describes couplings not only of the bare electron and proton masses with the gravitational field (3) but also couplings of electron kinetic and potential energies with the field. And finally, the Hamiltonians $\hat{H}_2(\hat{\mathbf{p}},\mathbf{r})$ and $\hat{H}_3(\hat{\mathbf{P}},\hat{\mathbf{p}},\tilde{\mathbf{R}},r)$ describe only the tidal effects.

Let us strictly derive the Hamiltonian (28) and (34), which has already been semi-quantitatively derived, from the more general Hamiltonian (35)–(39). As already mentioned, we use the approximation, where $m_p \gg m_e$, and, therefore, $\mu = m_e$. In particular, this allows us to consider proton as a heavy classical particle. We recall that we need to derive the Hamiltonian of the atom, whose center-of-mass is at rest with respect to the Earth. Thus, we can omit the center-of-mass kinetic energy and center-of-mass momentum. As a result, the first two contributions to electron part of the total Hamiltonian (35)–(39) can be written in the following way:

$$\hat{H}_0(\hat{\mathbf{p}},r) = m_e c^2 + \frac{\hat{\mathbf{p}}^2}{2m_e} - \frac{e^2}{r}, \tag{40}$$

and

$$\hat{H}_1(\hat{\mathbf{p}},r) = \left\{m_e c^2 + \left[3\frac{\hat{\mathbf{p}}^2}{2m_e} - 2\frac{e^2}{r}\right]\right\}\left(\frac{\phi}{c^2}\right), \tag{41}$$

where we place the center-of-mass of the atom at point $\tilde{\mathbf{R}} = 0$. Now, let us study the first tidal term (38) in the total Hamiltonian (35). At first, we pay attention that $|\mathbf{g}| \simeq |\phi|/R_0$. Then, as is well known, in a hydrogen atom $|\mathbf{r}| \sim \hbar/|\mathbf{p}| \sim r_B$ and $\mathbf{p}^2/(2m_e) \sim e^2/r_B$. These values allow us to evaluate the first tidal term (38) in the Hamiltonian (35) as $H_2 \sim (r_B/R_0)(|\phi|/c^2)(e^2/r_B) \sim 10^{-17}(|\phi|/c^2)(e^2/r_B)$. Note that this value is $10^{-17}$ times smaller than $H_1 \sim (|\phi|/c^2)(e^2/r_B)$ and $10^{-8}$ times smaller than the second correction with respect to the small parameter $|\phi|/c^2$. Therefore, we can disregard the contribution (38) to the total Hamiltonian (35).

As to the second tidal term (39) in the total Hamiltonian, we pay attention that it is exactly zero in the case, where $\mathbf{P} = 0$, considered in this review. Therefore, we can conclude that the Hamiltonian (40) and (41), derived in this section, exactly coincides with in the one that was semi-quantitatively derived by us earlier [see Eqs. (28) and (34)].

### 4.3. *Gravitational field as a perturbation to the Hamiltonian*

It is important that the gravitational field (3), under the condition of our Gedanken experiment, can be considered as the following sudden perturbation, $\hat{U}_1(\mathbf{r}, t)$, to the Hamiltonian (2) in the absence of gravitational field:

$$\hat{U}_1(\mathbf{r}, t) = \left(\frac{\phi}{c^2}\right)\left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r} + \hat{V}(\mathbf{r})\right]\Theta(t), \quad (42)$$

where $\Theta(t)$ is the so-called step function. Then, a general solution of Eq. (28) can be written in the following way:

$$\Psi(r,t) = \exp\left(\frac{-i\tilde{m}_e c^2 t}{\hbar}\right)\Psi_1^1(r)\exp\left(\frac{-i\tilde{E}_1 t}{\hbar}\right) + \exp\left(\frac{-im_e c^2 t}{\hbar}\right)\sum_{n>1}^{\infty} a_n \Psi_n(r)\exp\left(\frac{-iE_n t}{\hbar}\right), \quad (43)$$

where the wave functions $\Psi_n(r)$ are solutions for the $nS$ orbitals in a hydrogen atom and are normalized,

$$\int [\Psi_1^1(r)]^2 d^3r = 1; \quad \int [\tilde{\Psi}_n(r)]^2 d^3r = 1, \quad n > 1. \quad (44)$$

[It is easy to show that perturbation (42) can result only in nonzero quantum transitions between $1S$ and $nS$ electron orbitals, therefore, we keep in Eq. (43) only $\Psi_n(r)$ wave functions which are real.]

According to the standard time-dependent perturbation theory [12], the corrected wave function of ground state, $\Psi_1^1(r)$, as well as the corrections to mass and energy of ground state in Eq. (43) can be written as

$$\Psi_1^1(r) = \Psi_1(r) + \left(\frac{\phi}{c^2}\right)\sum_{n>1}^{\infty}\frac{V_{n,1}}{E_1 - E_n}\Psi_n(r),$$
$$\tilde{m}_e = \left(1 + \frac{\phi}{c^2}\right)m_e, \quad \tilde{E}_1 = \left(1 + \frac{\phi}{c^2}\right)E_1, \quad (45)$$

where $V_{n,1}$ is matrix element of the virial operator (17):

$$V_{n,1} = \int \Psi_n(r)\left[-2\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r}\right]\Psi_1(r)d^3\mathbf{r}. \quad (46)$$

Note that the very last term in Eq. (45) corresponds to the so-called red shift in gravitational field. It is due to the expected contribution to passive gravitational

mass from electron binding energy in the atom. As to the coefficients $a_n$ with $n \neq 1$ in Eq. (43), they can be also written in terms of the virial operator matrix elements,

$$a_n = -\left(\frac{\phi}{c^2}\right)\left(\frac{V_{n,1}}{E_1 - E_n}\right), \tag{47}$$

and coincides with Eq. (24). Note that the wave function in (43)–(47), which corresponds to electron ground energy level in the presence of the gravitational field (3) (i.e. at $t > 0$), is a series of eigenfunctions of electron energy operator, taken in the absence of the field. Therefore, if we measure energy, in electron quantum state (43)–(47), we obtain the following quantized values for electron gravitational mass:

$$m_n = m_e + \frac{E_n}{c^2}, \tag{48}$$

where we omit the red shift effect. From Eqs. (43)–(48), we can state that the expected Einstein's equation, $m = m_e + \frac{E_1}{c^2}$, survives in our case with probability close to 1, whereas with the following small probabilities,

$$P_n = |a_n|^2 = \left(\frac{\phi}{c^2}\right)^2 \frac{V_{n,1}^2}{(E_n - E_1)^2}, \quad n \neq 1, \tag{49}$$

it is broken. The reason for this breakdown is that, the virial term (17) does not commute with the Hamiltonian (2) in the absence of gravitational field. As a result, electron wave functions with definite passive gravitational masses are not characterized by definite energies in the absence of gravitational field. It is important that our current results coincide with that obtained in Sec. 3 by a different method.

### 4.4. *Experimental aspects*

Here, let us describe another Gedanken experiment, where gravitational field is adiabatically switched on. To this end, we consider wave function (1) to be valid at $t \to -\infty$ and apply the following perturbation, due to the gravitational field (3), for the Hamiltonian (2):

$$\hat{U}_2(\mathbf{r}, t) = \left(\frac{\phi}{c^2}\right)\left[m_e c^2 - \frac{\hbar^2}{2m_e}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) - \frac{e^2}{r} + \hat{V}(\mathbf{r})\right]\exp(\lambda t), \quad \lambda \to 0. \tag{50}$$

Then, at $t \simeq 0$ (i.e. in the presence of the field), the electron wave function can be written as

$$\Psi(r, t) = \exp\left(\frac{-i\tilde{m}_e c^2 t}{\hbar}\right)\Psi_1^1(r)\exp\left(\frac{-i\tilde{E}_1 t}{\hbar}\right) + \exp\left(\frac{-im_e c^2 t}{\hbar}\right)\sum_{n>1}^{\infty} a_n \Psi_n(r)\exp\left(\frac{-iE_n t}{\hbar}\right). \tag{51}$$

Application of the standard time-dependent perturbation theory [12] in the case of adiabatic switching on gravitational field results in

$$\Psi_1^1(r) = \Psi_1(r) + \left(\frac{\phi}{c^2}\right) \sum_{n>1}^{\infty} \frac{V_{n,1}}{E_1 - E_n} \Psi_n(r),$$
$$\tilde{m}_e = \left(1 + \frac{\phi}{c^2}\right) m_e, \quad \tilde{E}_1 = \left(1 + \frac{\phi}{c^2}\right) E_1, \tag{52}$$

and

$$a_n = 0, \quad P_n = 0. \tag{53}$$

Thus, in adiabatic limit, the phenomenon of quantization of passive gravitational mass (15) and (48) disappears. This means that the possible experimental observation of the above mentioned phenomenon has to be done in quickly changing gravitational field. It is important that the step-like function, $\Theta(t)$, which was used to derive Eq. (48), does not mean motion of a source of gravity with velocity higher than the speed of light. We can use the step-like function if significant change of gravitational field happens quicker than the characteristic period of quasiclassical rotation of electron in a hydrogen atom. In the case under consideration, we need the time of about $\delta t \leq t_0 = \frac{2\pi\hbar}{E_2 - E_1} \sim 10^{-15} s$. Of course, there exist much more convenient quantum systems with higher values of the parameter $t_0$, where the above discussed phenomenon could be observed. We recall that all excited energy levels are quasistationary and, thus, decay with time by emitting photons. Therefore, it is much more efficient to detect emitted photons than to directly measure a weight. As to the relatively small probabilities (24) of the mass quantization, it is not too small and can be compensated by large value of the Avogadro number, $N_A \approx 6 \times 10^{23}$. In other words, for macroscopic number of the atoms, we may have large number of emitted photons. For instance, the number of excited electrons (i.e. emitted photons) for 1000 moles of the atoms is estimated as

$$N_n = 2.95 \times 10^8 \times \left(\frac{V_{n,1}}{E_n - E_1}\right)^2, \quad N_2 = 0.9 \times 10^8. \tag{54}$$

## 5. Inequivalence Between Passive Gravitational Mass and Energy at Macroscopic Level

In Sec. 5, we perform our Gedanken experiment, where we switch on the gravitational field (3) for $t > 0$, using the gravitational field as a sudden perturbation (42). We consider two different cases: macroscopic ensemble of stationary quantum states and macroscopic ensemble of coherent superpositions of stationary quantum states. In this section, we disregard small probabilities of the order of $\frac{\phi^2}{c^2}$ [see Eqs. (25) and (49)] and, thus, ignore mass quantization phenomenon.

### 5.1. *Equivalence between passive gravitational mass and energy of stationary quantum states*

Suppose that, at $t < 0$, there is no gravitational field and we have macroscopic ensemble of hydrogen atoms with electrons being in their ground states (1). At $t > 0$, we perform our Gedanken experiment: we switch on gravitational field, which is treated as the perturbation (42) in inertial system. Let us for the moment consider one atom. At $t > 0$, general solution for electron wave function is

$$\Psi(r,t) = \exp\left(\frac{-i\tilde{m}_e c^2 t}{\hbar}\right)\Psi_1^1(r)\exp\left(\frac{-i\tilde{E}_1 t}{\hbar}\right) + \exp\left(\frac{-im_e c^2 t}{\hbar}\right)\sum_{n>1}^{\infty} a_n \Psi_n(r)\exp\left(\frac{-iE_n t}{\hbar}\right). \tag{55}$$

If we disregard small probabilities $a_n$ for $n > 1$, which were considered in Secs. 3 and 4, we can rewrite Eq. (55) as

$$\Psi(r,t) = \exp\left(\frac{-i\tilde{m}_e c^2 t}{\hbar}\right)\Psi_1^1(r)\exp\left(\frac{-i\tilde{E}_1 t}{\hbar}\right). \tag{56}$$

In accordance with the quantum perturbation theory [12], first-order correction to energy of wave function (56) can be written as

$$\tilde{m}_e = \left(1 + \frac{\phi}{c^2}\right) m_e, \quad \tilde{E}_1 = \left(1 + \frac{\phi}{c^2}\right) E_1, \tag{57}$$

which is a well known redshift [1]. It is important that, in Eq. (57), there is no correction due to the quantum virial term (42) and (46). The virial term correction is zero due to the so-called quantum virial theorem [12], which claims that for any value of $n$, including $n = 1$,

$$V_{n,n} = \int \Psi_n(r)\hat{V}(\mathbf{r})\Psi_n(r)d^3\mathbf{r} = 0. \tag{58}$$

Equation (57) directly demonstrates the equivalence of gravitational mass and energy at macroscopic level.

### 5.2. *Inequivalence between passive gravitational mass and energy for macroscopic coherent ensemble of superpositions of stationary quantum states*

Suppose that, in the absence of gravitational field (i.e. at $t < 0$), we have macroscopic ensemble of coherent superpositions of two wave functions, corresponding to ground state wave function, $\Psi_1(r)$, and first excited energy level wave function, $\Psi_2(r)$, in a hydrogen atom:

$$\Psi(r,t) = \frac{1}{\sqrt{2}}\exp\left(\frac{-im_e c^2 t}{\hbar}\right)\left[\exp\left(\frac{-iE_1 t}{\hbar}\right)\Psi_1(r) + \exp\left(\frac{-iE_2 t}{\hbar}\right)\Psi_2(r)\right]. \tag{59}$$

Coherent ensemble of such wave functions, where the difference between phases of functions $\Psi_1(r)$ and $\Psi_2(r)$ is fixed, is possible to create by using lasers. We perform the same Gedanken experiment and, therefore, suddenly switch on the gravitational field (3) at $t > 0$ [see the corresponding perturbation (42) to the Hamiltonian (2)]:

$$U_1(\mathbf{r}, t) = \frac{\phi}{c^2}[m_e c^2 + \hat{H}_0(\mathbf{r}) + \hat{V}(\mathbf{r})]\Theta(t). \tag{60}$$

Then, if we disregard small probabilities of the order of $\frac{\phi^2}{c^4}$ [see Eq. (25)] and, thus, don't take into account the mass quantization phenomenon (15), we can consider wave function (59) as wave function of two-level system and can use the corresponding variant of the time-dependent perturbation theory. In accordance with this theory [12], the wave functions in the gravitational field (3) can be written as

$$\Psi^1(r,t) = \exp\left(\frac{-im_e c^2 t}{\hbar}\right)\left[\exp\left(\frac{-iE_1 t}{\hbar}\right) a_1(t)\Psi_1(r) + \exp\left(\frac{-iE_2 t}{\hbar}\right) a_2(t)\Psi_2(r)\right]. \tag{61}$$

Using the results of the time-dependent perturbation theory, it is possible to find equations to determine the functions $a_1(t)$ and $a_2(t)$:

$$\begin{aligned} \frac{da_1(t)}{dt} &= -iU_{11}(t)a_1(t) - iU_{12}(t)\exp\left[-i\frac{(E_2 - E_1)t}{\hbar}\right] a_2(t), \\ \frac{da_2(t)}{dt} &= -iU_{22}(t)a_2(t) - iU_{21}(t)\exp\left[-i\frac{(E_1 - E_2)t}{\hbar}\right] a_1(t), \end{aligned} \tag{62}$$

where

$$\begin{aligned} U_{11}(t) &= \Theta(t)\frac{\phi}{c^2}\int \Psi_1^*(r)[m_e c^2 + \hat{H}(\mathbf{r}) + \hat{V}(\mathbf{r})]\Psi_1(r)d^3\mathbf{r} = \Theta(t)\frac{\phi}{c^2}(m_e c^2 + E_1), \\ U_{12}(t) &= \Theta(t)\frac{\phi}{c^2}\int \Psi_1^*(r)[m_e c^2 + \hat{H}(\mathbf{r}) + \hat{V}(\mathbf{r})]\Psi_2(r)d^3\mathbf{r} = \Theta(t)\frac{\phi}{c^2}V_{12}, \\ U_{22}(t) &= \Theta(t)\frac{\phi}{c^2}\int \Psi_2^*(r)[m_e c^2 + \hat{H}(\mathbf{r}) + \hat{V}(\mathbf{r})]\Psi_2(r)d^3\mathbf{r} = \Theta(t)\frac{\phi}{c^2}(m_e c^2 + E_2), \\ U_{21}(t) &= \Theta(t)\frac{\phi}{c^2}\int \Psi_2^*(r)[m_e c^2 + \hat{H}(\mathbf{r}) + \hat{V}(\mathbf{r})]\Psi_1(r)d^3\mathbf{r} = \Theta(t)\frac{\phi}{c^2}V_{21}, \end{aligned} \tag{63}$$

where $V_{ij}$ are matrix elements of the virial operator (17). After substitution of Eq. (63) in Eq. (62), it possible to find that the function (59) is

$$\Psi^1(r,t) = \exp\left(\frac{-im_e c^2 t}{\hbar}\right)[\Psi_1^1(r,t) + \Psi_2^1(r,t)], \tag{64}$$

where

$$\Psi_1^1(r,t) = \frac{1}{\sqrt{2}} \exp\left[-i\frac{(m_e c^2 + E_1)\phi t}{c^2\hbar}\right] \exp\left(-i\frac{E_1 t}{\hbar}\right) \left[1 - \frac{\phi V_{12}}{c^2(E_2 - E_1)}\right] \Psi_1(r) + \frac{1}{\sqrt{2}} \exp\left(-i\frac{E_2 t}{\hbar}\right) \frac{\phi V_{12}}{c^2(E_2 - E_1)} \Psi_1(r), \tag{65}$$

and

$$\Psi_2^1(r,t) = \frac{1}{\sqrt{2}} \exp\left[-i\frac{(m_e c^2 + E_2)\phi t}{c^2\hbar}\right] \exp\left(-i\frac{E_2 t}{\hbar}\right) \left[1 - \frac{\phi V_{21}}{c^2(E_1 - E_2)}\right] \Psi_2(r) + \frac{1}{\sqrt{2}} \exp\left(-i\frac{E_1 t}{\hbar}\right) \frac{\phi V_{21}}{c^2(E_1 - E_2)} \Psi_2(r). \tag{66}$$

It is possible to demonstrate that with accuracy to the first order of the small parameter, $\frac{|\phi|}{c^2} \ll 1$, the wave function (64)–(66) can be written as

$$\Psi^1(r,t) = \frac{1}{\sqrt{2}} \exp\left[-i\frac{(m_e c^2 + E_1)(1+\phi)t}{c^2\hbar}\right] \left\{ \left[1 - \frac{\phi V_{12}}{c^2(E_2 - E_1)}\right] \Psi_1(r) + \frac{\phi V_{21}}{c^2(E_1 - E_2)} \Psi_2(r) \right\} + \frac{1}{\sqrt{2}} \exp\left[-i\frac{(m_e c^2 + E_2)(1+\phi)t}{c^2\hbar}\right] \left\{ \left[1 - \frac{\phi V_{21}}{c^2(E_1 - E_2)}\right] \Psi_2(r) + \frac{\phi V_{12}}{c^2(E_2 - E_1)} \Psi_1(r) \right\}, \tag{67}$$

where the wave function is normalized with the same accuracy:

$$\int [\Psi^1(r,t)]^* \Psi^1(r,t) d^3 r = 1 + O\left(\frac{\phi^2}{c^4}\right). \tag{68}$$

Taking into account that we consider macroscopic coherent ensemble of superposition of quantum states (59) and (67), it is easy to calculate the expectation value of energy per one electron in gravitational field for wave function (67):

$$\langle E \rangle = m_e c^2 \left(1 + \frac{\phi}{c^2}\right) + \frac{(E_1 + E_2)}{2c^2} \left(1 + \frac{\phi}{c^2}\right) + V_{12}\frac{\phi}{c^2}. \tag{69}$$

Note that the first term and the second one are expected. On the other hand, the last term contains contribution to the weight of macroscopic coherent ensemble from the virial term (17) and breaks the equivalence of passive gravitational mass and energy for quantum superposition of stationary state.

Note that so far we have considered macroscopic coherent ensemble of superposition of stationary wave functions, which is characterized by constant difference of phases, $\alpha = 0$, between the first and the second quantum states. If we introduce

more general macroscopic coherent ensemble,

$$\Psi(r,t) = \frac{1}{\sqrt{2}} \exp\left(\frac{-im_e c^2 t}{\hbar}\right) \left[\exp\left(\frac{-iE_1 t}{\hbar}\right) \Psi_1(r) + \exp(i\alpha) \exp\left(\frac{-iE_2 t}{\hbar}\right) \Psi_2(r)\right], \tag{70}$$

the expectation value of energy in the gravitational field (3) is changed,

$$\langle E \rangle = m_e c^2 \left(1 + \frac{\phi}{c^2}\right) + \frac{(E_1 + E_2)}{2c^2} \left(1 + \frac{\phi}{c^2}\right) + V_{12} \cos\alpha \frac{\phi}{c^2}. \tag{71}$$

It is important that Eqs. (69) and (71) directly demonstrate the breakdown of the equivalence between gravitational mass and energy at macroscopic level for coherent ensemble of superposition of stationary quantum state. On the other hand, for the noncoherent ensembles, phase $\alpha$ is not fixed in Eq. (71) and, thus, the last virial terms quickly averages to zero.

## 6. Inequivalence Between Active Gravitational Mass and Energy at Macroscopic Level

In this section, we review our results [5,10], where we showed that active gravitational mass and energy were inequivalent to each other at macroscopic level for coherent ensembles of quantum superpositions of stationary states.

### 6.1. *Active gravitational mass in classical physics*

Here, we determine electron active gravitational mass in a classical model of a hydrogen atom, which takes into account electron kinetic and potential energies [7]. More specifically, we consider a particle with small bare mass $m_e$, moving in the Coulomb electrostatic field of a heavy particle with bare mass $m_p \gg m_e$. Our task is to find gravitational potential at large distance from the atom, $R \gg r_B$, where $r_B$ is the the so-called Bohr radius (i.e. effective "size" of a hydrogen atom). Below, we use the so-called weak field gravitational theory [1,7], where the post-Newtonian gravitational potential can be represented as [5,10]

$$\phi(R,t) = -G\frac{m_p + m_e}{R} - G\int \frac{\Delta T^{\rm kin}_{\alpha\alpha}(t,\mathbf{r}) + \Delta T^{\rm pot}_{\alpha\alpha}(t,\mathbf{r})}{c^2 R} d^3\mathbf{r}, \tag{72}$$

where $\Delta T^{\rm kin}_{\alpha\beta}(t,\mathbf{r})$ and $\Delta T^{\rm pot}_{\alpha\beta}(t,\mathbf{r})$ are contributions to stress–energy tensor density, $T_{\alpha\beta}(t,\mathbf{r})$, due to kinetic and the Coulomb potential energies, respectively. We point out that, in Eq. (72), we disregard all retardation effects. Thus, in the above-discussed approximation, electron active gravitational mass is equal to

$$m^a_e = m_e + \frac{1}{c^2}\int [\Delta T^{\rm kin}_{\alpha\alpha}(t,\mathbf{r}) + \Delta T^{\rm pot}_{\alpha\alpha}(t,\mathbf{r})] d^3\mathbf{r}. \tag{73}$$

Let us calculate $\Delta T^{\rm kin}_{\alpha\alpha}(t,\mathbf{r})$, using the standard expression for stress–energy tensor density of a moving relativistic point mass [1,2]:

$$T^{\alpha\beta}_{\rm kin}(\mathbf{r},t)=\frac{m_e v^{\alpha}(t)v^{\beta}(t)}{\sqrt{1-\dfrac{v^2(t)}{c^2}}}\delta^3[\mathbf{r}-\mathbf{r}_e(t)], \tag{74}$$

where $v^{\alpha}$ is a four-velocity, $\delta^3(\ldots)$ is the three-dimensional Dirac $\delta$-function, and $\mathbf{r}_e(t)$ is a three-dimensional electron trajectory.

From Eqs. (73) and (74), it directly follows that

$$\Delta T^{\rm kin}_{\alpha\alpha}(t)=\int \Delta T^{\rm kin}_{\alpha\alpha}(t,\mathbf{r})d^3\mathbf{r}=\frac{m_e[c^2+v^2(t)]}{\sqrt{1-\dfrac{v^2(t)}{c^2}}}-m_e c^2. \tag{75}$$

Note that, although calculations of the contribution from potential energy to stress–energy tensor are more complicated, they are straightforward and can be done by using the standard formula for stress–energy tensor of electromagnetic field [2],

$$T^{\mu\nu}_{\rm em}=\frac{1}{4\pi}\left[F^{\mu\alpha}F^{\nu}_{\ \alpha}-\frac{1}{4}\eta^{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}\right], \tag{76}$$

where $\eta_{\alpha\beta}$ is the Minkowski metric tensor, $F^{\alpha\beta}$ is the so-called tensor of electromagnetic field [2]. In this paper, we use approximation, where we do not take into account magnetic field and keep only the Coulomb electrostatic field. In this approximation, we can simplify Eq. (76) and obtain from it the following expression:

$$\Delta T^{\rm pot}_{\alpha\alpha}(t)=\int \Delta T^{\rm pot}_{\alpha\alpha}(t,\mathbf{r})d^3\mathbf{r}=-2\frac{e^2}{r(t)}, \tag{77}$$

where $e$ is the electron charge. As directly follows from Eqs. (75) and (77), electron active gravitational mass can be represented in the following way:

$$m^a_e=\frac{\left[\dfrac{m_e c^2}{\left(1-\dfrac{v^2}{c^2}\right)^{1/2}}-\dfrac{e^2}{r}\right]}{c^2}+\frac{\left[\dfrac{m_e v^2}{\left(1-\dfrac{v^2}{c^2}\right)^{1/2}}-\dfrac{e^2}{r}\right]}{c^2}. \tag{78}$$

We note that the first term in Eq. (78) is the expected one. Indeed, it is the total energy contribution to the mass, whereas the second term is the so-called relativistic virial one [12]. It is important that it depends on time. Therefore, in classical physics, active gravitational mass of a composite body depends on time too. Nevertheless, in this situation, it is possible to introduce averaged over time electron active gravitational mass. This procedure results in the expected equivalence between averaged over time active gravitational mass and energy [7]:

$$\langle m^a_e\rangle_t=\frac{\left\langle\dfrac{m_e c^2}{\left(1-\dfrac{v^2}{c^2}\right)^{1/2}}-\dfrac{e^2}{r}\right\rangle_t}{c^2}+\frac{\left\langle\dfrac{m_e v^2}{\left(1-\dfrac{v^2}{c^2}\right)^{1/2}}-\dfrac{e^2}{r}\right\rangle_t}{c^2}=m_e+\frac{E}{c^2}. \tag{79}$$

We point out that, in Eq. (79), the averaged over time virial term is zero due to the classical virial theorem. It is easy to show that for nonrelativistic case, our Eqs. (78) and (79) can be simplified to

$$m_e^a = m_e + \frac{\left(\dfrac{m_e v^2}{2} - \dfrac{e^2}{r}\right)}{c^2} + \frac{\left(2\dfrac{m_e v^2}{2} - \dfrac{e^2}{r}\right)}{c^2}, \tag{80}$$

and

$$\langle m_e^a \rangle_t = m_e + \frac{\left\langle \dfrac{m_e v^2}{2} - \dfrac{e^2}{r} \right\rangle_t}{c^2} + \frac{\left\langle 2\dfrac{m_e v^2}{2} - \dfrac{e^2}{r} \right\rangle_t}{c^2} = m_e + \frac{E}{c^2}. \tag{81}$$

### 6.2. *Active gravitational mass in quantum physics*

In this section, we consider the so-called semiclassical theory of gravity [14], where, in the Einstein's field equation, gravitational field is not quantized but the matter is quantized:

$$R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} = \frac{8\pi G}{c^4} \langle \hat{T}_{\mu\nu} \rangle. \tag{82}$$

Here, $\langle \hat{T}_{\mu\nu} \rangle$ is the expectation value of quantum operator, corresponding to the stress–energy tensor. To make use of Eq. (82), we have to rewrite Eq. (80) for electron active gravitational mass using momentum, instead of velocity. Then, we can quantize the obtained result:

$$\hat{m}_e^a = m_e + \frac{\left(\dfrac{\hat{\mathbf{p}}^2}{2m_e} - \dfrac{e^2}{r}\right)}{c^2} + \frac{\left(2\dfrac{\hat{\mathbf{p}}^2}{2m_e} - \dfrac{e^2}{r}\right)}{c^2}. \tag{83}$$

Note that Eq. (83) represents electron active gravitational mass operator. As directly follows, the expectation value of electron active gravitational mass can be written as

$$\langle \hat{m}_e^a \rangle = m_e + \frac{\left\langle \dfrac{\hat{\mathbf{p}}^2}{2m_e} - \dfrac{e^2}{r} \right\rangle}{c^2} + \frac{\left\langle 2\dfrac{\hat{\mathbf{p}}^2}{2m_e} - \dfrac{e^2}{r} \right\rangle}{c^2}, \tag{84}$$

where third term is the virial one.

#### 6.2.1. *Equivalence of the expectation values of active gravitational mass and energy for stationary quantum states*

Now, we consider a macroscopic ensemble of hydrogen atoms with each of them being in the $n$th energy level. For such ensemble, the expectation value of the mass (83) is

$$\langle \hat{m}_e^a \rangle = m_e + \frac{E_n}{c^2}. \tag{85}$$

In Eqs. (84) and (85), we take into account that the expectation value of the virial term is equal to zero in stationary quantum states due to the quantum virial theorem [12]. Thus, we can make the following important conclusion: in stationary quantum states, active gravitational mass of a composite quantum body is equivalent to its energy at a macroscopic level [5,10].

### 6.2.2. *Inequivalence between active gravitational mass and energy for macroscopic coherent ensemble of quantum superpositions of stationary states*

Below, we introduce the simplest macroscopic coherent ensemble of quantum superpositions of the following stationary states in a hydrogen atom,

$$\Psi(r,t) = \frac{1}{\sqrt{2}} \exp\left(-i\frac{m_e c^2 t}{\hbar}\right) \left[\Psi_1(r) \exp\left(-i\frac{E_1 t}{\hbar}\right) + \exp(i\alpha)\Psi_2(r) \exp\left(-i\frac{E_2 t}{\hbar}\right)\right], \tag{86}$$

where $\Psi_1(r)$ and $\Psi_2(r)$ are the normalized wave functions of the ground state (1S) and first excited state (2S), respectively. We stress that it is possible to create the coherent superposition, where $\alpha = \text{const}$ for all macroscopic ensemble, by using lasers. It is easy to show that the superposition (86) corresponds to the following constant expectation value of energy in the absence of gravitational field,

$$\langle E \rangle = m_e c^2 + \frac{E_1 + E_2}{2}. \tag{87}$$

Nevertheless, as seen from Eq. (84), the expectation value of electron active gravitational mass operator for the wave function (86) is not constant and exhibits time-dependent oscillations:

$$\langle \hat{m}_e^a \rangle = m_e + \frac{E_1 + E_2}{2c^2} + \frac{V_{1,2}}{c^2} \cos\left[\alpha + \frac{(E_1 - E_2)t}{\hbar}\right], \tag{88}$$

where $V_{1,2}$ is matrix element of the virial operator,

$$V_{1,2} = \int \Psi_1(r) \left(2\frac{\hat{\mathbf{p}}^2}{2m_e} - \frac{e^2}{r}\right) \Psi_2(r)\, d^3\mathbf{r}, \tag{89}$$

between the above-mentioned two stationary quantum states. It is important that the oscillations (88) and (89) directly demonstrate breakdown of the equivalence between the expectation values of active gravitational mass and energy for coherent quantum superpositions of stationary states [5,10]. We pay attention to the fact that such quantum time-dependent oscillations are very general and are not restricted by the case of a hydrogen atom. They are of a pure quantum origin and do not have classical analogs.

### 6.3. *Experimental aspects*

In this section, we suggest an idealized experiment, which allows to observe quantum time-dependent oscillations of the expectation values of active gravitational mass (88). In principle, it is possible to create a macroscopic ensemble of the coherent quantum superpositions of electron stationary states in some gas with high density. It is important that these superpositions have to be characterized by the feature that each atom (or molecule) has the same phase difference between two wave function components, $\Psi_1(r)$ and $\Psi_2(r)$. In this case, the macroscopic ensemble of the atoms (or molecules) generates gravitational field, which oscillates in time similar to Eq. (88), which, in principle, can be measured. It is important to use such geometrical distributions of the molecules and a test body, where oscillations (88) are "in phase" and, thus, do not cancel each other.

## 7. Summary

In Conclusion, in this paper, we have discussed in detail the breakdown of the equivalence between active and passive gravitational masses of an electron and its energy in a hydrogen atom. We stress that the considered phenomena are very general and are not restricted by atomic physics and the Earth's gravitational field. In other words, the above discussed phenomena exist for any quantum system with internal degrees of freedom and at any gravitational field. In this paper, we also have improved several drawbacks of the original pioneering works.

## Acknowledgments

We are thankful to N. N. Bagmet (Lebed), V. A. Belinski, Steven Carlip, Fulvio Melia, Jan Rafelski, Douglas Singleton, Elias Vagenas and V. E. Zakharov for fruitful and useful discussions.

## References

1. C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W.H. Freeman and Co, New York, USA, 1973).
2. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields*, 4th edn. (Butterworth-Heinemann, Oxford, 2003).
3. M. Zych and C. Brukner, *Nat. Phys.* **14** (2018) 1027.
4. G. Rosi, D'Amico, L. Cacciapuoti, F. Sorrentino, M. Prevedelli, M. Zych, C. Brukner and G. M. Tito, *Nat. Commun.* **8** (2016) 15529.
5. A. G. Lebed, *J. Phys. Conf. Ser.* **738** (2016) 012036.
6. C. W. Misner and P. Putnam, *Phys. Rev.* **116** (1959) 1045.
7. K. Nordtvedt, *Class. Quantum Grav.* **11** (1994) A119.
8. S. Carlip, *Am. J. Phys.* **66** (1998) 409.
9. M. Zyck, L. Rudnicki and I. Pikovski, arXiv:1808.05831v1.

10. A. G. Lebed, *Int. J. Mod. Phys.* **26** (2017) 1730022.
11. A. G. Lebed, Inequivalence between gravitational mass and energy due to quantum effects in general relativity, in *Proc. 15th Marcel Grossmann Meeting on General Relativity*, E. Battistelli, R. T. Jantzen and R. Ruffini (eds.) (World Scientific, Singapore, 2019).
12. D. Park, *Introduction to the Quantum Theory*, 3rd edn. (Dover Publications, New York, USA, 2005).
13. E. Fischbach, B. S. Freeman and W. K. Cheng, *Phys. Rev. D* **23** (1981) 2157.
14. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*, 3rd edn. (Cambridge University Press, Cambridge, UK, 1982).

# Recent results and perspectives on cosmic backgrounds from radio to far-infrared*

Burigana, Carlo;[1,2,3,a] De Zotti, Gianfranco;[4,b] Fialkov, Anastasia;[5,6,c]
Qadir, Asghar;[7,d] Trombetti, Tiziana;[1,e] Bonato, Matteo;[1,f]
Negrello, Mattia;[8,g] Norgaard-Nielsen, Hans Ulrik;[9,h] Tahir, Noraiz[7,i]

[1] *INAF–IRA, Via Piero Gobetti 101, I-40129 Bologna, Italy* †
[2] *Dipartimento di Fisica e Scienze della Terra, Università degli Studi di Ferrara, Via Giuseppe Saragat 1, I-44122 Ferrara, Italy*
[3] *INFN, Sezione di Bologna, Via Irnerio 46, I-40126, Bologna, Italy*
[4] *INAF–Osservatorio Astronomico di Padova, Vicolo dell'Osservatorio 5, I-35122 Padova, Italy*
[5] *Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, MA 02138, USA*
[6] *Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*
[7] *School of Natural Sciences, National University of Science and Technology, H-12 Islamabad, 46000, Pakistan*
[8] *School of Physics and Astronomy, Cardiff University, The Parade, Cardiff CF24 3AA, UK*
[9] *DTU Space, Elektrovej, DK - 2800 Kgs. Lyngby, Denmark*

[a] *burigana@ira.inaf.it* – [b] *gianfranco.dezotti@inaf.it* – [c] *anastasia.fialkov@gmail.com*
[d] *asgharqadir46@gmail.com* – [e] *trombetti@ira.inaf.it* – [f] *bonateo@yahoo.it*
[g] *NegrelloM@cardiff.ac.uk* – [h] *hunn@space.dtu.dk* – [i] *noraiztahir78637@gmail.com*

Cosmological and astrophysical surveys from radio to far-infrared, in both temperature and polarization, offer a unique view of the Universe properties and of the formation and evolution of its structures. The last release, close to be finalized, of the *Planck* mission results sets the scene for cosmological models and parameters, while the comparison with other types of data sets raises the issue of possible tensions about some parameters, first of all the Hubble constant. At the same time, on the extragalactic side, *Planck* carried out the deepest systematic all-sky survey of SZ galaxy clusters and detected thousands of dusty galaxies and many hundreds of extragalactic radio sources, also allowing us to investigate many specific topics, including molecular hydrogen clouds in galactic halos. The exploitation of future generation of CMB missions and the next radio facilities will allow us to deeply investigate several topics in cosmology and astrophysics, from the existence of primordial gravitational waves to the energy releases in the primeval plasma, from the dawn ages and the epoch of reionization to the formation and evolution of early galaxies and clusters, while a wide set of open astrophysical problems can be studied with future IR missions.

*Keywords*: Cosmology; Background radiations; Radio, microwave; Origin and formation of the Universe; Observational cosmology; Large scale structure of the Universe; Galaxies.

## 1. Introduction

The past and recent surveys from radio to far-infrared have offered fundamental probes to answer to many questions of cosmology and on the formation and evolution of the cosmic structures at different scales and times. In particular, the accurate observations of cosmic microwave background (CMB) properties sanctioned

*Based on talks presented at the Fourteenth Marcel Grossmann Meeting on General Relativity, Rome, July 2015.
†Istituto Nazionale di Astrofisica – Istituto di Radioastronomia, Via Piero Gobetti 101, I-40129 Bologna, Italy

the passage to the so-called precision cosmology, calling for a further progress to unravel the mysteries of the early Universe and of its evolution up to the present epoch and beyond. At the same time, the information contained in the radio sky is of increasing relevance to unveil the complexity of the early stages of stars and galaxies stored in the imprints left in the cosmic background by the 21-cm redshifted line from neutral hydrogen from the dark and dawn ages to the epoch of reionization (EoR), accessible thanks to the enormous improvement of recent and future radio projects and facilities. Sub-millimetre to infrared (IR) observations offer powerful diagnostics for a broad set of astrophysical mechanisms and conditions, also because, due to the reprocessing and the cosmic expansion, a large fraction of the optical and ultraviolet radiation from stars in galaxies comes to us at longer wavelengths. The exploitation of these surveys also revealed the close connection between our comprehension of the global properties of the Universe and the detailed understanding of the astrophysical processes operating in bounded objects and in the diffuse medium at different scales.

The structure of this paper is as follows. In Sect. 2.1 we report on some of the main cosmological results from the *Planck* mission, focussing on the precise determination of cosmological parameters in the $\Lambda$ cold dark matter ($\Lambda$CDM) model and in some extensions, and discuss the issue of possible tensions about some parameters, first of all the Hubble constant. The estimation of a particular parameter combination based on the analysis of the diffuse thermal Sunyaev-Zeldovich (SZ) effect coming from all the galaxies clusters and the related so-called bias problem are discussed. In Sect. 3 we briefly describe recent proposals of future CMB missions and their main scientific purposes. Sect. 4 is dedicated to some specific topics in astrophysical cosmology, namely the study of CMB spectral distortions and of cosmic IR background (CIB) spectrum through the analysis of the cosmic dipole frequency behaviour aimed at constraining the (early and late) thermal and star formation histories, the cosmological evolution of neutral hydrogen at high redshifts through the analysis of the 21-cm redshifted line and the imprints of black holes high mass X-ray binaries microquasars in cosmic backgrounds. In Sect. 5 we move to extragalactic astrophysics. After a brief description of the *Planck* contribution in the field, we illustrate two new promising research paths, the study of early phases of galaxy evolution exploiting the strong gravitational lensing and of the cluster evolution, in relation to the prospects of future CMB experiments. We then report on an investigation of the molecular hydrogen clouds in galactic halos using *Planck* observations of the Doppler asymmetry in CMB radiation. We finally illustrate a wide set of open astrophysical problems that can be investigated with a future IR space mission.

## 2. Cosmological parameters

### 2.1. *Planck results*

The *Planck* satellite[a], launched in 2009, was equipped with a Gregorian dual reflector telescope of 1.5m diameter, whose focal plane was able to collect the microwave photons owing to two wide band cooled receiver arrays, the Low Frequency Instrument (LFI) made up of radiometers operative at 20 K, and the High Frequency Instrument (HFI) composed of bolometers working at 100 mK. The two instruments observed the sky in 9 frequency bands, performing respectively 8 and 5 nearly all sky surveys, LFI at 30, 44 and 70 GHz and HFI at 100, 143, 217, 353, 545, 857 GHz, with telescope optical axis pointing at 85° from the spin axis[1,2]. *Planck* observed the sky from a large Lissajous orbit around the Sun-Earth L2 Lagrange point in order to offer stable observing conditions and minimize contamination from side lobe pick-up of stray radiation from the Sun, Earth, and Moon, scanning the sky while spinning at 1 rpm in almost great circles. The *Planck* full width half maximum (FWHM) resolution ranges from $33.3'$ to $4.3'$ from 30 GHz to 857 GHz, and the final sensitivity per FWHM$^2$ resolution element span $\sim 2-14$ $\mu$K/K in terms of $\delta T/T$ for frequencies $\nu \leq 353$GHz.

Reprocessing LFI and HFI data, the *Planck* Collaboration released maps in temperature at the nine frequency bands and in polarization at the seven polarization sensitive bands up to 353 GHz, and extracted Stokes $T, Q, U$ parameter maps of CMB anisotropies on the whole sky with unprecedented sensitivity and resolution. For many investigations, the anisotropy signals of billions of pixels are compressed in several thousand of numbers, namely the correlators $C_\ell^{XY}$, where $X$ and $Y$ stands for $T, E, B$, at the multipole $\ell \sim 180°/\theta$, $\theta$ being the angular scale of interest. A linear combination of $Q, U$ allows us to define, in Fourier space, the pure polarization angular power spectra (APS), the so-called $E$-modes (gradient component) and $B$-modes (curl component). Together with temperature anisotropies, four APS, $C_\ell^{TT}$, $C_\ell^{EE}$, $C_\ell^{BB}$, $C_\ell^{TE}$, the latter being the temperature polarization cross correlation, were extracted. Under the assumption of random Gaussian fluctuations, the CMB APS $C_\ell^{XY}$ contain all the relevant statistical information that was indeed mainly analyzed through Markov chain Monte Carlo methods ingesting the corresponding theoretical predictions of Boltzmann codes to derive the cosmological parameters characterizing sets of models, marginalizing also over instrument and foreground residual parametrizations[3].

Cosmological results based on *Planck* observations were presented in three (2013, 2015 and 2018) subsequent releases. The temperature and polarization CMB APS are consistent with the standard spatially-flat six-parameter ΛCDM cosmology with a power-law spectrum of adiabatic scalar perturbations (denoted "base ΛCDM").

[a] *Planck* hosted instruments provided by two scientific Consortia funded by European Space Agency (ESA) member states (in particular the lead countries: France and Italy) with contributions from National Aeronautics and Space Administration (NASA), USA, and telescope reflectors provided in a collaboration between ESA and a scientific Consortium led and funded by Denmark.

Table 1. Estimation and 68 % C.L. intervals for the six-parameters of the base ΛCDM model and five derived parameters from the 2018 *Planck* release of CMB APS in $TT$ and low-$\ell$ $EE$ modes (second column) and in $TT$, $EE$ and $TE$ modes in combination with CMB lensing reconstruction and BAO (third column).

| Six-parameters of the base ΛCDM model | | |
|---|---|---|
| $\Omega_{\rm b}h^2$ | $0.02212 \pm 0.00022$ | $0.02242 \pm 0.00014$ |
| $\Omega_{\rm c}h^2$ | $0.1206 \pm 0.0021$ | $0.11933 \pm 0.00091$ |
| $100\theta_{\rm MC}$ | $1.04077 \pm 0.00047$ | $1.04101 \pm 0.00029$ |
| $\tau$ | $0.0522 \pm 0.0080$ | $0.0561 \pm 0.0071$ |
| $\ln(10^{10}A_{\rm s})$ | $3.040 \pm 0.016$ | $3.047 \pm 0.014$ |
| $n_{\rm s}$ | $0.9626 \pm 0.0057$ | $0.9665 \pm 0.0038$ |
| Derived parameters | | |
| $H_0$ [km/s/Mpc] | $66.88 \pm 0.92$ | $67.66 \pm 0.42$ |
| $\Omega_{\rm m}$ | $0.321 \pm 0.013$ | $0.3111 \pm 0.0056$ |
| $\sigma_8$ | $0.8118 \pm 0.0089$ | $0.8102 \pm 0.0060$ |
| $10^9 A_{\rm s}e^{-2\tau}$ | $1.884 \pm 0.014$ | $1.881 \pm 0.010$ |
| Age of the Universe [Gyr] | $13.830 \pm 0.037$ | $13.787 \pm 0.020$ |

Cosmological parameter results[4] from the final full-mission *Planck* measurements[5–7] have been presented in the 2018 release, combining information from the temperature and polarization maps and the lensing reconstruction and including the information from baryonic acoustic oscillations (BAO) (see Tab. 1).[b]

Compared to the 2015 results[3], the improved measurement of large-scale polarization, foreground cleaning and likelihood analysis allow us to measure the value of $\tau$ (see also Sects. 4.1 and 4.2) with higher precision, suggesting a mid-point reionization redshift in the range $\simeq 7.1 - 7.8$ only slightly dependent on the explored data set (namely, $z_{\rm re} = 7.50 \pm 0.82$ or $z_{\rm re} = 7.8 \pm 0.7$ with reference to the cases in Tab. 1), in agreement with astrophysical observations of quasar absorption lines and supporting models in which reionization happened relatively fast and late. Extensions to more general reionization models indicate that the results found for the other cosmological parameters are essentially insensitive to uncertainties in the reionization history.

The refined analysis of the small-scale polarization provided a more robust determination of many parameters, the uncertainties in the treatment of various types of residuals affecting their estimation only at the $0.5\,\sigma$ level. A remarkable, good consistency of the six parameters of the base ΛCDM model was found when considering

[b] According to ref.[4] $h = H_0/(100\,\mathrm{km/s/Mpc})$, $H_0$ being the Hubble constant; $\Omega_{\rm b}$, $\Omega_{\rm c}$ and $\Omega_{\rm m}$ are respectively the baryon, cold dark matter (CDM) and (global, non-relativistic) matter density parameters; $\tau$ is the Thomson optical depth of cosmological reionization; $\theta_{\rm MC}$ is an approximation to the acoustic scale angle $\theta_* \equiv r_*/D_{\rm M}$ where $r_*$ is the comoving sound horizon at recombination quantifying the distance the photon-baryon perturbations can influence, and $D_{\rm M} = (1+z)D_{\rm A}$, where $D_{\rm A}$ is the usual angular diameter distance, is the comoving angular diameter distance that maps this distance into an angle on the sky; $A_{\rm s}$ and $n_{\rm s}$ are the amplitude and spectral index of the power spectrum of primordial scalar perturbations; $\sigma_8$ is the matter fluctuation amplitude on a scale of $8h^{-1}$Mpc.

polarization, temperature, and lensing data, separately and in combination. Even considering many typical extensions of this model, the results found for these parameters remain substantially stable, with only a weak dependence on the adopted cosmological model and only slightly increased uncertainties in their estimation.

### 2.2. *Model extensions and the tension on the Hubble constant*

Various types of extensions of the base ΛCDM model were considered in the *Planck* 2018 release, finding no compelling evidence supporting extensions to the base ΛCDM model.

Relativistic particle influence CMB anisotropies, mainly at intermediate and small angular scales. *Planck* data in combination with BAO measurements (and considering single-parameter extensions) constrain the effective extra relativistic degrees of freedom to be $N_{\rm eff} = 2.99 \pm 0.17$, in agreement with the Standard Model prediction $N_{\rm eff} = 3.046$, and tightly constrain the neutrino mass to $\sum m_\nu < 0.12$ eV.

Standard big-bang nucleosynthesis predictions for the helium and deuterium abundances for the base ΛCDM cosmology are in excellent agreement with observations.

The CMB spectra prefer higher lensing amplitudes than predicted in base ΛCDM model at over $2\,\sigma$, which pulls some parameters that affect the lensing amplitude away from the ΛCDM model; however, this is not supported by the lensing reconstruction[8] or (in models that also change the background geometry) BAO data.

The joint constraint with BAO measurements on spatial curvature is consistent with a flat universe, $\Omega_K = 0.001 \pm 0.002$, while combining with Type Ia supernovae (SNe), the dark energy (DE) equation of state parameter is measured to be $w_0 = -1.03 \pm 0.03$, with no significant deviations from an interpretation in terms of a cosmological constant.

At the scales measured by *Planck* the primordial spectrum has been found to be well described by a pure power-law. The joint analysis of CMB polarization from *Planck*, BICEP2 and Keck Array data, in combination with measurements from BAO, sets the ratio between primordial tensor and scalar perturbations to $r < 0.07$ at the wavenumber $k = 0.002$, constraining inflationary models[9] and setting the scene for future projects (see Sect. 3).

The *Planck* base ΛCDM results are in good agreement with BAO, SNe, and some galaxy lensing observations, but in slight tension with the Dark Energy Survey's combined-probe results including galaxy clustering (which prefer lower fluctuation amplitudes or matter density parameters).

The value of the $H_0$ for the *Planck* base ΛCDM cosmology results in substantial tension, at $\simeq 3.6\,\sigma$ level, with the local determination through Type Ia SNe by ref.[10] and even more, at $\simeq 4.4\,\sigma$ level, with the most recent one by ref.[11] which implies a higher value, $H_0 = (74.03 \pm 1.5)$km/s/Mpc, while the *Planck* measurement is in excellent agreement with independent inverse-distance ladder measurements using BAO, SNe, and element abundance results. Model extensions have been

investigated in ref.[4] to resolve the tension with the value of $H_0$ from Type Ia SNe, including, other than those mentioned above, also modified gravity, DE models, recombination history, primordial scalar perturbations with running spectral index possibly combined with tensor modes, but none of these models convincingly resolves this problem.

An attempt to solve this tension involves the observational evidence of the existence of local radial inhomogeneities extending in different directions up to a redshift $z \simeq 0.07$ that could affect a significant fraction ($\sim 40\%$) of the Cepheids used for SNe calibration. Different methods used in ref.[12] to compute the effects of these inhomogeneities on the low-redshift luminosity and angular diameter distance, using an exact solution of the Einstein's equations, linear perturbation theory and a low-redshift expansion, point towards the relevance of the non-relativistic Doppler redshift correction (proportional to the volume averaged density contrast and to the comoving distance from the center). Using a new direct relation between the luminosity distance and the monopole of the density contrast, which does not involve any metric perturbation, the monopole of the density field was reconstructed through a new inversion method from the deviations of the redshift uncorrected observed luminosity distance respect to the ΛCDM prediction based on cosmological parameters obtained from large scale observations. The method suggests the existence of inhomogeneities effects not previously taken into account in density field maps at $z \leq 0.06$, used to obtain the peculiar velocity for redshift correction, i.e. at a scale not large enough for detecting inhomogeneities up to $z = 0.07$, underlying the relevance of accurately normalizing the density field respect to the average large scale density of the Universe. Of course, the CMB high-$z$ measure of $H_0$ are insensitive to local inhomogeneities.

Other methods explore observations at intermediate redshifts. For example, in ref.[13] the joint use of four different observables (SZ effect combined with X-ray surface brightness of galaxy clusters, ages of old high-redshift galaxies, observational measurements of the Hubble parameter, and BAO peak) proposed in ref.[14] was revised reconsidering the sample of galaxy clusters and performing a new characterization of systematic uncertainties, and a value of $H_0$ fully consistent with *Planck* results was inferred.

Forthcoming cosmological surveys in various bands, including those aimed at filling the gap between low and high redshifts measures of $H_0$, will allow us to firmly understand if the tension comes from observational or data analysis issues or reflects instead the imprint of physics beyond ΛCDM possibly involving, e.g., specific neutrino sector[15] or early DE[16] models.

### 2.3. *Constraints from thermal SZ power spectrum*

The *Planck* mission performed multifrequency surveys of the SZ effect, the Second *Planck* catalogue of SZ galaxy clusters (PSZ2) being the largest and deepest all-sky selected galaxy clusters archive, enclosing 1653 sources[17], $\sim 73\%$ of them with

counterparts in external data sets (see also ref.[5]). Remarkably, the *Planck* surveys allowed us to study the diffuse thermal SZ (tSZ) effect coming from all the galaxies clusters in the observable Universe through the analysis of the fluctuations of the Comptonization parameter, $y$, representing an important tool to probe clusters physics and evolution as well as cosmology[18]. In particular, the angular power spectrum of the tSZ[19], $C_\ell^{\rm tSZ} = \int {\rm d}z[{\rm d}V/({\rm d}z{\rm d}\Omega)] \int {\rm d}M({\rm d}n/{\rm d}M)\,|y_\ell\,(M,z)|^2$, depending on the halo mass function, ${\rm d}n/{\rm d}M$, and the squared amplitude of two dimensional Fourier transform of the electron pressure profile, $|y_\ell|^2$, integrated over the redshift, with the differential volume element, and over the masses, provides information on certain combinations of cosmological parameters[20].

At large scales ($\ell \lesssim 10^3$) $C_\ell^{\rm tSZ}$ reflects the Poisson random distribution of the sources and the evolution of the halo mass function related to the amplitude of matter clustering, i.e., to $\sigma_8$. The electron pressure profile is instead relevant at all scales, affecting the $C_\ell^{\rm tSZ}$ amplitude and, at smaller scales, the $C_\ell^{\rm tSZ}$ shape, because of the halos morphology, resulting in a peak in $\ell(\ell+1)C_\ell^{\rm tSZ}$ at $\ell \simeq 3000$. The determination of the electron pressure profile dependence on the cluster mass is based on the relation between the SZ flux and the mass. The latter is typically inferred from X-ray observations and suffers from a relevant mismatch with that derived from lensing observations, when available, that are in principle sensitive to the true cluster mass, thus calling for the introduction of a bias parameter, $B = (1-b)^{-1}$, in the formula of the electron pressure profile, $(1-b)$ being the usual mass bias factor accounting for any difference between the X-ray mass proxies used to establish the scaling relations and the true (halo) mass.

At $\ell \lesssim 10^3$, $C_\ell^{\rm tSZ}$ is found to be mainly dependent on the parameter combination $F = \sigma_8\,(\Omega_{\rm m}/B)^{0.40}\,h^{-0.21}$ through the approximate relation $\ell(\ell+1)C_\ell^{\rm tSZ} \propto F^{8.1}$, and $F$ can be then estimated from data with likelihood methods including also proper parametrizations of noise and the other ("foreground") contributions to $C_\ell$. The value $F = 0.460 \pm 0.012$ at 68% CL retrieved in this way can be compared with that of $FB^{0.40} = 0.568 \pm 0.015$ at 68% CL derived from *Planck* CMB temperature and low-$\ell$ polarization anisotropy for the base $\Lambda$CDM model in the second release[3] (or only slightly changed when using the values in Tab. 1), thus requiring $B = 1.71 \pm 0.17$ at 68% CL, i.e. $1 - b \simeq 0.58$, consistent with results from SZ number counts[21], i.e. the mass estimated from X-rays is about 40% lower than their true mass[22]. Including deviations from hydrostatic equilibrium by non-thermal pressure could partially alleviate this problem, while forthcoming X-ray measurements by, e.g., eROSITA, will greatly reduce uncertainties in X-ray data and, together with next SZ observations, e.g., from NIKA2, could solve this bias problem, making galaxies clusters competitive to constrain some cosmological parameters, or, maybe, reveal new physics[20].

## 3. Future CMB missions

Recent proposals of CMB missions are based on concepts and designs aimed at improving the measure of polarization anisotropies and at minimizing the impact of foregrounds using a large number of frequency channels (see also the CM1 session, and related contributions in these proceedings, for their complementarities with ground-based and balloon-borne projects).

The Primordial Inflation Explorer[23] (PIXIE) is an Explorer-class mission submitted to NASA in 2011 aimed at the mapping with high signal to noise ratio in absolute intensity and linear polarization the CMB and the diffuse astrophysical foregrounds over the full sky, from 30 GHz to 6 THz (1 cm to 50 $\mu$m) with degree resolution. It is designed to perform absolute spectroscopy measurements and to simultaneously observe the large-scale CMB B-modes. It will provide crucial constraints on Universe ionization history, on CIB spectrum and anisotropies. Great hopes on CMB spectral distortion parameters (see also the CM3 session and related contributions in these proceedings) are expected from PIXIE, thanks to its absolute calibration of the CMB temperature measure with a precision about $10^3$ times better than COBE/FIRAS.

LiteBIRD[24] is a degree resolution Lite (Light) Satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection at the extremely early Universe proposed to the Japan Aerospace Exploration Agency (JAXA) in 2015. It is a highly-targeted mission, with the primary aim of the B-modes detection at the level of $r \sim 10^{-3}$, about one order of magnitude lower than that achievable from the ground, performed in combination with ground based, high-resolution telescopes helping the accurate subtraction of the lensing contribution to B-mode.

The Cosmic ORigins Explorer (CORE)[25], submitted to the ESA in 2016 in response to a call for future medium-sized space mission proposals for the M5 launch opportunity of ESA Cosmic Vision programme, is a satellite dedicated to microwave polarization. CORE is targeted to provide and exploit definitive maps of CMB polarization anisotropies at large and medium angular scales, with a suppression of all systematic effects at an extremely accurate level over 19 frequency channels spanning the 60-600 GHz range, relevant to characterize astrophysical foreground emissions.

The Probe of Inflation and Cosmic Origin (PICO)[26] is currently one of the 8 Probe-Scale – \$ 400M – \$ 1000 M – space missions proposed to NASA whose study is being funded. The concept and performance of PICO are similar to those of CORE, but with 21 frequency bands in the 21-799 GHz range, i.e. a larger frequency coverage for a better control of foreground signals and astrophysical studies.

CMB Bharat[c] is a satellite, proposed in response to the Announcement of Opportunity for future Astronomy missions by the Indian Space Research Organisation

[c] http://cmb-bharat.in/

(ISRO), dedicated to microwave polarization with concept and performance similar to those of CORE but with a proposed frequency coverage up to about 900 GHz to further improve CIB science and foreground mitigation, and, possibly, equipped with a dedicated spectrometer to perform absolute spectrum measurements and useful also for high level calibration of the differential polarization signal.

Remarkably, the next generation of CMB missions, and, if funded, a large mission, as, e.g., in the view proposed for the Polarized Radiation Imaging and Spectroscopy Mission (PRISM)[27], far from being limited to the scientific outcomes driven by their main goals, will allow us to significantly improve the current results in parameter estimation and large scale structure, to characterize a wide set of cosmological model extensions, to answer fundamental physics questions, and, in the case of a detection of primordial gravitational waves, to determine the shape of their spectrum with great implications for our understanding of the early Universe (see, e.g., refs.[28,29] for detailed forecasts based on CORE specifications), offering at the same time exquisite all-sky surveys for astrophysical studies, both Galactic and extragalactic (see Sect. 5.1).

## 4. Astrophysical cosmology

### 4.1. *Thermal and star formation history from cosmic dipoles*

The analysis of cosmic dipoles is crucial in cosmology, being linked with the isotropy and homogeneity of the Universe at the largest scales. The observed dipole comes from different contributions, of both cosmological and astrophysical (extragalactic and Galactic) origin. The frequency spectral behaviour of the dipole effect due to the observer motion with respect to the CMB rest frame offers the opportunity to be investigated to constrain the CMB and CIB frequency spectra thanks to future missions, described in Sect. 3, without requiring an independent absolute calibration (see Trombetti, these proceedings).

#### 4.1.1. *CMB and CIB spectra: from monopoles to dipoles*

The precise interpretation of tiny CMB spectral distortions predicted at different cosmic times provides information on cosmological and astrophysical processes, some of them unavoidable, otherwise unexplored. The analysis of CIB spectrum, still not well known, can provide instead a better understanding of the dust-obscured star-formation phase of galaxy evolution. Since FIRAS, no remarkable improvements have been achieved in the knowledge of CMB spectrum at $\nu \gtrsim 30$ GHz, while, in spite of its absolute calibration precision of 0.57 mK, the FIRAS characterization CIB amplitude and shape still presents a substantial uncertainty.

At early times, a Bose-Einstein (BE) like distorted spectrum with a positive (dimensionless) chemical potential, $\mu$, is produced by the dissipation of primordial perturbations at small scales[30,31]. Moreover, the matter temperature faster decrease with respect to that of radiation temperature in an expanding Universe produces

a spectral distortion mainly characterized by a BE-like spectrum with a negative chemical potential. The chemical potential quantifies the fractional energy, $\Delta\varepsilon/\varepsilon_i$, exchanged in the plasma during the interaction; for small distortions, $\mu \simeq 1.4\Delta\varepsilon/\varepsilon_i$. The photon occupation number of the BE spectrum is $\eta_{BE} = 1/(e^{x_e+\mu} - 1)$, where $x_e = x/\phi(z)$, $\phi(z) = T_e(z)/T_{CMB}(z)$, being $T_e(z)$ the electron temperature, $T_{CMB} = T_0(1 + z)$ and $x = h\nu/(kT_{CMB})$ a redshift independent, dimensionless frequency. The electron heating during the cosmological reionization associated with the early stages of structure and star formation is responsible for late type distortions, described, at high frequency, by a Comptonization spectrum[32] with Comptonization parameter $u(t) = \int_{t_i}^{t}[(\phi - \phi_i)/\phi](k_B T_e/m_e c^2) n_e \sigma_T c dt$. For small energy injections and integrating over the relevant epochs $u \simeq (1/4)\Delta\varepsilon/\varepsilon_i$. CMB spectrum studies can also constrain many types of non-standard processes, i.e. non evaporating BH spin, small scale magnetic fields power spectra, vacuum energy density and particle decay.

The analytic form of the CIB spectrum, observed at present time, is[33] $\eta_{CIB} = [c^2/(2h\nu^3)]I_{CIB}(\nu) = I_0\,[k_B T_{CIB}/(h\nu_0)]^{k_F}\, x_{CIB}^{k_F}/(e^{x_{CIB}} - 1)$, with $T_{CIB} = (18.5 \pm 1.2)$ K, $x_{CIB} = h\nu/k_B T_{CIB} = 7.78(\nu/\nu_0)$, $\nu_0 = 3 \times 10^{12}$ Hz and $k_F = 0.64 \pm 0.12$. Here $I_0$ sets the CIB spectrum amplitude, its best-fit value being $(1.3 \pm 0.4) \times 10^{-5}$. Indeed, the direct determination of the CIB spectrum is not trivial, requiring absolute intensity measurements and being limited by foreground signals.

A relative velocity between an observer and the CMB rest frame induces a dipole (the so called $\ell = 1$ anisotropy) in the CMB sky temperature through the Doppler effect, dominated by the velocity of the Solar System, $\vec{\beta_S} = \vec{v_S}/\vec{c}$, with respect to the CMB (Solar dipole), with a seasonal modulation due to the velocity of the Earth or the satellite, $\vec{\beta_o}$, with respect to the Sun (orbital dipole). Neglecting the orbital dipole (useful for calibration aims), we will denote with $\vec{\beta}$ the relative velocity of the Solar dipole.

The dipole amplitude is directly proportional to the first (logarithmic) derivative with respect to the frequency of the photon occupation number[34]. More in general, the full pattern at a given observational frequency, $\nu_{obs}$, can be computed from the photon distribution function, $\eta^{BB,dist}$, for the assumed type of distortions (BB, CIB, BE or Comptonization (C)) at the frequency $\nu_{obs}$ multiplied by $(1-\hat{n}\cdot\vec{\beta})/(1-\beta^2)^{1/2}$ to account for all the possible sky directions with respect to the observer peculiar velocity[35]. The observed signal map in thermodynamic temperature is given by $T_{therm}^{BB/dist}(\nu, \hat{n}, \vec{\beta}) = xT_0/\ln(1/(\eta(\nu, \hat{n}, \vec{\beta})^{BB,dist} + 1))$, where $\eta(\nu, \hat{n}, \vec{\beta}) = \eta(\nu')$ with $\nu' = \nu((1 - \hat{n}\cdot\vec{\beta})/(1 - \beta^2)^{1/2}$. Decomposing the maps into spherical harmonics and reproducing them from the $a_{\ell m}$ up to a desired multipole $\ell_{max}$, it is possible to compute the signal map for each $\ell$: at the accuracies of the discussed projects, it is found to be important for the dipole, possibly appreciable for the quadrupole and of strongly decreasing relevance at higher multipoles.

4.1.2. *Results and perspectives*

Detailed simulations allow us to quantify the sensitivity of a given project to measure spectral distortion parameters and CIB amplitude in the presence of only ideal noise or including also potential (calibration and foreground) residuals through $\Delta\chi^2$ analyses based on the comparison of the adopted input sky with different model assumptions, possibly applying suitable sky masks like those publicly available from the *Planck* Legacy Archive (PLA)[d]. Calibration errors and foreground residuals can be modeled generating random fluctuations, depending also on sky signal, parametrizing their overall fractional levels at a certain pixel scale with two key parameters (namely $E_{\rm cal}$ and $E_{\rm for}$).

The joint exploitation of all frequency channels (assumed independent from each other) and of the combinations of the maps differences from pairs of frequency bands is found to be particularly advantageous, being sensitive to both amplitude and spectral shape of the considered dipole signal, while the use of a suitable mask can improve parameters reconstruction. Tab. 2 summarizes the main results, obtained for the CORE specifications, expressed in terms of a factor characterizing the improvement with respect to FIRAS. Even under pessimistic assumptions, a clear improvement in the CIB spectrum amplitude and a marginal detection of the energy release associated to astrophysical reionization models are foreseen, while improvements in foreground mitigation and calibration will allow us to derive more

Table 2. Predicted improvement in the recovery of CIB spectrum amplitude and the CMB distortion parameters with respect to FIRAS for different calibration and foreground residual assumptions. "P76" stands for the *Planck* common mask (in temperature) of the PR2–2015 release covering about 76% of the sky and "P76ext" denotes its extension that excludes all the pixels at $|b| \leq 30^\circ$. When not stated, all values refer to $E_{\rm cal}$ and $E_{\rm for}$ at $N_{\rm side} = 64$; the values in round brackets possibly refer to the corresponding frequency range (in GHz). From ref.[35] .

| | $E_{\rm cal}$ (%) | $E_{\rm for}$ (%) | CIB amplitude | BE | C |
|---|---|---|---|---|---|
| Ideal case, all sky | - | - | $\simeq 4.4 \times 10^3$ | $\simeq 10^3$ | $\simeq 6.0 \times 10^2$ |
| All sky | $10^{-4}$ | $10^{-2}$ | $\simeq 15$ | $\simeq 42$ | $\simeq 18$ |
| P76 | $10^{-4}$ | $10^{-2}$ | $\simeq 19$ | $\simeq 42$ | $\simeq 18$ |
| P76ext | $10^{-2}$ | $10^{-2}$ | $\simeq 17$ | $\sim 4$ | $\sim 2$ |
| P76ext | $10^{-4}$ | $10^{-2}$ | $\simeq 22$ | $\simeq 47$ | $\simeq 21$ |
| P76ext | $10^{-4}$ | $10^{-3}$ | $\simeq 2.1 \times 10^2$ | $\simeq 2.4 \times 10^2$ | $\simeq 1.1 \times 10^2$ |
| P76ext | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-2}$ | $\simeq 19$ | $\simeq 26$ | $\simeq 11$ |
| P76ext | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-3}$ | $\simeq 48$ | $\simeq 35$ | $\simeq 15$ |
| P76ext, $N_{\rm s} = 128$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-2}$ | $\simeq 38$ | $\simeq 51$ | $\simeq 23$ |
| P76ext, $N_{\rm s} = 128$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-3}$ | $\simeq 43$ | $\simeq 87$ | $\simeq 39$ |
| P76ext, $N_{\rm s} = 256$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-2}$ | $\simeq 76$ | $\simeq 98$ | $\simeq 44$ |
| P76ext, $N_{\rm s} = 256$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-3}$ | $\simeq 85$ | $\simeq 1.6 \times 10^2$ | $\simeq 73$ |

[d]http://pla.esac.esa.int/pla/

precise results or constraints on dissipation mechanisms occurring at any epoch, including primeval processes.

### 4.2. *21-cm signal of neutral hydrogen from high cosmological redshifts*

The evolution history of the neutral Universe in the first few hundred million years after the Big Bang, as well as the processes of the primordial star formation and reionization, remains poorly constrained. Theoretical modelling bridges between existing observations of the two dimensional surface of recombination at redshift $z \sim 1100$ probed by the CMB radiation, and the epoch of galaxy formation with the most distant detected galaxy located at $z = 11.1$ [36].

The 21-cm signal of neutral hydrogen provides a unique way to constrain the first few hundred million years (see Fialkov, these proceedings). The hyper-fine splitting of the lowest hydrogen energy level gives rise to the rest-frame $\nu_{21} = 1.42$ GHz radio signal with the equivalent wavelength of 21 cm (see ref. [37] for a recent review). The cumulative signal of neutral intergalactic medium observed against the background radiation depends on the processes of cosmic heating and ionization and, to the leading order, scales as $T_{21} \propto x_{\rm HI}\left(1 - T_{\rm rad}T_{\rm S}^{-1}\right)$ where $T_{\rm rad}$ is the brightness temperature of the background radiation at 1.42 GHz and $T_{\rm S}$ is the spin temperature of the transition which at cosmic dawn is close to the kinetic temperature of the gas, $T_{\rm K}$. Owing to its dependence on the underlying astrophysics and cosmology, this signal is a powerful tool to characterize the formation and the evolution of the first populations of astrophysical sources and, potentially, properties of dark matter (DM), across cosmic time. Uncertainties in the high-redshift astrophysical processes result in a large variety of the possible 21-cm signals (see grey lines in Fig. 1).

Observations with the High-Band antenna of the Experiment to Detect the Global EoR Signature (EDGES) [39–41] and the Shaped Antenna measurement of the background RAdio Spectrum2 (SARAS2) [42,43] yield upper limits on the 21-cm signal from redshifts $z \sim 6-15$. In the framework of standard astrophysical modelling, the data require star formation in small DM halos [41] and rule out extremely non-efficient X-ray heating sources [41–43].

Recently, the first detection of the sky-averaged (global) 21-cm signal of neutral hydrogen from $z \sim 13 - 27$ has been claimed based on two years of observations with the EDGES Low-Band antenna in the $50 - 100$ MHz frequency range [44]. If confirmed, this is the only existing observation from the intermediate redshift range and is the first observational evidence of the primordial star formation at $z \sim 20$ ($\sim 180$ million years after the Big Bang) and early X-ray heating. The reported cosmological 21-cm signal is centered at $z = 17.2$ (which corresponds to $\nu = 78.2$ MHz), and features an absorption trough of $T_{21} = -500^{+200}_{-500}$ mK, where the error corresponds to 99% confidence including both thermal and systematic noise. The depth of the feature is at least twice as strong as predicted in standard astrophysical scenarios (based on the assumption of $\Lambda$CDM cosmology and hierarchical structure

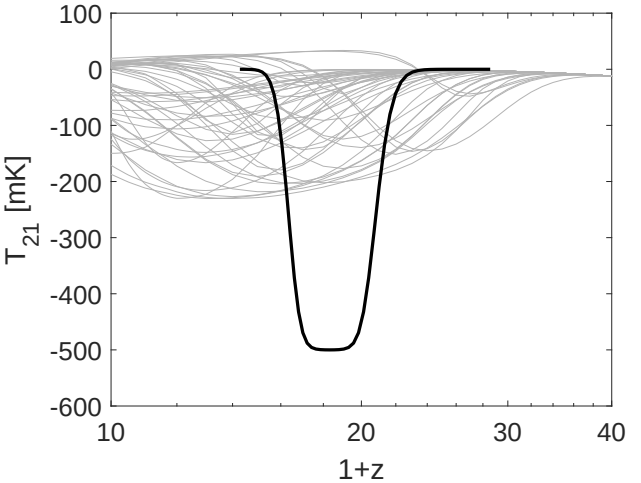



Fig. 1. The variety of plausible predicted 21-cm signals (grey lines – computed here for 50 models according to ref.[38]) and the best-fit profile of the EDGES Low-band detection (thick black line).

formation) where the strongest possible feature at $z \simeq 17$ is $\simeq -209$ mK, assuming the CMB as the background radiation with $T_{CMB} \simeq 49.5$ K and the coldest possible temperature of the intergalactic medium (IGM) of $\sim 7.2$ K obtained in the absence of X-ray heating sources. The observed $T_{21} < -300$ mK requires either the gas to be much colder (around 5.2 K) or the background radiation to be much stronger (around 67.2 K), which is hard to explain by astrophysics alone. The frequency of the detected feature can be related to the formation of the first stars which couple the spin temperature to the temperature of the gas via absorption and re-emission of Ly-$\alpha$ photons; while the narrowness of the feature indicates a relatively rapid onset of X-ray heating.

To explain the depth of the reported feature exotic mechanisms have to be added to the standard picture. The amplitude of the observed 21-cm feature can be explained by non-gravitational interaction between DM and baryons, e.g., via Rutherford-like scattering which could drain excess energy from the gas lowering its kinetic temperature (see, e.g., ref.[45]). Even though this scenario is strongly constrained by observations, it is not completely ruled out and is still plausible for a narrow range of parameters including the DM mass, electric charge of DM particle and cross-section (see, e.g., ref.[46]). A smoking gun signature of the baryon-dark matter (b-dm) scattering with a velocity-dependent cross-section, first proposed in ref.[45] as a solution to the EDGES Low-Band anomaly, was shown to be an enhanced pattern of BAO in the 21-cm power spectrum[47]. The power spectrum itself is boosted by as much as three orders of magnitude which renders the fluctuations detectable by telescopes such as the low band antennas of LOFAR.

An excess radio background at the rest-frame 1.42 GHz at $z = 17$ would also explain the large contrast observed between the background temperature and the spin temperature of the 21-cm transition[44,48]. Interestingly, evidence of excess radio background above the CMB at low radio frequencies was detected by ARCADE2 at $3 - 90$ GHz[49] and recently confirmed by LWA1 at $40 - 80$ MHz[50]. However, it is still not clear what part of the observed excess is extragalactic[51]. If it exists, the excess radio background is also expected to boost the 21-cm power spectrum[52].

Verification of the EDGES-Low signal is on the way with instruments such as the EDGES Mid-Band antenna, SARAS3, LEDA, and the LOFAR Low Band.

#### 4.2.1. *Black holes high mass X-ray binaries microquasars and cosmic backgrounds*

A possible explanation of the EDGES excess involving black hole high mass X-ray binary microquasars (BH-HMXB-MQs) has been recently proposed in ref.[53]. There is indeed an increasing evidence of a high formation rate of BH-HMXBs at high redshifts, theoretically advanced in ref.[54]. According to this model, a significant fraction of Pop III stars with mass $\gtrsim 20$ $M_\odot$ and low metallicity, $Z \leq 0.001$ $Z_\odot$, in binaries or larger multiple systems is predicted to collapse with no energetic SN kicks and to remain in situ, finally resulting into BH-HMXB-MQs of Pop III. These objects are likely sources of a substantial synchrotron radio emission, with a relatively smooth spectrum, able to produce a cosmic background that could significantly enhance the absorption feature of the redshifted 21-cm line by neutral hydrogen produced at redshifts $z \sim 17$. The absence of a cosmic far-IR thermal background associated with a cosmic radio background can be explained in this model by the fact that BH-HMXB-MQs in stellar clusters of Pop III are formed before the appearance of SN explosions, neutron stars and dust. BH-HMXB-MQs promptly inject hard X-rays and relativistic jets in the cold hydrogen that enshrouds the slowly expanding HII regions ionized by the most massive progenitor stars of Pop III. The high column depths for the cold hydrogen IGM largely block the X-rays from inner sources but are essentially transparent to the radio synchrotron emission.

Crucial tests of this scenario will be based on the next generation of multi-wavelength / multi-messenger surveys, including interferometric observations of the HI redshifted 21-cm line able to provide an accurate tomographic view of HI inhomogeneities, X-ray missions that will allow us a better understanding of the HMXBs role, and gravitational waves observatories providing a larger statistics of HMXBs and BBHs useful for cosmological investigations and, hopefully, a measure of the spectrum of gravitational wave background that, in this model, is predicted to have a flattening of the spectral index at frequencies as low as $\sim 30$ Hz.

## 5. Extragalactic astrophysics

### 5.1. *Extragalactic point sources*

Although the main goal of *Planck* was to provide accurate all-sky CMB temperature and polarisation maps, it has supplied unprecedented data of astrophysical interest in several fields. On the extragalactic side, *Planck* carried out the deepest systematic all-sky survey of galaxy clusters seen via the SZ effect. It also detected thousands of dusty galaxies as well as many hundreds of extragalactic radio sources in a spectral range difficult or impossible to explore from the ground and only lightly surveyed by other space missions.

Because of its quite limited angular resolution ($\simeq 5'$ at sub-mm wavelengths) *Planck* was confusion limited at quite bright flux density levels, so that the overwhelming majority of dusty galaxies it detected are very local, at distances generally below 100 Mpc[55]. *Planck* offered the first opportunity to accurately determine their sub-mm properties[56].

Even more interestingly, *Planck* opened new paths to the study of early phases of galaxy evolution, exploiting the strong gravitational lensing (Sect. 5.1.1), and of cluster evolution (Sect. 5.1.2); see De Zotti et al., these proceedings.

#### 5.1.1. *Strongly lensed galaxies*

Somewhat unexpectedly (but see the predictions by ref.[57]) *Planck* also succeeded at detecting some high-$z$ galaxies whose flux densities were boosted by extreme gravitational magnifications[58] (see the left panel of Fig. 2.)

The 11 strongly lensed galaxies discovered on *Planck* maps[58] have redshifts in the range 2.2–3.6. This offers the exciting possibility of exploiting high spatial and spectral resolution follow-up to address major, still open issues on galaxy formation and evolution: which are the main physical mechanisms shaping the galaxy properties: in situ processes? interactions? mergers? cold flows from the intergalactic medium? How do feedback processes work? To settle these issues we need direct information on the structure and the dynamics of high-$z$ galaxies. But these are compact, with typical sizes of 1–2 kpc (e.g., ref.[59]), corresponding to angular sizes of 0.1–0.2 arcsec at $z \simeq 2$–3. Thus they are hardly resolved even by ALMA and by the HST. If they are resolved, high enough S/N ratios per resolution element are achieved only for the brightest galaxies, probably not representative of the general population. Strong lensing offers a way out of all these difficulties.

#### 5.1.2. *Proto-clusters*

Classical techniques for detecting galaxy clusters (optical/near-IR "red sequence", X-ray emission, SZ effect) preferentially or exclusively select evolved objects, with mature galaxy populations and a hot intra-cluster medium. As a result most known clusters are at $z < 1.5$, i.e. below the peak of global star-formation activity.

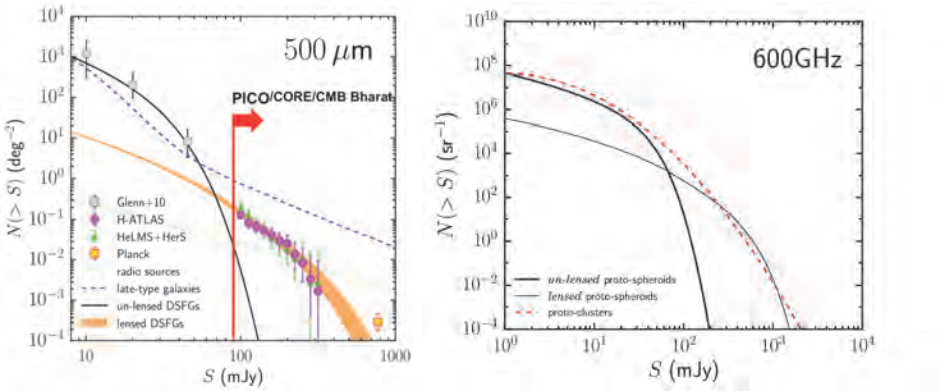



Fig. 2. **Left panel:** integral counts of the main extragalactic source populations at 500 $\mu$m (600 GHz). The counts of unlensed (solid black line) and strongly lensed (orange band) high-$z$ galaxies (labelled DSFGs) are compared with observational determinations from *Herschel* surveys. The yellow square on the bottom-right corner shows the estimated counts of strongly lensed galaxies detected by *Planck*. The dashed blue line shows the counts of late-type (normal plus starburst) galaxies. All these lines show predictions of the model by ref.[60]. The dotted green line shows the counts of radio sources predicted by the C2Ex model of ref.[61]. The vertical red line shows the estimated detection limits of planned next generation CMB experiments, PICO, CORE and CMB Bharat. Adapted from ref.[62] [©SISSA Medialab Srl. Reproduced by permission of IOP Publishing. All rights reserved]. **Right panel:** integral counts of proto-clusters at 600 GHz, predicted by the model of ref.[63], compared with the counts of high-$z$ lensed and unlensed proto-spheroidal galaxies.

At higher redshifts cluster members enter the active star-formation phase and the hot intergalactic gas is no longer necessarily in place. But then strong galaxy overdensities stand out as intensity peaks in the low angular resolution maps of CMB experiments, as clearly demonstrated by *Planck*[64]. Thus CMB experiments offer the possibility of extending the investigation of cluster evolution up to much higher redshifts than would be possible by other means.

However, *Planck*'s resolution was too poor to detect individual proto-clusters[63]. As illustrated by the right panel of Fig. 2, next generation experiments will detect many tens of thousands of these objects as peaks in sub-mm maps, in addition to the evolved ones, detected by the SZ effect. This will constitute a real breakthrough in the observational validation of models for structure formation.

### 5.2. *Virial clouds in the galactic halos*

The rotation of the galactic objects and disks of various nearby edge-on spiral galaxies has been studied by using the Doppler asymmetry in the CMB radiation[65–70]. It was proposed that the asymmetry is caused by molecular hydrogen clouds which are exactly merged with the CMB, i.e. they are at the CMB temperature[71]. An objection was raised that the effect might not be due to pure molecular hydrogen clouds, it might be due to clouds completely made of interstellar dust or clouds

containing some fraction of $H_2$ and dust. In order to address the issue in this paper we only modeled the clouds with the assumption that they are composed of pure $H_2$ molecules (single fluid model). After modeling the clouds we tried to estimate the luminosity caused by the clouds (see Qadir et al., these proceedings). Further, we want to extend the analysis (in a later work) to a two-fluid model. The hope is that by comparing it with the *Planck* data we can constrain the missing baryonic distribution in the galactic halos.

Since, the clouds are immersed in a heat bath which is CMB, so, they are isothermal and we called them "virial clouds". They are formed due to the Jeans instability, so, they should have a precise Jeans mass and radius. The density of the clouds should be flat at the center and then it goes exactly zero at the the Jeans radius in order to give a definite boundary. The density distribution can be given by the isothermal Lane-Emden equation but the density profile in that case will not be exactly zero at the boundary. In this case the density distribution had to be put on ad-hoc basis. We do not have to do this. Instead of using the isothermal Lane-Emden equation we used the canonical ensemble distribution to obtain the mass density profile. We also introduced the speed of sound $c_s$ which is given by $c_s = \sqrt{(\gamma k T_{CMB})/m_H}$, where, $\gamma$ is the adiabatic factor. At the CMB temperature no higher degrees of freedom would excite, so, we only have the translational degree, making the ideal gas approximation extremely good and so $\gamma = 5/3$. For our purpose, $m_H \approx 2.016$ g/mol and $T \approx 2.726\,\mathrm{K}$, so $c_s \simeq 1.110 \times 10^4$ km s$^{-1}$. The virial theorem $2K + U = 0$ gives the Jeans mass squared as[72]

$$M_J^2 \simeq \left(\frac{81}{32\pi\rho_c}\right)\left(\frac{3c_s^2}{5G}\right)^3 , \tag{1}$$

and the corresponding Jeans radius squared as

$$R_J^2 = \frac{27c_s^2}{20\pi\rho_c G} , \tag{2}$$

where $\rho_c$ is the central density of the virial cloud. We obtained a differential equation as

$$r\frac{d\rho(r)}{dr} - r^2\left(\frac{4\pi G m_H}{kT_{CMB}}\right)\rho^2(r) - \rho(r)\ln\left(\frac{\rho(r)}{\zeta}\right) = 0 . \tag{3}$$

where, $\zeta = 8m_H^{5/2}(G\rho_c/3kT_{CMB})^{3/2}$. We have the initial condition $\rho'(r)|_{r=0} = 0$ and the boundary condition that $\rho(R_J) = 0$. As such, we can solve it numerically. The resulting density profile is shown in Fig. 3. The obtained central density of the cloud is $\rho_c \simeq 1.60 \times 10^{-18}$ kg m$^{-3}$, the radius is $R_J \simeq 0.030$pc, and the mass is $M_J \simeq 0.798 M_\odot$.

Since the virial clouds are at the CMB temperature, their luminosity is the CMB luminosity, except for their Doppler shift. Thus, we only need to look at the differential frequency shift, $\Delta\nu/\nu = -(v_{rot}/c)\cos\theta$, where $v_{rot}$ is the galactic halo rotational velocity and the luminosity of a single cloud, $L_c$, could be given by $L_c = (\Delta\nu/\nu)L \simeq 1.486 \times 10^{-4} L\cos\theta$, being $\theta$ the angle between the direction of the

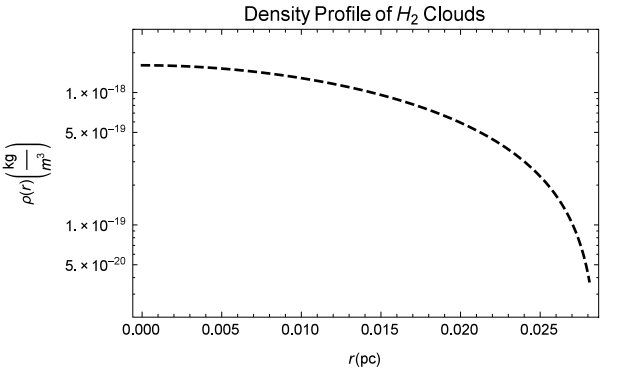



Fig. 3. The density of the virial molecular hydrogen cloud. It decreases monotonically from the central density, $\rho_c \simeq 1.60 \times 10^{-18}$ kg m$^{-3}$, to zero at $r = R_J \simeq 0.030$ pc.

velocity and the line of sight and $L$ the observed luminosity. This would directly give an estimate of the total number of clouds being seen.

### 5.3. *Perspectives from SPICA*

IR observations allowed us to study the obscured Universe, providing information about the physical state and energy balance of the cool matter relevant for the study of many physical processes occurring during the formation and evolution of galaxies and stars. Indeed, a large fraction ($\sim 50\%$ or more) of the energy emitted by stars in galaxies in form of optical and ultraviolet radiation is absorbed and reprocessed by dust grains and re-emitted in the IR, and, for distant objects, the cosmic expansion implies a further substantial redshift of the observed photons. For these reasons, galactic evolution, with a peak of activity at z$\sim$1-4 corresponding to a Universe age of about 2-3 Gyr, can be profitably investigated in the (mid- and far-) IR domain. Furthermore, improving our understanding of fundamental processes of star formation and evolution, including dust and gas physics of pre-stellar cores and proto-planetary discs, that can be more accurately studied in our galaxy in the IR band, allows us a better modeling of high-$z$ phenomena.

The physical conditions, such as excitation, density and metallicity, of nuclei of galaxies as well as of stars and planets formation regions can be investigated analyzing ionic, atomic, molecular and dust features in the IR. The star formation rate, the stellar type and the density of the gas in HII regions around hot young stars can be investigated using ionic fine structure lines, e.g., [NeII], [SIII], [OIII], while the AGN accretion rate is traced by spectral lines of highly ionized species, e.g., [OIV], [NeV] (see the expected potentials of far-IR spectroscopic observations for studying galaxy/AGN co-evolution presented in ref.[73]). Photo-dissociation regions can be studied via the [CII] and [OI] lines and the emission from small dust grains and PAHs. Also, the cooling of diffuse warm gas in galaxies, through, e.g., [NII],

probing ISM physics, is accessible to IR observations. Pure rotational lines from $H_2$, HD and OH, mid- to high-J CO and $H_2O$ lines, and PAH emission features appear in the object rest-frame IR band, the latter, containing several per cent of of the star-forming galaxies IR emission, being crucial for dust-obscured objects redshift determination. Other interesting dust features from minerals, e.g., olivine, calcite and dolomite, and $CO_2$ ice, $C_2H_2$ molecules and fullerenes are also observable in the IR.

The above tracers have been only partially studied with previous IR missions, mainly because of limitations in telescope sizes and cryogenic performances.

To overcome these limits, a mid- and far-IR mission concept, the Space Infrared Telescope for Cosmology and Astrophysics (SPICA)[74], was proposed to JAXA and ESA as a collaboration mission between Japanese and European scientists. It was selected in May 2018 by ESA as a finalist for the next Medium class Mission 5 of the Cosmic Vision programme for a proposed launch date of 2032. The spacecraft is designed to be deployed in a halo orbit around the Sun-Earth L2 Lagrange point to minimize systematic effects. The mission foresees a 2.5-meter diameter telescope cooled to below 8 K. A combination of passive cooling and mechanical coolers, that will cool the telescope and the instruments, will enable a mission lifetime significantly beyond a nominal requirement of three years. Combining a large and cold telescope with instruments employing state-of-the-art ultra-sensitive detectors, an improvement in IR spectroscopic sensitivity (achieving the – $5\sigma$/1hr – level of $\sim 5 \times 10^{-20}$ W/m$^2$) of more than two orders of magnitude with respect to Herschel, Spitzer and SOFIA is planned. SPICA will offer spectral resolving power ranging from $R \sim$50 through 11000 in the 17-230 $\mu$m domain as well as $R \sim$28.000 spectroscopy between 12 and 18 $\mu$m. Furthermore, SPICA will perform efficient 30-37 $\mu$m broad band mapping, and small field spectroscopic and polarimetric imaging in the 100-350 $\mu$m range.

A wide set of questions in astrophysics, e.g., the role in galaxy evolution of star formation, BHs and SNe explosions, metal and dust history, the matter cycle from galactic scale to larger structures and IGM, the formation of first galaxies and BHs, will get benefit from deep analyses of the IR tracers made possible by such a mission.

## Acknowledgments

CB, GDZ and TT acknowledge partial support from ASI/INAF agreement n. 2014-024-R.1 for the *Planck* LFI Activity of Phase E2 and from the ASI/Physics Department of the university of Roma–Tor Vergata agreement n. 2016-24-H.0 for study activities of the Italian cosmology community; MB, CB and TT acknowledge partial support from the INAF PRIN SKA/CTA project FORECaST; AF acknowledges the support by the Royal Society for her University Research Fellow at Sussex University; AQ and NT acknowledge the Salento University of Lecce for hospitality; AQ is grateful to IUPAP, ESA, ICRANet and AS-ICTP for support to participate in MG15. Some of the results in this paper have been derived using the HEALPix[75]

package. The use of the *Planck* Legacy Archive is acknowledged. It is a pleasure to thank all the speakers of the CM2 session and their collaborators.

## References

1. J. A. Tauber, H. U. Norgaard-Nielsen, P. A. R. Ade *et al.*, Planck pre-launch status: The optical system, *A&A* **520**, p. A2 (September 2010).
2. Planck Collaboration, Planck early results. I. The Planck mission, *A&A* **536**, p. A1 (December 2011).
3. Planck Collaboration, Planck 2015 results. XIII. Cosmological parameters, *A&A* **594**, p. A13 (September 2016).
4. Planck Collaboration, Planck 2018 results. VI. Cosmological parameters, *arXiv e-prints* (July 2018).
5. Planck Collaboration, Planck 2018 results. I. Overview and the cosmological legacy of Planck, *arXiv e-prints* (July 2018).
6. Planck Collaboration, Planck 2018 results. II. Low Frequency Instrument data processing, *arXiv e-prints* (July 2018).
7. Planck Collaboration, Planck 2018 results. III. High Frequency Instrument data processing and frequency maps, *arXiv e-prints* (July 2018).
8. Planck Collaboration, Planck 2018 results. VIII. Gravitational lensing, *arXiv e-prints* (July 2018).
9. Planck Collaboration, Planck 2018 results. X. Constraints on inflation, *arXiv e-prints* (July 2018).
10. A. G. Riess, S. Casertano, W. Yuan, L. Macri, B. Bucciarelli, M. G. Lattanzi, J. W. MacKenty, J. B. Bowers, W. Zheng, A. V. Filippenko, C. Huang and R. I. Anderson, Milky Way Cepheid Standards for Measuring Cosmic Distances and Application to Gaia DR2: Implications for the Hubble Constant, *ApJ* **861**, p. 126 (July 2018).
11. A. G. Riess, S. Casertano, W. Yuan, L. M. Macri and D. Scolnic, Large Magellanic Cloud Cepheid Standards Provide a 1% Foundation for the Determination of the Hubble Constant and Stronger Evidence for Physics Beyond LambdaCDM, *arXiv e-prints* (March 2019).
12. A. E. Romano, Hubble trouble or Hubble bubble?, *International Journal of Modern Physics D* **27**, p. 1850102 (2018).
13. G. P. da Silva and A. G. Cavalcanti, A More Accurate and Competitive Estimative of H $_0$ in Intermediate Redshifts, *Brazilian Journal of Physics* **48**, 521 (October 2018).
14. J. A. S. Lima and J. V. Cunha, A 3% Determination of H $_0$ at Intermediate Redshifts, *ApJL* **781**, p. L38 (February 2014).
15. C. D. Kreisch, F.-Y. Cyr-Racine and O. Doré, The Neutrino Puzzle: Anomalies, Interactions, and Cosmological Tensions, *arXiv e-prints* (February 2019).
16. V. Poulin, T. L. Smith, T. Karwal and M. Kamionkowski, Early Dark Energy Can Resolve The Hubble Tension, *arXiv e-prints* (November 2018).

17. Planck Collaboration, Planck 2015 results. XXVII. The second Planck catalogue of Sunyaev-Zeldovich sources, *A&A* **594**, p. A27 (September 2016).
18. Planck Collaboration, Planck 2015 results. XXII. A map of the thermal Sunyaev-Zeldovich effect, *A&A* **594**, p. A22 (September 2016).
19. E. Komatsu and T. Kitayama, Sunyaev-Zeldovich Fluctuations from Spatial Correlations between Clusters of Galaxies, *ApJL* **526**, L1 (November 1999).
20. B. Bolliet, Cosmological constraints from the thermal Sunyaev Zeldovich power spectrum?, *arXiv e-prints* (June 2018).
21. Planck Collaboration, Planck 2015 results. XXIV. Cosmology from Sunyaev-Zeldovich cluster counts, *A&A* **594**, p. A24 (September 2016).
22. B. Bolliet, B. Comis, E. Komatsu and J. F. Macías-Pérez, Dark energy constraints from the thermal Sunyaev-Zeldovich power spectrum, *MNRAS* **477**, 4957 (July 2018).
23. A. Kogut, D. J. Fixsen, D. T. Chuss *et al.*, The Primordial Inflation Explorer (PIXIE): a nulling polarimeter for cosmic microwave background observations, *JCAP* **7**, p. 025 (July 2011).
24. H. Ishino, Y. Akiba, K. Arnold *et al.*, LiteBIRD: lite satellite for the study of B-mode polarization and inflation from cosmic microwave background radiation detection, in *Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series*, , Proceedings of the SPIEVol. 9904July 2016.
25. J. Delabrouille, P. de Bernardis, F. R. Bouchet *et al.*, Exploring cosmic origins with CORE: Survey requirements and mission design, *JCAP* **4**, p. 014 (April 2018).
26. S. Hanany, M. Alvarez, E. Artis *et al.*, PICO: Probe of Inflation and Cosmic Origins, *arXiv e-prints* (February 2019).
27. P. André, C. Baccigalupi, A. Banday *et al.*, PRISM (Polarized Radiation Imaging and Spectroscopy Mission): an extended white paper, *JCAP* **2**, p. 006 (February 2014).
28. E. Di Valentino, T. Brinckmann, M. Gerbino *et al.*, Exploring cosmic origins with CORE: Cosmological parameters, *JCAP* **4**, p. 017 (April 2018).
29. F. Finelli, M. Bucher, A. Achúcarro *et al.*, Exploring cosmic origins with CORE: Inflation, *JCAP* **4**, p. 016 (April 2018).
30. W. Hu, D. Scott and J. Silk, Power spectrum constraints from spectral distortions in the cosmic microwave background, *ApJL* **430**, L5 (July 1994).
31. J. Chluba and R. A. Sunyaev, The evolution of CMB spectral distortions in the early Universe, *MNRAS* **419**, 1294 (January 2012).
32. Y. B. Zel'dovich, A. F. Illarionov and R. A. Sunyaev, The Effect of Energy Release on the Emission Spectrum in a Hot Universe, *Soviet Journal of Experimental and Theoretical Physics* **35**, p. 643 (1972).
33. D. J. Fixsen, E. Dwek, J. C. Mather, C. L. Bennett and R. A. Shafer, The Spectrum of the Extragalactic Far-Infrared Background from the COBE FIRAS Observations, *ApJ* **508**, 123 (November 1998).

34. L. Danese and G. de Zotti, Dipole anisotropy and distortions of the spectrum of the cosmic microwave background, *A&A* **94**, p. L33 (February 1981).
35. C. Burigana, C. S. Carvalho, T. Trombetti *et al.*, Exploring cosmic origins with CORE: Effects of observer peculiar motion, *JCAP* **4**, p. 021 (April 2018).
36. P. A. Oesch, G. Brammer, P. G. van Dokkum *et al.*, A Remarkably Luminous Galaxy at z=11.1 Measured with Hubble Space Telescope Grism Spectroscopy, *ApJ* **819**, p. 129 (March 2016).
37. R. Barkana, The rise of the first stars: Supersonic streaming, radiative feedback, and 21-cm cosmology, *Physics Reports* **645**, 1 (July 2016).
38. A. Cohen, A. Fialkov, R. Barkana and M. Lotem, Charting the parameter space of the global 21-cm signal, *MNRAS* **472**, 1915 (December 2017).
39. R. A. Monsalve, A. E. E. Rogers, J. D. Bowman and T. J. Mozdzen, Results from EDGES High-band. I. Constraints on Phenomenological Models for the Global 21 cm Signal, *ApJ* **847**, p. 64 (September 2017).
40. R. A. Monsalve, B. Greig, J. D. Bowman, A. Mesinger, A. E. E. Rogers, T. J. Mozdzen, N. S. Kern and N. Mahesh, Results from EDGES High-band. II. Constraints on Parameters of Early Galaxies, *ApJ* **863**, p. 11 (August 2018).
41. R. A. Monsalve, A. Fialkov, J. D. Bowman, A. E. E. Rogers, T. J. Mozdzen, A. Cohen, R. Barkana and N. Mahesh, Results from EDGES High-Band: III. New Constraints on Parameters of the Early Universe, *arXiv e-prints* (January 2019).
42. S. Singh, R. Subrahmanyan, N. Udaya Shankar, M. Sathyanarayana Rao, A. Fialkov, A. Cohen, R. Barkana, B. S. Girish, A. Raghunathan, R. Somashekar and K. S. Srivani, First Results on the Epoch of Reionization from First Light with SARAS 2, *ApJL* **845**, p. L12 (August 2017).
43. S. Singh, R. Subrahmanyan, N. Udaya Shankar, M. Sathyanarayana Rao, A. Fialkov, A. Cohen, R. Barkana, B. S. Girish, A. Raghunathan, R. Somashekar and K. S. Srivani, SARAS 2 Constraints on Global 21 cm Signals from the Epoch of Reionization, *ApJ* **858**, p. 54 (May 2018).
44. J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen and N. Mahesh, An absorption profile centred at 78 megahertz in the sky-averaged spectrum, *Nature* **555**, 67 (March 2018).
45. R. Barkana, Possible interaction between baryons and dark-matter particles revealed by the first stars, *Nature* **555**, 71 (March 2018).
46. E. D. Kovetz, V. Poulin, V. Gluscevic, K. K. Boddy, R. Barkana and M. Kamionkowski, Tighter limits on dark matter explanations of the anomalous EDGES 21 cm signal, *Phys.Rev.D* **98**, p. 103529 (November 2018).
47. A. Fialkov, R. Barkana and A. Cohen, Constraining Baryon-Dark-Matter Scattering with the Cosmic Dawn 21-cm Signal, *Physical Review Letters* **121**, p. 011101 (July 2018).
48. C. Feng and G. Holder, Enhanced Global Signal of Neutral Hydrogen Due to Excess Radiation at Cosmic Dawn, *ApJL* **858**, p. L17 (May 2018).

49. D. J. Fixsen, A. Kogut, S. Levin, M. Limon, P. Lubin, P. Mirel, M. Seiffert, J. Singal, E. Wollack, T. Villela and C. A. Wuensche, ARCADE 2 Measurement of the Absolute Sky Brightness at 3-90 GHz, *ApJ* **734**, p. 5 (June 2011).
50. J. Dowell and G. B. Taylor, The Radio Background below 100 MHz, *ApJL* **858**, p. L9 (May 2018).
51. R. Subrahmanyan and R. Cowsik, Is there an Unaccounted for Excess in the Extragalactic Cosmic Radio Background?, *ApJ* **776**, p. 42 (October 2013).
52. A. Fialkov and R. Barkana, Signature of Excess Radio Background in the 21-cm Global Signal and Power Spectrum, *arXiv e-prints* (February 2019).
53. I. F. Mirabel, Black Hole High Mass X-ray Binary Microquasars at Cosmic Dawn, *arXiv e-prints* (February 2019).
54. I. F. Mirabel, M. Dijkstra, P. Laurent, A. Loeb and J. R. Pritchard, Stellar black holes at the dawn of the universe, *A&A* **528**, p. A149 (April 2011).
55. M. Negrello, M. Clemens, J. Gonzalez-Nuevo *et al.*, The local luminosity function of star-forming galaxies derived from the Planck Early Release Compact Source Catalogue, *MNRAS* **429**, 1309 (February 2013).
56. M. S. Clemens, M. Negrello, G. De Zotti *et al.*, Dust and star formation properties of a complete sample of local galaxies drawn from the Planck Early Release Compact Source Catalogue, *MNRAS* **433**, 695 (July 2013).
57. M. Negrello, F. Perrotta, J. González-Nuevo, L. Silva, G. de Zotti, G. L. Granato, C. Baccigalupi and L. Danese, Astrophysical and cosmological information from large-scale submillimetre surveys of extragalactic sources, *MNRAS* **377**, 1557 (June 2007).
58. R. Cañameras, N. P. H. Nesvadba, D. Guery *et al.*, Planck's dusty GEMS: The brightest gravitationally lensed galaxies discovered with the Planck all-sky survey, *A&A* **581**, p. A105 (September 2015).
59. S. Fujimoto, M. Ouchi, K. Kohno *et al.*, ALMA 26 Arcmin$^2$ Survey of GOODS-S at One Millimeter (ASAGAO): Average Morphology of High-z Dusty Star-forming Galaxies in an Exponential Disk ($n \simeq 1$), *ApJ* **861**, p. 7 (July 2018).
60. Z.-Y. Cai, A. Lapi, J.-Q. Xia, G. De Zotti, M. Negrello, C. Gruppioni, E. Rigby, G. Castex, J. Delabrouille and L. Danese, A Hybrid Model for the Evolution of Galaxies and Active Galactic Nuclei in the Infrared, *ApJ* **768**, p. 21 (May 2013).
61. M. Tucci, L. Toffolatti, G. de Zotti and E. Martínez-González, High-frequency predictions for number counts and spectral properties of extragalactic radio sources. New evidence of a break at mm wavelengths in spectra of bright blazar sources, *A&A* **533**, p. A57 (September 2011).
62. G. De Zotti, J. González-Nuevo, M. Lopez-Caniego *et al.*, Exploring cosmic origins with CORE: Extragalactic sources in cosmic microwave background maps, *JCAP* **4**, p. 020 (April 2018).
63. M. Negrello, J. Gonzalez-Nuevo, G. De Zotti, M. Bonato, Z.-Y. Cai, D. Clements, L. Danese, H. Dole, J. Greenslade, A. Lapi and L. Montier, On

the statistics of proto-cluster candidates detected in the Planck all-sky survey, *MNRAS* **470**, 2253 (September 2017).
64. Planck Collaboration Int. XXXIX, Planck intermediate results. XXXIX. The Planck list of high-redshift source candidates, *A&A* **596**, p. A100 (December 2016).
65. F. de Paolis, V. G. Gurzadyan, G. Ingrosso, P. Jetzer, A. A. Nucita, A. Qadir, D. Vetrugno, A. L. Kashin, H. G. Khachatryan and S. Mirzoyan, Possible detection of the M 31 rotation in WMAP data, *A&A* **534**, p. L8 (October 2011).
66. F. De Paolis, V. G. Gurzadyan, A. A. Nucita, G. Ingrosso, A. L. Kashin, H. G. Khachatryan, S. Mirzoyan, E. Poghosian, P. Jetzer, A. Qadir and D. Vetrugno, Planck confirmation of the disk and halo rotation of M 31, *A&A* **565**, p. L3 (May 2014).
67. F. De Paolis, V. G. Gurzadyan, A. A. Nucita, G. Ingrosso, A. L. Kashin, H. G. Khachatryan, S. Mirzoyan, G. Yegorian, P. Jetzer, A. Qadir and D. Vetrugno, Planck revealed bulk motion of Centaurus A lobes, *A&A* **580**, p. L8 (August 2015).
68. V. G. Gurzadyan, F. De Paolis, A. A. Nucita, G. Ingrosso, A. L. Kashin, H. G. Khachatryan, S. Sargsyan, G. Yegorian, P. Jetzer, A. Qadir and D. Vetrugno, Planck view of the M 82 galaxy, *A&A* **582**, p. A77 (October 2015).
69. F. De Paolis, V. G. Gurzadyan, A. A. Nucita, L. Chemin, A. Qadir, A. L. Kashin, H. G. Khachatryan, S. Sargsyan, G. Yegorian, G. Ingrosso, P. Jetzer and D. Vetrugno, Triangulum galaxy viewed by Planck, *A&A* **593**, p. A57 (September 2016).
70. V. G. Gurzadyan, F. De Paolis, A. A. Nucita, A. L. Kashin, A. Amekhyan, S. Sargsyan, G. Yegorian, A. Qadir, G. Ingrosso, P. Jetzer and D. Vetrugno, Messier 81's Planck view versus its halo mapping, *A&A* **609**, p. A131 (February 2018).
71. F. de Paolis, G. Ingrosso, P. Jetzer, A. Qadir and M. Roncadelli, Observing molecular hydrogen clouds and dark massive objects in galactic halos., *A&A* **299**, p. 647 (July 1995).
72. S. Chandrasekhar, *An introduction to the study of stellar structure.* 1957.
73. M. Bonato, M. Negrello, Z.-Y. Cai *et al.*, Exploring the relationship between black hole accretion and star formation with blind mid-/far-infrared spectroscopic surveys, *MNRAS* **444**, 3446 (November 2014).
74. P. R. Roelfsema, H. Shibai, L. Armus *et al.*, SPICA-A Large Cryogenic Infrared Space Telescope: Unveiling the Obscured Universe, *PASA* **35**, p. e030 (August 2018).
75. K. M. Górski, E. Hivon, A. J. Banday, B. D. Wandelt, F. K. Hansen, M. Reinecke and M. Bartelmann, HEALPix: A Framework for High-Resolution Discretization and Fast Analysis of Data Distributed on the Sphere, *ApJ* **622**, 759 (April 2005).

# Understanding accretion flow properties of black hole candidates after implementation of the TCAF solution in XSPEC

Dipak Debnath[1]*, Sandip K. Chakrabarti[1], Santanu Mondal[1], Arghajit Jana[1], Debjit Chatterjee[1], Aslam Ali Molla[1], Kaushik Chatterjee[1], Riya Bhowmick[1]

*1. Indian Centre For Space Physics, 43 Chalantika, Garia Station Road, Kolkata, 700084, India; *E-mail: dipakcsp@gmail.com*

Understanding accretion flow dynamics around black hole (BH) sources one needs to make detailed study of spectral as well as temporal properties of these objects during their active phases. We believe that studying accretion flow properties of BHs, it is necessary analyze observed data with a generalized physical model. Recently after the implementation of the generalized two Component Advective Flow (TCAF) solution as a local additive table model into HeaSARC's spectral analysis package XSPEC, a more clear picture about the flow nature around the BHs are obtained. TCAF model fitted spectral parameters (two types of accretion rates: Keplerian disk and sub-Keplerian halo, and two types of shock parameters: location and strength of the shock) allow us to find a better physical picture about the physical processes associated with BHs. One can also predict frequency of primary dominating QPOs from TCAF model fitted shock parameters. Based on evolution of halo to disk accretion rate ratio (ARR), quasi-periodic oscillation (QPOs; if present) frequencies, a physical understanding of different observed spectral states and their transitions are done. We are able to estimate intrinsic source parameters (mass, spin), jet X-ray fluxes, etc. from spectral analysis with TCAF model. This makes this model as an ideal tool to study BHs.

*Keywords*: X-Rays:binaries – stars:black holes – stars individual: (H 1743-322, MAXI J1659-152, MAXI J1836-194, MAXI J1543-564, MAXI J1535-571, Swift J1753.5-0127, XTE J1118+480) – accretion, accretion disks – shock waves – radiation:dynamics

## 1. Introduction

Compact objects such as black hole (BH), neutron star (NS), white dwarf (WD) are the end products of the main sequence stars. BHs are more massive objects which are invisible due to their high gravity. They are only observable when they accrete matter from their massive companions. Depending upon their flaring activity, they are classified mainly into two types: transient and persistent. Transient black hole candidates (BHCs) are very interesting objects to study in X-rays as they exhibit rapid evolutions in their temporal and spectral properties during outbursts. In the literature many works are reported based on the variation of spectral and temporal properties of these objects (see, e.g., 1, 3, 4 and references therein). During an outburst, transient BHCs generally show four primary spectral states: *hard (HS), hard-intermediate (HIMS), soft-intermediate (SIMS)* and *soft (SS)*. According to Debnath et al.[4], classical or type-I transient BHCs show all four spectral states in the sequence : HS→HIMS→SIMS→SS→SIMS →HIMS→HS, while harder or type-II sources do not show SS (sometimes no SIMS). These transient BHCs also exhibit low and high frequency quasi-periodic oscillations (QPOs) in some of these spectral states (see, 3 and references therein). The frequencies of QPOs are found to evolve with time during the rising and declining phases (more precisely, during HS and

HIMS) of the outbursts (see, 3, 5 and references therein). Jets or outflows are also found to be associated with hard and intermediate spectral states.

Studying spectral data with the informations of temporal nature provides a better picture about the physics of accretion around a BH. One may fit spectra with in-build or local phenomenological or theoretical model(s) in commonly used HeaSARC's spectral analysis software package XSPEC. But we believe that to find real physical picture about the accretion flow dynamics of BHs, one must study with a generalized theoretical model which should address all aspects of the observational data. Two component advective flow (TCAF) solution (see, 6, hereafter CT95) is a generalized transonic flow solution of radiative transfer equations. It consists of two types of accretion (Keplerian and sub-Keplerian) rates separated by the critical viscosity. Formation of shock (where flowing velocity suddenly jumps from supersonic to subsonic branch) is a natural solution of the model. As matter slows down at the shock location, post-shock region puffs up in the vertical direction, and forms the CENtrifugal pressure dominated BOundary Layer or CENBOL. This CENBOL acts as 'hot' Compton cloud where thermal seed photon from the Keplerian disk upscatters via inverse-Compotonization and becomes hard.

The TCAF solution has been implemented as an additive table model in HeaSARC's spectral analysis software package XSPEC to fit BH spectra (7, 8), which requires only four physical flow parameters (Keplerian disk rate $\dot{m}_d$ and sub-Keplerian halo rate $\dot{m}_h$, shock location $X_s$ and compression ratio $R = \rho_+/\rho_-$, where $\rho_+$, $\rho_-$ are post- and pre-shock matter densities respectively) excluding mass of the BH ($M_{BH}$) and normalization ($N$) to fit a spectrum from a BHC. Detailed study of several transient as well as persistent BHCs with the model, allowed us to find accurate picture of the accretion flow dynamics of these sources. Physical explanation of various observed spectral states during an outburst of a transient BHC has been done quite successfully based on evolution of the TCAF model fitted flow parameters. TCAF model fitted shock parameters ($X_s$ and $R$) also allow us to estimate frequencies of the dominating type-C QPOs (if observed in PDS; see, 7, 9), which is a temporal property. Since, in TCAF model $M_{BH}$ is an important model input parameter, one may predict mass of an unknown BHC by keeping it as a free parameter while fitting spectra. Masses of several BHCs are estimated with better accuracy from our spectral study with the TCAF model *fits* file[4,9–12]. In 13, we introduced a new method to separate jet X-ray fluxes from total X-rays based on the variation of the TCAF model normalization[13–15].

## 2. Implementation of TCAF Solution as a Table Model in XSPEC

TCAF solution has been implemented successfully as an additive local model into HeaSARC's spectral analysis software package XSPEC in 2014 (see, 7, 8). To generated table model *fits* file, a large number ($\sim 10^6$) of theoretical model spectra are used. They are generated by varying five model input parameters ($M_{BH}$, $\dot{m}_d$, $\dot{m}_h$, $X_s$, $R$) in the modified CT95 code as described in 8. A direct spectral fit

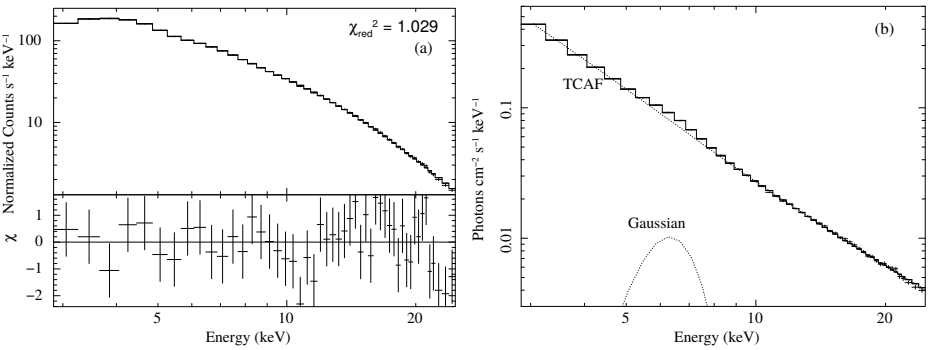



Fig. 1. (a) TCAF model fitted 2.5 − 25 keV PCA spectrum of GX 339-4 (Observation ID = 95409-01-14-04; MJD = 55300) with variation of $\Delta\chi$ is shown in the left panel. The value of the model fitted reduced $\chi^2$ is written down. (b) The unfolded model components of the spectral fit are shown on the right panel. Adopted from Debnath, Chakrabarti & Mondal (2014).

could be done using TCAF model program as a local model in XSPEC, where it will generate spectra while fitting. But this method is time consuming as to fit one spectrum may take time up to few minutes to hours. In Fig. 1(a-b), TCAF model fitted GX 339-4 spectrum of observation ID = 95409-01-14-04 and MJD = 55300 is shown. While fitting spectra, we found that TCAF model normalization ($N$) for a BHC generally do not vary (if observed with a particular instrument), since it is a function of $M_{BH}$, distance ($D$), and disk inclination angle ($i$) of the system.

## 3. Physical Explanation of Various Observational Aspects

Successful implementation of the generalized accretion flow solution TCAF into XSPEC, allowed us to find informations about the physical flow parameters directly from spectral fits. Both spectral and timing properties of BHCs are now understandable from physical point of views. One may find a more clear physical picture about the accretion processes around BHs. Measurement of intrinsic source parameters are also could be done from spectral analysis with the TCAF model.

### 3.1. *Spectral states and their evolution*

After the successful implementation of the TCAF solution as an additive table model in XSPEC, we studied accretion flow properties of several transient and persistent BHCs (see, 7, 16, 8, 4, 17, 12, 9, 10, 11, 18, 15, 19, 20, 21). Transient sources show rapid transition between different spectral states, where as persistent sources prefer to stay longer time in one states (although there has some exceptions, for e.g., GRS 1915+105). Evolution of TCAF model fitted flow parameters allow us to find physical explanation behind these nature of the sources.

Studying spectral properties with phenomenological models, for example with a combined thermal disk blackbody and non-thermal power-law models gives rough idea about the variation of thermal and non-thermal flux components without

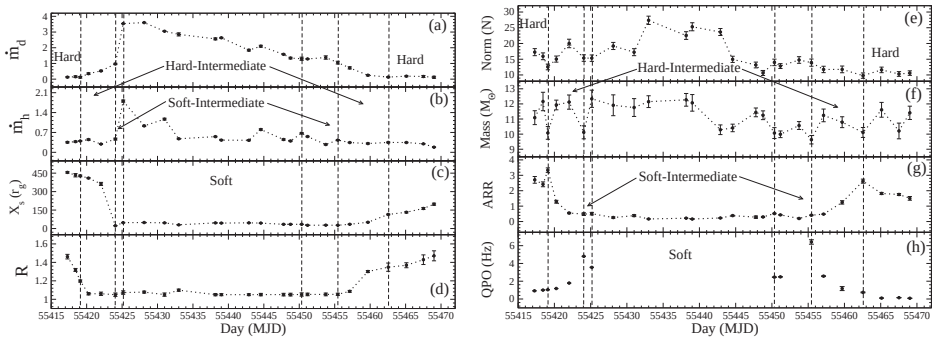



Fig. 2. In left panel, variations of TCAF model fitted parameters : (a) Keplerian disk rate $\dot{m}_d$ (in $\dot{M}_{Edd}$), (b) sub-Keplerian halo rate $\dot{m}_h$ (in $\dot{M}_{Edd}$), (c) shock location $X_s$ (in Schwarzchild radius $r_g$), and (d) compression ratio $R$ with time (in MJD) for 2010 outburst of H 1732-322 are shown. In right panel, variation of (e) TCAF model normalization $N$, (f) mass of the BH ($M_{BH}$ in $M_\odot$), (g) ARR ($=\dot{m}_h/\dot{m}_d$), and (h) observed QPO frequency (in Hz) are shown. Adopted from Mondal, Debnath & Chakrabarti (2014) and Molla et al. (2017).

knowing real physical cause of their origin. One may classify an outburst into different spectral states based on degree of importance of thermal and non-thermal fluxes and nature (shape, frequency, $Q$ value, rms% etc.) of QPOs (if present). But variation of the TCAF model fitted flow parameters, accretion rate ratio (ARR=$\dot{m}_h/\dot{m}_d$), nature of QPOs (if present) provide better physical picture on the accretion flow properties. Properties of different spectral states (HS, HIMS, SIMS, SS) and their transitions are understood more physically (see, 22 for a review).

The variation of TCAF fitted parameters ($\dot{m}_d$, $\dot{m}_h$, $X_s$, $R$, $N$, $M_{BH}$) during the 2010 outburst of H 1743-322 are shown in Fig. 2(a-f). In Fig. 2(g-h), variation of ARR and observed QPO frequency are shown. During the rising phase, as day progresses, both accretion rates increase with inward movement of the shock wave with reducing strength. Due to this process, ARR reaches maximum on HS (rising) to HIMS (rising) transition day and then rapidly reaches to a lower value on HIMS (rising) to SIMS (rising) transition day. Also, on this transition (HIMS to SIMS) day, observed frequency of the monotonically increasing QPO becomes maximum. The opposite is true during the declining phase of the outburst, where both the accretion rates decrease with time with outward movement of the shock and increasing shock strength. Here also ARR becomes maximum on HIMS (declining) to HS (declining) transition day and QPO frequency starts to decrease monotonically with time from the SIMS (declining) to HIMS (declining) transition day. During SS, no QPOs are observed with lower (roughly constant) ARR value. A somewhat similar nature of variations of TCAF model fitted physical flow parameters are observed for different spectral states of other transient BHCs as well.

### 3.2. *Prediction of type-C QPO frequency*

In 7, 9, we showed that temporal properties such as frequency of the dominating type-C QPOs could be predicted from the TCAF model fitted shock parameters

(location $X_s$ and compression ratio $R$). This is because the same shock which defines the CENBOL boundary, i.e., the size of the 'hot' Compton cloud, also causes low frequency QPOs as it oscillates. The frequency of the QPO is inversely proportional to the infall time ($t_{infall}$) in post-shock region when resonance condition between cooling and infall time scales is satisfied[23,24]. The frequency ($\nu_{QPO}$) of the observed QPOs determined by the relation,

$$\nu_{QPO} \sim t_{infall}^{-1} = C/[R\ X_s(X_s-1)^{1/2}],$$

where $C$ is a constant $= M_{BH} \times 10^{-5}$.

### 3.3. *Mass estimation of unknown BH sources*

Knowledge of the mass of the compact object (here BH) is very important while studying these objects. In TCAF model $M_{BH}$ is an important input parameter. So if $M_{BH}$ is not well known, one may keep it as free while fitting spectrum. Each spectral fit with the model *fits* file gives us to one best-fitted mass value of the source. The variation of the $M_{BH}$ (see Fig. 2e; not actual change in mass of the source, since here it is a model fitted parameter, which may vary due to fitting or data quality errors) from spectral fitting of several observations allows us to predict probable mass range of the source. In this way, recently we measured masses of several BHCs (for example: MAXI J1659-152[10], MAXI J1543–564[9], MAXI J1836-194[12], H 1743-322[11], Swift J1753.5-0127[4], XTE J1118+480[15], MAXI J1535-571[20], Cygnus X-1[21]) quite successfully in a better accuracy (narrow range).

### 3.4. *Estimation of jet X-ray fluxes*

A model normalization is a multiplicative 'factor' that converts observed spectrum to match with the theoretical model spectrum emitted from the source. In general in spectral fits, it varies. But unlike other models, TCAF model normalization does not vary in observations as it is a function of intrinsic source parameters ($M_{BH}$, $D$, $i$). In TCAF, flow parameters not only takes care of the shape of the observed spectrum but also changes in intensity or flux. Although one may require higher $N$ values if there is a jet activity or precession in disk or presence of any dominating physical processes whose effects are not included in the model *fits* file.

If variation of $N$ are not constant i.e., not in a narrow range, we look at the variation of $N$ and radio flux ($F_R$). If on the minimum $N$ observation day $F_R$ also at its minimum, then we assume that requirement of higher $N$ values in other observations, are due to excess X-ray flux from jets. Then we refit all spectra by keeping $N$ frozen at its minimum observed value and calculate fluxes, which are defined as flux contribution from accretion disk or inflowing matter ($F_{inf}$). Now taking difference of $F_{inf}$ from $F_X$ (which is the earlier calculated flux when all model parameters are kept free), we get flux contribution from jets or outflows ($F_{ouf}$).

Jana et al.[13,14] estimated jet X-ray fluxes for two Galactic BHCs: Swift J1753.5-0127, and MAXI J1836-194. Jet X-rays are found to be higher in intermediate

spectral states (here HIMS) than HS. This is consistent with the theory (see, 25). The maximum contribution of jet X-rays are found to be 32% during 2005 outburst of Swift J1753.5-0127, and 86% during 2011 outburst of MAXI J1836-194. The correlation study of $F_{ouf}$, $F_X$, $F_{inf}$ with $F_R$ using relation $F_R \sim F_X^b$, where $b$ is the correlation index. The $b$ values are found at $0.59 \pm 0.11$ for Swift J1753.5-0127 and $0.61 \pm 0.08$ for MAXI J1836-194. The correlation values are within the standard correlation limit $(0.6 - 0.7)$. Further strong correlations are found during entire outburst of MAXI J1836-194 and HS of Swift J1753.5-0127, implies compact jet. A weak correlation is found in HIMS of the 2005 outburst of Swift J1753.5-0127, implies nature of jet as blobby or discrete.

## 4. Concluding Remarks

After the successful implementation of the TCAF solution as an additive table model in XSPEC, one gets direct estimation of flow parameters. This parameters enable us not only to understand spectral properties of the BHCs but also to predict other timing and physical parameters associated with the BHs. One can see direct correlation between spectral and timing properties of transient BHCs during an outburst with the variation of TCAF model fitted/derived physical flow parameters. Most importantly we require only four parameters (assuming mass and normalization are constant) to fit spectrum. It confirms our decade long conjecture that TCAF is the most accurate theoretical model or tool to understand physics of accretion processes around a BH.

## Acknowledgments

D.D. acknowledges support from Higher Education Dept. of the Govt. of West Bengal, India; and research funds from ISRO (ISRO/RES/2/418/17-18), DST (GITA/DST/TWN/P-76/2017 and EMR/2016/003918) of the Govt. of India.

## References

1. D. Debnath, S.K. Chakrabarti and A. Nandi, et. al., *BASI*, **36**, 151 (2008)
2. D. Debnath, S.K. Chakrabarti and A. Nandi, *A&A*, **520**, 98 (2010)
3. D. Debnath, S.K. Chakrabarti and A. Nandi, *AdSpR*, **52**, 2143 (2013)
4. D. Debnath, A. Jana and S.K. Chakrabarti, et al., *ApJ*, **850**, 92 (2017)
5. A. Nandi, D. Debnath and S. Mandal, et. al., *A&A*, **542**, 56 (2012)
6. S.K. Chakrabarti and L.G. Titarchuk, *ApJ*, **455**, 623 (1995)
7. D. Debnath, S. Mondal and S.K. Chakrabarti, *MNRAS*, **440**, L121 (2014)
8. D. Debnath, S. Mondal and S.K. Chakrabarti, *MNRAS*, **447**, 1984 (2015)
9. D. Chatterjee, D. Debnath, and S.K. Chakrabarti, et al., *ApJ*, **827**, 88 (2016)
10. A.A. Molla, S.K. Chakrabarti and D. Debnath et al., *MNRAS*, **460**, 3163 (2016)
11. A.A. Molla, S.K. Chakrabarti and D. Debnath, et al., *ApJ*, **834**, 88 (2017)
12. A. Jana, D. Debnath and S.K. Chakrabarti, et al., *ApJ*, **819**, 107 (2016)

13. A. Jana, S.K. Chakrabarti and D. Debnath, *ApJ*, **850**, 91 (2019)
14. A. Jana, D. Debnath and S.K. Chakrabarti, et al., *MNRAS* (2019) (submitted)
15. D. Chatterjee, D. Debnath and A. Jana, et al., *Ap&SS*, **364**, 14 (2019)
16. S. Mondal, D. Debnath and S.K. Chakrabarti, *ApJ*, **786**, 4 (2014)
17. D. Debnath, A.A. Molla and S.K. Chakrabarti, et al., *ApJ*, **803**, 59 (2015)
18. A. Bhattacharjee, I. Banerjee and A. Banerjee, et al., *MNRAS*, **466**, 1372 (2017)
19. S.K. Chakrabart, D. Debnath and S. Nagarkoti, *AdSpR*, **63**, 3749 (2019)
20. J.R. Shang, D. Debnath and D. Chatterjee, et al., *ApJ*, **875**, 4 (2019)
21. I. Banerjee, et al., *MNRAS* (2019) (submitted)
22. D. Debnath, Chapter in Book "Exploring the Universe: From Near Space to Extra-Galactic", Springer, *ASSP*, **53**, 229 (2018)
23. D. Molteni, H. Sponholz and S.K. Chakrabarti, *ApJ*, **457**, 805 (1996)
24. S.K. Chakrabarti, S., Mondal and D. Debnath, *MNRAS*, **452**, 3451 (2015)
25. S.K. Chakrabarti, *A&A*, **351**, 185 (1999)

# Polarimetry and strong gravity effects from spots orbiting near a black hole

Vladimír Karas,* Michal Dovčiak, Jiří Svoboda, & Wenda Zhang

*Astronomical Institute, Czech Academy of Sciences, Boční II 1401, CZ-14100 Prague, Czech Republic*

Giorgio Matt

*Dip. Matematica e Fisica, Università Roma Tre, Via della Vasca Navale 84, I-00146, Rome, Italy*

Andreas Eckart, & Michal Zajaček†

*I. Physikalisches Institut, Universität zu Köln, Zülpicher Str. 77, D-50937 Cologne, Germany*
*Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, D-53121 Bonn, Germany*

We study the modulation of the observed radiation flux and the associated changes in the polarization degree and angle that are predicted by the orbiting spot model for flares from accreting black holes. The geometric shape of the emission region influences the resulting model lightcurves, namely, the emission region of a spiral shape can be distinguished from a simpler geometry of a small orbiting spot.

We further explore this scheme for the observed flares from the supermassive black hole in the context of Galactic center (Sgr A*). Our code simulates the lightcurves for a wide range of parameters. The energy dependence of the changing degree and angle of polarization should allow us to discriminate between the cases of a rotating and a non-rotating black hole.

*Keywords*: Gravitation; Black Holes; Galactic center

## 1. Introduction

Relativistic corrections to a signal from orbiting spots can lead to large rotation in the plane of observed X-ray polarization. When integrated over an extended surface of the source, this can diminish the observed degree of polarization. Such effects are potentially observable and can be used to distinguish among different models of the source geometry and the radiation mechanisms responsible for the origin of the polarized signal. The idea was originally proposed in the 1970s,[4,5,17] however, its observational confirmation and practical use in observations are still is a challenging task.

The geometrical effects of strong gravitational fields act on photons independently of their energy. The gravitational field is described by the metric of Kerr black hole[16]

$$ds^2 = -\frac{\Delta}{\Sigma}\Big(dt - a\sin^2\theta\, d\phi\Big)^2 + \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \frac{\sin^2\theta}{\Sigma}\Big[a\, dt - \big(r^2+a^2\big)\, d\phi\Big]^2 \quad (1)$$

*E-mail: vladimir.karas@cuni.cz
†New address: Center for Theoretical Physics, Polish Academy of Sciences, Warsaw, Poland

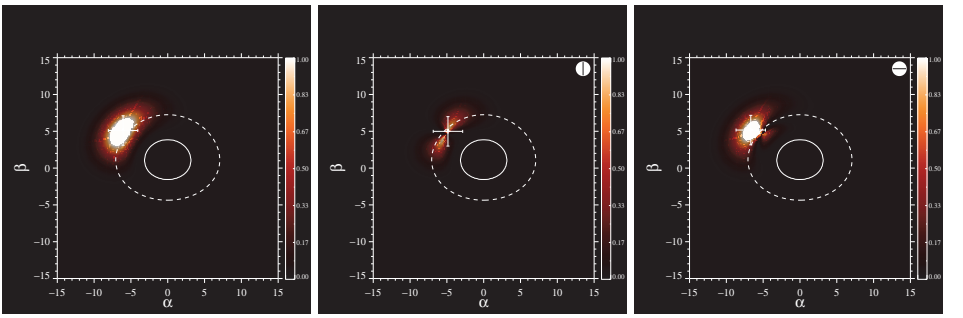



Fig. 1. An exemplary snapshot of a spot orbiting at constant radius just above the marginally stable orbit ($r = 1.1 r_{\rm ISCO}$). Left panel: the total intensity image shows the observer plane $(\alpha, \beta)$ near a non-rotating black hole, as observed at a moderate view angle ($\theta_{\rm o} = 45$ deg). The horizon radius (solid curve) and the ISCO (dashed curve) are indicated. Middle and right panels: The spot emission is assumed to be intrinsically polarized and recorded in two polarization channels, rotated by 90 degrees with respect to each other.[20,21]

in Boyer-Lindquist (spheroidal) coordinates $t$, $r$, $\theta$, $\phi$. The metric functions $\Delta(r)$ and $\Sigma(r, \theta)$ are known in an explicit form. The event horizon occurs at the roots of equation $\Delta(r) = 0$; the outer solution is found given by $r = R_+ = 1 + (1 - a^2)^{1/2}$.

Let us note that there are some similarities as well as differences between the expected manifestation of relativistic effects in polarization changes in X-rays and in other spectral bands (NIR). We show these interrelations in the associated poster and point out that the near-infrared polarization measurements of the radiation flares from the immediate vicinity of the horizon have been studied in detail from the Galactic Center (Sagittarius A*) supermassive black hole.[15,20,21]

## 2. Polarization from an Orbiting Spot

Within the scheme of the spot scenario the spots are considered to represent regions of enhanced emission on the disc surface rather than massive clumps that would decay due to shearing motion in the disc. The observed signal is modulated by relativistic effects. Doppler and gravitational lensing influence the observed radiation flux and this can be computed by GR ray-tracing. Such an approach has been developed to compute also strong gravity effects acting on polarization properties.[11]

Our code (KY) is publicly available.[7] The current version[a] allows the user to include the polarimetric resolution and to compute the observational consequences of strong-gravity effects from a Kerr black hole accretion disc. Within the XSPEC notation, this polarimetric resolution is encoded by a switch defining which of the four Stokes parameters is returned in the photon count array at the moment of the output from the model evaluation. This way one can test and combine various models, and pass the resulting signal through the response matrices of different

[a] http://stronggravity.eu/results/models-and-data/

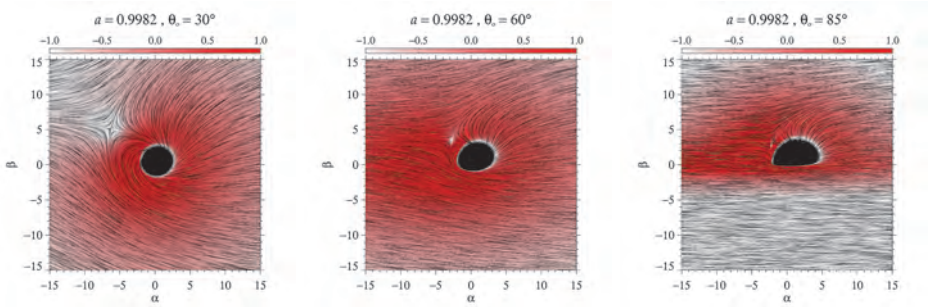



Fig. 2. A map of the distribution of the polarization angle (flow lines) and the intensity of the polarized light (colour coded in the logarithmic scale; arbitrary units are indicated by the colour bar), for a large value of spin $a$ of the Kerr black hole and different view angles of the observer, as indicated on top of the panels. This time the model has been computed for the case of thermal radiation of an equatorial accretion disk polarized via Comptonization; see Dovčiak et al. (2008).[8] A critical point is clearly seen in the flow-line structure through which then the orbiting spot circles.

instruments. As an illustration, figure 1 shows a single time-frame from a simulation of a spot rotating rigidly at constant radius.

The detected polarization degree is expected to decrease mainly in the part of the orbit where the spot moves close to the region where the photons are emitted perpendicularly to the disc and the polarization angle changes rapidly. The drop of the observed polarization degree occurs also in those parts of the orbit where the magnetic field points approximately along light ray (see fig. 2 for a map showing the distribution of the polarization effects within the disk plane of a rapidly rotating Kerr black hole). For the more realistic models the resulting polarization exhibits a very complex behaviour due to, e.g., Faraday rotation effects interfering with strong-gravity.[10] Among generic features is the peak in polarization degree for the spin parameter $a \to M$ and large inclinations, caused by the lensing effect at a particular position of the spot in the orbit. This effect disappears in the non-rotating $a = 0$ case.

To conclude, let us note that the above described modelling has recently gained new impetus in the context of forthcoming X-ray polarimetry satellite missions that have been under active consideration. The enhanced X-ray Timing and Polarimetry (eXTP) mission[9,22] is the concept of novel Chinese X-ray mission that has currently reached the extended phase A with a possible contribution from Europe and the anticipated launch date by 2025.

The Large Area Detector (LAD) aboard the new mission will be a suitable instrument for performing time-resolved spectroscopy with polarimetric resolution, large effective area, and moderate energy resolution to explore bright flares from accreting black holes, possibly in the connection with the processes of tidal disruption of stars and similarly violent events.[6,23–25]

Let us just remind the reader that the revealing role of X-ray polarimetry clearly extends to all scales far away from the central black hole. This has been widely discussed in the context of reflection nebulae surrounding Galactic center.[3,14] The combination of the polarization signal from the Sgr B and Sgr C complexes strengthens the need for an imaging detector with a fine spatial resolution to resolve the structures as small as the Bridge clouds. Therefore, the X-ray polarimetry is needed to explore variety of physical processes operating near supermassive black hole.

## 3. Conclusions

We explored the approach based on mapping the Kerr black hole equatorial plane onto the observer's plane at radial infinity. Orbiting spots are projected onto the disc plane (hence imposing the vertically averaged approximation) and then their image is transported towards a distant observer. The strong gravity effects can be seen as the predicted (time-dependent) direction of polarization is changed by light propagation through the curved spacetime. What can be foreseen in the near future is the tracking of the wobbling image centroid that a spot produces. With the polarimetric resolution, this wobbling can provide the evidence of orbiting features.

In conclusion, the rotating spots are a viable scenario capable to explain the occurrence about once per day of modulated flares from within a few milli-arcseconds of the Sagittarius A* supermassive black hole.[12,18] However, the astrophysically realistic scenarios have to account not only for the time-scales of gradual dispersion of the orbiting features by tidal effects[1,19] but also for the effects of plasma influencing the light propagation through a non-vacuum spacetime near the black hole.[2,13]

## Acknowledgments

The authors acknowledge the Czech Ministry of Education, Youth and Sports (MŠMT) project INTER-INFORM (ref. LTI 17018), titled "Promotion and Development of International Scientific Cooperation in Relativistic Astrophysics and Preparation of X-ray Space Missions".

## References

1. Abramowicz M. A., Lanza A., Spiegel E. A. and Szuszkiewicz E. (1992), *Nature*, 356, 41.
2. Bičák J. and Hadrava P. (1975), *A&A*, 44, 389.
3. Churazov E., Sunyaev R., & Sazonov S. (2002), *MNRAS*, 330, 81.
4. Connors P. A and Stark R. F. (1977), *Nature* 269, 128.
5. Connors P. A., Stark R. F. and Piran T. (1980), *ApJ*, 235, 224.
6. De Rosa A., Uttley P., Gou L. J. and Liu Y (1989), *Science China – Physics, Mechanics & Astronomy*, 62, id. 29504.
7. Dovčiak M., Karas V. and Matt G. (2004), *MNRAS*, 355, 1005.

8. Dovčiak M., Muleri F., Goosmann R. W., Karas V. and Matt G. (2008), *MNRAS*, 391, 32.
9. Feroci M., Ahangarianabhari M., Ambrosi G., Ambrosino F. et al. (2018), *Proceedings of the SPIE*, 10699, id. 106991C.
10. Jiménez-Rosales A. and Dexter J. (2018), *MNRAS*, 478, 1875.
11. Karas V., Kojima Y. and Kunneriath D. (2018), *Contrib. of the Astron. Obs. Skalnaté Pleso*, 48, 446 (arXiv:1801.10203).
12. Karssen G. D., Bursa M., Eckart A., Valencia-S M., Dovčiak M., et al. (2017), *MNRAS*, 472, 4422.
13. Kimpson T., Wu Kinwah and Zane S. (2019), *MNRAS*, accepted (arXiv:1901.03733).
14. Marin F., Karas V., Kunneriath D., & Muleri F. (2014), *MNRAS*, 441, 3170.
15. Meyer L., Eckart A., Schödel R., Duschl W. J., Mužić K., et al. (2006), *A&A*, 460, 15.
16. Misner C. W., Thorne K. S. and Wheeler J. A. (1973), *Gravitation* (Freeman, San Francisco).
17. Pineault S. (1977), *MNRAS*, 179, 691.
18. Shahzamanian B., Eckart A., Zajaček M., Valencia-S. M. and Sabha N. (2018), *Galaxies*, 6, 13.
19. Zakharov A. F. (1994), *MNRAS*, 269, 283.
20. Zamaninasab M., Eckart A., Dovčiak M., Karas V., Schödel R. et al. (2011), *MNRAS*, 413, 322.
21. Zamaninasab M., Eckart A., Witzel G., Dovčiak M., Karas, V., et al. (2010), *A&A*, 510, A3.
22. Zhang S. N., Feroci M., Santangelo A., Dong Y. W. et al. (2016), *Proceedings of the SPIE*, 9905, id. 99051Q.
23. Zhang W., Yu W., Karas V. and Dovčiak M. (2016), *ApJ*, 807, id. 89.
24. Zhang W., Yu W., Karas V. and Dovčiak M. (2019), *ApJ*, submitted.
25. Zhang S. N., Santangelo A., Feroci M., and Xu Y. P. (2019), *Science China – Physics, Mechanics & Astronomy*, 62, id. 29502.

# New insights into the plasma region near black holes from hard X-ray observations

Piotr Lubiński

*Institute of Physics, University of Zielona Góra,*
*Licealna 9, PL-65-417 Zielona Góra, Poland*
**E-mail: P.Lubinski@if.uz.zgora.pl*

Two examples of the analysis of the hard X-ray/soft $\gamma$-ray emission from accreting black holes observed with the *INTEGRAL* satellite are presented. The first study is based on data covering 16 years of monitoring of the bright persistent Cyg X-1 system. This huge data set allowed us to get more insights into the geometry of the plasma region in different spectral states of the object. In particular, we identify two substates of the hard state and provide new information on the plasma behaviour during the jet emission. The second study is based on data taken during the 2015 outburst of transient black hole binary V404 Cygni. The results of detailed spectral analysis show that a single Componization component, purely thermal or hybrid, is not sufficient to explain the hard X-ray emission during a strong radio emission period.

*Keywords*: Accreting black holes; spectral states; hard X-rays; $\gamma$-rays.

## 1. Introduction

The properties of the plasma residing in the centre of the black hole (BH) accreting systems remain an open question despite several decades of observational and theoretical efforts[1]. Since the plasma is emitting mainly in the hard X-ray/soft $\gamma$-ray band, the INTEGRAL satellite[2] operating in that band is currently one of the best tools for a direct diagnosis of the plasma properties. In this talk a review of the results of two examples of such studies will be presented. The first study is related to the spectral states of the well-known persistent BH binary Cyg X-1. The second study involves the hard X-ray emission observed from the V404 Cygni transient source during its recent outburst in 2015.

## 2. Spectral states of Cyg X-1

Characterization of the spectral states of the BH systems is a common method to get deeper insights into to the nature of these systems[1]. In general, both transient and persistent systems evolve between two main states, hard and soft, with several intermediate state classes observed. The hard state corresponds to the emission from a hot plasma dominating in the X-ray band, well modeled with a thermal Comptonization of some seed photons. In the soft state the plasma emission appears weaker and rather non-thermal, and the soft band is dominated by a thermal emission from the accretion disk. Spectral states are usually classified through one or two parameters derived from the soft X-ray observations, i.e., in the band below $\approx$10 keV. In the case of X-ray monitors providing count rate in several energy bands the state is identified with the use of the count rate and ratio of count rates in two energy bands (hardness ratio). These quantities are commonly plotted in a form of so-called hardness-intensity diagram (HID) showing the spectral evolution

of the object. The HID data allow for a quite good separation of the states as demonstrated for example for Cyg X-1[4].

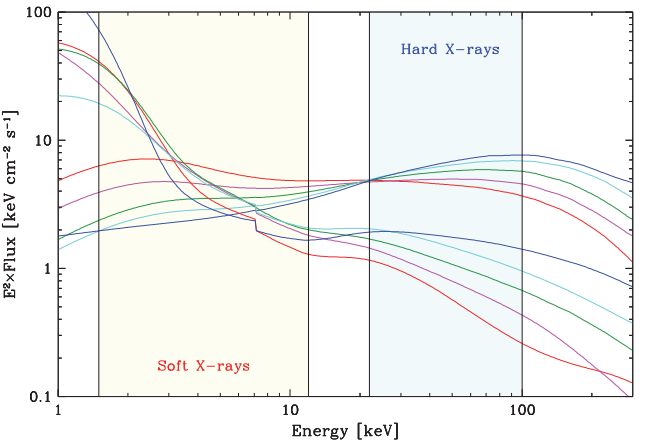



Fig. 1. Models fitted to twelve INTEGRAL's summed spectra of Cyg X-1, with data selected according to the hard X-ray flux and hardness. Yellow region shows typical energy range used to classify the spectral states. Cyan region is the hard X-ray band used in our study.

The soft X-ray band is a natural choice for spectral states studies because of a large number of photons emitted from the source and relatively low instrumental background. On the other hand, the X-ray/soft $\gamma$ spectra of accreting BHs are quite complex because besides the two dominating components, i.e., the disk and plasma emission, there are other processes involved, such as absorption or reprocessing in surrounding media, including Compton reflection and iron line emission. These additional components affect mostly the soft X-ray band, whereas for objects with a limited reflection the hard X-ray emission is mainly governed by the plasma. Therefore, the plasma properties can be studied in a quite unambiguous way using the emission above $\approx$ 20 keV, not affected by the disk emission and absorption.

In Fig. 1 we present models fitted to the Cyg X-1 spectra collected with the *INTEGRAL*'s IBIS and JEM-X telescopes. These spectra were summed over observations done in the 2002–2017 period, with a selection based on the count rates in the 22–40 and 40–100 keV bands. It can be clearly seen that the hard X-ray band enables a better separation of the hard and soft spectral states than the soft X-ray band. This motivated us to do a massive analysis of several spectral parameters obtained in that band with the ISGRI detector over 16 years of *INTEGRAL*'s observations of Cyg X-1. Since the ISGRI calibration evolved over so many years, the spectral fitting with responses taking into account that evolution was done to extract the photon index $\Gamma_{\rm H}$ and flux $F_{\rm H}$ for the 22–100 keV band. In total, we extracted the data for 7821 observations lasting 0.5–2.0 hours (so-called science windows). These results are presented in Fig. 2. The soft and hard state are separated by the flux level of about $75{\times}10^{-10}$ erg sm$^{-2}$ s$^{-1}$. In addition, for the hard state we

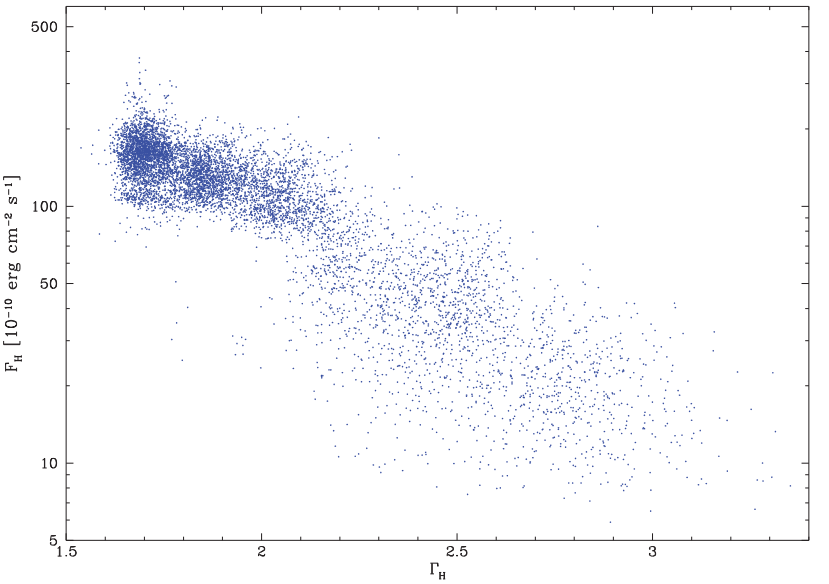



Fig. 2. Hard X-ray photon index $\Gamma_H$ - hard X-ray flux $F_H$ diagram. Both quantities were determined for the 22–100 keV band.

observe a clear clustering around $\Gamma_H$ equal to 1.68, 1.85 and 2.02. We called these substates pure hard, transitional hard and hard intermediate, respectively.

To characterize better these states we did a Spearman rank-order test for the $\Gamma_H$-$F_H$ data over each single *INTEGRAL* orbit (about 3 days), determining the correlation coefficient $r_S$ and null hypothesis probability $P_S$. As shown in Fig. 3 the flux - photon index correlation varies with increasing $\Gamma_H$. For pure hard state as well as for the softest data we do not observe correlation whereas it is quite strong for the transitional hard state and then slowly decreases when the spectra become softer. Such behavior implies that the pure hard state and softest state are physically different from the rest. Interpreted in terms of the Comptonization scenario a lack of correlation points toward self-synchrotron source of the seed photons instead of the disk photons.

Cyg X-1 was quite frequently observed during the studied period at the 15 GHz radio band by the Ryle telescope[5] and its successor, the AMI Array[10]. The radio flux $F_R$ plotted against the hard X-ray photon index determined for contemporary *INTEGRAL* spectra shows interesting features. For all three hard substates we observe a stratification of $F_R$, with the relatively low range seen for the pure hard state, then higher values for the transitional state and the highest radio emission during the hard intermediate state. For softer spectra the radio flux slowly decreases and vanishes for the softest state. Such an interdependence reinforces our findings for Cyg X-1: the hard X-ray data reveal several substates of the hard state in that object. In the forthcoming paper[3] we will present details of our analysis and a deeper interpretation of the results.

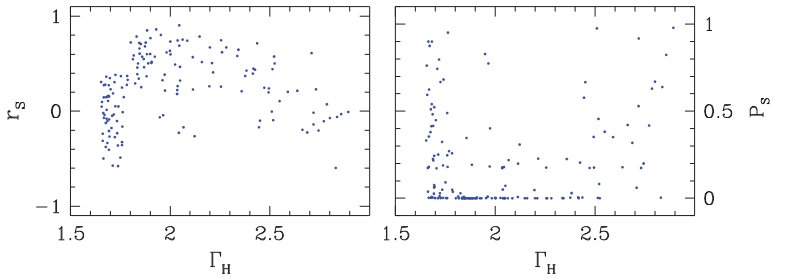



Fig. 3. Correlation between the hard X-ray photon index and flux in a function of the photon index. Left panel shows the Spearman correlation coefficient $r_S$, right panel corresponding null hypothesis probability $P_S$.

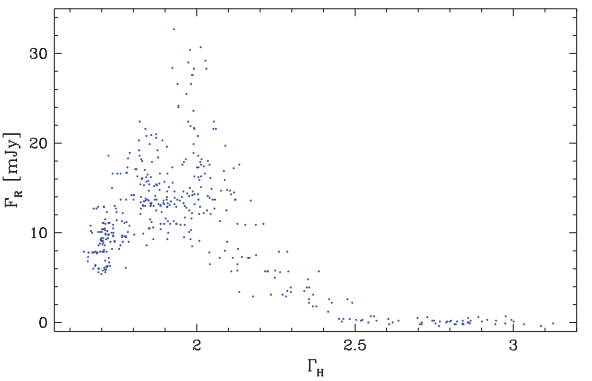



Fig. 4. Radio flux at 15 GHz measured by the Ryle and AMI observatories plotted against the hard X-ray photon index.

## 3. Evolution of V404 Cygni emission during the 2015 outburst

The transient BH system V404 Cygni enters an outburst phase every $\approx 30$ years, showing some properties specific to this object. The emission is highly variable and strongly absorbed when compared to the other transient systems[11]. Its recent outburst in 2015 was observed by many observatories[9], what resulted in dozens of papers presenting interesting results. Analysis of the hard X-ray spectra from *INTEGRAL* was already presented, however, the research was done in a rather limited band, without the soft X-rays[7,8]. Here we present preliminary results of a broad-band analysis of spectra from the JEM-X, ISGRI and PICsIT detectors, covering the 3–1000 keV band.

Figure 5 presents the hard X-ray and soft $\gamma$-ray light curves of the 2015 outburst. An interesting feature is that the emission was extremely variable during the satellite orbits 1554, 1554 and first part of 1556, whereas later the count rates varies rather smoothly, with a clear detection of emission above 250 keV. Taking into account the single science window spectra we have summed up the data for periods with similar overall spectral shape. The resulting 18 spectra were fitted with the complex

model including hybrid Comptonization, Compton reflection and several absorption components. Examples of these spectra together with the fitted model are shown in Fig. 6. The $\gamma$-ray emission during the bright phase of the 1557 orbit is detected up to almost 1 MeV, revealing a hard power-law like tail in addition to the hybrid plasma model.

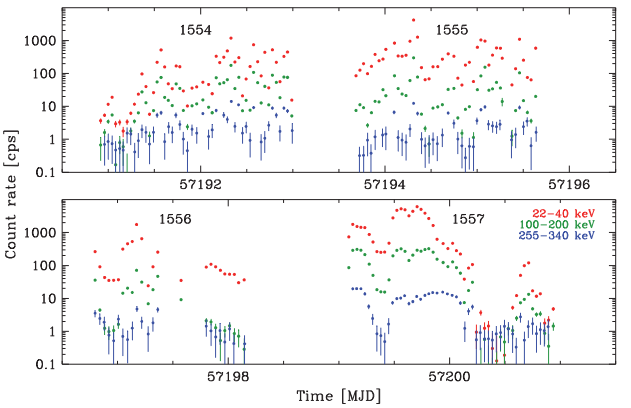



Fig. 5. Variability of the V404 Cygni hard X-ray/*gamma* emission during the 2015 outburst. The 22–40 keV and 100–200 keV count rates were measured by *INTEGRAL*/ISGRI and the 255–340 keV count rate by *INTEGRAL*/PICsIT.

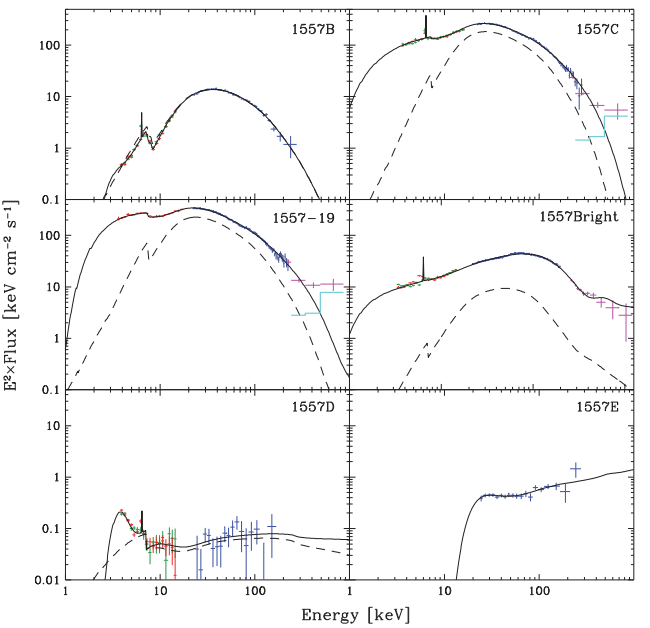



Fig. 6. Examples of V404 Cygni spectra collected by the *INTEGRAL* satellite in the 3–1000 keV band. Total fitted model is shown with a solid line and the reflection component with a dashed line. Cyan steps show the 3-$\sigma$ detection limits for the PICsIT detector data (magenta).

Besides the hard tail seen in several cases, all the spectra are well modeled with the Comptonization model, either thermal or hybrid with a non-thermal component. A presence of the additional tail component appears to be correlated with an enhanced radio emission[9], indicating its relation to the strong jet. Almost all spectra are dominated by the Compton reflection component. Its low ionization state points toward reprocessing in a dense and cold medium what seems to be puzzling taking into account the X-ray brightness of the object. We do not observe a correlation of the absorbing column density with the flux level, thus the strong variability seems to be driven by some phenomenon other than varying absorber. Further details of this research will be presented in a forthcoming paper[6].

## Acknowledgments

The author have been supported by the Polish National Science Centre (NCN) grant UMO-2014/13/B/ST9/00570. This research was based on observations with *INTEGRAL*, an ESA project with instruments and science data centre funded by ESA member states (especially the PI countries: Denmark, France, Germany, Italy, Switzerland, Spain), the Czech Republic, and Poland and with the participation of Russia and the USA.

## References

1. T. M. Belloni, in *Lecture Notes in Physics* **794**, 53 (Berlin Springer Verlag), (2010).
2. C. Winkler, *et al.*, *A&A* **411**, L1 (2003).
3. A. Filothodoros, *et al.*, *ApJ*, submitted.
4. V. Grinberg, *et al.*, *A&A* **554**, A88 (2013).
5. M.E. Jones, in *IAU Colloq. 131: Radio Interferometry. Theory, Techniques, and Applications ASPC* **19**, 395 (1991).
6. P. Lubiński, *et al.*, *ApJ*, in preparation.
7. J.-P. Roques, *et al.*, *ApJ* **813**, 22 (2015).
8. C. Sánchez-Fernández, *et al.*, *A&A* **602**, A40 (2017).
9. A.J. Tetarenko, *et al.*, *MNRAS* **469**, 3141 (2017).
10. J.T.L. Zwart, *et al.*, *MNRAS* **391**, 1545 (2008).
11. P.T. Życki, *et al.*, *MNRAS* **309**, 561 (1999).

# Accretion around low mass and supermassive black holes with TCAF

S. Mondal*

*Ben-Gurion University of the Negev,*
*Beer Sheva, Ben-Gurion, Israel*
** E-mail: santanu@post.bgu.ac.il*

S. K. Chakrabarti

*Indian Centre for Space Physics,*
*Kolkata, West Bengal 700084, India*
*www.csp.res.in*

P. Nandi

*S. N. Bose National Centre for Basic Sciences,*
*JD Block, Sector III, Kolkata, West Bengal 700106, India*
*www.bose.res.in*

Observational evidences show that during the outburst, black hole binaries move to different spectral states depending on accretion flow that feeds these objects. However proper physical picture is still unclear, what causes these states change and how the quasi periodic oscillation (QPO) in observed lightcurve is related with the disk properties? In a Two Component Advective Flow (TCAF) solution it is established that low angular momentum and low viscous sub-Keplerian flow produces hot corona, which up-scatters soft photons from the Keplerian disk. We analyze the outbursts data of black holes of different scales ranging from low mass to super massive observed by different satellites with TCAF and discuss the outburst features and their spectral states.

*Keywords*: X-Rays:binaries; Black Holes; shock waves; accretion; accretion disks; Radiation:dynamics

## 1. Introduction

Accreting black hole spectra consist different components, e.g. blackbody, power-law and an iron line at around 6.4 keV. These distinct components are mainly from the optically thick and thin components of the flow. It is well established that the spectra change their shape during the outburst from hard to soft through intermediate states. Several attempts were made to explain changes in the spectral shape and its variation (Remillard & McClintock 2006 for a review). It is believed that the changes in the accretion rate might be responsible for the changes in the spectral states (Maccarone & Coppi 2003). However, causes of the change in accretion rate on a progressive days, the origin of corona and its temperature, optical depth etc. were unclear.

The problems were satisfactorily resolved when proper usage of the solution of transonic flows in presence of cooling, heating and viscosity were used. In a Two Component Advective Flow (TCAF, Chakrabarti & Titarchuk 1995, hereafter CT95) solution, a Keplerian disk which arises out of higher viscosity is flanked by a hot sub-Keplerian flow of lower viscosity. This sub-Keplerian, low angular

momentum hot matter forms an axisymmetric shock due to the balance of the centrifugal force with gravity (Chakrabarti 1989). The subsonic region between the shock (centrifugal barrier) to the inner sonic point is hot and puffed up and behaves as the Compton cloud, which reprocesses intercepted soft photons from the standard disk. This region could be oscillatory when its cooling time scale roughly matches with the infall (compression) timescale inside CENBOL (Molteni et al. 1996). Recently, Chakrabarti et al. (2015) applied the resonance condition for H 1743-322 black hole candidate and showed that the low frequency quasi-periodic oscillations (LFQPOs) are produced when cooling time scale roughly matches with the heating time scale. Transonic solution by Mondal & Chakrabarti (2013; CT95) shows that cooling mechanism is also responsible for the change in spectral states during the outburst.

After the implementation of TCAF in XSPEC (Debnath et al. 2014) and fitting data of several black hole candidates (BHCs), one obtains physical parameters of the flow, such as the accretion rates of the disk and halo components, the shock location and the shock compression ratio. If the mass of the candidate is unknown, this will also be found out from the spectral fitting (e.g., Molla et al. 2016; 2017). A plot of photon count variation with accretion rate ratio (ARR = halo rate/disk rate) gives the so call ARR intensity diagram (ARRID, Mondal et al. 2014; Jana et al. 2016) and directly shows why the spectral state changes. The changes in accretion rates on a progressive days is due to changes in viscosity parameter ($\alpha$) of the flow during the outburst (Mondal et al. 2017). The presence of double horn Iron line is also observed from the TCAF (Mondal et al. 2016) during the fitting of high resolution NuSTAR data of GX 339-4 BHC, which helps to determine the spin parameter of the candidate. It should be noted that so far TCAF is implemented for low mass black hole binaries as an additive table model. However, very recently, the model is directly implemented in XSPEC as a local model to analyze the active galactic nuclei (AGNs, Nandi et al. 2019) data.

The *paper* is organized in the following way: in the next Section, briefly we discuss the observation and data analysis procedure for both Swift and NuSTAR satellites. In §3, we present our results of spectral fitting for both low mass and super massive BHs. We also discuss how viscosity is responsible for the change in spectral states. Finally we draw our concluding remarks.

## 2. Observation and data analysis

In the present manuscript, we analyze both Swift/XRT (Gehrel et al., 2004) and NuSTAR (Harrison et al. 2013) satellite observations of the BHC Swift J1357.2-0933 during its 2017 outburst. The Swift/XRT observation we analyzed 00031918066 (Windowed Timing mode, WT) with exposure time 0.7 ksec and NuSTAR observations with observation ID: 90301005002, with exposure time 45ksec. To fit the data with the TCAF model in XSPEC, we use a TCAF model generated *fits* file (Debnath et al. 2014). We use the absorption model *TBabs* with hydrogen column

density fixed at $1.3\times 10^{21}$ atoms cm$^{-2}$ throughout the analysis for the SwiftJ1357.2. We analyze *NuSTAR* data with exposure times longer than 5 ksec for NGC 4151 with observation date from November 11 to 14, 2012. The observation IDs are 60001111002, 60001111003 and 60001111005 respectively. For the spectral fitting of AGN data with TCAF, we run TCAF directly in XSPEC as a local model (see Nandi et al. 2019). The model was originally introduced by Chakrabarti in 1995 for AGN and a discussions of satisfactory fits of AGN spectra are present in the literature (Mandal & Chakrabarti 2008). For the data processing and analysis, we follow the standard tasks discussed in Mondal et al. (2016).

## 3. Results and Discussions

In this section we discuss about the model fitted results and the geometry of the flow on the basis of the fitted parameters. We consider Swift J1357.2-0933, a low mass candidate and NGC 4151, an AGN for our study.

### 3.1. *Swift J1357.2-0933*

The source was first detected in 2011 by the *Swift* Burst Alert Telescope (Krimm et al. 2011). We fit the observation of Swift J1357.2-0933 with TCAF solution based fits file which uses five physical parameters (if mass is unknown): The parameters are as follows (i) Mass of the black hole, (ii) disk accretion rate, (iii) halo accretion rate, (iv) location of the shock, and (v) shock compression ratio. All the parameters collectively give the electron density and temperature, photon spectrum and density, the fraction of soft photons intercepted by the Compton cloud from the Keplerian disk, as well as the reflection of hard photons from the Compton cloud by the disk. In Chakrabarti (1997) showed a detailed spectral variation with different model parameters. In Fig. 1 (left panel), we present the TCAF model fitted spectrum for this candidate which covers a broadband energy range from 0.5 - 70 keV. The data fits well for $M_{\odot} = 4.01$, $\dot{m}_d = 0.019$, $\dot{m}_h = 0.213$, $X_s = 36.7\ r_g$, R=2.0 with $\chi^2/dof$=473.1/446.

The model fitted parameters show that the halo rate is higher than the disk rate and the shock compression ratio is in the regime of strong mass loss (see Chakrabarti 1999). This is an indication of hard spectral state. If we follow the previous observed day and fit the data with model, it shows that, at some point in time after the first observation day, viscosity may have started to go up, and the Keplerian disk rate also started to increase (see Mondal & Chakrabarti 2019). However, this was not enough so as to change the spectral state (as a minimum viscosity is required for such changes (Mondal et al. 2017)). We also fit the data with power-law (PL) model. The PL model fitted photon index (=1.8) also indicates a hard spectral state. Following the above model understanding and the values of the model fitted parameters, we conclude that the source was in rising hard state of the outburst during the observations. We also estimate some physical parameters of the disk

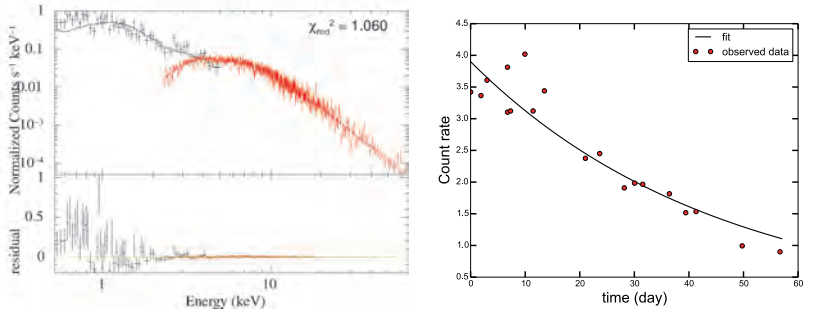



Fig. 1. Left panel: Swift/XRT and NuSTAR (focal plane module A, FPMA) data from 0.5-70.0 keV energy range are fitted with an absorbed TCAF model. Right panel: Variation of count rate with time (day). Black solid line shows the model fit and filled circles are observed count rates. Data are fitted with $H/R = 0.024$ (from above), $P = 2.8$ hrs, $q = 0.03$ and SS73 $\alpha = 0.25$. Data and figure are presented from MC19.

and check the stability of the disk to produce a sufficient viscosity to explain the observed lightcurve from the TCAF model fitted parameters. From the Kepler's law, one can derive a relation between orbital period ($P$) and orbital separation $a$. Outer disk radius ($r_{out}$) of the primary star is calculated from the Roche lobe radius of the primary following Eggleton (1983). In this work, we consider that the outer edge of the disk ($r_{out}$) is 70% of the Roche lobe radius. which appears to be $r_{out} = 0.68 \times 10^{11}$. The steps followed to fit the observed lightcurve are: (i) we have mass and disk accretion rate from the model fit, (ii) using (i) we estimate the temperature of the disk thus the sound speed, (iii) outer disk radius is estimated from orbital period and mass ratio from the literature, (iv) once (ii) and (iii) are known one can estimate height of the disk, Keplerian angular frequency, and disk kinematic viscosity to estimate the viscous time scale, and (v) finally, we extract SS73-$\alpha$ parameter ( Shakura & Sunyaev 1973) for which the derived and the fitted $\tau$ values are consistent. Here, $\alpha$ takes the value 0.25 to give a decay timescale ($\tau$) of $\sim$ 45 days. To fit the observed lightcurve using decay timescale, we use an exponential decay function, $f = A\ exp(-t/\tau)$, where $A$ is a normalization constant, which takes the value of $3.91 \pm 0.16$ with an exponential decay timescale ($\tau$) around $45.28 \pm 4.78$ *days*. The estimated SS73 disk viscosity parameter ($\alpha = 0.25$) satisfies the same order of magnitude obtained for other observed candidates (Mondal et al. 2017). In Fig. 1 (right panel), we show exponential decay function fitted with the observed lightcurve. As the halo rate is higher than the disk rate and the shock compression ratio is always greater than unity for this outburst, inverse-Comptonization process is dominant which cools down the system and the shock moves inward (Mondal & Chakrabarti 2013; CT95). The velocity of the shock movement is $\sim$ 0.15 m/s for this candidate which is similar to the shock velocity in other outbursts (Mondal et al. 2015 for H 1743-322 and references therein). This indicates that the outburst probably remained in the rising phase and no other

spectral state has been missed in between these $\sim 40$ *days*. Our model fitted disk rate is $\sim 2\%$ of $\dot{M}_{Edd}$. As the shock compression ratio is intermediate, the jets and outflows are expected to be strong with mass outflow/inflow rate ratio 3.4-4.2%.

### 3.2. *NGC 4151*

The Seyfert 1.5 galaxy NGC 4151 (z=0.00332, de Voucouleurs et al. 1991), sometimes considered as a Seyfert 1 type AGN, is one of the most popular sources for which many AGN phenomena were first characterized (for more details, see Ulrich 2000). Here we study spectral properties of NGC 4151 with TCAF using three *NuSTAR* observations during 2012. The outburst data considered in this work is fitted with TCAF in addition with a power-law (PL) model. In Fig. 2(a-c), we show the TCAF model fitted 3.0-70.0 keV spectra along with residual in the bottom panel for three observation IDs. In Fig. 2d, unabsorbed model spectra are shown which are used to fit the observed data.

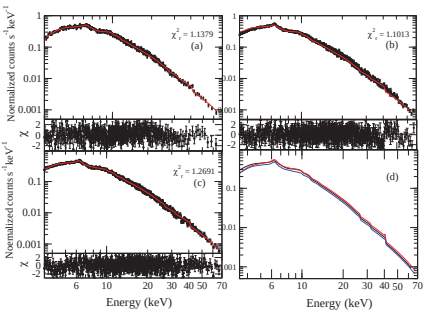



Fig. 2. TCAF model fitted 3.0 - 70 keV *NuSTAR* spectra with variation of $\chi^2$, for three IDs. (a) 60001111002, (b) 60001111003 and (c) 60001111005. In (d), unabsorbed TCAF+PL model generated spectra are shown, which are used to fit the observation. Figure is adopted from Nandi et al. (2019).

From the spectral fit we estimate the mass of the central black hole to be $3.03^{+0.26}_{-0.26} \times 10^7 M_\odot$ which appears to be in the same ball park as obtained by previous studies. Furthermore, unlike other models, we obtained the two accretion rates and found that the halo rate is much larger as compared to the disk accretion rate, consistent with the fact that the object was always in the hard state and the fact that AGN usually accretes low angular momentum matter, unless a part is also converted to Keplerian by a sufficient viscosity. Analysis also shows that in all these three observations, the $X_s$ (i.e. corona) was more or less constant $\sim 150\ r_s$. This is partly due to the fact that in a supermassive black hole system short term variabilities are rare. However to get the more detail understanding about the evolution of the accretion flow, we need to have a long term monitoring. The details about the model parameters variation can be seen in Nandi et al. (2019).

**Acknowledgments**

This work was supported by the IAEC-UPBC joint research foundation (grant No. 300/18), and by the Israel Science Foundation (grant No. 1769/15).

**References**

1. Chakrabarti, S. K., 1989, *MNRAS*, **340**, 7
2. Chakrabarti, S.K. & Titarchuk, L.G., 1995, *ApJ*, **455**, 623
3. Chakrabarti, S.K., 1999, *A&A*, **351**, 185
4. Chakrabarti, S.K. Accretion Disks in Active Galaxies: The Sub-Keplerian Paradigm, in Ann. NY Acad. Sci., 759, Seventeenth Texas Symposium on Relativistic Astrophysics and Cosmology (1995) ed. H. Bohringer, G.E. Morfil and J. Trumper, **546**
5. Chakrabarti, S. K., 1997, *ApJ*, **484**, 313
6. Chakrabarti, S.K., Mondal, S., & Debnath, D., 2015, *MNRAS*, **452**, 3451
7. de Voucouleurs, G., de Voucouleurs, A., Corwin, H. G., Jr., et al. 1991, Third Reference Catalogue of Bright Galaxies (Berlin:Spinger)
8. Debnath, D., Chakrabarti, S. K., & Mondal, S., 2014, *MNRASL*, **440**, 121
9. Eggleton, P. P. 1983, *ApJ*, **268**, 368
10. Gehrels et al., 2004, *ApJ*, **611**, 1005
11. Harrison, F. A., et al., 2013, *ApJ*, **770**, 103
12. Jana, A., Debnath, D., Chakrabarti, S. K., et al. 2016, *ApJ*, **819**, 107
13. Krimm, H. A., et al. 2011, The Astronomer's Telegram, **3138**
14. Mandal S., & Chakrabarti S. K., 2008, *ApJ*, **689**, L17
15. Molteni, D., Sponholz, H. & Chakrabarti, S.K., *ApJ*, **457**, 805 (1996)
16. Maccarone, T. C., & Coppi, P. S. 2003, *MNRAS*, **338**, 189
17. McClintock, J. E. & Remillard, R. A. 2006, in Compact Stellar X-ray Sources, Cambridge, Astrophysical Ser., vol. 39, ed. W. Lewin & M. van der Klis (Cambridge Univ. Press), 157
18. Molla, A. A., et al., 2016, *MNRAS*, **460**, 3163
19. Molla A. A., Chakrabarti S. K., Debnath D., Mondal S., 2017, *ApJ*, **834**, 88
20. Mondal, S., & Chakrabarti, S. K., 2013, *MNRAS*, **431**, 2716
21. Mondal, S., Debnath, D. & Chakrabarti, S. K., 2014, *ApJ*, **786**, 4
22. Mondal, S., Chakrabarti, S.K., & Debnath, D., 2015, *ApJ*, **798**, 57
23. Mondal, S., Chakrabarti, S. K., & Debnath, D. 2016, *Ap&SS*, **361**, 309
24. Mondal, S., Chakrabarti, S.K., Nagarkoti, S. et al., 2017, *ApJ*, **850**, 47
25. Mondal, S., & Chakrabarti, S. K., 2019, *MNRAS*, **431**, 2716 (MC19)
26. Nandi, P., Chakrabarti, S. K., & Mondal, S., 2019, *ApJ*, **877**, 65
27. Shakura, N. I., & Sunyaev, R. A. 1973, *A&A*, **24**, 337
28. Ulrich, M.-H., 2000, *A&ARv*, **10**, 135

# ULXs as magnetized sub-Eddington advective accretion flows around stellar mass black holes

Banibrata Mukhopadhyay

*Department of Physics, Indian Institute of Science,*
*Bangalore 560012, India*
*E-mail: bm@iisc.ac.in*
*www.iisc.ac.in*

Ultra-luminous X-ray sources (ULXs) have been puzzling us with a debate whether they consist of an intermediate mass black hole or super-Eddington accretion by a stellar mass black hole. Here we suggest that in the presence of large scale strong magnetic fields and non-negligible vertical motion, the luminosity of ULXs, particularly in their hard states, can be explained with sub-Eddington accretion by stellar mass black holes. In this framework of 2.5D magnetized advective accretion flows, magnetic tension plays the role of transporting matter (equivalent to viscous shear via turbulent viscosity) and we neither require to invoke an intermediate mass black hole nor super-Eddington accretion. Our model explains the sources, like, NGC 1365 X1/X2, M82 X42.3+59, M99 X1 etc. which are in their hard power-law dominated states.

*Keywords*: accretion disks; ULXs; magnetic fields; black holes.

## 1. Introduction

While the existence of stellar mass black holes of mass $M \lesssim 80M_\odot$ and supermassive black holes of mass $M \gtrsim 10^6 M_\odot$ are confirmed, there is no direct evidence for black holes of mass in between. Scientists believing in the continuous mass distribution argue for the existence of such black holes, called intermediate mass black hole. However, many others argue that there is no such obvious expectation as the origins of stellar mass and intermediate mass black holes are completely different. Nevertheless, there are ultra-luminous X-ray sources (ULXs) observed in galaxies around, which apparently cannot be explained by the conventional idea of stellar mass black holes accreting at a sub-Eddington limit. Hence, the proposal is that the sources harbor an intermediate mass black hole, particularly when they reveal lower temperature in the underlying multicolor black hole spectra [1]. However, there is another idea behind ULXs that they are stellar mass black holes only but accreting at a super-Eddington rate: candidates for slim accretion disk [2].

Nevertheless, none of the above ideas is a conventional one. There are significant evidences that X-ray binaries are sub-Eddington accretors and there is no direct evidence yet of galactic black hole mass $M \gtrsim 100M_\odot$ (though the detection of gravitational wave argues for the black hole mass larger than that determined in X-ray astronomy). Here our story lines start. We show that ULXs in hard states can be explained by a stellar mass black hole accreting at a sub-Eddington rate with advection in the presence of large scale strong magnetic field. Hence, by the interplay between magnetic field and advection, X-ray binaries could be quite luminous in the hard state. For that we neither require an intermediate mass black hole nor super-Eddington accretion. Hence, while still the existence of intermediate

black hole, even appeared as ULX, is not ruled out, some ULXs in hard states, e.g. NGC 1365 X1/X2, M82 X42.3+59, M99 X1 etc., are suggested to be highly magnetized stellar mass black hole sources only.

We model a combined disk-outflow coupled system with the inclusion of vertical velocity and large scale magnetic stress explicitly. This is essentially a 2.5D magnetized advective accretion disk model. We show that energetics and luminosities of such a flow are in accordance with ULXs.

## 2. Magnetized disk-outflow coupled system

We consider a magnetized, viscous, advective disk-outflow/jet symbiotic system with cooling around black holes. We consider the large scale magnetic and turbulent viscous stresses both and depending on the field strength one of them may dominate over other. Here we assume a steady and axisymmetric flow and all the flow parameters: radial velocity ($v_r$), specific angular momentum ($\lambda$), outflow or vertical velocity ($v_z$), fluid pressure ($p$), mass density ($\rho$), radial ($B_r$), azimuthal ($B_\phi$), and vertical ($B_z$) components of magnetic field, are functions of both radial and vertical coordinates. Throughout we express length variables in units of $GM_{BH}/c^2$, where $G$ is the Newton's gravitational constant, $M_{BH}$ the mass of BH, and $c$ the speed of light. Accordingly, we also express other variables. Hence, the continuity and momentum balance equations are respectively

$$\nabla.(r\rho\mathbf{v}) = 0, \text{ and } (\mathbf{v}.\nabla)\mathbf{v} = \mathbf{F} - \frac{1}{\rho}\nabla\left(p + \frac{B^2}{8\pi}\right) + \frac{(\mathbf{B}.\nabla)\mathbf{B}}{4\pi\rho} + \frac{1}{\rho}\nabla.\mathbf{W}, \quad (1)$$

where $\mathbf{v}$ and $\mathbf{B}$ are velocity and magnetic field vectors respectively, $|\mathbf{F}|$ is the magnitude of the gravitational force for a BH in the pseudo-Newtonian framework[3]. The importance of generalized viscous shearing stress tensor ($\mathbf{W} = W_{ij}$) is taking care explicitly in this formalism. Various components of $W_{ij}$ are written in terms of $\alpha$-prescription[4] with appropriate modifications[5]. We also have to supplement the above equations with the equations for no magnetic monopole and induction, as respectively

$$\nabla.\mathbf{B} = 0 \text{ and } \nabla \times (\mathbf{v} \times \mathbf{B}) + \nu_m \nabla^2 \mathbf{B} = 0, \quad (2)$$

where $\nu_m$ is the magnetic diffusivity. We consider equation (2) in the very large magnetic Reynolds number ($\propto 1/\nu_m$) limit, which is the case for an accretion disk. We further have to supply the energy balance equations for ions and electrons by taking into account the detailed balance of heating, cooling and advection. The magnetized energy equations for ions and electrons read as

$$\Gamma_3' \left[ v_r \left\{ \frac{\partial p}{\partial r} - \Gamma_1 \frac{p}{\rho}\frac{\partial \rho}{\partial r} \right\} + v_z \left\{ \frac{\partial p}{\partial z} - \Gamma_1 \frac{p}{\rho}\frac{\partial \rho}{\partial z} \right\} \right] = Q^+ - Q^{ie}, \quad (3)$$

where

$$\Gamma_1 = \frac{32 - 24\beta - 3\beta^2 + \frac{2\beta(4-3\beta)}{3\beta_M}}{24 - 21\beta}, \text{ and } \Gamma_3' = \frac{24 - 21\beta}{2(4-3\beta)},$$

$$\Gamma_3' \left[ v_r \left\{ \frac{\partial p_e}{\partial r} - \Gamma_1 \frac{p_e}{\rho} \frac{\partial \rho}{\partial r} \right\} + v_z \left\{ \frac{\partial p_e}{\partial z} - \Gamma_1 \frac{p_e}{\rho} \frac{\partial \rho}{\partial z} \right\} \right] = Q^{ie} - Q^-, \tag{4}$$

where $Q^+$ represents the viscous and magnetic (Ohmic) heats generated in the flow, $Q^{ie}$ the Coulomb coupling estimating the amount of heat transferred from ions to electrons, and finally $Q^-$ the radiative cooling rate through electrons via different cooling processes including bremsstrahlung, synchrotron and inverse Comptonization of soft photons supplied from the Keplerian disk. Various cooling formalisms are adopted from past works[6–8]. In order to solve the equations semi-analytically, we make a reasonable hypothesis in the disk-outflow symbiotic region that the vertical variation of any dynamical variable (say, $A$) is much less than that with radial variation, that allows us to introduce $\partial A/\partial z \approx sA/z$, where $s$ is just the degree of scaling and is a small number.

## 3. Disk hydromagnetics and energetics

Figure 1 shows disk-outflow hydromagnetics revealing that large scale strong magnetic fields are able to transport angular momentum adequately rendering further significant $v_r$ and $v_z$ with decreasing $r$. The angular momentum profile turns out to be similar to that obtained based purely on $\alpha-$viscosity and hence $W_{ij}$ when the field is weak. However, the benefit with large scale magnetic stress is that it renders significant vertical outflow along with radial inflow. It is confirmed from Figs. 1d,e that close to the black hole magnetic field could be even $\sim 10^7$ G with an efficient magnetic shear compared to $\alpha-$viscosity induced viscous shear.

Now energetics can be estimated based on above hydromagnetism. The energy equation in conservative form under steady state condition is given by

$$\nabla.\mathcal{F} = 0 \text{ with } \mathcal{F}_i = \rho v_i \left( \frac{v^2}{2} + \frac{\Gamma_1}{\Gamma_1 - 1} \frac{p}{\rho} + \frac{B^2}{8\pi} + \Phi \right) + v_j M_{ij} - v_j W_{ij}, \tag{5}$$

where $i, j$ correspond to $r$ or $\phi$ or $z$, indicating radial or azimuthal or vertical component of the respective variables with $v^2 = v_r^2 + \lambda^2/r^2 + v_z^2$, $\Phi$ is the gravitational potential, and $M_{ij}$ is the magnetic stress tensor with standard definition, given by

$$M_{ij} = \frac{B^2}{8\pi}\delta_{ij} - \frac{B_i B_j}{4\pi}. \tag{6}$$

The outflow power extracted from the disk is computed at the disk-outflow surface region. It defines as[5]

$$P_j(r) = \int 4\pi r \left[ \rho v_z \left\{ \frac{v^2}{2} + \frac{\Gamma_1}{\Gamma_1 - 1} \frac{p}{\rho} + \Phi - \left( \frac{\lambda}{r} W_{\phi z} + v_r W_{rz} \right) \right\} + \frac{v_z}{4\pi} \left( B_r^2 + B_\phi^2 - \frac{v_r}{v_z} B_r B_z - \frac{\lambda}{r v_z} B_\phi B_z \right) \right]_h dr. \tag{7}$$

This accretion induced outflow power contains contributions from mechanical and enthalphy powers, and those of viscous and Poynting parts. Our model is restricted

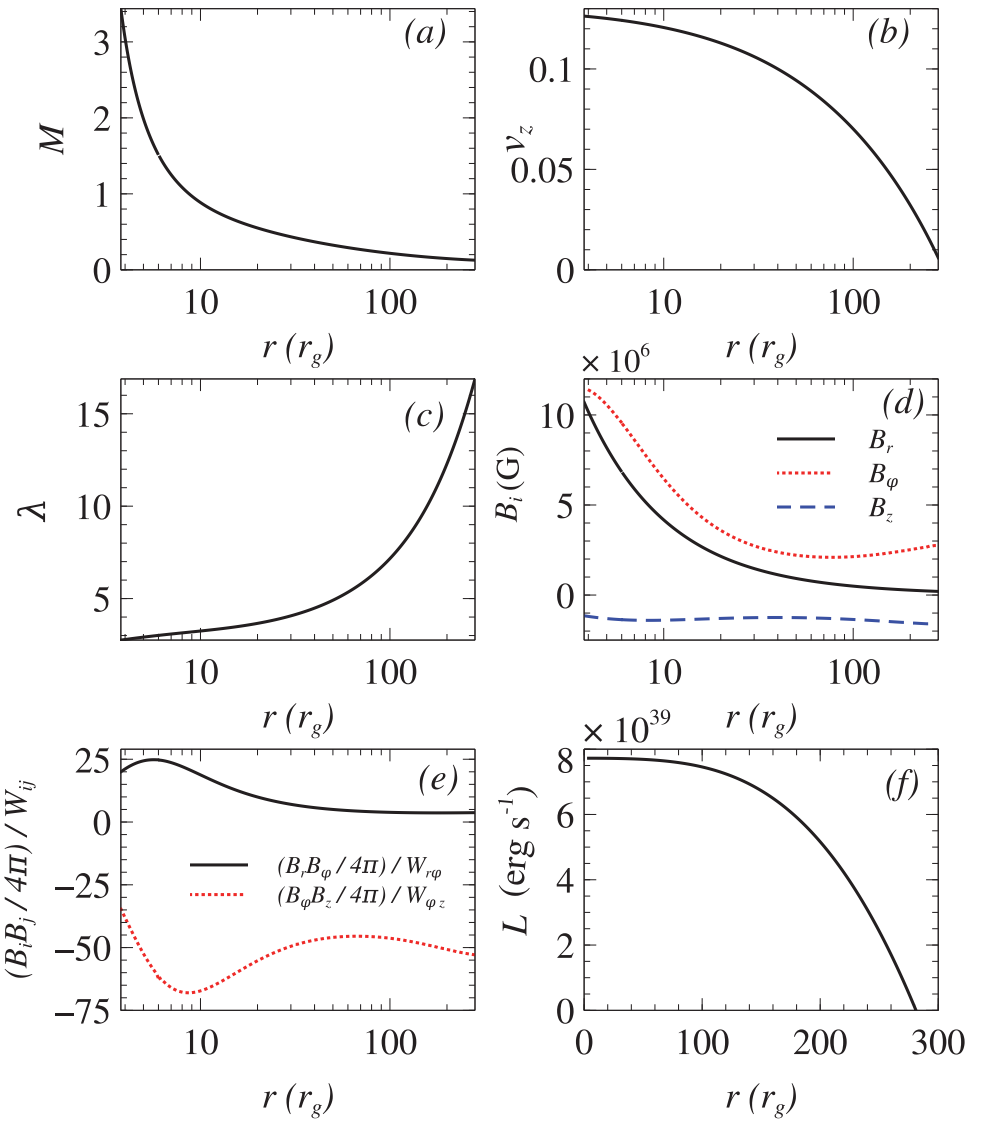



Fig. 1. Variation of (a) Mach number, (b) outflow speed, (c) specific angular momentum, (d) magnetic field components, (e) ratio of magnetic to viscous shearing stresses, and (f) luminosity, as functions of distance from the black hole. Other parameters are $M_{BH} = 20M_{\odot}$, $\dot{m} = 0.05$ Eddington rate.

vertically up to the disk-outflow coupled region, above which outflow may decouple and accelerate. Hence this computed power is basically the initial power of any astrophysical jets at the launching region. Also the disk luminosity can be computed from the cooling mechanisms and can be defined as

$$L = \int \left( \int_0^h Q^- 4\pi r \; dz \right) \; dr. \tag{8}$$

The variation of disk luminosity, whose magnitude is the most important observable in the present context, is shown in Fig. 1f. At an arbitrary $r$, the luminosity is obtained by integrating from the outer disk radius $r_{out}$ to that corresponding $r$. For the case of a stellar-mass black hole of mass $M_{BH} = 20\ M_\odot$ with total mass accretion rate $\dot{m} = 0.05$ Eddington rate, the maximum attainable luminosity, based on the integration over whole disk, is $L \sim 8 \times 10^{39}$ erg s$^{-1}$. This value is quite adequate to explain observed luminosities of ULXs in hard states. Table 1 enlists some ULXs with their respective power-law indices, indicating their harder nature. It is very interesting that the luminosity of the sources $L \sim 10^{40}$ erg s$^{-1}$, which can be explained by a stellar mass black hole accreting at a sub-Eddington rate in the presence of strong magnetic fields, as described in Fig. 1f.

Table 1: Some ULX sources in a hard power-law dominated state.

| Source | $\Gamma$ | $L_{0.3-10\ \mathrm{keV}}$ ($10^{40}$ erg s$^{-1}$) |
|---|---|---|
| M99 X1 [9] | $1.7^{+0.1}_{-0.1}$ | 1.9 |
| Antennae X-11 [10] | $1.76^{+0.05}_{-0.05}$ | 2.11 |
| Holmberg IX X-1 [11] | $1.9^{+0.1}_{-0.02}$ | 1.0 |
| NGC 1365 X1 [12] | $1.74^{+0.12}_{-0.11}$ | 2.8 |
| NGC 1365 X2 [12] | $1.23^{+0.25}_{-0.19}$ | 3.7 |
| M82 X42.3+59 [13] | $1.44^{+0.09}_{-0.09}$ | 1.13 |

## 4. Summary

ULXs and the question of plausible existence of intermediate mass black hole in the universe are both puzzling us for quite sometime. Some authors argue ULXs to be the sources of an intermediate mass black hole. Some others argue ULXs to be super-Eddington acrretors by a stellar mass black hole. The later group further argues that there is no need to expect a continuous mass distribution of black holes from stellar mass to supermassive scales. We suggest quite differently and uniquely. We show that at least some of ULXs are nothing but the highly magnetized accreting sources of stellar mass black holes accreting at a sub-Eddington rate only. The required field magnitude is of the order of $10^7$ G to explain ULXs in hard states, which is well below the underlying Eddington value. Therefore, at least some of

ULXs could just be stellar mass black holes. While this suggestion leaves the question for the existence of intermediate mass black hole wide open, it argues for the power of magnetically dominated/arrested accretion flows to explain enigmatic astrophysical sources.

**Acknowledgment**

The author thanks Tushar Mondal of IISc for discussion and drawing the figure.

**References**

1. J.M. Miller, A.C. Fabian, M.C. Miller, *ApJ* **614**, L117 (2004).
2. M.A. Abramowicz, B. Czerny, J.P. Lasota, and E. Szuszkiewicz, *ApJ* **332**, 646 (1988).
3. B. Mukhopadhyay, *ApJ* **581**, 427 (2002).
4. N.I. Shakura and R.A. Sunyaev *A&A* **24**, 337 (1973).
5. T. Mondal and B. Mukhopadhyay *MNRAS* **482**, L24 (2019).
6. R. Narayan and I. Yi *ApJ* **452**, 710 (1995).
7. S. Mandal and S.K. Chakrabarti *A&A* **434**, 839 (2005).
8. S.R. Rajesh and B. Mukhopadhyay *MNRAS* **402**, 961 (2010).
9. R. Soria and D.S. Wong *MNRAS* **372**, 1531 (2006).
10. H. Feng H. and P. Kaaret *ApJ* **696**, 1712 (2009).
11. P. Kaaret and H. Feng *ApJ* **702**, 1679 (2009).
12. R. Soria, G. Risaliti, M. Elvis, G. Fabbiano, S. Bianchi, and Z. Kuncic *ApJ* **695**, 1614 (2009).
13. H. Feng H., F. Rao, and P. Kaaret 2010 *ApJ* **710**, L137 (2010).

# Eccentric equatorial trajectories around a Kerr black hole as a QPO model for M82X-1

Prerna Rana[1,†] and A. Mangalam[1,‡]

[1] *Indian Institute of Astrophysics, Sarjapur Road, 2nd Block Koramangala, Bangalore, 560034, India*

*E-mail: prernarana@iiap.res.in*[†] *, mangalam@iiap.res.in*[‡]

We study the bound orbit conditions for equatorial and eccentric orbits around a Kerr black hole both in the parameter space ($E$, $L$, $a$) representing the energy, angular momentum of the test particle, and spin of the black hole, and also ($e$, $\mu$, $a$) space representing the eccentricity, inverse-latus rectum of the orbit, and spin. We apply these conditions and implement the relativistic precession (RP) model to M82X-1, which is an Intermediate-mass black hole (IMBH) system, where two high-frequency Quasi-Periodic Oscillations (HFQPOs) and a low-frequency QPO were simultaneously observed. Assuming that the QPO frequencies can also be generated by equatorial and eccentric trajectories, we calculate the probability distributions to infer $e$, $a$, and periastron distance, $r_p$, of the orbit giving rise to simultaneous QPOs. We find that an eccentric orbit solution is possible in the region between innermost stable circular orbit (ISCO) and the marginally bound circular orbit (MBCO) for $e = 0.2768^{+0.0657}_{-0.0451}$, $a = 0.2897 \pm 0.0087$, and $r_p = 4.6164^{+0.0694}_{-0.1259}$.

*Keywords*: Classical black holes; Relativity and Gravitation; Infall, accretion and accretion disks; Kerr black hole; QPOs.

## 1. Introduction

Black holes have been discovered as part of various systems in galaxies; as one of the components of the transient binary systems known as black hole X-ray binaries (BHXRB) with black hole mass in the range $5 - 30 M_\odot$, as intermediate mass black holes (IMBH) with mass range 300 to 600 $M_\odot$, and as galactic nuclei of mass $10^6 - 10^{10} M_\odot$ (Ref. 1). X-ray emission from such systems gives evidence of the presence of a compact object as X-rays originate very close to the black hole, and hence are important probes of the strong gravity regime. One of the important features discovered in the X-ray emission from these systems is quasi-periodic oscillations (QPOs), which are broad peaks in the Fourier power density spectrum. QPOs have been discovered with their frequencies ranging from mHz-kHz (Ref. 2), in BHXRB, mHz to a few Hz in IMBH (Ref. 3), and expected with the time scales of hours to days for AGN. The most interesting features among these are the high-frequency QPOs (HFQPOs) in the range 100-500Hz (Refs. 5, 6), detected in various BHXRB which may be used to calculate spin $a$ and mass $M_\bullet$ of the black hole when detected simultaneously. Some of these objects have also shown two simultaneous HFQPOs with their centroid frequencies having ratios ∼3:2 or 5:3 (Ref. 2), indicating a resonance. There are also some cases, for example, GROJ1655-40 (Ref. 7), and M82 X-1 (IMBH) (Ref. 3), which are known to show the detection of three simultaneous QPOs, two HFQPOs and one low-frequency QPO (LFQPO). There are various existing models that attempt to explain the origin of QPOs but perhaps the most

widely accepted is the relativistic precession (RP) model (Ref. 8), which interprets LFQPO as the nodal precession frequency given by ($\nu_\phi$-$\nu_\theta$), and the two HFQPOs as associated with the azimuthal frequency, $\nu_\phi$, and the periastron precession frequency ($\nu_\phi$-$\nu_r$) of a particle orbiting around a black hole. RP model has been previously applied to the cases of BHXRB (Ref. 7), assuming that the frequencies correspond to circular orbits of an equatorial system.

In this article, we first present the necessary bound orbit conditions both in the ($E$, $L$, $a$) and in the ($e$, $\mu$, $a$) space of the corresponding orbit for the equatorial bound orbits. Using these conditions, we classify various bound orbits in both spaces. Then, by applying the RP model to the case of an IMBH M82X-1, we find that an eccentric equatorial orbit can also originate three simultaneous QPOs. We also discuss the possible region around the black hole giving rise to these QPOs.

## 2. Bound orbit conditions for equatorial orbits

In this section, we write the bound orbit conditions in ($E$, $L$) and ($e$, $\mu$) space for a particle orbiting a Kerr black hole of a given spin, $a = J/M_\bullet$. Throughout this article, we have scaled $E$, $L$, and $a$ parameters by mass of the black hole $M_\bullet$, and have used geometrical units ($G = c = 1$).

### 2.1. *Dynamical parameter space* $\{E, L, a\}$

The equation defining radial motion of a particle (unit mass, $m_0 = 1$) in the equatorial plane of a Kerr black hole having spin $a$ is given by (Ref. 9)

$$\frac{\left(E^2-1\right)}{2} = \frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2 - \frac{1}{r} + \frac{L^2 - a^2\left(E^2-1\right)}{2r^2} - \frac{(L-aE)^2}{r^3}, \tag{1}$$

where $r$ is the radial distance from black hole and $\tau$ is the proper time. The radial kinetic energy vanishes at the turning points of the orbit, $\mathrm{d}r/\mathrm{d}\tau = 0$, which gives

$$\left(1-E^2\right)r^3 - 2r^2 + \left(L^2 - a^2\left(E^2-1\right)\right) r - 2(L-aE)^2 = 0. \tag{2}$$

We apply the Cardano's method (Ref. 10) to obtain real roots of Eq. (2) in terms of dynamical parameters $\{E, L, a\}$, which corresponds to the case when a bound orbit exists between the first two turning points. The three real roots of Eq. (2) can be expressed as (Ref. 14)

$$r_1 = r_a = \frac{2\left[1 + \left[4 - 3\left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right)\right]^{1/2}\cos\left(\frac{\varphi}{3}\right)\right]}{3\left(1-E^2\right)}, \tag{3a}$$

$$r_2 = r_p = \frac{2\left[1 + \left[4 - 3\left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right)\right]^{1/2}\cos\left(\frac{\varphi-2\pi}{3}\right)\right]}{3\left(1-E^2\right)}, \tag{3b}$$

$$r_3 = \frac{2\left[1 + \left[4 - 3\left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right)\right]^{1/2}\cos\left(\frac{\varphi+2\pi}{3}\right)\right]}{3\left(1-E^2\right)}, \tag{3c}$$

where $\varphi$ is defined by

$$\cos\varphi = \frac{\left[8 - 9\left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right) + 27\left(1-E^2\right)^2\left(L-aE\right)^2\right]}{\left[4 - 3\left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right)\right]^{3/2}}; \quad (3d)$$

where $r_1 > r_2 > r_3$, $r_1$ and $r_2$ are the apastron and periastron points of the eccentric orbit respectively. It follows that

$$-1 < \cos\varphi < +1 \quad \text{holds for eccentric orbits.} \quad (3e)$$

As per the Cardano's method, the condition for three real roots is described by the discriminant of the cubic equation, which we obtain for Eq. (2), given by (Ref. 14)

$$\begin{aligned}\Delta = \; & 27\left(1-E^2\right)^2 x^4 - \left(L^2 - a^2\left(E^2-1\right)\right)^2 - 18x^2\left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right) \\ & +16x^2 + \left(1-E^2\right)\left(L^2 - a^2\left(E^2-1\right)\right)^3, \end{aligned} \quad (4)$$

where $x = L - aE$. Hence, the condition on $\{E, L, a\}$ to get three real and distinct roots of $r$, which is possible only when the orbit is an eccentric bound orbit, is given by

$$\Delta < 0 \qquad \text{and} \qquad 0 < E < 1. \quad (5)$$

The case of two equal roots and one distinct real root is possible when $\Delta = 0$, which corresponds to the stable circular, and the separatrix orbit or the unstable circular orbit, which are classified using $\cos\varphi$ by the following conditions:

$$\text{Stable circular orbit when } \cos\varphi = -1 \;\Rightarrow\; r_1 = r_2 \text{ and } 0 < E < 1, \quad (6a)$$

$$\text{Separatrix orbit when } \cos\varphi = +1 \;\Rightarrow\; r_2 = r_3 \text{ and } 0 < E < 1, \quad (6b)$$

$$\text{Unstable circular orbit when } \cos\varphi = +1 \;\Rightarrow\; r_2 = r_3 \text{ and } E > 1. \quad (6c)$$

In this way, we are able to classify the various bound orbits in the $\{E, L, a\}$ space. Fig. 1(a) shows this classification as bound orbit region, we call it the $\Delta$ region, in the ($E$, $L$) plane which is bounded by the curves representing the stable circular orbits Eq. (6a), separatrix orbits Eq. (6b), and $E = 1$. The details of results presented in this section will be provided in paper in preparation (Ref. 14).

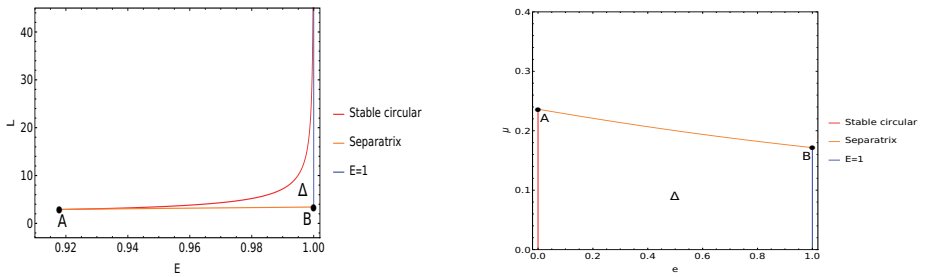
### 2.2. *Conic parameter space* $\{e, \mu, a\}$

Next, we extend and translate the bound orbit conditions to the conic parameter space which is useful for the geometric study of the bound trajectories. The eccentricity and inverse-latus rectum are defined as

$$e = \frac{r_a - r_p}{r_a + r_p}, \quad \mu = \frac{r_a + r_p}{2 r_a r_p}. \quad (7)$$

Using the transformation relations between ($E$, $L$) and ($e$, $\mu$) (Refs. 11, 12), we can express $\Delta$, Eq. (4), in terms of $\{e, \mu, a\}$ as (Ref. 14)

$$\Delta = \frac{-e^2}{\mu^2}\left[1 - x^2\mu^2\left(1+e\right)\left(3-e\right)\right]^2\left[1 - x^2\mu^2\left(1-e\right)\left(3+e\right)\right]^2 = \frac{-e^2}{\mu^2}\Delta_1^2\Delta_2^2. \quad (8)$$




(a) (b)

Fig. 1. The bound orbit condition, $\Delta \leq 0$, is shown as the $\Delta$ region (a) bounded by stable circular Eq. (6a), separatrix orbits Eq. (6b), and $E = 1$ in the ($E$, $L$) plane, and (b) bounded by stable circular Eq. (9a), separatrix orbits Eq. (9b), and $e = 1$ in the ($e$, $\mu$) plane for $a = 0.5$, where points A and B represents the innermost stable circular orbit (ISCO) and marginally bound circular orbit (MBCO) respectively.

It is seen that $\Delta \leq 0$ is valid for all the real values of $x$, and $\Delta = 0$ if any one of the following conditions is satisfied:

$$e = 0, \tag{9a}$$

$$\Delta_1 = 0 \Rightarrow \mu^2 x^2 (3 - e)(1 + e) = 1, \tag{9b}$$

$$\Delta_2 = 0 \Rightarrow \mu^2 x^2 (3 + e)(1 - e) = 1, \tag{9c}$$

where Eq. (9a) corresponds to the stable circular orbits, Eq. (9b) corresponds to the separatrix orbits and they together with $e = 1$ define the boundaries of the $\Delta$ region in the ($e$, $\mu$) plane; see Fig. 1(b). Hence, the condition in $\{e, \mu, a\}$ space for an eccentric bound orbit is given by $r_3 < r_p$, which corresponds to (Ref. 14)

$$\mu^2 x^2 (3 - e)(1 + e) \leq 1, \tag{10}$$

and is the operative condition for eccentric bound orbits representing the ordering of roots, $r_1 > r_2 > r_3$. The expression of $\cos\varphi$, Eq. (3d), in the $\{e, \mu, a\}$ space is given by (Ref. 14)

$$\cos\varphi = \frac{\left[1 - 3x^2\mu^2\left(1 - e^2\right)\right] \cdot \left[-1 + 3e + 3x^2\mu^2\left(1 - e - e^2 + e^3\right)\right] \cdot \left[1 + 3e - 3x^2\mu^2\left(1 + e - e^2 - e^3\right)\right]}{\left\{4 - 3\left(1 - e^2\right)\left[1 - x^2\mu^2\left(1 - e^2\right)\right]\left[1 + \mu^2 x^2\left(3 + e^2\right)\right]\right\}^{3/2}}. \tag{11}$$

## 3. Parameter estimation from RP model for QPOs

Now, we take a more generalized approach to the RP model by considering $e \neq 0$. We consider that three simultaneous QPOs can also be originated by a self-emitting blob orbiting in an equatorial eccentric trajectory around a Kerr black hole. Then, we calculate the parameters ($e$, $r_p$) of the orbit and spin $a$ of the black hole using the fundamental frequency formulae for equatorial orbits derived in Refs. (11, 12), and presented in Table 1. We search for the parameters in the allowed $\Delta$ region

Table 1. Fundamental frequencies of equatorial eccentrc orbits, Refs. 11, 12, 13.

| | |
|---|---|
| $\nu_\phi\,(e,\mu,a)$ | $\dfrac{c^3\cdot\{a_1\text{EllipticPi}\left[-{p_2}^2,\pi/2,m^2\right]+b_1\text{EllipticPi}\left[-{p_3}^2,\pi/2,m^2\right]\}}{2\pi GM\left\{\begin{array}{l}a_2\left[\dfrac{{p_1}^2\text{EllipticE}\left[\pi/2,m^2\right]}{2\left(1+{p_1}^2\right)\left(m^2+{p_1}^2\right)}-\dfrac{\text{EllipticF}\left[\pi/2,m^2\right]}{2\left(1+{p_1}^2\right)}\right]+c_2\text{EllipticPi}\left[-{p_2}^2,\pi/2,m^2\right]\\ \quad+\text{EllipticPi}\left[-{p_1}^2,\pi/2,m^2\right]\left\{a_2\dfrac{\left[{p_1}^4+2{p_1}^2\left(1+m^2\right)+3m^2\right]}{2\left(1+{p_1}^2\right)\left(m^2+{p_1}^2\right)}+b_2\right\}+d_2\text{EllipticPi}\left[-{p_3}^2,\pi/2,m^2\right]\end{array}\right\}}$, |
| $\nu_r\,(e,\mu,a)$ | $\dfrac{c^3}{2GM\left\{\begin{array}{l}a_2\left[\dfrac{{p_1}^2\text{EllipticE}\left[\pi/2,m^2\right]}{2\left(1+{p_1}^2\right)\left(m^2+{p_1}^2\right)}-\dfrac{\text{EllipticF}\left[\pi/2,m^2\right]}{2\left(1+{p_1}^2\right)}\right]+c_2\text{EllipticPi}\left[-{p_2}^2,\pi/2,m^2\right]\\ \quad+\text{EllipticPi}\left[-{p_1}^2,\pi/2,m^2\right]\left\{a_2\dfrac{\left[{p_1}^4+2{p_1}^2\left(1+m^2\right)+3m^2\right]}{2\left(1+{p_1}^2\right)\left(m^2+{p_1}^2\right)}+b_2\right\}+d_2\text{EllipticPi}\left[-{p_3}^2,\pi/2,m^2\right]\end{array}\right\}}$ |
| $\nu_\theta\,(e,\mu,a)$ | $\dfrac{2\nu_r a\sqrt{1-E^2}z_+\mu^{1/2}\text{EllipticF}\left[\dfrac{\pi}{2},m^2\right]}{\pi\left[1-\mu^2x^2\left(3-e^2-2e\right)\right]^{1/2}}$ |
| Constants | $m^2=\dfrac{4\mu^2ex^2}{\left[1-\mu^2x^2\left(3-e^2-2e\right)\right]}$, ${p_1}^2=\dfrac{2e}{1-e}$, ${p_2}^2=\dfrac{2ea^2\mu}{\left(a^2\mu-a^2\mu e-r_+\right)}$, ${p_3}^2=\dfrac{2ea^2\mu}{\left(a^2\mu-a^2\mu e-r_-\right)}$, ${z_+}^2=\dfrac{\left[a^2\left(1-E^2\right)+L^2\right]}{a^2\left(1-E^2\right)}$,<br>$a_1=\dfrac{\mu^{1/2}\left[La^2-2xr_+\right]}{\sqrt{1-a^2}\left(a^2\mu-a^2\mu e-r_+\right)\sqrt{1-\mu^2x^2\left(3-e^2-2e\right)}}$, $b_1=\dfrac{\mu^{1/2}\left[-La^2+2xr_-\right]}{\sqrt{1-a^2}\left(a^2\mu-a^2\mu e-r_-\right)\sqrt{1-\mu^2x^2\left(3-e^2-2e\right)}}$,<br>$a_2=\dfrac{2E}{\mu^{3/2}\left(1-e\right)^2\sqrt{1-\mu^2x^2\left(3-e^2-2e\right)}}$, $b_2=\dfrac{4E}{\mu^{1/2}\left(1-e\right)\sqrt{1-\mu^2x^2\left(3-e^2-2e\right)}}$,<br>$c_2=\dfrac{2a^2\mu^{1/2}\left(-La+2Er_-\right)}{r_-\sqrt{\left[1-\mu^2x^2\left(3-e^2-2e\right)\right]}\sqrt{1-a^2}\left(a^2\mu-a^2\mu e-r_+\right)}$, $d_2=\dfrac{2a\mu^{1/2}\left(-2Lr_-\sqrt{1-a^2}-2Er_-a+La^2\right)}{r_-\sqrt{\left[1-\mu^2x^2\left(3-e^2-2e\right)\right]}\sqrt{1-a^2}\left(a^2\mu-a^2\mu e-r_-\right)}$. |

(shown in Fig. 1) and defined by the bound orbit condition, Eq. (10). We apply this method to the case of an IMBH M82X-1, which has shown three simultaneous QPOs at $\nu_1 = 5.07 \pm 0.06$Hz, $\nu_2 = 3.32 \pm 0.06$Hz, and $\nu_3$ =204.8±6.3mHz (Ref. 3, 4) to find $\{e, r_p, a\}$. We fix the mass of black hole to be $M_\bullet = 428 M_\odot$ given in Ref. 3, and simultaneously solve the equations $\nu_\phi = \nu_1$, $\nu_\phi - \nu_r = \nu_2$, and $\nu_\phi - \nu_\theta = \nu_3$ to find the exact solutions for the RP model given by $\{e_0, r_{p_0}, a_0\}$. We assume the QPO frequencies as Gaussian distributed with their mean values at $\nu_1$, $\nu_2$ and $\nu_3$. We estimate the $1\sigma$ errors in the solution $\{e_0, r_{p_0}, a_0\}$ using the Jacobian of the transformation to $\{\nu_1, \nu_2, \nu_3\}$ space whose formulae are given in Table 1, and find the corresponding normalized joint probability density distribution in the ($e$, $r_p$, $a$) space. As the probability density $P\,(e, r_p, a)$ is three dimensional, we collapse the profile of the probability density in two dimensions and fit a Gaussian function to find the mean values, $\{e_0, r_{p0}, a_0\}$, and variances, $\{\sigma_e, \sigma_{r_p}, \sigma_a\}$ in other dimension.

Using this procedure, we found the exact solution at $e = 0.2768^{+0.0657}_{-0.0451}$, $a = 0.2897 \pm 0.0087$, and $r_p = 4.6164^{+0.0694}_{-0.1259}$ for M82X-1. We see that the probability density naturally peaks at non-zero eccentricity suggesting that the most probable solution might not be restricted to circular orbits as was previously assumed in Ref. 7. Fig. 2 shows this solution in the ($r_p$, $a$) plane suggesting an eccentric orbit having its periastron point in the region between ISCO and MBCO is the most probable orbit for the generation of three simultaneous QPOs. We have implemented a similar approach to various BHXRB to include non-equatorial and eccentric trajectories in our paper in preparation (Ref. 13).

## 4. Summary and discussion

We discussed the bound orbit conditions for the eccentric equatorial orbits in the $\{E, L, a\}$, and $\{e, \mu, a\}$ spaces. We applied it to the specific case of M82X-1, where three simultaneous QPOs were discovered (Ref. 3, 4), to find the parameters $e = 0.2768^{+0.0657}_{-0.0451}$, $a = 0.2897 \pm 0.0087$, and $r_p = 4.6164^{+0.0694}_{-0.1259}$ for the orbit generating these QPOs using the RP model. We find that the eccentric orbit solutions are possible in the region between ISCO and MBCO, as shown in Fig. 2. Hence, by assuming that the accretion disk ends near ISCO, we conclude that the blobs that originate near ISCO and follow equatorial and eccentric trajectories in this region that produce HFQPOs. In a paper in preparation (Ref. 13), we also discuss the non-equatorial eccentric trajectories as possible solutions for QPOs assuming the RP model.

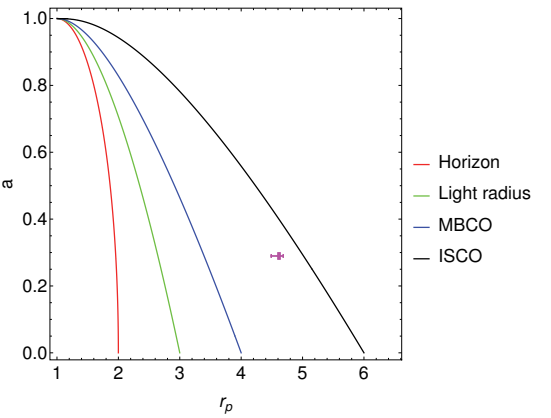



Fig. 2. An equatorial eccentric orbit solution (the point in magenta color) shown in the $(r_p,\ a)$ plane for the simultaneous QPOs discovered in M82X-1. The solution exists in the region between ISCO and MBCO.

We acknowledge the support from the SERB project CRG 2018/003415.

## References

1. R. Narayan and J. E. McClintock, Observational Evidence for Black Holes, *arXiv e-prints* , p. arXiv:1312.6698 (Dec 2013).
2. T. M. Belloni and L. Stella, Fast Variability from Black-Hole Binaries, *Space Science Reviews* **183**, 43 (September 2014).
3. D. R. Pasham, T. E. Strohmayer and R. F. Mushotzky, A 400-solar-mass black hole in the galaxy M82, *Nature* **513**, 74 (September 2014).
4. D. R. Pasham and T. E. Strohmayer, On the Nature of the mHz X-Ray Quasi-periodic Oscillations from Ultraluminous X-Ray Source M82 X-1: Search for Timing-Spectral Correlations, *ApJ* **771**, 2 101, *doi: 10.1088/0004-637X/771/2/101* (July 2013), *arXiv:1308.1677*.
5. T. E. Strohmayer, Discovery of a 450 HZ Quasi-periodic Oscillation from the Microquasar GRO J1655-40 with the Rossi X-Ray Timing Explorer, *ApJ* **552**, L49 (May 2001).
6. R. A. Remillard, G. J. Sobczak, M. P. Muno and J. E. McClintock, Characterizing the Quasi-periodic Oscillation Behavior of the X-Ray Nova XTE J1550-564, *ApJ* **564**, 962 (January 2002).

7. S. E. Motta, T. M. Belloni, L. Stella, T. Muñoz-Darias and R. Fender, Precise mass and spin measurements for a stellar-mass black hole through X-ray timing: the case of GRO J1655-40, *MNRAS* **437**, 2554 (January 2014).
8. L. Stella, M. Vietri and S. M. Morsink, Correlations in the Quasi-periodic Oscillation Frequencies of Low-Mass X-Ray Binaries and the Relativistic Precession Model, *ApJ* **524**, L63 (October 1999).
9. B. Carter, Global Structure of the Kerr Family of Gravitational Fields, *Physical Review* **174**, 1559 (October 1968).
10. G. Cardano, *English Translation: The Great Art or the Rules of Algebra, T R Witmer* (MIT Press, Cambridge, MA, 1968).
11. P. Rana and A. Mangalam, Astrophysically relevant bound trajectories around a Kerr black hole, *Classical and Quantum Gravity* **36**, p. 045009 (Feb 2019).
12. P. Rana and A. Mangalam, Astrophysically relevant bound trajectories around a Kerr black hole, *arXiv e-prints* , p. arXiv:1901.02730 (January 2019).
13. P. Rana and A. Mangalam, A geometric origin for QPO frequencies in BHXRB, *submitted to ApJ* (2020).
14. P. Rana and A. Mangalam, (*in preparation,* 2020).

# Polarization of emission from black hole accretion disks: including returning radiation

Roberto Taverna,*; Michal Bursa, Michal Dovčiak, Vladimir Karas, Frédéric Marin, Giorgio Matt, Romana Mikušincová and Wenda Zhang

*Department of Mathematics and Physics, University of Roma Tre, Rome, Italy 00146
*E-mail: taverna@fis.uniroma3.it*
*www.matfis.uniroma3.it*

Accretion disks around stellar-mass black holes (BHs) are observed to emit radiation peaking in the soft X-rays (at about 1 keV) when the source is in the thermal state. This emission is expected to be polarized, because of electron scattering in the disk. However, photons will experience a rotation of the polarization plane due to General Relativity effects. X-ray polarization measurements can then be crucial in assessing strong gravity effects around accreting stellar-mass BHs, providing independent constraints on their spin. In this work, we present the simulated polarization spectra (obtained according to the model by Dovčiak et al.[1]) for the X-ray, thermal radiation emitted by the accretion disk, as expected to be measured by IXPE, an X-ray polarimeter which will be launched in 2021. We also illustrate the code, currently under development, incorporate the contribution of returning radiation and account for absorption effects by the disk material.

*Keywords*: polarization – relativity – instrumentation: polarimeters – X-rays: binaries.

## 1. Introduction

Stellar-mass black holes (BHs) are believed to be born mainly in the gravitational core-collapse of very massive stars (with $M \gtrsim 25$–$30\,M_\odot$), while supermassive black holes ($M \approx 10^6$–$10^9 M_\odot$) are hosted in the center of most galaxies (including our own). Despite the fact that even light cannot escape from their event horizon, BHs have been observed so far as X-ray sources, through the electromagnetic radiation emitted from accretion disks in their surroundings, as well as, in the case of supermassive BHs, by studying the orbits of nearby stars influenced by the presence of their gravitational field. Very recently, BHs have been in the spotlight thanks to the results achieved by the LIGO/VIRGO collaboration, which detected for the first time the gravitational waves emitted in BH-BH merging events.[2] Nevertheless, electromagnetic observations still play a key role in understanding the physics of BHs. Moreover, the recent developments in X-ray polarimetry techniques gave new impetus to this field of research.[3]

In this work we revisit the issue of modeling the spectral and polarization properties of radiation emitted from the accretion disk around a stellar mass BH, where general relativistic effects play a crucial role. Theoretical models to investigate spectra and polarization from these sources have been developed by many authors in the past (see e.g. Ref. 4 for a more comprehensive list of references). We revisit in particular the work by Dovčiak et al.,[1] which addressed the contribution of photons

that reach the observer without interacting anymore with the disk since their emission (the so-called "direct radiation"). As already discussed in Ref. 5 the "returning radiation" (i.e. the emitted photons which are forced by the strong gravity effects to return to the disk, where they may be scattered to eventually arrive at infinity) may provide an important contribution, and therefore we decided to include it in the KYN code. However, contrary to Schnittman & Krolik, we use an observer-to-emitter approach, and we also aim to explore non-trivial prescriptions for the disk albedo. The predictions we will obtain for the polarization observables can be promptly tested by the new-generation X-ray polarimeter IXPE,[6] that will be launched in 2021.

In section 2 we briefly set out the theoretical model and present our main assumptions. An overview of the numerical implementation is described in section 3, while some preliminary plots are discussed in section 4. Finally, conclusions are presented in section 5.

## 2. Basic theoretical model

Accretion disks around stellar-mass BHs are observed to emit radiation peaking in the soft X-rays (at about 1 keV) when the source is in its thermal state. Photons propagating in the space-time around a BH (which is described by the Kerr metric, for different values of the dimensionless angular momentum per unit mass $a$) will experience general relativistic effects. More in detail, photons emitted close to a BH are redshifted due to the strong gravitational field, and their trajectory follows null geodesics, deviating from straight lines. As a consequence, the paths of photons emitted from regions of the disk closer to the central BH can be bent by its gravitational field. These photons may still reach the observer at infinity (direct radiation), but part of them can return towards the disk surface and scatter before to eventually arrive at infinity (returning radiation), depending on the emission point on the disk and on the emission direction.

Disk radiation is also expected to be linearly polarized, because of Thomson scatterings that occur onto the electrons which are present in the atmospheric layer assumed to cover the disk (see e.g. Ref. 7). For symmetry reasons, the polarization vector of photons must be either parallel or perpendicular to the disk symmetry axis projected in the plane of the sky. In particular, if no absorption is assumed in the atmosphere, polarization will be parallel to this axis for lower values of the atmospheric optical depth ($\tau \lesssim 1$) and perpendicular for higher ones. As for the photon energy and trajectory, strong gravity can also influence the polarization state of radiation.[8–10] In fact, the photon polarization plane should be parallely transported along the curved trajectory, and this results in a rotation of the polarization vector with respect to the direction assumed at the emission. Moreover, if the photon propagates around a rotating BH, it can be shown that the polarization vector undergoes a further rotation (also called the Gravitational Faraday rotation, see Ref. 11), the amount of which depends on $a$.

For these reasons, the expected polarization fraction $P_{\rm o}$ and polarization angle $\chi_{\rm o}$ will be strongly influenced by the general relativistic effects associated to the central BH gravitational field. In particular, the change $\Psi$ in polarization angle for a photon propagating along a null geodesic is given by

$$\tan\Psi = \frac{Y}{X}\,, \tag{1}$$

where, in the case of a Petrov type D space-time, the quantities $Y$ and $X$ can be expressed in terms of the components $\kappa_1$ and $\kappa_2$ of the Walker-Penrose constant,[12] which is a constant of motion along a null geodesic, as

$$\begin{aligned} X &= -(\alpha - a\sin\theta_{\rm o})\kappa_1 - \beta\kappa_2 \\ Y &= (\alpha - a\sin\theta_{\rm o})\kappa_2 - \beta\kappa_1\,. \end{aligned} \tag{2}$$

In equation (2), $\alpha$ and $\beta$ are the impact parameters which identify the $x$- and $y$-axes, respectively, of the observer's sky reference frame in the plane perpendicular to the line-of-sight, and $\theta_{\rm o}$ is the observer inclination.[1]

## 3. Numerical implementation

Calculations are performed for a standard disk[13,14] focussing, for the sake of simplicity, on the case in which the emitted radiation follows locally a simple blackbody distribution. Photons are assumed to be scattered by the electrons of a geometrically-thin atmospheric layer above the disk surface, which polarizes the emitted radiation. In particular, the intrinsic polarization properties in the diffusion limit ($\tau \gg 1$) are extracted from Chandrasekhar's analytical expressions[7] for a geometrically thin and optically thick atmosphere with infinite optical depth. It can be shown that this is a good approximation as long as $\tau \gtrsim 5$.[1] On the other hand, for lower values of $\tau$ the intrinsic polarization pattern is numerically evaluated using the Monte Carlo code STOKES, firstly developed by Goosmann & Gaskell[15] (see also Refs. 16 and references therein for subsequent updates).

### 3.1. *Direct radiation*

KYN[1] is a ray-tracing code based on an observer-to-emitter approach, contrary to the Monte Carlo code developed in Ref. 5. The observer-to-emitter method has the advantage of a much less computational load, even if, on the other hand, the inverse approach may prove to be more useful when electron scattering in the disk is considered.[4]

In its original version, KYN takes into account only the contribution of direct radiation. The code starts with calculating the local specific Stokes parameters $\boldsymbol{s}_{\rm l} = [i_{\rm l}, q_{\rm l}, u_{\rm l}]$ according to the formulae by Ref. 7, for photons emitted from different points on the disk surface, each characterized by the distance $r$ from the central BH and the azimuth $\phi$. Then, once selected the polar angles $\theta_{\rm e}$ and $\varphi_{\rm e}$ that the photon

emission direction makes with the disk normal, the code computes the integrated Stokes parameters at the observer,[1]

$$\begin{aligned}\Delta i_{\rm o} &= \int {\rm d}S \int {\rm d}E_{\rm l}\, i_{\rm l} G \\ \Delta q_{\rm o} &= \int {\rm d}S \int {\rm d}E_{\rm l}\, [q_{\rm l}\cos(2\Psi) - u_{\rm l}\sin(2\Psi)] G \\ \Delta u_{\rm o} &= \int {\rm d}S \int {\rm d}E_{\rm l}\, [u_{\rm l}\cos(2\Psi) + q_{\rm l}\sin(2\Psi)] G\,,\end{aligned} \tag{3}$$

where $\Psi$ and $G$ are the change in polarization angle (see equation 1) and the transfer function,[1,17] respectively, which account for the general relativistic effects that photons experience moving along the geodesic which connects the observer to the emission point, while $E_{\rm l}$ is the local photon energy.

### 3.2. *Returning radiation*

What is new in this work with respect to Ref. 1 is the addition of a specific module for including in the KYN code the contribution of returning radiation, motivated by the fact that in previous works[5] it turned out to be important. Calculations are performed using the C++ code SELFIRR:

- for each point on the disk surface (at the distance $\bar{r}_{\rm i}$ from the central BH), all the possible directions along which photons can return on the disk are sampled on a discrete $(\bar{\theta}_{\rm i}, \bar{\varphi}_{\rm i})$ angular mesh, with $\bar{\theta}_{\rm i}$ and $\bar{\varphi}_{\rm i}$ the polar angles with respect to the disk normal;
- then, for each point of the $(\bar{r}_{\rm i}, \bar{\theta}_{\rm i}, \bar{\varphi}_{\rm i})$ grid, the code computes the null geodesics, if any, that arrive to the disk at the selected distance $\bar{r}_{\rm i}$ with incidence direction given by $(\bar{\theta}_{\rm i}, \bar{\varphi}_{\rm i})$ from another point of the disk surface[a];
- finally, the values of the emission radius $\bar{r}_{\rm e}$, angles of emission $\bar{\theta}_{\rm e}$ and $\bar{\varphi}_{\rm e}$ (with respect to the disk normal), transfer function $\bar{G}$ and change in polarization angle $\bar{\Psi}$ which characterize each of the computed geodesics are stored in an external file, to be read by KYN.

The specific Stokes parameters of the photons originally emitted from the disk are still given according to a Novikov-Thorne profile[14] and using the formulae by Ref. 7 as in the case of direct radiation (Sec. 3.1). At the incidence point, the contributions of the different geodesics are summed together, in order to obtain the specific Stokes parameters due to returning radiation in each point of the disk surface.

Once returned to the disk, photons are considered to be reflected; for the sake of simplicity, in this preliminary version of the code we used a 100% albedo prescription, i.e. photons which return to the disk are all reflected towards infinity. For calculating the Stokes parameters of reflected photons we used the Chandrasekhar's

[a] We stress that SELFIRR naturally accounts for photons that, due to lensing, can travel along more than one geodesic between two points on the disk also with the same direction of incidence.

formulae for diffuse reflection (see Ref. 7 for more details). The integrated Stokes parameters at the observer are eventually obtained as mentioned in section 3.1 (see equations 3). The code provides in output the values of the Stokes parameters as functions of the photon energy at the observer, as well as the energy-dependent polarization degree and angle. Besides the outputs for the separate contributions of direct and returning radiation, the code is set to return also the "total" spectra and polarization observable behaviors, obtained by summing together the Stokes parameters of the two components.

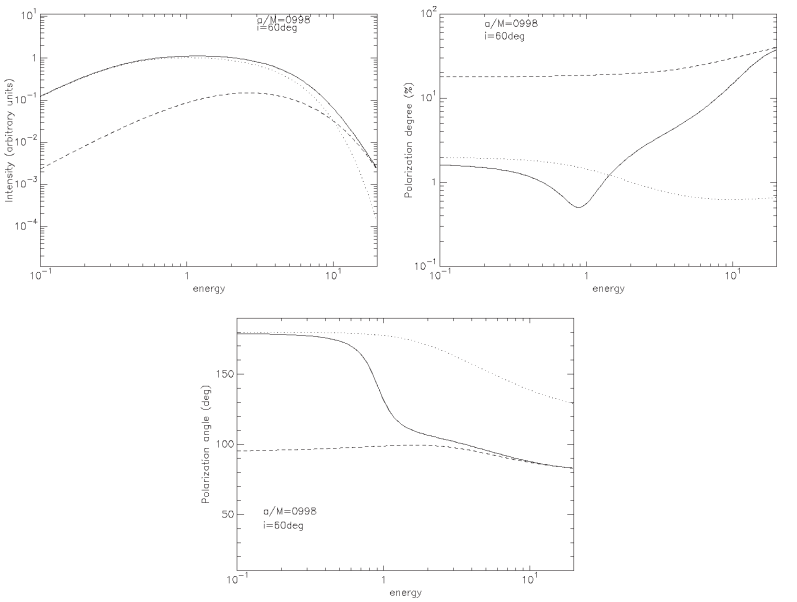



Fig. 1. Spectrum (top-left), linear polarization fraction (top-right) and polarization angle (bottom) produced by a KYN run for a maximally rotating ($a = 0.998$), stellar-mass BH ($M = 10\,M_{\odot}$, $\dot{M} = 3.5 \times 10^{17}$ g/s) with inclination angle $i = 60°$ (solid lines). Dotted and dashed lines represent the separate contributions of direct and returning radiation, respectively.

## 4. Preliminary results

The simplifying assumption of 100% albedo at the disk surface allowed us to compare directly our preliminary results with those obtained in Ref. 5. However, we aim to relax this assumption in future applications, using non-trivial (although still simple) prescriptions for the disk albedo and considering the mutual interplay between the strong gravity effects and those of the disk vertical structure on the returning radiation. As an example, we show in Figure 1 a typical output of our code in the case of a $10\,M_{\odot}$, maximally rotating BH (spin $a = 0.998$), for a value of the accretion rate $\dot{M} = 3.5 \times 10^{17}$ g/s, which roughly corresponds to a disk luminosity of order 10% of the Eddington limit (see Refs. 5, 18). Chandrasekhar's formulae

for a geometrically thin, optically thick disk are used to derive the polarization properties at the emission (see Sec. 3).

The plots show the spectrum, polarization fraction and polarization angle[b] as functions of the photon energy at infinity for an inclination angle of 60° (solid lines), together with the separate contributions of direct (dotted lines) and returning (dashed line) radiation. As a direct comparison with Figure 6 in Ref. 5 suggests, the results obtained from our code are compatible with those originally discussed by Schnittman & Krolik. In particular, the addition of the returning radiation component to the direct one modifies dramatically the polarization pattern expected when only direct radiation is considered. At high photon energies, where the collected photons come from the inner part of the disk and returning radiation effects are indeed more important, the polarization fraction can also exceed 10% and photons are predominantly polarized perpendicularly to the disk surface. On the other hand, polarization properties of low-energy photons, which come from the outermost regions of the disk, tend to follow more closely the Chandrasekhar's prescription.[7] In between these two regimes a transition occurs, where the contributions of direct and returning radiation are equivalent. Here the polarization direction turns out to rotate from parallel to perpendicular and correspondingly the polarization fraction attains a minimum, below 1%.

## 5. Conclusions and future prospects

We presented a numerical method to simulate the spectral and polarization properties of the X-ray radiation emitted by accreting stellar-mass BHs, including the effects of returning radiation alongside those of direct radiation. The preliminary plots shown in Figure 1 and obtained in the simplified 100% albedo prescription turn out to be in good agreement with the results already discussed in literature (see e.g. Ref. 5), although an inverse (observer-to-emitter) approach with respect to previous investigations has been used. Future improvements of our work will then consist in producing more complete simulations, in which the albedo is calculated by using more realistic, self-consistent disk models.

The first results obtained using the basic model illustrated in sections 2 and 3 strengthen the view that X-ray polarization measurements can be actually crucial in assessing strong gravity effects around accreting stellar-mass BHs. Whilst some pioneering efforts in measuring X-ray polarization from stellar-mass BHs were made with the OSO-8 satellite,[19] the poor sensitivity and quasi mono-energetic approach of past instrumentation did not allow to achieve conclusive results. On the other hand, the new techniques available with new-generation polarimeters like IXPE promise to strongly increase the sensitivity of polarization measurements in the 2–8 keV energy range. Numerical simulations, which exploit the IXPE effective

[b]In Figure 1 (as in the plots shown in Ref. 5) a polarization angle of 90° corresponds to photons polarized perpendicularly to the disk surface.

area and modulation factor, have been already performed[20] to test the feasibility of such polarization measurements for radiation emitted by stellar-mass BH accretion disks. These simulations have confirmed the capability of the instrument in constraining the physics of BHs, showing how polarization measurements may be used to determine the BH spin parameter $a$.

## References


1. M. Dovčiak, F. Muleri, R. W. Goosmann, V. Karas and G. Matt, Thermal disc emission from a rotating black hole: X-ray polarization signatures, *MNRAS* **391**, p.32 (2008).
2. B. P. Abbott, et al., Properties of the Binary Black Hole Merger GW150914, *Phys. Rev. Lett.* **116**, p.241102 (2016).
3. R. Bellazzini, E. Costa, G. Matt, and G. Tagliaferri, X-ray Polarimetry: A New Window in Astrophysics, *Cambridge University Press, Cambridge* (2010).
4. J. D. Schnittman and J. H. Krolik, A Monte Carlo code for relativistic radiation transport around Kerr black holes, *ApJ* **777**, p. 11 (2013).
5. J. D. Schnittman and J. H. Krolik, X-ray polarization from accreting black holes: the thermal state, *ApJ* **701**, p. 1175 (2009).
6. M. C. Weisskopf et al., in UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XVIII, *Proc. SPIE* **8859** (2013).
7. S. Chandrasekhar, Radiative transfer, *Dover, New York* (1960).
8. P. A. Connors and R. F. Stark, Observable gravitational effects on polarised radiation coming from near a black hole, *Nature* **269**, p. 128 (1977).
9. R. F. Stark and P. A. Connors, Observational test for the existence of a rotating black hole in CYG X-1, *Nature* **266**, p. 429 (1977).
10. P. A. Connors, T. Piran and R. F. Stark, Polarization features of X-ray radiation emitted near black holes, *ApJ* **235**, p. 224 (1980).
11. H. Ishihara, M. Takahashi and A. Tomimatsu, Gravitational Faraday rotation induced by a Kerr black hole, *Phys. Rev. D* **48**, p.472 (1988).
12. M. Walker and R. Penrose, On quadratic first integrals of the geodesic equations for type { 22} spacetimes, *Comm. in Math. Phys.* **18**, p. 265 (1970).
13. N. I. Shakura and R. A. Sunyaev, Black holes in binary systems. Observational appearance, *A&A* **24**, p. 337 (1973).
14. I. D. Novikov and K. S. Thorne, Astrophysics of black holes, in Black Holes (Les Astres Occlus), *C. Dewitt,m and B. S. Dewitt*, p. 343–450 (1973).
15. R. W. Goosmann and C. M. Gaskell, Modeling optical and UV polarization of AGNs. I. Imprints of individual scattering regions, *A&A* **465**, p. 129 (2007).
16. F. Marin, Modeling optical and UV polarization of AGNs, *A&A* **615**, p. 171 (2018).
17. M. Dovciak, Radiation of Accretion Discs in Strong Gravity (astro-ph/0411605), PhD Thesis, *Charles Univ., Prague* (2004).

18. H. Krawczynski, Tests of general relativity in the strong-gravity regime based on X-ray spectropolarimetric observations of black holes in X-ray binaries, *ApJ* **754**, p. 133 (2012).
19. M. C. Weisskopf, E. H. Silver, H. L. Kestenbaum, K. S. Long and R. Novick, A precision measurement of the X-ray polarization of the Crab Nebula without pulsar contamination, *ApJ* **220**, p. L117 (1978).
20. M. C. Weisskopf, An Overview of X-Ray Polarimetry of Astronomical Sources, *Galaxies* **6**, p. 33 (2018).

# GR simulations of the Rossby Wave Instability: what impacts HFQPOs' observables

Peggy Varniere* and Fabien Casse

*APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité,*
*10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France*
*Laboratoire AIM, CEA/IRFU-CNRS/INSU-Université Paris Diderot, CEA DRF/IRFU/DAp, F-91191 Gif-sur-Yvette, France.*
**E-mail: varniere@ apc.univ-paris7.fr*

Frederic H. Vincent

*LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cité 5 place Jules Janssen, 92195 Meudon, France*

The Rossby-Wave Instability (RWI) has been proposed to be at the origin of the high-frequency QPOs observed in black-hole systems. Here we are presenting the first full GR simulations of the instability around a Kerr black-hole which allow us to explore the impact of the spin on the instability. Those simulations, coupled with a full GR ray-tracing, allow us to directly compare our simulation with the observables we get through X-ray observations.

*Keywords*: black hole binaries; MG15 Proceedings; HFQPOs

## 1. Introduction: NOVAs as a tool for theoretician and observer alike

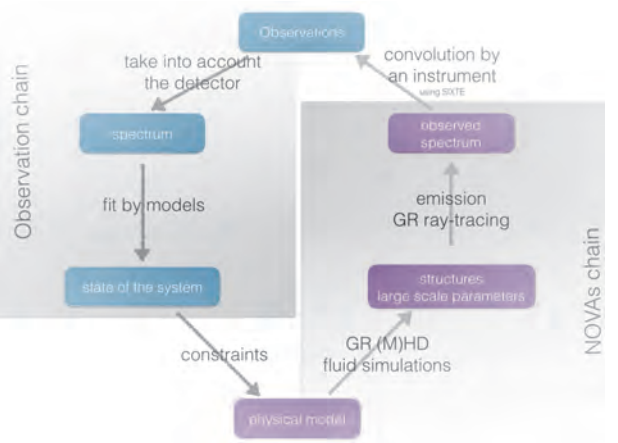



Fig. 1. How the 'standard' observation chain compare with the `NOVAs` chain.

By combining smoothly two GR codes, one providing a full hydrodynamical solution [1, 2][a] and one providing the ray-tracing of the emission[8][b], we now have a fully functional numerical observatory which allows us to obtain spectrums and lightcurves of theoretical models with limited hypotheses. Further linking the output of NOVAs with SIXTE allows us to also test the capacity of new instruments to distinguish between models and explore new possibilities.

Using this tool, we can explore the observables coming from different models. In particular we are interested in the fast variability of microquasars for which there has been a long string of efforts to understand their origin. This is especially true for the origin of the rather elusive High-Frequency Quasi-Periodic Oscillation (HFQPO) in systems containing (or thought to contain) a black hole (BH).

## 2. The RWI in the full Kerr metric

Here we focus on the model based on the Rossby-Wave Instability[3–5] which is predicted to occur when the inner edge of the disk is close to the last orbit. It has been previously proposed to explain several phenomena occurring in the vicinity of black holes[1,2,9] but its existence in an extreme Kerr environment had never been fully demonstrated until recently.

Indeed, using `NOVAs` we performed the first simulations of the RWI in the Kerr metric up to high spin of $a = 0.99$ proving its existence as well as confirming the conditions necessary for it to arise in black-hole disks. From this first set of simulations following the RWI from a Keplerian setup to a near extreme Kerr black-hole, we were able to study in more detail how the location of the RWI influences its physics and what changes, if any, we should be looking for in observations.

Fig. 2. summarizes our findings by showing, for multiple spins, the evolution of the growth rate, saturation level and time to reach saturation of the instability as function of its position with respect to the black hole. On the top part of Fig. 2. we see that the time to reach saturation, while relatively constant at larger $r_b$, the small differences coming mostly from the differences in the setup we used, increased rapidly when the RWI develops inside $\sim 4r_g$. This rise shows the impact of the lapse on the propagation of the instability. While the change is about 40% of the period, the actual numerical value is extremely small as those objects are very fast. It is not possible to detect such differences in observations.

At the bottom of Fig. 2. we see that the growth rate of the instability goes through a maximum for $r_b \sim 5r_g$. This, once again is too small of a change to have any detectable effect as one needs too long of an observation to detect HFQPOs, so we cannot follow the 'instantaneous' evolution of its rms.

[a] Freely available at `https://github.com/amrvac/amrvac`
[b] Freely available at `http://gyoto.obspm.fr`

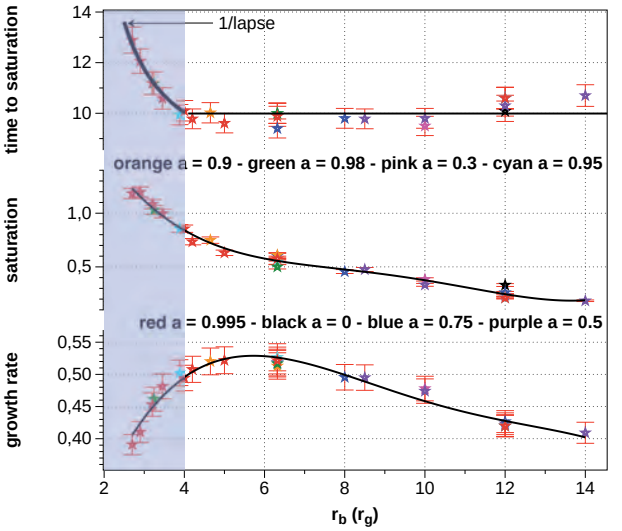



Fig. 2. Evolution of the growth rate of the instability, its saturation level and the time to reach saturation for different spins and positions of the instability ($r_b$).

More interestingly, we see on the middle plot, that the closer to $1r_g$ the RWI develops, the higher its saturation is. This quantity is related to the maximum strength a QPO can attain, hence to how much modulation there is in the observed flux. While we cannot directly test that with observations, indeed, the limitation being the very low number of spin determination for HFQPO sources, this is something we will be exploring, especially with NICER.

## 3. Potential Link with observations

Using `NOVAs`, we look into more detail at ways to compare the RWI with 'standard' observations. Indeed, while a direct comparison of the numerical simulation with observations is impossible, using the NOVAs framework we can produce lightcurves and spectrums that we can then compare with observables.

### 3.1. *What is at the origin of the different HFQPO pairs/ peak selection in the PDS*

The RWI has been proposed to explain the high-frequency QPOs, partially because of its ability to have multiple modes present at the same time. A question that can easily be answered using `NOVAs` is what causes the different modes to be detected. To achieve that we performed numerical simulations having slightly different setups and then produced the associated PDS to see when we get a certain type of pair.

Depending on the setup we get different modes of the RWI and in turn those create different patterns in the PDS depending on which modes dominate and this translates in different peak distributions in the PDS as seen on the left of Fig. 3.

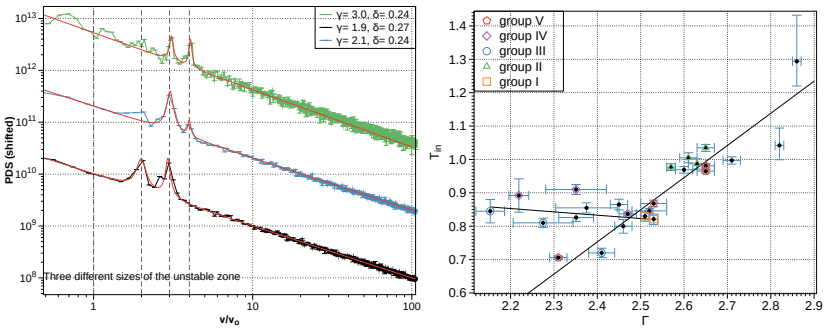



Fig. 3. Left: three PDSs obtained for the same black-hole spin but different local conditions. Right: scatter plot of the power-law slope, Γ, versus the temperature at the inner edge of the disk, $T_{in}$, for all the HFQPO observations in XTE J1550-564[7].

From those simulations we did not yet get a clear criteria for the different peak distributions but we were able to reproduce all the observed peak distributions[7] while having the parameters of the system change by about 30%. Using all the observations of XTE J1550-564 with a published HFQPO we looked at how much variation exists in the system parameters. We see on the right of Fig. 3. that the observed changes in the system parameters are compatible with the kind of changes needed by the RWI to explain the observed PDSs.

### 3.2. *Evolution of HFQPOs with spin/ rms amplitude of the modulation as function of spin*

One of the first things we looked at is the growth of the RWI in the case of different black-hole spins. As you can see on the left of Fig. 4. we get higher saturation levels, and hence detectabilities, as the spin increases. While this is not a direct observable, we can still look at how the maximum rms amplitude of the observed HFQPOs behaves as function of spin. The right graph of Fig. 4. compiled all the HFQPOs' rmses published for black-holes with known spins. Even if this is a limited sample, we see a similar trend as the one we have in our simulations, namely that high spin systems have higher rmses. We will be continuing to look for new HFQPO/spin couples to add to that plot as soon as new data become available.

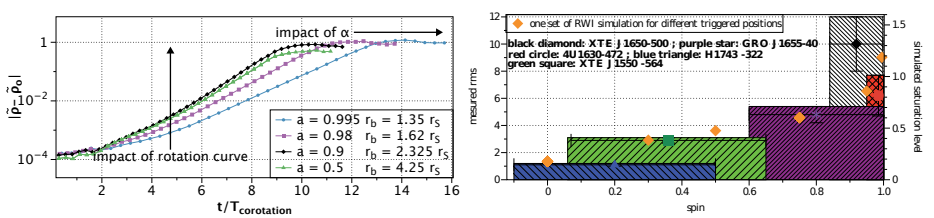



Fig. 4. Left: growth of the RWI as function of the time for different spins. Right: distribution of the rms amplitude of HFQPOs as function of spin for all the known couples.

## 4. Conclusion

Using our new `NOVAs` framework we are able to show that the RWI does indeed develop at the inner edge of a disk at its last stable orbit when full GR is taken into account.

We then explored several observables related to the RWI as a model for HFQPOs and compared them with actual data. We first look at the number and relation between the dominant peaks in Fourier space, showing that they are dependent on the local disk conditions where the RWI is developing. On top of that, we were able to reproduce all of the observed HFQPOs' distributions as of today and showed that the required changes in the local disk is compatible with the observed values.

Also, when computing the RWI for the full range of spins we saw that, under similar conditions, the higher spins will have a higher saturation level. While we do not have observations of different systems under the same conditions, we looked at the maximum rms observed as function of spin and showed that higher spin systems tend to get HFQPOs with higher rms.

All in all, using `NOVAs` allowed us to strengthen the RWI as a possible explanation for the HFQPO.

## Acknowledgements

We acknowledge the financial support of the UnivEarthS Labex program at Sorbonne Paris Cite (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02)..

## References

1. Casse, F., Varniere, P. & Meliani, Z. Impact of the gravity of a Schwarzschild black hole upon the Rossby wave instability. mnras **464**, 3704–3712 (2017). DOI 10.1093/mnras/stw2572.
2. Casse, F. & Varniere, P. Impact of the Spin on the Rossbby Wave Instability. mnras (2018).
3. Tagger, M. & Varniere, P., Accretion-Ejection Instability, MHD Rossby Wave Instability, Diskoseismology, and the High-Frequency QPOs of Microquasars, apJ, **652**, 1457-1465 (2006). DOI 10.1086/508318
4. Varniere, P. and Tagger, M. and Rodriguez, J., A microquasar classification from a disk instability perspective A&A , **525**, A87 (2011). DOI 10.1051/0004-6361/201015028
5. Varniere, P. and Tagger, M. and Rodriguez, J., A possible interpretation for the apparent differences in LFQPO types in microquasars A&A, **545**, A40 (2012). DOI 10.1051/0004-6361/201116698
6. Varniere, P., Casse, F. & Vincent, F. H. Gr simulations of the rwi: What impacts hfqpos observables. to be submitted (2019).
7. Varniere, P. & Rodriguez, J. Looking for the elusive 3:2 ratio in xte j1550-564. ApJ (2018).

8. Vincent, F. H., Paumard, T., Gourgoulhon, E. & Perrin, G. GYOTO: a new general relativistic ray-tracing code. Classical and Quantum Gravity **28**, 225011 (2011). DOI 10.1088/0264-9381/28/22/225011. 1109.4769.
9. Vincent, F. H. and Meheut, H. and Varniere, P. and Paumard, T., Flux modulation from the Rossby wave instability in microquasars' accretion disks: toward a HFQPO model, A&A, **551**, A54 (2013) DOI 10.1051/0004-6361/201220695

# Disk spectrum evolution as function of spin: impact on model fitting

Peggy Varniere* and Fabien Casse

*APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France*
*Laboratoire AIM, CEA/IRFU-CNRS/INSU-Université Paris Diderot, CEA DRF/IRFU/DAp, F-91191 Gif-sur-Yvette, France.*
**E-mail: varniere@ apc.univ-paris7.fr*

Frederic H. Vincent

*LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cité 5 place Jules Janssen, 92195 Meudon, France*

Fabien Casse

*APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France*
*Laboratoire AIM, CEA/IRFU-CNRS/INSU-Université Paris Diderot, CEA DRF/IRFU/DAp, F-91191 Gif-sur-Yvette, France.*

With NICER up and running and the ATHENA mission in preparation there is need to study in more detail what is the influence, if any, of the spin on the observed spectrum. Using the newly developed NOVAs framework which is comprised of a GR code, GRAMRVAC, coupled with the ray-tracing GR code, gyoto, we computed a series of simulated disk spectrums for a variety of black-hole spins. Using standard models from xspec we fitted those spectrums and looked at how they performed to give us the parameters of the simulated spectrums. Here we report on our findings and how we could improve the existing models to get better insights into the physical large scale disk parameters.

*Keywords*: black hole binaries; MG15 Proceedings; energy spectrum

`NOVAs`, our Numerical Observatory of Violent Accreting systems[a], can have a lot of uses, such as preparing observations, thanks to being easily coupled to `SIXTE`, but also trying to understand the origin of some observed features. Here we show a few examples of what `NOVAs` can do in particular related to the energy spectrum of accretion disks around black-holes.

## 1. Computing a simulated disk spectrum using `NOVAs`

### 1.1. *Introduction: Forming the Inner edge of the disk*

The first step to compute the energy spectrum of an accretion disk around a black-hole is to obtain the shape of the disk. Using `NOVAs`, in particular the `GRAMRVAC`

[a]See other proceedings by the same author in this volume.

step, we ran a series of simulations following the formation of a stable accretion disk reaching its last stable orbit. We focused on the shape of the temperature profile as function of the spin, needed to compute the emission, and also on the position of the inner edge.

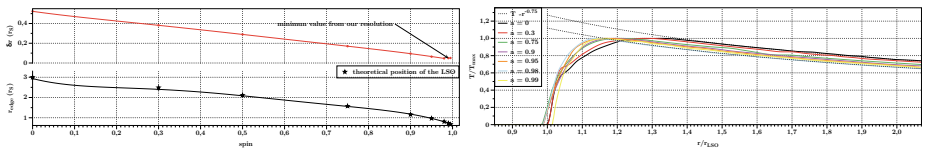



Fig. 1. Left: position and shape of the inner edge of the simulated disk forms as compared to the predicted last stable orbit. We see that, as the spin increases, the inner edge of the disk gets sharper and thinner. Right: temperature profile as function of the radius in the disk (normalized to the last stable orbit) for increasing spins. We see that a higher spin means a sharper and thinner edge.

As expected, we recover the position of the last stable orbit very precisely (left Fig.1) but the curvature of the temperature profile depends on the spin with lower spins having a larger turnover zone (right Fig.1). This means that higher spins imply a sharper and thinner inner edge of the disk.

### 1.2. *simulated disk spectrum*

Once we obtain the temperature in every point of our disk we can compute its blackbody emission and ray-trace it back to the observer using the `gyoto` step of `NOVAs`. One of the advantages of `NOVAs` is that we can turn on/off certain aspects of the process in order to test its impact.

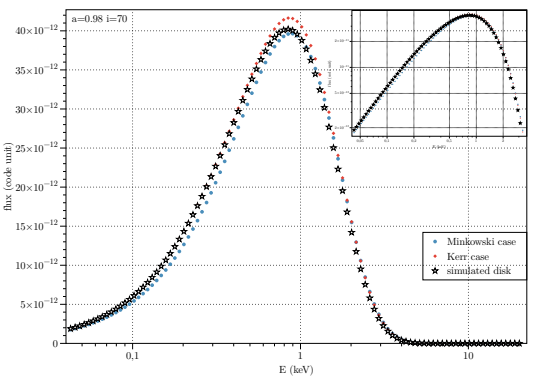



Fig. 2. Comparison between the full GR simulated spectrum (star) with the cases of a Shakura&Sunyaev equivalent disk with either a Minkowski ray-tracing (diskbb equivalent) or Kerr. The corner plot is the same plot in log-log.

The first test was to see which aspect of the full GR treatment was necessary and when. To that respect we compared in Fig.2 the cases with full GR simulation (stars), GR only for the fluid dynamical part but Minkowski space for the ray-tracing back to the observer (Minkowski case) and with the cases of a Shakura&Sunyaev equivalent disk with either a Minkowski ray-tracing (diskbb equivalent) or Kerr.

As expected from the temperature profile which becomes closer to the equivalent Shakura&Sunyaev case at higher spin (sharper) we see in Fig.2 (in log-lin and log-log), that for high spins, the simulated disk spectrum is close to both the SS-Kerr and SS-Minkowski cases, which is an acceptable approximation of the full GR simulation in particular if we are looking for a fast fit of the overall disk parameters.

### 1.3. *Impact of inclination*

Another advantage of a numerical observatory is that we can look at the same system under different conditions such as for example the inclination. This allows us to test the impact of parameters related to the position of the system with respect to the observer.

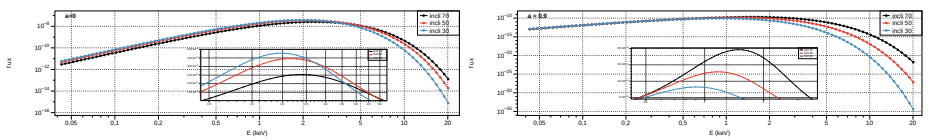


Fig. 3. Energy spectrum of the simulated disk for different inclinations of the system for the case of zero spin (left) and a higher spin of $a = 0.9$ (right).

Fig.3 shows the impact of inclination on the shape of the spectrum in the case of no spin (a=0) and a higher spin (a=0.9). The most noticeable effect is on the energy of the maximum, especially for high inclination systems where the Doppler effect is the strongest.

## 2. From numerical simulation to spectral fitting

A great advantage of a numerical observatory is that we know the 'correct' value from the simulation and therefore we can evaluate the error made on the parameters determination coming from each spectral fitting method applied on our 'observations'. As the output of `NOVAs` is compatible with `xspec` we can use the standard fitting procedure on our numerical observations and assess how close the fitted parameters are to those input in the simulation.

### 2.1. *Error on the disk parameters*

Often the main results from the spectral fitting are the overall disk parameters such as the position of the inner edge of the disk and its temperature. Here we

are using energy spectrums coming from our simulations and then use `xspec` to fit them with `diskbb, ezdiskbb, diskpbb` to see how well each of those performs when compared with the value input in the simulation.

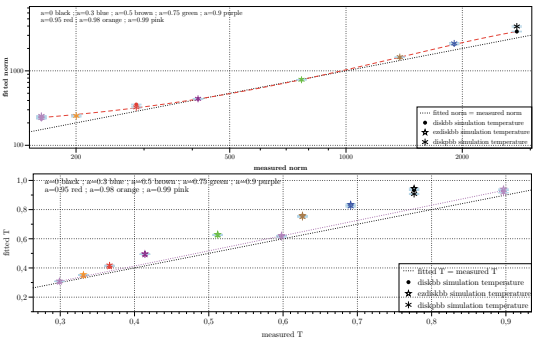



Fig. 4. Comparison of the fitted parameters of the disk, namely the norm and inner temperature, compared to the measured value in the simulations for different spins values.

We see on Fig.4 that for all those cases the norm was reasonably recovered for all spins especially for mid-level spins, while the temperature was much better recovered for high spins, with an error close to 25% at low spin. This can be related to the difference in shape of the temperature profile which is much closer to the diskbb model at higher spin than it is at lower spin.

## 2.2. *Testing the efficiency of the spin determination*

Using the simulation of an accretion disk formed around different spin black-holes we created the 'observations' to plug into `xspec` and used the `kerrbb` model to fit the spin value from the energy spectrum. To simplify the test we input the exact value of inclination, mass and distance of the `NOVAs` observation, leaving only the spin to be fitted.

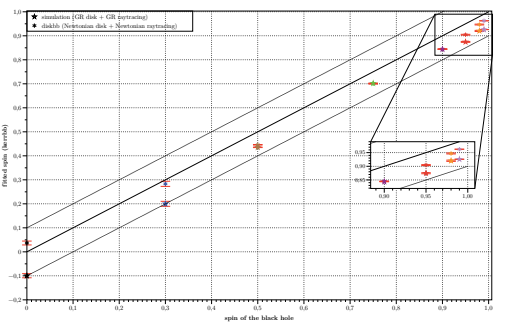



Fig. 5. Comparison between the spin of the numerical observation and its fitted value with `kerrbb`.

The resulting fitted spins are represented in Fig.5 with respect to the spin of the numerical observation. For all the spins the fitted value is within 0.1 of the actual spin of the numerical observation and the quality of the fitted value improves as the spin increases. It is interesting to note that, for all of the cases explored here, the fitted value was always lower than the actual spin of the object. This is a preliminary study and more cases need to be explored but it shows a good efficiency of the spin fitting method if we know well the system's parameters.

### 2.3. *Understanding the origin of some observed features*

Another possible application of the `NOVAs` framework is to look for the cause of some observed features. Here we are looking at what could be the reason behind the fact that it is harder to obtain good $\chi^2$ for the spectral fit of high-spin systems, and that for every state.

In order to explore the difference between high and low spin we look at the shape of the spectrum, or more precisely at how the shape of the spectrum compares with the `diskbb` spectrum which is often used to fit spectral data.

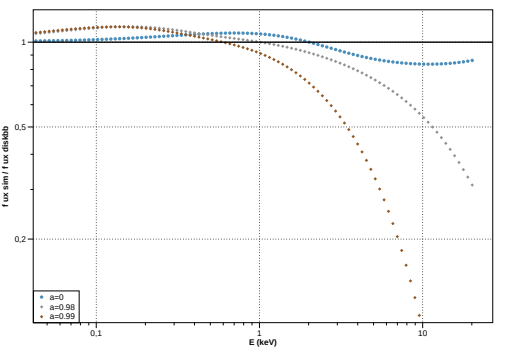



Fig. 6. Evolution as function of energy of the simulated flux over the `diskbb` flux for the same system.

Fig.6 shows that for low spin (blue points) the full GR simulated flux is relatively close to the `diskbb` flux that it is fitted against but as soon as we go to high spins the overall shape of the spectrum starts to diverge. This would then lead to a worse $\chi^2$ for higher spins than can be achieve for lower spins.

## 3. Conclusion

Here we focus on a few cases where `NOVAs`, our Numerical Observatory of Violent Accreting systems[b], have helped us understand observations. Indeed, using `NOVAs` we can now explore the influence of hard-to-measure parameters, such as the spin

[b]See other proceedings by the same author in this volume.

of the black-hole or the inclination of the system, in the hope of finding some observational constraints.

Among the numerous applications we first looked at how well some commonly-used `xspec` models performed to retrieve the information from our simulated spectrums and found that most models have similar pitfalls but give adequate results. We also explore how well the `kerrbb` model constrains the spin and found that for our range of spin and accretion rates the `kerrbb` model was able to get the spin within 0.1 of its actual value.

Finally, we propose the relativistic effects on the disk emission to be responsible for the difficulty to get good $\chi^2$ for high spin sources. This emphasizes the need to improve our disk modelisation in `xspec` for newer missions.

**Acknowledgements**

We acknowledge the financial support of the UnivEarthS Labex program at Sorbonne Paris Cite (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02)..

**References**

1. Casse, F., Varniere, P. & Meliani, Z. Impact of the gravity of a Schwarzschild black hole upon the Rossby wave instability. mnras **464**, 3704–3712 (2017). DOI 10.1093/mnras/stw2572.
2. Casse, F. & Varniere, P. Impact of the Spin on the Rossbby Wave Instability. accepted by mnras (2018).
3. Varniere, P. and Tagger, M. and Rodriguez, J., A microquasar classification from a disk instability perspective A&A , **525**, A87 (2011). DOI 10.1051/0004-6361/201015028
4. Varniere, P. and Tagger, M. and Rodriguez, J., A possible interpretation for the apparent differences in LFQPO types in microquasars A&A, **545**, A40 (2012). DOI 10.1051/0004-6361/201116698
5. Vincent, F. H., Paumard, T., Gourgoulhon, E. & Perrin, G. GYOTO: a new general relativistic ray-tracing code. Classical and Quantum Gravity **28**, 225011 (2011). DOI 10.1088/0264-9381/28/22/225011. `1109.4769`.
6. Vincent, F. H. and Meheut, H. and Varniere, P. and Paumard, T., Flux modulation from the Rossby wave instability in microquasars' accretion disks: toward a HFQPO model, A&A, **551**, A54 (2013) DOI 10.1051/0004-6361/201220695

# NOVAs: A Numerical Observatory of Violent Accreting systems

Peggy Varniere* and Fabien Casse

*APC, AstroParticule et Cosmologie, Université Paris Diderot, CNRS/IN2P3, CEA/Irfu, Observatoire de Paris, Sorbonne Paris Cité, 10, rue Alice Domon et Léonie Duquet, 75205 Paris Cedex 13, France*
*Laboratoire AIM, CEA/IRFU-CNRS/INSU-Université Paris Diderot, CEA DRF/IRFU/DAp, F-91191 Gif-sur-Yvette, France.*
**E-mail: varniere@ apc.univ-paris7.fr*

Frederic H. Vincent

*LESIA, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, Univ. Paris Diderot, Sorbonne Paris Cité 5 place Jules Janssen, 92195 Meudon, France*

Here we are presenting NOVAs, a Numerical Observatory of Violent Accreting systems, which couples a GR AMR MPI fluid dynamics code (GRAMRVAC) able to follow accretion around a Kerr Black-hole with the ray-tracing code GYOTO. Together, they allow us to test different models by running the simulation and obtaining spectral energy distributions and power-density spectrums from which we can extract the same observables as for 'real' observations, hence making it a Numerical Observatory.

*Keywords*: black hole binaries; MG15 Proceedings; HFQPOs

## 1. Introduction: why a numerical observatory?

While there are a lot of observational data and theoretical models trying to explain black-hole system, it is hard to bridge the gap between observables, such as the energy spectrum for example, and an analytical or numerical model. This is especially true when looking at highly relativistic systems such as the inner region of an accreting black-hole.

The idea behind NOVAs is to create consistently, from numerical GR-(M)HD simulations the same outputs as we have from observations, *i.e.* lightcurve, energy spectrum and Power-Density Spectrum, in a compatible format to be analyzed by software like `xspec`. It makes use of several existing or in development codes which are:

`GRAMRVAC` All the general relativistic (GR) fluid dynamics are done with the general relativistic version of MPI-AMRVAC[a]. In most of the simulations presented here we used a Harten, Lax and van Leer (HLL) solver linked to a Koren slope limiter. See [1, 2] for the full code description and some applications.

`GYOTO` For all the GR ray-tracing computations, we use the open-source[b] GYOTO code. Photons are traced by integrating the geodesic equation using a

[a]Freely available at `https://github.com/amrvac/amrvac`
[b]Freely available at `http://gyoto.obspm.fr`

Runge-Kutta-Fehlberg adaptive-step integrator at order 7/8 (meaning that the method is 8th order, with an error estimation at 7th order). From such maps of specific intensity, the light curve (flux as a function of time) is derived by summing all pixels weighted by the element of the solid angle, which is subtended by each pixel. See [7] for the code presentation.

`SIXTE` In order to add instrumental effects we use the SIXTE[c] package for X-Ray telescope observation simulations. It allows us to undertake instrument performance analyses and to produce simulated event files for mission- and analysis studies.

On top of developping the GR formalism for `GRAMRVAC` we have added to the codes, when needed, new outputs and formats so that one can smoothly go from the fluid simulation to the spectrum observed by an instrument.

## 2. How does NOVAs works

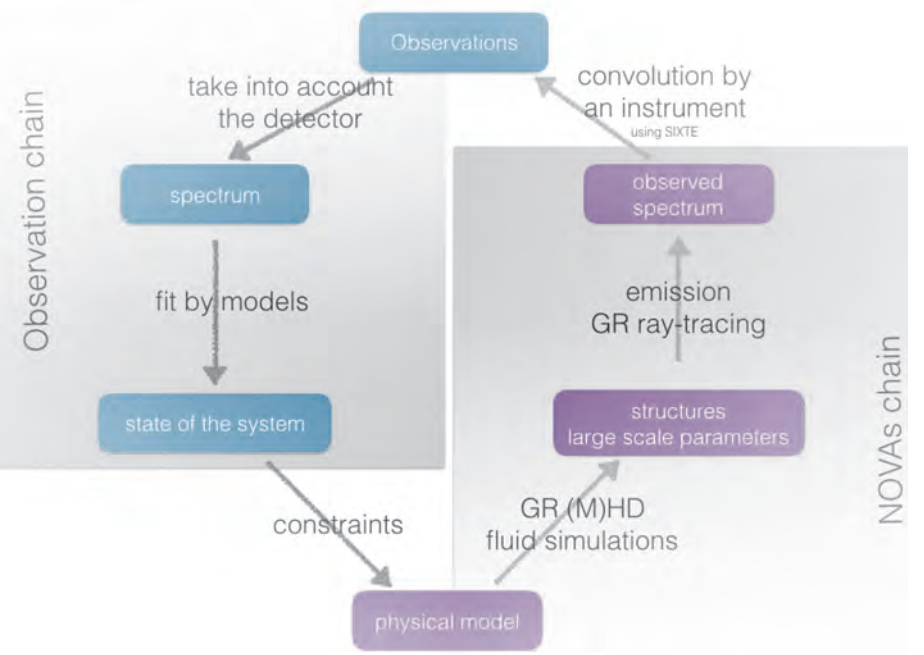



Fig. 1. How the 'standard' observation chain compares with the `NOVAs` chain.

Fig. 1 shows how the numerical observatory of `NOVAs` works when compared with 'standard' observations. This emphasizes how a numerical observatory is complementary to 'standard' observations.

[c]Freely available at `https://www.sternwarte.uni-erlangen.de/research/sixte/index.php`

To explain in more detail how the NOVAs chain works, lets look at the case of an accretion disk orbiting around a Kerr black hole, of spin a, mass M:

– The NOVAs chain starts from a physical model. In our example below we have matter orbiting around a Kerr black hole of spin a, mass M. The first step is to use `GRAMRVAC` to evolve the system in order to obtain a self consistent disk. Here we are particularly interested in the shape of the inner edge formed, left of Fig. 2, as it will first impact the direct emission of the disk but also as it can lead to disk instabilities such as the Rossby-Wave Instability [3, 4, 5, 8].

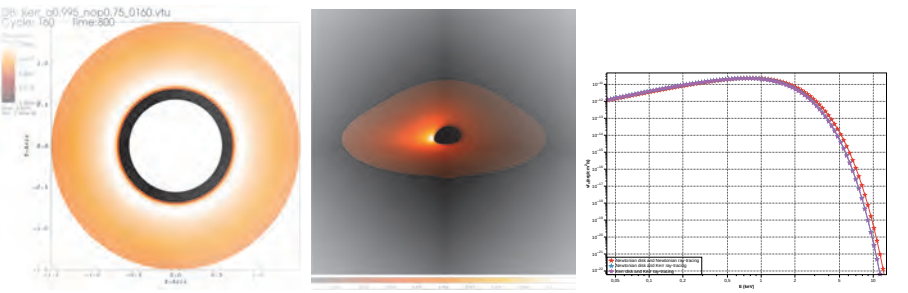


Fig. 2. Left: snapshot from the GR-HD simulation leading to the creation of a disk. Middle: the emission ray-traced back to the observed for a perfect instrument. Right: energy spectrum of the same snapshot.

– Once we have the overall structure of the system from the fluid simulation and the large scale parameters of the system, such as its distance and inclination, we use `GYOTO` to ray-trace back to the observer the emitted flux as a function of energy. A zoom in on the inner portion of the disk is shown in the middle figure of Fig. 2.

– Lastly, those can in turn can be translated into a lightcurve and an energy spectrum for the source (directly or going through a non-perfect instrument using `SIXTE`) as shown on the right of Fig. 2. Then those can be reduced through the standard analysis process to access observables.

There are numerous applications to that numerical observatory, among those we can make predictions from models, test how certain observed features can be explained, test the impact of parameters that are hard to pinpoint such as the inclination, and also test the detectability of certain features with new instruments using `SIXTE`.

## 3. First Applications

Our first application of this Observatory is to study the ability of the Rossby-Wave Instability to reproduce the High Frequency Quasi-Periodic Oscillations detected as peaks in the Power-density spectrum of some black-hole binaries. This is the first

time a full GR study of the RWI has been done with an extreme Kerr black-hole and the associated PDS is computed.

### 3.1. *The RWI around an extreme Kerr black-hole*

While the RWI has been demonstrated to exist in a variety of setups this has only been done in Newtonian and Pseudo-Newtonian approaches until recently. For that reason, the 'first light' of NOVAs was to confirm that the RWI develops as expected in the Kerr metrics, even for high spins.

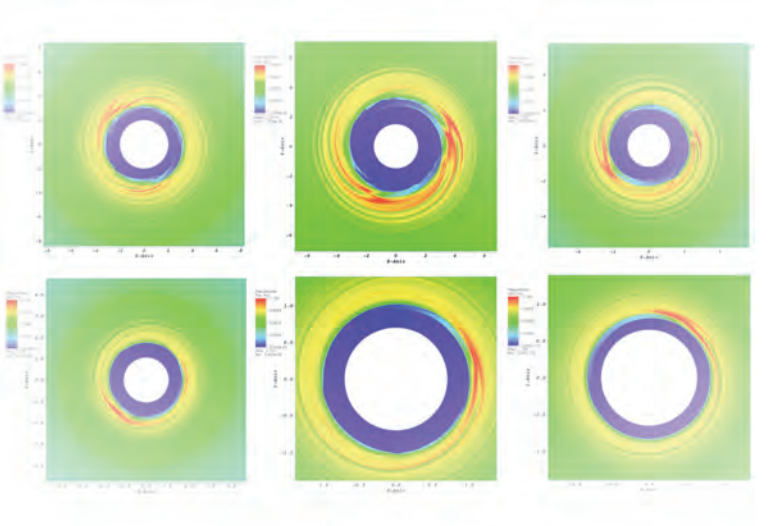


Fig. 3. Snapshots from the GR-HD simulation with the RWI active at the inner edge of a disk orbiting a black hole of spin a={0, 0.3, 0.5, 0.75, 0.95, 0.98}.The x-axis is scaled in $r_S$.

As we can see on Fig. 3., in every simulation the RWI develops close to the last stable orbit,as expected from linear study [3]. The exact position depending on the density profile at the inner edge of the disk, hence on the spin of the black hole. The full impact of the spin on the instability has been explored in [1, 2].

### 3.2. *HFQPOs created from the RWI*

Thanks to `NOVAs` having a full observatory capability we can go one step further and make the lightcurve associated with those simulations as it would be observed with an X-ray satellite and then use standard reduction tools to obtain the power-density spectrum. This allows us to use the most direct way to check if the RWI can indeed produce the High-Frequency Quasi-Periodic Oscillations it has been proposed to explain. Indeed those HFQPOs are detected in the form of peaks in the PDS and, thanks to `NOVAs`, we can produce the PDS associated with our simulations.

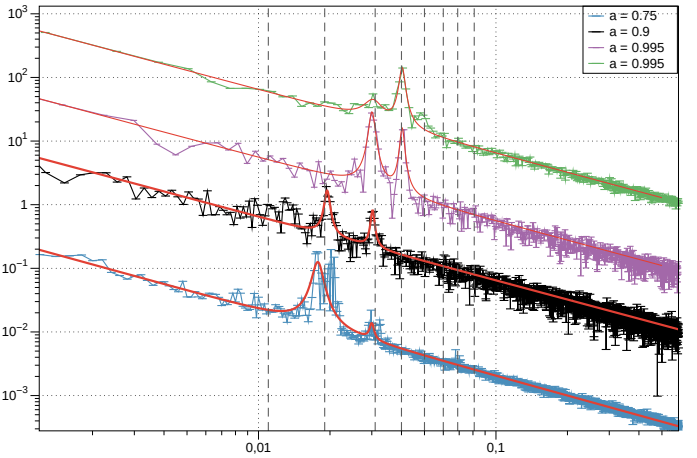



Fig. 4. a small selection of PDSs from our simulations showcasing similar peak ratios as is observed in microquasars.

We see in Fig. 4 that the RWI does not only produces peaks in the PDS and therefore confirm itself as a candidate for the HFQPOs, but it also sometimes exhibits 3:2 and 4:3 ratios for the peaks similar to what is observed in sources such as XTE J1550-564 [6]. In the context of the RWI what fixes the strength of each peak is the local conditions. Indeed, the two upper curves of Fig. 4 are for the same black hole spin but different local disk conditions, leading to different ratios of the strength between the peaks.

## 4. Conclusion

By combining smoothly two GR codes, one providing a full fluid dynamical solution and one providing the ray-tracing of the emission back to the observer, we now have a fully functional numerical observatory, `NOVAs`, which allows us to obtain spectrums and lightcurves of theoretical models with limited hypotheses. Further linking the output of `NOVAs` with SIXTE allows us to also test the capacity of new instruments to distinguish between models and explore new theoretical features.

Among the numerous applications we first looked at the RWI to demonstrate that it will be trigged near the last stable orbit of any spin black-hole as it was linearly predicted. We then computed the lightcurves and associated PDSs and proved the ability of the RWI to produce peaks in the PDS similar to what is observed in microquasars.

## Acknowledgements

We acknowledge the financial support of the UnivEarthS Labex program at Sorbonne Paris Cite (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02)..

## References


1. Casse, F., Varniere, P. & Meliani, Z. Impact of the gravity of a Schwarzschild black hole upon the Rossby wave instability. mnras **464**, 3704–3712 (2017). DOI 10.1093/mnras/stw2572.
2. Casse, F. & Varniere, P. Impact of the Spin on the Rossbby Wave Instability. accepted by mnras (2018).
3. Tagger, M. & Varniere, P., Accretion-Ejection Instability, MHD Rossby Wave Instability, Diskoseismology, and the High-Frequency QPOs of Microquasars, apJ, **652**, 1457-1465 (2006). DOI 10.1086/508318
4. Varniere, P. and Tagger, M. and Rodriguez, J., A microquasar classification from a disk instability perspective A&A , **525**, A87 (2011). DOI 10.1051/0004-6361/201015028
5. Varniere, P. and Tagger, M. and Rodriguez, J., A possible interpretation for the apparent differences in LFQPO types in microquasars A&A, **545**, A40 (2012). DOI 10.1051/0004-6361/201116698
6. Varniere, P. and Rodriguez, J., Looking for the elusive 3:2 ratio of high-frequency quasi-periodic oscillations in the microquasar XTE J1550−564, accepted by ApJ (2018). DOI
7. Vincent, F. H., Paumard, T., Gourgoulhon, E. & Perrin, G. GYOTO: a new general relativistic ray-tracing code. Classical and Quantum Gravity **28**, 225011 (2011). DOI 10.1088/0264-9381/28/22/225011. 1109.4769.
8. Vincent, F. H. and Meheut, H. and Varniere, P. and Paumard, T., Flux modulation from the Rossby wave instability in microquasars' accretion disks: toward a HFQPO model, A&A, **551**, A54 (2013) DOI 10.1051/0004-6361/201220695

# Four tensors determining thermal and electric conductivities of degenerate electrons in magnetized plasma

Maria V. Glushikhina[1**] and Gennady S. Bisnovatyi-Kogan[1,2*]

[1] *Space Research Institute of Russian Academy of Sciences, Moscow, 117997, Russia*
[*] *E-mail: gkogan@iki.rssi.ru*
[**] *E-mail: m.glushikhina@iki.rssi.ru*

[2] *National Research Nuclear University (MEPhI), Kashirskoe Shosse, 31, Moscow, 115409, Russia*

A solution of the Boltzmann equation is obtained for a magnetized plasma with strongly degenerate electrons and nondegenerate nuclei. The components of the diffusion, thermal diffusion, and diffusion thermoeffect tensors in a nonquantizing magnetic field are calculated in the Lorentz approximation without allowance for electron-electron collisions, which is asymptotically accurate for plasma with strongly degenerate electrons. Asymptotically accurate analytical expressions for the electron diffusion, thermal diffusion, and diffusion thermoeffect tensors in the presence of a magnetic field are obtained. The expressions reveal a considerably more complicated dependence on magnetic field than analogous dependences derived in the previous publications on this subject.

*Keywords*: neutron stars; dense matter; kinetics.

## 1. Introduction

The kinetic coefficients in the crusts of neutron stars and the cores of white dwarfs play an important role in the evolution of these stars. The heat fluxes and current densities are determined by heat conduction, diffusion, thermal diffusion, and the diffusion thermoeffect. To calculate these coefficients, it is necessary to know the transport properties of a dense stellar matter, in which electrons are strongly degenerate and form a nearly ideal Fermi gas, while ions are nondegenerate and form either a Coulomb liquid or a Coulomb crystal. Under such conditions, electrons are the main carriers of charge and heat, and the kinetic coefficients are mainly determined by electron scattering from ions. Knowing the distributions of heat and current, one can calculate the magnetothermal evolution of a neutron star with the crust that forms a Coulomb crystal [1]. The kinetic coefficients of degenerate electrons in neutron stars and white dwarfs in a magnetic field were analyzed in [2,3]. The ratio between the electric conductivities along and perpendicular to the magnetic field was phenomenologically obtained in [2]. It is given by

$$\frac{\sigma_\perp}{\sigma_\parallel} = \frac{1}{1+(\omega\tau)^2} \tag{1}$$

and was used in [3] to calculate the thermal and electroconductivity coefficients of a degenerate electron gas in the crusts of neutron stars. Here, $\omega = eB/(m_e c)$ is the cyclotron electron frequency, $\tau$ is the time between electron collisions, e is the

electron charge, $m_e$ is the electron mass, B is the magnetic field strength, and c is the speed of light. The influence of the magnetic field on the electron thermal conductivity and electric conductivity in form (1) was used in many subsequent papers.

The electron thermal conductivity tensor in a dense magnetized matter was calculated in [4] by solving the Boltzmann equation by the Chapman-Enskog method for nondegenerate electrons and in the Lorentz approximation for strongly degenerate non-relativistic electrons. Similar calculations had been done in [5] for tensors of diffusion, thermal diffusion and diffusion thermoeffect.

In the present contribution, we briefly present results of [5]. The expressions reveal a substantially more complicated dependence on magnetic field than those obtained in earlier publications on this subject. The use of the kinetic coefficients calculated in this paper makes it possible to take into account the processes occurring in the crust of a neutron star more accurately. The obtained expressions can also be used to describe the transport coefficients in other magnetized objects containing free degenerate electrons.

## 2. Boltzmann equations and transfer equations

We use a Boltzmann equation for electrons, in a magnetic field, with an allowance of arbitrary degeneracy, and assuming them as non-relativistic. We consider the electron gas in a crystal lattice of heavy nuclei, and take into account the interaction of the electrons with a nondegenerate nuclei and with one another. If the mass of the nucleus $m_N$ is much greater, than the electron mass $m_e$, then to the terms $\sim m_e/m_N$ the details of the distribution function $f_{N0}$ are unimportant, and the calculations can be made for arbitrary $f_{N0}$.

Boltzmann equation, which describes the time variation of the electron distribution function $f$ in presence of the electric and magnetic fields is written as[6,7]

$$\frac{\partial f}{\partial t} + c_i \frac{\partial f}{\partial r_i} - \frac{e}{m_e}(E_i + \frac{1}{c}\varepsilon_{ikl} c_k B_l)\frac{\partial f}{\partial c_i} + J = 0. \tag{2}$$

Here $(-e), m_e$ are the charge (negative) and the mass of the electron, $E_i, B_i$ are the strength of the electric field, and magnetic induction, $\varepsilon_{ikl}$ is the totally antisymmetric Levi-Civita tensor, $c$ is the speed of the light. The collision integral $J$ for arbitrary degenerate electrons, from[8,9], is written in the form,

$$J = J_{ee} + J_{eN} = R \int [f^{'} f_1^{'}(1-f)(1-f_1) - f f_1 (1-f^{'})(1-f_1^{'})] \times g_{ee} b\, db\, d\varepsilon dc_{1i}$$
$$+ \int [f^{'} f_N^{'}(1-f) - f f_N (1-f^{'})] \times g_{eN} b\, db\, d\varepsilon dc_{Ni}. \tag{3}$$

Here, the impact parameter $b$, and $\varepsilon$ are geometrical parameters of particle collisions with relative velocities $g_{ee}, g_{eN}, R = \frac{2m_e^3}{h^3}$. The integration in electron part of the

collision integral in (3) is performed over the phase space of the incoming particles ($dc_{1i}$), and the physical space of their arrival ($b\,db\,d\varepsilon$)[8]. The velocity functions after collision are marked with touches.

The Boltzmann equation for electrons with a binary collision integral (3) may be applied in conditions, when the electron gas can be considered as ideal, meaning that the kinetic energy of the electrons is much larger than the energy of electrostatic interactions. It is applicable for plasma at small density. In the neutron stars and white dwarfs we have an opposite conditions of plasma at very large density, when it is important to take into account the electrons degeneracy. It is known from the statistical physics, that a gas of strongly degenerate electrons becomes ideal, because large Fermi energy substitutes here the thermal energy[10]. The calculations in this paper are applied to the high density plasma with degenerate electrons.

The collision integral similar to $J_{ee}$ from (3) for strongly degenerate neutrons in nuclear matter had been found analitically in[9], see also[11]. It was found that that in the presence of non-degenerate heavy nuclei and strongly degenerate neutron, the input of collisions between them in the heat transfer and diffusion coefficients is negligibly small, in comparison with neutron-nuclei collisions. The same situation we have for the strongly degenerate electrons. Therefore for strongly degenerate electrons the Lorentz approximation, with account of collisions between light and heavy particles only, is asymptotically exact. So for our consideration we can neglect $J_{ee}$ in comparison with $J_{eN}$, and we can equate $J = J_{eN}$ in the (3).

Lets introduce the thermal velocity of electrons, $v_i = c_i - c_{0i}$, where $c_{0i}$ is the mass-average velocity. So we can write the Boltzmann equation with respect to the thermal velocity in the form[7]

$$\frac{df}{dt} + v_i \frac{\partial f}{\partial r_i} - \left[ \frac{e}{m_e}(E_i + \frac{1}{c}\varepsilon_{ikl} v_k B_l) + \frac{dc_{0i}}{dt} \right] \frac{\partial f}{\partial v_i} - \frac{e}{m_e c}\varepsilon_{ikl} v_k B_l \frac{\partial f}{\partial v_i} - \frac{\partial f}{\partial v_i} v_k \frac{\partial c_{0i}}{\partial r_k} + J = 0. \tag{4}$$

The transfer equations for the electron concentration, total momentum, and electron energy, in the two-component mixture of electrons and nuclei, can be obtained in a usual manner from the Boltzmann equation in a quasi-neutral plasma [6–8,12].

## 3. Kinetic coefficients of strongly degenerate electrons in magnetic field

We solve the Boltzmann equation by Chapman-Enskog successive approximation method[8]. In the zero-order approximation, the electron distribution function $f_0$ in thermodynamic equilibrium is determined by the Fermi-Dirac function. In the first order approximation, we seek the distribution function $f$ in the form: $f = f_0[1 + \chi(1 - f_0)]$, where $\chi$ is a small correction.

The function $\chi$ allows representing the solution in the form[4]

$$\chi = -A_i \frac{\partial \ln T}{\partial r_i} - n_e D_i d_i \frac{G_{5/2}}{G_{3/2}}, \tag{5}$$

$$d_i = \frac{\rho_N}{\rho}\frac{\partial \ln P_e}{\partial r_i} - \frac{\rho_e}{P_e}\frac{1}{\rho}\frac{\partial P_N}{\partial r_i} + \frac{en_e}{P_e}(E_i + \frac{1}{c}\varepsilon_{ikl}c_{0k}B_l) - \frac{m_e}{kT}F_i. \tag{6}$$

The functions $A_i$ and $D_i$ describe the heat flux and particle diffusion, respectively. It was demonstrated in[6,7] that polar vectors $A_i$ and $D_i$ in the presence of a magnetic field with an axial vector $B_i$ can be sought for in the form:

$$\begin{aligned} A_i &= A^{(1)}v_i + A^{(2)}\varepsilon_{ijk}v_jB_k + A^{(3)}B_i(v_jB_j), \\ D &= D^{(1)}v_i + D^{(2)}\varepsilon_{ijk}v_jB_k + D^{(3)}B_i(v_jB_j), \end{aligned} \tag{7}$$

Introducing functions

$$\xi_A = A^{(1)} + iBA^{(2)}, \quad \xi_D = D^{(1)} + iBD^{(2)}, \tag{8}$$

and dimentionless velocity $u_i = v_i\sqrt{m_e/2kT}$, we obtain from the (4) the equations for $\xi_A$ and $\xi_D$

$$\begin{aligned} f_0(1-f_0)(u^2 - \frac{5G_{5/2}}{2G_{3/2}})u_i &= -iBf_0(1-f_0)\frac{e\xi_A}{m_ec}u_i + I_{eN}(\xi_A u_{Ni}), \\ f_0(1-f_0)u_i &= -iBf_0(1-f_0)\frac{e\xi_D}{m_ec}u_i + I_{eN}(\xi_D u_{Ni}), \end{aligned} \tag{9}$$

where

$$I_{eN}(\xi u_{Ni}) = \int f_0 f_{N0}(1-f_0^{'})(\xi u_i - \xi^{'}u_i^{'})g_{eN}bdbd\varepsilon dc_{Ni}. \tag{10}$$

Detailed description of (9) solving can be obtained from the [4] and[5]. Here we introduce computed kinetic coefficients for three tensors and show the comparison of our results with those, obtained in the subsequent papers.

General expressions for the heat flux $q_i$, and the average directional (diffusion) electron velocity $\langle v_i \rangle$ are given by

$$\begin{aligned} q_i &= -\lambda_{ij}\frac{\partial T}{\partial x_j} - n_e\frac{G_{5/2}}{G_{3/2}}\nu_{ij}d_j = q_i^{(A)} + q_i^{(D)} \\ \langle v_i \rangle &= -\mu_{ij}\frac{\partial T}{\partial x_j} - n_e\frac{G_{5/2}}{G_{3/2}}\eta_{ij}d_j = \langle v_i^{(A)} \rangle + \langle v_i^{(D)} \rangle, \end{aligned} \tag{11}$$

The indices (A) and (D) correspond to the heat flux and diffusion velocity of electrons determined by the temperature gradient $\partial T/\partial x$ and diffusion vector $d_j$.

$$\begin{aligned} q_i = q_i^{(A)} + q_i^{(D)} = &-\left(\lambda^{(1)}\delta_{ij} + \lambda^{(2)}\varepsilon_{ijk}B_k + \lambda^{(3)}B_iB_j\right)\frac{\partial T}{\partial x_j} \\ &-n_e\frac{G_{5/2}}{G_{3/2}}\left(\nu^{(1)}\delta_{ij} + \nu^{(2)}\varepsilon_{ijk}B_k + \nu^{(3)}B_iB_j\right)d_j, \end{aligned} \tag{12}$$

$$\langle v_i \rangle = \langle v_i^{(A)} \rangle + \langle v_i^{(D)} \rangle = -n_e \frac{G_{5/2}}{G_{3/2}} \left( \eta^{(1)} \delta_{ij} + \eta^{(2)} \varepsilon_{ijk} B_k + \eta^{(3)} B_i B_j \right) d_j - \left( \mu^{(1)} \delta_{ij} + \mu^{(2)} \varepsilon_{ijk} B_k + \mu^{(3)} B_i B_j \right) \frac{\partial T}{\partial x_j}. \quad (13)$$

where $\lambda_{ij}$ $\nu_{ij}$ are tensors for heat conductivity and diffusion thermoeffect, respectively, while $\mu_{ij}$ and $\eta_{ij}$ are the thermal diffusion and diffusion tensors.[11,13].

The frequency of electron-ion collisions $\nu_{ei}$ was expressed in[14] and in the limiting case for degenerate matter can be expressed as

$$\nu_{ei} = \frac{32\pi^2}{3} m_e \frac{Z^2 \mathrm{e}^4 \Lambda n_N}{h^3 n_\mathrm{e}} \qquad (D) \qquad \tau_d = 1/\nu_d. \quad (14)$$

Using expression for $\tau_d$, the components for thermal diffusion tensor, diffusion tensor and diffusion thermoeffect tensor can be written in this form:

$$\mu^{(1)} = \frac{4\pi^3 k^2 T (3n_e)^{1/3} \tau_d}{3 n_e h^2 \pi^{1/3}} \left[ \frac{1}{1+\omega^2\tau_d^2} - 2\frac{\omega^2\tau_d^2}{(1+\omega^2\tau_d^2)^2} - \frac{\pi^2}{6} \left( \frac{1}{1+\omega^2\tau_d^2 \frac{y^3}{x_0^3}} \right)'' |_{y=x_0} \right] \quad (15)$$

$$\mu^{(2)} = -\frac{4\pi^3 k^2 T (3n_e)^{1/3}}{3 n_e h^2 \pi^{1/3}} \frac{\omega\tau_d^2}{B} \left[ \frac{2}{1+\omega^2\tau_d^2} - 2\frac{\omega^2\tau_d^2}{(1+\omega^2\tau_d^2)^2} - \frac{\pi^2}{6} \left( \frac{1}{1+\omega^2\tau_d^2 \frac{y^3}{x_0^3}} \right)'' |_{y=x_0} \right] \quad (16)$$

$$B^2 \mu^{(3)} = \mu^{(1)}(B=0) - \mu^{(1)}. \quad (17)$$

$$\eta^{(1)} = \frac{kT}{n_e m_e} \tau_d \left( \frac{1}{1+\omega^2\tau_d^2} + \frac{\pi^2}{6} \left( \frac{1}{1+\omega^2\tau_d^2 (y^3/x_0^3)} \right)'' |_{x=x_0} \right), \quad (18)$$

$$\eta^{(2)} = -\frac{kT}{n_e m_e} \frac{\omega\tau_d^2}{B} \left( \frac{1}{1+\omega^2\tau_d^2} + \frac{\pi^2}{6} \left( \frac{1}{1+\omega^2\tau_d^2 (y^3/x_0^3)} \right)'' |_{x=x_0} \right), \quad (19)$$

$$B^2 \eta^{(3)} = \eta^{(1)}(B=0) - \eta^{(1)}. \quad (20)$$

$$\nu^{(1)} = \frac{kTh^2}{8m_e^2} \left( \frac{3n_e}{\pi} \right)^{2/3} \tau_d \left( \frac{1}{1+\omega^2\tau_d^2} + \frac{\pi^2}{6} \left( \frac{1}{1+\omega^2\tau_d^2 (y^3/x_0^3)} \right)'' |_{x=x_0} \right), \quad (21)$$

$$\nu^{(2)} = -\frac{kTh^2}{8m_e^2} \left( \frac{3n_e}{\pi} \right)^{2/3} \frac{\omega\tau_d^2}{B} \left( \frac{1}{1+\omega^2\tau_d^2} + \frac{\pi^2}{6} \left( \frac{1}{1+\omega^2\tau_d^2 (y^3/x_0^3)} \right)'' |_{x=x_0} \right), \quad (22)$$

$$B^2\nu^{(3)} = \nu^{(1)}(B=0) - \nu^{(1)}. \tag{23}$$

Here $x_0 = \mu/kT$ is rate of degeneracy, $\mu$ is the chemical potential of electrons , and $y = u^2$ is squared dimentionless velocity.

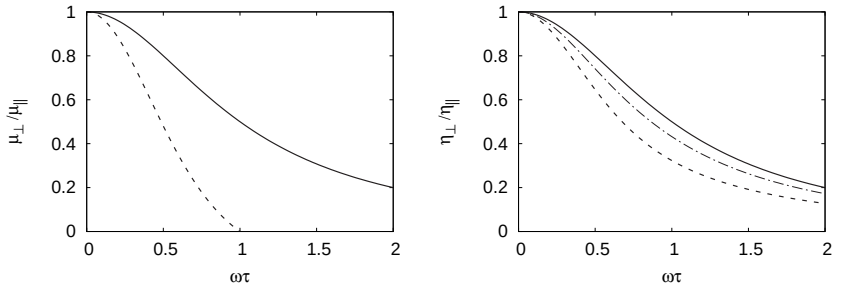



Fig. 1. The left figure is the ratio $\mu_\perp/\mu_\parallel$ as a function of $\omega\tau$. For comparison, the curves representing the phenomenologically determined thermal diffusion coefficient (solid curve) and that obtained from the asymptotic solution to the Boltzmann equation (dashed curve) are shown. The right figure is the ratio $\eta_\perp/\eta_\parallel$ as a function of $\omega\tau$. For diffusion, the phenomenologically obtained solid curve coincides with the curve obtained by solving the Boltzmann equation in the case of strong degeneracy. If terms on the order of smanless of $1/x_0^2$ are retained in the exact solution, then the solution obtained from the Boltzmann equation differs from the phenomenological one given by (1). The dash-dotted and dashed curves correspond to the solutions obtained with the allowance for small terms at $kT = 0.11E_f \quad (x_0 = 9)$ and $kT = 0.2E_f \quad (x_0 = 5)$ respectively.

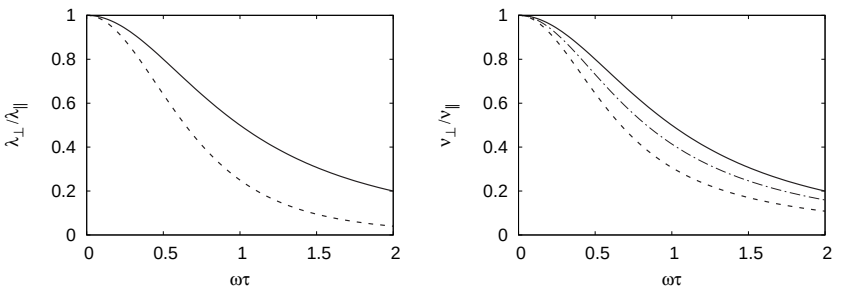



Fig. 2. The left figure is ratio $\lambda_\perp/\lambda_\parallel$ as a function of $\omega\tau$. For comparison, the curves representing the phenomenologically determined thermal conductivity (solid curve) and the asymptotic thermal conductivity obtained by solving the Boltzmann equation (dashed curve) are shown. The right figure is ratio $\nu_\perp/\nu_\parallel$ as a function of $\omega\tau$. Similar to diffusion, the phenomenologically obtained solid curve coinsides with the asymptotic solution for the diffusion thermoeffect. If terms on the order of smalness of $1/x_0^2$ are retained in the solution obtained from the Boltzmann equation, then the plot of $\nu_\perp/\nu_\parallel$ differs from the phenomenological curve. The dash-dotted and dashed curves correspond to the solutions obtained at $kT = 0.11E_f \quad (x_0 = 9)$ and $kT = 0.2E_f \quad (x_0 = 5)$, respectively.

Components for heat conductivity tensor $\lambda_{ij}$ were obtained in the work [4], and the left plot at Fig.2 of $\lambda_\perp/\lambda_\parallel$ is given here for completeness.

## 4. Discussion

In this contribution tensors of diffusion, thermal diffusion and diffusion thermoeffect are presented for strongly degenerate electrons in presence of a non-quantizing magnetic field. The kinetic coefficients are obtained asymptotically exactly in Lorentz approximation, when the electron-electron collision may be neglected in comparison with electron-nuclei collisions at nondegenerate nuclei. In most works considering kinetic coefficients in astrophysical objects, in the neutron stars in particular, following Flowers and Itoh[2], the influence of the magnetic field on the heat flux and on the average directional electron velocity was taken into account phenomenologically, using the coefficient $1/(1+\omega^2\tau^2)$, which decreases the heat flux and the diffusion in the direction perpendicular to the direction of a magnetic field. Our results, obtained by the solution of Boltzmann equation show, that the influence of the magnetic field on the coefficients of heat conductivity is stronger, and has a more complicated character. The new coefficients can be used for calculation of heat fluxes and electric current in white dwarfs, on the surface and in the crust of magnetized neutron star, as well as in the magnetized matter accreting to the magnetized neutron star.

## References


1. E. E. Salpeter, Astrophys. J., **134**, 669 (1961).
2. E. Flowers, N. Itoh, Astrophys. J., **206**, 218 (1976).
3. D. G. Yakovlev, V. A. Urpin, Astron. Zh., **57**, 526 (1980).
4. G. S. Bisnovatyi-Kogan, M. V. Glushikhina, Plasma Physics Reports, **44**, 405 (2018).
5. G. S. Bisnovatyi-Kogan, M. V. Glushikhina, Plasma Physics Reports, **44**, 971 (2018).
6. G. S. Bisnovatyi-Kogan, Transport properties of partially ionized two-temperature plasma and isotropic corrections to Maxwellian distribution function (in Russian), Diploma Thesis, Moscow Institute of Physics and Technology (Soviet Union, 1964).
7. G. S. Bisnovatyi-Kogan, J. Appl. Mech. Tech. Phys., No. 3, 43 (1964).
8. *Chapmen S., Cowling T.G.* // Mathematical Theory of Nonuniform Gases, Cambrige, (1952).
9. S. Z. Tomonaga, Z. Phys. **110**, 573 (1938).
10. *Landau L. D., Lifshitz E. M.* // Statistical Physics. **5**, 3rd ed., Butterworth-Heinemann, 1980.
11. G. S.Bisnovatyi-Kogan, M. M. Romanova, JETP, **56**, 243 (1983).
12. W. Marshall, The kinetic theory of an ionized gas, At. Energy Res Estable, N T/R, 2419 (1960).

13. *Kalikhman L. E.* // Elements of magnetohydrodynamics, Atomizdat, Moscow, 1964.
14. *Ginzburg V. A., Rukhadze A. A.* // Waves in Magnetoactive Plasma, M.: Nauka, 1970.

# Hexagonal close pack mesh for fluid dynamics

Atsushi Miyauchi

*Research Organization for Information Science and Technology (RIST),*
*1-5-2 Minatojima-minami, Kobe 650-0047, Japan*
*E-mail: miyauchi@rist.or.jp*

Although regular orthogonal mesh is usually used in numerical relativity, isotropic mesh is preferable for complex fluid motions from the points of numerical accuracy and stability. Hexagonal close pack(HCP) mesh is thought to be an ideal for those purposes. However, HCP has not been used widely so far since it is difficult to implement HCP with data structures existing programming language such as C/C++ or Python provides. In this paper, I propose novel coordinate system suitable to realize HCP with usual array data. In addition, Voronoi tessellation and mesh refinement of HCP is described briefly.

*Keywords*: Hexagonal Close Pack; Coordinate System; Mesh Refinement; Voronoi tessellation; Finite Volume Method.

## 1. Twisted Cartesian Coordinate

In case of HCP mesh, three principal axes of unit cell strongly skew. If we employ unit cell as building block of three-dimensional array, the region defined by the array spans oblong too much. As a result, it wastes large margins to contain a cubic object. Indirect access using list vectors is a typical solution. However, indirect access is poorly efficient in performance for modern processors. Then I have proposed twisted Cartesian coordinate to make HCP implemented easily and efficiently with three-dimensional arrays[1]. Keep in mind, hereafter I treat tri-layered type (i.e. Face-Centered Cubic) of HCP which is better in symmetry than dual-layered type. I found the way to make my algorithm simpler recently. Figure 1 shows a improved allocation algorithm of nearest neighbors using regular three dimensional array indexed by (i,j,k).

Definition of Greek-symbols in the figure follows. Percent symbols mean remainder after division.

$$\xi \equiv \lfloor (k\%3)/2 \rfloor = \begin{cases} 1 & \textit{if } k\%3 = 2 \\ 0 & \textit{otherwise} \end{cases} \qquad \alpha \equiv j\%2 = \begin{cases} 1 & \textit{if } j\%2 = 1 \\ 0 & \textit{otherwise} \end{cases}$$

$$\eta \equiv \lfloor ((k+1)\%3)/2 \rfloor = \begin{cases} 1 & \textit{if } k\%3 = 1 \\ 0 & \textit{otherwise} \end{cases} \qquad \beta \equiv (j+1)\%2 = \begin{cases} 1 & \textit{if } j\%2 = 0 \\ 0 & \textit{otherwise} \end{cases}$$

$$\zeta \equiv \lfloor ((k+2)\%3)/2 \rfloor = \begin{cases} 1 & \textit{if } k\%3 = 0 \\ 0 & \textit{otherwise}. \end{cases} \tag{A.1}$$

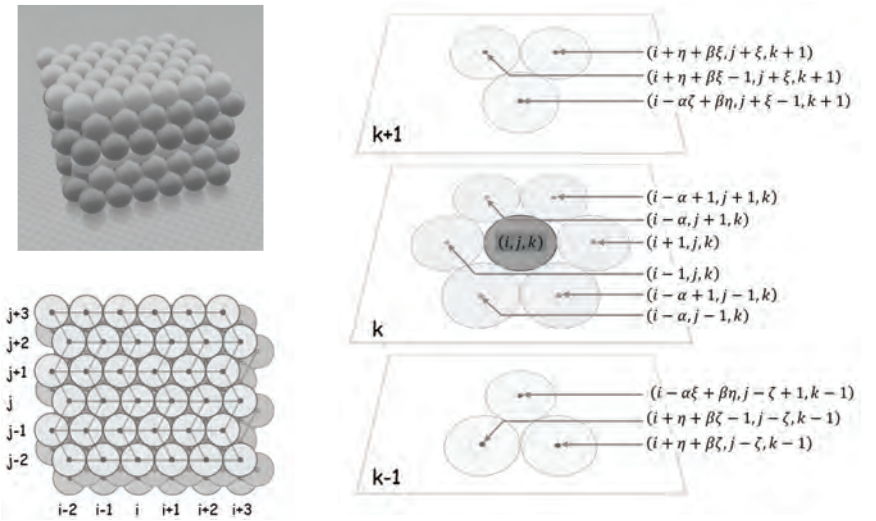



Fig. 1. Stereographic configuration (top-left), arrangement of grid points in a horizontal plane (bottom-left) and a stencil of nearest neighbors surrounding a grid point $(i, j, k)$ (right).

## 2. Voronoi Tessellation

In computational fluid dynamics, finite volume method is frequently used to calculate flow fields over complicated geometry[2]. It calculates a change of quantity in a volume by summing up flux across surrounding surfaces. Accordingly, to employ FCC for finite volume method, a partition of space without any opening is required. After straightforward but tedious calculations of Voronoi tessellation, we can find Rhombic dodecahedron serves as a such partitioned volume. You can also consult Wikipedia to be informed[3]. Although it is not a platonic solid, rhombic dodecahedron has fairly good symmetry since its surface is comprised of twelve identical rhomboids. In addition, it is noteworthy that surface-to-volume ratio is 3.67/d which is nearly half of cube, where d stands for edge length. Empirically, small surface ratio is preferable from the point of numerical stability.

## 3. Mesh Refinement

Recursive mesh refinement of HCP is possible through the steps in what follows. First, make all voxels shrunk half in length around its center, then fill generated space between voxels in each layer with additional voxels. Finally, insert reciprocal layers to fit void between existing layers. Although not shown here, current version of interpolation procedure of data from parents to children still remains partly asymmetric and left room to be improved. However, agglomeration procedure of children to parents is straightforward and symmetric.

## 4. Conclusion

A concise and easy way of identifying nearest neighbors in HCP (FCC actually) is presented which thought to be promising for precise simulation of gravitational waves through finite volume method. However, mesh refinement algorithm still needs improvement.

## References

1. A. Miyauchi, K. Iwamoto, S. N. V. Arjunan and K. Takahashi, pSpatiocyte: A Parallel Stochastic Method for Particle Reaction-Diffusion Systems, *arXiv:1605.03726* (2016).
2. C. Hirsch, *Numerical Computation of Internal and External Flows* $2^{nd}$ *ed.*, (Elsevier, 2007).
3. Rhombic dodecahedron (Jan. 11, 2019, 13:16 UTC). In *Wikipedia: The Free Encyclopedia*. Retrieved from `https://en.wikipedia.org/wiki/Rhombic_dodecahedron`.

# Magnetized 1.5-dimensional advective accretion flows around black holes

Tushar Mondal* and Banibrata Mukhopadhyay†

*Department of Physics, Indian Institute of Science, Bengaluru 560012, India*
**E-mail: mtushar@iisc.ac.in*
†*E-mail: bm@iisc.ac.in*

We address the role of large scale magnetic stress in the angular momentum transport, as well as the formation of different kinds of magnetic barrier in geometrically thick, optically thin, vertically averaged 1.5-dimensional advective accretion flows around black holes. The externally generated magnetic fields are captured by the accretion process from the environment, say, companion stars or interstellar medium. This field becomes dynamically dominant near the event horizon of a black hole due to continuous advection of the magnetic flux. In such magnetically dominated accretion flows, the accreting matter either decelerates or faces magnetic barrier in vicinity of the black hole depending on the magnetic field geometry. We find that the accumulated strong poloidal fields along with certain toroidal field geometry help in the formation of magnetic barrier which may knock the matter to infinity. When matter is trying to go back to infinity after getting knocked out by the barrier, in some cases it is prevented being escaped due to cumulative action of strong magnetic tension and gravity, and hence another magnetic barrier. We suggest, this kind of flow may be responsible for the formation of episodic jets in which magnetic field can lock the matter in between these two barriers. We also find that for the toroidally dominated disc, the accreting matter rotates very fast and decelerates towards the central black hole.

*Keywords*: accretion, accretion discs – black hole physics – MHD (magnetohydrodynamics) – gravitation – X-rays: binaries – galaxies: jets

## 1. Introduction

In 1972, Bekenstein proposed an idealized engine, namely Geroch-Bekenstein engine[1], that makes use of the extreme gravitational potential of a black hole (BH) to convert mass to energy with almost perfect efficiency. However, the practical realization of this engine is very difficult in astrophysical systems. Generally astrophysical BHs do convert mass to energy via accretion process with modest efficiencies. Later the idea came to address the importance of large scale dipole magnetic field in an accretion flow[2]. Following this idea, it was suggested and also verified numerically[3] that this efficient conversion is practically possible in the presence of large scale poloidal magnetic field. Such an efficient accretion phenomenon is named as Magnetically Arrested Disc (MAD)[4].

The origin of strong magnetic field near the event horizon of a BH is as follows. For the case of advective accretion flow, the magnetic fields are captured from accreting medium or companion star. These fields are dragged inward with continuous accretion process and become dynamically dominant through flux freezing in vicinity of a BH. Theoretical models suggest the importance of large-scale magnetic field in accretion and in the formation of strong outflows/jets, as well as high-energy radiation[5–7]. Observations of all well-known jetted sources also indicate

the presence of dynamically important magnetic field at the jet-footprint[8]. In this proceeding, we explore the underlying role of such large-scale strong magnetic fields on the disc flow behaviours for optically thin, geometrically thick, advective accretion flows around BHs. Unlike MAD, here the advection of both poloidal and toroidal magnetic fields are considered. We address the possible origin of different kinds of magnetic barriers depending on field geometry.

## 2. Model equations

We address magnetized, optically thin, advective, axisymmetric, vertically averaged, steady-state accretion flow around BHs in the pseudo-Newtonian framework with Mukhopadhyay potential[9]. The flow parameters, namely, radial velocity ($v$), specific angular momentum ($\lambda$), fluid pressure ($p$), mass density ($\rho$), the radial ($B_r$) and toroidal ($B_\phi$) components of magnetic fields, are solved simultaneously as functions of radial coordinate $r$. Throughout in our computation, we express the radial coordinate in units of $r_g = GM_{BH}/c^2$, where $G$ is Newton's gravitational constant, $M_{BH}$ is the mass of the black hole, and $c$ is speed of light. We also express $\lambda$ in units of $GM_{BH}/c$ and velocity in units of $c$, and the other variables accordingly to make all the variables dimensionless. Hence, the continuity equation, the components for momentum balance equation, and the energy equation can read as, respectively,

$$\frac{d}{dr}\left(r\rho h v\right) = 0, \tag{1}$$

$$v\frac{dv}{dr} - \frac{\lambda^2}{r^3} + \frac{1}{\rho h}\frac{d(hp)}{dr} + F = -\frac{\sqrt{h}B_\phi}{4\pi\rho h r}\frac{d}{dr}\left(r\sqrt{h}B_\phi\right), \tag{2}$$

$$v\frac{d\lambda}{dr} = \frac{1}{r\rho h}\frac{d}{dr}\left(r^2 W_{r\phi} h\right) + \frac{\sqrt{h}B_r}{4\pi\rho h}\frac{d}{dr}\left(r\sqrt{h}B_\phi\right), \tag{3}$$

$$\frac{hv}{\Gamma_3 - 1}\left(\frac{dp}{dr} - \frac{\Gamma_1 p}{\rho}\frac{d\rho}{dr}\right) = Q^+ - Q^- = f_m Q^+ = f_m(Q^+_{vis} + Q^+_{mag}), \tag{4}$$

where $F$ is the magnitude of the gravitational force corresponding to the pseudo-Newtonian potential. $W_{r\phi}$ is the viscous shearing stress, which can be expressed using Shakura-Sunayev[10] $\alpha$-viscosity prescription with appropriate modifications[11] due to advection as given by $W_{r\phi} = \alpha(p + \rho v^2)$. Following the vertical momentum balance equation, the disc half-thickness can be written as

$$h = r^{1/2}F^{-1/2}\sqrt{\left(p + \frac{B^2}{8\pi}\right)/\rho}. \tag{5}$$

The energy equation is written by taking care of the proper balance of heating, cooling and advection. $Q^+$ is the energy generated per unit area due to magnetic ($Q^+_{mag}$) and viscous dissipation ($Q^+_{vis}$), whereas $Q^-$ infers the energy radiated out per unit area through different cooling mechanisms[7]. However for the present purpose, we do not incorporate cooling processes explicitly. The factor $f_m$ measures the degree of cooling and it varies from 0 to 1 for two extreme cases of efficient cooling and no cooling respectively. The details of dissipation terms are followed

from the Ref. 12. Note that in the energy equation, we do not include the heat generated and absorbed due to nuclear reactions[13].

The presence of magnetic field provides two other fundamental equations, namely, the equation for no magnetic monopole and the induction equation. These are respectively

$$\nabla.\mathbf{B} = 0, \text{ and } \nabla \times (\mathbf{v} \times \mathbf{B}) + \nu_m \nabla^2 \mathbf{B} = 0, \tag{6}$$

where $\mathbf{v}$ and $\mathbf{B}$ are the velocity and magnetic field vectors respectively and $\nu_m$ is the magnetic diffusivity. For this accretion disc solution, we consider the induction equation in the limit of very large magnetic Reynolds number ($\propto 1/\nu_m$).

## 3. Solution procedure

The set of six coupled differential equations $(1) - (6)$ for six flow variables $v$, $\lambda$, $p$, $\rho$, $B_r$, and $B_\phi$ are solved simultaneously to obtain the solutions as functions of the independent variable $r$. The appropriate boundary conditions are as follows. Very far away from the BH, matter is sub-sonic and the transition radius between the Keplerian to sub-Keplerian flow is the outer boundary of our solutions. Very near the BH, matter is super-sonic and the event horizon where matter velocity reaches speed of light is the inner boundary. In between these two boundaries, matter becomes tran-sonic where matter velocity is equal to (or similar to) medium sound speed and this location is called as sonic/critical point. We use this point as one of the boundary. Different types of sonic/critical points, say, saddle, nodal, and spiral, are described in the Ref. 12.

## 4. Results

The angular momentum transport in accretion physics had been a long standing issue until Balbuas & Hawley[14] described the importance of magnetorotational instability (MRI), particularly in an ionized medium in the presence of weak magnetic fields. Later Mukhopadhyay & Chatterjee[15] showed that the efficient transport of angular momentum is also possible by large-scale magnetic field in geometrically thick, advective accretion flow, even in the complete absence of $\alpha$-viscosity. Here, we address the importance of large-scale strong magnetic field in accretion geometry. When the magnetic field is strong enough, presumably that corresponds to the upper limit to the amount of magnetic flux which a disc around a BH can sustain, it disrupts the accretion flow. Then the force associated with the magnetic stress becomes comparable to the strong gravitational force of the BH. As a consequence, the magnetic field arrests the infalling matter in vicinity of the BH. Here we plan to understand such magnetic activity depending on different kinds of field geometry.

In Figure 1 we show three different natures of accretion depending on different magnetic field geometries. The left column indicates the Mach number ($M$), defined as the ratio of the radial velocity to the medium sound speed, whereas the right

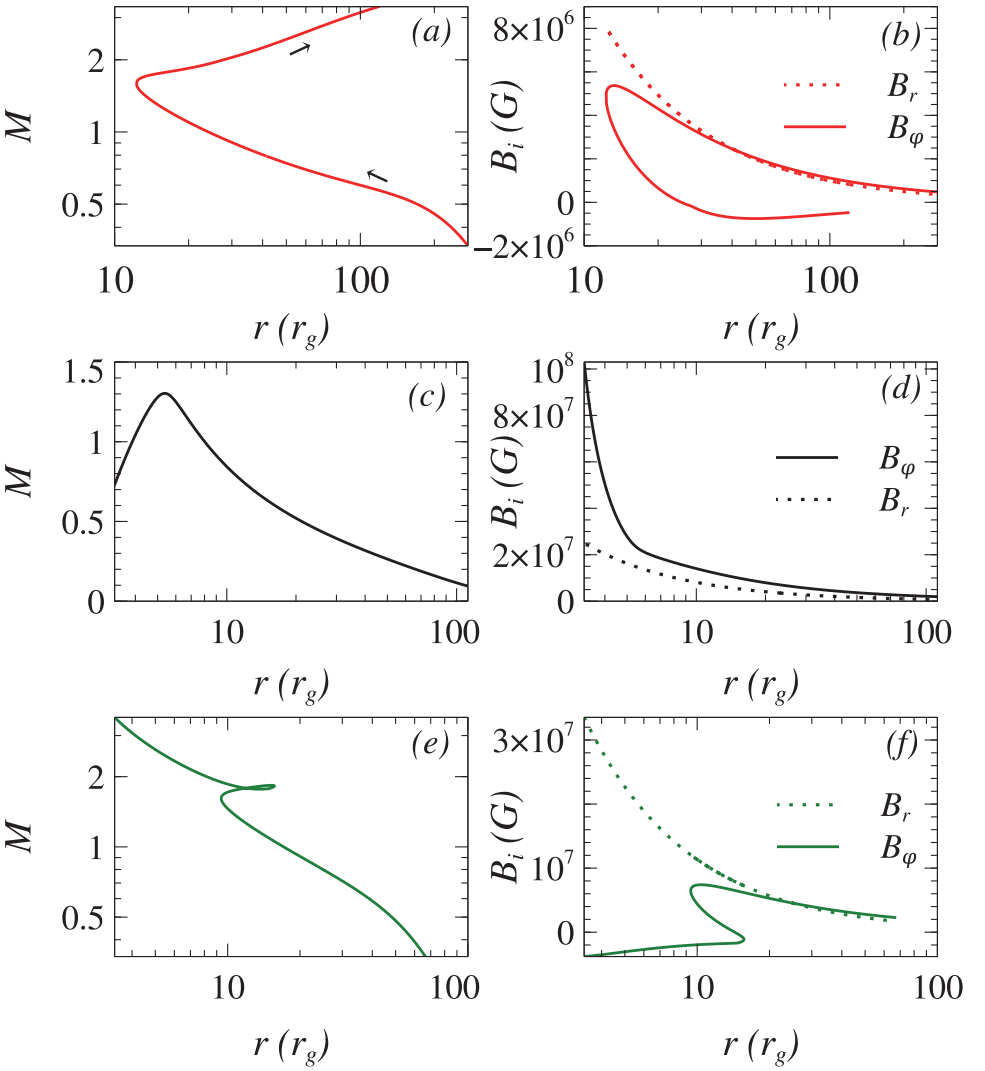



Fig. 1. Variation of Mach number ($M$) and the corresponding magnetic field components ($B_i$) for different relative field strengths in the disc. The model parameters are $M_{\rm BH} = 10M_\odot$, $\dot{M} = 0.01\dot{M}_{\rm Edd}$ and $f_m = 0.5$.

column represents the corresponding magnetic field components ($B_i$). In Figure 1 ($a$), the Mach number increases monotonically towards the central BH until it faces magnetic barrier at $r \simeq 12$. After facing such a barrier, matter goes back to infinity. To explain the origin of such barrier, we focus on the underlying role of magnetic force along with strong gravity, centrifugal force, and force due to gas pressure. When matter falls towards central BH, the forces supporting gravity may suppress by the forces acting outward. In such circumstances, following radial momentum

balance equation, the condition for which barrier appears as

$$-\frac{1}{4\pi\rho h}\left(\sqrt{h}B_\phi\frac{d(\sqrt{h}B_\phi)}{dr}\right)-\frac{1}{\rho h}\frac{d(hp)}{dr}+\frac{\lambda^2}{r^3}>F+\frac{1}{4\pi\rho h}\frac{(\sqrt{h}B_\phi)^2}{r}, \qquad (7)$$

where the first term appears from magnetic pressure and the last term from magnetic tension. Simply it suggests the essential condition to appear magnetic barrier is $\frac{d}{dr}(\sqrt{h}B_\phi)\ll-\frac{\sqrt{h}B_\phi}{r}$. On the other hand, the equation for no magnetic monopole leads to $\frac{d}{dr}(\sqrt{h}B_r)=-\frac{\sqrt{h}B_r}{r}$. These combine conditions suggest that the matter faces magnetic barrier only when the field strength is strong enough to satisfy equation (7) and the disc is poloidal magnetic field dominated. Figure 1($b$) infers such field conditions.

In Figure 1($c$), the Mach number initially increases monotonically and finally slows down to the central BH. In this case, the disc is toroidal magnetic field dominated as shown in Figure 1($d$). In this situation matter is rotating very fast compared to the inward dragging. In Figure 1($e$), the Mach number increases up to $r\simeq 9.5$, where the magnetic barrier appears due to dominant poloidal magnetic field. After knocking out by such a barrier, matter tries to go away from the BH, but faces another barrier at $r\simeq 15.6$. This is because of the cumulative action of the inward strong gravity force and the inward force due to magnetic tensions along the field lines. In between these two barriers matter faces spiral-type critical point and again falls back to the central BH. The corresponding field components are shown in Figure 1($f$). Note that the magnetic field strength within plunging region is few factor times $10^7$ G for $10M_\odot$ BH for all the cases mentioned above.

## 5. Discussions

Accretion disc can carry small as well as large-scale strong magnetic fields. Small-scale field may generate, locally, through some physical scenario, Biermann batery mechanism, dynamo process etc. However, the origin of large-scale strong magnetic fields is still not well understood. It was suggested that the externally generated magnetic field either from companion star or from interstellar medium may capture and drag inward through continuous accretion process. This field becomes dynamically dominant in vicinity of a BH due to flux freezing. In such circumstances, strong magnetic fields arrest the infalling matter and disrupt the conventional accretion flow properties.

In this proceeding, we address three different types of accretion flow properties depending on field geometry in the presence of large scale strong magnetic fields. First, when the disc is poloidal field dominated, the infalling matter faces magnetic barrier and tries to go back to infinity. On its way away from the BH, matter may completely lose its angular momentum and at the same time the combined effects of inward strong gravity and the magnetic tension along field lines can prevent the matter being escaped. In this particular case, matter again falls back to the BH by facing second magnetic barrier. The important application of such a flow

behaviour is the formation of episodic jets in which magnetic field can lock the infalling matter in between these two magnetic barriers. Second, for such a poloidal field dominated case, matter can completely go back to infinity after knocked out by the first magnetic barrier. Magnetic tension is not strong enough to prevent the matter being escaped in this case. This type of flows is the building block to produce unbound matter and hence strong continuous outflows/jets. Third, when the disc is toroidal field dominated, the infalling matter rotates very fast rather than being dragged inward in vicinity of the BH. In such case, matter falls slowly into the BH. For all these scenarios the maximum magnetic field strength near the stellar mass BH is few factor times $10^7$ G, which is well below the Eddington magnetic field limit[6] and hence perfectly viable. The other important aspect of this quasi-spherical advective flow is the angular momentum transport. The outward transport of angular momentum occurs through large-scale magnetic stress. The specific angular momentum is quite below the local Keplerian value.

In the other proceedings of this volume, we discuss more general disc-outflow symbiotic model[16] and its observational implication to ultra-luminous X-ray sources[17].

## Acknowledgments

TM acknowledges P. V. Lakshminarayana travel grant ODAA/INT/18/11, Office of Development and Alumni Affairs, Indian Institute of Science, Bengaluru, India.

## References

1. J. D. Bekenstein, *Nuovo Cimento, Lett.* **4**, 773 (1972).
2. G. S. Bisnovatyi-Kogan, A. A. Ruzmaikin, *Ap&SS* **28**, 45 (1974).
3. A. Tchekhovskoy, R. Narayan, J. C. McKinney, *MNRAS* **418**, L79 (2011).
4. R. Narayan, I. V. Igumenshchev, M. A. Abramowicz, *PASJ* **55**, L69 (2003).
5. D. L. Meier, S. Koide, Y. Uchida, *Science* **291**, 84 (2001).
6. T. Mondal, B. Mukhopadhyay, *MNRAS* **482**, L24 (2019).
7. T. Mondal, B. Mukhopadhyay, *MNRAS* **486**, 3465 (2019).
8. M. Zamaninasab, et al., *Nature* **510**, 126 (2014).
9. B. Mukhopadhyay, *ApJ* **581**, 427 (2002).
10. N. Shakura, R. Sunyaev, *A&A* **24**, 337 (1973).
11. S. K. Chakrabarti, *ApJ* **464**, 664 (1996).
12. T. Mondal, B. Mukhopadhyay, *MNRAS* **476**, 2396 (2018).
13. S. R. Datta, B. Mukhopadhyay, *MNRAS* **486**, 1641 (2019).
14. S. A. Balbus, J. F. Hawley, *Rev. Mod. Phys.* **70**, 1 (1998).
15. B. Mukhopadhyay, K. Chatterjee, *ApJ* **807**, 43 (2015).
16. T. Mondal, B. Mukhopadhyay, *this volume.*
17. B. Mukhopadhyay, *this volume.*

# Simulation of the dense protostellar cores formation in the collision of molecular clouds

Boris P. Rybakin[1] and Sergey G. Moiseenko[2]

[1] *SRISA RAS, MSU M.V.Lomonosov*
*Moscow, Russia*
*E-mail: rybakin@vip.niisi.ru*
[2] *Space Research Institute, Profsoyuznaya str. 84/32, Moscow 117997, Russia*
*E-mail: moiseenko@iki.rssi.ru*

We represent results of simulation protostellar molecular clouds(MC) collision process. The interaction of the MC with each other leads to the emergence of extremely dense, gravitationally bounded regions, which can be sources of the formation of new stars and star systems. Computer modeling was carried out on high-resolution grids up to $2048 \times 1024 \times 1024$. Oscillations of the MC surface, which are caused by the formation of a dense lenticular structure at the interface of the colliding clouds in the process of collision, are found.

*Keywords*: Molecular clouds collision; Numerical simulation; Kelvin-Helmholtz instability.

## 1. Introduction

Collisions of molecular clouds and their interaction with shock waves play a large role in the chain of events leading to the formation of stars.

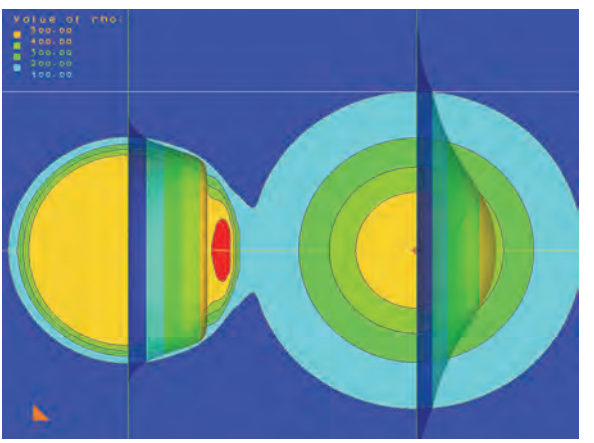



Fig. 1. Density dictribution in the collision of two MC

When molecular clouds collide or when shock waves pass through them, the gas density in MC increases by many orders of magnitude. This can lead to fragmentation and restructuring of clouds. As a result of such processes, a system of dense filaments and proto-cores is formed. At relatively small scales (about 1 parsec), star

formation processes occur in numerous dense nuclei in the filaments. The resulting turbulent motions become the dominant factor in the formation of high-density gas regions, some of which may become protostar. In Fig. 1 shows an example of the collision of two molecular clouds, which is a trigger for the formation of new stars.[1].

This article presents the results of numerical modeling of the processes of collision MO. Turbulent motions that occur during a collision are the dominant cause of the formation of areas with high density, some of which can become a protostar.

## 2. Formulation of the problem

Modeling is carried out with using high-resolution difference schemes, on sufficiently detailed grids. The paper presents the processes of formation, evolution and interaction between each other filament structures, the emergence of supersonic turbulence, etc.As known[2,3] molecular clouds are often gravitationally bound objects. The appearance of the Jeans instability and the formation of dense proto-cores is often the result of such a gravitational coupling. Further gravitational contraction of the proto-core can lead to the emergence of new stars.

The interstellar medium(ISM) consists of a relatively warm substance $\sim 10^4 K$. The ISM contains molecular clouds whose densities are much higher than the density of the ISM. The temperature in the molecular cloud is less than the ISM $T_{cloud} \sim 10^2 K$ temperature. The shock wave formed after the supernova explosion hits a molecular cloud that is in the interstellar medium. Fig. 1 schematically shows the geometry of the computational domain. The cross-section of the three-dimensional computational domain by a plane perpendicular to the axis OZ is shown. In Fig. 1 shows the initial (t = 0) position of two molecular clouds in a central collision. In the sections shown, the initial density distribution in the clouds is shown.

The computational domain is a parallelepiped with sizes ranging from $2048 \times 1024 \times 1024$ to $1024 \times 512 \times 512$ cells along the axes $X$, $Y$ and $Z$, respectively. In the above calculations, the cell size along the axes $\triangle\, x, \triangle\, y, \triangle\, z$ was assumed to be the same: $\triangle\, x = \triangle\, y = \triangle\, z$. Thus, each environment has its own sound speed equal to $a_i = \sqrt{\gamma_i \frac{p_i}{rho_i}}$, $i = 0.1$, ($i = 0$ for ISM, $i = 1$ for MC)

Mach numbers $M_i = u_i/a_i$ are also different for these media. However,[5,6] showed that it is possible to ignore these differences and use the same value of $\gamma = 5/3$ for the entire computational domain.

For modeling, a system of laws of conservation of mass, momentum and energy in a three-dimensional Cartesian coordinate system is used.

When modeling the process of collision of two molecular clouds or the interaction of a strong shock wave with a molecular cloud, we will use the equation of state in the form:

$$p = \rho(\gamma - 1)\Big[e - \frac{1}{2}(u^2 + v^2 + w^2)\Big], \tag{1}$$

here $\gamma$ is the adiabatic index, $\rho$ density, $\mathbf{v} = \{u, v, w\}$ velocity vector and $p$ pressure[4].

Initially, the cloud is in dynamic equilibrium with the background gas. Thermal conductivity and radiation losses in the interaction can be neglected. In the work $\gamma = 5/3$. The density of the interstellar medium is assumed to be $rho_a = 2.15 * 10^{-25}\ g/sm^3$, the temperature is $T_a = 10^4 K, u_a = 0.0$. The cloud density is $\rho_c = 1.075 * 10^{-22} g/sm^3$, temperature is $T_c = 100K$, $u_c = 0.0$. The parameters of the gas behind the shock wave are determined from the Rankin Hugoniot equations.

One of the most important parameters of the process of interaction of the shock wave with the MC is the ratio of the density of MC to the ISM density:

$$\chi = \frac{\rho_{mc}}{\rho_{ism}} \tag{2}$$

Here $\rho_{mc}$ is the density of the molecular cloud, $\rho_{ism}$ is the density of the interstellar medium, respectively. We assume that the shock wave does not undergo radiation cooling and the Mach number of the shock wave is sufficiently large.

## 3. Results

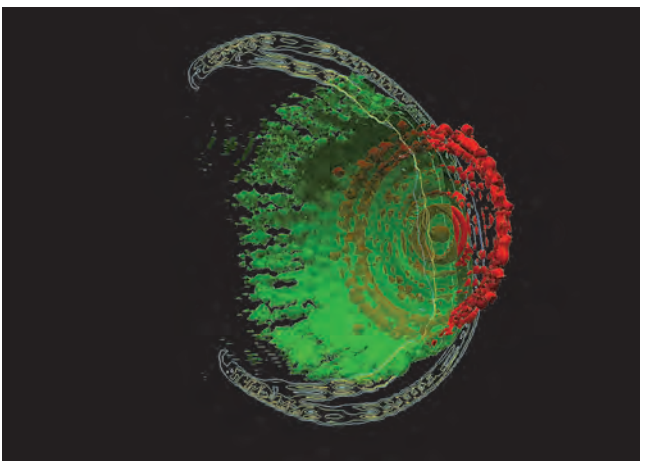


Fig. 2. The stage of passage by the cloud of MC1 zone at the exit from the area opposite to the MC2. Given the distribution of fragmented condensation of gas, with a contrast of 8000.

The results of the central collision of two MOs at a speed of 0.5 to 15 km/s are given. Under such conditions, there is a significant fragmentation of clouds, the spread of gas from the center to the periphery, the formation of dense areas in the area of impact. Depending on the initial density and speed of impact, areas of increased density appear in the contact zone and are constantly increasing, which over time can become gravitationally bound objects. In Fig. 2 shows the results of molecular collision calculations at a mutual velocity of 5 km/s. The moments of

dimensionless time are given, at which a smaller MC penetrated into large one more than about half the radius of a small cloud (right fragment). The center shows the contact area, which quickly changes in shape and mass.

This process is presented in more detail in Fig. 3. The development in time of the field of a rapidly changing density gradient in the core (red) and in the shell (green) is given. The maximum density contrast, when colliding at relatively low speeds, increases by two orders of magnitude compared to the initial contrast $\chi = 500$. A change in the density, shape, and position of the contact surface leads to the appearance of oscillations that occur on the midsections (equators) of the MO and the appearance of oscillations in the interstellar medium. With further interaction, the MC of a shock-compressed, lens-like layer forms superdense regions (clumps) in which ring waves propagate.

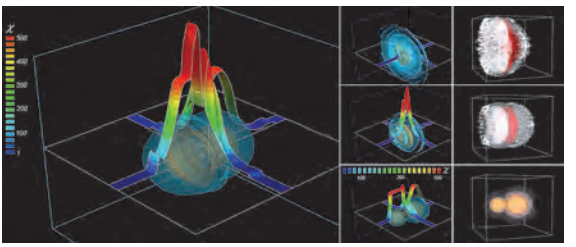


Fig. 3. On the left is a plot of the density distribution in the collision of two MOs, on the right - successive stages of the collision process (bottom-up)

In Fig. 3 shows the graphs of the density distribution at different stages of a collision (left). The density distribution in the lens-like contact area is shown in the center. On the right - the destruction of the shells of the MO and the formation of filaments at successive time moments.

## 4. Conclusions

Numerical simulation of collisions of initially spherical clouds in a hot interstellar medium (MS) revealed significant details of the increase in density in the boundary regions between the molecular clouds and the MC substance. This process is accompanied by the Kelvin-Helmholtz instability and disturbance of the gas density above the disturbed surface layers of the clouds. Oscillations caused by extremely strong density changes at the contact boundary of the MC become clearly observable. It is very likely that this process is generated through the exchange of energy in high-gradient outer layers and deformation of the lens-like surface at the contact boundary. The nonlinear instability of thin shells in the gas phase can play a crucial role in the launch of this process. The Kelvin-Helmholtz instability apparently, lead to the formation of surface oscillations and deformation of the MC.

This work was supported by the Program of the Presidium of the Academy of Sciences of Russia, No. 26.

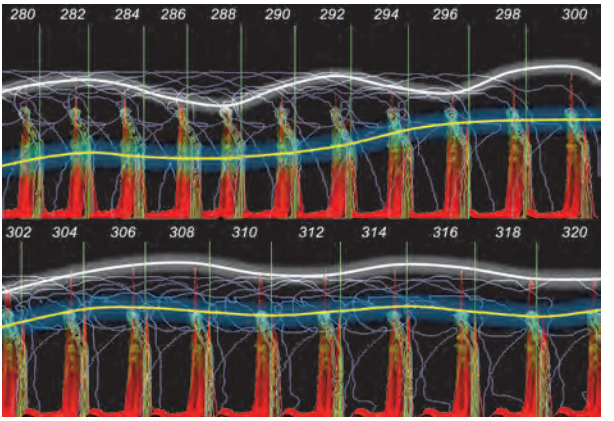



Fig. 4. Illustration of pulsation changes in the core for the contrast ratio variant $\chi = 500/100$ with U1 = 0.5 U2 = 0.25

## References

1. Truelove J. K., Klein R. I., McKee C. F., Holliman J. H., Howell L. H., Greenough J. A., Woods D. T. Self-gravitational hydrodynamics with 3-D adaptive mesh refinement: methodology and applications to molecular cloud collapse and fragmentation. The Astrophysical Journal, 495, 821 (1998).
2. Hartquist, T.W., Dyson, J.E. 1988, Ap&SS, 144, 615.
3. Hartquist, T.W., Dyson, J.E., Williams, R.J.R., 1997, ApJ, 482, 182.
4. Rybakin B. Modeling of III-D Problems of Gas Dynamics on Multiprocessing Computers and GPU // Elsevier, Computers & Fluids. DOI information: 10.1016/j.compfluid.2012.01.016, 31-JAN-2012.
5. Picone, J.M., Boris, J.P.: Vorticity generation by shock propagation through bubbles in a gas. J. Fluid Mech. 189, 23–51 (1988).
6. Bagabir, A., Drikakis, D.: Mach number effects on shock–bubble interaction. Shock Waves 11, 209–218 (2001).
7. Rybakin, B.P., Shider .. Parallel algorythm construction for the solutuon of the problems of gravitational gas dynamics// Computational methods and Programming. 2010, **11**. 388-394.
8. B. Rybakin, V. Goryachev Modeling of density stratification and filamentous structure formation in molecular clouds after shock wave collision. Computers and Fluids, Pergamon Press Ltd. (United Kingdom), 2018, v. 173, pp. 189-194.

# The bow shock pulsar wind nebulae propagating through a non-uniform ISM

O. D. Toropina*

*Space Research Institute, Russian Academy of Sciences*
*Profsoyuznaya 84/32, Moscow 117997, Russia*
**E-mail: toropina@iki.rssi.ru*

M. M. Romanova and R. V. E. Lovelace

*Department of Astronomy, Cornell University,*
*NY, 14853, Ithaca, USA*

We model the propagation of pulsars through the inhomogeneous ISM using non-relativistic axisymmetric magneto-hydrodynamic (MHD) simulations. We take into account the wind from the star, which carries predominantly azimuthal magnetic field, and investigate the PWN at different levels of magnetization (the ratio of magnetic to matter energy-densities) in the wind. We consider the interaction of PWN with large-scale and small-scale imhomogeneities in the ISM at different values of magnetization. We conclude that the inhomogeneities in the ISM can change the shapes of the bow shocks and magnetotails at different values of the magnetization.

*Keywords*: neutron stars; magnetic field; PWN; ISM; MHD.

## 1. Introduction

Pulsars emit winds of relativistic particles and magnetic fields, and are often surrounded by the PWN. Many pulsars have high velocities and propagate supersonically through the ISM, and their PWN interact with the ISM, forming bow shocks and magnetotails. The bow shocks are often observed in the $H_\alpha$ spectral line[3]. Many interesting structures (bow shocks, very long tails, and jet-like features) are observed in the X-ray[6] and radio[9] wavebands.

One remarkable PWN is the Guitar Nebula, which is powered by the pulsar PSR B2224+65 that travels at a high velocity of about 1600 km/sec. The high-resolution observations in the $H_\alpha$ line show that the shape of the Nebula's head becomes wider with time. The variation in shape may be connected with the variation in the density of the ISM[5]. Figure 1 shows another interesting example of PWN, the $H_\alpha$ pulsar bow shock connected with the pulsar PSR J0742-2822[3]. This PWN shows multiple irregularities in its shape, which suggests that the pulsar may be travelling through small-scale fluctuations in the ISM.

In the X-ray band, many of the PWN show irregularities in their shapes, as well as very long tails (see review by Kargaltsev at al.[8]). Even longer tails are observed in the radio band[9]. Figure 2 shows two examples of PWN, observed in the X-ray and radio bands: the PWN associated with the pulsar PSR J1509-5850 and the Mouse Nebula, powered by the pulsar PSR J1747-2958[7]. One can see that both PWN have long tails in the X-ray band (red colour), and even longer tails in the radio band (blue colour).

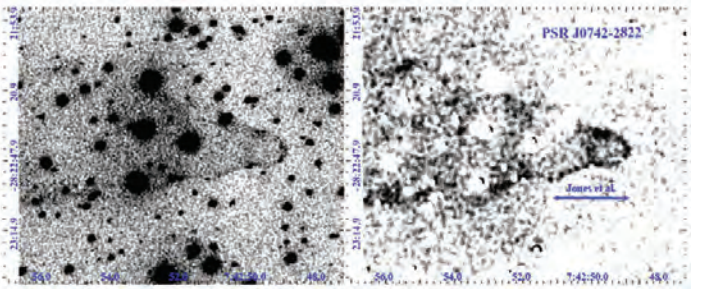



Fig. 1. PWN of the pulsar PSR J0742-2822, observed in the $H_\alpha$ line by (image taken by Brownberger & Romani[3]. *Left panel:* A median-filtered 3 x 600 W012 SOI image of PSR J0742.2822, smoothed with a 0.45" Gaussian. *Right panel:* Same image, but with the scaled continuum image subtracted and a 0.9" top-hat smoothing. The arrow indicates the extent of the previous nebula detection.

If a pulsar propagates supersonically through the ISM, then the PWN interacts with the ISM, forming a bow shock and magnetotails. In the bow shock, the energy of accelerated particles may dominate over magnetic energy-density. However, in the magnetotails, the magnetic energy-density may be comparable to or larger than the energy-density of the particles. Long, magnetically-dominated magnetotails are expected to form in the PWN. They may be visible, if the accelerated particles propagate into the magnetotails, or invisible otherwise. In modelling the supersonic PWN, it is important to take into account both, the matter and the magnetic field components of the PWN.

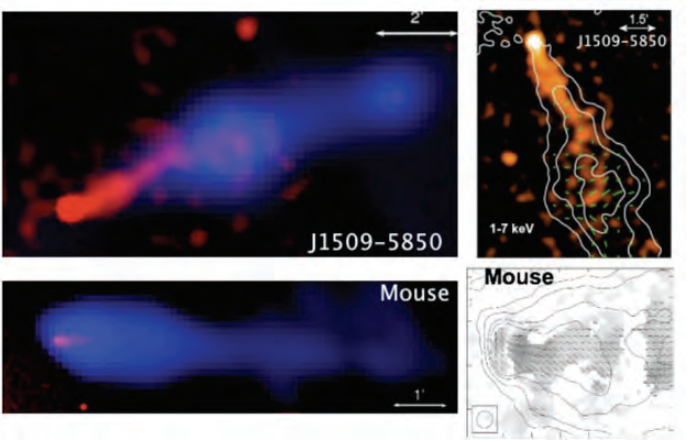



Fig. 2. X-ray and radio images of the very long pulsar tails, by Kargaltsev and Pavlov[7]. Right panels show the radio contours and the direction of the magnetic field. The red and blue colours in the left panels correspond to X-ray and radio, respectively.

Supersonic propagation of pulsars through the ISM has been studied in a number of axisymmetric non-relativistic and relativistic hydrodynamic simulations (e.g., Bucciantini[4], Barkov[2]). Many of the above simulations of the bow shocks propagating through an inhomogeneous ISM medium were performed in hydrodynamic approximation. Our goal is to study this problem in axisymmetric MHD simulations at different magnetization values of the pulsar wind.

## 2. Numerical Model

We performed MHD simulations to investigated the supersonic propagation of magnetized neutron star through the non-uniform ISM. We used an axisymmetric, resistive MHD code. The code incorporates the methods of local iterations and flux-corrected-transport[12]. The flow is described by the resistive MHD equations:

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\ \mathbf{v}) = 0\ ,$$

$$\rho\frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla p + \frac{1}{c}\mathbf{J}\times\mathbf{B} + \mathbf{F}^g\ ,$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla\times(\mathbf{v}\times\mathbf{B}) + \frac{c^2}{4\pi\sigma}\nabla^2\mathbf{B}\ ,$$

$$\frac{\partial(\rho\varepsilon)}{\partial t} + \nabla\cdot(\rho\varepsilon\mathbf{v}) = -p\nabla\cdot\mathbf{v} + \frac{\mathbf{J}^2}{\sigma}\ . \qquad (1)$$

We assume axisymmetry ($\partial/\partial\phi = 0$), but calculate all three components of $\mathbf{v}$ and $\mathbf{B}$. The equation of state for an ideal gas, $p = (\gamma-1)\rho\varepsilon$, where $\gamma = 5/3$ is the specific heat ratio and $\varepsilon$ is the specific internal energy of the gas. The equations incorporate Ohm's law $\mathbf{J} = \sigma(\mathbf{E}+\mathbf{v}\times\mathbf{B}/c)$, where $\sigma$ is the electric conductivity. The associated magnetic diffusivity, $\eta_{\rm m} \equiv c^2/(4\pi\sigma)$, is assumed to be a constant. The gravitational force, $\mathbf{F}^g = -GM\rho\mathbf{R}/R^3$, is due to the star because the self-gravity of the accreting gas is negligible.

We use a cylindrical, inertial coordinate system $(r,\phi,z)$ with the $z-$axis parallel to the star's dipole moment $\mu$ and rotation axis $\mathbf{\Omega}$. The vector potential $\mathbf{A}$ is calculated so that automatically $\nabla\cdot\mathbf{B} = 0$ at all times. The intrinsic magnetic field of the star is taken to be an aligned dipole, with vector potential $\mathbf{A} = \mu\times\mathbf{R}/R^3$. A detailed description of numerical model can be found in[10],[11].

We measure length in units of the Bondi radius $R_B \equiv GM/c_s^2$, with $c_s$ the sound speed at infinity, density in units of the density of the interstellar medium $\rho$, and magnetic field strength in units of $B_0$ which is the field at the pole of the numerical star. We measure time in units of $t_0 = (Z_{max} - Z_{min})/v_\infty$, which is the crossing time of the computational region with the star's velocity, $v_\infty$.

## 3. Propagation of PWN through the non-uniform ISM

We investigate the propagation of a PWN through an ISM with an inhomogeneous matter distribution. Observations point to large and small-scale inhomogeneities, and therefore we consider two types of imhomogeneities: (1) large-scale clouds, which are much larger than the size of the bow shock, (2) small-scale clouds, whose size is comparable with the width of the bow shock.

### 3.1. *Large-scale cloud*

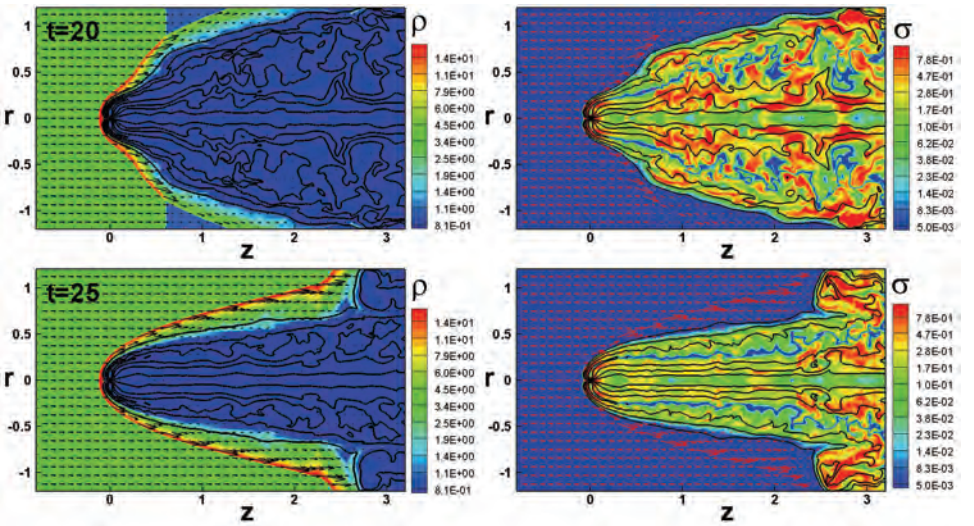



Fig. 3. *Left panels:* Propagation of a PWN of moderate magnetization through a large cloud of density $\rho_{\rm cloud} = 3\rho_0$ at two moments in time. The background represents the logarithm of density. The solid lines are magnetic field lines. *Right panels:* Same, but the background shows magnetization $\sigma$.

Fig. 3 shows an example of propagation of a large cloud of density $\rho_{\rm cloud} = 3\rho_0$ through the PWN in the model with a moderate magnetization. The left panels show that the cloud compresses the bow shock and the bow shock has a smaller opening angle (the Mach cone). The cloud is in pressure balance with the rest of the ISM, so that the sound speed in the cloud is three times lower than in the rest of the ISM. Therefore, the Mach number of the star inside the cloud is $M_{\rm new} = v_p/c_s = 3M_{\rm old} = 60$. This is why we observe a smaller opening angle of the Mach cone. The bottom left panel shows that after entering the cloud, the KH instability develops in the region of CD. This is an expected result: at larger Mach number, there is a larger difference between velocities of matter from two sides of the contact discontinuity. The right panels of Fig. 3 show that the region of high magnetization becomes more narrow after the passage of the cloud. Also,

the region of high magnetization becomes less homogeneous (more patchy) due to the action of the KH instability. These simulations demonstrate that propagation through a large-scale cloud leads to a reconstruction of the magnetic field in space and to change in the value of the magnetic field. The tangled magnetic field may reconnect and the magnetic energy can be converted to the energy of accelerated particles and radiate. This may be a possible reason for the re-brightening of PWN magnetotails observed in X-ray [8].

#### 3.1.1. *Small-scale inhomogeneities*

Our simulations show that the KH instability frequently forms at the CD and produces some wavy structure of the bow shock. However, a wavy structure could also result from the interaction of the bow shock with the small-scale inhomogeneities in the ISM. To model the interaction of the small-scale inhomogeneities with the bow shock, we took models where the KH instability is weak and does not modify the shape of the bow shock.

We model small-scale inhomogeneities as a set of small clouds with Gaussian density distribution, with the maximum density of $\rho_{\max}$ and half-width of $\Delta r = 0.2$. To keep the pressure balance between the clouds and the rest of the ISM, we take the temperature in the cloud to be $T_{\rm cloud} = T_{\rm ISM} * \rho_0/\rho_{\rm cloud}$. We experimented with clouds of different densities, and found that the shape of the bow shock varies significantly if the density in the cloud is $\rho_{\rm cloud} \gtrsim 2\rho_0$. At $\rho_{\rm cloud} = 2\rho_0$, a wavy structure starts to become visible, while at $\rho_{\rm cloud} = 3\rho_0$, the bow shock changes its local shape significantly.

Fig. 4 shows a wavy variation of density in the bow shocks that appears after the propagation of clouds of density $\rho_{\rm cloud} = 3\rho_0$ in the models with low and moderate magnetizations. Such an interaction with small-scale inhomogeneities may explain the wiggles in the shape of the bow shock observed in PSR J0742-2822 (see Fig. 1). In PSR J0742-2822, the estimated stand-off distance is $R_{\rm sd} \approx 4.2 \times 10^{16} d/2{\rm kpc}$ cm. The size of clouds in our model $R_{\rm cl} \approx (2-4)R_{\rm sd} \approx (0.8-1.6) \times 10^{17} d/2{\rm kpc}$ cm. It may correspond to one of scales in turbulent ISM matter. We should note that an alternative explanation of these small-scale features in bow shocks can be connected with the KH instability. Currently, we cannot distinguish between these two mechanisms.

## 4. Conclusions

We performed MHD simulations of the bow shock PWN propagating through a uniform and non-uniform ISM at different levels of magnetization. Our main findings are the following:

**1.** The interaction of the bow shock with a large-scale, dense cloud leads to the compression of the bow shock and the formation of a new bow shock with a smaller opening angle. The level of compression (the ratio between the original

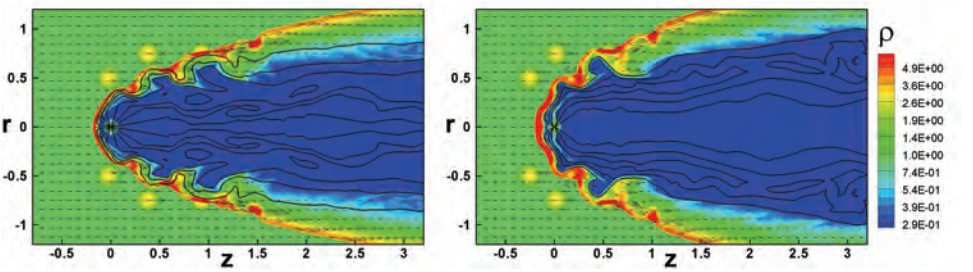



Fig. 4. Interaction of the bow shock with small-scale clouds of maximum density $\rho_{\rm cloud} = 3\rho_0$ in the model with low $\sigma$ (left panel) and in the model with medium $\sigma$ (right panel). The background represents the logarithm of density. The solid lines are magnetic field lines.

and final widths of the bow shock) is approximately the same for PWN of different magnetization.

**2.** The interaction of the bow shock with the small-scale inhomogeneities in the ISM leads to a wavy structure in the bow shock. The amplitude of the "waves" increases with the ratio $\rho_{\rm cloud}/\rho_0$. For example, variation in the shape of the bow shock of pulsar PSR J0742-2822 can be explained by the propagation through a series of clouds with a density ratio of $\rho_{\rm cloud}/\rho_0 \approx 3$. Small-scale clouds produce wavy shapes of the bow shocks in PWN of different magnetization.

## References


1. Arzoumanian, Z., Chernoff, D. F., & Cordes, J. M. 2002, ApJ, 568, 289
2. Barkov, M., & Lyutikov, M. 2018, MNRAS, eprint arXiv:1804.07327
3. Brownsberger, S., Romani, R.W., 2014, ApJ, Volume 784, Issue 2, article id. 154, 14 pp.
4. Bucciantini, N., 2002, A&A, 387, 1066
5. Chatterjee S. & Cordes J. M., 2004, ApJ, Volume 600, Issue 1, pp. L51
6. Kargaltsev, O., & Pavlov, G.G. 2010, AIPC, 1248, 25
7. Kargaltsev, O., Durant, M., Pavlov, G., Garmire, G. 2012, ApJ Supplement, Volume 201, Issue 2, article id. 37, 7 pp.
8. Kargaltsev, O., Pavlov, G.G., Klingler, N., & Rangelov, B. 2017, J. Plasma Phys., Vol. 83, pp. 1-25
9. Ng, C.-Y. et al. 2010, ApJ, 712, 596
10. Toropina, O.D., Romanova, M.M., Toropin, Yu.M.,& Lovelace, R. V. E. 2001, ApJ, 561, 964
11. Toropina, O.D., Romanova, M.M., & Lovelace, R. V. E. 2012, MNRAS, Volume 420, Issue 1, pp. 810-816
12. Zhukov, V. T., Zabrodin, A. V., & Feodoritova, O. B. 1993, Comp. Maths. Math. Phys., 33, No. 8, 1099

# MHD simulation of magnetized laboratory jets

Olga D. Toropina, Gennadiy S. Bisnovatyi-Kogan and Sergey G. Moiseenko

*Space Research Institute, Russian Academy of Sciences*
*Profsoyuznaya 84/32, Moscow 117997, Russia*
** E-mail: toropina@iki.rssi.ru*

We present the results of MHD simulation of the laboratory experiment for creating plasma jets. In the experiment on a NEODIM laser installation the plasma jet is formed as a result of the action of a powerful laser on the target. We simulated plasma flow, we chose a numerical method, boundary and initial conditions. We investigated the picture of the flow and compared it with the experiment.

*Keywords*: MHD; plasma; laboratory jets.

## 1. Introduction

Astrophysical jets are observed at various astronomical scales and are found in active galactic nuclei, quasars, compact objects in binary systems and young stellar objects. The emergence of jets is possible with the explosion of collapsing supernovae. The processes of formation and evolution of astrophysical jets are one of the most interesting problems of astrophysics[1].

The study of astrophysical jets is conducted in two ways: observations in the optical, X-ray and radio bands and multidimensional numerical MHD modeling. A simulation must take into account the entire complex nature of relativistic jets, including the gravitational field and the magnetic field of the central object. Recently, methods of laboratory astrophysics were added to the methods of investigating relativistic jets. They allow to create of plasma jets by the laser and to investigate the structure, evolution, and features of jets in the laboratory conditions.

The equations of magnetic hydrodynamics describing the formation and evolution of jets are nonlinear even in the one-dimensional case. For the theoretical description of the processes occurring in jets, it is necessary to apply methods of multidimensional numerical simulation. Multidimensional numerical calculations make it possible to obtain a detailed picture of the flow of matter in the jet and the shock-wave structure of the flow. Numerical simulation allows calculations to be performed both for initial data corresponding to astrophysical jets, and for a set of parameters from a laboratory experiment.

Laboratory research is conducted by several groups of scientists. Fua, Liang, Tzeferacos and Lamb work on the creation of a magnetized supersonic laboratory jet. They perform numerical simulation of the experiment using the two-dimensional FLASH code[2]. Experiments on laboratory jets were carried out at the Kurchatov Institute on the installation of the plasma focus PF-3 by Krauz at al.[3]. Pikuz at al.[4] presented the results of a laboratory laser experiment and the numerical simulation of a magnetized jet expiring from a young stellar object. The collimated jet in these calculations is obtained due to the presence of a poloidal magnetic field.

## 2. Experiment

The purpose of the study is numerical modeling and analysis of the results of a laboratory experiment simulating the formation of cosmic relativistic plasma jets on the 10 TW picosecond NEODIM laser installation (TsNIIMASH, Korolev, Russia)[7].

The laser unit has the following parameters of the laser pulse: energy up to 10 J, wavelength 1.055 mkm, duration 1.5 ps, laser radiation contrast about $10^7$. The focusing system based on an off-axis parabolic mirror with a focal length of 20 cm provides a concentration of at least 40% of the energy of the laser beam into a 15 $\mu m$ spot and, correspondingly, a peak intensity of $2 \times 10^{18} W/cm^2$.

The laboratory experiment consists of the following: a cylindrical chamber is placed as a target a foil of $Cu$ or $Ta$ with a thickness of 30 $\mu m$ and 50 $\mu m$. On the opposite side are track detectors CR-39. The energy $10\,J = 6.24 \times 10^{19}$ eV is uniformly and instantaneously applied to the central region of the target with a diameter of 10 m. As a result of heating the target, two directional plasma ejections appear in the form of symmetrical jets. In our numerical simulation, only one of them is considered.

In the experiment, clearly distinct ring structures formed on track detectors by protons with energies in the range of $0.8 \div 1.7$ MeV are observed. An example of images of proton beams is shown in Figure 1.

The scale of physical processes in astrophysical jets is very different from those in jets that occur when a laser pulse is applied to a target. However, the application of similarity criteria makes it possible to obtain a correspondence between a number of characteristics of an astrophysical and laboratory jet. The use of such similarity criteria is convenient both for direct comparison of laboratory and observational data, and for multidimensional numerical simulation. As a result of scaling, we get a satisfactory similarity of a number of jets parameters expiring from the young stars and the jet obtained on the NEODIM laser installation. A detailed description of the similarity criteria and scaling can be found in our recent papers[5,6].

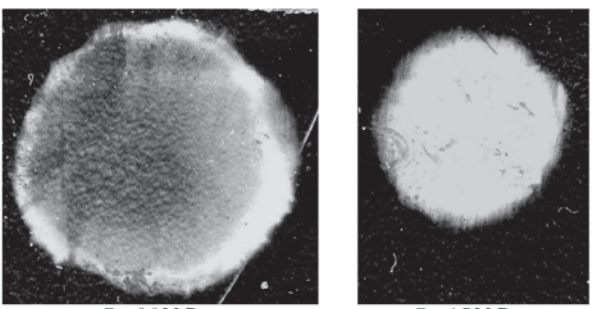



Fig. 1. Images of proton beams with different energies on track detectors CR-39 for a target of Cu thickness 50 m.

## 3. Numerical model

We did not consider the processes of laser heating, melting, and evaporation of the target material in detail during the simulation. A considerable amount of work has been devoted to the study of laser ablation processes, for example, Mazhukin et al.[8] We believe that the ablation process has already happened, the substance has evaporated from the surface of the target and is taken into account in determining the initial and boundary conditions. For numerical simulation of the experiment, the system of magnetohydrodynamics equations[9] was used. In the future, to more accurately describe the processes occurring with the target material as a result of high-power laser irradiation, it seems necessary to combine MHD and the kinetic approach. But at the initial stage of modeling, we can confine ourselves to the system of MHD equations.

We take into account an external poloidal magnetic field directed along the normal to the plate. Simulations were carried out for different values of the magnetic field. The emission from the plasma target was simulated under the influence of laser radiation. A single-component plasma flow was studied - the region and shape of the flow at various distances from the target and at different times, the distribution of plasma density, energy, and the magnetic field. The analysis of the experiment showed that the formation of the jet is associated with the heating of the plate, and the geometry of the incident laser beam is irrelevant. The model of jet formation in the experiment can be constructed in axisymmetric approximation, assuming that the spot heated by the laser has the form of a circle. An inertial cylindrical coordinate system $(r, \phi, z)$ is used. The problem assumes axial symmetry in the distribution of all the macroscopic quantities $\rho$, $T$, $v$, $H$, $\partial/\partial\phi = 0$, but all components of the velocity $v$ and magnetic fields $H$ are calculated. The target (the central region of the plate to which the energy is applied) is a cylinder of radius $R_d$, of thickness $Z_d$. The scheme of the modeling domain is shown in Figure 2.

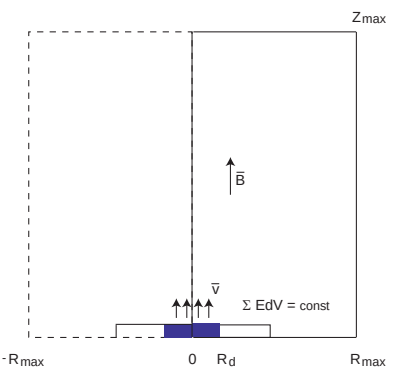



Fig. 2. Simulation region

We use an axisymmetric, resistive MHD code. The code incorporates the methods of local iterations and flux-corrected-transport developed by Zhukov, Zabrodin, Feodoritova[10]. The flow is described by the resistive MHD equations[9]:

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\ \mathbf{v}) = 0\ , \tag{1}$$

$$\rho\frac{\partial \mathbf{v}}{\partial t} + \rho(\mathbf{v}\cdot\nabla)\mathbf{v} = -\nabla p + \frac{1}{c}\mathbf{J}\times\mathbf{H} \tag{2}$$

$$\frac{\partial \mathbf{H}}{\partial t} = \nabla\times(\mathbf{v}\times\mathbf{H}) + \frac{c^2}{4\pi\sigma}\nabla^2\mathbf{H}\ , \tag{3}$$

$$\frac{\partial(\rho\varepsilon)}{\partial t} + \nabla\cdot(\rho\varepsilon\mathbf{v}) = -p\nabla\cdot\mathbf{v} + \frac{\mathbf{J}^2}{\sigma}\ , \tag{4}$$

$$\frac{4\pi}{c}\mathbf{J} = \nabla\times\mathbf{H}\ . \tag{5}$$

The magnetic field is the poloidal field ($H_z$, $H_r$). In order to fulfill exactly the condition $\nabla\cdot\mathbf{H} = 0$, in the calculations instead of the poloidal component of the magnetic field strength $H_z$ and $H_r$, the toroidal component $A_\phi$ vector potential $\mathbf{A}$ is used, $\mathbf{H} = \nabla\times\mathbf{A}$.

We use a set of characteristic dimensionless parameters. First parameter $\beta = 8\pi P_{0jet}/H_0^2$ determines the relation between the characteristic values of the pressure of the matter $P_{0jet}$ and the pressure of the magnetic field. The second parameter of the numerical model is the characteristic dimensionless magnetic viscosity $\tilde{\eta}_M = \eta_M/L_0V_{A0}$ , where $\sigma$ is the constant effective conductivity of the plasma; $L_0$ is the characteristic size and $V_{A0}$ is the characteristic Alfven velocity.

A detailed description of the numerical model, boundary and initial conditions can be found in our recent papers[5,6].

## 4. Results of numerical simulation

We performed MHD simulation of the laboratory experiment for creating plasma jets taking into account the constant poloidal magnetic field directed along the normal to the target $H$. We performed a series of calculations for different values of the parameter $\beta$. The characteristic parameters for demonstrated case are: $\beta = 10^{-1}$ and $\tilde{\eta}_M = 10^{-5}$

The initial density of matter in the whole calculation region, with the exception of the target, is homogeneous and has a small value: $\rho = \rho_0$. The initial velocity of the substance in the calculation region, with the exception of the target, is zero $v_z = v_r = 0$. The substance has no angular momentum, $v_\phi = 0$. In the target region $0 \le r \le R_d$, $0 \le z \le Z_d$, the density $\rho \approx 300\rho_0$ and the supersonic velocity of the particles emitted from the target $v \ge c_s$ are set, here $c_s = \sqrt{\gamma\frac{P_0}{\rho_0}}$ is the sound

speed in the unperturbed region. For simplicity of test calculations, the velocity of the emitted matter at the initial instant of time has only the z-component $v_z$; $v_r = 0$. The magnetic field is defined as follows: $A_\phi = rH$, $A_r = 0$, $A_z = 0$, $H_\phi = 0$.

The results of modeling the release of matter for this case are shown in Figure 3. The color background corresponds to the logarithm of the density, the arrows indicate the velocity vectors, the length of the arrow is proportional to the velocity value. Solid lines represent magnetic field lines. It can be seen that the target substance under the action of the thermal energy of the laser begins to fly along the axis $Z$ to the opposite wall of the chamber. After the release of matter it begins to expand in all directions, but then the flow is collimated by a magnetic field. The influence of the external gas on the collimation of the flow in this case is insignificant. Inside the flow, a shock wave is formed at a distance of $z \approx 1.6$.

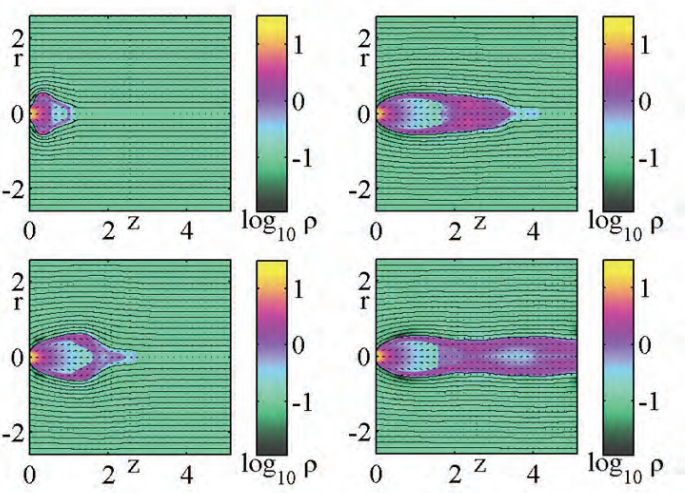



Fig. 3. Picture of the matter flow for MHD case $\beta = 10^{-1}$ at different moments.

We can see from the simulations that near the detector a cavity is formed. It is bounded by a shock wave. The size of the cavity depends on the magnitude of the magnetic field H. Similar results were obtained in work Albertazzi B. et al.[4].

This leads to the formation of ring structures, which are observed in the experiment. Figure 4 is shown the angle of these structures depends on the magnitude of the magnetic field. With a weak field (and also with its absence), this angle is large (left figure). With an increase in the field, the dimension of the cavity and, accordingly, the angle of the ring structure decreases, and then a solid spot is observed instead of the ring (right figure).

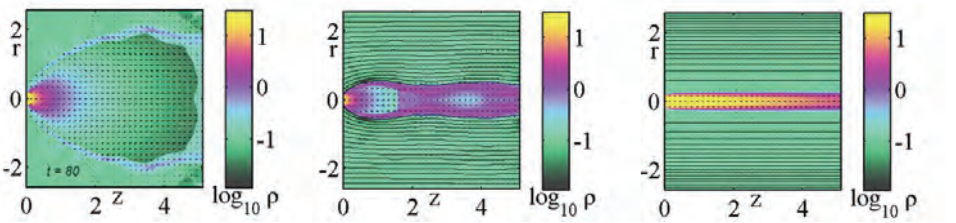



Fig. 4. Picture of the matter flow for HD case and MHD cases $\beta = 10^{-1}$ and $\beta = 10^{-3}$

## 5. Conclusions

The paper presents the results of numerical simulation of magnetized supersonic jets. Formation and development of jet in a laboratory laser experiment is a complex physical phenomenon involving a large number of different physical processes. Numerical simulation of such an experiment, allowing to account for all possible physical processes, is a complex problem and is currently difficult to implement. We used the approximation of magnetohydrodynamics, as the first step in the study of this problem.

## References

1. Bisnovatyi-Kogan G.S. Proc. 6th Int. Workshop of the Astronomical Observatory of Capodimonte (OAC 6), Capri, Italy, September 18-21, 1991, Dordrecht: Kluwer. Eds. L. Errico, and Alberto A. Vittone. Astrophys. Space Sci. Library, **186**, 369, 1993
2. Fua W., Liang E.P., Tzeferacos P., Lamb D.Q., High Energy Density Physics, **17**, 42, 2015
3. Mitrofanov, K. N. et.al. Astronomy Reports, 61, 138, 2017
4. Albertazzi B. et al. Science, **346**, 325, 2014
5. Belyaev, V. S.; Bisnovatyi-Kogan, G. S.; Gromov, A. I.; Zagreev, B. V.; Lobanov, A. V.; Matafonov, A. P.; Moiseenko, S. G.; Toropina, O. D., 2018 Astronomy Reports, Volume 62, Issue 3, pp.162-182
6. Toropina, O. D.; Bisnovatyi-Kogan, G. S.; Moiseenko, S. G., 2018 International Journal of Modern Physics D, Volume 27, Issue 10, id. 1844017-382
7. Belyaev V. S., Vinogradov V. I. , Matafonov A. P. Nuclear physics, **71**, 466, 2008
8. Mazhukin V.I., Shapranov A.V., Demin M.M., Samokhin A.A., Zubko A.E. Mathematica Montisnigri, **38**, 3, 2017
9. Landau, L. D., Lifshitz, E. M. 1960, Electrodynamics of Continuous Media (New York: Pergamon), chap. 8
10. Zhukov, V. T., Zabrodin, A. V., Feodoritova, O. B. 1993, Comp. Maths. Math. Phys., 33, No. 8, 1099

# Probing the inner accretion disk around a spinning black hole: Revisiting the Bardeen-Petterson effect

Srimanta Banerjee* and Sudip Bhattacharyya

*Department of Astronomy and Astrophysics, Tata Institute of Fundamental Research, Mumbai, 400005, India*
**E-mail: srimanta.banerjee@tifr.res.in*

Chandrachur Chakraborty

*Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing, 100871, China*

We study a thin warped accretion disk around a spinning black hole in the viscous regime (i.e., $\alpha > H/R$; $\alpha$ is the Shakura-Sunyaev viscosity parameter, $H$ is the disk thickness and $R$ is the radial distance), and calculate the steady state radial profile of the disk tilt angle for a wide range of relevant parameters of the system (like the Kerr parameter of the black hole). Although the inner part of such a disk was proposed to become aligned with the spin direction of the black hole by the Bardeen-Petterson effect, we show that for a reasonable range of the parameters of the system, the inner disk can stay significantly tilted with respect to the black hole spin. A tilt in the inner accretion disk can affect the observed X-ray spectral and timing features, and hence it makes the inner accretion disk particularly useful for probing the strong gravity region.

*Keywords*: Accretion disks; Black hole physics; Relativistic processes; X-ray binaries.

## 1. Introduction

Accretion disks around spinning black holes are sometimes found to be warped, i.e., disks with changing orbital planes with the radius. For example, warped disks around spinning black holes were predicted for the AGN NGC 4258[1], the X-ray binaries SS433[2] and GRO J1655-40[3]. When there is a non-axisymmetric force, like the Lense-Thirring (LT) precession[4] induced by a misaligned black hole, acting on the disk, a warp is usually generated. Bardeen and Petterson[5] developed the theory of a thin warped disk around a Kerr black hole and showed that the inner part of the disk aligns itself with the spin direction of the black hole by the LT precession. This is known as the Bardeen-Petterson effect. As the LT torque falls off roughly as $1/R^3$ ($R$ is the radial distance), the outer part of the disk remains misaligned[5]. In between these two regions, the disk orients itself from the aligned inner part to the misaligned outer part creating a warped region in the disk. Although several inconsistencies[6] were found in their calculations, the inner part of the thin warped disk was mostly considered to be aligned with the black hole spin direction up to a certain radius, $R_w$ [7–10]. However, recently X-ray spectral and timing features of the Galactic accreting black hole H1743–322[11] has indicated that the inner disk could be tilted. On the theoretical side, Lodato and Pringle[12] found a possibility of tilt in the inner disk in their numerical investigation of a warped thin disk around a spinning black hole.

In this work[13], we revisit the case of a thin viscous warped disk around a Kerr black hole, and study the steady state behavior of the entire disk as a function of the relevant parameters of the astrophysical system. We take into account the contribution of the inner disk, which was mostly ignored in the past due to the alignment of the inner accretion disk, and explore the possibility of a tilt in the inner accretion disk by scanning over a wide range of the parameter space.

## 2. Warped accretion disk around a Kerr Black Hole in Steady State: Formalism

In this work[13], we consider an astrophysical system composed of a slowly rotating Kerr black hole (of mass $M$ and spin $a$) at the center, and a geometrically thin (disk aspect ratio $H/R << 1$, $H$ is the disk thickness and $R$ is the radial distance from the centre) and Keplerian accretion disk around the black hole. The spin of the black hole is directed along the z-axis in our coordinate system. The accretion disk far from the center is slightly tilted with respect to the spin axis of the black hole. We assume the disk to be sufficiently viscous, i.e., $\alpha > H/R$ such that the warp is transmitted diffusively in the disk[6]. Thus, the angular momentum density conservation equation on each annulus of the disk subjected to viscous and LT torques in steady state can be shown to be given by[7]

$$\frac{1}{R}\frac{\partial}{\partial R}\left[\left(\frac{3R}{L}\frac{\partial}{\partial R}\left(\nu_1 L\right)-\frac{3}{2}\nu_1\right)\mathbf{L}+\frac{1}{2}\nu_2 RL\frac{\partial \mathbf{l}}{\partial R}\right]+\frac{\boldsymbol{\omega}_p\times\mathbf{L}}{R^3}=0, \tag{1}$$

where $\mathbf{L}=\sqrt{GMR}\Sigma(R)\mathbf{l}$ is the angular momentum density vector, $\nu_1$ and $\nu_2$ are the kinematic viscosities acting along the plane (horizontal) of the disc and normal (vertical) to the disc respectively, $\omega_p/R^3$ is the LT precession frequency or nodal precession frequency with $\omega_p = 2GJ/c^2$ ($J$ is the total angular momentum of the black hole, $G$ is the gravitational constant and $c$ is the speed of light in vacuum), $\mathbf{l} \simeq (l_x, l_y, 1)$ is the unit tilt vector in the direction of angular momentum density and $\Sigma$ is the surface density of the disk. Here, $\nu_1$ is the usual Shakura-Sunyaev viscosity[14], and $\nu_2$ can be shown to be related to $\nu_1$ and $\alpha$ for small amplitude warp in the following way[15]

$$\frac{\nu_2}{\nu_1}=\frac{1}{2\alpha^2}.\frac{4(1+7\alpha^2)}{4+\alpha^2}. \tag{2}$$

We assume these two viscosities to be constant in the disk.

The modulus of steady state angular momentum density vector can be obtained by taking the scalar product of $\mathbf{l}$ with the equation (1). This gives[7,13]

$$L(R)=C_2R^{1/2}-2C_1, \tag{3}$$

where $C_1$ and $C_2$ are the integration constants. The expression of these two constants can be obtained by using the boundary conditions, $\Sigma(R_{\rm in}) = \Sigma_{\rm in}$ (where $R_{\rm in}$ is the ISCO radius, $R_{\rm ISC0}$, for a Kerr black hole), and $\Sigma \to \infty$ as $R \to \infty$. One obtains the following expressions upon using the above mentioned boundary

conditions[13]

$$C_2 = \sqrt{GM}\Sigma_\infty, \tag{4}$$

and

$$C_1 = \frac{1}{2}\sqrt{GMR_{\rm in}}\left(\Sigma_\infty - \Sigma_{\rm in}\right). \tag{5}$$

One finally arrives at the warped disk equations[13] under the influence of LT precession around a spinning black hole by substituting the expression of L (equation (3)) into equation (1)

$$\frac{\partial}{\partial R}\left(3\nu_1 C_1 l_x + \frac{1}{2}\nu_2 RL\frac{\partial l_x}{\partial R}\right) = \omega_p \frac{L}{R^2} l_y, \tag{6}$$

and

$$\frac{\partial}{\partial R}\left(3\nu_1 C_1 l_y + \frac{1}{2}\nu_2 RL\frac{\partial l_y}{\partial R}\right) = -\omega_p \frac{L}{R^2} l_x, \tag{7}$$

where $\boldsymbol{\omega}_p \times \mathbf{l} = (-\omega_p l_y, \omega_p l_x, 0)$ in our construction. The above two equations can also be combined in the following way[7,13]

$$\frac{\partial}{\partial R}\left(3\nu_1 C_1 W + \frac{1}{2}\nu_2 RL\frac{\partial W}{\partial R}\right) = -i\omega_p \frac{L}{R^2} W, \tag{8}$$

where $W = l_x + il_y = \beta e^{i\gamma}$. Here, $\beta = \sqrt{l_x^2 + l_y^2}$ and $\gamma = \tan^{-1}(l_y/l_x)$ are the tilt and twist angles respectively. We would use the following boundary conditions for solving the equations[13] (6,7)

$$l_x(R_{\rm in}) = \beta_i \cos(\gamma_i),\ l_y(R_{\rm in}) = \beta_i \sin(\gamma_i), \tag{9}$$

and

$$l_x(R_f) = \beta_f,\ l_y(R_f) = 0, \tag{10}$$

where $\gamma_i$, $\beta_i$, $R_f$ and $\beta_f$ are the twist angle at the inner boundary, the tilt angle at the inner edge, the outer edge radius and the outer edge tilt angle, respectively. The twist angle at the outer edge can be assumed to be zero as far from the black hole the effect of LT precession is negligibly small. Thus, $l_y$ becomes zero at the outer edge (see the definition of $l_y$ below equation (8)). This above boundary conditions can be shown to generate non-zero viscous torques at the inner edge[13].

As the inner part of a warped accretion disk was mostly assumed to be aligned with the black hole spin due to the BP effect, the contribution of the inner disk was ignored. Hence the terms associated with $C_1$ in the warped disk equations (6,7) were dropped as $C_1$ is related to the inner disk. In this work[13], we numerically solve the full warped disk equation(s) (6,7) with considering the effect of the terms associated with $C_1$ and the general boundary conditions in order to obtain the radial profile of disk tilt angle as a function of the parameters of the system. Since, our analysis is valid for the entire disk, we explore the behavior of the inner disk for different ranges of the relevant parameters (like the Kerr parameter $a$, $\nu_2$ and $\beta_{\rm in}$), and study how the inner disk alignment depends on these parameters.

## 3. Results and Discussion

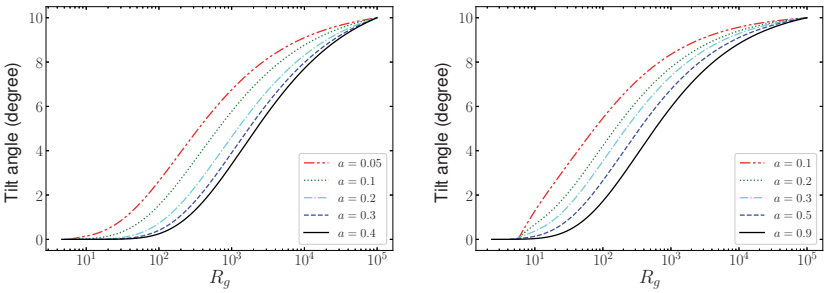



Fig. 1. Radial profiles of the disk tilt angle for different values of the Kerr parameter $a$ for the viscosities $\nu_2 = 10^{14}$ cm$^2$ $s^{-1}$ (Left panel) and $\nu_2 = 10^{15}$ cm$^2$ $s^{-1}$ (Right panel). Here, $R_g = GM/c^2$.

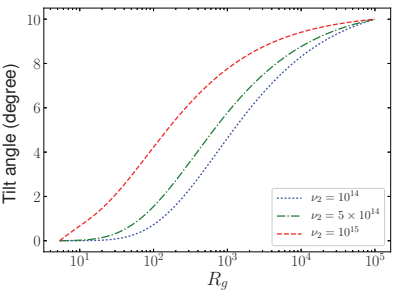



Fig. 2. Radial profiles of the disk tilt angle for different values of $\nu_2$ for the Kerr parameter $a = 0.2$. The viscosity values in the panel are given in cm$^2$ $s^{-1}$. Here, $R_g = GM/c^2$.

We investigate in detail the behavior of the radial profile of the disk tilt angle for a wide range of parameters $a$, $\nu_2$ (or $\nu_1$ and $\alpha$; see equation (2)) and the disk inner edge tilt angle $\beta_i$, and also pay close attention to how the inner disk responds to the changes in those parameters. For the purpose of demonstration, we choose the black hole mass $M = 10M_\odot$, i.e., a typical value of mass of a black hole in Galactic X-ray binaries[16], viscosity $\nu_2$ in the range[17] $10^{14} - 10^{15}$ cm$^2$ s$^{-1}$ and $\alpha = 0.15$[3]. We set the inner edge twist angle and the ratio between outer edge and inner edge surface densities to $\gamma_i = 1^\circ$ and $z_{\rm in} = 0.75$ respectively. Since, we are considering a slightly tilted disk, we set the outer edge disk tilt angle to $10^\circ$. Also, we consider only the case of prograde rotation in this work.

We see from the Fig.1 that as the Kerr parameter takes a higher value, the angle between the inner disk angular momentum vector and the black hole spin falls as the LT torque, being proportional to $a$, becomes more dominant for the fixed values

of the other parameters. But the viscous torque on the plane of the disk ($\nu_2$ is associated with this torque) hinders the process of reduction of tilt angle in the inner disk (see Fig.2), and hence a higher value of $\nu_2$ reduces the dominance of the LT torque in the inner disk (Fig.2). Thus, the alignment of the inner disk is determined by the interplay between the LT torque and the viscous torque on the plane, and their relative strength decides for what values of the Kerr parameter the disk will be aligned with the black hole spin for a given value of $\nu_2$. We see from the Fig.1 that for $\nu_2 = 10^{14}$ cm$^2$ $s^{-1}$, the inner disk alignment occurs if $a \gtrsim 0.3$[13], where for $\nu_2 = 10^{15}$ cm$^2$ $s^{-1}$ the alignment happens when $a \gtrsim 0.8$[13]. The radius up to which the disk is aligned was earlier approximated by the warp radius[3,7,10,12], $R_w = \omega_p/\nu_2$, and this warp radius is defined as the distance at which the warp diffusion time-scale ($R^2/\nu_2$) equals the LT precession time-scale ($R^3/\omega_p$). But, in reality the disk can stay significantly tilted at the warp radius (see Fig.3), and the tilt angle at $R_w$ is roughly 0.13 times the tilt angle at the outer edge[7]. In this work, we find that the alignment radius, i.e., the radius up to which the tilt angle is less than 0.01°, can be approximated by 0.094 $R_w$[13].

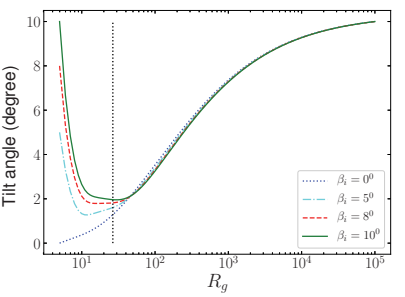



Fig. 3. Radial profiles of the disk tilt angle for different values of the disk inner edge tilt angle $\beta_i$. The other parameter values are $a = 0.3$, $M = 10M_\odot$, $\nu_2 = 10^{15}$ cm$^2$ $s^{-1}$. The vertical line refers to the warp radius as estimated by the relation $\omega_p/\nu_2$. This figure exhibits the dip feature, and also there is a nonzero tilt at the inner disk for all the values of $\beta_i$. Here, $R_g = GM/c^2$.

We have so far set the inner edge tilt angle to zero. Now we explore how different values of the inner edge tilt angle affect the radial profile of the disk tilt angle. Taking a nonzero disk inner edge tilt angle makes the LT torque more dominant near the inner edge. Thus, the more one increases $\beta_i$, the LT torque near the inner edge increases the more, and thus the mismatch between the black hole spin and the angular momentum vector near the inner edge is reduced more rapidly (see Fig.3). But the LT torque may not be reduce the inner disk tilt angle to zero as the viscous torque on the plane inhibits this process. Also after a certain value of radial distance, the LT torque becomes weak and the viscous torque on the plane of the disk takes over the control producing a dip in the radial profile of the disk tilt angle.

## 4. Conclusion

In our work[13], we solve the full prograde thin warped disk equation in the viscous regime in steady state, and obtain the radial profile of the disk tilt angle as a function of various parameters of the system. We take into account the contribution of the inner disk which was ignored in most earlier works, and also consider physically relevant boundary conditions for solving the the warped disk equation(s) (6,7). We find that the relative dominance between the LT torque and viscous torque on the plane of the disk decides the alignment of the inner disk. Thus, there exists a critical value of the Kerr parameter depending upon $M$ and $\nu_2$ beyond which the BP alignment switches on. As a tilt in the inner disk can have observational implication on the timing and spectral features of X-ray emanating from the disk, our analysis may be useful for studying the region close to the black hole.

## References


1. Herrnstein J. R., Greenhill L. J. and Moran J. M. 1996, *The Astrophysical Journal Letters* **468**, L17 (1996).
2. Begelman M. C., King, A. R. and Pringle J. E., 2006, *Mon. Not. R. Astron. Soc.* **370**, 399 (2006).
3. Martin R. G., Tout C. A. and Pringle J. E., *Mon. Not. R. Astron. Soc.* **387**, 188 (2008).
4. Lense J. and Thirring H., *Phys. Z.* **19**, 156 (1918).
5. Bardeen J. M. and Petterson J. A., *The Astrophysical Journal* **195**, L65 (1975).
6. Papaloizou J. C. B. and Pringle J. E., 1983, *M*on. Not. R. Astron. Soc. **202**, 1181 (1983).
7. Scheuer P. A. G. and Feiler R., *Mon. Not. R. Astron. Soc.* **282**, 291 (1996).
8. Natarajan P. and Armitage J. P. 1999, *Mon. Not. R. Astron. Soc.* **309**, 961 (1999).
9. Natarajan, P. and Pringle J. E., *The Astrophysical Journal Letters* **506**, L97 (1998).
10. Martin R. G., Pringle J. E. and Tout, C. A., 2007, *Mon. Not. R. Astron. Soc.* **381**, 1617 (2007).
11. Ingram A., van der Klis M., Middleton M., Done C., Altamirano D., Heil L., Uttley P., and Axelsson M., *Mon. Not. R. Astron. Soc.* **461**, 1967 (2016).
12. Lodato G. and Pringle J. E., *Mon. Not. R. Astron. Soc.* **368**, 1196 (2006).
13. Banerjee S., Chakraborty C. and Bhattacharyya S., *The Astrophysical Journal* **870**, 95 (2019).
14. Shakura N. I. and Sunyaev R. A., *Astronomy & Astrophysics* **24**, 337 (1973).
15. Ogilvie G. I., *Mon. Not. R. Astron. Soc.* **304**, 557 (1999).
16. Fragos T., and McClintock J. E. 2015, *The Astrophysical Journal* **800**, 17 (2005)
17. Frank J., King A. R. and Raine D., Accretion Power in Astrophysics, 3rd Edition, (CUP 2002)

# Relativistic and Newtonian fluid tori with electric charge

Vladimír Karas, Ondřej Kopáček

*Astronomical Institute, Czech Academy of Sciences,*
*Boční II 1401, CZ-14100 Prague, Czech Republic*
*E-mail: vladimir.karas@cuni.cz*

Devaky Kunneriath

*North American ALMA Science Center, National Radio Astronomy Observatory,*
*Charlottesville, VA 22903, USA*

Audrey Trova

*Center of Applied Space Technology and Microgravity (ZARM), University of Bremen,*
*Am Fallturm, D-28359 Bremen, Germany*

Jiří Kovář, Petr Slaný and Zdeněk Stuchlík

*Institute of Physics, Faculty of Philosophy and Science, Silesian University in Opava,*
*Bezručovo nám. 13, CZ-74601 Opava, Czech Republic*

We discuss the effects of electric charging on the equilibrium configurations of magnetized, rotating fluid tori around black holes of different mass. In the context of gaseous/dusty tori in galactic nuclei, the central black hole dominates the gravitational field and it remains electrically neutral, while the surrounding material acquires some electric charge that can separate in space and exhibit non-negligible self-gravitational effects. The structure of the torus is thus influenced by the balance between the gravitational and electromagnetic forces. A cusp can develop near the inner rim even in Newtonian tori due to the charge distribution. Furthermore, it is interesting to note that stable polar clouds can emerge near the rotation axis in reminiscence of "lamppost" scenario. For the latter, an appropriate distribution of the angular momentum and electric charge are essential.

*Keywords*: Gravitation; Black Holes; Accretion; Magnetic fields

## 1. Introduction

Within the framework of General Relativity, black holes are fully described by a small number of parameters[1]. The most relevant from the viewpoint of astrophysical applications are the mass of the black hole, $M_\bullet$, and angular momentum $J_\bullet$ or, in a dimension-less form, the spin parameter $a^* \equiv a/M_\bullet = J_\bullet c/GM_\bullet^2$. The classical black-hole solution of Kerr-Newmann can be equipped by electric charge, although any significant value appears to be astrophysically unlikely over an extended duration because of continued process of neutralization by a selective charge accretion from the surrounding plasma into any cosmic black holes must be embedded.

The Kerr-Newman solution the line element adopts the form adopts the form[2]

$$ds^2 = -\frac{\Delta}{\Sigma}\left(dt - a\sin\theta d\phi\right)^2 + \frac{\Sigma}{\Delta}dr^2 + \Sigma d\theta^2 + \frac{\sin^2\theta}{\Sigma}\left[\left(r^2+a^2\right)d\phi - adt\right]^2, \quad (1)$$

in dimensionless Boyer-Linquist coordinates $(t, r, \phi, \theta)$, where $\Delta(r) = r^2 - 2r + a^2 + e^2$ and $\Sigma(r, \theta) = r^2 + a^2 \sin^2 \theta$. Quantities $a$ and $e$ are the rotational and charge parameters of the spacetimes. The associated non-zero components of the antisymmetric electromagnetic field tensor $F_{ij} = A_{j,i} - A_{i,j}$ are

$$F_{rt} = \frac{e(r^2 - a^2 \cos^2 \theta)}{\Sigma^2}, \qquad F_{r\phi} = \frac{-ae \sin^2 \theta (r^2 - a^2 \cos^2 \theta)}{\Sigma^2}, \tag{2}$$

$$F_{\theta t} = \frac{-a^2 er \sin 2\theta}{\Sigma^2}, \qquad F_{\theta\phi} = \frac{aer \sin 2\theta (r^2 + a^2)}{\Sigma^2}. \tag{3}$$

The effective potential $W(r, \theta)$ for the motion of a particle with the specific charge $\tilde{q} = q/m$ satisfies the relation

$$XW = Y + \sqrt{Y^2 - XZ}, \tag{4}$$

where the functions $X(r, \theta) = (r^2 + a^2)^2 - \Delta a^2 \sin^2 \theta$, $Y(r, \theta) = (\tilde{L}a + \tilde{q}er)(r^2 + a^2) - \tilde{L}a\Delta$, and $Z(r, \theta) = (\tilde{L}a + \tilde{q}er)^2 - \Delta\Sigma - \Delta\tilde{L}^2/\sin^2 \theta$; here, $\tilde{L} = L/m$ is the conserved axial component of specific angular momentum. The effective potential (4) develops local minima outside the equatorial plane $\theta = \pi/2$, which suggests the possibility of halo orbits.[a] The minima occur provided that the electric charge of the accreted matter is non-zero and interacts with the global magnetic field of the central object and/or of the external origin due to currents flowing outside the horizon.

In the other words, the stable configuration allows for a distribution of charged matter near the equator, in lobes located symmetrically above and below the equatorial plane, or even around the polar axis (see Fig. 1). This represents general-relativistic version of Störmer's halo orbits that have been explored in connection with the motion of electrically charged dust grains in planetary magnetospheres.[5] In the case of Kerr-Newman solution, it can be found that the particular form of mutually connected gravitational and electromagnetic fields do not allow existence of stable halo orbits above the outer horizon; an additional external magnetic field is required.

The prevailing process of neutralization, however, may not prevent the surrounding material to undergo electric charging in the environment of complex plasma that is irradiated by energetic X-rays, as is often the case in active galactic nuclei. This mechanism leads to charge separation within the dusty tori. Therefore, we have proposed the process of charging as mechanism that can maintain geometrical thickness of the dusty tori in the vertical direction.[6]

Furthermore, the existence of mutually detached regions (lobes) of bound stable orbits can provide a framework to trigger oscillations; it allows us to imagine a situation when the bulk motion along the circular trajectories is superposed by

[a] Within the particle approximation the halo orbits are formed by stable bound trajectories which do not cross the equator. In the fluid analogy they correspond to toroidal structures with the pressure maxima at $\theta \neq \pi/2$ (i.e. $z \neq 0$).[3,4]

a fraction o material that performs oscillatory motion between the disconnected lobes. Halo orbits could have outstanding consequences for the radiation signal from accreting black holes. On the other hand, general relativity effects have a tendency to bring the halo orbits closer to the equatorial plane compared to corresponding Newtonian solutions. Also, to ensure the intrinsic self-consistency of the model we need to include self-gravity of the medium in the consideration, as this can further diminish the vertical thickness of the figures of equilibrium of electrically charged fluid.[8,9,11,12]

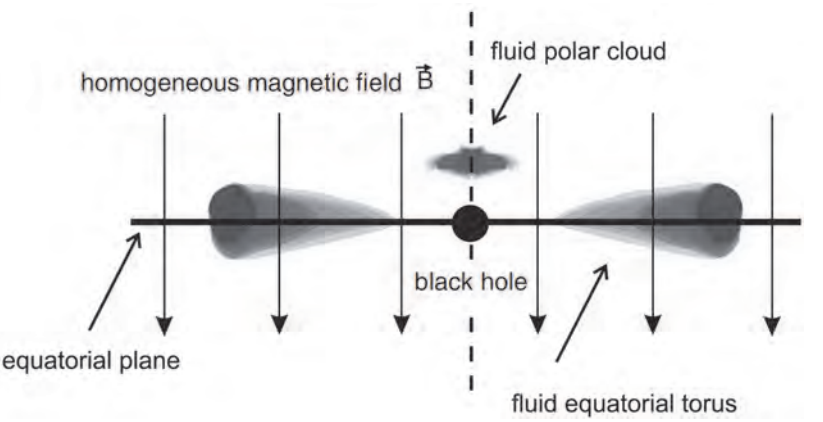



Fig. 1. A black hole is immersed in asymptotically homogeneous magnetic field $\vec{B}$ and encircled by a torus-like equatorial fluid configuration, and an off-equatorial polar cloud. The charge distribution within the fluid interacts with the magnetic field and helps to levitate the fluid above the equatorial plane. Similar effect occurs for trajectories of electrically charged particles. Regions of stability can emerge near the equatorial plane as well as near the polar axis, depending on the angular momentum and electric charge distribution within the fluid.[6–8]

## 2. Electrically charged tori near a magnetized black hole

We describe the fluid by polytropic equation of state, $p = \kappa\rho^{\Gamma}$, with $\kappa$ and $\Gamma$ being parameters of the polytrope. This relation ensures conservation of entropy, as is appropriate for a perfect fluid. Neglecting the electrostatic corrections to the equation of state we can use the polytropic equation consistently even to described an electrically charged state. In the case of sufficiently small $\rho$ the medium is non-relativistic and the contribution of the specific internal energy $\epsilon/\rho - 1$ to the total energy density becomes negligible (i.e., $\epsilon \approx \rho$), however, this assumption is not accurate enough once the density increases; in such a case the self-electromagnetic field cannot be ignored.

The configuration of the fluid is determined by its rotation in the $\phi$-direction with 4-velocity $U^{\alpha} = (U^t, U^{\phi}, 0, 0)$, specific angular momentum $\ell = -U_{\phi}/U_t$ and

angular velocity $\omega = U^\phi/U^t$. These are related by

$$\omega = -\frac{\ell g_{tt} + g_{t\phi}}{\ell g_{t\phi} + g_{\phi\phi}}, \quad (U_t)^2 = \frac{g_{t\phi}^2 - g_{tt} g_{\phi\phi}}{\ell^2 g_{tt} + 2\ell g_{t\phi} + g_{\phi\phi}}. \tag{5}$$

The shape of a charged fluid torus with a charge density profile $q$ and energy density $\epsilon$ is determined by isobaric surfaces of the pressure $p$ profile (equi-pressure surfaces), which can be derived from the conservation equation

$$\nabla_\beta T^{\alpha\beta} = F^{\alpha\beta} J_\beta, \tag{6}$$

where $J^\alpha$ is the fluid 4-current density and $T^{\alpha\beta}$ is the stress-energy tensor,

$$T^{\alpha\beta} = (\epsilon + p) U^\alpha U^\beta + p g^{\alpha\beta}. \tag{7}$$

From the above-given relation two partial differential equations follow for the radial and latitudinal distribution of pressure,

$$\frac{\partial p}{\partial r} = -(p+\epsilon)\mathcal{R}_1 + q\mathcal{R}_2 \equiv \mathcal{R}_0, \quad \frac{\partial p}{\partial \theta} = -(p+\epsilon)\mathcal{T}_1 + q\mathcal{T}_2 \equiv \mathcal{T}_0, \tag{8}$$

where $\mathcal{R}_0 = \mathcal{R}_0(r,\theta)$ and $\mathcal{T}_0 = \mathcal{T}_0(r,\theta)$ denote the right hand sides of the two latter equations.[6]

As a specific example of the imposed electromagnetic field we can write the static case of an asymptotically uniform magnetic field with intensity $B$. Such settings can be described by the special case (zero rotational parameter) of Wald's test-field solution of Einstein-Maxwell equations, which in dimensionless units reads

$$A_t = -\frac{e}{r}, \quad A_\phi = \tfrac{1}{2} B r^2 \sin^2\theta, \tag{9}$$

describing the electromagnetic field in the background of Schwarzschild geometry. A generalization of the adopted approach to the case of a rotating (Kerr) black hole is straightforward but technically more complicated and beyond the limited space of the present contribution.

## 3. Conclusions

We further explored off-equatorial configurations of electrically charged fluid near magnetized black holes. By comparison with previous works we showed that the conditions of existence of these configurations are diverse and they can be established in varied circumstances during the black hole accretion. By the mass-scaling paradigm for cosmic black-hole sources, similar effects can occur also in the context of stellar-mass black holes, where a massive torus can form as a remnant of a tidally disrupted companion.

Furthermore, electrically charged particles and fluid elements can be set on escaping trajectories and eventually released at a relativistic outflow velocity escaping from the atmosphere of the otherwise stable configuration. The case of charged particles differs from charged dust grains by their charge-to-mass ratio, but the acceleration mechanism operates in a similar manner.[10]

Let us conclude by emphasising the fact that electrically charged toroidal structures can adopt varied shapes from the rich topological family. They exist not only in the equatorial plane as a single torus, but also as two tori coupled and joint through a cusp, existing individually even out of the equatorial plane (as a pair of off-equatorial tori), or just around the polar axis (as a kind of realisation of the lamp-post model). Here adopted the fluid description; in the limit of very low-density the approach of kinetic theory turns out to be better suited.[13]

**Acknowledgments**

V.K. and A.T. acknowledge the Czech Science Foundation (GAČR) – German Research Foundation (DFG) collaboration project No. 19-01137J.

**References**

1. Carter B., *Physical Review* **174**, 1559 (1968).
2. Misner C. W., Thorne K. S. and Wheeler J. A., *Gravitation* (Freeman, San Francisco, 1973).
3. Dullin H. R., Horányi M. and Howard J. E., *Physica D* **171**, 178 (2002).
4. Kovář J., Stuchlík Z., and Karas V., *Class. Quantum Grav.* **25**, id. 095011 (2008).
5. Störmer C., *The Polar Aurora* (Clarendon Press, Oxford, 1955).
6. Kovář J., Stuchlík Z., Karas V., Cremaschini C. and Miller J. C., *Physical Review D* **84**, id. 084002 (2011).
7. Kovář J., Slaný P., Cremaschini C., Stuchlík Z., Karas V. and Trova A., *Physical Review D* **93**, id. 124055 (2016).
8. Schroven K., Trova A., Hackmann E. and Lämmerzahl C., *Physical Review D* **98**, id. 023017 (2018).
9. Karas V., Huré J.-M. and Semerák O., *Class. Quantum Grav.* **21**, R1 (2004).
10. Kopáček O. and Karas V., *The Astrophys. Journal* **853**, id. 53 (2018).
11. Trova A., Karas V., Kovář and Slaný P., *The Astrophys. Journal, Suppl. Series* **226**, id. 12 (2016).
12. Trova A., Schroven K., Hackmann E., Karas V., Kovář J. and Slaný, *Physical Review D* **97**, id. 104019 (2018).
13. Cremaschini C., Kovář J., Slaný P., Stuchlík Z. and Karas V., *The Astrophys. Journal, Suppl. Series* **209**, id. 15 (2013).

# Implicit Lagrangian numerical approach for magnetorotational supernova simulations

Sergey G. Moiseenko[1], Gennady S.Bisnovatyi-Kogan[1,2]

[1]*Space Research Institute, Profsoyuznaya str. 84/32, Moscow, 117993,*
[2]*National Research Nuclear University MEPhI (Moscow Engineering Physics Institute), Kashirskoe Shosse 31, Moscow 115409,Russia*
[*]*E-mail: moiseenko@iki.rssi.su, gkogan@iki.rssi.ru*

Nikolai V.Ardelyan

*Department of Computational Mathematics and Cybernetics,Moscow State University, Vorobjevy Gory, Moscow B-234, Russia,*
*E-mail: ardel@cs.msu.ru*

We discuss numerical method based on Implicit completely conservative Lagrangian operator-difference scheme on triangular grid of variable structure. The method was successfully applied for the simulations of magnetorotational supernova explosion.

*Keywords*: magnetohydrodynamics, supernovae, numerical methods.

## 1. Introduction

The numerical simulations is one of the important tools for study astrophysical problems. A number of astrophysical problems can be simulated by numerical solution of the magnetohydrodynamical(MHD) equations. There lot of different numerical approaches for the simulation of MHD equations. Many astrophysical problems are charactrised by huge variety of spatial and temporal parameters. Application of In this paper we consider the completely conservative Lagrangian operator-difference scheme on a triangular grid of variable structure suggested in ([1,2]).

Lagrangian numerical schemes are free of the artificial transfer of angular momentum, but in the case when in the flow there are nonuniform contraction or expansion, vortexes or shear flows the Lagrangian cells become flattened soon after the beginning of the simulations, thus significantly reducing the accuracy of calculations and preventing continuing the simulations. In such situation the grid remapping allows to overcome this drawback of Lagrangian schemes.

## 2. Completely conservative operator-difference method on a triangular grid of variable structure

The operator-difference scheme used in our simulation is a first-order implicit conservative Lagrangian scheme on an irregular triangular grid. For further details of the scheme, see[1,2]. We briefly describe the idea of the method for 2D cylindrical coordinates case. Define in the computational domain $\Omega$ a triangular grid $\omega = \omega_\times + \omega_\triangle$, which consists of a set of knots $\omega_\times$ and a set of triangular cells $\omega_\triangle$; define in particular the set of boundary knots $\omega_\gamma$. Introduce linear spaces $B_{\times,1}\,(B_{\gamma,1})$ of the vector functions defined in the knots ( boundary knots) $\overline{x}_j \in \omega_\times\,(\overline{x}_q \in \omega_\gamma)$ and

linear spaces $B^0_{\triangle,0}$ ($B_{\gamma,0}$) of scalar functions defined in the cells (boundary knots) $\triangle_i \in \omega_\triangle$ ($\overline{x}_q \in \omega_\gamma$). Introduce also the linear space $B_{\triangle,0}$ of scalar functions defined in the cells and boundary knots. The following relation is valid for the introduced spaces: $B_{\triangle,0} = B^0_{\triangle,0} \oplus B_{\gamma,0}$.

Define linear operators ${\nabla_\times}^0$, $\nabla_\triangle \cdot$ and $\Phi_\gamma$ in the following form:

$$\left(\nabla_\triangle \cdot \mathbf{v}\right)_i = \frac{1}{3V_i}\sum_{l=1}^{3} N_{l,l+1}(\mathbf{v}_l R_{l,l+1} + \mathbf{v}_{l+1} R_{l+1,l}), \tag{1}$$

$$\mathbf{v} \in B_{\times,1},\ \triangle_i \in \omega_\triangle. \tag{2}$$

$$\left({\nabla_\times}^0 p\right)_j = \frac{1}{3V_j}\sum_{k=1}^{K_j} {N_k}' p_k, \quad p \in B^0_{\triangle,0},\ \overline{x}_j \in \omega_\times. \tag{3}$$

$$(\Phi_j p)_q = \frac{1}{3V_q} N'_q p_q, \quad p \in B_{\gamma,0},\ \overline{x}_q \in \omega_\gamma, \tag{4}$$

where $R_{j_1,j_2} = r_{j\,1} + r_{j\,2}/2$,

$$N_{j_1,j_2} = \begin{pmatrix} z_{j\,2} - z_{j\,1} \\ r_{j\,1} - r_{j\,2} \end{pmatrix} \text{ is a normal vector to the segment} [x_{j\,1}, x_{j\,2}], \tag{5}$$

$${N_k}' = -(N_{j,k} R_{j,k} + N_{j,k+1} R_{j,k+1}),\ {N_q}' = (N_{q,q+1} R_{q,q+1} + N_{q-1,q} R_{q,q-1}). \tag{6}$$

Introduce also the grid operator $\nabla_\times$ in the following form: $\forall \tilde{p} = p + p_\gamma \in B_{\triangle,0}, \quad p \in B^0_{\triangle,0}, \quad p_\gamma \in B_{\gamma,0}$ :

$$\nabla_\times \tilde{p} = \begin{cases} {\nabla_\times}^0 p + \Phi_\gamma p_\gamma, & \overline{x}_j \in \omega_\times, \\ {\nabla_\times}^0 p, & \overline{x}_j \in \omega_\times \backslash \omega_\gamma. \end{cases} \tag{7}$$

The operator $\nabla_\triangle \cdot$ is a grid analog for the differential operator div , $\nabla_\times$ is a grid analog for grad. The following relation is valid for the introduced operators:

$$(\nabla_\times p, \mathbf{v})_\times + (\nabla_\triangle \cdot \mathbf{v}, p)_\triangle = (\Phi_\gamma p_\gamma, \mathbf{u}_\gamma)_\gamma. \tag{8}$$

where $p \in B^0_{\triangle,0}$, $\mathbf{v} \in B_{\times,1}$, $p_\gamma \in B_{\gamma,0}$, $\mathbf{v}_\gamma \in B_{\gamma,1}$.

The grid linear operators $\nabla^0_\times \cdot$, $\nabla_\triangle$, $\Phi_\gamma \cdot$ are defined analogously to (2), (3), (3′). Assume that $\mathbf{x}$, $\mathbf{u}, \Phi$ are defined in knots and $p, \rho$ – in cells. Define the time grid $\omega_\tau = \{t_n = n\triangle t,\ n = 0, 1, \ldots\}$; for the arbitrary grid function defined on $\omega_\tau$, introduce the following specifications: $f(t_n) = f^n = f, f^{n+1} = \hat{f}, f_t = \frac{\hat{f} - f}{\tau}, \quad f^{(\sigma)} = \sigma\hat{f} + (1-\sigma)f$.

For simplicity we consider set of gas dynamical equations with self-gravitation:

$$\begin{aligned} \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \mathbf{u}, \quad \frac{\mathrm{d}\rho}{\mathrm{d}t} + \rho \mathrm{div}\mathbf{u} = 0, \\ \rho\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = -\mathrm{grad}p - \rho\mathrm{grad}\Phi, \\ \rho\frac{\mathrm{d}\varepsilon}{\mathrm{d}t} + p\mathrm{div}\mathbf{u} = 0, \quad \eta = \frac{1}{\rho} = \frac{T\Re}{p}, \quad \varepsilon = \frac{T\Re}{\gamma - 1}, \\ \Delta\Phi = 4\pi G\rho. \end{aligned} \tag{9}$$

where $\frac{\mathrm{d}}{\mathrm{d}t}$ is total time $t$ derivative , $\mathbf{x} = (r, z)$, $\mathbf{u} = (u^r,\ u^\varphi,\ u^z)$ is velocity vector, $p$ is pressure, $\varepsilon$ is internal energy, $\Phi$ is gravitational potential, $\rho$ is density, $T$ is temperature, $G$ is gravitational constant, $\Re$ is universal gas constant, $\gamma$ is adiabatic index. Axial and cylindrical symmetry are assumed.

On the basis of the introduced operators, the following implicit difference scheme for (9) can be written[4]:

$$\rho_i V_i = {\rho_i}^0 {V_i}^0 = m_i, \quad \forall \triangle_i \in \omega_\triangle,\ \mathbf{x}_{jt} = {\mathbf{u}_j}^{(0.5)}, \quad \forall \overline{x}_j \in \omega_\times,$$

$$\rho_j \mathbf{u}_{jt} = -(\nabla_\times g)_j - q\rho_j \big(S_\times(\nabla_\triangle \Phi)\big)_j, \quad \forall \overline{x}_j \in \omega_\times,$$

$$\rho_i \varepsilon_{it} = -g_i \big(\nabla_\triangle \cdot \mathbf{u}^{(0.5)}\big)_i, \quad \forall \triangle_i \in \omega_\triangle,\ g = p^{(\alpha)} + \omega, \quad \omega = -\frac{\nu}{\eta}\eta_t, \quad \forall \triangle_i \in \omega_\triangle,$$

$$\eta_i = 1/\rho_i = T_i/p_i, \quad \varepsilon_i = T_i/(\gamma - 1), \quad \forall \triangle_i \in \omega_\triangle,$$

$$(\nabla_\times \cdot \nabla_\triangle \Phi)_j = \rho_j,\ \rho_j = (S_\times \rho)_j = \frac{1}{3V_j^\times} \sum_{k=1}^{K_j} \rho_k V_k^\triangle, \quad \forall \overline{x}_j \in \omega_\times.$$

Here $\nu$ is a coefficient of artificial viscosity. The total time derivatives for the velocity vector in the difference form are:

$$\mathbf{u}_{jt} = \big\{u_{jt}^r - \frac{u_j^{\varphi(0.5)}}{r_j^{(0.5)}} u_j^{\varphi(0.5)}, \quad u_{jt}^\varphi + \frac{u_j^{\varphi(0.5)}}{r_j{}^{(0.5)}} u_j^{r(0.5)}, \quad u_{jt}^z\big\}.$$

Substituting the boundary values known from the boundary conditions into (4), we get a set of non-linear operator equations for the grid functions $\mathbf{x}, \mathbf{u} \in B_{\times,1}$, $\rho, T \in B_{\triangle,0}$ at every time step, which is solved by a specially adjusted variant of the Newton method. The gravitational potential $\Phi$ was calculated explicitly at the beginning of every time step.

## 3. Triangular grid structure reconstruction procedure

One of the important problems for application of difference schemes in Lagrangian variables is grid distortion which prevents to continue the simulations and could lead also to the loss of accuracy of the results. In order to overcome this problem in the papers[2,3] and some other papers of these authors the grid remapping procedure for a triangular grid was developed.

Here the remapping procedure was modified what allows to carry out simulations in the presence of large differences in flow parameters[6,7].

The procedure for the remapping the grid, in turn, consists of two stages.The first is a local change in the structure of the grid. The second is to determine the values of the grid functions in the areas of the changed structure.The local change in the structure of the computational grid is reduced to three local operations:

(1) the diagonal of a quadrangle formed by two triangles is replaced by another diagonal
(2) combining two adjacent grid nodes
(3) adding a new node to the middle of the link between neighboring nodes

The structure of the grid is corrected using the first two operations. Mesh adaptation occurs through operations 2 and 3.

## 4. Grid functions remapping method for completely conservative Lagrangian operator-difference scheme for astrophysical MHD problems

At the second stage of the of the grid remapping, the values of the grid functions in the nodes and cells involved in changing the structure mesh are calculated. For the simulation of astrophysical problems, a method of conditional minimization of functionals was proposed and implemented, which ensures the exact implementation of conservation laws. To calculate the density in new cells, it was suggested to look for the conditional minimum of the following functional[6]:

$$\min_{\tilde{\rho}_i} \sum_{\tilde{I}} \frac{(\tilde{\rho}_i - \rho_{ci})^2}{\rho_{ci}/\tilde{V}_i}, \tag{10}$$

under the condition of mass conservation of vicinity of the remapping region:

$$\sum_{\tilde{I}} \tilde{\rho}_i \tilde{V}_i = \sum_{I} \rho_i V_i = M_{\text{vicinity}}. \tag{11}$$

Here $\rho_{ci}$ - is a linear interpolation of the density to a new grid cells. Applying the Lagrange multiplier method to minimize the conditional functional, we obtain:

$$\tilde{\rho}_i = \rho_{ci} \frac{M_{\text{vicinity}}}{\sum\limits_{\tilde{I}} \rho_{ci} \tilde{V}_i}. \tag{12}$$

It is easy to show that the density recalculation by the formula (12)allows us to fulfill the law of conservation of mass in the remapping region.

For the calculations of new values of the poloidal magnetic field components $H_r$, $H_z$ we use the following conditional minimization of the functionals[6]: $\min_{\tilde{H}_{ri}} \sum_{\tilde{I}} \{(\tilde{H}_{ri} - H_{rci})^2\} \tilde{V}_i$, $\min_{\tilde{H}_{zi}} \sum_{\tilde{I}} \{(\tilde{H}_{zi} - H_{zci})^2\} \tilde{V}_i$, under the conditions of conservation of every component of the magnetic energy in the remapping region: $\sum_{\tilde{I}} \tilde{H}_{ri}^2 \tilde{V}_i = \sum_I H_{ri}^2 V_i = E_{mag,\, r\ \text{vicinity}}$, $\sum_{\tilde{I}} \tilde{H}_{zi}^2 \tilde{V}_i = \sum_I H_{zi}^2 V_i = E_{mag,\, z\ \text{vicinity}}$. Here $H_{rci}$, $H_{zci}$ - is a linear interpolation of poloidal magnetic field components to a new grid cells.

Conditional minimum of these functionals is found by the Lagrange multiplier method.

For the calculation of the poloidal components of the magnetic field $H_r$, $H_z$ we get following formulae:

$$\tilde{H}_{ri} = H_{rci}\sqrt{\frac{E_{mag,\,r\text{ vicinity}}}{\sum\limits_{\tilde{I}} H_{rci}^2\tilde{V}_i}},\ \tilde{H}_{zi} = H_{zci}\sqrt{\frac{E_{mag,\,z\text{ vicinity}}}{\sum\limits_{\tilde{I}} H_{zci}^2\tilde{V}_i}},$$

For the calculations of the toroidal component of the magnetic field $H_\varphi$ on a new grid the following conditional minimum of the functional is calculated:

$$\min_{\tilde{H}_{\varphi i}} \sum_{\tilde{I}} \frac{(\tilde{H}_{\varphi i} - H_{\varphi ci})^2}{H_{\varphi ci}}\tilde{S}_i, \tag{13}$$

under the condition of the conservation of the toroidal magnetc flux $H_\varphi S$ in the remapping region: $\sum_{\tilde{I}} \tilde{H}_{\varphi i}\tilde{S}_i = \sum_I H_{\varphi i}S_i = F_{mag,\,\varphi\text{ vicinity}}$. Finally for the calculation if the toroidal components of the magnetic field $\tilde{H}_{\varphi i}$ we get the following formula:

$$\tilde{H}_{\varphi i} = H_{\varphi ci}\frac{F_{mag,\,\varphi\text{ vicinity}}}{\sum\limits_{\tilde{I}} H_{\varphi ci}\tilde{S}_i}.$$

Using this formula one can calculate the vales of the toroidal component of the magnetic field in a new cells preserving the toroidal magnetic flux.

In order to find the values of the toroidal component of the magnetic field in the cells of the remapping region, with conservation of the toroidal magnetic energy the conditional minimum of the following functional is sought (by analogy with the poloidal components of the magnetic field): $\min_{\tilde{H}_{\varphi i}} \sum_{\tilde{I}}\{(\tilde{H}_{\varphi i} - H_{\varphi ci})^2\}\tilde{V}_i$, provided that the toroidal magnetic energy is preserved in the remapping region: $\sum_{\tilde{I}} \tilde{H}_{\varphi i}^2\tilde{V}_i = \sum_I H_{\varphi i}^2 V_i = E_{mag,\,\varphi\text{ vicinity}}$. In this case, to calculate the toroidal components of the magnetic field $\tilde{H}_{\varphi i}$ we get the following formula:

$$\tilde{H}_{\varphi i} = H_{\varphi ci}\sqrt{\frac{E_{mag,\,\varphi\text{ vicinity}}}{\sum\limits_{\tilde{I}} H_{\varphi ci}^2\tilde{V}_i}}.$$

To find the new values of the temperature $\tilde{T}_i$ (the pressure $\tilde{p}_i$) in the remapping region, we seek the conditional minimum of the following functional: $\min_{\tilde{T}_{\varphi i}} \sum_{\tilde{I}}(\tilde{T}_{\varphi i} - T_{\varphi ci})^2\tilde{V}_i$, under the condition of preserving of the internal energy in the remapping region $\sum_{\tilde{I}} \tilde{\varepsilon}_i\tilde{V}_i = \sum_I \varepsilon V_i = E_{int,\,\text{vicinity}}$, We solve this system of equations by Newton method.

## 5. Application to the magnetorotational processes in astrophysics

The idea of the MR mechanism of core-collpased supernovae was suggested by G.S. Bisnovatyi-Kogan in[10]. During the nonuniform collapse of the rotating iron core differentially rotating configuration is formed. Internal parts of the collapsed core rotate faster than its periphery. In the case of presence even weak initial poloidal

magnetic field the magnetic field will be twisted due to the differential rotation. Increasing magnetic pressure will lead to the formation of MHD shock wave. The shock will move outwards producing supernova explosion. For the numerical study of the MR supernova explosion a set of ideal MHD equations with self-gravitation were numerically solved using described above numerical scheme[6,7]. During the simulations of MR supernova explosion the development of the special type of MR instability was found[11].

## References


1. N.V. Ardelyan, K.V. Kosmachevskii, S.V. Chernigovskii, *Problems of construction and research of conservative difference schemes for magneto-gasdynamics.*(MSU, Moscow,1987) (in Russian).
2. N.V. Ardelyan, K.V. Kosmachevskij, "Implicit free-Lagrange method for computing two-dimensional magnetogas-dynamic flows," Computational Mathematics and Modelling **6**, 4, 209–224 (1995).
3. N.V. Ardelyan, K.V. Kosmachevskii, S.V. Chernigovskii, Some features of the calculation of two-dimensional gasdynamic problems on irregular triangular meshes.//In: Computational mathematics and computer software. Moscow State University, 1985, p.10 ( in Russian).
4. N.V. Ardeljan, K.V. Kosmachevskii, S.V. Chernigovskii, *Problems of construction and research of conservative difference schemes for magneto-gasdynamics.*(MSU, Moscow,1987) (in Russian).
5. V.M. Goloviznin, A.A. Samarskii, A.P. Favorskii "On artificial viscosity and stability of difference schemes of hydrodynamics.", *Preprint of the Institute of Applied Mathematics of the Russian Acad. Sci.* No. 70, 1976 (in Russian).
6. N.V. Ardeljan, G.S. Bisnovatyi-Kogan, S.G. Moiseenko, "Magnetorotational supernovae," Monthly Notices of the Royal Astronomical Society **359**, 333–344 (2005).
7. S.G. Moiseenko, G.S. Bisnovatyi-Kogan, N.V. Ardeljan, "A magnetorotational core-collapse model with jets," Monthly Notices of the Royal Astronomical Society **370**, 501–512 (2006).
8. Ardeljan N.V., Bisnovatyi-Kogan G.S., Kosmachevskii K.V., Moiseenko S.G., Simulation of the collapse of a rotating gas cloud on triangular reconstructing lagrangian grid, 1996, Astron &Astrophys Suppl, 115, 573.
9. Ardeljan N.V., Bisnovatyi-Kogan G.S., Moiseenko S.G., Nonstationary magnetorotational processes in a rotating magnetized cloud, 2000, Astron & Astrophys, 355, 1181.
10. G.S. Bisnovatyi-Kogan, "The Explosion of a Rotating Star As a Supernova Mechanism," Soviet Astronomy **14**, 652–655 (1971).
11. S.G. Moiseenko, G.S. Bisnovatyi-Kogan, "Development of the magneto-differential-rotational instability in magnetorotational supernova," Astronomy Reports **7**, 573–580 (2015).

# Magnetized disc-outflow symbiotic model around black holes

Tushar Mondal* and Banibrata Mukhopadhyay†

*Department of Physics, Indian Institute of Science, Bengaluru 560012, India*
**E-mail: mtushar@iisc.ac.in*
*†E-mail: bm@iisc.ac.in*

Theoretical and observational inferences reveal that accretion and outflows/jets are strongly correlated through fundamental conservation laws. All well-known jetted sources with accreting black hole systems indicate a signature of hot advective accretion flow and a dynamically important magnetic fields at the jet footprint. We inter-connect here the disc and outflow/jet for magnetized, viscous, optically thin, geometrically thick, 2.5-dimensional hot advective flows around black holes. The magnetic fields are captured from the environment, say interstellar medium or companion star, and dragged inward with continuous accretion process. This disc becomes magnetically dominated in vicinity of a black hole due to flux freezing. The accretion geometry as well as energetics of the accretion induced outflows/jets strongly depend on the mass and spin of the black hole, accretion rate, and magnetic field strength. We suggest, this magnetically dominated accretion powered system can easily explain the energetics of the ultraluminous X-ray sources in their hard power-law dominated states.

*Keywords*: accretion, accretion discs – black hole physics – MHD (magnetohydrodynamics) – gravitation – X-rays: binaries – galaxies: jets

## 1. Introduction

The observational evidences show strong outflows and jets in various black hole (BH) accreting systems, in quasars or active galactic nuclei, microquasars, and gamma-ray bursts. These outflows and jets are more prone to emanate from the inner region of the accretion flow, particularly when the flow is hot, puffed-up, and advective. Observational evidences predict that all BH X-ray binaries to exhibit strong radio jets in its low/hard spectral state[1]. The low/hard state basically indicates the accretion regime with quite low mass accretion rate and the disc is radiatively inefficient. However, the proper mechanisms for the formation, collimation, and acceleration of relativistic jets are still not well understood.

Theoretical models suggest the role of large-scale magnetic fields in the angular momentum transport in accretion, in the formations of strong outflows/jets, and in high-energy radiation mechanism[2,3]. The large-scale open magnetic fields threaded the disc can extract energy and angular momentum in the form of outflows by tapping gravitational potential energy of a BH[4]. The energy extraction is also possible from the rotational energy of a BH in the presence of sufficient amount of magnetic flux through Blandford and Znajek mechanism[5]. This is verified numerically in general relativistic magnetohydrodynamics framework for magnetically arrested disc (MAD) paradigm, where poloidal magnetic fields are playing the underlying role[6]. Also, vertically inflated toroidal fields can enhance outflow power through the formation of 'magnetic tower jets'[7]. Observational evidences of jetted sources also suggest the presence of dynamically important magnetic fields at the jet launching

region in the vicinity of a BH[8]. All these insights motivate us to model disc-outflow symbiosis in the presence of large-scale strong magnetic fields for hot, geometrically thick, advective accretion flow around BHs. Here the advection of both poloidal and toroidal components of magnetic fields occur. We compute the power associated with the magnetically driven accretion induced outflows in this proceeding.

## 2. Model equations

We propose a magnetized, advective, optically thin, geometrically thick disc-outflow/jet symbiotic model around BHs in the framework of pseudo-Newtonian potential. Here we use cylindrical coordinate assuming a steady and axisymmetric flow, i.e., $\partial/\partial t \equiv \partial/\partial\phi \equiv 0$. Hence the flow variables, namely, radial velocity ($v_r$), specific angular momentum ($\lambda$), vertical velocity ($v_z$), the radial ($B_r$), azimuthal ($B_\phi$), and vertical ($B_z$) components of magnetic field, fluid pressure ($p$), and mass density ($\rho$) are functions of both radial ($r$) and vertical ($z$) coordinates. Throughout in our computation, we express radial and vertical coordinates in units of $r_g = GM_{\rm BH}/c^2$, where $G$ is Newton's gravitational constant, $M_{\rm BH}$ is the mass of the BH, and $c$ is the speed of light. We also express other variables accordingly (e.g. velocity in units of $c$, $\lambda$ in units of $GM_{\rm BH}/c$ etc.) to make all the variables dimensionless. Hence, the continuity equation, the momentum balance equations, the equation for no magnetic monopole, the components of induction equation, and the energy equation are respectively

$$\frac{1}{r}\frac{\partial}{\partial r}\left(r\rho v_r\right) + \frac{\partial}{\partial z}\left(\rho v_z\right) = 0, \tag{1}$$

$$v_r\frac{\partial v_r}{\partial r} + v_z\frac{\partial v_r}{\partial z} - \frac{\lambda^2}{r^3} + \frac{1}{\rho}\frac{\partial p}{\partial r} + F = \frac{1}{4\pi\rho}\left[-\frac{B_\phi}{r}\frac{\partial}{\partial r}\left(rB_\phi\right) + B_z\left(\frac{\partial B_r}{\partial z} - \frac{\partial B_z}{\partial r}\right)\right], \tag{2}$$

$$v_r\frac{\partial \lambda}{\partial r} + v_z\frac{\partial \lambda}{\partial z} = \frac{r}{\rho}\left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 W_{r\phi}\right) + \frac{\partial W_{\phi z}}{\partial z}\right] + \frac{r}{4\pi\rho}\left[\frac{B_r}{r}\frac{\partial}{\partial r}\left(rB_\phi\right) + B_z\frac{\partial B_\phi}{\partial z}\right], \tag{3}$$

$$v_r\frac{\partial v_z}{\partial r} + v_z\frac{\partial v_z}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial z} + \frac{Fz}{r} = \frac{1}{4\pi\rho}\left[B_r\left(\frac{\partial B_z}{\partial r} - \frac{\partial B_r}{\partial z}\right) - B_\phi\frac{\partial B_\phi}{\partial z}\right], \tag{4}$$

$$\frac{1}{r}\frac{\partial}{\partial r}\left(rB_r\right) + \frac{\partial B_z}{\partial z} = 0, \tag{5}$$

$$\frac{\partial}{\partial z}\left[r\left(v_zB_r - v_rB_z\right)\right] = 0, \tag{6}$$

$$\frac{\partial}{\partial r}\left(v_rB_\phi - \frac{\lambda B_r}{r}\right) = \frac{\partial}{\partial z}\left(\frac{\lambda B_z}{r} - v_zB_\phi\right), \tag{7}$$

$$\frac{\partial}{\partial r}\left[r\left(v_zB_r - v_rB_z\right)\right] = 0, \tag{8}$$

$$\frac{v_r}{\Gamma_3 - 1}\left[\frac{\partial p}{\partial r} - \Gamma_1\frac{p}{\rho}\frac{\partial \rho}{\partial r}\right] + \frac{v_z}{\Gamma_3 - 1}\left[\frac{\partial p}{\partial z} - \Gamma_1\frac{p}{\rho}\frac{\partial \rho}{\partial z}\right] = Q^+ - Q^- = fQ^+, \tag{9}$$

where $F$ is the magnitude of the gravitational force for a rotating BH corresponding to the pseudo-Newtonian potential[9]. The generalized viscous stress tensor ($W_{ij}$) is written in terms of Shakura-Sunayev[10] $\alpha$-viscosity parameter with proper modifications[11] due to advection and is given by $W_{r\phi} = \alpha(p + \rho v_r^2)$ and $W_{\phi z} \approx \frac{z}{r}W_{r\phi}$[12]. We consider here the induction equation in the limit of very large magnetic Reynolds number. The energy generation rate per unit volume ($Q^+$) due to magnetic and viscous dissipations are followed from the Ref. 13. $Q^-$ represents the radiative cooling rate per unit volume through different cooling mechanisms[3]. However in this

proceeding, we do not consider cooling processes explicitly. The factor '$f$' included here to represent the degree of cooling which varies from 0 (for very efficient cooling) to 1 (for no cooling). In this disc-outflow symbiosis region, we also assume that within the disc flow region the vertical variation of any flow variables is much less compared to the corresponding radial variation. This allow us to introduce the vertical variation as $\partial/\partial z \approx s/z$, where the scale variable '$s$' is very small number ($|s| \sim 0.01$).

## 3. Solution procedure

The set of nine coupled differential equations $(1) - (9)$ for eight flow variables $v_r$, $\lambda$, $v_z$, $p$, $\rho$, $B_r$, $B_\phi$, and $B_z$, along with vertical scale height ($z$) are solved simultaneously to obtain the solutions as functions of independent variable $r$. The sub-sonic matter at very far away from the BH, passes through the sonic/critical location and becomes super-sonic near the event horizon of a BH for such BH accretion process. The outer boundary for our solution is the transition radius between Keplerian to sub-Keplerian flow. The event horizon where matter velocity reaches the speed of light, is the inner boundary. We also use the critical point as one of the boundaries. The different types of sonic/critical points, say, saddle, spiral, and nodal, are described in the Ref. 14.

## 4. Results

We address here the role of large-scale magnetic fields in the outward transport of angular momentum, as well as in the formation of outflows/jets for this hot advective disc-outflow symbiosis around BHs. Fig. 1 describes the variation of flow parameters, whereas Fig. 2 indicates the variation of magnetic field properties and outflow power as functions of $r$ along the disc-outflow surface for both non-rotating (solid line) and fast rotating (dotted line) BHs. Fig. 1($a$) describes the profile for Mach number ($M$) defined as the ratio of radial velocity to the sound speed. The critical location and the corresponding specific angular momentum are $r_c = 6.0$, $\lambda_c = 3.0$ for BHs with spin parameter $a = 0$, and $r_c = 4.2$, $\lambda_c = 1.8$ for BHs with $a = 0.998$, respectively. Very far away from the BH, matter is sub-sonic, whereas near the event horizon matter becomes super-sonic ($M > 1$). The outward transfer of angular momentum is shown in Fig. 1($b$). The vertical velocity and the sound speed are higher for rapidly spinning BHs as shown in Figs. 1($c$) and ($d$) respectively. The vertical velocity at the horizon increases from 0.065 when $a = 0$ to 0.145 when $a = 0.998$. The increase of $c_s$ with spin indicates that the temperature of the coupled disc-induced outflow is higher for rapidly spinning BHs.

The magnetic field components and their relative strength are shown in Figs. 2($a$) and ($b$) respectively. The advection of both the toroidal and poloidal field components occur in our computation. Very far away from the BH, the disc is toroidal field dominated due to presence of differential rotation. Within the plunging region,

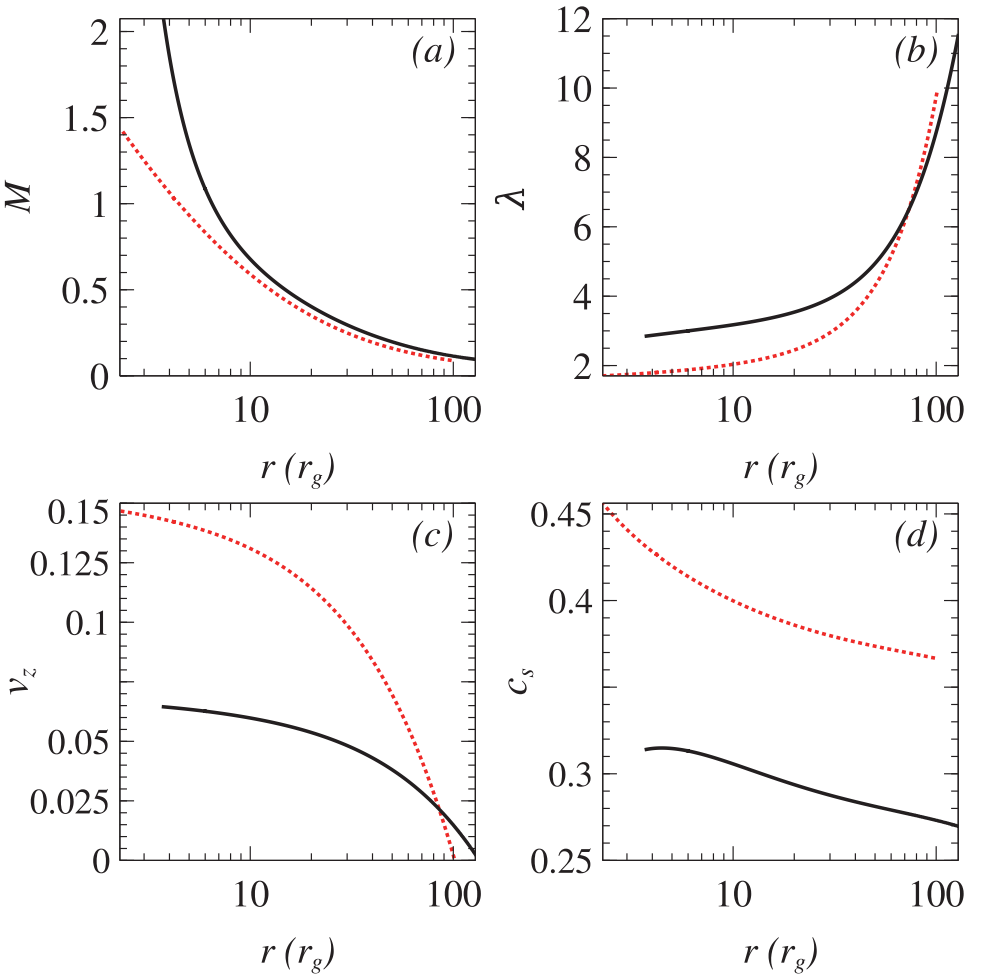



Fig. 1. The variation of ($a$) Mach number, ($b$) specific angular momentum, ($c$) vertical/outflow velocity, and ($d$) sound speed, as functions of $r$. Solid lines are for non-rotating ($a = 0$) BHs and dotted lines are for fast spinning ($a = 0.998$) BHs. The model parameters are $M_{\rm BH} = 20M_{\odot}$, $\dot{M} = 0.05\dot{M}_{\rm Edd}$, $f = 0.5$, and $\alpha = 0.01$.

the poloidal field enhances more compared to the toroidal one due to dominant inward advection and lack of differential rotation. Note that the maximum magnetic field strength near the BH is few factor times $10^7$ G for $20M_{\odot}$ BH. The ratios of the magnetic- to viscous-stresses are shown in Fig. 2($c$). The dominant nature of large-scale magnetic stress over viscous one indicates that the magnetic stresses are playing the underlying role in angular momentum transport, as well as in outflow formation. The corresponding outflow power is shown in Fig. 2($d$). The extracted power from the disc-outflow surface is the combination of mechanical, enthalpy, viscous, and Poynting parts[13]. The maximum reachable power in such configuration is $6 \times 10^{39}$ erg/s for BHs with $a = 0$, and $8.65 \times 10^{39}$ erg/s for BHs with $a = 0.998$.

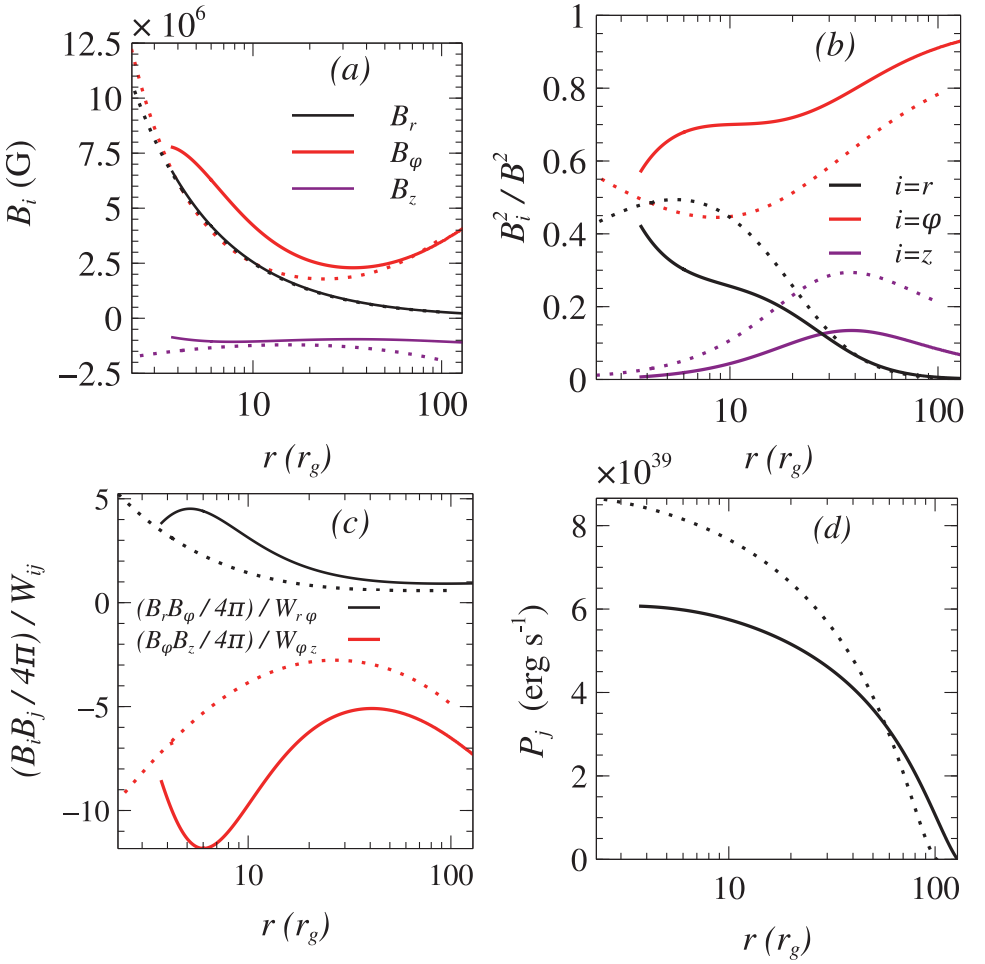



Fig. 2. The variation of ($a$) magnetic field components, ($b$) magnetic- to viscous-stress ratios, ($c$) relative magnetic field strength, and ($d$) the outflow power, as functions of $r$. Solid lines are for non-rotating ($a = 0$) BHs and dotted lines are for fast spinning ($a = 0.998$) BHs. The model parameters are same as in Fig 1 (colour online).

## 5. Discussions

In this proceeding, we discuss the disc flow behaviours and the formation of strong outflows from 2.5-dimensional magnetized, viscous, advective disc-outflow symbiotic model around BHs. The disc is hot puffed-up with (ion) temperature $\sim 10^{11} - 10^{12}$ K for non-rotating to fast-rotating BHs. The magnetic field strength near the stellar mass BHs is few factor times $10^7$ G, which is well below the Eddington magnetic field limit [13]. The components of magnetic stress dominate over the corresponding viscous ones, and hence the outward transfer of angular momentum is mainly controlled by large-scale magnetic stress. However, the origin of such strong magnetic

field in the vicinity of a BH is not clear yet. It was suggested that the externally generated fields are captured from environment, say, interstellar medium or companion star, and dragged inward with continuous accretion process[15]. The fields become dynamically dominant near the event horizon of a BH due to flux freezing in this quasi-spherical advective process. In our computation, the advection of both the toroidal and poloidal components of magnetic field occur.

We compute the power associated with magneto-centrifugally driven outflows from this magnetically dominated disc-outflow surface around BHs. Note that the effect of BHs' spins does not appear to play efficient role in such power measurement in this model. This is because we introduce the spin effect through the pseudo-Newtonian potential around rotating BHs. To operate Blandford and Znajek mechanism[5] properly, we have to formulate this disc-outflow/jet symbiosis in general relativistic magnetohydrodynamics framework. The observational implication to ultra-luminous X-ray sources is discussed in the other proceeding[16] of this volume based on such magnetically dominated disc-outflow symbiosis including cooling.

## Acknowledgments

TM acknowledges P. V. Lakshminarayana travel grant ODAA/INT/18/11, Office of Development and Alumni Affairs, Indian Institute of Science, Bengaluru, India.

## References

1. R. P. Fender, T. M. Belloni, E. Gallo, *MNRAS* **355**, 1105 (2004).
2. D. L. Meier, S. Koide, Y. Uchida, *Science* **291**, 84 (2001).
3. T. Mondal, B. Mukhopadhyay, *MNRAS* **486**, 3465 (2019).
4. R. D. Blandford, D. G. Payne, *MNRAS* **199**, 883 (1982).
5. R. D. Blandford, R. L. Znajek, *MNRAS* **179**, 433 (1977).
6. A. Tchekhovskoy, R. Narayan, J. C. McKinney, *MNRAS* **418**, L79 (2011).
7. D. Lynden-Bell, *MNRAS* **341**, 1360 (2003).
8. R. P. Eatough, et al., *Nature* **501**, 391 (2013).
9. B. Mukhopadhyay, *ApJ* **581**, 427 (2002).
10. N. Shakura, R. Sunyaev, *A&A* **24**, 337 (1973).
11. S. K. Chakrabarti, *ApJ* **464**, 664 (1996).
12. S. Ghosh, B. Mukhopadhyay, *RAA* **9**, 157 (2009).
13. T. Mondal, B. Mukhopadhyay, *MNRAS* **482**, L24 (2019).
14. T. Mondal, B. Mukhopadhyay, *MNRAS* **476**, 2396 (2018).
15. G. S. Bisnovatyi-Kogan, A. A. Ruzmaikin, *Ap&SS* **28**, 45 (1974).
16. B. Mukhopadhyay, *this volume.*

# Multi-accretion events from corotating and counterrotating SMBHs tori

Daniela Pugliese* and Zdenek Stuchlík

*Institute of Physics and Research Centre of Theoretical Physics and Astrophysics, Faculty of Philosophy&Science, Silesian University in Opava, Bezručovo náměstí 13, CZ-74601 Opava, Czech Republic*

*E-mail: d.pugliese.physics@gmail.com;zdenek.stuchlik@physics.cz*

Ringed accretion disks (**RADs**) are aggregates of corotating and counterrotating toroidal accretion disks orbiting a central Kerr super-massive Black Hole (**SMBH**) in **AGNs**. The dimensionless spin of the central **BH** and the fluids relative rotation are proved to strongly affect the **RAD** dynamics. There is evidence of a strict correlation between **SMBH** spin, fluid rotation and magnetic fields in **RADs** formation and evolution. Recently, the model was extended to consider **RADs** constituted by several magnetized accretion tori and the effects of a toroidal magnetic field in **RAD** dynamics have been investigated. The analysis poses constraints on tori formation and emergence of **RADs** instabilities in the phases of accretion onto the central attractor and tori collision emergence. Magnetic fields and fluids rotation are proved to be strongly constrained and influence tori formation and evolution in **RADs**, in dependence on the toroidal magnetic fields parameters. Eventually, the **RAD** frame investigation constraints specific classes of tori that could be observed around some specific **SMBHs** identified by their dimensionless spin

*Keywords*: Accretion; Accretion disks; Black holes; Active Galactic Nucleai (AGN)

## 1. Introduction

Ringed accretion disks (**RADs**) are aggregates of axi-symmetric tori orbiting Kerr **SMBHs**. The orbiting agglomerate is composed by both corotating and counterrotating tori. These configurations might be generated after different periods of the attractor life, where the angular momentum of the material in accretion on the central attractor in different phases of the **BH** life could have very different direction and magnitude. The **RAD** model was first developed considering tori centered on the equatorial plane of the central **BH**. **RAD** can be associated to jet emission in many ways, because of the inner edge of accreting disks jet launch point correlation and because the model adopted for each toroidal disk component of the aggregate provides an open solution, proto-jets, consisting of shells funnels of matter along the axis of the **BH**[2,3,5]. It should be noted that the occurrence of double accretion in the **RAD** is restricted to the case of two tori only, for a couple made by an inner corotating torus and by an outer counterrotating one. Consequently, the double jet shell emission in this system is limited to a jet from external counterrotating matter and an inner jet from the internal corotating matter[2]. It was also provided a general classification of the **RAD** and of the attractors according to their dimensional spin, providing an indication on the configurations and situation where observational evidences can be found. In Ref. 6 a large part of modellization was dedicated to describing the **RAD** as a single disk: in fact it has been shown that such kind od agglomerate is generically a geometrically thin disk, axis-symmetric, with knobby surface, and a very varied internal activity–Figs (1). The **RAD** could

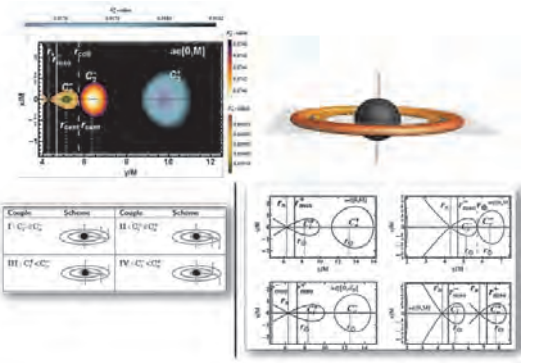


Fig. 1. Upper Left: **RAD** of the order 3, 2D-GRHD density profiles. Upper Right: **RAD** of the order 2, 3D-GRHD density profiles of corotating colliding tori. Below Left: schemes for the $\ell$corotating ($\ell$ **c**) or $\ell$counterrotating ($\ell$ **r**) couples. Below Right: correspondent **RAD** density profiles. Upper line is for the ($\ell$ **c**) of counterrotating (left) and corotating (right) tori [1,3].

be disguised as one disk characterized buy articulated phases of super-Eddington accretion, which could explain the masses problem of super-massive **BH** at high redshift, after a series of interrupted phase of accretion [1].

The model was introduced in Ref. 8 and fully developed in Ref. 6. In Ref. 5 the unstable configurations were analyzed and particularly constrains on locations of inner and outer edges of the **RAD** and the toroidal components were given. A detail analysis of **RAD** tori sequences as remnants of multiple accreting periods of Kerr **SMBHs**, in active galactic nuclei (**AGNs**) was considered in Ref. 3. In Ref. 4, the constraints on double configurations, constituted by two tori centered on the **BH** were given considering the possibility of collision emergence. Energetics associated to the **RAD** processes occurring in the ringed structure between its components and particularly the evaluation of **BH** accretion rates is presented in Ref. 1. Here we present the main results of the hydrodynamic (HD) and magnetized **RAD** based on the investigation in Refs 6, 5, 4, 3, 2, 1, 7.

**Overview** In Sec. (2) General considerations on hydrodynamic **RAD** models are discussed. In Sec. (3) constraints are discussed. In Sec. (4) we discuss the effects of toroidal magnetic fields on **RADs**. Concluding remarks follow in Sec. (5).

## 2. Orbiting Axi-symmetric tori in a Kerr spacetime

Because of the stationarity (time $t$ independence) and the axis-symmetry ($\phi$-independence) of each toroid of the aggregate, each torus is regulated by the Euler equation only with a barotropic equation of state $p = p(\rho)$, there is

$$\int_z^p \frac{dp}{p+\rho} = W(p) - W(0) = -\ln \frac{U_t}{(U_t)_{in}} + \int_{l_in}^{l} \frac{\Omega dl}{1-\Omega l}, \tag{1}$$

where $\Omega$ is the fluid relativistic angular frequency, $\ell$ specific angular momenta, here assumed constant and conserved inside each torus but not in the **RAD**, $(t, r, \phi, \theta)$

are Boyer-Lindquist (B-L) coordinates. We consider $\ell > 0$ for *corotating* and $\ell < 0$ for *counterrotating* fluids, within the notation $(\mp)$ respectively. $a = J/M \in ]0, M]$ is the **BH** spin, $J$ is the total angular momentum of the gravitational source and $M$ is the gravitational mass parameter. For the tori couple $(C_{(i)}, C_{(o)})$, orbiting in the equatorial plane of a given Kerr **BH** with specific angular momentum $(\ell_{(i)}, \ell_{(o)})$, we introduce the concept of $\ell$*corotating* ($\ell$**c**) disks, defined by the condition $\ell_{(i)}\ell_{(o)} > 0$, and $\ell$*counterrotating* ($\ell$**r**) disks defined by the relations $\ell_{(i)}\ell_{(o)} < 0$. The two ($\ell$**c**) tori can be both corotating, $\ell a > 0$, or counterrotating, $\ell a < 0$, with respect to the central attractor $a > 0$. Each **RAD** torus is a General Relativistic (GR) model based on the Boyer theory of the equi-pressure surfaces. Boyer surfaces are constant pressure surfaces and $\Sigma_i = \Sigma_j$ for $i, j \in (p, \rho, \ell, \Omega)$. Toroidal surfaces correspond to the equipotential surfaces, critical points of $V_{eff}(\ell)$ as function of $r$, thus solutions of $W: \ \ln(V_{eff}) = \mathrm{c} = \text{constant}$ or $V_{eff} = K =$constant **C**–cross sections of the closed Boyer surfaces (equilibrium quiescent torus); $\mathbf{C_x}$ –cross sections of the closed cusped Boyer surfaces (accreting torus); $\mathbf{O_x}$–cross sections of the open cusped Boyer surfaces, generally associated to proto-jet configurations. In the following we use the notation () to indicate a configuration which can be closed, $C$, or open $O$–Figs (1). The model in constructed investigating the angular momentum distribution inside the disk (which is not constant):

$$\pm\ell_n^{\mp} = \pm \left. \frac{a^3 + a r_n (3r_n - 4M) \mp \sqrt{r_n^3 \Delta_n^2}}{a^2 - (r_n - 2M)^2 r_n} \right|_{r_n^*}, \quad \Delta_n \equiv r_n^2 - 2M r_n + a^2 \qquad (2)$$

where $n$ is for the toroidal **RAD** component, and $\Delta_n$ is here a metric factor. In fact an essential part of the **RAD** analysis is the characterization of the boundary conditions on each torus in the agglomerate and of the disk. We considered also the function

$$\mathrm{K}^{\pm}(\Delta_{(\pm)}, r) \equiv \frac{1}{2} \left. \sqrt{\frac{r\left[(\Delta_- - \Delta_+)^2 + 4(r-2)r\right]}{2\Delta_-^2 - \Delta_-\Delta_+ r + r^3}} \right|_{\ell_n^{\mp}},$$
$$\ell = \frac{\Delta_- + \Delta_+}{2}, \ a = \frac{(\Delta_+ - \Delta_-)}{2}, \qquad (3)$$

(dimensionless units), provides constrains on the matter density distribution inside the disk and stitching together the **RAD** tori.

## 3. Constraints

In the Bondi quasi spherical accretion, the fluid angular momentum is everywhere smaller than the Keplerian one and therefore dynamically unimportant. In this analysis, however, we consider a full GR model for each **RAD** toroid where in fact there exists an extended region where the fluids angular momentum in the torus is larger (in magnitude) than the Keplerian (test particle) angular momentum. More precisely as first canvas model we adopt for each toroid a thick, opaque (high optical depth) and super-Eddington, radiation pressure supported accretion disk (in the

toroidal disks, pressure gradients are crucial) cooled by advection with low viscosity. As a consequence, during the evolution of dynamical processes, the functional form of the angular momentum and entropy distribution depends on the initial conditions of the system and on the details of the dissipative processes. From these considerations, using the distribution of relativistic specific angular momentum in the **RAD** as in Eq. (2) we can fix, as in Refs 6, 5, 4, 1, the constraints on the range of variation of the inner edge of accreting torus, $r_x$, and on the point of maximum density (pressure) in each torus, $r_{cent}$, in dependence from the range of variation of the specific angular momentum in the disk. Precisely, constraints on the angular momentum $\ell$ ranges are as follow:

**[Range- L1]** $\mp\mathbf{L1}^{\pm} \equiv [\mp\ell^{\pm}_{mso}, \mp\ell^{\pm}_{mbo}[$ where topologies $(C_1, C_\times)$ are possible, with accretion point in $r_\times \in ]r_{mbo}, r_{mso}]$ and center with maximum pressure $r_{cent} \in ]r_{mso}, \rho_{mbo}]$;
**[Range- L2]** $\mp\mathbf{L2}^{\pm} \equiv [\mp\ell^{\pm}_{mbo}, \mp\ell^{\pm}_{\gamma}[$ where topologies $(C_2, O_\times)$ are possible, with unstable point $r_j \in ]r_\gamma, r_{mbo}]$ and center with maximum pressure $r_{cent} \in ]\rho_{mbo}, \rho_\gamma]$;
**[Range- L3]** $\mp\mathbf{L3}^{\pm} \equiv \ \ell \geq \mp\ell^{\pm}_{\gamma}$ where only equilibrium torus $C_3$ is possible with center $r_{cent} > \rho_\gamma$; (in the following indices $i \in \{1, 2, 3\}$ refer to the ranges of angular momentum $\ell \in \mathbf{Li}$) being $mso$ =marginally stable orbit and $mbo$ =marginally bounded orbit, $\gamma$ =marginally circular orbit (photon orbit.) Alongside the geodesic structure of the Kerr spacetime represented by the set of radii $R \equiv (r^{\pm}_{mso}, r^{\pm}_{mbo}, r^{\pm}_{\gamma})$, and $r^{\pm}_{\mathcal{M}}$ solution of $\partial^2_r \ell = 0$, relevant to the location of the disk center and outer edge are $R_\rho \equiv (\rho^{\pm}_{mbo}, \rho^{\pm}_{\gamma}, \rho^{\pm}_{\mathcal{M}})$ ::

$$r^{\pm}_{\gamma} < r^{\pm}_{mbo} < r^{\pm}_{mso} < \rho^{\pm}_{\rm mbo} < \rho^{\pm}_{\gamma} \quad \text{where} \quad \rho^{\pm}_{\rm mbo} : \ \ell_{\pm}(r^{\pm}_{mbo}) = \ell_{\pm}(\rho^{\pm}_{\rm mbo}) \equiv \ell^{\pm}_{\mathbf{mbo}},$$
$$\rho^{\pm}_{\gamma} : \ell_{\pm}(r^{\pm}_{\gamma}) = \ell_{\pm}(\rho^{\pm}_{\gamma}) \equiv \ell^{\pm}_{\gamma}, \quad \rho^{\pm}_{\mathcal{M}} : \ell_{\pm}(\rho^{\pm}_{\mathcal{M}}) = \ell^{\pm}_{\mathcal{M}}. \tag{4}$$

This expanded structure rules good part of the geometrically thick disk physics and multiple structures. The presence of these radii stands as one of the main effects of the presence of a strong curvature of the background geometry [1,3,5].

### 3.1. *The RAD constraints*

The **RAD** of the order $n = 2$ (the order $n$ is the number of toroidal components) can be composed by the the following toridal couples:

**i)** $C^{\pm}_{\times} < C^{\pm}$, **ii)** $C^{+}_{\times} < C^{\pm}$, **iii)** $C^{-}_{\times} < C^{\pm}$ and **iv)** $C^{-}_{\times} < C^{+}_{\times}$. We indicated with $<$ the relative rotation of maximum density points in the tori–Figs (1). The situation concerning the emergence of more accretion points in **RAD** and presence of double accretion and screening tori, is summarized as follows:

$$\text{for} \quad a \in ]0, M[: \quad \underbrace{\mathbf{C^{-}_{x}}}_{\mathbf{(a)}} < \underbrace{... < C_{-} < ...}_{\mathbf{(b)}} < \underbrace{\mathbf{C^{+}_{x}}}_{\mathbf{(c)}} < ... \underbrace{< C_{\pm} <}_{\mathbf{(d)}} ... \tag{5}$$

where (**a**) is the inner ring in accretion, (**b**) is the inner subsequence of corotating tori in equilibrium, (**c**) is the outer accreting tori and (**d**) is the outer (mixed or separated) subsequence composed by equilibrium tori only. Further constraints (for the

specific angular momentum, elongation and number $n$ of tori) on the subsequences (**b**) and (**d**) depend on the attractor spin mass ratio–Figs (1). More precisely:

**The RAD canvas**

• There is a maximum $n = 2$ of accreting tori of the king $C_x^- < C_x^+$ around supermassive **BHs** with $a \neq 0$. Such tori are strongly constrained in their values of $\ell$ and $K$ depending on the spin-mass ratio of the attractor.

•"Screening"-tori located between the accreting disk and the central **BH** in a **RAD** sequence as $C_x^- < C^- < ... < C_x^+ < C^\pm$ are possible only as corotating, quiescent $C^-$ inner tori. A screening torus is a corotating (non-accreting) torus eventually detectable by X-ray spectra emission obscuration.

• If a counterrotating torus of a **RAD** is accreting onto the central **BH**, then a **RAD** with a corotating outer $C_1^-$ torus, is as $()_x^- < C_x^+ < C_1^- < C^\pm$, only orbiting "slow" **SMBHs** ($a < 0.46M$).

**[•-Corotating Tori]** A corotating torus can be the outer of a couple with an inner counterrotating accreting torus. Then the outer torus may be corotating (non accreting), or counterrotating in accretion or quiescent. Both the inner corotating and the outer counterrotating torus of the couple can accrete onto the attractor.

**[•-Counterrotating Tori]** A counterrotating torus can therefore reach the instability being the inner of a ($\ell$**r**)-($\ell$**c**), or the outer torus of a ($\ell$**r**) couple. If the *accreting* torus is *counterrotating*, $C_\times^+$, with respect to the Kerr attractor, there is *no* inner counterrotating torus, but there is $C_\times^+ < C^\pm$ or $()^- < C_\times^+$.

## 4. Influence of toroidal magnetic field in multi-accreting tori

We considered a toroidal magnetic field contribution in each **RAD** component where the magnetic field is[9]:

$$B^\phi = \sqrt{\frac{2p_B}{g_{\phi\phi} + 2\ell g_{t\phi} + \ell^2 g_{tt}}} \quad \text{with} \quad p_B = \mathcal{M}\left(g_{t\phi}g_{t\phi} - g_{tt}g_{\phi\phi}\right)^{q-1}\omega^q \tag{6}$$

the magnetic pressure, $\omega$ is the fluid enthalpy, $q$ and $\mathcal{M}$ (magnitude) are constant; $V_{eff}$ is a function of the metric and the angular momentum $\ell$. Euler equation (1) is modified by the term:

$$\partial_\mu \tilde{W} = \partial_\mu \left[\ln V_{eff} + \mathcal{G}\right] \quad \mathcal{G}(r,\theta) = \mathcal{S}\left(g_{t\phi}g_{t\phi} - g_{tt}g_{\phi\phi}\right)^{q-1}, \; \mathcal{S} \equiv \frac{q\mathcal{M}\omega^{q-1}}{q-1}, \tag{7}$$

(where $a \neq 0$) $q = 1$, is a singular values for the magnetic parameter $\mathcal{S}$, which is negative at $q < 1$, in this case excretion tori are possible. We here concentrate on $q > 1$. We therefore consider the equation for the $\tilde{W} = K$. For $\mathcal{S} = 0$ (or $\mathcal{M} = 0$) this reduces to the HD case. The **RAD** angular momentum distribution is:

$$\widetilde{\ell}^\mp \equiv \frac{a^3 + ar\left[4\mathcal{Q}(r-M)\mathcal{S}\Delta^{\mathcal{Q}} + 3r - 4\right] \mp \sqrt{r^3\Delta^2\left[1 + [2\mathcal{Q}(r-1)^2 r\mathcal{S}\Delta^{\mathcal{Q}-1}(2\mathcal{Q}\mathcal{S}\Delta^{\mathcal{Q}} + 1)]\right]}}{\Delta^{-1}[a^4 - a^2(r-3)(r-2)r - (r-2)r\left[2\mathcal{Q}(r-1)\mathcal{S}\Delta^{\mathcal{Q}+1} + (r-2)^2 r\right]]} \tag{8}$$

$\mathcal{Q} \equiv q - 1$ (dimensionless units). However the introduction of a toroidal magnetic field $B$, makes the study of the momentum distribution within the disk rather

complicated, instead In Ref. 7 it was adopted the function derived from $\mathcal{S}$-**RAD** parameter:

$$\mathcal{S}_{crit} \equiv -\frac{\Delta^{-\mathcal{Q}}}{\mathcal{Q}} \frac{a^2(a-\ell)^2 + 2r^2(a-\ell)(a-2\ell) - 4r(a-\ell)^2 - \ell^2 r^3 + r^4}{2r(r-1)\left[r(a^2-\ell^2) + 2(a-\ell)^2 + r^3\right]} \tag{9}$$

(dimensionless units) capable of setting the location of maximum density points in the disk and the existence and location of the instability points. Eq. (9) clearly enucleates the magnetic field contribution in the $\mathcal{Q}$ term, while interestingly highlights the role of the parameters $\ell$ versus $a$ [2,8]. The effects of the toroidal magnetic field in the **RAD** composition are evident in Fig. (2). In Ref. 7 it is noted that a **RAD** can be formed by having generally a sufficiently small parameter ($\mathcal{S}q$). Profiles of ($\ell$**c**) cases are similar independently by the corotation or counterrotation of the fluids in the **RAD** with the respect to the central Kerr **SMBH**. Generally the inner torus has maximum values of the $\mathcal{S}$ smaller the maximum $\mathcal{S}$ found in the outer tori. The most interesting results perhaps emerge in the case of ($\ell$**r**) couples where it is clear that the magnetic profiles for the couple $C_- < C_+$ (where double accretion occurs) are radically different from the case $C_+ < C_-$. The analysis shows also the importance of the coupling between the toroidal component of the magnetic field and the fluid angular momentum, particularly in the counter-rotating case, $\ell < 0$; for this case, for values $q < 1$ excretion can arise [7,10].

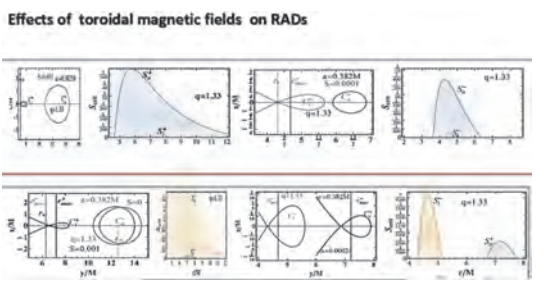



Fig. 2. Magnetized **RAD** of the order $n = 2$. Density profiles versus $\mathcal{S}$ functions profiles. Upper line features the $\ell$corotating couples of corotating (-) and counterrotating (+) tori. Bottom line features the $\ell$counterrtating couples made by $C_+ < C_-$ (left) or $C_- < C_+$ (right) [7].

## 5. Concluding Remarks and Future perspectives

The **RAD** investigation has revealed a very rich scenario with interesting phenomenology linked to the activity inside the agglomeration, correlated to the processes of instability and interaction between tori and the tori and central **SMBH** attractor. In this scenario constrains on the presence of screening and obscuring tori were discussed. The analysis of the screening and obscuring tori in particular could lead to observational evidences of a double tori **RAD** system from the emission

spectrum as X-ray emission screening, showing as fingerprints of the discrete radial profile of the **RAD**. In Ref. 11 then relatively indistinct excesses of the relativistically broadened emission-line components were predicted arising in a well-confined radial distance in the accretion structure originating by a series of episodic accretion events. Another interesting aspect of this model is the possibility of having inter-disk activity resolved in tori collision or double accretion phase with a double jet emission phase. From the **RAD** investigation this activity is limited to only two specific tori of aggregate made by a special couple constituted by a corotating and a counterrotating torus. As sideline result we provided a full characterization of the counterrotating tori in the multi-accreting systems. This model is designed for an extension to a dynamic GRMHD setup. Currently the toroidal model adopted to picture each **RAD** components is used as initial configuration for such systems. Another significant aspect is the possibility of inter-correlate the oscillations of the **RAD** components with **QPOs**. Generally instabilities of such configurations, can reveal to be of great significance for the High Energy Astrophysics related to accretion onto supermassive **BHs**, in Quasars and **AGNs**. Such activities could be targeted by the planned X-ray observatory `ATHENA`[a].

*Acknowledgments*

D.P. acknowledges support from the Junior GACR grant of the Czech Science Foundation No:16-03564Y.

## References

1. D. Pugliese and Z. Stuchlik, Class. Quant. Grav. **35**, 18, 185008 (2018).
2. D. Pugliese and Z. Stuchlik, Class. Quant. Grav. **35**, 10, 105005 (2018).
3. D. Pugliese, & Z. Stuchlík, JHEAp **17** 1-37, (2018).
4. D. Pugliese, &Z. Stuchlík, APJS **229** 2, 40, (2017).
5. D. Pugliese, &Z. Stuchlík, APJS, **223**, 2, 27, (2016).
6. D. Pugliese Z. &Z. Stuchlík, APJS, **221** 25 (2015).
7. D. Pugliese &G. Montani, MNRAS, **476**, 4, P4346-4361 (2018).
8. D. Pugliese, & G. Montani,Phys.Rev. D **91**, 8, 083011 (2015).
9. S. S. Komissarov, MNRAS, **368**, 993 (2006); K. Adamek & Z. Stuchlik, Class. Quantum Grav. **30**, 205007 (2013).; D. Pugliese &G. Montani, EPL, **101**, 1, 19001 (2013).
10. Z. Stuchlik, P. Slany, J. Kovar, Class. Quantum Grav., **26**, 215013 (2009).
11. V. Sochora, V. Karas, et al., MNRAS, **418**, 276-283 (2011) ; J. Schee &Z. Stuchlik, Gen. Rel. Grav., **41**, 1795 (2009). ;J. Schee &Z. Stuchlik, JCAP, **1304**, 005 (2013).

[a] `http://the-athena-x-ray-observatory.eu/`

# The effects of vertical self-gravity on neutrino-dominated accretion disks: A magnetized case

Narjes Shahamat

*Department of Physics, School of Science, Ferdowsi University of Mashhad, Mashhad, PO Box 91775-1436 Iran*

Shahram Abbassi

*Department of Physics, School of Science, Ferdowsi University of Mashhad, Mashhad, PO Box 91775-1436 Iran*
** E-mail:abbassi@um.ac.ir*

In this work we consider the effects of vertical self-gravity on a magnetized neutrino-dominated accretion disk, which is supposed to be a candidate for central engine of gamma-ray bursts (GRBs). We study some of the physical timescales that are considered to play a crucial role in the disk's late-time activity, such as viscous, cooling, and diffusion timescales. We are also interested to probe the emission of X-ray flares' probability, observed in GRBs' extended emission by an investigation on the "magnetic barrier" and "fragmentation". Our results approve the self-gravity as an amplifier for Blandford–Payne luminosity (BP power) and the magnetic field produced through the accretion process, but a suppressor for neutrino luminosity and magnetic barrier. The latter takes place as a result of the fragmentation enhancement in the outer disk, which is more likely to happen for the higher mass accretion rates.

*Keywords*: Gamma-ray Burst; Accretion Discs; Neutrino; Magnetic Field; Self-gravity.

## 1. Physical Model

### 1.1. *Basic Formalism*

We study a steady and axisymmetric magnetized NDAF ($\partial/\partial t = 0$, $\partial/\partial\phi = 0$) in which self-gravity has been taken into account, vertically. Considering the magnetic field influence in both large scale (magnetic braking mechanism) and small scale (viscous dissipation effects), yields the following results for the continuity and angular momentum equations

$$\dot{M} = -2\pi R\Sigma v_R = constant \tag{1}$$

$$\dot{M} = \frac{2\pi\alpha R^2 \Pi}{\Omega_k R^2 - j} + \frac{B_\phi B_z R^2}{\frac{\partial}{\partial R}(R^2\Omega_k)} \tag{2}$$

The energy balance equation reads

$$Q_{vis} = Q_{adv} + Q_\nu^- + Q_B^- \tag{3}$$

The neutrino cooling rate is expressed by a bridging formula (e.g., Di Matteo et al. [1]; Kohri et al. [3]) as follows:

$$Q_\nu = \sum_i \frac{(7/8)\sigma T^4}{(3/4)[\frac{\tau_{\nu_i}}{2} + \frac{1}{\sqrt{3}} + \frac{1}{3\tau_{a,\nu_i}}]} \tag{4}$$

For $Q_B$, we have

$$Q_B^- = 2R\Omega(B_\phi B_z/4\pi) \tag{5}$$

The equation of state (EOS) is written as (e.g., Di Matteo et al. [1]; Liu et al. [5])

$$p = p_{gas} + p_{rad} + p_{deg} + p_\nu \tag{6}$$

Last term denotes the neutrino pressure $P_\nu = \frac{u_\nu}{3}$, in which $u_\nu = \sum_i \frac{(7/8)aT^4(\tau_{\nu_i}/2+1/\sqrt{3})}{\tau_{\nu_i}/2+1/\sqrt{3}+1/(3\tau_{a,\nu_i})}$ is the neutrino-energy density (Di Matteo et al. [1]; Kohri et al. [3]).

Moreover, we consider the polytropic equation of state in vertical direction $p = K\rho^{4/3}$, where $K$ is a constant (Liu et al. [5]).

The hydrostatic balance equation in vertical direction reads

$$4\pi G\Sigma_z + \frac{\partial\Psi}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial z} + \frac{1}{8\pi\rho}(\frac{\partial B_\varphi^2}{\partial z} + \frac{\partial B_R^2}{\partial z}) - \frac{1}{4\pi\rho}B_R\frac{\partial B_z}{\partial R} = 0 \tag{7}$$

We adopt the pseudo-Newtonian potential, written by Paczynski and Wiita (Paczynski et al. [7]):

$$\Psi = \frac{-GM}{\sqrt{R^2+z^2}-R_g} \tag{8}$$

Regarding some mathematical considerations and relations between magnetic field components achieved by simulations (table 2 of Stone et al. [11]) gives

$$\begin{aligned} &2\pi G\rho_0^2H^2 + \Omega^2\frac{\rho_0H^2}{2} + p_0(e^{-1/2}-1) - 10^{-5}H\frac{dp_0}{dR} \\ &-2p_0(\frac{\beta_2}{1-\beta_2}\gamma_2e^{-2/3} + \frac{\beta_1}{1-\beta_1}\gamma_1) - \frac{10^{-5}p_0H}{R} \\ &+p_0(\frac{\beta_2}{1-\beta_2}e^{-2/3} + \frac{\beta_1}{1-\beta_1}) = 0 \end{aligned} \tag{9}$$

where $\beta$s and $\gamma$s are constants.

Furthermore, the magnetic viscosity equation is

$$\frac{B_RB_\phi}{4\pi} = -\frac{3}{2}\alpha p. \tag{10}$$

In order to achieve a better understanding of the effects of magnetic field and self-gravity through an analogy among the three cases: self-gravitating magnetized NDAF, magnetized case and self-gravitating one, we may get the required equations by ignoring the terms associated with the magnetic field and self-gravity in the above mentioned equations. For instance, considering the hydrostatic balance equation (7), the magnetized NDAF can be obtained via ignoring the first term (which reflects the self-gravity impact)

$$\frac{\partial\Psi}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial z} + \frac{1}{8\pi\rho}(\frac{\partial B_\varphi^2}{\partial z} + \frac{\partial B_R^2}{\partial z}) - \frac{1}{4\pi\rho}B_R\frac{\partial B_z}{\partial R} = 0. \tag{11}$$

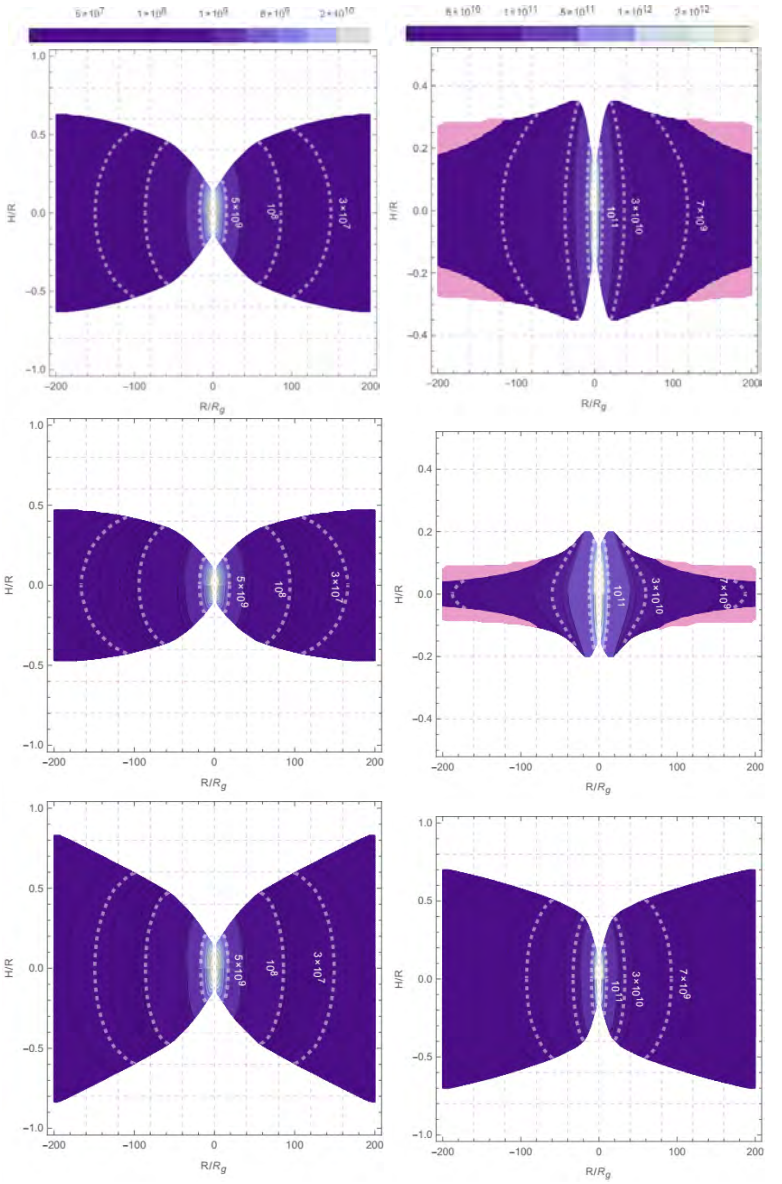



Fig. 1. Contours of density for $\dot{M} = 0.1 M_{\odot}/s$ (three left plots) and $\dot{M} = 10 M_{\odot}/s$ (three right plots), with three orders of density are highlighted through dashed lines ($5 \times 10^9$, $10^8$, $3 \times 10^7$ $gr/cm^3$ for the left panels, and $10^{11}$, $3 \times 10^{10}$, $7 \times 10^9$ $gr/cm^3$ for the right ones ). The two top panels are devoted to the self-gravitating magnetized NDAF, the tow middle plots are related to the self-gravitating case and the two bottom ones show the magnetized NDAF. Additionally, the shaded areas in pink show the gravitationally unstable zones.

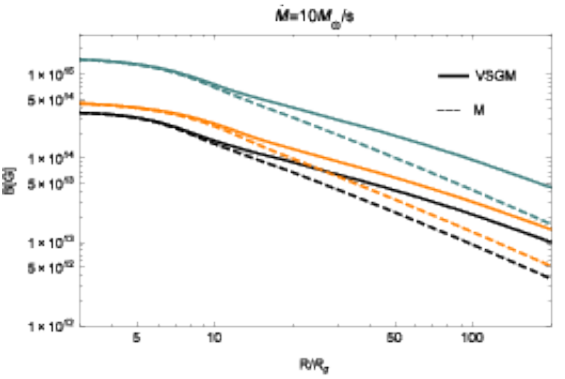



Fig. 2. Magnetic field components, $B_R, B_\phi, B_z$, for two cases of magnetized NDAF: with (VSGM) and without self-gravity (M) (solid and dashed curves). The radial, azimuthal and poloidal components are plotted in orange, blue and black, respectively.

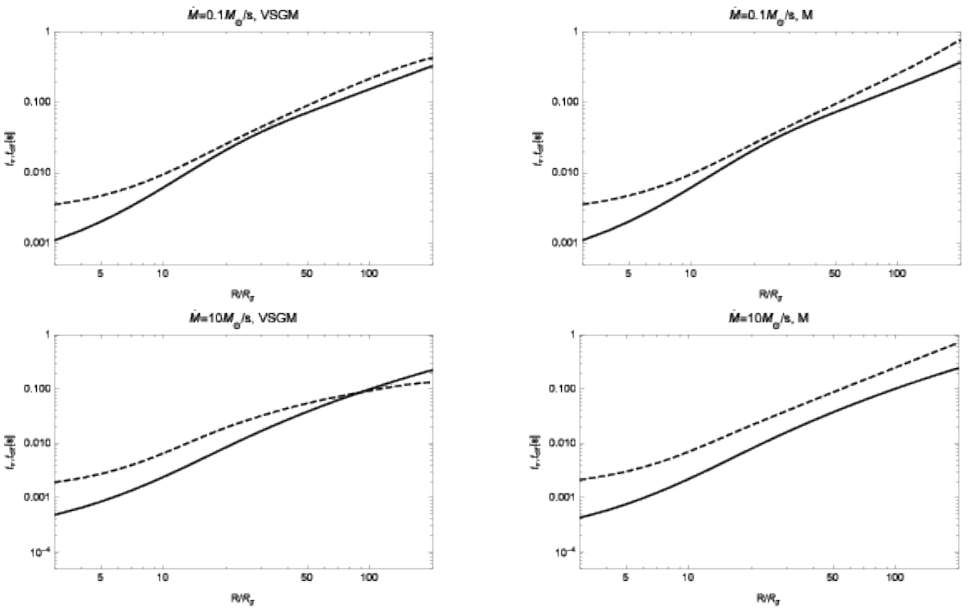



Fig. 3. Viscous and diffusion time scales for $\dot{M} = 0.1, 10 M_\odot/s$ illustrated by solid and dashed lines. Both magnetized (M) and vertically self-gravitating magnetized (VSGM) cases are considered.

The self-gravitating case is also achievable through an elimination of the magnetic terms

$$4\pi G\Sigma_z + \frac{\partial \Psi}{\partial z} + \frac{1}{\rho}\frac{\partial p}{\partial z} = 0. \tag{12}$$

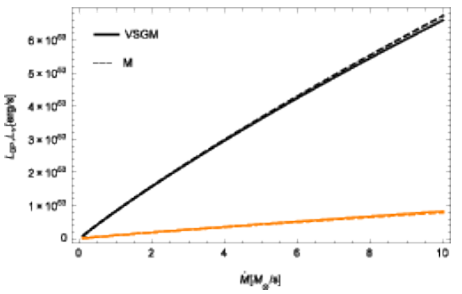



Fig. 4. Neutrino luminosity and BP power (black and orange lines) versus mass accretion rate. Solid lines represent the magnetized NDAF with vertical self-gravity (VSGM), but in dashed lines it is ignored (M).

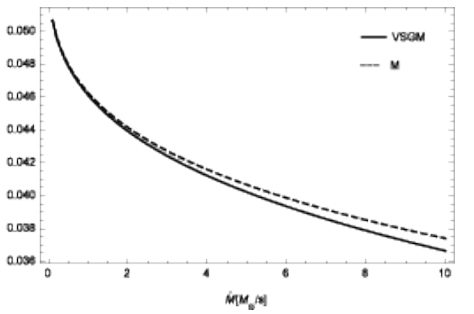



Fig. 5. Neutrino cooling efficiency versus mass accretion rate. Solid and dashed curves display the magnetized NDAF considering (VSGM) and ignoring self-gravity (M), respectively.

It can be written in the form of

$$2\pi G\rho_0^2 H^2 + \Omega^2 \frac{\rho_0 H^2}{2} + p_0(e^{-1/2} - 1) = 0. \tag{13}$$

Toomre parameter is as follows (e.g., Shu [10])

$$Q_M = Q(1 + \frac{\beta}{1-\beta}) \tag{14}$$

The magnetic field buoyancy and its rising time toward the disc surface can be estimated as

$$t_{dif} \approx \frac{H}{v_A} \tag{15}$$

Besides, the viscous time $t_\nu$ can present the timescale of the magnetic flux accumulating in the vicinity of the black-hole, then we have

$$t_\nu = \int_{3R_g}^{R} \frac{1}{v_R} dR \tag{16}$$

In case of $t_\nu < t_{diff}$ one may expect magnetic barrier to happen. In the gravitational context, fragmentation is possible if (Gammie [2]; Perna et al. [8])

$$t_{cool} < t_{cirt} \approx 3\Omega^{-1}, \tag{17}$$

where cooling timescale is denoted by $t_{cool} \approx (H/R)^2 t_\nu$ (Pringle [9]).

The neutrino luminosity, $L_\nu$, is expressed as

$$L_\nu = 2\pi \int_{R_{in}}^{R_{out}} Q_\nu R dR \tag{18}$$

The BP power output from a disc is equal to the power of disc magnetic braking and can be calculated as (Livio et al. [6]; Lee et al. [4])

$$L_{BP} = 2\pi \int_{R_{in}}^{R_{out}} Q_B R dR. \tag{19}$$

## 2. Conclusions

We study the structure and evolution of neutrino dominated accretion discs, in which the consideration of self-gravity and magnetic field provide us with a more realistic picture of these central engines of GRBs. We find self-gravity a booster of magnetic field, mainly in the outer disc. Such an effect, especially in higher accretion rates, enhances the BP power and descends the neutrino luminosity fraction. The latter is against Liu et al.[5] outcomes but seems a natural result of the strong magnetic field presence as we discussed formerly. On the other hand, the probable fragmentation process, lessens the magnetic barrier possibility by a decrease in magnetic field diffusion time scale in higher mass accretion rates. This result can also be deduced from the estimations, we have already made, and their comparison with the observational evidences.

On the whole, we find both MHD and neutrino processes effective enough ($10^{50-54} erg/s$) to produce GRBs' spectrum. Of course, in the case of higher accretion rates, the drop in neutrino efficiency (Fig. 5), as a result of a growth in neutrino opacity, may confirm a decrease in neutrinos capability to transport the energy outside. In the context of late time X-ray flares, the magnetic barrier process would be more probable to power such extended emissions in low accretion rates, because fragmentation is less likely to happen. Yet, this might not be the case for higher accretion rates as fragmentation can overcome magnetic barrier to produce energetic X-ray flares.

## References


1. Di Matteo, T., Perna, R., Narayan, R., 2002, ApJ, 579, 706.
2. Gammie, C.F., 2001, ApJ, 553, 174.
3. Kohri, K., Narayan, R., Piran, T., 2005, ApJ, 629, 341.
4. Lee H.K., Wijers, R.A.M.J., Brown G.E., 2000, Phys. Rev., 325, 83.
5. Liu, T., Yu, X.F., Gu, W.M., Lu, J.F., 2014, ApJ, 791, 69.
6. Livio M., Ogilvie G.I., Pringle J.E., 1999, ApJ, 512, 100.
7. Paczynsky, B., Wiita, P.J., 1980, A&A, 88, 23.
8. Perna, R., Armitage, P.J., Zhang, B., 2006, ApJ, 636L, 29.
9. Pringle, J.E., 1981, ARA&A, 19, 137.
10. Shu, F.H., 1992, The Physics of Astrophysics, Vol. 2 (Mill Valley: Univ. Science Books).
11. Stone J.M., Hawley J.F., Gammie C.F., Balbus, S.A., 1996, ApJ, 463, 656.

# Unified dark energy and dark matter from dynamical space time

David Benisty

*Frankfurt Institute for Advanced Studies (FIAS), Ruth-Moufang-Strasse 1, 60438 Frankfurt am Main, Germany,*
*Physics Department, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel*
*E-mail: benidav@post.bgu.ac.il*

Eduardo I. Guendelman

*Frankfurt Institute for Advanced Studies (FIAS), Ruth-Moufang-Strasse 1, 60438 Frankfurt am Main, Germany,*
*Physics Department, Ben-Gurion University of the Negev, Beer-Sheva 84105, Israel*
*Bahamas Advanced Study Institute and Conferences, 4A Ocean Heights, Hill View Circle, Stella Maris, Long Island, The Bahamas*
*E-mail: guendel@bgu.ac.il*

A unification of dark matter and dark energy based on a dynamical space time theory is discussed. By introducing a dynamical space time vector field $\chi_\mu$ as a Lagrange multiplier, a conservation of an energy momentum tensor $T^{\mu\nu}_{(\chi)}$ is implemented in addition to the conservation of the metric energy momentum tensor. This Lagrangian generalizes the Unified dark energy and dark matter from a scalar field different from quintessence which did not consider a Lagrangian formulation. This generalization allows the solutions which were found previously, in addition to a-non singular bouncing solutions that rapidly approach to the ΛCDM model. The dynamical time vector field exactly coincides with the cosmic time for the a ΛCDM solution and suffers a slight shift (advances slower) with respect to the cosmic time in the region close to the bounce for the bouncing non singular solutions. For some exponential potential which gives a possible interaction between DE and DM and could explain the coincidence problem.

*Keywords*: Unified Dark Energy and Dark Matter; Dynamical time

## 1. Introduction

Dark energy and Dark matter constitute most of the observable Universe. Yet the true nature of these two phenomena is still a mystery. One fundamental question with respect to those phenomena is the coincidence problem which is trying to explain the relation between dark energy and dark matter densities.

A model, which also continues to be valid after GW170817 event, for a unification of dark energy and dark matter from a single scalar field $\phi$, was suggested by Gao, Kunz, Liddle and Parkinson[1]. Their model is close to traditional quintessence, and gives dynamical dark energy and dark matter, but introduces a modification of the equations of motion of the scalar field that apparently are impossible to formulate in the framework of an action principle. The basic stress energy tensor which was considered in addition to Einstein equation was:

$$T^{\mu\nu} = -\frac{1}{2}\phi^{,\mu}\phi^{,\nu} + U(\phi)g^{\mu\nu} \tag{1}$$

where $\phi$ is a scalar field and $U(\phi)$ is the potential for that scalar. Assuming homogeneous and isotropic behavior the scalar field should be only time dependent $\phi = \phi(t)$.

Then the kinetic term $-\frac{1}{2}\phi^{,\mu}\phi^{,\nu}$ is parameterizing the dark matter because it contains only energy density with no pressure and $U(\phi)g^{\mu\nu}$ is parameterizing the dark energy. The basic requirement for this stress energy tensor is it's conservation law $\nabla_\mu T^{\mu\nu} = 0$. By assuming a constant potential $U(\phi) = \mathbf{Const}$ the model provides from the potential the traditional cosmological constant and the kinetic term of the scalar field is shown to provide, from the conservation law of the energy momentum tensor, that the kinetic term dependence has a dust like behavior.

$$-\frac{1}{2}\nabla_\mu(\phi^{,\mu}\phi^{,\nu}) = 0 \quad \Rightarrow \quad \dot{\phi}^2 \sim \frac{1}{a^3} \tag{2}$$

This simple case refers to the classical ΛCDM model. The special advantage of this model is a unification of dark energy and dark matter from one scalar field and has an interesting possibility for exploring the coincidence problem.

The lack of an action principle for this model brought us to reformulate the unification between dark energy and dark matter idea put forward by Gao, Kunz, Liddle and Parkinson[1] in the framework of a Dynamical Space Time Theory[2,3] which forces a conservation of energy momentum tensor in addition to the covariant conservation of the stress energy momentum tensor that appears in Einstein equation. In the next chapter we explore the equations of motion for these theories. In the third chapter we solve analytically the theory for constant potentials which reproduce the ΛCDM model with a bounce, which gives a possibility to solve the initial big bang singularity. In the last chapter we solve the theory for an exponential potential which gives a good possibility for solving the coincidence problem.

## 2. Dynamical Space Time Theory

One from the basic features in the standard approach to theories of gravity is the local conservation of an energy momentum tensor. In the field theory case it's derived as a result rather than a starting point. For example, the conservation of energy can be derived from the time translation invariance principle. The local conservation of an energy momentum tensor can be a starting point rather than a derived result. Let's consider a 4 dimensional case where a conservation of a symmetric energy momentum tensor $T^{\mu\nu}_{(\chi)}$ is imposed by introducing the term in the action:

$$\mathcal{S}_{(\chi)} = \int d^4x\sqrt{-g}\,\chi_{\mu;\nu}T^{\mu\nu}_{(\chi)} \tag{3}$$

where $\chi_{\mu;\nu} = \partial_\nu\chi_\mu - \Gamma^\lambda_{\mu\nu}\chi_\lambda$. The vector field $\chi_\mu$ called a dynamical space time vector, because of the energy density of $T^{\mu\nu}_{(\chi)}$ is a canonically conjugated variable to $\chi_0$, which is what we expected from a dynamical time:

$$\pi_{\chi_0} = \frac{\partial\mathcal{L}}{\partial\dot{\chi}^0} = T^0_0(\chi) \tag{4}$$

If $T^{\mu\nu}_{(\chi)}$ is being independent of $\chi_\mu$ and having $\Gamma^\lambda_{\mu\nu}$ being defined as the Christoffel connection coefficients (the second order Formalism), then the variation with respect

to $\chi_\mu$ gives a covariant conservation law:

$$\nabla_\mu T^{\mu\nu}_{(\chi)} = 0 \tag{5}$$

From the variation of the action with respect to the metric, we get a conserved stress energy tensor $G^{\mu\nu}$ (in appropriate units), which is well known from Einstein equation:

$$G^{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta\sqrt{-g}}{\delta g^{\mu\nu}}[\mathcal{L}_\chi + \mathcal{L}_m]\,, \quad \nabla_\mu G^{\mu\nu} = 0\,. \tag{6}$$

where $G^{\mu\nu}$ is Einstein tensor, $\mathcal{L}_\chi$ is the Lagrangian in (3) and $\mathcal{L}_m$ is an optional action that involve other contributions.

## 3. DE-DM Unified Theory from Dynamical Space-Time

A suggestion of an action which can produce DE-DM unification takes the form:[5]

$$\mathcal{L} = \frac{1}{2}R + \chi_{\mu;\nu}T^{\mu\nu}_{(\chi)} - \frac{1}{2}g^{\alpha\beta}\phi_{,\alpha}\phi_{,\beta} - V(\phi) \tag{7}$$

Consisting of an Einstein Hilbert action ($8\pi G = 1$), quintessence and Dynamical space-time action, when the original stress energy tensor $T^{\mu\nu}_{(\chi)}$ is the same as the stress energy tensor (1) Gao and colleagues used:

$$T^{\mu\nu}_{(\chi)} = -\frac{1}{2}\phi^{,\mu}\phi^{,\nu} + U(\phi)g^{\mu\nu} \tag{8}$$

The action depends on three different variables: the scalar field $\phi$, the dynamical space time vector $\chi_\mu$ and the metric $g_{\mu\nu}$. Therefore there are 3 sets in for the equation of motions. For the solution we assume homogeneity and isotropy, therefore we solve our theory with a FLRW metric:

$$ds^2 = -dt^2 + a(t)^2(\frac{dr^2}{1-Kr^2} + r^2 d\Omega^2) \tag{9}$$

According to this ansatz, the scalar field is just a function of time $\phi(t)$ and the dynamical vector field will be taken only with a time component $\chi_\mu = (\chi_0, 0, 0, 0)$, where $\chi_0$ is also just a function of time. A variation with respect to the dynamical space time vector field $\chi_\mu$ will force a conservation of the original stress energy tensor, which in FRWM gives the relation:

$$\ddot{\phi} + \frac{3}{2}\mathcal{H}\dot{\phi} + U'(\phi) = 0 \tag{10}$$

Compared with the equivalent equation which comes from quintessence model, this model gives a different and smaller friction term, as compared to the canonical scalar field. Therefore for increasing redshift, the densities for the scalar field will increase slower than in the standard quintessence.

The second variation with respect to the scalar field $\phi$ gives a non-conserved current:

$$\chi^\lambda_{;\lambda}U'(\phi) - V'(\phi) = \nabla_\mu j^\mu \tag{11a}$$

$$j^{\mu} = \frac{1}{2}\phi_{,\nu}(\chi^{\mu;\nu} + \chi^{\nu;\mu}) + \phi^{,\mu} \tag{11b}$$

and the derivatives of the potentials are the source of this current. For constant potentials the source becomes zero, and we get a covariant conservation of this current. In a FLRW metric this equation of motion takes the form:

$$\ddot{\phi}(\dot{\chi}_0 - 1) + \dot{\phi}[\ddot{\chi}_0 + 3\mathcal{H}(\dot{\chi}_0 - 1)] = U'(\phi)(\dot{\chi}_0 + 3\mathcal{H}\chi_0) - V'(\phi) \tag{12}$$

The last variation, with respect to the metric, gives the stress energy tensor that is defined by the value of the Einstein's tensor:

$$\begin{gathered} G^{\mu\nu} = g^{\mu\nu}(\frac{1}{2}\phi_{,\alpha}\phi^{,\alpha} + V(\phi) + \frac{1}{2}\chi^{\alpha;\beta}\phi_{,\alpha}\phi_{,\beta} + \chi^{\lambda}\phi_{,\lambda}U'(\phi)) \\ -\frac{1}{2}\phi^{,\mu}((\chi^{\lambda}_{;\lambda} + 2)\phi^{,\nu} + \chi^{\lambda;\nu}\phi_{,\lambda} + \chi^{\lambda}\phi^{\ ;\nu}_{;\lambda}) - \frac{1}{2}(\chi^{\lambda}\phi^{\ ;\mu}_{;\lambda}\phi^{,\nu} + \chi^{\lambda;\mu}\phi_{,\lambda}\phi^{,\nu}) \end{gathered} \tag{13}$$

For the specially homogeneous, cosmological case, the energy density and the pressure of the scalar field are:

$$\rho = (\dot{\chi}_0 - \frac{1}{2})\dot{\phi}^2 + V(\phi) \tag{14a}$$

$$p = \frac{1}{2}\dot{\phi}^2(\dot{\chi}_0 - 1) - V(\phi) - \chi_0\dot{\phi}U'(\phi) \tag{14b}$$

## 4. A bouncing ΛCDM solution

In order to compute the evolution of the scalar field and to check whether it is compatible with observable universe, we have to specify a form for the potentials. Let's take a simplified case of constant potentials:

$$U(\phi) = C, \quad V(\phi) = \Omega_{\Lambda} \tag{15}$$

Overall, in the equations of motions only the derivative the potential $U(\phi)$ appears, not the potential itself. Therefore a constant part of the potential $U(\phi)$ does not contribute to the solution. However $V(\phi)$, as we shall see below, gives the cosmological constant. The conservation of the stress energy tensor from equation (10) gives:

$$\dot{\phi}^2 = \frac{2\Omega_m}{a^3} \tag{16}$$

where $\Omega_m$ is an integration constant which appears from the solution. From the second variation, with respect to the scalar field $\phi$, a conserved current is obtained, which from equation (12) gives the exact solution of the dynamical time vector field:

$$\dot{\chi}_0 = 1 - \kappa\, a^{-1.5} \tag{17}$$

where $\kappa$ is another integration of constant. Eventually, the densities and the pressure for this potentials are given by (18). By substituting the solutions for the scalar $\dot{\phi}$ and the vector $\dot{\chi}_0$ (in units with $\rho_c = \frac{8\pi G}{3H_0^2} = 1$) we get:

$$\rho = \Omega_{\Lambda} - \frac{\Omega_{\kappa}}{a^{4.5}} + \frac{\Omega_m}{a^3} \quad , \quad p = -\Omega_{\Lambda} - \frac{1}{2}\frac{\Omega_{\kappa}}{a^{4.5}} \tag{18}$$

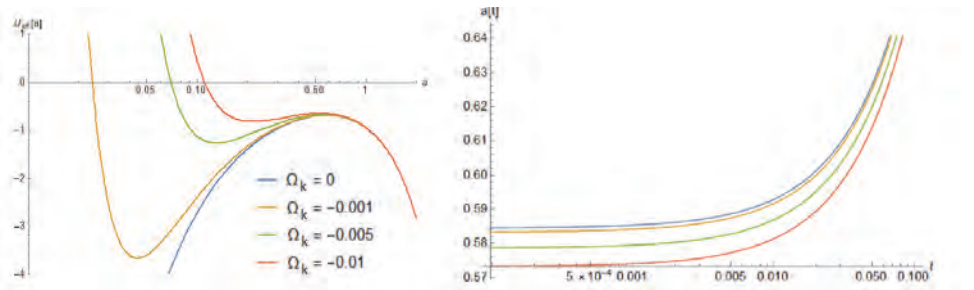



Fig. 1. Left hand side: Plot of the effective potential. For $\Omega_\kappa \neq 0$, there is a bouncing universe with dynamical dark energy. Right hand side: Plot of the scale parameter vs. the cosmic time. In any case $\dot{\chi}_0 \approx 1$

where $\Omega_\kappa = \kappa\Omega_m$. Notice that $\Omega_m, \Omega_\kappa$ are integration constants the solution contains and $\Omega_\Lambda$ is parameter from the action of the theory. We can separate the result into three different "dark fluids": dark energy ($\omega = -1$), dark matter ($\omega = 0$) and an exotic part ($\omega = \frac{1}{2}$), which is the responsible for the bounce (for $\kappa > 0$). From Eq. (16) the solution produces a positive $\Omega_m$ since it's proportional to $\dot{\phi}^2$. For $\Omega_\Lambda$ the measurements for the late universe forces the choice of this parameter to be positive. However for another solutions (in the context of Anti de-Sitter space, for instance) this parameter could be negative from the beginning. In Fig. 1 we can see the effective potential for different values of $\Omega_\kappa$. For $\Omega_\kappa = 0$ the solution returns to the known $\Lambda$CDM model. However for $\Omega_\kappa < 0$ we obtain a bouncing solution which also returns to the $\Lambda$CDM for late time expansion. The scale parameter evolution depicted in Fig.(1) can show us the initial conditions where $\dot{a}(t) = 0$, because at that point $a(t)$ is a minimum. In addition, for all cases the initial condition for the scale parameter is not zero $a(0) \neq 0$. These features imply a bouncing universe solutions.

In addition to those solutions, there is a strong correspondence between the zero component of the dynamical space time vector field and the cosmic time. For $\Lambda$CDM there is no bouncing solution $\kappa = 0$ and therefore from equation (17) we get $\chi_0 = t$ that implies that the dynamical time is exactly the cosmic time. For bouncing $\Lambda$CDM we obtain a relation between the dynamical and the cosmic time with some delay between the dynamical time and the cosmic time for the early universe (in the bouncing region). For the late universe the dynamical time returns back to run as fast as the cosmic time again. This relation between the dynamical and the cosmic time may have interesting application in the solution to "the problem of time" in quantum cosmology which will discussed elsewhere. Notice that the dynamical time is a field variable while the cosmic time is a coordinate.

## 5. Discussion

Here "unified dark energy and dark matter from a scalar field different from quintessence" is formulated through an action principle. Introducing the coupling

of a dynamical space time vector field to an energy momentum tensor that appears in the action, determines the equation of motion of the scalar field from the variation of the dynamical space time vector field or effectively from the conservation law of an energy momentum tensor, as in Ref.[1]. The energy momentum tensor that is introduced in the action is related but not in general the same as the one that appears in the right hand side of the gravitational equations, as opposed to the non-Lagrangian approach of[1], so our approach and that of[1] are not equivalent. In many situations the solutions studied in[1] can be also obtained here asymptotically. However, there are some solutions, as the non-singular bounce solutions, which are not predicted by[1].

In those solutions the dynamical time behaves as the cosmic time. In particular, an exact ΛCDM solution, the cosmic time and the dynamical time exactly the same. This correspondence could be solve the problem of time in quantum cosmology. Possible signatures for this model or for more generalized forms could be could be identified from cosmological perturbations theory. For instance, the perturbation for the scalar field is clear. However, the perturbation for the vector field could be represented with more degrees of freedom which can reproduce a different power spectrum for the Cosmic Microwave Background Anisotropies for instance. But more over than this, the model that was suggested in the last part was only with an exponential potential. However many combinations of potentials are applicable for testing the evolution for the energy densities, and using data fitting for those models. The benefits for this models are that they still preserve the speed of gravity equal to the speed of light, and also that arises from an action principle. Researching those families of solutions with more general potentials could help solve the coincidence problem.

## References


1. C. Gao, M. Kunz, A. R. Liddle and D. Parkinson, Phys. Rev. D **81**, 043520 (2010) doi:10.1103/PhysRevD.81.043520 [arXiv:0912.0949 [astro-ph.CO]].
2. E. I. Guendelman, Int. J. Mod. Phys. A **25**, 4081 (2010) doi:10.1142/S0217751X10050317 [arXiv:0911.0178 [gr-qc]].
3. D. Benisty and E. I. Guendelman, Mod. Phys. Lett. A **31**, no. 33, 1650188 (2016) doi:10.1142/S0217732316501881 [arXiv:1609.03189 [gr-qc]].
4. E. Guendelman, R. Herrera, P. Labrana, E. Nissimov and S. Pacheva, Gen. Rel. Grav. **47**, no. 2, 10 (2015) doi:10.1007/s10714-015-1852-1 [arXiv:1408.5344 [gr-qc]].
5. D. Benisty and E. I. Guendelman, Phys. Rev. D **98** (2018) no.2, 023506 doi:10.1103/PhysRevD.98.023506 [arXiv:1802.07981 [gr-qc]].
6. D. Benisty and E. I. Guendelman, Phys. Rev. D **98**, no. 4, 043522 (2018) doi:10.1103/PhysRevD.98.043522 [arXiv:1805.09314 [gr-qc]].

# Hamiltonian analysis in new general relativity

Daniel Blixt[1,*], Manuel Hohmann[1], Martin Krššák[1,2] and Christian Pfeifer[1]

[1] *Laboratory of Theoretical Physics, Institute of Physics, University of Tartu, Tartu, 50411, Estonia*

[2] *Center for Gravitation and Cosmology, College of Physical Science and Technology, Yangzhou University, Yangzhou 225009, China*

[*] *E-mail: blixt@ut.ee*

*http://kodu.ut.ee/˜blixt/*

It is known that one can formulate an action in teleparallel gravity which is equivalent to general relativity, up to a boundary term. In this geometry we have vanishing curvature, and non-vanishing torsion. The action is constructed by three different contractions of torsion with specific coefficients. By allowing these coefficients to be arbitrary we get the theory which is called "new general relativity". In this note, the Lagrangian for new general relativity is written down in ADM-variables. In order to write down the Hamiltonian we need to invert the velocities to canonical variables. However, the inversion depends on the specific combination of constraints satisfied by the theory (which depends on the coefficients in the Lagrangian). It is found that one can combine these constraints in 9 different ways to obtain non-trivial theories, each with a different inversion formula.

*Keywords*: Teleparallel gravity; New general relativity; ADM-variables.

## 1. Conventions

Greek indices denote global coordinate indices running from 0 to 3, small Latin indices are spatial coordinate indices running from 1 to 3, whereas capital Latin indices denote Lorentz indices running from 0 to 3. We are always dealing with Lorentzian metrics. Sign convention for the Minkowski metric is $\eta_{AB} = \mathrm{diag}(-1, 1, 1, 1)$.

## 2. Introduction

Gravity is conventionally described with the Levi-Civita connection which is induced by a pseudo-Riemannian metric. This means that the covariant derivative of the metric is zero, and the connection is torsion-free but has curvature. However, there are equivalent theories to general relativity[1]. We will focus on teleparallel gravity[2] where we have vanishing curvature, but non-vanishing torsion.

In particular we will perform the Hamiltonian analysis of "new general relativity" (NGR)[a]. For discussions of certain issues with these theories see[4–6] Previous work on the Hamiltonian analysis on teleparallel gravity theories have been performed in[6–18]. However, the full Hamiltonian analysis of NGR has not been performed. NGR is described by the following action:

$$S_{\mathrm{NGR}} = m_{Pl}^2 \int |\theta| \left( a_1 T^{\mu}{}_{\nu\rho} T_{\mu}{}^{\nu\rho} + a_2 T^{\mu}{}_{\nu\rho} T^{\rho\nu}{}_{\mu} + a_3 T^{\mu}{}_{\rho\mu} T^{\nu\rho}{}_{\nu} \right) \mathrm{d}^4 x, \tag{1}$$

[a] With NGR, we refer to the more general three-parameter teleparallel gravity in contrast to the special one-parameter teleparallel gravity theory which NGR originally referred to[3].

where $m_{Pl}$ is the Planck mass, $T^{\mu}{}_{\nu\rho} = \Gamma^{\mu}{}_{\rho\nu} - \Gamma^{\mu}{}_{\nu\rho}$ is the torsion component with $\Gamma^{\mu}{}_{\nu\rho} = e_A{}^{\mu}\partial_\rho\theta^A{}_{\nu} + e_A{}^{\mu}\left(\Lambda^{-1}\right)^A{}_D\,\partial_\rho\Lambda^D{}_B\theta^B{}_{\nu}$, with $\theta$ being the tetrad, $e$ its inverse and $\Lambda$ is a Lorentz matrix. Global spacetime indices are raised and lowered with $g_{\mu\nu} = \theta^A{}_{\mu}\theta^B{}_{\nu}\eta_{AB}$, while Lorentz indices are raised and lowered with $\eta_{AB}$. A theory equivalent to general relativity is obtained by setting $a_1 = \frac{1}{4},\ a_2 = \frac{1}{2},$ and $a_3 = -1$.

Alternatively, the NGR action can be written down in the so-called axial, vector, and tensor decomposition[19]. Then

$$S_{\rm NGR} = m_{\rm Pl}^2 \int |\theta|\left(c_1 T_{\rm ax} + c_2 T_{\rm ten} + c_3 T_{\rm vec}\right), \tag{2}$$

with $a_1 = -\frac{1}{3}(c_1 + 2c_2),\ a_2 = \frac{2}{3}(c_1 - c_2),\ a_3 = \frac{2}{3}(c_2 - c_3)$, and

$$\begin{aligned} T_{\rm vec} &= T^{\rho}{}_{\rho\mu}T_{\nu}{}^{\nu\mu}, \\ T_{\rm ax} &= \frac{1}{18}\left(T_{\rho\mu\nu}T^{\rho\mu\nu} - 2T_{\rho\mu\nu}T^{\mu\rho\nu}\right), \\ T_{\rm ten} &= \frac{1}{2}\left(T_{\rho\mu\nu}T^{\rho\mu\nu} + T_{\rho\mu\nu}T^{\mu\rho\nu}\right) - \frac{1}{2}T^{\rho}{}_{\rho\mu}T_{\nu}{}^{\nu\mu}\,. \end{aligned} \tag{3}$$

## 3. Method

In order to go from the Lagrangian to the Hamiltonian analysis we need to identify the velocities, derive the conjugate momenta and express everything in canonical variables. We may decompose the torsion scalar in the ADM variables[18] lapse $\alpha$, shift $\beta^i$ and the spatial components of the tetrad $\theta^A{}_i$:

$$\begin{aligned} \mathbb{T} = {} & \frac{1}{2\alpha^2}T^A{}_{i0}T^B{}_{j0}M^{i\ j}_{\ A\ B} \\ & + \frac{1}{\alpha^2}T^A_{i0}T^B_{kl}\left[M^{i\ l}_{\ A\ B}\beta^k + 2\alpha a_2 h^{il}\xi_B\theta_A{}^k + 2\alpha a_3 h^{il}\xi_A\theta_B{}^k\right] \\ & + \frac{1}{\alpha^2}T^A{}_{ij}T^B{}_{kl}\left[\frac{1}{2}M^{\ i\ k}_{A\ B}\beta^j\beta^l + 2\alpha a_2 h^{jl}\xi_A\theta_B{}^i\beta^k + 2\alpha a_3 h^{jl}\xi_A\theta_B{}^k\beta^i\right] + {}^3\mathbb{T}, \end{aligned} \tag{4}$$

where $h_{ij} = \theta^A{}_i\theta^B{}_j\eta_{AB}$ is the induced metric, which is used to raise and lower spatial indices, $\xi^A = -\frac{1}{6}\epsilon^A{}_{BCD}\theta^B{}_i\theta^C{}_j\theta^D{}_k\epsilon^{ijk}$ ,

$$M^{i\ j}_{\ A\ B} = -2a_1 h^{ij}\eta_{AB} + (a_2 + a_3)\xi_A\xi_B h^{ij} - a_2\theta_A{}^j\theta_B{}^i - a_3\theta_A{}^i\theta_B{}^j, \tag{5}$$

and

$$\begin{aligned} {}^3\mathbb{T} \equiv {} & a_1\eta_{AB}T^A{}_{ij}T^B{}_{kl}h^{ik}h^{jl} + a_2\eta_{AC}\theta^C{}_m h^{im}\eta_{BD}\theta^D{}_p h^{jp}T^A{}_{kj}T^B{}_{li}h^{kl} \\ & + a_3\eta_{AC}\theta^C{}_m h^{im}\eta_{BD}\theta^D{}_p h^{jp}h^{kl}T^A{}_{ki}T^B{}_{lj}. \end{aligned} \tag{6}$$

Without any loss of generality[7] we can restrict ourselves to the Weitzenböck gauge for which the torsion components are expressed as $T^A{}_{\mu\nu} = \partial_\nu\theta^A{}_\mu - \partial_\mu\theta^A{}_\nu$, and hence the conjugate momenta become,

$$\alpha\frac{\pi^i_A}{\sqrt{h}} = T^B_{j0}M^{i\ j}_{\ A\ B} + T^B_{kl}\left[M^{i\ l}_{\ A\ B}\beta^k + 2\alpha a_2 h^{il}\xi_B\theta_A{}^k + 2\alpha a_3 h^{il}\xi_A\theta_B{}^k\right]. \tag{7}$$

The velocities can now be inverted and expressed in canonical variables using

$$S_A^i = \dot{\theta}^B_{\ j} M^{i\ \ j}_{\ A\ B}, \tag{8}$$

with

$$\begin{aligned} S_A^i = & D_j \left(\alpha \xi^B + \beta^m \theta^B_{\ m}\right) M^{i\ \ j}_{\ A\ B} \\ & - T^B_{kl} \left[M^{i\ \ l}_{\ A\ B} \beta^k + 2\alpha a_2 h^{il} \xi_B \theta_A^{\ k} + 2\alpha a_3 h^{il} \xi_A \theta_B^{\ k}\right] + \alpha \frac{\pi_A^i}{\sqrt{h}}, \end{aligned} \tag{9}$$

where $D_i$ is the Levi-Civita covariant derivative with respect to the induced metric. However, $M$ in equation (8) is singular for certain combinations of parameters of the theory and can hence only be inverted by the Moore-Penrose pseudo-inverse matrix[12]. This is apparent if one decomposes the equation into irreducible representations of the rotation group, which generates the following constraints,

$$2a_1 + a_2 + a_3 =: {}^{\mathcal{V}}A = 0 \implies {}^{\mathcal{V}}C^i := S_A^i \xi^A = 0, \tag{10}$$

$$2a_1 - a_2 =: {}^{\mathcal{A}}A = 0 \implies {}^{\mathcal{A}}C_{ij} := S_A^k \theta^A_{\ [j} h_{i]k} = 0, \tag{11}$$

$$2a_1 + a_2 =: {}^{\mathcal{S}}A = 0 \implies {}^{\mathcal{S}}C_{ij} := S_A^k \theta^A_{\ (j} h_{i)k} - \frac{1}{3} S_A^k \theta^A_{\ k} h_{ij} = 0, \tag{12}$$

$$2a_1 + a_2 + 3a_3 =: {}^{\mathcal{T}}A = 0 \implies {}^{\mathcal{T}}C := S_A^i \theta^A_{\ i} = 0. \tag{13}$$

These are primary constraints, since these constrain both the tetrad field and their conjugate momenta, which also can be decomposed into irreducible parts. In the axial, vector, tensor decomposition we have that

$$2c_2 + c_3 \propto {}^{\mathcal{V}}A = 0, \tag{14}$$

$$2c_1 + c_2 \propto {}^{\mathcal{A}}A = 0, \tag{15}$$

$$c_2 \propto {}^{\mathcal{S}}A = 0, \tag{16}$$

$$c_3 \propto {}^{\mathcal{T}}A = 0. \tag{17}$$

In this language the primary constraints get some further geometrical meaning. Equations (14) and (15) together imposes the teleparallel equivalent to general relativity and impose invariance of the Lagrangian under pure tetrad local Lorentz transformations[9]. This is, however, not more apparent from the axial, vector, tensor decomposition we made. What is more interesting are the constraints imposed by equations (16) and (17). In this decomposition of the torsion scalar they exactly correspond to putting $T_{\text{ten}}$ and $T_{\text{vec}}$ to zero respectively.

## 4. Results

Different combinations of (10)-(13) yield 9 non-trivial classes of theories:

| Theory | Constraints | Location in figure 1 |
|---|---|---|
| $A_I \neq 0 \; \forall I \in \{\mathcal{V}, \mathcal{A}, \mathcal{S}, \mathcal{T}\}$ | No constraints | white area |
| $A_\mathcal{V} = 0$ | ${}^\mathcal{V}C_i = 0$ | red line |
| $A_\mathcal{A} = 0$ | ${}^\mathcal{A}C_{ji} = 0$ | black line |
| $A_\mathcal{S} = 0$ | ${}^\mathcal{S}C_{ji} = 0$ | vertical green line |
| $A_\mathcal{T} = 0$ | ${}^\mathcal{T}C = 0$ | horizontal blue line |
| $A_\mathcal{V} = A_\mathcal{A} = 0$ | ${}^\mathcal{V}C_i = {}^\mathcal{A}C_{ji} = 0$ | turquoise point |
| $A_\mathcal{A} = A_\mathcal{S} = 0$ | ${}^\mathcal{A}C_{ji} = {}^\mathcal{S}C_{ji} = 0$ | purple points (perimeter) |
| $A_\mathcal{A} = A_\mathcal{T} = 0$ | ${}^\mathcal{A}C_{ji} = {}^\mathcal{T}C = 0$ | orange point |
| $A_\mathcal{V} = A_\mathcal{S} = A_\mathcal{T} = 0$ | ${}^\mathcal{V}C_i = {}^\mathcal{S}C_{ji} = {}^\mathcal{T}C = 0$ | gray point (center) |

Any other solutions would be trivial ($c_1 = c_2 = c_3 = 0$). Excluding these trivial solutions we can normalize our parameters to

$$\tilde{c}_i = \frac{c_i}{\sqrt{c_1^2 + c_2^2 + c_3^2}}, \tag{18}$$

for $i = 1, 2, 3$, which means that we can make a 2-dimensional plot to visualize these theories in the normalized parameter-space. This can be nicely visualized in polar coordinates $(\theta, \phi)$ on the unit sphere with

$$\tilde{c}_1 = \cos\theta, \quad \tilde{c}_2 = \sin\theta\cos\phi, \quad \tilde{c}_3 = \sin\theta\sin\phi. \tag{19}$$

Every pair of antipodal points on the sphere corresponds to a ray in the 3-dimensional parameter space, whose elements describe the same theory. Hence, it suffices to display only the upper half sphere $\tilde{c}_1 \geq 0$, which is done in figure 1. However, note that points on the equator $\tilde{c}_1 = 0$ still appear twice, and both copies should be identified with each other. This applies in particular to the two purple points in figure 1, both describing the class of theories defined by pure vector torsion $\tilde{c}_1 = \tilde{c}_2 = 0$. The Hamiltonian is found to always appear with four Lagrange multipliers (linearity in lapse and shifts) with,

$$H = \alpha\mathcal{H}\left(\theta, M^{-1}\right) + \beta^k\mathcal{H}_k\left(\theta, M^{-1}\right) + D_i\left[\left(\alpha\xi^A + \beta^j\theta^A_{\ j}\right)\pi^i_A\right], \tag{20}$$

in the unconstrained case[7].

## 5. Discussion

One can distinguish 9 different classes of NGR theories by the presence or absence of primary constraints appearing in their Hamiltonian formulation. What remains to be determined is how many secondary constraints are induced by demanding closure of the constraint algebra. Some considerations in this direction have been studied in[6,18], however, our work invites for further investigation. The theories satisfying $A_I \neq 0, \forall\ I \in \{\mathcal{V}, \mathcal{A}, \mathcal{S}, \mathcal{T}\}$ can be parameterized by two free parameters (and a global rescaling of the Lagrangian, fixing the value of the Planck mass, which does not affect the presence or absence of primary constraints). Models which exhibit one primary constraint $A_I = 0$ have one free parameter left, while for those with

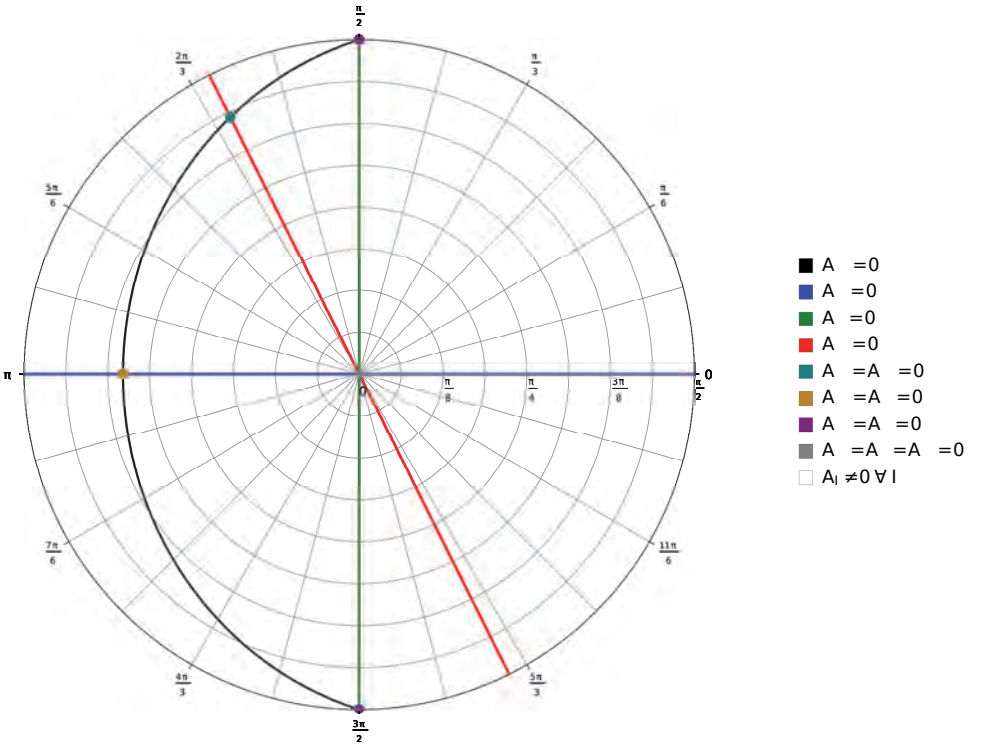



Fig. 1. Visualization of the parameter space of new general relativity in coordinates reflecting the axial, vector, tensor decomposition of the Lagrangian, colored by the occurrences of primary constraints. The radial axis shows the zenith angle $\theta$, while the (circular) polar axis shows the azimuth angle $\phi$, following the definition (19).

more primary constraints all parameters are fixed. The free parameters might affect the vanishing, or non-vanishing of certain Poisson brackets, which therefore have to be calculated in order to obtain the number of degrees of freedom.

The number of degrees of freedom can be compared with polarization modes in gravitational waves[20]. Furthermore, it can be compared with the linear level in order to find out if the theories are strongly coupled. One may extend this analysis to $f(T_{\rm ax}, T_{\rm ten}, T_{\rm vec})$[19] or include parity violating terms.

## References

1. J. Beltrán Jiménez, L. Heisenberg and T. S. Koivisto, "The Geometrical Trinity of Gravity," arXiv:1903.06830 [hep-th].
2. M. Krssak, R. J. Van Den Hoogen, J. G. Pereira, C. G. Boehmer and A. A. Coley, "Teleparallel Theories of Gravity: Illuminating a Fully Invariant Approach," arXiv:1810.12932 [gr-qc].
3. K. Hayashi and T. Shirafuji, "New General Relativity," Phys. Rev. D **19** (1979) 3524 Addendum: [Phys. Rev. D **24** (1982) 3312].

4. J. M. Nester, "Is there really a problem with the teleparallel theory?, Class. Quant. Grav. 5. (1988) no. 7, 10.1088
5. W. Kopczynski, "Problems with the metric-teleparallel theories of gravitation", J. Phys. A: Math. Gen. 15. (1982) 493
6. W. H. Cheng, D. C. Chern and J. M. Nester, "Canonical Analysis of the One Parameter Teleparallel Theory," Phys. Rev. D **38** (1988) 2656.
7. D. Blixt, M. Hohmann and C. Pfeifer, "Hamiltonian and primary constraints of new general relativity," Phys. Rev. D **99** (2019) no.8, 084025 [arXiv:1811.11137 [gr-qc]].
8. D. Blixt, M. Hohmann and C. Pfeifer, "On the gauge fixing in the Hamiltonian analysis of general teleparallel theories," arXiv:1905.01048 [gr-qc].
9. M. Blagojevic and I. A. Nikolic, "Hamiltonian structure of the teleparallel formulation of GR," Phys. Rev. D **62** (2000) 024021 [hep-th/0002022].
10. J. W. Maluf and J. F. da Rocha-Neto, "Hamiltonian formulation of general relativity in the teleparallel geometry," Phys. Rev. D **64** (2001) 084014.
11. J. W. Maluf, "Hamiltonian formulation of the teleparallel description of general relativity," J. Math. Phys. **35** (1994) 335.
12. R. Ferraro and M. J. Guzmán, "Hamiltonian formulation of teleparallel gravity," Phys. Rev. D **94** (2016) no.10, 104045 [arXiv:1609.06766 [gr-qc]].
13. M. Li, R. X. Miao and Y. G. Miao, "Degrees of freedom of $f(T)$ gravity," JHEP **1107** (2011) 108 [arXiv:1105.5934 [hep-th]].
14. R. Ferraro and M. J. Guzmán, "Hamiltonian formalism for f(T) gravity," Phys. Rev. D **97** (2018) no.10, 104028 [arXiv:1802.02130 [gr-qc]].
15. Y. C. Ong and J. M. Nester, "Counting Components in the Lagrange Multiplier Formulation of Teleparallel Theories," Eur. Phys. J. C **78** (2018) no.7, 568 [arXiv:1709.00068 [gr-qc]].
16. A. Okołów, "ADM-like Hamiltonian formulation of gravity in the teleparallel geometry: derivation of constraint algebra," Gen. Rel. Grav. **46** (2014) 1636 [arXiv:1309.4685 [gr-qc]].
17. A. Okołów, "ADM-like Hamiltonian formulation of gravity in the teleparallel geometry," Gen. Rel. Grav. **45** (2013) 2569 [arXiv:1111.5498 [gr-qc]].
18. A. Okolow and J. Swiezewski, "Hamiltonian formulation of a simple theory of the teleparallel geometry," Class. Quant. Grav. **29** (2012) 045008 [arXiv:1111.5490 [math-ph]].
19. S. Bahamonde, C. G. Böhmer and M. Krššák, "New classes of modified teleparallel gravity models," Phys. Lett. B **775** (2017) 37 [arXiv:1706.04920 [gr-qc]].
20. M. Hohmann, M. Krššák, C. Pfeifer and U. Ualikhanova, "Propagation of gravitational waves in teleparallel gravity theories," Phys. Rev. D **98** (2018) no.12, 124004 [arXiv:1807.04580 [gr-qc]].

# Towards non-singular metric solutions in ghost-free nonlocal gravity

Luca Buoninfante

*Dipartimento di Fisica "E.R. Caianiello", Università di Salerno, I-84084 Fisciano (SA), Italy*
*INFN - Sezione di Napoli, Gruppo collegato di Salerno, I-84084 Fisciano (SA), Italy*
*Van Swinderen Institute, University of Groningen, 9747 AG, Groningen, The Netherlands*
*E-mail: lbuoninfante@sa.infn.it*

In this short paper we study how black hole singularities can be tackled in the context of nonlocal ghost-free gravity, in which the action is characterized by the presence of non-polynomial differential operators containing infinite order covariant derivatives. The ghost-freeness condition can be preserved by requiring that such nonlocal operators are made up of exponential of entire functions, thus avoiding the emergence of extra unhealthy poles in the graviton propagator. We will mainly focus on how infinite order derivatives can regularize the singularity at the origin by making explicit computations in the linear regime. In particular, we will show that this kind of non-polynomial operators can smear out point-like distribution and that the Schwarzschild metric can not be a solution of the field equations in the ghost-free infinite derivative gravity.

*Keywords*: : Nonlocal gravity; higher-derivative gravity; quantum gravity; spacetime singularities; black holes.

## 1. Introduction

Gravitational theories whose actions are made up of terms quadratic in the curvature invariants, besides the Einstein-Hilbert piece, are perturbatively renormalizable[1], but incomplete due to the presence of a massive spin-2 ghost mode which makes the Hamiltonian unstable at the classical level classically, while from a quantum mechanical point of view breaks the unitarity condition[a].

The problem of ghosts in higher derivative theories has been known for long time[9], but only recently it was realized that the emergence of pathological degrees of freedom can be prevented when the order of the derivatives in the action is not finite but *infinite*. Indeed, by choosing certain non-polynomial analytic differential operators in the action one can preserve the number of physical degrees of freedom of the local theory, without introducing any unhealthy mode[10–13]. Having non-polynomial operators means that the order of derivatives is infinite and this makes the action *nonlocal*.

In Refs.[14–17], this kind of non-polynomial operators were used in a gravitational context to construct a quadratic curvature theory of gravity which is classically stable and unitary at the quantum level. It was noticed that the presence of nonlocality through infinite order derivatives can regularize infinities and many progresses ave been made in order to resolve black hole[15,16,18–29] and cosmological[14,30–33] singularities. At the quantum level, this class of ghost-free shows an improved UV behaviour[15,34–36] but the renormalizability has been shown only for

[a]See Refs.[2–6,8,8], and references therein, for astrophysical and cosmological applications involving quadratic curvature gravitational actions.

a peculiar subclass of non-polynomial operators[35]; see also Refs.[34,37–40] for nonlocal extensions of the standard model of particle physics, and Refs.[41,42] for applications beyond standard quantum mechanics.

In this manuscript we will review the main aspects regarding avoidance of singularities in ghost-free infinite derivative gravity, showing how infinite order derivatives can smear out point-like sources and, thus, regularize singularities. The paper is organized as follows: in Section 2, we introduce the gravitational action and, by working in the linear regime, we derive the form of the spacetime metric for a static point-like source; in Section 3 we show the regularizing properties of infinite derivative operators and explain why the Schwarzschild metric can not be a solution in ghost-free nonlocal gravity.

Throughout the paper we work in Natural Units, $\hbar = 1 = c$, and adopt the mostly positive metric convention, $\eta = \mathrm{diag}(-1, +1, +1, +1)$.

## 2. Nonlocal gravitational action and linearized metric solution

For simplicity we will only work with one of the simplest model of quadratic nonlocal gravity which captures all the relevant properties of nonlocality. Let us consider the following gravitational action[15,16,18]

$$S = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}\,\{\mathcal{R} + G_{\mu\nu}\mathcal{F}(\Box)\mathcal{R}^{\mu\nu}\}\,, \tag{1}$$

where $G_{\mu\nu} = \mathcal{R}_{\mu\nu} - g_{\mu\nu}\mathcal{R}/2$ is the Einstein tensor, $\kappa^2 \equiv 8\pi G$ with $G = 1/M_p^2$ being the Newton constant.

In local theories the form factor $\mathcal{F}(\Box)$ is a polynomial of $\Box$, while in the case of nonlocal theories it is a non-polynomial function of $\Box$. We consider one of the simplest example[16]:

$$\mathcal{F}(\Box) = \frac{1 - e^{-\Box/M_s^2}}{\Box} \tag{2}$$

which is analytic and modify the short-distance (UV) behavior of Einstein's GR as only positive power of $\Box$ appear when Taylor expanding; $M_s$ is needed to make the argument of the exponential dimensionless, while from a physical point of view it is a new fundamental energy scale at which nonlocal effects should manifest.

If we perturb the Minkowski metric by small a perturbation,

$$g_{\mu\nu} = \eta_{\mu\nu} + \kappa h_{\mu\nu}\,, \tag{3}$$

where $h_{\mu\nu}$ is the graviton perturbation, the action can be written up to order $\mathcal{O}(h_{\mu\nu}^2)$ as follows[16]:

$$S^{(2)} = \frac{1}{4}\int d^4x\, h_{\mu\nu}\, e^{-\Box/M_s^2}\mathcal{O}^{\mu\nu\rho\sigma}\, h_{\rho\sigma}, \tag{4}$$

where

$$\begin{aligned}\mathcal{O}^{\mu\nu\rho\sigma} \equiv{}& \frac{1}{4}\left(\eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho}\right)\Box - \frac{1}{2}\eta^{\mu\nu}\eta^{\rho\sigma}\Box \\ &+\frac{1}{2}\left(\eta^{\mu\nu}\partial^\rho\partial^\sigma + \eta^{\rho\sigma}\partial^\mu\partial^\nu - \eta^{\mu\rho}\partial^\nu\partial^\sigma - \eta^{\mu\sigma}\partial^\nu\partial^\rho\right)\end{aligned} \tag{5}$$

is a four-rank operator $\mathcal{O}^{\mu\nu\rho\sigma}$ which is totally symmetric in all its indices. By inverting the kinetic operator $e^{-\Box/M_s^2}\mathcal{O}^{\mu\nu\rho\sigma}$ one can obtain the graviton propagator around Minkowski; see Refs. [10,13,15,16,43,44] for details. While, by variating the action in Eq. (4) we obtain the linearized field equations, which we wish to solve in presence of a Delta Dirac distribution placed at $r = 0$, i.e. the stress-energy tensor is given by $T_{\mu\nu} = m\delta^0_\mu\delta^0_\nu\delta^{(3)}(\vec{r})$. By working in the conformal Newtonian gauge, we can write the perturbed metric (3) in isotropic coordinates as follows:

$$ds^2 = -(1 + 2\Phi(r))dt^2 + (1 - 2\Psi(r))(dr^2 + r^2 d\Omega^2), \tag{6}$$

where $r = \sqrt{x^2 + y^2 + z^2}$ is the isotropic radial coordinate, so that the field equations reads

$$e^{-\nabla^2/M_s^2}\nabla^2\Phi = e^{-\nabla^2/M_s^2}\nabla^2\Psi = 4\pi G\delta^{(3)}(\vec{r}) \tag{7}$$

where $T \equiv \eta^{\mu\nu}T_{\mu\nu}$, and we have used $\kappa h_{00} = -2\Phi$, $\kappa h_{ij} = -2\Psi\delta_{ij}$, $\kappa h = 2(\Phi - 3\Psi)$ and $\Box \simeq \nabla^2$.

The infinite derivative differential equation in Eq. (7) can be easily solved by going to Fourier space and then anti-transforming back to coordinate space:

$$\Phi(r) = \Psi(r) = -4\pi Gm\int d^3k\frac{e^{-k^2/M_s^2}}{k^2}e^{i\vec{k}\cdot\vec{r}} = -\frac{Gm}{r}\mathrm{Erf}\left(\frac{M_s r}{2}\right), \tag{8}$$

where $\mathrm{Erf}(x) := \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt$ is the so called error-function.

The gravitational potential in Eq. (8) turns out to be non-singular at $r = 0$, $\Phi(0) = 2GmM_s/\sqrt{\pi} < \infty$; while for large distances it reduces to Newton's law $\Phi(r) \sim -Gm/r$, as expected. From a physical point of view, nonlocality is able to regularize the point-like source at $r = 0$ smearing it out on a region of size $1/M_s$ thanks to the infinite order derivatives. Also, the linearized curvature invariants turn out to be non-singular at $r = 0$; in particular, the metric becomes conformally-flat, since all the components of the Weyl tensor vanish at the origin [23].

## 3. Towards non-linear and non-singular metric solutions

The metric solution in Eq. (8) is linear and says nothing about the non-linear regime where singularity may still appear. We now wish to move towards the non-linear regime and explain why nonlocality may help also in this case, in particular we want to understand whether the Schwarzschild singularity is admissible in ghost-free nonlocal gravity.

First of all, let us remind that in Einstein's general relativity the Schwarzschild metric, whose components contain the mass $m$, is derived by imposing a boundary condition at the origin, i.e. by putting a delta-Dirac distribution at $r = 0$, which means that on the right-hand side of the Einstein field equations we have vacuum everywhere except at $r = 0$ [45]. Indeed, strictly speaking the Schwarzschild metric is not a vacuum solution of the Einstein's field equation.

To check whether the Schwarzschild metric is a solution in ghost-free infinite derivative gravity we need to substitute $\mathcal{R} \sim \delta^{(3)}(\vec{r})$ and $\mathcal{R}_{\mu\nu} \sim \delta^{(3)}(\vec{r})$ in the full non-linear field equations[18]. In such a case, we do not obtain exactly zero on the left-hand side but we will have some complicated expression involving derivatives of delta. For example, in a fourth order theory of gravity one would generate up to third order derivatives of the delta-Dirac, so that the Schwarzschild metric is still an exact solution which couples to a more complicated point-source which is not just a simple delta, and is not positive definite, which means that it is not a physical solution. We can understand this feature by looking at the linear regime for the action in Eq. (1). The modified Poisson equation for the gravitational potential would be of the following kind:

$$[1 - \mathcal{F}(\nabla^2)\nabla^2]\nabla^2\Phi(r) = 4\pi G\rho(r). \tag{9}$$

In fourth order gravity we have a constant form-factor $\mathcal{F}(\Box) = -\alpha$, with $\alpha > 0$, from which it is clear that there always exists a source which generates the Schwarzschild solution, $\Phi(r) \sim 1/r$, and such a source is given by

$$\rho(r) = \frac{1}{4\pi G}[1 + \alpha\nabla^2]\nabla^2\left(-\frac{Gm}{r}\right) = -\frac{m}{2\pi r}\left[\delta'(r) + \alpha\,\delta'''(r)\right] \tag{10}$$

which is not positive defined. A similar scenario will also hold for any local higher derivative theory of gravity: for any gravitational theory with a polynomial form factor $\mathcal{F}$ there exist always a source which generates a the $1/r$ Schwarzschild singularity. For instance, in sixth order gravity, $\mathcal{F}(\Box) = -\alpha\Box$, the $1/r$ potential is generated by the density distribution $\rho(r) = -m/(2\pi r)[\delta'(r) + \alpha\,\delta^{(5)}(r)]$.

In ghost-free infinity derivative gravity, the form factor is a non-polynomial function, therefore one has infinite order derivatives acting on $1/r$, or in other words, by looking at the full field equations they will act on the Schwarzschild curvature $\mathcal{R} \sim \delta^{(3)}(\vec{r})$. Unlike the local scenario in which we have a finite number of terms, nonlocal differential operators would generated a complicated expression on the left-hand side involving infinite order derivatives of the delta-Dirac, which generically would not correspond to a point-like source but to an object with a non-point support.

A-priori the Schwarzschild metric could be still a solution but coupled to a source which is not point-like. However, we argue that in infinite derivative gravity there exist *no* source which can couple to the Schwarzschild metric. We have a strong hint coming from the linear regime in favour of our claim. Indeed, by looking at the modified Poisson equation in Eq. (9) with the form factor in Eq. (2) and using the method of Fourier transform, we obtain:

$$\rho(r) = \frac{1}{4\pi G}e^{-\nabla^2/M_s^2}\left(-\frac{Gm}{r}\right) = \frac{1}{2\pi^2 r}\int_0^\infty dk\,k\sin(k\,r)\,e^{+k^2/M_s^2} = \infty. \tag{11}$$

From the last equation, we can notice that, at least in the linear regime, it is impossible to define any source $\rho(r)$ which can generated the Schwarzschild metric

with $1/r$ behaviour. In fact, the only option could be $\rho(r) \sim e^{-\nabla^2/M_s^2}\delta(\vec{r})$, but it turns out that it cannot be mathematically defined since $e^{-\nabla^2/M_s^2}\delta(\vec{r}) = \infty$.

Hence, non-locality through infinite order derivatives does not allow the Schwarzschild metric as an exact solution, and this is a very strong hint which allow us to claim that the same feature will be also present in the full non-linear regime[25]; see also Ref.[46] for further developments regarding the non-linear level.

**Acknowledgments**

The author is grateful to A. S. Koshelev, G. Lambiase, J. Marto and A. Mazumdar for the fruitful collaboration and all the discussions on the subject.

**References**

1. K. S. Stelle, Phys. Rev. D **16**, 953 (1977); Gen. Rel. Grav. **9**, 353 (1978).
2. A. A. Starobinsky, Sov. Astron. Lett. **9**, 302 (1983).
3. S. Capozziello, G. Lambiase, M. Sakellariadou and An. Stabile, Phys. Rev. D **91**, 044012 (2015).
4. G. Lambiase, M. Sakellariadou, A. Stabile and An. Stabile, JCAP **1507**, 003 (2015).
5. G. Lambiase, M. Sakellariadou and A. Stabile, JCAP **1312**, 020 (2013).
6. G. Lambiase, A. Stabile and An. Stabile, Phys. Rev. D **95**, 084019 (2017). L. Buoninfante, G. Lambiase, L. Petruzziello and A. Stabile, Eur. Phys. J. C **79**, no. 1, 41 (2019).
7. M. Blasone, G. Lambiase, L. Petruzziello and A. Stabile, Eur. Phys. J. C **78**, no. 11, 976 (2018).
8. M. Blasone, S. Capozziello, G. Lambiase and L. Petruzziello, Eur. Phys. J. Plus **134**, 169 (2019).
9. M. Ostrogradsky, Mem. Acad. St. Petersbourg **6**, no. 4, 385 (1850).
10. N. V. Krasnikov, Theor Math. Phys. 73 1184, 1987, Teor. Mat. Fiz. 73, 235 (1987).
11. Yu. V. Kuzmin, Yad. Fiz. 50, 1630-1635 (1989).
12. J. W. Moffat, Phys. Rev. D 41, 1177 (1990). D. Evens, J. W. Moffat, G. Kleppe and R. P. Woodard, Phys. Rev. D 43, 499 (1991).
13. E. T. Tomboulis, hep-th/9702146.
14. T. Biswas, A. Mazumdar and W. Siegel, JCAP **0603**, 009 (2006).
15. L. Modesto, Phys. Rev. D **86**, 044005 (2012).
16. T. Biswas, E. Gerwick, T. Koivisto and A. Mazumdar, Phys. Rev. Lett. **108**, 031101 (2012).
17. T. Biswas, A. S. Koshelev and A. Mazumdar, Fundam. Theor. Phys. **183**, 97 (2016). T. Biswas, A. S. Koshelev and A. Mazumdar, Phys. Rev. D **95**, no. 4, 043533 (2017).
18. T. Biswas, A. Conroy, A. S. Koshelev and A. Mazumdar, Class. Quant. Grav. **31**, 015022 (2014), Erratum: [Class. Quant. Grav. **31**, 159501 (2014)].

19. J. Edholm, A. S. Koshelev and A. Mazumdar, Phys. Rev. D **94**, no. 10, 104033 (2016).
20. V. P. Frolov, A. Zelnikov and T. de Paula Netto, JHEP **1506**, 107 (2015)
21. V. P. Frolov, Phys. Rev. Lett. **115**, no. 5, 051102 (2015).
22. A. S. Koshelev and A. Mazumdar, Phys. Rev. D **96**, no. 8, 084069 (2017).
23. L. Buoninfante, A. S. Koshelev, G. Lambiase and A. Mazumdar, JCAP **1809** (2018) no.09, 034.
24. A. S. Koshelev, J. Marto and A. Mazumdar, Phys. Rev. D **98** (2018) no.6, 064023.
25. L. Buoninfante, A. S. Koshelev, G. Lambiase, J. Marto and A. Mazumdar, JCAP **1806** (2018) no.06, 014.
26. L. Buoninfante, G. Harmsen, S. Maheshwari and A. Mazumdar, Phys. Rev. D **98** (2018) no.8, 084009.
27. L. Buoninfante, A. S. Cornell, G. Harmsen, A. S. Koshelev, G. Lambiase, J. Marto and A. Mazumdar, Phys. Rev. D **98**, no. 8, 084041 (2018).
28. E. Kilicarslan, Phys. Rev. D **98**, no. 6, 064048 (2018).
29. L. Buoninfante and A. Mazumdar, arXiv:1903.01542 [gr-qc].
30. T. Biswas, T. Koivisto and A. Mazumdar, JCAP **1011**, 008 (2010).
31. T. Biswas, A. S. Koshelev, A. Mazumdar and S. Y. Vernov, JCAP **1208**, 024 (2012).
32. A. S. Koshelev and S. Y. Vernov, Phys. Part. Nucl. **43**, 666 (2012).
33. A. S. Koshelev, K. Sravan Kumar and A. A. Starobinsky, JHEP **1803**, 071 (2018).
34. T. Biswas and N. Okada, Nucl. Phys. B **898**, 113 (2015).
35. L. Modesto and L. Rachwal, Nucl. Phys. B **889**, 228 (2014).
36. S. Talaganis, T. Biswas and A. Mazumdar, Class. Quant. Grav. **32**, no. 21, 215017 (2015).
37. A. Ghoshal, A. Mazumdar, N. Okada and D. Villalba, Phys. Rev. D **97**, no. 7, 076011 (2018).
38. L. Buoninfante, G. Lambiase and A. Mazumdar, Nucl. Phys. B **994**, 114 (2019).
39. L. Buoninfante, A. Ghoshal, G. Lambiase and A. Mazumdar, Phys. Rev. D **99**, no. 4, 044032 (2019).
40. L. Buoninfante, G. Lambiase and M. Yamaguchi, arXiv:1812.10105 [hep-th].
41. L. Buoninfante, G. Lambiase and A. Mazumdar, Nucl. Phys. B **931**, 250 (2018).
42. L. Buoninfante, G. Lambiase and A. Mazumdar, Eur. Phys. J. C **78**, no. 1, 73 (2018).
43. T. Biswas, T. Koivisto and A. Mazumdar, "Nonlocal theories of gravity: the flat space propagator," arXiv:1302.0532 [gr-qc].
44. L. Buoninfante, Master's Thesis (2016), arXiv:1610.08744v4 [gr-qc].
45. H. Balasin and H. Nachbagauer, Class. Quant. Grav. **10**, 2271 (1993).
46. L. Buoninfante, A.S. Koshelev, J. Marto, K.S. Kumar, A. Mazumdarn, *in preparation.*

# Horizon quantum mechanics and the inner side of black holes

Roberto Casadio[a] and Andrea Giusti[b]

*Dipartimento di Fisica e Astronomia, Università di Bologna,*
*via Irnerio 46, 40126 Bologna, Italy*
*I.N.F.N., Sezione di Bologna, IS - FLAG*
*via B. Pichat 6/2, I-40127 Bologna, Italy*
[a] *E-mail: casadio@bo.infn.it*
[b] *E-mail: andrea.giusti@bo.infn.it*

Andrea Giugno

*Arnold Sommerfeld Center, Ludwig-Maximilians-Universität*
*Theresienstraße 37, 80333 München, Germany*
*E-mail: A.Giugno@physik.uni-muenchen.de*

Octavian Micu

*Institute of Space Science, Bucharest, Romania*
*P.O. Box MG-23, RO-077125 Bucharest-Magurele, Romania*
*E-mail: octavian.micu@spacescience.roe*

The Horizon Quantum Mechanics allows one to analyse the gravitational radius of spherically symmetric systems and compute the probability that a given quantum state is a black hole. We first review the global formalism and show that it reproduces a gravitationally inspired GUP relation but also leads to unacceptably large fluctuations in the horizon size of astrophysical black holes if one insists in describing them as (smeared) central singularities. On the other hand, if they are extended systems, like in the corpuscular models, no such issue arises and one can in fact extend the formalism to include asymptotic mass and angular momentum with the harmonic model of rotating corpuscular black holes. The Horizon Quantum Mechanics then shows that, in simple configurations, the appearance of the inner horizon is suppressed and extremal (macroscopic) geometries seem highly disfavoured.

*Keywords*: Black holes; Quantum physics

## 1. HQM for spherical sources

The world as we know it, is best described by quantum physics, and black holes should be represented as quantum objects as well. A first non-trivial question that follows is how much of the classical description of black holes given by general relativity we can still keep at the quantum level. One feature that should presumably live up into the quantum realm is that black holes are "gravitational bound states". Nonetheless, most of the existing literature simply analyses quantum effects on classical black hole space-times, and the proper quantum nature of the background itself is not fully accounted for. Finally, one might notice that a classical quantity that characterises such bound states (in the particular case when spherical symmetry is preserved) is the so called "gravitational radius",

$$R_{\rm H} = 2\,G_{\rm N}\,M\ , \tag{1}$$

where $M$ here can be the local Misner-Sharp mass $M = M(r)$ or the total ADM mass of the system. A quantum treatment of the gravitational radius could therefore be a good starting point for developing a fully quantum theory of black holes, much like quantising the position of the electron proved to be a good starting point for a quantum description of the hydrogen atom.

The Horizon Quantum Mechanics (HQM)[1–13] was precisely proposed with the purpose of describing the gravitational radius of spherically symmetric compact sources and for determining the existence of a horizon in a quantum mechanical fashion. In a classical spherically symmetric system, the gravitational radius (1) uniquely determines the location of the trapping surfaces where the null geodesic expansion vanishes. The latter surfaces are proper horizons in a time-independent configuration, which is the case we shall consider here for simplicity. It is then straightforward to uplift this description of the causal structure of space-time to the quantum level by simply imposing the relation between the gravitational radius and the Misner-Sharp mass as a constraint to be satisfied by the physical states of the system[2], that is

$$0 = \left( \hat{M} - \frac{\hat{R}_{\rm H}}{2\,G_{\rm N}} \right) \mid \Psi \,\rangle = \sum_{\alpha,\beta} \left( E_\alpha - \frac{R_{{\rm H}\beta}}{2\,G_{\rm N}} \right) C(E_\alpha, R_{{\rm H}\beta}) \mid E_\alpha \,\rangle \mid R_{{\rm H}\beta} \,\rangle \ , \tag{2}$$

where $\mid E_\alpha \,\rangle$ are eigenstates "Hamiltonian" operator $\hat{M}$, that is

$$\hat{M} = \sum_\alpha E_\alpha \mid E_\alpha \,\rangle\langle\, E_\alpha \mid \ , \tag{3}$$

and likewise $\mid R_{{\rm H}\alpha} \,\rangle$ are eigenstates of the gravitational radius operator,

$$\hat{R}_{\rm H} = \sum_\alpha R_{{\rm H}\alpha} \mid R_{{\rm H}\alpha} \,\rangle\langle\, R_{{\rm H}\alpha} \mid \ . \tag{4}$$

Solutions to the constraint (2) are given by

$$C(E_\alpha, R_{{\rm H}\beta}) = C(E_\alpha, 2\,G_{\rm N}\,E_\alpha)\,\delta_{\alpha\beta} \equiv C_{\rm S}(R_{{\rm H}\alpha}/2\,G_{\rm N})\,\delta_{\alpha\beta} \ , \tag{5}$$

and one can then define the horizon wave-function (HWF)[1] as

$$\psi_{\rm H}(R_{{\rm H}\alpha}) = \langle\, R_{{\rm H}\alpha} \mid \psi_{\rm H} \,\rangle = C_{\rm S}(R_{{\rm H}\alpha}/2\,G_{\rm N}) \ . \tag{6}$$

Once the HWF is introduced, one can define the probability that a given source with position wave-function $\psi_S = \psi_S(r)$ is a black hole as

$$P_{\rm BH} = \int_0^\infty \mathcal{P}_<(r < R_{\rm H})\,{\rm d}R_{\rm H} \ , \tag{7}$$

where $\mathcal{P}_<(r < R_{\rm H}) = P_{\rm S}(r < R_{\rm H})\,\mathcal{P}_{\rm H}(R_{\rm H})$ is the probability density that the source is found within its gravitational radius. The latter is in turn defined in terms of the probability

$$P_{\rm S}(r < R_{\rm H}) = 4\,\pi \int_0^{R_{\rm H}} |\psi_{\rm S}(r)|^2\,r^2\,{\rm d}r \tag{8}$$

that the source lies inside the radius $r = R_{\rm H}$ and the probability density

$$\mathcal{P}_{\rm H}(R_{\rm H}) = 4\,\pi\,R_{\rm H}^2\,|\psi_{\rm H}(R_{\rm H})|^2 \tag{9}$$

that $R_{\rm H}$ equals the gravitational radius.

Beside formal developments, we have applied the HQM to specific states of the harmonic black hole model[14], which can be considered as a working realisation of the corpuscular black holes proposed by Dvali and Gomez[15]. If the black hole contains only one quantum constituent, with energy $M \simeq \hbar/\lambda$, where $\lambda$ is the Compton length of the constituent, the results is a black hole only provided $\lambda \lesssim \ell_{\rm p}$ (the Planck length), corresponding to a mass $M \gtrsim m_{\rm p}$ (the Planck mass). For this case one also recovers a Generalised Uncertainty Principle[3]

$$\Delta r \sim \frac{\hbar}{\Delta p} + \Delta R_{\rm H} \sim \frac{\hbar}{\Delta p} + \Delta p\ , \tag{10}$$

and a quantum hoop conjecture[4]. However, the fluctuations in the horizon size are given by

$$\langle\,\Delta \hat{R}_{\rm H}^2\,\rangle \sim \lambda^{-2} \sim \Delta p^2\ , \tag{11}$$

which becomes extremely large for macroscopic black holes with $M \simeq \hbar/\lambda \gg m_{\rm p}$. This result supports the idea that a black hole must instead be made by a large number $N$ of constituents of energy $\varepsilon \simeq \hbar/\lambda$ (such that $M \simeq N\,\varepsilon \sim \sqrt{N}\,m_{\rm p}$ and $\lambda \sim R_{\rm H} \sim \sqrt{N}\,\ell_{\rm p}$). For the latter case, horizon fluctuations can be very small, of order $1/N$, and the quantum harmonic model is simple enough that one can determine explicitly the probability that the chosen states are indeed black holes[16].

## 2. HQM for rotating sources

In a space-time generated by an axially-symmetric rotating source, we exploit the fact that the asymptotic behaviour of the system is described by the Kerr metric. We can then uplift to a quantum condition for the physical states the classical relation that determines the two horizon radii

$$R_\pm = M \pm \sqrt{M^2 - a^2} \tag{12}$$

from the mass $M$ and angular momentum $J = a^2\,M^2$ of the source which holds in the (asymptotic) Kerr metric[10]. The formalism described in the previous section now doubles with the introduction of two gravitational radius operators, $\hat{R}_\pm$, and corresponding HWF's we can denote as $\psi_\pm = \psi_{\rm H}(R_\pm)$. The probability $P_{\rm BH} = P_+$ that the whole source is a black hole is then obtained from $\psi_+$, and the probability $P_-$ that the inner horizon is also realised is likewise determined from $\psi_-$.

The harmonic black hole model[14] also allows one to compute the above two probabilities for states of a large number $N$ of gravitons with non-vanishing total angular momentum[10]. The main result is that, for extremal configurations with $M^2 \simeq a^2$, one finds

$$P_+ \simeq P_{1+}^N \ll 1\ , \tag{13}$$

where $P_{1+} \lesssim 1$ is the probability that each constituent is inside $R_+$. For instance, we typically obtain $P_1 \simeq 0.7$ and, since $N \sim M^2/m_{\rm p}^2 \gg 1$ for astrophysical sources, it appears that such configurations are very unlikely to have a horizon and be black holes. Moreover, the inner horizon $R_-$ also has a small probability to exist for large sources, to wit

$$P_- \simeq P_{1-}^N \ll 1 \,, \tag{14}$$

where now $P_{1-} \ll 1$ is the probability that each constituent is inside the inner horizon $R_-$. A typical value is $P_{1-} \simeq 0.05$ and $N \gg 1$.

The overall conclusion is that quantum fluctuations for rotating geometries, although negligibly small for the horizon of non extremal configurations, appear to be large enough to spoil the causal structure of extremal configurations and eliminate the presence of inner horizons. We recall that the latter are in fact rather problematic in the semiclassical description, because of the effect called mass inflation[17]. We remark that the HQM seems to provide a solution to this problem in general, since the inner horizon is removed also for electrically charged sources[5].

## Acknowledgments

R.C. and A.G. are partially supported by the INFN grant FLAG. The work of R.C. and A.G. has also been carried out in the framework of activities of the National Group of Mathematical Physics (GNFM, INdAM). O.M. was supported by the grant LAPLAS 4.

## References

1. R. Casadio, "Localised particles and fuzzy horizons: A tool for probing Quantum Black Holes," arXiv:1305.3195 [gr-qc]; Springer Proc. Phys. **170** (2016) 225 [arXiv:1310.5452 [gr-qc]].
2. R. Casadio, A. Giugno and A. Giusti, Gen. Rel. Grav. **49** (2017) 32 [arXiv:1605.06617 [gr-qc]].
3. R. Casadio and F. Scardigli, Eur. Phys. J. C **74** (2014) 2685 [arXiv:1306.5298 [gr-qc]].
4. R. Casadio, O. Micu and F. Scardigli, Phys. Lett. B **732** (2014) 105 [arXiv:1311.5698 [hep-th]].
5. R. Casadio, O. Micu and D. Stojkovic, JHEP **1505** (2015) 096 [arXiv:1503.01888 [gr-qc]];
6. R. Casadio, O. Micu and D. Stojkovic, Phys. Lett. B **747** (2015) 68 [arXiv:1503.02858 [gr-qc]].
7. R. Casadio, A. Giugno and O. Micu, Int. J. Mod. Phys. D **25** (2016) 1630006 [arXiv:1512.04071 [hep-th]].
8. X. Calmet and R. Casadio, Eur. Phys. J. C **75** (2015) 445 [arXiv:1509.02055 [hep-th]].

9. R. Casadio, R. T. Cavalcanti, A. Giugno and J. Mureika, Phys. Lett. B **760** (2016) 36 [arXiv:1509.09317 [gr-qc]].
10. R. Casadio, A. Giugno, A. Giusti and O. Micu, Eur. Phys. J. C **77** (2017) 322 [arXiv:1701.05778 [gr-qc]].
11. R. Casadio and O. Micu, Eur. Phys. J. C **78** (2018) 852 [arXiv:1806.05944 [gr-qc]].
12. R. Casadio, A. Giusti and R. Rahim, EPL **121** (2018) 60004 [arXiv:1801.05621 [gr-qc]].
13. R. Casadio, A. Giugno, A. Giusti and M. Lenzi, "Quantum Formation of Primordial Black holes," arXiv:1810.05185 [gr-qc].
14. R. Casadio and A. Orlandi, JHEP **1308** (2013) 025 [arXiv:1302.7138 [hep-th]].
15. G. Dvali and C. Gomez, JCAP **01** (2014) 023 [arXiv:1312.4795 [hep-th]]. "Black Hole's Information Group", arXiv:1307.7630; Eur. Phys. J. C **74** (2014) 2752 [arXiv:1207.4059 [hep-th]]; Phys. Lett. B **719** (2013) 419 [arXiv:1203.6575 [hep-th]]; Phys. Lett. B **716** (2012) 240 [arXiv:1203.3372 [hep-th]]; Fortsch. Phys. **61** (2013) 742 [arXiv:1112.3359 [hep-th]]; G. Dvali, C. Gomez and S. Mukhanov, "Black Hole Masses are Quantized," arXiv:1106.5894 [hep-ph].
16. R. Casadio, A. Giugno, O. Micu and A. Orlandi, Entropy **17** (2015) 6893 [arXiv:1511.01279 [gr-qc]].
17. E. Poisson and W. Israel, Phys. Rev. Lett. **63** (1989) 1663 .

# The non perturbative gyro-phase is the *Kaluza-Klein* $5^{th}$ dimension

C. Di Troia*

*ENEA, Fusion and Nuclear Safety Department, C. R. Frascati, 00044 (RM) Italy*
**E-mail: claudio.ditroia@enea.it*

The curve in space occupied by the mass during time evolution, is a geodesic on the space-time manifold curved by the presence of masses: the mass can only follow its trajectory consistently with the underlying gravitational field. Is it possible to think at the charge trajectory in a similar fashion? Is it possible to say that the charge trajectory, the curve in phase-space occupied by the charge during time evolution, is the geodesic on the extended phase-space curved by the presence of charges? If yes then it should be possible to obtain an Einstein's equation also for electromagnetism. This is done by considering a metric on the whole extended phase-space, which is the phase-space $(x, v)$ plus the time, $t$. It is proposed to add a Hilbert-Einstein term in the lagrangian when velocities are considered as dynamical variables. Here, it will be analyzed what happens if the (non perturbative) guiding center description of motion is adopted [1]. In such case, a similar mechanism to the one proposed by Kaluza and Klein (KK) [3,4] a century ago is found. The advantage of using the present description is that, now, there is no need of looking for a compactification scheme as required in the original KK mechanism. Indeed, the extra-dimension that appears in the guiding center transformation is a physical and, in principle, measurable variable being the gyro-phase, the angle obtained when the velocity space is described by a sort of cylindrical transformation of velocities coordinates. Regardless of the equations that are really similar to the one seen in the KK mechanism, the new claim is in the interpretation of the extra dimension as a coordinate coming from the velocity space. Until now, all the compactification mechanisms have been shown to give problems, like the inconsistency of the scale of masses with observations. Instead, without a compactification at the Planck scale length and giving a physical meaning to the extra-coordinate, it seems that the KK mechanism can finally be accepted as a realistic explanation of the presence of gravitation and electromagnetism treated in a unified manner in general relativity theory extended to higher dimensions.

*Keywords*: Kaluza Klein, Guiding Center Transformation, Einstein's equation in higher dimensions

## 1. Introduction

The thesis of the present work is that through the guiding center transformation [1,2] it is possible to support the Kaluza-Klein model [3–6] without the need to use new physical dimensions with respect to those of the phase space extended to time, $(t, x, v)$. What turns out is that the extra dimension comes from the velocity space. There are two constants of motion, the single particle lagrangian that does not change value with respect to the guiding center transformation and the magnetic moment, $\mu$, that does not change value with respect to the variation of the gyrophase, $\gamma$. In the non perturbative guiding center transformation it is assumed that there exist a reference point, the guiding center $(X, U)$, from where the particle motion is seen to be closed and periodic. Thus the motion of a charged particle is represented by the product of the guiding center orbit times the circle with gyroradius, $\rho$. Such

helicoidal motion is implicitly written in a flat phase space as

$$x = X + \rho(t, X, \gamma; \mu, \varepsilon) \tag{1}$$
$$u = U(t, X; \mu, \varepsilon) + \nu(t, X, \gamma; \mu, \varepsilon),$$

where, $\varepsilon$ is the energy, $u = x'$ is the relativistic velocity of the particle, $U = X'$ is the relativistic velocity of the guiding center and $\nu = \rho'$ is the difference of the formers or the proper time derivative of the gyroradius. The relativistic lagrangian is:

$$L = -p_\alpha u^\alpha, \tag{2}$$

being the co-momentum $p_\alpha = u_\alpha + (e/m)A_\alpha(x^\beta)$, with $A_\alpha$ the e.m. potential. Also the guiding center has a co-momentum defined as $P_\alpha = U_\alpha + (e/m)A_\alpha(X^\beta)$, in such a way that the lagrangian in (2) is

$$L = -p_\alpha u^\alpha = -P_\alpha U^\alpha - (e/m)g'. \tag{3}$$

If the gauge function $g$ is proportional to the product of the magnetic moment, $\mu$, times the gyro-phase, $\gamma$, it is called the *guiding center gauge function* : $g = (m/e)\mu\gamma$. Such gauge function doesn't alter the equation of motion but it is useful for preserving the same value of the lagrangian along the motion. The link with the KK model starts from denoting the values $z^0 = t$, $z = X$, $z^4 = \gamma$ and $w_0 = P_0 = U_0 + (e/m)\Phi$, $w = P = U + (e/m)A$ and $w_4 = (m/e)\mu$, then

$$L = -w_a z'^a, \qquad \text{for a=0,1,2,3,4.} \tag{4}$$

which is equivalent to a scalar product in a space-time of five dimensions. Moreover, if you require that $w_5 = w_6 = 0$ and $z^5 = \mu$, $z^6 = \varepsilon$ is the particle energy (per unitary mass), then you can also write

$$L = -w_A z'^A, \qquad \text{for A=0,1,2,3,4,6.} \tag{5}$$

The latter is what is called the phase-space lagrangian from which it is possible to find the *Hamilton's equations*. It is better to denote with a *hat* the guiding center phase-space lagrangian: $L = \hat{L}(z^A, z'^B)$, for $A, B = 0, 1, 2, 3, 4, 6$. As said in[7], the reason for the vanishing of $w_5$ and $w_6$ is due to the fact that $\varepsilon$ is the conjugate coordinate of $t$ and $(m/e)\mu$ is the conjugate coordinate of $\gamma$.

Now, the lagrangian is invariant at a glance with respect to general non-canonical phase-space coordinates transformations, that include also the gauge transformations.

## 2. Kaluza-Klein solution

The coordinates $z^A$ with $A = 0, 1, 2, 3, 4, 5, 6$, just introduced, belong to the extended phase space, *e.g.* $z^A = (t, X, \gamma, \mu, \varepsilon)$ are the guiding center coordinates. As for general relativity, where a geometry is given to the space-time, in this section a geometry is given to the extended phase-space.

Let's start from the *Poincaré-Cartan* one-form from (5): $\hat{L}d\hat{s} = -w_A dz^A$, for $A = 0, 1, 2, 3, 4, 5, 6$. The same one-form can be written as

$$\hat{L}d\hat{s} = -\hat{g}_{AB} w^B dz^A, \tag{6}$$

being $\hat{L}$ a scalar quantity and where $\hat{g}_{AB}$ is the metric tensor on the extended phase-space with the property that $w_A \equiv \hat{g}_{AB} w^B$. Thus, $w^B$ are the *contra-variant* momenta. Once the metric tensor on the extended phase-space is appeared, it is possible to apply a variational principle for finding it. For this reason, we consider a *lagrangian density over the extended phase space* where the single particle lagrangian is multiplied for the distribution of masses and, then, added to the HE lagrangian in extended dimensions. In the following the lagrangian density on the extended phase-space is

$$\ell a = f_m \hat{L} - \frac{\hat{\mathcal{R}}}{16\pi\hat{G}}, \tag{7}$$

Thus, the action is:

$$S = \int \ell a \, d\mathcal{M}, \tag{8}$$

which is a definite integration in a domain $\partial\mathcal{M}$ of the extended phase space. In the guiding center transformation the volume element of the extended phase-space is $d\mathcal{M} = \tilde{J}_p dt d^3X d\gamma d\mu d\varepsilon$.

It is possible to separate in $\ell a$ the effects of different contributions: a *matter*, a *field* and an *interaction* lagrangian densities on the extended phase-space. Concerning the *field action*, $S_{\rm f}$, we have:

$$\begin{aligned} S_{\rm f} &= -\int \frac{\hat{\mathcal{R}}}{16\pi\hat{G}} \sqrt{|\hat{g}|} d^7 z = \\ &= -\int \frac{F_{\alpha\beta}F^{\alpha\beta}}{4} \sqrt{-g} dt d^3X - \int \frac{R}{16\pi G} \sqrt{-g} dt d^3X. \end{aligned} \tag{9}$$

In order to obtain the latter result the KK mechanism is applied setting the following fifth dimensional metric tensor:

$$\tilde{g}_{ab} = \begin{vmatrix} g_{\alpha\beta} - 16\pi G A_\alpha A_\beta & -16\pi G (m/e)^2 \mu A_\alpha \\ -16\pi G (m/e)^2 \mu A_\beta & -16\pi G (m/e)^4 \mu^2 \end{vmatrix}. \tag{10}$$

The gravitational constant, $\hat{G}$, is obtained from

$$\frac{\hat{G}}{G} = 4 \int (m/e)^2 \sqrt{\pi G} \mu \tilde{J}_{\mathcal{P}} d\gamma d\varepsilon d\mu. \tag{11}$$

At the same time, the action due to the interaction is

$$S_{\rm id} = \int f_m (1 + \hat{L}) \sqrt{|\hat{g}|} d^7 z = -\int A_\alpha J^\alpha \sqrt{-g} dt d^3X, \tag{12}$$

where $J^\alpha$ is the charge four-current density which is a field depending on $(t, X)$. Thus the sources of fields are the guiding centers. The former equation is obtained through the non perturbative guiding center transformation and through the *misleading* symmetry, below described. It is worth noticing that once integrated in the velocity space, the obtained lagrangian density is the one used for describing the presence of matter as source of a gravitational field, which gives the *Einstein*'s equation, together with a charge four-current density as source of an e.m. field, which gives the *Maxwell*'s equations. It is worth noticing that after the integration over the velocity space the model equations have lost locality, being $X$ the position of the guiding center and not of the particle. Without locality all the singularities of gravitation, *e.g.* the problems concerning black holes treated as singularities, and of electrodynamics, *e.g.* the self-energy problem, are ruled out.

Moreover, from (7) it is also possible to deduce a *Boltzmann*'s equation.

#### 2.0.1. Boltzmann*'s equation from geometry*

By writing $\hat{L} = s'$, as the proper time derivative of the *principal Hamilton function*, $s$, then it is possible to recognize a *Legendre*'s transformation in equation (7) that explicitly becomes

$$\ell a = f_m s' - \frac{\hat{\mathcal{R}}(f_m, s)}{16\pi\hat{G}}. \tag{13}$$

Surprisingly, the distribution of masses is the conjugate momentum of the single particle action. By varying such lagrangian density on extended phase-space with respect to $s$ and $s'$, then the *Euler-Lagrangian* equation is

$$f'_m = \frac{\nabla_s \hat{\mathcal{R}}}{16\pi\hat{G}}. \tag{14}$$

The former is a *Boltzmann*'s equation, describing the evolution of the distribution function, with a $RHS$ that depends on the geometry instead of a collision operator. However, it is not possible to arrive at a Fokker-Planck's equation without considering fluctuations, not considered here.

### 2.1. *The* misleading *symmetry*

In the relativistic case, it is chosen to preserve the product $u^\alpha A_\alpha(x^\beta) = U^\alpha A_\alpha(X^\beta)$ that allows to write $L = -1 - (e/m)u^\alpha A_\alpha(x^\beta) = \hat{L}$ with

$$\hat{L} = -1 - (e/m)U^\alpha A_\alpha(X^\beta), \tag{15}$$

which is the same form of $L$. The required condition is reached if

$$(m/e)\mu\gamma' = 1 - U^\alpha U_\alpha. \tag{16}$$

The latter relation is also more interesting if $(m/e)\mu\gamma' = U^4 U_4$, where $U^4 = z^{4\prime} = \gamma'$ and $U_4$ is firstly defined as $U_4 \equiv w_4 = (m/e)\mu$. In such way that

$$U^a U_a = 1, \qquad \text{for } a = 0, 1, 2, 3, 4. \tag{17}$$

Moreover, if the relation $w_a = U_a + (e/m)A_a$ is used, then $A_4 = 0$ for consistency: there is not a $5^{th}$ component of the e.m. potential. The symmetry that leaves invariant the form of $L = -1 - (e/m)u^\alpha A_\alpha(x^\beta) = -1 - (e/m)U^\alpha A_\alpha(X^\beta)$ is said *misleading* because there is no way, starting from the dynamics, *e.g.* from the lagrangian, to distinguish particle's coordinates from guiding center's coordinates. The only chance for appreciating the difference is by measuring the dispersion relation: from kinematics, the particle has $u^\alpha u_\alpha = 1$ whilst the guiding center doesn't, $U^\alpha U_\alpha \neq 1$, that means that the guiding center is a virtual particle. It is worth noticing that in such extended phase-space approach, the space-time is defined as the domain where the e.m. four-potential is defined, in such a way that electromagnetism occurs for separating the space-time where uncharged particles moves from the whole extended phase-space.

## 3. Conclusion

We have just seen that the guiding center transformation, which is a particular *local* translation in the extended phase space, *e.g.* see (1), is a symmetry because it leaves the same lagrangian form. In analogy to what happens for the *local* translation in spacetime, the conserved quantity for the present symmetry should be called the *extended energy-momentum* tensor $\hat{T}_{AB}$. Now, the *Einstein tensor* for the extended phase space is obtained from the variation of $-\hat{\mathcal{R}}/16\pi\hat{G}$ with respcet to $\delta\hat{g}^{AB}$:

$$\hat{G}_{AB} = \hat{Ric}_{AB} - \hat{R}\,\hat{g}_{AB}/2, \tag{18}$$

and the *Einstein equation* can be written also for the extended phase space,

$$\hat{G}_{AB} = 8\pi\hat{G}\,\hat{T}_{AB}. \tag{19}$$

It is worth noticing that, if confirmed, we have just obtained gravitation and electromagnetism from a geometrical perspective.

## References


1. C. Di Troia, *Journal of Modern Physics*, **9** (2018) 701.
2. J. R. Cary and A. J. Brizard, *Rev. Mod. Phys.* **81** (2009) 693.
3. Th. Kaluza, *Sitzungsber. Preuss Akad. Wiss. Berlin Math. Phys.*, **k1** (1921) 374.
4. O. Klein, *Z. Phys.* **37** (1926), 895.
5. D. Bailin and A. Love, *Annals physics* **151** (1983) 1.
6. J. M. Overduin and P. S. Wesson, *Rep. Prog. Phys.* **50** (1987) 1087.
7. J. R. Cary and R. G. Littlejohn, *Annals physics* **151** (1983) 1.

# Symmetry of Brans-Dicke gravity as a novel solution-generating technique

Valerio Faraoni

*Dept. of Physics & Astronomy and STAR Research Cluster, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*
*E-mail: vfaraoni@ubishops.ca*

Dilek K. Çiftci

*Dept. of Physics & Astronomy and STAR Research CLuster, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*

*Department of Physics, Namık Kemal University, Tekirdağ, Turkey*
*E-mail: dkazici@nku.edu.tr*

Shawn D. Belknap-Keet

*Dept. of Physics & Astronomy, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*
*E-mail: sbelknapkeet02@ubishops.ca*

A little known symmetry group of Brans-Dicke gravity in the presence of conformally invariant matter (including electrovacuo) is used as a solution-generating technique, starting from a known solution as a seed. This novel technique is applied to generate, as examples, new spatially homogeneous and isotropic cosmologies, a 3-parameter family of spherical time-dependent spacetimes conformal to a Campanelli-Lousto geometry, and a family of cylindrically symmetric geometries.

*Keywords*: Brans-Dicke gravity, exact solutions, symmetry.

## 1. Introduction

Scalar-tensor gravity is the prototype theory alternative to general relativity (GR). The main motivation to modify gravity comes from cosmology: to explain the present acceleration of the universe, the GR-based ΛCDM model introduces a completely *ad hoc* dark energy.[1] Modifying gravity is an alternative.[2,3] The most popular approach is metric $f(R)$ gravity, which is an $\omega = 0$ Brans-Dicke (BD) theory (with a potential) in disguise. But this is not the only fundamental motivation: all attempts to quantize gravity introduce corrections to GR, consisting of quadratic terms in the curvature (giving rise to Starobinsky inflation[4]), scalar fields, or nonlocal terms (for example, the low-energy limit of bosonic string theory is an $\omega = -1$ BD theory[5,6]). What is more, Dirac's idea of varying fundamental "constants" of physics[7] is partially realized in scalar-tensor gravity, where the scalar degree of freedom $\phi \sim G^{-1}$ is dynamical. More recently, with quantum gravity in mind, it has been found that generalized BD solutions describe asymptotically Lifschitz black holes.[8]

When available, analytic solutions provide insight into various aspects of a theory, but they are relatively rare in scalar-tensor gravity. Therefore, it is valuable

to find new solution-generating techniques. We explore a new one based on a 1-parameter symmetry group of BD theory. As an application, we found three new families of solutions of BD theory with potential $V \propto \phi^\beta$, and of $f(R) = R^n$ gravity.[9] They include Friedmann-Lemaître-Robertson-Walker (FLRW) cosmologies with power-law or exponential scale factor; spherical, time-dependent, asymptotically FLRW solutions; and axially symmetric (cosmic string-like) geometries.

## 2. Symmetry Group of Brans-Dicke Theory

The action of vacuum BD theory with a potential

$$S_{BD} = \int d^4x\sqrt{-g}\left[\phi R - \frac{\omega}{\phi}\, g^{ab}\nabla_a\phi\nabla_b\phi - V(\phi)\right] \tag{1}$$

is invariant in form under the transformation $(g_{ab}, \phi) \to \left(\tilde{g}_{ab}, \tilde{\phi}\right)$, where

$$\tilde{g}_{ab} = \Omega^2 g_{ab} = \phi^{2\alpha} g_{ab}\,, \tag{2}$$

$$\tilde{\phi} = \phi^{1-2\alpha}\,, \qquad \alpha \neq 1/2 \tag{3}$$

and where we follow the notation of Ref.[10]. Using the standard transformation properties under conformal transformations

$$\tilde{g}^{ab} = \Omega^{-2} g^{ab}\,, \qquad \sqrt{-\tilde{g}} = \Omega^4\sqrt{-g}\,, \tag{4}$$

$$\tilde{R} = \Omega^{-2}\left(R - \frac{6\Box\Omega}{\Omega}\right)\,, \tag{5}$$

one obtains

$$R = \phi^{2\alpha}\tilde{R} - \frac{6\alpha(1-\alpha)}{(1-2\alpha)^2}\,\phi^{6\alpha-2}\tilde{g}^{ab}\tilde{\nabla}_a\tilde{\phi}\tilde{\nabla}_b\tilde{\phi} + \frac{6\alpha}{1-2\alpha}\,\phi^{4\alpha-1}\tilde{\Box}\tilde{\phi}\,. \tag{6}$$

The last term contributes only a total divergence to $\sqrt{-g}\,\phi R$ in the action (1),

$$\frac{6\alpha}{1-2\alpha}\,\sqrt{-\tilde{g}}\,\tilde{\Box}\tilde{\phi} = \frac{6\alpha}{1-2\alpha}\,\partial_\mu\left(\sqrt{-\tilde{g}}\,\tilde{g}^{\mu\nu}\partial_\nu\tilde{\phi}\right) \tag{7}$$

and the BD action (1) then becomes

$$S_{BD} = \int d^4x\sqrt{-\tilde{g}}\left\{\tilde{\phi}\tilde{R} - \left[\frac{\omega}{(1-2\alpha)^2} + \frac{6\alpha(1-\alpha)}{(1-2\alpha)^2}\right] \cdot\frac{\tilde{g}^{ab}}{\tilde{\phi}}\tilde{\nabla}_a\tilde{\phi}\tilde{\nabla}_b\tilde{\phi} - \tilde{\phi}^{\frac{-4\alpha}{1-2\alpha}}V(\phi)\right\}\,. \tag{8}$$

By redefining the BD coupling and the potential according to

$$\tilde{\omega}(\omega,\alpha) = \frac{\omega + 6\alpha(1-\alpha)}{(1-2\alpha)^2}\,, \tag{9}$$

$$\tilde{V}(\tilde{\phi}) = \tilde{\phi}^{\frac{-4\alpha}{1-2\alpha}}V\left(\tilde{\phi}^{\frac{1}{1-2\alpha}}\right)\,, \tag{10}$$

we write

$$S_{BD} = \int d^4x\sqrt{-\tilde{g}}\left[\tilde{\phi}\tilde{R} - \frac{\tilde{\omega}}{\tilde{\phi}}\,\tilde{g}^{ab}\tilde{\nabla}_a\tilde{\phi}\tilde{\nabla}_b\tilde{\phi} - \tilde{V}(\tilde{\phi})\right], \tag{11}$$

*i.e.*, the action (1) is invariant in form. The transformations (3) form a 1-parameter Abelian group.[11,12]

A special case is given by a power-law potential $V(\phi) = V_0\phi^n$, which becomes $\tilde{V}(\tilde{\phi}) = V_0\tilde{\phi}^{\tilde{n}}$ with $\tilde{n} = \frac{n-4\alpha}{1-2\alpha}$ (and is invariant if $n = 2$).

### 2.1. *Electrovacuum Brans-Dicke theory*

Electrovacuum BD theory is described by the action

$$S_{BD} = \int d^4x\sqrt{-g}\left[\phi R - \frac{\omega}{\phi}\,g^{ab}\nabla_a\phi\nabla_b\phi - V(\phi) - \frac{1}{4}\,F^{ab}F_{ab}\right]; \tag{12}$$

since $\tilde{F}_{ab} = F_{ab}$ and $\sqrt{-g}\,F^{ab}F_{ab} = \sqrt{-\tilde{g}}\,\widetilde{F^{ab}F_{ab}}$, also $\sqrt{-g}\,\mathcal{L}_{(m)}$ remains invariant under the transformations (3).

### 2.2. *Conformally invariant matter*

For simplicity, let us use now the field equations

$$R_{ab} - \frac{1}{2}\,g_{ab}R = \frac{8\pi}{\phi}\,T_{ab} + \frac{\omega}{\phi^2}\left(\nabla_a\phi\nabla_b\phi - \frac{1}{2}\,g_{ab}g^{cd}\nabla_c\phi\nabla_d\phi\right) + \frac{1}{\phi}\left(\nabla_a\nabla_b\phi - g_{ab}\Box\phi\right) - \frac{V}{2\phi}\,g_{ab}\,, \tag{13}$$

$$\Box\phi = \frac{1}{2\omega+3}\left[\frac{8\pi T}{\phi} + \phi\,\frac{dV}{d\phi} - 2V\right]. \tag{14}$$

The symmetry transformation (3) gives

$$\tilde{\Box}\tilde{\phi} = \frac{1}{2\tilde{\omega}+3}\left[\frac{8\pi}{1-2\alpha}\,\tilde{\phi}^{\frac{-4\alpha}{1-2\alpha}}T + \tilde{\phi}\,\frac{d\tilde{V}}{d\tilde{\phi}} - 2\tilde{V}\right] \tag{15}$$

and the field equations are conformally invariant only if $T = 0$. Moreover,

$$\tilde{R}_{ab} - \frac{1}{2}\,\tilde{g}_{ab}\tilde{R} = \frac{8\pi}{\tilde{\phi}^{\frac{1}{1-2\alpha}}}T_{ab} + \frac{\tilde{\omega}}{\tilde{\phi}^2}\left(\tilde{\nabla}_a\tilde{\phi}\tilde{\nabla}_b\tilde{\phi} - \frac{1}{2}\,\tilde{g}_{ab}\tilde{g}^{cd}\tilde{\nabla}_c\tilde{\phi}\tilde{\nabla}_d\tilde{\phi}\right) + \frac{1}{\tilde{\phi}}\left(\tilde{\nabla}_a\tilde{\nabla}_b\tilde{\phi} - \tilde{g}_{ab}\tilde{\Box}\tilde{\phi}\right) - \frac{\tilde{V}}{2\tilde{\phi}}\,\tilde{g}_{ab} \tag{16}$$

where $\tilde{T}_{ab} = \Omega^{-2}\,T_{ab}$, so the first term in the right hand side becomes $8\pi\tilde{T}_{ab}/\tilde{\phi}$ and it is invariant. Hence, the field equations are invariant for arbitrary $V(\phi)$ but only for conformally invariant matter.

## 3. Example: Brans-Dicke Cosmology

In the case of FLRW cosmology, the Lagrangian reduces to a point-like one, but the symmetry (3) is not a Noether nor a Hojman symmetry. Moreover, the classical symmetry is broken by Wheeler-DeWitt quantization in minisuperspace: quantum effects cause an anomalous symmetry breaking similar to that occuring in condensed matter systems.[13]

This time our "seed" is[14,15]

$$ds^2 = -dt^2 + S^2(t)\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2_{(2)}\right) \tag{17}$$

with power-law scale factor $S(t)$. We have vacuum BD theory with $V \equiv 0$ and

$$S(t) = S_0 t^p\,, \tag{18}$$

$$\phi(t) = \phi_0 t^q\,. \tag{19}$$

The symmetry transformation yields

$$d\tilde{s}^2 = -t^{2\alpha q}dt^2 + S_0^2 t^{2(p+\alpha q)}\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2_{(2)}\right)\,; \tag{20}$$

we introduce a new time $\tau$ with $t = (\alpha q+1)^{\frac{1}{\alpha q+1}}\,\tau^{\frac{1}{\alpha q+1}}$, then the new solution is recast in the form

$$d\tilde{s}^2 = -d\tau^2 + \tilde{S}_0^2 \tau^{\frac{2(p+\alpha q)}{\alpha q+1}}\left(\frac{dr^2}{1-kr^2} + r^2 d\Omega^2_{(2)}\right)\,, \tag{21}$$

$$\tilde{\phi}(\tau) = (\alpha q+1)^{\frac{q(1-2\alpha)}{\alpha q+1}}\,\phi_0^{1-2\alpha}\tau^{\frac{q(1-2\alpha)}{\alpha q+1}}\,, \tag{22}$$

or $\tilde{S}(\tau) = \tilde{S}_0\tau^{\tilde{p}}$, $\tilde{\phi}(\tau) = \tilde{\phi}_0\tau^{\tilde{q}}$ with

$$\tilde{p} = \frac{p+\alpha q}{\alpha q+1}\,, \qquad \tilde{q} = \frac{q(1-2\alpha)}{\alpha q+1}\,, \tag{23}$$

$$\tilde{S}_0 = (\alpha q+1)^{\frac{p+\alpha q}{\alpha q+1}}\,S_0\,, \quad \tilde{\phi}_0 = (\alpha q+1)^{\frac{q(1-2\alpha)}{\alpha q+1}}\,\phi_0^{1-2\alpha}\,. \tag{24}$$

## 4. A New Family of Spherical Time-Dependent Solutions

Begin now from a special case of a family of spherical, time-dependent solutions of vacuum BD gravity conformal to the Fonarev[16] spacetime of GR[17]

$$\begin{aligned} ds^2 = &-A(r)^{\frac{1}{\sqrt{1+4d^2}}(2d-\frac{1}{\sqrt{|2\omega+3|}})}\,\mathrm{e}^{4dat(2d-\frac{1}{\sqrt{|2\omega+3|}})}dt^2 \\ &+\mathrm{e}^{2at(1-\frac{2d}{\sqrt{|2\omega+3|}})}\Big[A(r)^{\frac{-1}{\sqrt{1+4d^2}}(2d+\frac{1}{\sqrt{|2\omega+3|}})}dr^2 \\ &+A(r)^{1-\frac{1}{\sqrt{1+4d^2}}(2d+\frac{1}{\sqrt{|2\omega+3|}})}r^2 d\Omega^2_{(2)}\Big]\,, \end{aligned} \tag{25}$$

$$\phi(t,r) = \phi_0\,\mathrm{e}^{\frac{4dat}{\sqrt{|2\omega+3|}}}A(r)^{\frac{1}{\sqrt{|2\omega+3|(1+4d^2)}}}\,, \tag{26}$$

where $A(r) = 1 - 2m/r$, $V(\phi) = V_0\phi^\beta$, $\beta = 2\left(1 - d\sqrt{|2\omega+3|}\right)$. Our "seed" is the special case with $a \neq 0$ and with time dependence eliminated by the parameter choice

$$d = \left(2\sqrt{|2\omega+3|}\right)^{-1} = \sqrt{|2\omega+3|}/2 \tag{27}$$

*simultaneously*, which gives $\omega = -1$ (this is the low-energy limit of the bosonic string, so presumably this solution has a stringy analogue). The scalar field remains time-dependent, $\beta = 1$, and $V(\phi) = V_0\phi$ (corresponding to a cosmological constant). Then

$$ds^2 = -dt^2 + A(r)^{-\sqrt{2}}dr^2 + A(r)^{1-\sqrt{2}}r^2 d\Omega^2_{(2)}\,, \tag{28}$$

$$\phi(t,r) = \phi_0\, \mathrm{e}^{2at} A(r)^{1/\sqrt{2}}\,, \tag{29}$$

which is a special case[17] of the Campanelli-Lousto geometry.[18] The symmetry transformation (3) applied to this seed now generates a new solution with

$$\tilde{V}(\tilde{\phi}) = V_0\, \tilde{\phi}^{\frac{1-4\alpha}{1-2\alpha}}\,, \qquad \tilde{\omega} = \frac{6\alpha(1-\alpha)-1}{(1-2\alpha)^2} \tag{30}$$

given by

$$\begin{aligned} d\tilde{s}^2 = &-\mathrm{e}^{4\alpha a t} A(r)^{\alpha\sqrt{2}} dt^2 + \mathrm{e}^{4\alpha a t}\left[A(r)^{-\sqrt{2}(1-\alpha)}dr^2\right. \\ &\left. + A(r)^{1-\sqrt{2}(1-\alpha)} r^2 d\Omega^2_{(2)}\right]\,, \end{aligned} \tag{31}$$

$$\tilde{\phi}(t,r) = \tilde{\phi}_0\, \mathrm{e}^{2a(1-2\alpha)t} A(r)^{\frac{1-2\alpha}{\sqrt{2}}}\,, \qquad \tilde{\phi}_0 = \phi_0^{1-2\alpha}\,. \tag{32}$$

If $a \neq 0$, we can define the new time $\tau = \frac{\mathrm{e}^{2\alpha a t}}{2\alpha a}$ to obtain

$$d\tilde{s}^2 = -A(r)^{\alpha\sqrt{2}}d\tau^2 + (2\alpha a\tau)^2\left[A(r)^{-\sqrt{2}(1-\alpha)}dr^2 + A(r)^{1-\sqrt{2}(1-\alpha)}r^2 d\Omega^2_{(2)}\right]\,, \tag{33}$$

$$\tilde{\phi}(\tau,r) = \phi_*\, \tau^{\frac{1-2\alpha}{\alpha}} A(r)^{\frac{1-2\alpha}{\sqrt{2}}}\,, \qquad \tilde{\phi}_* = \left[(2\alpha a)^{1/\alpha}\,\phi_0\right]^{1-2\alpha}\,. \tag{34}$$

If $m \to 0$, or for $r \gg m$, this geometry reduces to a spatially flat FLRW universe with linear scale factor $S(\tau) = 2\alpha a\tau$ and the scalar field

$$\tilde{\phi}(\tau,r) = \phi_*\, \tau^{\frac{1-2\alpha}{\alpha}} \tag{35}$$

acts asymptotically as a perfect fluid with $P = -\rho/3$.

## 5. Conclusions

The 1-parameter Abelian symmetry group formed by the transformations (2) and (3) as $\alpha$ varies offers a new solution-generating technique for (electro)vacuum BD gravity with a potential $V(\phi)$. The symmetry transformation (3) introduces one new parameter $\alpha$ in addition to those already present in the seed solution. We have

found a new 2-parameter family of spherical time-dependent, asymptotically FLRW solutions and also new FLRW solutions with power-law (or exponential[9]) scale factor. New cylindrical (cosmic string-like) solutions have also been found and they are reported in Ref.[9].

By using the BD representation of $f(R)$ gravity, it can be shown that these new solutions are also solutions of $f(R) = R^n$ theory.[9] The mathematical technique exposed does not, of course, guarantee that the new solutions are physically interesting, but it constitutes a new tool. Future work will explore new applications of the solution-generating technique presented here.

## Acknowledgments

V.F. is supported by the Natural Sciences and Engineering Research Council of Canada (Grant No. 2016- 03803), and D.K.C. by the Scientific and Technological Research Council of Turkey TUBITAK (Programme BDEB-2219).

## References

1. L. Amendola and S. Tsujikawa, *Dark Energy, Theory and Observations* (Cambridge University Press, Cambridge, UK, 2010).
2. S. Capozziello, S. Carloni, and A. Troisi, *Rec. Res. Dev. Astron. Astrophys.* **1**, 625 (2003) [arXiv:astro-ph/0303041].
3. S.M. Carroll, V. Duvvuri, M. Trodden, and M.S. Turner, *Phys. Rev. D* **70**, 043528 (2004).
4. A.A. Starobinsky, *Phys. Lett. B* **91**, 99 (1980).
5. C.G. Callan, D. Friedan, E.J. Martinez, and M.J. Perry, *Nucl. Phys. B* **262**, 593 (1985).
6. E.S. Fradkin and A.A. Tseytlin, *Nucl. Phys. B* **261**, 1 (1985).
7. P.A.M. Dirac, *Nature (London)* **139**, 323 (1937); *Proc. Roy. Soc. Lon. A* **165**, 199 (1938); **333**, 403 (1973).
8. H. Maeda and G. Giribet, *J. High Energy Phys.* **11**, 015 (2011).
9. V. Faraoni, D.K. Çiftci, and S. Belknap-Keet, *Phys. Rev. D* **97**, 064004 (2018).
10. R.M. Wald, *General Relativity* (Chicago University Press, Chicago, 1984).
11. V. Faraoni, *Phys. Lett. A* **245**, 26 (1998).
12. V. Faraoni, *Phys. Rev. D* **59**, 084021 (1999).
13. S. Pal, *Phys. Rev. D* **94**, 084023 (2016).
14. Y. Fujii and K. Maeda, *The Scalar-Tensor Theory of Gravity* (Cambridge University Press, Cambridge, 2003).
15. V. Faraoni, *Cosmology in Scalar Tensor Gravity*, Fundamental Theories of Physics Series vol. 139 (Kluwer Academic, Dordrecht, 2004).
16. O.A. Fonarev, *Class. Quantum Grav.* **12**, 1739 (1995).
17. V. Faraoni and S. Belknap-Keet, *Phys. Rev. D* **96**, 044040 (2017).
18. M. Campanelli and C. Lousto, *Int. J. Mod. Phys. D* **02**, 451 (1993).

# The relativity of spacetime and geometric relativistic dynamics

Yaakov Friedman*

*Jerusalem College of Technology,*
*Departments of Mathematics and Physics,*
*P.O.B. 16031 Jerusalem 91160, Israel*
**E-mail: friedman@jct.ac.il*

Tzvi Scarr

*Jerusalem College of Technology,*
*Department of Mathematics,*
*P.O.B. 16031 Jerusalem 91160, Israel*

Shmuel Stav

*Jerusalem College of Technology,*
*Department of Physics,*
*P.O.B. 16031 Jerusalem 91160, Israel*

We apply *Relativistic Newtonian Dynamics* ($RND$), a Lagrangian-based, metric theory to a static, spherically symmetric gravitational field. Using a variational principle and conserved momenta, we construct several metrics, analytic everywhere except at $r = 0$, which have $g_{01} \neq 0$ yet still leads to the same trajectories as in the Schwarzschild model. These metrics passes all classical test of $GR$. However, this model and $GR$ predict different velocities on the trajectories, both for massive objects and massless particles. The total time for a radial round trip of light in $RND$ is the same as in the Schwarzschild model, but $RND$ allows for light rays to have different speeds propagating toward and away from the massive object. One of theses metrics keeps the speed of light toward the object to be $c$. We present possible experiments to test whether $g_{01}$=0. $RND$ extends to multiple non-static forces, each of which obeys an inverse square law and whose field propagates at the speed of light.

*Keywords*: Schwarzschild metric; Geometric dynamics; Spherically symmetric metric; One-way light speed.

## 1. Introduction

The *relativity of spacetime* means that spacetime is an object-dependent notion. The potential of an electric force depends on the charge of the particle, and the particle's acceleration depends on its charge-to-mass ratio. This was recognized in the geometric approach of [1]. Given a particle, we assume that there is a metric of the particle's spacetime, defined in a preferred inertial frame, which is a function of the total potential of the forces which affect the particle.

Moreover, we assert that the particle moves *freely* in its spacetime. In other words, the force field creates the geometry of the particle's spacetime in such a way that the particle's acceleration is *zero* in its spacetime. The basis for this assumption is the fact that an inanimate object has no internal mechanism with which to change its velocity. Thus, in its world, its acceleration is zero. Note that $GR$ makes the same claim about an object freely falling in a gravitational field.

The object moves along a *geodesic* and thus has zero acceleration. We generalize this idea from *GR* and claim that *every* object is in free fall in *its spacetime*. We call this principle the *Generalized Principle of Inertia (GPI)*. Explicitly, it states that an **inanimate object moves freely, that is, with uniform motion, in *its spacetime***.

Note that if gravity is the only force around, we do not perceive the relativity of spacetime. It is masked by the Equivalence Principle, which implies that a gravitational field is felt equally by all objects, independent of their mass. In the specific case of gravity, the landscape is the same for everybody, and spacetime is not object dependent. However, in light of our goal to extend our approach to *all* forces (in particular, the electromagnetic force), we must take the relativity of spacetime into account.

Since a geodesic has minimal length, it is advantageous to use a variational principle. Thus, we will derive Euler-Lagrange type equations and obtain conservation of certain momenta. With these tools we can compute the metric.

We always have a preferred frame of reference, in which we place an observer far away from the sources of the field, where the potential is zero. For example, if we are considering the motion of Mercury, we use a frame with the sun at the center and measure distances with respect to the background of fixed stars. If we are dealing with motion within a distant galaxy, we will use the frame represented by the microwave background radiation [2, 3].

## 2. The Variational Principle and Conserved Momenta

We present the preferred frame, Lagrangian approach to geometrize the motion of an object of mass $m$, obeying the $GPI$, in a force field. Let $ds^2 = g_{\alpha\beta}(q)dq^\alpha dq^\beta$ be the metric of the object's spacetime, where $q^\alpha, \alpha = 0, 1, 2, 3$ are the preferred-frame coordinates. In order to define the length of the trajectory, we introduce a Lagrangian: $L(q, \acute{q}) = \sqrt{g_{\alpha\beta}(q)\acute{q}^\alpha \acute{q}^\beta}$. By the $GPI$, the object's trajectory $q(\sigma)$ satisfies the Euler-Lagrange equations

$$\frac{\partial L}{\partial q^\mu} - \frac{d}{d\sigma}\frac{\partial L}{\partial \acute{q}^\mu} = 0, \tag{1}$$

where $\acute{q} = \frac{dq}{d\sigma}$. It is known that the length of the trajectory does not depend on the parametrization. We choose $\sigma$ to be $\sigma = \tau = c^{-1}s$, called *proper time*, which is proportional to $s$ and reduces to the coordinate time $t$ in the classical limit. Then equation (1) becomes

$$mc\frac{\partial L}{\partial q^\mu} = \frac{dp_\mu}{d\tau}, \quad \text{where} \quad p_\mu = mc\frac{\partial L}{\partial \dot{q}^\mu} = mg_{\mu\beta}\dot{q}^\beta \tag{2}$$

is the conjugate momentum and the dot denotes differentiation by $\tau$. This imply:
**Proposition 1** If the metric coefficients $g_{\alpha\beta}$ do not depend on the coordinate $q^\mu$, then the $\mu$ component $p_\mu$ of the conjugate momentum is conserved on the trajectory.

### 3. The metric of a static, spherically symmetric gravitational field

We consider the motion of an object of mass $m$ in a static, spherically symmetric gravitational field. We use spherical coordinates $ct, r, \theta, \varphi$ in our preferred inertial frame and place the origin of our frame at the center of the symmetry of the field. The Newtonian potential is $U(r) = -GmM/r$. Let

$$u(r) = -\frac{2U(r)}{mc^2} = \frac{r_s}{r} \tag{3}$$

denote the corresponding dimensionless potential, where $r_s = \frac{2GM}{c^2}$ is the Schwarzschild radius.

From spherical symmetry and conservation of angular momentum, the $RND$ metric in our preferred frame is of the form [4]

$$ds^2 = g_{00}(r)c^2dt^2 - g_{11}(r)dr^2 - 2cg_{01}(r)dtdr - r^2d\Omega^2, \tag{4}$$

where $d\Omega^2 = d\theta^2 + \sin^2\theta d\varphi^2$. The metric coefficient $g_{01}$ may differ from zero, since we do not assume that the speed of light toward the source is the same as the speed of light away from the source. From the symmetry of the problem, we may also assume, without loss of generality, that the motion is in the plane $\theta = \pi/2$. Hence, we may replace $d\Omega^2$ with $d\varphi^2$.

Using the Newtonian limit, we can show [5] that $g_{00}(r) = 1 - u$. Conservation of angular momentum implies $r^2\dot{\varphi} = J$. Since the metric is static, the conjugate momenta $p_t$ is conserved and

$$(1-u)c\dot{t} - g_{01}\dot{r} = K. \tag{5}$$

Introducing (the negative of) the determinant $g = g_{00}g_{11} + g_{01}^2$, the *RND evolution equation for massive and massless particles* is

$$g\dot{r}^2 = -(1-u)\left(c^2\varepsilon + \frac{J^2}{r^2}\right) + K^2, \tag{6}$$

where the symbol $\varepsilon$ is defined to equal 1 for massive objects and 0 for massless particles. If $g = 1$ (and only if), the $RND$ evolution equation (6) leads to the same trajectories as in Schwarzschild spacetime. Thus, in order to predict the correct anomalous precession of Mercury, we assume from now on that $g = 1$, and the $RND$ metric of a static, spherically symmetric gravitational field becomes

$$ds^2 = (1-u(r))c^2dt^2 - \frac{1-g_{01}^2(r)}{1-u(r)}dr^2 - 2cg_{01}(r)dtdr - r^2d\varphi^2, \tag{7}$$

with $u(r)$ defined by (3). Note that the dynamics based on the metric (7) predicts the observed precession of Mercury, periastron advance of a binary star and gravitational lensing, regardless of the value of $g_{01}$.

## 4. Same trajectories but different velocities

The $RND$ metric (7) produces the same values for $\dot{r} = \frac{dr}{d\tau}$ and the same trajectories as in Schwarzschild spacetime, for both massive and massless particles. However, $RND$ and the Schwarzschild model predict different velocities $dr/dt$ in the preferred frame. From (5), we obtain $\dot{t} = (K + g_{01}\dot{r})/(c(1-u))$, implying that the radial velocity $v = \dot{r}/\dot{t}$ depends on $g_{01}$.

Hence, we expect that the magnitude of the velocity when moving from perigee to apogee is not the same as when moving in the other direction. Since the difference depends on $\dot{r}$, this can be observed for orbits with high eccentricity, and, in order that $g_{01}$ not be too small, we must be close to the Schwarzschild radius. So we expect to observe this difference for stars with high eccentricity in the neighborhood of a black hole, like the stars $S2$ and $S14$ of Sagittarius A*.

Since for light $ds^2 = 0$, the speed of light $v_\downarrow(r) \leq 0$ toward the massive object and the speed of light $v_\uparrow(r) \geq 0$ away from the source, is, using (7),

$$v_\uparrow(r) = c\frac{1-u(r)}{1+g_{01}(r)},\ v_\downarrow(r) = -c\frac{1-u(r)}{1-g_{01}(r)}. \tag{8}$$

Having measured the two one-way speeds, one then computes

$$g_{01} = \frac{v_\downarrow + v_\uparrow}{v_\downarrow - v_\uparrow}. \tag{9}$$

Since the metric is, by assumption, asymptotically flat, we see that as $r$ approaches infinity, the magnitude of both $v_\downarrow(r)$ and $v_\uparrow(r)$ approach the inertial frame speed of light $c$, as expected.

$GR$ and $RND$ disagree on the one-way speeds, but the $RND$ time of a round trip for light is the same as in the Schwarzschild spacetime [4]. This means that a round-trip speed of light experiment cannot distinguish between $GR$ and $RND$. For this, we need an experiment which measures both one-way speeds of light. We have also shown that $RND$ predicts the observed the Shapiro time delay if $g_{01}(r) \ll 1$. Here, even though $g_{01} \neq 0$, the $RND$ prediction agrees with $GR$. The reason is that the trajectory of light as it passes the Sun is similar to a round trip, since the light first approaches the Sun and then travels away from the Sun.

Let $g_{01}(r) = u^n(r)$, for $n = 1, 2, 3, ....$ Then the metric coefficients are polynomials of degree $2n-1$, and the metric is analytic everywhere except at the origin. Of particular interest is the first-order polynomial metric, for which $g_{01}(r) = u(r)$. In this case $ds^2 = \eta_{\mu\nu}dx^\mu dx^\nu - u(cdt + dr)^2$. For this metric, the two directional radial light velocities are

$$v_\uparrow(r) = \frac{1-u(r)}{1+u(r)}c \quad \text{and} \quad v_\downarrow(r) = -c. \tag{10}$$

The speed of light $v_\uparrow$ vanishes at the Schwarzschild radius. However, unlike the Schwarzschild model, light travelling toward the source does not slow down. We see this as more natural than the Schwarzschild prediction.

## 5. Testing the $g_{01} = 0$ prediction of $GR$ for a spherically symmetric gravitational field

It can be shown that Einstein's field equations imply that $g_{01} = 0$. Here, we propose two experiments, one for massive bodies and one for light, to test the $GR$ prediction that $g_{01} = 0$ and to determine the value of $g_{01}$ in Nature.

The first experiment, as mentioned above, is to observe orbital velocities of stars with high eccentricity near black holes and calculate $g_{01}$ from equation (9).

The second experiment tests the influence of the gravitational field of the Earth on the one-way speed of light. We have designed an experiment using two $193nm$ sub-mW lasers [6] of Optica Photonics, with coherence length $> 150m$, see Figure 1. For the first-order polynomial metric, from (10), we would expect the "toward" light (a) and the "away from" light (b) from laser $A$ to produce different phase shifts at $d = 10m$ of $0.3\pi$, which should be detectable by interference with the laser $B$. The influence of the frequency of the light, caused by the gravitational redshift, predicted by $GR$ and tested in the Pound-Rebka experiment [7], should be negligible with respect to the effect we are testing.

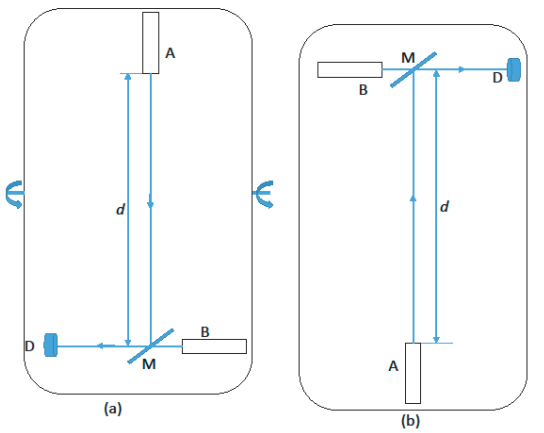



Fig. 1. Experimental setup.

## 6. More forces and more sources

First, we extend $RND$ to an arbitrary static, conservative force . Let $U(\mathbf{x})$ denote the potential of the field at the spatial point $\mathbf{x}$. Introduce, at each $\mathbf{x}$ where $\nabla U(\mathbf{x}) \neq \mathbf{0}$, a normalized vector $\mathbf{n}(\mathbf{x})$ in the direction of $\nabla U(\mathbf{x})$. Let $d\mathbf{x}_n = (d\mathbf{x} \cdot \mathbf{n})\mathbf{n}$ and $d\mathbf{x}_{tr} = d\mathbf{x} - (d\mathbf{x} \cdot \mathbf{n})\mathbf{n}$, respectively, denote the projections of the spatial increment

$d\mathbf{x}$ in the parallel and transverse directions to $\mathbf{n}(\mathbf{x})$. In this notation, as in (7), the $RND$ metric is

$$ds^2 = (1 - u(\mathbf{x}))c^2dt^2 - \frac{1 - g_{01}^2(\mathbf{x})}{1 - u(\mathbf{x})}d\mathbf{x}_n^2 - 2cg_{01(\mathbf{x})}dtd\mathbf{x}_n - d\mathbf{x}_{tr}^2. \tag{11}$$

It can be shown that the metric can be continuously extended to points where $\nabla U$ vanishes.

Consider now motion of an object in a field generated several sources affecting it. We assume that the field of each source $j$ obeys an inverse square law with potential $U(r) = -k_j/r$ in its comoving frame and propagates with the speed of light in our preferred frame. We use Lorentz covariance and an approach similar to that of the Liénard-Wiechert potentials (see, for example, [8], pp. 174-6) and [9]. We find that the four-potential per unit mass at space-time $x$ is

$$A_\alpha(x) = \Sigma_j \frac{-k_j(w_j)_\alpha}{(x_j'(\tau_j(x)) - x) \cdot w_j}, \tag{12}$$

where $x_j'(\tau_j(x)) - x$ is the relative position of the source $j$ at the retarded time $\tau_j(x)$, and $w_j$ is its four-velocity the source at that time. Note that this four-potential tends to 0 when $x$ is far removed from the sources.

Since each individual four-potential is a timelike vector, the total four-potential is also timelike. Therefore, we can apply a Lorentz transformation $\Lambda$ from $K$ to an inertial frame $K'$ such that $\Lambda(A) = (U, 0, 0, 0)$. In $K'$, the four-potential behaves like a "static" potential. Therefore, in $K'$, the metric is defined by formula (11), with $u$ defined by (3). To obtain the metric in $K$, apply the inverse Lorentz transformation $\Lambda^{-1}$ to the metric in $K'$.

## References


1. Duarte, C.: *Int. J. Geom. Meth. Mod. Phys.* **12** (2015).
2. A. A. Penzias and R. W. Wilson: *Ap. J.* **142** 419 (1965).
3. P.J.E. Peebles: Principles of Physical Cosmology. Princeton University Press, London (1993).
4. Friedman, Y. and Stav, S.: New metrics of a spherically symmetric gravitational field passing classical tests of General Relativity (2019) https://www.researchgate.net/publication/328504942.
5. Friedman, Y. and Scarr, T.: *Europhys. Lett.* (2019).
6. https://www.toptica.com/products/customized-solutions/193-nm-sub-mw.
7. Pound, R. V. and Rebka Jr., G. A.: *Phys. Rev. Lett.* **4** 337 (1960).
8. Landau, L. and Lifshitz, E.: The Classical Theory of Fields. Course of Theoretical Physics Vol. 2 Pergamon, Addison-Wesley (1971).
9. Mashhoon, B.: Gravitoelectromagnetism in Pascual-Sanchez, J.-F., Floria, L., San Miguel, A., Vicente, F. eds. Reference Frames and Gravitomagnetism. World Scientific 121 (2001).

# Doubly-foliable space-times and gauge-fixing of perturbations in scalar-tensor gravity theories

Cecília Gergely*, Zoltán Keresztes and László Árpád Gergely

*Institute of Physics, University of Szeged, Hungary*
** E-mail: lawrencesterne92@gmail.com*
*http://www.staff.u-szeged.hu/ gergely/group/group.html*

The direct detection of gravitational waves and gamma-ray counterparts has confirmed that gravitational waves propagate with the speed of light, disruling some of the scalar-tensor gravity models. A huge class of Horndeski theories (those with generic G2 and G3, and with G4 depending only of the scalar field) however survived the test. The study of perturbations of such models is important to establish the ghost-free and instability-free parameter regimes. This has been investigated for a wide range of scalar-tensor theories in the spherically symmetric setup, exploring a 2+1+1 decomposition of space-time based on an orthogonal double foliation. The orthogonality however consumed one gauge degree of freedom, allowing the discussion of only the odd sector of the scalar part of perturbations. In order to describe the even sector perturbations, we worked out a novel 2+1+1 decomposition of space-time and gravitational dynamics, based on a non-orthogonal double foliation. We explore this new formalism for the perturbations of both the spherically symmetric metric tensor and scalar field in generic scalar-tensor theories, achieving an unambiguous gauge-fixing. This opens up the way for the discussion of the full spectrum of perturbations of spherically symmetric scalar-tensor gravity, including both the odd and even sectors of the scalar part of perturbations.

*Keywords*: nonorthogonal 2+1+1 decomposition; scalar-tensor gravity, gauge-fixing.

## 1. Introduction

While General Relativity has been a temendously successful theory in establishing the compatibility of the gravitational interaction with the Special Relativistic constraint for signal propagation and has been tested successfully in all sorts of observations both on the Solar System scale and beyond, its extensions have been also systematically investigated.

Their necessity emerged due to A) the expectation to combine the geometric theory of gravity with quantum considerations, B) the need to understand the mysterious, yet at present dominating component of the Universe, dark energy, which drives its accelerating expansion, and C) the yet unsuccessful attempts to identify dark matter through its weak interaction, rending it at present entirely in the gravitational sector.

Possible extensions arise from both modifying the dynamics of the metric tensor and from including secondary fields. Such extensions could be tested, constrained, or falsified through precision cosmological observations or analysis of detected gravitational waves.

Indeed, the recent eleven-fold detection of gravitational waves by the LIGO Scientific Collaboration and Virgo Collaboration[1–7] confirmed the propagation speed of gravitational waves (tensorial perturbations) to be the speed of light within $-3 \times 10^{-15}$ and $+7 \times 10^{-16}$ accuracy.[8] The surviving theories could be

further subjected to stability analysis, demanding the avoidance of Ostrogradsky-instabilities, ghosts, Laplacian instabilities or tachyons.

In the framework of scalar-tensor gravity the Horndeski theories[9,10] are known to ensure second-order dynamics for both the metric tensor and the scalar field. Demanding the speed of light to be the speed of propagation of tensorial modes, the $L_5$ contribution to the Lagrangian has been disruled, together with the kinetic term dependence of $L_4$.[11–14] The latter could be kept for a class of the beyond Horndeski models,[15,16] which allow for higher-than-second order spatial derivatives, but no such time derivatives.

Both classes of theories were excessively analysed in a cosmological context. Black hole perturbations were also explored. For perturbations of spherically symmetric, static black holes the odd and even sector perturbations decouple. Both were analysed for the Horndeski class of models,[17,18] while for the beyond Horndeski models only the stability analysis of the odd sector has been carried on,[19] based on an orthogonal double foliation.[20,21] The orthogonality requirement however resulted in a gauge fixing, which was unambiguous for the odd sector, nevertheless generated an arbitrary function of all coordinates but the radial for the even sector, hampering the physical interpretation of the even sector perturbations.

In order to avoid this shortcoming we proposed[22] to develop another decomposition of space-time and all affected geometrical quantities along a nonorthogonal double foliation. The details of this work involve a quadruple set of embedding variables (extrinsic curvatures, normal fundamental forms and normal fundamental scalars).[23] These variables provide the set of velocity phase-space variables complementing the induced 2-metric, radial shift and radial lapse. A Hamiltonian analysis established the canonical coordinates and momenta in terms of these geometric variables in a General Relativistic setup, also the Hamiltonian and diffeomorphism constraints.[23] The complementary metric variables, the temporal shift and lapse are left arbitrary, hence an unambiguous gauge-fixing became possible.

## 2. The nonorthogonal double foliation

The space-time $\mathcal{B}$ is foliated with the 3-dimensional hypersurfaces $\mathcal{S}_t$ (spacelike, with normal $n^a$) and $\mathfrak{M}_\chi$ (timelike, with normal $l^a$) , their intersection being the 2-dimensional surface $\Sigma_{t\chi}$ (with $g_{ab}$ the induced metric, normals $k^a$ along $\mathfrak{M}_\chi$ and $m^a$ along $\mathcal{S}_t$). The 4-dimensional metric is 2+1+1 decomposed in two equivalent ways, employing the orthonormal bases $(n^a, m^a)$ and $(k^a, l^a)$:

$$\tilde{g}_{ab} = -n_a n_b + m_a m_b + g_{ab} \ , \tag{1}$$

$$\tilde{g}_{ab} = -k_a k_b + l_a l_b + g_{ab} \ . \tag{2}$$

The evolution vectors along a temporal coordinate $t$ and a spatial one $\chi$ running in $\mathfrak{M}_\chi$ can also be decomposed in these bases:

$$\left(\frac{\partial}{\partial t}\right)^a = N n^a + N^a + \mathcal{N} m^a = \frac{N}{\mathfrak{c}} k^a + N^a \ , \tag{3}$$

$$\left(\frac{\partial}{\partial\chi}\right)^a = Mm^a + M^a = M\left(-\mathfrak{s}k^a + \mathfrak{c}l^a\right) + M^a \ , \tag{4}$$

Here $N^a$ is a 2-dimensional shift, $M^a$ is a radial shift; $N$ and $M$ the lapse and radial lapse, $\mathfrak{s} = \sinh\phi$, $\mathfrak{c} = \cosh\phi$. The new element of the formalism is the third component $\mathcal{N} = N\tanh\phi$ of the 3-dimensional shift vector, which in the orthogonal limit vanishes.

The 2+1+1 decomposition of the covariant derivatives of the basis vectors

$$\tilde{\nabla}_a n_b = K_{ab} + 2m_{(a}\mathcal{K}_{b)} + m_a m_b\mathcal{K} + n_a m_b\mathcal{L}^* - n_a\mathfrak{a}_b \ , \tag{5}$$

$$\tilde{\nabla}_a l_b = L_{ab} + 2k_{(a}\mathcal{L}_{b)} + k_a k_b\mathcal{L} + l_a k_b\mathcal{K}^* + l_a\mathfrak{b}_b \ , \tag{6}$$

$$\tilde{\nabla}_a k_b = K^*_{ab} + l_a\mathcal{K}^*_b + l_b\mathcal{L}_a + l_a l_b\mathcal{K}^* + k_a l_b\mathcal{L} - k_a\mathfrak{a}^*_b \ , \tag{7}$$

$$\tilde{\nabla}_a m_b = L^*_{ab} + n_a\mathcal{L}^*_b + n_b\mathcal{K}_a + n_a n_b\mathcal{L}^* + m_a n_b\mathcal{K} + m_a\mathfrak{b}^*_b \ , \tag{8}$$

define a set of geometrical quantities, which caracterize the embedding of $\Sigma_{t\chi}$: extrinsic curvatures ($K_{ab},\ L_{ab},\ K^*_{ab},\ L^*_{ab}$); accelerations ($\mathfrak{a}_a,\ \mathfrak{a}^*_a$) of the timelike congruences ($n^a,\ k^a$), and similar "accelerations" ($\mathfrak{b}_a,\ \mathfrak{b}^*_a$) of the spacelike congruences ($l^a,\ m^a$); normal fundamental scalars ($\mathcal{K},\ \mathcal{L},\ \mathcal{K}^*,\ \mathcal{L}^*$); normal fundamental forms ($\mathcal{K}_a,\ \mathcal{L}_a$) and similarly defined quantities ($\mathcal{K}^*_a,\ \mathcal{L}^*_a$) for the basis vectors ($m^a,\ k^a$) with vorticities

$$\begin{aligned}\hat{\omega}^{(\mathbf{k})}_{ab} l^b &= \frac{1}{2}D_a\phi + \frac{\mathfrak{s}}{2\mathfrak{c}}\left(\mathfrak{a}_a + \mathfrak{b}_a\right) \ , \\ \check{\omega}^{(\mathbf{m})}_{ab} n^b &= \frac{1}{2}D_a\phi - \frac{\mathfrak{s}}{2\mathfrak{c}}\left(\mathfrak{a}_a + \mathfrak{b}_a\right) \ .\end{aligned} \tag{9}$$

These are all depicted on Fig. 1. They exhibit a complex set of interdependence relations, meticuously explored.[23]

## 3. Gauge fixing in scalar-tensor gravitational theories

In this section we summarize the main result[23] in the gauge-fixing of the perturbations of the spherically symmetric, static background in an unspecified scalar-tensor theory. Similarly to the unitary gauge in cosmology, here we aim to achieve a "radial unitary" gauge, thus to absorb the perturbation of the scalar field in a suitable choice of coordinates. On the spherically symmetric, static background (where quantities carry an overbar) Helmholtz-like theorems assure the decompositions

$$\delta N_a = \bar{D}_a P + E^b{}_a\bar{D}_b Q \, , \tag{10a}$$

$$\delta M_a = \bar{D}_a V + E^b{}_a\bar{D}_b W \, , \tag{10b}$$

$$\delta g_{ab} = \bar{g}_{ab}A + \bar{D}_a\bar{D}_b B + \frac{1}{2}\left(\, E^c{}_a\bar{D}_c\bar{D}_b + E^c{}_b\bar{D}_c\bar{D}_a\right) C \, , \tag{10c}$$

with $E_{ab} = \sqrt{\bar{g}}\,\varepsilon_{ab}$ the 2-dimensional Levi-Civita tensor and $\varepsilon_{\theta\varphi} = 1$. The scalars $(N, M, \mathcal{N}, P, V, A, B)$ and $(Q, W, C)$ form the even and odd sectors of the perturbations, respectively, and they decouple. The latter represent the tensor degrees of freedom.[17]

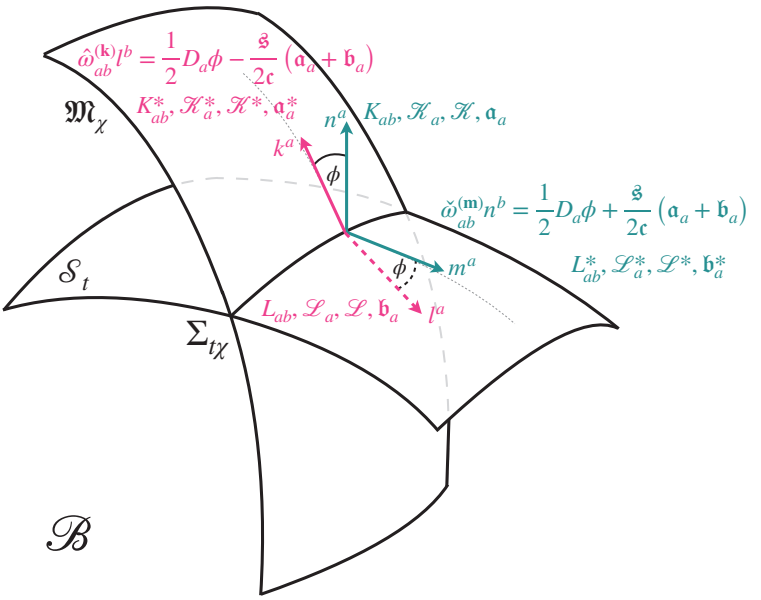



Fig. 1. The bases adapted to the two foliations, their embedding variables and the nonvanishing vorticities.

Next, we explore the diffeomorphism gauge freedom, with generators

$$\left(\xi^t, \xi^\chi, \xi^a = \bar{D}^a \xi + E^{ba} \bar{D}_b \eta\right) \quad , \qquad (a = \theta, \varphi) \tag{11}$$

also Helmholtz-decomposed. The initial metric and scalar perturbations are related to those (denoted by overhat) emerging after the diffeomorphism as

$$\mathfrak{L}_\xi \tilde{g}_{ab} = \delta\tilde{g}_{ab} - \widehat{\delta\tilde{g}_{ab}} \quad , \quad \mathfrak{L}_\xi \phi = \delta\phi - \widehat{\delta\phi} \quad , \tag{12}$$

which generate the respective variations of the even and odd sector variables.[19] In the framework of the nonorthogonal double foliation we managed to unambiguously fix the gauge through the choice

$$\xi^t = \frac{P + \dot{\xi}}{\bar{N}^2} , \qquad \xi^\chi = \frac{\delta\phi}{\bar{\phi}'} , \qquad \xi = -\frac{B}{2} , \qquad \eta = -\frac{C}{2} , \tag{13}$$

which ensure the radial unitary gauge $\widehat{\delta\phi} = 0$, the perturbation of the 2-metric being a conformal rescaling ($\widehat{B} = 0 = \widehat{C}$) and $\widehat{P} = 0$, a condition impossible to achieve in the ortogonal double foliation formalism, which consumed the respective degree of freedom.

Hence the odd sector consist of $\left(\widehat{Q}, \widehat{W}\right)$, while the even sector is formed by $\left(\widehat{N}, \widehat{M}, \widehat{\mathcal{N}}, \widehat{V}, \widehat{A}\right)$. The former has been extensively discussed,[19] while the study of the latter became possible with the unambiguous gauge choice emerging in the nonorthogonal double foliation formalism and will be presented elsewhere.

**Acknowledgments**

The participation of the authors at the MG15 meeting was supported by the Hungarian National Research Development and Innovation Office (NKFIH) in the form of the grant 123996. This work has been carried out in the framework of COST actions CA15117 (CANTATA) and CA16104 (GWverse), supported by COST (European Cooperation in Science and Technology). During the preparation of this manuscript C.G. was supported by the UNKP-18-2 New National Excellence Program of the Ministry of Human Capacities of Hungary while Z.K. was supported by both the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the UNKP-18-4 New National Excellence Program of the Ministry of Human Capacities of Hungary.

**References**

1. LIGO Scientific Collaboration and Virgo Collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, 061102 (2016) [arXiv:1602.03837 [gr-qc]].
2. LIGO Scientific Collaboration and Virgo Collaboration, *GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence*, Phys. Rev. Lett. **116**, 241103 (2016) [arXiv:1606.04855 [gr-qc]].
3. LIGO Scientific Collaboration and Virgo Collaboration, *GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2,* Phys. Rev. Lett. **118**, 221101 (2017) [arXiv:1706.01812 [gr-qc]].
4. LIGO Scientific Collaboration and Virgo Collaboration, *GW170608: Observation of a 19-Solar-Mass Binary Black Hole Coalescence,* Astrophys. J. Lett. **851**, L35 (2017) [arXiv:1711.05578 [astro-ph.HE]].
5. LIGO Scientific Collaboration and Virgo Collaboration, *GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence,* Phys. Rev. Lett. **119**, 141101 (2017) [arXiv:1709.09660 [gr-qc]].
6. LIGO Scientific Collaboration and Virgo Collaboration, *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral,* Phys. Rev. Lett. **119**, 161101 (2017) [arXiv:1710.05832 [gr-qc]].
7. LIGO Scientific Collaboration and Virgo Collaboration, *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, (2018) [arXiv:181112907 [astro-ph.HE]].
8. LIGO Scientific Collaboration and Virgo Collaboration, Fermi Gamma-ray Burst Monitor, and INTEGRAL*, Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB170817A,* Astrophys. J. Lett. **848**, L13 (2017).
9. G.W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, Int. J. Theor. Phys. **10**, 363–384 (1974).

10. C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, *From k-essence to generalized Galileons*, Phys. Rev. D **84**, 064039 (2011).
11. T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki, *Strong constraints on cosmological gravity from GW170817 and GRB 170817A*, Phys. Rev. Lett. **119**, 251301 (2017) [arXiv:1710.06394 [astro-ph.CO]].
12. J. M. Ezquiaga and M. Zumalacárregui, *Dark Energy after GW170817: Dead ends and the road ahead*, Phys. Rev. Lett. **119**, 251304 (2017) [arXiv:1710.05901 [astro-ph.CO]].
13. P. Creminelli and F. Vernizzi, *Dark Energy after GW170817 and GRB170817A*, Phys. Rev. Lett. **119**, 251302 (2017) [arXiv:1710.05877 [astro-ph.CO]].
14. J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, Phys.Rev.Lett. **119**, 251303 (2017) [arXiv:1710.05893 [astro-ph.CO]].
15. J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Healthy theories beyond Horndeski*, Phys. Rev. Lett. **114**, 211101 (2015) [arXiv:1404.6495 [hep-th]].
16. J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Essential building blocks of dark energy*, J. Cosmol. Astropart. Phys. 08 (2013) 025 [arXiv:1304.4840 [hep-th]].
17. T. Kobayashi, H. Motohashi and T. Suyama, *Black hole perturbation in the most general scalar-tensor theory with second-order field equations I: The odd-parity sector,* Phys. Rev. D **85**, 084025 (2012) [arXiv:1202.4893 [gr-qc]].
18. T. Kobayashi, H. Motohashi and T. Suyama, *Black hole perturbation in the most general scalar-tensor theory with second-order field equations II: the even-parity sector*, Phys. Rev. D **89**, 084042 (2014) [arXiv:1402.6740 [gr-qc]].
19. R. Kase, L. Á. Gergely and S. Tsujikawa, *Effective field theory of modified gravity on spherically symmetric background: leading order dynamics and the odd mode perturbations*, Phys. Rev. D **90**, 124019 (2014) [arXiv:1406.2402 [hep-th]].
20. L. Á. Gergely and Z. Kovács, *Gravitational dynamics in s+1+1 dimensions*, Phys. Rev. D **72**, 064015 (2005) [arXiv:gr-qc/0507020].
21. Z. Kovács and L. Á. Gergely, *Gravitational dynamics in s+1+1 dimensions II. Hamiltonian theory*, Phys. Rev. D **77**, 024003 (2008) [arXiv:0709.2131 [gr-qc]].
22. C. Gergely, Z. Keresztes and L. Á. Gergely, *Hamiltonian Dynamics of Doubly-Foliable Space-Times*, Universe, 4(1), 9 (2018).
23. C. Gergely, Z. Keresztes and L. Á. Gergely, *Gravitational dynamics in a 2+1+1 decomposed spacetime along nonorthogonal double foliations: Hamiltonian evolution and gauge fixing*, submitted (2019).

# Shock-waves in the gravitational wave compatible Horndeski theories with linear kinetic term

László Árpád Gergely* and Bence Racskó

*Institute of Physics, University of Szeged, Hungary*
**E-mail: laszlo.a.gergely@gmail.com*
*http://www.staff.u-szeged.hu/ gergely/group/group.html*

Propagating shock-waves can be discussed in terms of junction conditions between space-time regions separated by a hypersurface. Recent observations of gravitational waves and their electromagnetic counterparts established that the former also propagate with the speed of light. Hence energetic gravitational waves could be perceived as shock-waves on null hypersurfaces. The most generic scalar-tensor theories with at most second order dynamics, the Horndeski-theories were severely constrained. We derive junction conditions across a null hypersurface for the subclass of allowed Horndeski-theories with linear kinetic term dependence, exploring a formalism based on a transverse null vector. We obtain a 2+1 decomposed generalised Lanczos equation, with the jump of the transverse curvature induced by both the distributional energy-momentum tensor of the wavefront of the shock-wave, and by the jump in the transverse derivative of the scalar. The surface density, current and pressure of the distributional light-like shock-wave and the transverse derivative of the scalar are also constrained by a scalar junction equation.

*Keywords*: Scalar-tensor gravity; junction conditions; null-hypersurfaces.

## 1. Introduction

In general relativity, matching two solutions of the Einstein field equations accross a hypersurface separating them is realized through the Israel junction conditions, provided the hypersurface is timelike or spacelike.[1] Across a boundary hypersurface, they prescribe the continuity of the induced metric and of the extrinsic curvature. The discontinuity of the extrinsic curvature is also admissible for hypersurfaces referred to as thin shells. The Lanczos equation relates the jump of the extrinsic curvature to the distributional stress-energy tensor $\mathcal{T}^{\mu\nu}$ of the thin shell.

Israel's procedure breaks down if the hypersurface is null. This is a physically relevant case, as impulsive shock-waves traveling at the speed of light are described by null hypersurfaces, as are event horizons. The generalization of the Israel junction conditions to arbitrary hypersurfaces has been developed by Barrabés and Israel.[2] If the hypersurface is specified to be null, then in Poisson's reformulation, the junction conditions take a particularily simple form.[3]

The most general scalar-tensor theory with second order equations of motion for both the scalar and the metric tensor was given by Horndeski.[4,5] Recent gravitational wave observations[6–12] have severely constrained this class. The subclass guaranteeing the propagation of the tensorial modes with the speed of light[13–16] has a Lagrangian for the scalar field often referred to as kinetic gravity braiding theory.[17,18] The coupling of the curvature with the scalar can be through an unspecified function of the scalar field, in the Jordan frame. Adding these contributions together we get a generalised kinetic gravity braiding class of scalar-tensor theories.

Junction conditions across timelike and spacelike hypersurfaces in the full Horndeski class have been found,[19,20] but the null case has not been discussed. This case is important, as all electromagnetic and gravitational waves propagate on the light-cone.

In this work we address the junction conditions accross null hypersurfaces in a subset of the generalised kinetic gravity braiding class of theories, namely those in which the Lagrangian is linear in the kinetic term $X = -(\nabla\phi)^2/2$:

$$L = B(\phi)X + V(\phi) - 2\xi(\phi)\Box\phi X + \frac{1}{2}F(\phi)R \tag{1}$$

Here $B, V, \xi, F$ are arbitrary smooth functions of the dynamical scalar field $\phi$. In the following we denote by $\langle Q\rangle$ and $[Q]$ the average and jump of the quantity $Q$ across the null hypersurface. Greek and capital latin indices denote spacetime and two-dimensional spatial indices, respectively.

## 2. Junction conditions

Following Poisson,[3] we employ a pseudo-orthonormal basis with two null vectors $N^\mu$ and $L^\mu$, normalized as $L^\mu N_\mu = -1$. Here $N^\mu$ is the normal (surface gradient, which is also tangent), while $L^\mu$ plays the role of the transverse vector, with respect to which we perform a (2+1)+1 decomposition. Across the hypersurface, the continuity of both the metric tensor $g_{\mu\nu}$ and scalar $\phi$ are required. Their first derivatives can have jumps $c_{\mu\nu}$ and $\zeta$ given by

$$[\partial_\kappa g_{\mu\nu}] = -N_\kappa c_{\mu\nu}, \tag{2}$$

$$[\partial_\mu \phi] = -N_\mu \zeta, \tag{3}$$

hence the second order derivatives of the metric and scalar fields acquire distributional contributions along the thin shell. The distributional equations of motion take the form

$$\mathscr{E}^{\mu\nu} = \frac{1}{2}\mathscr{T}^{\mu\nu} \tag{4}$$

$$\mathscr{E}^{\phi} = 0, \tag{5}$$

where $\mathscr{E}^{\mu\nu}$ and $\mathscr{E}^{\phi}$ are the singular parts of the tensorial and scalar Euler-Lagrange expressions, respectively. The distributional stress-energy-momentum tensor of the shockwave may be decomposed with respect to the basis $\{L, N, e_2, e_3\}$ (where $e_2, e_3$ are spatial basis vectors orthogonal to both $L$ and $N$) as

$$\mathscr{T}^{\mu\nu} = \rho N^\mu N^\nu + j^A(N^\mu e^\nu_A + e^\mu_A N^\nu) + p^{AB} e^\mu_A e^\nu_B, \tag{6}$$

with the expressions $\rho, j^A, p^{AB}$ interpreted as surface energy density, surface current and surface stress, respectively. These quantities could entirely be given in terms of internal hypersurface coordinates, hence they are calculatable even if the bulk coordinates are discontinuous across the hypersurface.

The explicit form of these components for the Lagrangian (1) have been derived[21] as

$$\rho = F(\phi)[\mathcal{K}_{AB}]q^{AB} + F'(\phi)[\phi_L] - 2\xi(\phi)[\phi_L^2]\phi_N, \tag{7}$$

$$j^A = -F(\phi)[\mathcal{K}_N{}^A] + 2\xi(\phi)\phi_N\phi^A[\phi_L], \tag{8}$$

$$p^{AB} = pq^{AB}, \tag{9}$$

$$p = F(\phi)[\mathcal{K}_{NN}] - 2\xi(\phi)[\phi_L]\phi_N^2, \tag{10}$$

where $\phi_L, \phi_N$ are derivatives along $L^\mu$ and $N^\mu$ respectively, $\phi_A$ is the derivative along $e_A^\mu$, $q_{AB} = g(e_A, e_B)$ is the two-dimensional induced metric, $q^{AB}$ its inverse, capital latin indices are raised and lowered with the $q$ metrics, and $\mathcal{K}$ is the "transverse curvature" defined as

$$\mathcal{K}_{ab} = e_a^\mu e_b^\nu \nabla_\mu L_\nu, \tag{11}$$

where $a, b = 1, 2, 3$ and $e_1^\mu \equiv e_N^\mu = N^\mu$. The jump of the transverse curvature is related to the jump of the metric derivative as $[\mathcal{K}_{ab}] = \frac{1}{2}e_a^\mu e_b^\nu c_{\mu\nu}$ and the jump of the scalar field's transverse derivative is $[\phi_L] = \zeta$. Equations (7-10) represent the 2+1 decomposition of the tensorial equation (4), a generalization of the Lanczos equation.

The scalar equation (5) for the Lagrangian (1) becomes

$$\begin{aligned} 0 = {} & \xi(\phi)\phi_N^2 q^{AB}[\mathcal{K}_{AB}] - 2\xi(\phi)\phi_N\phi^A[\mathcal{K}_{NA}] + F'(\phi)[\mathcal{K}_{NN}] \\ & - 2\xi(\phi)[\phi_L]\phi_{NN} + 2\xi(\phi)\phi_N[\phi_L\mathcal{K}_{NN}] \end{aligned} \tag{12}$$

This equation contains the jumps of various components of the transverse curvature, which in principle are expressible[21] from the generalized Lanczos equations (7-10) in terms of the surface energy density, current, pressure and the jump of the transverse derivative of the scalar field. Hence, the scalar equation (12) could be perceived as a constraint on the distributional sources, mediated by the cubic derivative coupling $\xi$ and the jump of the transverse derivative of the scalar and of its square.

The expression $[\phi_L\mathcal{K}_{NN}]$ can be transformed away[21] by a suitable gauge in the tetrad choice. The normal vector field is autoparallel[3]

$$N^\nu\nabla_\nu N^\mu = \kappa N^\mu \tag{13}$$

with the *non-affinity parameter* $\kappa = \mathcal{K}_{NN}$. If the null fields are rescaled as $\bar{N}^\mu = e^\alpha N^\mu$ and $\bar{L}^\mu = e^{-\alpha}L^\mu$ with some function $\alpha$ defined on the hypersurface, the non-affinity parameter transforms as

$$\bar{\kappa} = e^\alpha \left(N^\nu\nabla_\nu\alpha + \kappa\right). \tag{14}$$

It is possible to set

$$[\phi_{\bar{L}}\bar{\kappa}] = 0 \tag{15}$$

by solving the differential equation

$$\frac{\partial \alpha}{\partial \lambda} = -\langle \kappa \rangle - \frac{\langle \phi_L \rangle}{[\phi_L]}[\kappa], \tag{16}$$

where $\lambda$ is a coordinate adapted to $N^\mu$. Therefore, with this choice of gauge, the last term of Eq. (12) vanishes.

## 3. Concluding remarks

We have presented generic tensorial and scalar junction conditions across null hypersurfaces in generalized kinetic gravity braiding theories with linear kinetic term dependence.

The tensorial equation emerges as a generalized Lanczos equation, with the 2+1 decomposition (7-10) based on a pseudo-orthonormal basis in which the null vector is the normal. The jump of the scalar derivative along a second null vector, which is transverse to the hypersurface, also enters the generalized Lanczos equation.

The scalar equation (12) further constrains the energy, current, pressure of the shock-wave and the jump of the transverse derivative of the scalar field. In the absence of the cubic derivative coupling $\xi$, this simply reduces to the condition of vanishing isotropic pressure of the shock-wave, a property already revealed in the case of Brans-Dicke theory in the Jordan frame.[22]

## Acknowledgments

The participation of LÁG at the MG15 meeting was supported by the Hungarian National Research Development and Innovation Office (NKFIH) in the form of the grant 123996. This work has been carried out in the framework of COST actions CA15117 (CANTATA) and CA16104 (GWverse), supported by COST (European Cooperation in Science and Technology). During the preparation of this manuscript B.R. was supported by the UNKP-18-3 New National Excellence Program of the Ministry of Human Capacities of Hungary.

## References

1. W. Israel, *Singular hypersurfaces and thin shells in general relativity*, Nouvo Cim. B, **44**, 1-14 (1966).
2. C. Barrabès, W. Israel, *Thin shells in general relativity and cosmology: The lightlike limit*, Phys. Rev. D, **43**, 1129-1142 (1991).
3. E. Poisson, *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics*, Cambridge University Press: Cambridge, UK (2004).
4. G.W. Horndeski, *Second-order scalar-tensor field equations in a four-dimensional space*, Int. J. Theor. Phys. **10**, 363–384 (1974).

5. C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, *From k-essence to generalized Galileons*, Phys. Rev. D **84**, 064039 (2011).
6. LIGO Scientific Collaboration and Virgo Collaboration, *Observation of Gravitational Waves from a Binary Black Hole Merger*, Phys. Rev. Lett. **116**, 061102 (2016) [arXiv:1602.03837 [gr-qc]].
7. LIGO Scientific Collaboration and Virgo Collaboration, *GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence*, Phys. Rev. Lett. **116**, 241103 (2016) [arXiv:1606.04855 [gr-qc]].
8. LIGO Scientific Collaboration and Virgo Collaboration, *GW170104: Observation of a 50-Solar-Mass Binary Black Hole Coalescence at Redshift 0.2,* Phys. Rev. Lett. **118**, 221101 (2017) [arXiv:1706.01812 [gr-qc]].
9. LIGO Scientific Collaboration and Virgo Collaboration, *GW170608: Observation of a 19-Solar-Mass Binary Black Hole Coalescence,* Astrophys. J. Lett. **851**, L35 (2017) [arXiv:1711.05578 [astro-ph.HE]].
10. LIGO Scientific Collaboration and Virgo Collaboration, *GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence,* Phys. Rev. Lett. **119**, 141101 (2017) [arXiv:1709.09660 [gr-qc]].
11. LIGO Scientific Collaboration and Virgo Collaboration, *GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral,* Phys. Rev. Lett. **119**, 161101 (2017) [arXiv:1710.05832 [gr-qc]].
12. LIGO Scientific Collaboration and Virgo Collaboration, *GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs*, (2018) [arXiv:181112907 [astro-ph.HE]].
13. T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki, *Strong constraints on cosmological gravity from GW170817 and GRB 170817A*, Phys. Rev. Lett. **119**, 251301 (2017) [arXiv:1710.06394 [astro-ph.CO]].
14. J. M. Ezquiaga and M. Zumalacárregui, *Dark Energy after GW170817: Dead ends and the road ahead*, Phys. Rev. Lett. **119**, 251304 (2017) [arXiv:1710.05901 [astro-ph.CO]].
15. P. Creminelli and F. Vernizzi, *Dark Energy after GW170817 and GRB170817A*, Phys. Rev. Lett. **119**, 251302 (2017) [arXiv:1710.05877 [astro-ph.CO]].
16. J. Sakstein and B. Jain, *Implications of the Neutron Star Merger GW170817 for Cosmological Scalar-Tensor Theories*, Phys. Rev. Lett. **119**, 251303 (2017) [arXiv:1710.05893 [astro-ph.CO]].
17. R. Kase and S. Tsujikawa, *Dark energy in Horndeski theories after GW170817: A review*, Int. J. Mod. Phys. D, in press (2019) [arXiv:1809.08735 [gr-qc]].
18. C. Deffayet, O. Pujolas, I. Sawicki, A. Vikman, *Imperfect Dark Energy from Kinetic Gravity Braiding*, JCAP, 10, 026 (2010) [arXiv:1008.0048 [hep-th]].
19. A. Padilla, V. Sivanesan, *Boundary Terms and Junction Conditions for Generalized Scalar-Tensor Theories*, JHEP, 2012:122 (2012) [arXiv:1206.1258 [gr-qc]].

20. S. Nishi, T. Kobayashi, N. Tanahashi, M. Yamaguchi, *Cosmological matching conditions and galilean genesis in Horndeski's theory*, JCAP, 03, 008N (2014) [arXiv:1401.1045 [hep-th]].
21. B. Racskó, L. Á. Gergely, in preparation (2019).
22. B. Racskó , L. Á. Gergely, *Light-Like Shockwaves in Scalar-Tensor Theories*, Universe, 4(4), 44 (2018).

# Degrees of freedom and local Lorentz invariance in $f(T)$ gravity

María José Guzmán[1,*] and Rafael Ferraro[2,3]

[1] *Instituto de Física de La Plata (IFLP, CONICET-UNLP), C. C. 67, 1900 La Plata, Argentina*
[2] *Instituto de Astronomía y Física del Espacio (IAFE, CONICET-UBA), Casilla de Correo 67, Sucursal 28, 1428 Buenos Aires, Argentina*
[3] *Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires, Ciudad Universitaria, Pabellón I, 1428 Buenos Aires, Argentina*
[*] *E-mail: mjguzman@fisica.unlp.edu.ar*

$f(T)$ gravity is a generalization of the teleparallel equivalent of general relativity (TEGR), where $T$ is the torsion scalar made up of the Weitzenböck connection. This connection describes a spacetime with zero curvature but with nonvanishing torsion, which fully encodes the gravitational phenomena. We will present recent results in $f(T)$ gravity related with the issue of the degrees of freedom of the theory. In particular, we discuss the recent finding that $f(T)$ gravity has one extra degree of freedom compared with TEGR, which was concluded through a detailed Hamiltonian analysis of the constraint structure of the theory. The physical interpretation of this result at the level of the trace of the equations of motion and its comparison with the $f(R)$ case is discussed.

*Keywords*: f(T) gravity; Teleparallel gravity; Constrained Hamiltonian systems.

## 1. Teleparallel gravity and the $f(T)$ paradigm

The teleparallel equivalent of general relativity is an alternative formulation of gravity that employs the tetrad field $\mathbf{e}_a = e_a^\mu \partial_\mu$ as a dynamical variable. Unlike general relativity, which is formulated in a Riemann spacetime equipped with the Levi-Civita connection, teleparallel theories of gravity are commonly defined in terms of the Weitzenböck connection. The Riemann tensor vanishes in this connection, therefore it describes a curvatureless spacetime with absolute parallelism (or *tele*-parallelism). However, the Weitzenböck connection $\Gamma^\rho{}_{\mu\nu} = e_a^\rho \partial_\mu E_\nu^a$ has nonvanishing torsion given by $T^\rho{}_{\mu\nu} = e_a^\rho(\partial_\mu E_\nu^a - \partial_\nu E_\mu^a)$, where $\mathbf{E}^a = E_\mu^a dx^\mu$ is the co-tetrad field. The action giving dynamical equations that are equivalent to Einstein equations is

$$S = \frac{1}{2\kappa}\int d^4x\; E\; T = \frac{1}{2\kappa}\int d^4x\; E S_\rho{}^{\mu\nu} T^\rho{}_{\mu\nu} \tag{1}$$

where $\kappa = 8\pi G$, $E = \det(E_\mu^a)$, and it is implicitly defined the torsion scalar $T$, and the so-called superpotential $S_\rho^{\mu\nu}$ is given by

$$S_\rho{}^{\mu\nu} = \frac{1}{4}(T_\rho{}^{\mu\nu} - T^{\mu\nu}{}_\rho + T^{\nu\mu}{}_\rho) + \frac{1}{2}T_\sigma{}^{\sigma\mu}\delta_\rho^\nu - \frac{1}{2}T_\sigma{}^{\sigma\nu}\delta_\rho^\mu. \tag{2}$$

The equations of motion for the action (1) are

$$4e\partial_\mu(E e_a^\lambda S_\lambda{}^{\mu\nu}) + 4e_a^\lambda T^\rho{}_{\mu\lambda} S_\rho{}^{\mu\nu} - e_a^\nu T = -2\kappa e_a^\sigma \mathcal{T}_\sigma{}^\nu, \tag{3}$$

where $\mathcal{T}_\sigma{}^\nu$ represents the energy-momentum tensor. The equivalence between GR and TEGR theories emerges, both at the levels of the equations of motion and

the Lagrangians, when the metric is expressed in terms of the co-tetrad as $g_{\mu\nu} = \eta_{ab}E^a_\mu E^b_\nu$, since it results

$$R = -T + 2\ e\ \partial_\mu(ET^\mu) \tag{4}$$

where $R$ is written in terms of the tetrad field and calculated with the Levi-Civita connection. Therefore, both Lagrangians only differ by a surface term that is integrated out when plugged into the action.

TEGR is a suitable framework for building deformations of GR, since its Lagrangian has only first order derivatives of the tetrad field, therefore any function of $T$ will give a gravitational theory with second order equations of motion. The first application of modified teleparallel gravities to high energy modifications of GR was through a Born-Infeld approach[1], where the main purpose was to obtain an early accelerated expansion of the Universe without resorting to an inflaton field. Later there were proposed low energy deformations to GR through the so-called $f(T)$ gravities, intended to be an alternative explanation of the late-time accelerated expansion of the Universe. The $f(T)$ gravity action is given by[2–4]

$$S = \frac{1}{2\kappa}\int d^4x\ E\ f(T), \tag{5}$$

and its equations of motion are

$$4e\partial_\mu(f'(T)Ee^\lambda_a S_\lambda{}^{\mu\nu}) + 4f'(T)e^\lambda_a T^\sigma_{\mu\lambda}S_\sigma{}^{\mu\nu} - e^\nu_a f(T) = -2\kappa e^\lambda_a \mathcal{T}_\lambda{}^\nu. \tag{6}$$

These dynamical equations have a very unusual peculiarity: they violate local Lorentz invariance. This happens in a very particular way: given a tetrad $\mathbf{E}^a$ that satisfies the equations (6), the local Lorentz transformed tetrad $\mathbf{E}^{a'} = \Lambda^{a'}_a\mathbf{E}^a$ does not necessarily satisfy them. This is commonly understood as the theory choosing a preferential frame that endows the spacetime with a determined parallelization. Other interpretations of the problem of Lorentz invariance and its relation with the additional degree(s) of freedom of the theory will be reviewed in the next section.

## 2. On local Lorentz invariance and Hamiltonian formalism

The issue about the number and nature of the additional degrees of freedom in modified teleparallel gravity theories is an open question, which has been faced through several approaches. One of the main strategies utilized is the study of the Hamiltonian formulation of $f(T)$ gravity. This has been studied in the past[5], with the outcome that the theory posesses $n-1$ extra d.o.f. when compared with TEGR. The authors suggested that the extra d.o.f. of $f(T)$ gravity would manifest in a kind of Higgs mechanism, through a massive vectorial field or a scalar field plus a massless vectorial field. So far, it has not been shown such physical equivalence between these hypotetical fields and the $f(T)$ Lagrangian. Moreover, no extra degrees of freedom appears at the level of cosmological perturbations[6–9]. This strongly suggests that the counting of degrees of freedom in[5] should be revised in the light of new work. Recently the issue about the Hamiltonian formalism of

$f(T)$ gravity has been revised by an independent Hamiltonian analysis [10], with the outcome that the theory possess only one extra degree of freedom compared with TEGR. This analysis was based in a Hamiltonian formulation of teleparallel gravity [11] built from a Lagrangian quadratic in first-order derivatives in the tetrad field and proportional to a *supermetric* [11] or *constitutive tensor* [12], a Lorentz invariant mathematical object that significantly simplifies the calculation of the Hamiltonian and the Poisson brackets. This formulation shows that the Hamiltonian structure of TEGR is very simple: the traditional Arnowitt-Deser-Misner constraint structure from general relativity has an additional subalgebra representing local Lorentz transformations that are gauge symmetries of the theory [11].

The Hamiltonian formulation of TEGR mentioned before has been used to develop the Hamiltonian formulation of $f(T)$ gravity in the Jordan frame representation of the theory [10]. This frame is obtained by defining the Legendre transform of the $f(T)$ action with the help of a scalar field, alike to the $f(R)$ gravity case. It is found a novel constraint structure where it appears a new constraint associated with the introduction of the scalar field, which does not commute with a subset of the constraints $\tilde{G}^{(1)}_{ab}$ associated with the Lorentz sector. In particular, the scalar field constraint $G^{(1)}_{\pi}$ has a nonvanishing Poisson bracket with only one linear combination of the $n(n-1)/2$ Lorentz constraints $\tilde{G}^{(1)}_{ab}$. Both constraints pair up to become second class, while the remaining constraints are first-class. This could be interpreted as a partial violation of Lorentz symmetry in a unique, undetermined combination of boosts and rotations. The counting of degrees of freedom gives that $f(T)$ gravity has $n(n-3)/2+1$ physical d.o.f., this means one additional d.o.f. when compared with TEGR [10].

Other meditations upon the degrees of freedom of $f(T)$ gravity can be found in the literature. For instance, the method of characteristics for hyperbolic partial differential equations has been applied for studying the qualitative behavior of the degrees of freedom [13]. The authors find evidence for an extra d.o.f., however they assume that the theory would have three extra d.o.f. [5]. The absence of the supposedly missing two extra d.o.f. leads to the authors to conclude that the theory would have superluminal modes and a bad posed Cauchy problem. It is mandatory to revisit these claims in the light of recent research. Other potentially helpful approaches in this respect must be explored, as the proposals of *covariant* approaches where the spin connection is taken different from zero [14–16]. Another helpful strategy could be the introduction of Lagrange multipliers that enforce the vanishment of the Riemann tensor, guaranteeing the null curvature condition [17]. Finally, considerations about remnant symmetries on the theory [18] and the search for solutions to $f(T)$ through a null tetrad approach [19,20] should be further addressed.

In what follows, we will explore one possible approach for understanding the nature of the extra d.o.f. in $f(T)$ gravity, through its comparison with the well known $f(R)$ gravity case. We will find convincing evidence for its existence in the trace of the equations of motion, however it is noteworthy to point out that the interpretation of the extra d.o.f. in both theories is radically different.

## 3. The nature of the extra d.o.f.

One of the most studied modifications of general relativity is $f(R)$ gravity. The action of this gravitational theory is given by

$$S_{f(R)} = \frac{1}{2\kappa}\int d^4x\ \sqrt{-g}f(R) + S_m(g_{\mu\nu}), \tag{7}$$

where $S_m$ is the action for matter minimally coupled to gravity. This action has fourth-order dynamical equations given by

$$f'(R)R_{\mu\nu} - \frac{1}{2}f(R)g_{\mu\nu} - [\nabla_\mu\nabla_\nu - g_{\mu\nu}\Box]f'(R) = \kappa T_{\mu\nu}. \tag{8}$$

By taking the trace of the equations of motion it is obtained a relation between the Ricci scalar $R$ and the trace $\mathcal{T}$ of the energy-momentum tensor that is algebraical, that is

$$f'(R)R - 2f(R) + 3\Box f'(R) = \kappa\mathcal{T}. \tag{9}$$

This equation shows the propagation of a new degree of freedom related with $f'(R)$ [22,23]. Moreover, it is possible to write (8) and (9) as second-order equations for the metric and the scalar object $f'(R)$ by changing the notation to

$$\phi \equiv f'(R), \qquad V(\phi) \equiv R\phi - f(R), \tag{10}$$

to rewrite Eq.(8) as

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \frac{\kappa}{\phi}\mathcal{T}_{\mu\nu} - \frac{g_{\mu\nu}}{2\phi}V(\phi) + \frac{1}{\phi}[\nabla_\mu\nabla_\nu\phi - g_{\mu\nu}\Box\phi]. \tag{11}$$

We recognize in Eq.(10) the Legendre transform of $f(R)$, hence it is also obtained that $R = V'(\phi)$, which together with (11) are the dynamical equations associated with the action

$$S_{JF}[g_{\mu\nu},\phi] = -\frac{1}{2\kappa}\int d^4x\sqrt{-g}[\phi R - V(\phi)] + S_{\text{matter}}. \tag{12}$$

For completing the Legendre transformations it is needed that $\phi$ be a function of the scalar curvature $R$ and moreover, that the Lagrangian in (12) is equivalent to the Legendre transform of the function $V(\phi)$, therefore it can be rewritten as a function of $R$ by defining $f(R) = \phi R - V(\phi)$. The trace of Eq. (11) can be regarded as a wave equation for a self-interacting scalar field $\phi$, since it satisfies

$$3\Box\phi - \phi^3[\phi^{-2}V(\phi)]' = \kappa\mathcal{T}. \tag{13}$$

In this way, $\phi$ looks as a scalar field minimally coupled to the metric $g_{\mu\nu}$, but also coupled to the matter through the trace of the energy-momentum tensor. Meanwhile, Eq.(11) are Einstein equations, on which the source are the matter and the scalar field. The strategy used through the redefinition (10) allows to rephrase $f(R)$

gravity as a scalar-tensor theory with second-order dynamical equations. The equations (11) describe the dynamics of the metric field, therefore they encompass two d.o.f. in four dimensions. Additionally, there is one extra d.o.f. described by (13).

A reasonable approach would be to perform the same procedure to the equations of motion of $f(T)$ gravity, that is to say, to develop an analogous calculation of the trace of the equations of motion. The analogy is not complete, due to the fact that, while the equations of motion of $f(R)$ gravity are second-order on $f'(R)$, the dynamical equations for $f(T)$ gravity are first-order in the analog variable $f'(T)$. We can clearly see this feature if we rewrite the equations of motion (6) through the Legendre transform

$$\phi = f'(T), \qquad V(\phi) = T\phi - f(T), \tag{14}$$

to become

$$4\phi^{-1}e\partial_\mu(\phi E e_a^\lambda S_\lambda{}^{\mu\nu}) + 4e_a^\lambda T^\sigma{}_{\mu\lambda}S_\sigma{}^{\mu\nu} - e_a^\nu T = -\frac{2\kappa}{\phi}e_a^\lambda \mathcal{T}_\lambda{}^\nu + e_a^\nu \frac{V(\phi)}{\phi}. \tag{15}$$

These dynamical equations keep the structure of the TEGR equations, except for the renormalization of the volume $E$, the gravitational constant $\kappa$, and an additional term $\phi^{-1}V(\phi)$ that behaves as a local cosmological constant. Then the trace of (15) easily writes as

$$2T^\mu\partial_\mu\phi + 2\phi e\partial_\mu(ET^\mu) - T\phi + 2V(\phi) = -\kappa\mathcal{T}, \tag{16}$$

or alternatively as

$$2T^\mu\partial_\mu\phi + 2V(\phi) + \phi R = -\kappa\mathcal{T}, \tag{17}$$

when substituting $2e\partial_\mu(ET^\mu)$ by $R + T$, as stated in Eq.(4). We return to the TEGR case when $\phi = 1$ and $V(\phi) = 0$; in this case the equation (17) implies $R = -\kappa\mathcal{T}$, the classical result obtained in GR. Therefore, it is fair to say that (17) encodes more information than merely GR, and could be interpreted as the description of the propagation of an extra scalar degree of freedom that was not present in TEGR. It has been suggested that the existence of an extra d.o.f. in $f(T)$ is connected with the loss of a gauge symmetry. This can be seen in terms of the symmetries of $T$, which do not remain invariant under general local Lorentz transformations. Henceforth, $f(T)$ will inherit a remnant gauge symmetry, which has an on-shell character[21]. The implications of this remnant symmetry and its relevance in cosmological solutions deserves further research.

## 4. Conclusions and future work

In this work we have introduced teleparallel gravity and its simplest modification, the so-called $f(T)$ gravity, and briefly reviewed the important issue about the degrees of freedom in this theory. We have discussed several approaches on the counting of d.o.f., in particular some results obtained through the Dirac-Bergmann algorithm for constrained Hamiltonian systems. Recent work suggests that $f(T)$ gravity

would have only one extra d.o.f., in oposition to previous research. These claims are checked through the comparison with the $f(R)$ gravity case. In both theories the calculation of the trace of the equations of motion reveals an additional scalar d.o.f.. Nonetheless, its interpretation in the $f(T)$ case is qualitatively different, as it satisfies a first-order differential equation. This equation shows evidence that the additional scalar field could be related to the proper parallelization of spacetime. It is strongly encouraged further research about the characterization of the d.o.f. of more general teleparallel gravities and its implications in cosmology and compact objects.

## References


1. R. Ferraro and F. Fiorini, Phys. Rev. D **75**, 084031 (2007).
2. R. Ferraro and F. Fiorini, Phys. Rev. D **78**, 124019 (2008).
3. G. Bengochea and R. Ferraro, Phys. Rev. D **79**, 124019 (2009).
4. R. Ferraro, AIP Conf. Proc. **1471**, 103 (2012)
5. M. Li, R. X. Miao and Y. G. Miao, J. High Energ. Phys. **07**, 108 (2011).
6. K. Izumi and Y. C. Ong, J. Cosmol. Astropart. Phys. **06** (2013) 029
7. B. Li, T. P. Sotiriou, and J. D. Barrow, Phys. Rev. D **83**, 104017 (2011)
8. S.-H. Chen, J. B. Dent, S. Dutta, and E. N. Saridakis, Phys. Rev. D **83**, 023508 (2011)
9. A. Golovnev and T. Koivisto, J. Cosmol. Astropart. Phys **1811**, no. 11, 012 (2018)
10. R. Ferraro and M. J. Guzmán, Phys. Rev. D **97**, 104028 (2018).
11. R. Ferraro and M. J. Guzmán, Phys. Rev. D **94**, 104045 (2016).
12. Y. Itin, F. W. Hehl and Y. N. Obukhov, Phys. Rev. D **95**, no. 8, 084020 (2017)
13. Y. C. Ong, K. Izumi, J. M. Nester and P. Chen, Phys. Rev. D **88**, 024019 (2013)
14. M. Krššák and E. N. Saridakis, Class. Quant. Grav. **33**, no. 11, 115009 (2016)
15. A. Golovnev, T. Koivisto, M. Sandstad, Class. Quant. Grav. 34 (2017) no.14, 145013.
16. M. Hohmann, L. Järv and U. Ualikhanova, Phys. Rev. D **97**, no. 10, 104011 (2018)
17. Y. C. Ong and J. M. Nester, Eur. Phys. J. C **78**, no. 7, 568 (2018)
18. R. Ferraro and F. Fiorini, Phys. Rev. D **91**, no. 6, 064019 (2015)
19. C. Bejarano, R. Ferraro and M. J. Guzmán, Eur. Phys. J. C **75**, 77 (2015).
20. C. Bejarano, R. Ferraro and M. J. Guzmán, Eur. Phys. J. C **77**, no. 12, 825 (2017)
21. R. Ferraro and M. J. Guzmán, Phys. Rev. D **98** , 124037 (2018)
22. T. P. Sotiriou and V. Faraoni, Rev. Mod. Phys. **82**, 451 (2010)
23. G. J. Olmo, Phys. Rev. D **75**, 023511 (2007)

# Born-Infeld gravity and cosmological singulartities

Sayan Kar

*Department of Physics and Centre for Theoretical Studies,*
*Indian Institute of Technology Kharagpur ,721302, India*
** E-mail: sayan@phy.iitkgp.ac.in*

Soumya Jana

*Theoretical Physics Division, Physical Research Laboratory,*
*Ahmedabad 380009, India*
*E-mail: sjana@prl.res.in*

In 2010, Banados and Ferreira (BF) constructed a variant of Born-Infeld (BI) gravity with a simple matter coupling and demonstrated how the standard background FRW cosmology could become free of curvature singularities (big-bang). Further investigations revealed many interesting consequences of this BF version of BI gravity. For a toy 3D version, we show a simple analytical solution exhibiting the removal of the big-bang singularity. Thereafter, we look at BI gravity coupled to scalar BI matter, where we are able to find non-singular (loitering and bounce types) background solutions with late as well as early time acceleration. Finally, we elevate the BI gravity parameter to a space-time dependent field (in a novel Brans-Dicke like way) and demonstrate how cosmologies without singularities and with late as well as early-time acceleration can indeed arise quite naturally.

## 1. The singularity problem

The fact that curvature singularities inevitably appear in many solutions of Einstein's field equations of General Relativity (GR) is well-known. Very general theorems on singularities (largely defined using geodesic incompleteness) with minimal assumptions (causal structure and energy conditions) were proven long ago by Penrose, Hawking and others. In electrodynamics too, a singular behaviour of the electric field at the location of the charge is known. It is not any huge embarrasment in electrodynamics. On the other hand a spacetime singularity (be it in the sense of diverging curvature scalars or geodesic incompleteness or both) is indeed problematic because one cannot extend the spacetime beyond that location. Apart from the fact that via Einstein equations, spacetime singularities corrrespond to infinities in energy density of matter, it is the inextendability of spacetime beyond such special points where some pathologies arise and/or curvature invariants diverge, which makes the theory unacceptable at those points.

It is believed that quantum gravity would cure classical GR of the singularity problem. And, indeed there are indications in various such attempts – i.e. in string theory and loop quantum gravity. On the other hand, it may be asked if any *other* classical theory of gravity can admit non-singular solutions. This is a valid question and it is along these lines we direct our presentation below.

## 2. Born-Infeld structures and Born-Infeld gravity

The appearance of a singularity at the location of the electric charge, is well known. A way out was sought in the work of Born and Infeld (BI)[1], which had links with earlier work due to Eddington[2]. BI proposed a nonlinear electrodynamics with an action given as:

$$S_{BI} = \int \frac{1}{\kappa^2}\left\{1 - \sqrt{-Det[\eta_{ij} + \kappa F_{ij}]}\right\} d^4x. \tag{1}$$

The field equations when solved for a point source lead to a resolution of the self-energy singularity in the electric field. $\kappa$ plays the role of a new parameter which yields an upper bound on the electric field of a point charge.

In 1998, Deser and Gibbons[3] first proposed a gravity action (in a metric formulation) building on BI electrodynamics and Eddington's old ideas (notably the *square root* and the *determinant* in the action). Later, in 2004, Vollick[4] worked out a Palatini formulation of the Deser-Gibbons theory, incorporating matter in a rather artificial way. More recently, Banados and Ferreira (BF)[5] have come up with a theory where the matter coupling is quite simple. The action proposed by BF is given as:

$$S_{BF}(g, \Gamma, \Psi) = \frac{1}{\kappa}\int d^4x \left[\sqrt{-|g_{ij} + \kappa R_{ij}|} - \lambda\sqrt{-g}\right] + S_M(g, \Psi), \tag{2}$$

where $|g_{ij}+\kappa R_{ij}| = Det[g_{ij}+\kappa R_{ij}]$ and $g = Det[g_{ij}]$. Note that the matter coupling here is standard, using the $g_{ij}$ and matter fields $\Psi$. For $\kappa R$ small we get back the Einstein-Hilbert action ($\Lambda = \frac{\lambda-1}{\kappa}$). When $\kappa R$ is large we recover the Eddington action[2]. Without matter, BF theory is equivalent to GR.

Variation w.r.t. $g_{ij}$ gives us the field equation:

$$\sqrt{-q}q^{ij} = \lambda\sqrt{-g}g^{ij} - \kappa\sqrt{-g}T^{ij}. \tag{3}$$

On varying w.r.t $\Gamma^i_{jk}$ we find that the $q_{ij}$ as obtained from solving the equation,

$$q_{ij} = g_{ij} + \kappa R_{ij}(q) \tag{4}$$

is a metric with a Christofel connection. The usual conservation law (covariant derivative using $g_{ij}$), $\nabla_j T^{ij} = 0$ holds. One may also consider the BF action as a bimetric gravity action where the two metrics are $q_{ij}$ and $g_{ij}$. $g_{ij}$ is the physical one while $q_{ij}$ is auxiliary.

The main problem in finding analytical solutions is in the inversion of the equation $q_{ij} = g_{ij}+R_{ij}(q)$. Since these equations cannot be *solved*, one needs to postulate $g_{ij}$ and $q_{ij}$ separately (with different unknown functions) and write down the field equations. Then only one may be able to find solutions.We shall now provide examples in cosmology. For more details on Born-Infeld inspired gravity theories and their consequences see a recent review[6].

### 2.1. *Cosmology*

We assume line elements of the form:

$$ds^2 = -dt^2 + a^2(t)\left(dx^2 + dy^2 + dz^2\right), \tag{5}$$

$$ds_q^2 = -U(t)dt^2 + a^2(t)V(t)\left(dx^2 + dy^2 + dz^2\right), \tag{6}$$

where $U(t)$, $V(t)$, $a(t)$ are three unknowns. The energy-momentum tensor $T^{ij}$ is that of a perfect fluid, i.e. $T^{00} = \rho$, $T^{\alpha\alpha} = p$ ($\alpha = 1, 2, 3$).

With these assumptions one writes down all the field equations stated earlier, for this case. Assuming $p = \frac{\rho}{3}$ one gets the Friedmann equation as:

$$3H^2(\bar{\rho}) = \frac{1}{\kappa}\left[\bar{\rho} - 1 + \frac{1}{3\sqrt{3}}\sqrt{(1+\bar{\rho})(3-\bar{\rho})^3}\right] \times \frac{(1+\bar{\rho})(3-\bar{\rho})^2}{(3+\bar{\rho}^2)^2} \tag{7}$$

where $\bar{\rho} = \kappa\rho$. For small $\bar{\rho}$ one gets get $H^2 \sim \frac{\rho}{3}$, which is the Friedmann equation in GR. Note that $H^2$ has zeros at $\bar{\rho} = 3$ and $\bar{\rho} = 0$ for $\kappa > 0$. For $\kappa < 0$ zeros are at $\bar{\rho} = 0$ and $\bar{\rho} = -1$. Note the difference with the usual FRW model where $3H^2 = \rho$ and the only zero is at $\rho = 0$.

One can find numerical solutions of the scale factor for $\kappa < 0$, $\kappa > 0$ (Fig. 1):

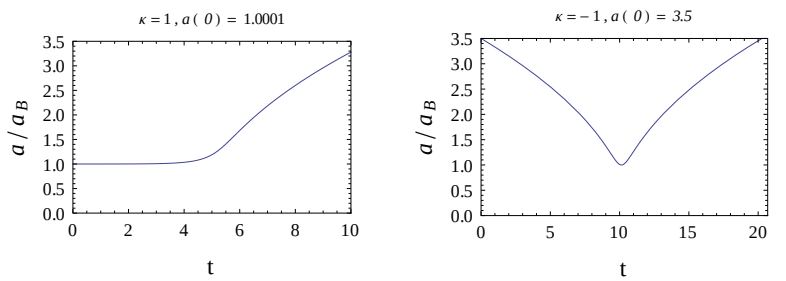



Fig. 1. Plot of the scale factor $a(t)$ for $\kappa = \pm 1$. $a_B$ is the nonzero minimum value of $a(t)$ corresponding to maximum density $\bar{\rho}_B$ and $t$ is an arbitrary time scale. $8\pi G = 1$, $c = 1$.

Note the nonzero minimum value of $a(t)$ which corresponds to the upper limit on $\rho$. There is no big-bang singularity or any singularity in the future. The absence of the singularity is largely controlled by the BI parameter $\kappa$ and the structure of BI gravity. In a toy $2+1$ dimensional model it is possible to find simple analytical solutions which exhibit non-singular features[7]. For example, in a $2+1$ version of the BF theory, the scale factor in the $2+1$ cosmology (with $\kappa < 0$ and $p = \frac{\rho}{2}$), has the form: $a(t) = \sqrt{\frac{t^2}{|\kappa|} + 1}$, which never vanishes and yields a toy, nonsingular universe.

## 3. BI gravity + BI matter: cosmologies

A useful question to ask is: what happens if there are BI structures in both the gravity and matter sectors in BF theory? We assume a simple scalar BI structure

which has been studied extensively as *tachyon matter* about a decade or so ago. The action we work with is $S = S_{BI}(g, \Gamma, \Psi) + S_M$, where

$$S_{BI} = \frac{c^3}{8\pi G\kappa} \int d^4x \left[ \sqrt{-|g_{\mu\nu} + \kappa R_{\mu\nu}(\Gamma)|} - \lambda\sqrt{-g} \right],$$

where $\lambda = \kappa\Lambda + 1$, and

$$S_M = -\frac{1}{c} \int \sqrt{-g}\alpha_T^2 \mathcal{V}(\phi) \sqrt{1 + \alpha_T^{-2} g^{\mu\nu} \partial_\mu\phi\partial_\nu\phi} d^4x,$$

where $\mathcal{V}(\phi)$ is the potential for the scalar field and $\alpha_T$ is the constant parameter. Analytical solutions are possible. More importantly, we ask: what are the features of the cosmological solutions? Here, in Fig. 2, the scale factor $a$ (for a constant negative pressure solution with the equivalent pressure and energy densities of the scalar field, $p_\phi = -\alpha_T^2 C_2 c^2$ and $\rho_\phi = C_2(a^{-3} + \alpha_T^2)$) is plotted as a function of the cosmological time ($\tau$). We choose $8\pi G = 1$, $c = 1$ and $\kappa = 0.5, \alpha_T^2 = 5.0, C_2 = 0.001$. Initial value of $a(\tau)$ is choosen as $a = 0.06$ at $\tau = 0.06$. Initial loitering is followed by deceleration and late-time acceleration. The zoomed version in Fig. 3 shows clearly the loitering phase and the transition into deceleration. The loitering phase also includes an acceleration, where the scale factor has an exponential growth ($a \sim a_0 \exp(2\sqrt{2}c\tau/\sqrt{3\kappa})$). For negative $\kappa$ we obtain a bounce solution (not shown).

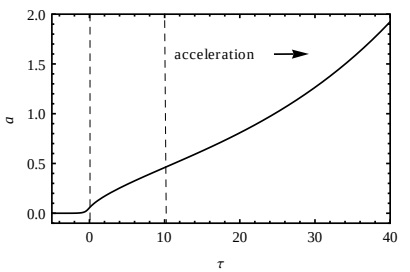



Fig. 2. Plot of scale factor $a(\tau)$.

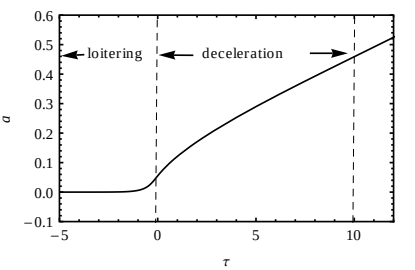



Fig. 3. Zoomed $a(\tau)$ near big-bang

We have fitted the late-time acceleration found in our solution, using Type Ia supernovae (Union2.1 compilation) data. The best fit values of the cosmological parameters for the positive $\kappa$ solution are given as:

$$q_0 = -0.605^{+0.026}_{-0.054} \; ; \; \Omega_{DM0} = 0.255^{+0.051}_{-0.021} \; ; \; \Omega_{DE0} = 0.745^{+0.016}_{-0.051}$$

Note that there is no separate dark matter or dark energy. We use a result due to Roy Chowdhury and Padmanabhan[9] to split the BI matter into a dark matter part and a dark energy part. This is characteristic of BI matter like the scalar BI used here. In summary, the BI matter coupled to BI gravity does produce a nonsingular universe with a late-time acceleration. More details are available in[8].

## 4. BI gravity with a Brans-Dicke scalar: cosmologies

Let us finally turn to another important issue which is related to the BI gravity parameter $\kappa$. From various proposals for tests of the BF theory one obtains various bounds on $\kappa$. This is rather problematic and we wish to avoid this in a Brans-Dicke like way. Recall that the original BD idea was to replace $G$ by a smooth field $\phi(x, y, z, t)$. Here we replace $\kappa$ with a smooth field $\kappa(x, y, z, t)$. The action is taken as:

$$S_{BI\kappa} = \int \left[\frac{1}{\kappa}\left(\sqrt{-|g_{\alpha\beta} + \kappa R_{\alpha\beta}(\Gamma)|} - \sqrt{-g}\right) - \sqrt{-g}\tilde{\omega}(\kappa) g^{\mu\nu}\partial_\mu\kappa\partial_\nu\kappa\right] d^4x + S_M\,(g, \Psi)\,. \quad (8)$$

Here, $\tilde{\omega}(\kappa)$ is a coupling function (like in scalar-tensor theories).

The three field equations ($\Gamma$, $g$ and $\kappa$ variation) are:

$$q_{\alpha\beta} = g_{\alpha\beta} + \kappa R_{\alpha\beta}(q), \quad (9)$$

$$\sqrt{-q}q^{\alpha\beta} - \sqrt{-g}g^{\alpha\beta} = -\kappa\sqrt{-g}\,T^{\alpha\beta}_{eff}, \quad (10)$$

where

$$T^{\alpha\beta}_{eff} = T^{\alpha\beta} - \tilde{\omega} g^{\alpha\beta} g^{\mu\nu}\partial_\mu\kappa\partial_\nu\kappa + 2\tilde{\omega} g^{\mu\alpha} g^{\nu\beta}\partial_\mu\kappa\partial_\nu\kappa, \quad (11)$$

$T^{\alpha\beta}$ is the usual stress-energy tensor, and finally,

$$2\kappa\tilde{\omega}(\kappa)\nabla_\mu\nabla^\mu\kappa + \kappa\tilde{\omega}'(\kappa)\nabla_\mu\kappa\nabla^\mu\kappa + \frac{1}{\kappa} + \frac{\sqrt{-q}}{\sqrt{-g}}\left(\frac{1}{2}q^{\alpha\beta}R_{\alpha\beta}(q) - \frac{1}{\kappa}\right) = 0. \quad (12)$$

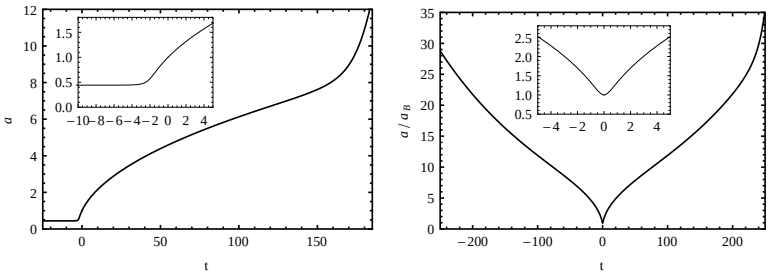
We now turn to FRW cosmologies. Assuming $\kappa(t) = \kappa_0 + \epsilon\, e^{\mu\, t}$ with $\kappa_0$, $\epsilon$, and $\mu$ as parameters, we consider several cases: $\rho, p = 0$ (vacuum); $p = 0$ (pressureless dust); $p = \frac{\rho}{3}$ (radiation). We investigate the nature of the scale factors for various ranges of $\kappa_0$, $\epsilon$ and $\mu$.

Let us illustrate this for the case of radiation: (a) $p = \frac{\rho}{3}$: $\kappa_0 > 0$, $\mu > 0$, $\epsilon < 0$ (Fig. 4(a)), (b) $p = \frac{\rho}{3}$ : $\kappa_0 < 0$, $\mu > 0$, $\epsilon < 0$ (Fig. 4(b)).

Generic features for $\kappa > 0$ (non-singular beginning), and $\kappa < 0$ (bounce) are retained in Figs. 4(a) and 4(b). Further details on this work is available in[10].

## 5. Concluding remarks

A recently proposed novel theory of gravity based on a Born–Infeld structure which matches with GR in vacuum is discussed with emphasis on its ability to resolve the cosmological singularity problem. We have first shown how the original BF version yields a nonsingular FRW universe. Incorporating BI structures in both the gravity and matter sectors we have analytically obtained non-singular FRW cosmologies with initial loitering and late time acceleration. We have also fitted the late-time cosmology with available Supernova data and obtained best-fit values for the cosmological parameters. Further, solutions with an initial bounce and late acceleration are found.




(a) (b)

Fig. 4. (a) plot of scale factor $a(t)$ for $\kappa_0 > 0$, $\mu > 0$, $\epsilon < 0$. The parameters used are $\kappa_0 = 1$, $\mu = 0.1$, and $\rho_0 = 0.1$ (in $\rho = \rho_0/a^4$). We choose $a(0) = 1$, $\kappa(0) = 0.999$ for the numerical solution. (b) plot of scale factor $a(t)$ for $\kappa_0 < 0$, $\mu > 0$, $\epsilon < 0$. The parameters used are $\kappa_0 = -1$, $\mu = 0.1$, and $\rho_0 = 0.01$. We choose $a(0) = (-\kappa_0\rho_0)^{1/4} = a_B$, $\kappa(0) = -1.00001$ for the numerical solution.

In the final part of this article, we have proposed a Brans-Dicke like modification of the BI gravity theory where the BI parameter is elevated to a scalar field. Analysing cosmological solutions we have observed that a nonsingular FRW universe is possible here too for $\kappa > 0$, while for $\kappa < 0$, we have an initial bounce.

Are these solutions useful as background cosmologies? Can we study cosmological perturbations, structure formation, early universe phenomenology, inflation, dark matter, CMB .... in the models we have discussed here? Unless that is done and more links with present-day cosmological observations are established, these models and solutions will merely remain elegant mathematical curiosities.

**Acknowledgments**

One of the authors (SK) thanks the organisers for giving him the opportunity to present this work at MG15. He also thanks them for their warm hospitality.

**References**

1. M. Born and L. Infeld, Proc. R. Soc. A **144**, 425 (1934)
2. A. Eddington, *The mathematical theory of relativity*, Cambridge University Press, Cambridge, England, 1924.
3. S. Deser and G. W. Gibbons, Class. Qtm. Grav. **15**, L35 (1998).
4. D. N. Vollick, Phys. Rev. **D 69**, 064030 (2004).
5. M. Banados and P. Ferreira, Phys. Rev. Lett. **105**, 011101 (2010).
6. J. B. Jimenez, L. Heisenberg, G. J. Olmo and D. Garcia Rubiera,Phys. Reps. **727**, 1-129 (2018).
7. S. Jana and S. Kar, Phys. Rev. **D 88**, 024013 (2013)
8. S. Jana and S. Kar, Phys. Rev. **D 94** , 064016 (2016)
9. T. Roy Chowdury and T. Padmanabhan, Phys. Rev. **D 66**, 081301 (2002).
10. S. Jana and S. Kar, Phys. Rev. **D 96** , 024050 (2017).

# Novel couplings between nonmetricity and matter

Tiberiu Harko

*Department of Physics, Babes-Bolyai University,*
*Kogalniceanu Street, Cluj-Napoca 400084, Romania, and*
*Department of Mathematics, University College London,*
*Gower Street, London WC1E 6BT, United Kingdom*
*E-mail: t.harko@ucl.ac.uk*

Tomi S. Koivisto

*Nordita, KTH Royal Institute of Technology and Stockholm University,*
*Roslagstullsbacken 23, 10691 Stockholm, Sweden*
*E-mail: timoko@kth.se*

Gonzalo J. Olmo

*Departamento de Física Teórica and IFIC, Centro Mixto Universidad de Valencia - CSIC.*
*Universidad de Valencia, Burjassot-46100, Valencia, Spain*
*E-mail: gonzalo.olmo@uv.es*

Francisco S.N. Lobo* and Diego Rubiera-Garcia**

*Instituto de Astrofísica e Ciências do Espaço,*
*Faculdade de Ciências da Universidade de Lisboa,*
*Edifício C8, Campo Grande, P-1749-016 Lisbon, Portugal*
** E-mail: fslobo@fc.ul.pt*
*** E-mail: drgarcia@fc.ul.pt*

We present a novel theory of gravity, namely, an extension of symmetric teleparallel gravity. This is done by introducing a new class of theories where the nonmetricity $Q$ is coupled nonminimally to the matter Lagrangian. This nonminimal coupling entails the nonconservation of the energy-momentum tensor, and consequently the appearance of an extra force. We also present several cosmological applications.

*Keywords*: Modified gravity; nonmetricity; cosmology.

## 1. Introduction

The fact that general relativity (GR) is facing so many challenges, namely, (i) the difficulty in explaining particular observations, (ii) the incompatibility with other well established theories, (iii) and the lack of uniqueness, may be indicative of a need for new gravitational physics. Thus, a promising approach is to assume that at large scales GR breaks down, and a more general action describes the gravitational field[1–5]. The physical motivations for these modifications of gravity include the following: (i) the possibility of a more realistic representation of the gravitational fields near curvature singularities; (ii) and to create some first order approximation for the quantum theory of gravitational fields.

An interesting model that is related to this work is the nonminimal curvature-matter coupling[6–9], in which the Lagrangian density is of the form $L \sim f_1(R) + [1 + \lambda f_2(R)]\, L_m$, where $f_i(R)$ (with $i = 1, 2$) are arbitrary functions of the Ricci

scalar $R$ and $L_m$ is the matter Lagrangian density. Applications have been explored such as in an effective dark energy or dark matter[9–12]. An interesting feature of these theories is the non-conservation of the energy-momentum tensor, so that the coupling between the matter and the higher derivative curvature terms describes an exchange of energy and momentum between both.

Indeed, it is possible to tackle gravitation through several approaches, namely, with the metric formalism[1–3], which consists on varying the action with respect to the metric and setting the Levi-Civita connection, or through the metric-affine formalism[13], where the metric and the affine connection are regarded as independent variables, or even through a hybrid approach[9,14,15]. Note that the metric $g_{\mu\nu}$ may be thought of as a generalization of the gravitational potential and is used to define notions such as distances, volumes and angles. On the other hand, the affine connection $\Gamma^{\mu}{}_{\alpha\beta}$ defines parallel transport and covariant derivatives.

In fact, a basic result in differential geometry is that the general affine connection may be decomposed into the following 3 independent components:

$$\Gamma^{\lambda}{}_{\mu\nu} = \left\{{}^{\lambda}{}_{\mu\nu}\right\} + K^{\lambda}{}_{\mu\nu} + L^{\lambda}{}_{\mu\nu}\,, \tag{1}$$

where the first term is the Levi-Civita connection of the metric $g_{\mu\nu}$, defined as:

$$\left\{{}^{\lambda}{}_{\mu\nu}\right\} \equiv \frac{1}{2} g^{\lambda\beta}\left(\partial_{\mu} g_{\beta\nu} + \partial_{\nu} g_{\beta\mu} - \partial_{\beta} g_{\mu\nu}\right)\,. \tag{2}$$

The second term $K^{\lambda}{}_{\mu\nu}$ is the contortion, given by

$$K^{\lambda}{}_{\mu\nu} \equiv \frac{1}{2} T^{\lambda}{}_{\mu\nu} + T_{(\mu}{}^{\lambda}{}_{\nu)}\,, \tag{3}$$

with the torsion tensor defined as $T^{\lambda}{}_{\mu\nu} \equiv 2\Gamma^{\lambda}{}_{[\mu\nu]}$. The third term is the disformation

$$L^{\lambda}{}_{\mu\nu} \equiv \frac{1}{2} g^{\lambda\beta}\left(-Q_{\mu\beta\nu} - Q_{\nu\beta\mu} + Q_{\beta\mu\nu}\right)\,, \tag{4}$$

which is defined in terms of the nonmetricity tensor, given by $Q_{\rho\mu\nu} \equiv \nabla_{\rho} g_{\mu\nu}$.

This implies that by making assumptions on the affine connection, one is essentially specifying a metric-affine geometry[16]. For instance, the standard formulation of GR assumes a Levi-Civita connection, which implies vanishing torsion and nonmetricity, while its teleparallel equivalent (TEGR), uses the Weitzenböck connection, implying zero curvature and nonmetricity[17]. More recently, the symmetric teleparallel theories of gravity were analysed in[18–21], which possess remarkable features. Here, we present an extension of the symmetric teleparallel gravity, by introducing a new class of theories where a general function of the nonmetricity $Q$ is coupled nonminimally to the matter Lagrangian, in the framework of the metric-affine formalism[22].

This work is outlined in the following manner: In Section 2, we present the formalism of an extended symmetric teleparallel equivalent of general relativity. In Section 3, we consider some cosmological applications and conclude in Section 4.

## 2. Nonmetricity-matter coupling

### 2.1. *Action and gravitational field equations*

Consider the action defined by two functions of the nonmetricity $Q$, given by

$$S = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{1}{2}f_1(Q) + f_2(Q)L_M\right], \tag{5}$$

where $L_M$ is a Lagrangian function for the matter fields.

We define the nonmetricity tensor and its two traces as follows:

$$Q_{\alpha\mu\nu} = \nabla_\alpha g_{\mu\nu}\,, \qquad Q_\alpha = Q_\alpha{}^\mu{}_\mu\,, \qquad \tilde{Q}_\alpha = Q^\mu{}_{\alpha\mu}\,. \tag{6}$$

It is also useful to introduce the superpotential

$$4P^\alpha{}_{\mu\nu} = -Q^\alpha{}_{\mu\nu} + 2Q_{(\mu\ \ \nu)}^{\ \ \alpha} - Q^\alpha g_{\mu\nu} - \tilde{Q}^\alpha g_{\mu\nu} - \delta^\alpha_{(\mu}Q_{\nu)}\,, \tag{7}$$

where one can readily check that $Q = -Q_{\alpha\mu\nu}P^{\alpha\mu\nu}$, which is a useful relation.

For notational simplicity, we introduce the following definitions

$$f = f_1(Q) + 2f_2(Q)L_M\,, \qquad F = f_1'(Q) + 2f_2'(Q)L_M\,, \tag{8}$$

and specify the following variations

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta(\sqrt{-g}L_M)}{\delta g^{\mu\nu}}\,, \qquad H_\lambda{}^{\mu\nu} = -\frac{1}{2}\frac{\delta(\sqrt{-g}L_M)}{\delta\Gamma^\lambda{}_{\mu\nu}}\,, \tag{9}$$

as the energy-momentum tensor and the hyper-momentum tensor density, respectively.

Varying the action (5) with respect to the metric, one obtains the gravitational field equations given by

$$\frac{2}{\sqrt{-g}}\nabla_\alpha\left(\sqrt{-g}FP^\alpha{}_{\mu\nu}\right) + \frac{1}{2}g_{\mu\nu}f_1 + F\left(P_{\mu\alpha\beta}Q_\nu{}^{\alpha\beta} - 2Q_{\alpha\beta\mu}P^{\alpha\beta}{}_\nu\right) = -f_2T_{\mu\nu}\,. \tag{10}$$

and the variation with respect to the connection, yields the following relation

$$\nabla_\mu\nabla_\nu\left(\sqrt{-g}FP^{\mu\nu}{}_\alpha - f_2H_\alpha{}^{\mu\nu}\right) = 0\,. \tag{11}$$

### 2.2. *Divergence of the energy-momentum tensor*

The divergence of the energy-momentum tensor is given by:

$$\begin{aligned}\mathcal{D}_\mu T^\mu{}_\nu + \frac{2}{\sqrt{-g}}\nabla_\alpha\nabla_\beta H_\nu{}^{\alpha\beta} = &-\frac{2}{\sqrt{-g}f_2}\left[\left(\nabla_\alpha\nabla_\beta f_2\right)H_\nu{}^{\alpha\beta} + 2f_{2,(\alpha}\nabla_{\beta)}H_\nu{}^{\alpha\beta}\right]\\ &-\left(T^\mu{}_\nu - \delta^\mu_\nu L_M\right)\nabla_\mu\log f_2\,,\end{aligned} \tag{12}$$

where $\mathcal{D}_\alpha$ denoted the metric covariant derivative with respect to the symbols (2). This shows that due to the coupling between the nonmetricity $Q$ and the matter fields, the matter energy-momentum tensor is no longer conserved. The first term of the right-hand-side line is due to the nonminimal coupling of the hypermomentum, and the second term is related to the nonminimal coupling of the energy-momentum tensor.

Consider a perfect fluid, given by the energy-energy tensor, $T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + p g_{\mu\nu}$, so that the energy and momentum balance equations gives

$$\dot{\rho} + 3\mathcal{H}\,(\rho + p) = \mathcal{S}, \tag{13}$$

where $\dot{} = u_\mu \mathcal{D}^\mu$, and $\mathcal{H} = (1/3)\mathcal{D}^\mu u_\mu$. For simplicity, one can decompose the energy source term as $\mathcal{S} = \mathcal{S}_\mathcal{T} + \mathcal{S}_\mathcal{H}$, where $\mathcal{S}_\mathcal{T}$ is defined by

$$\mathcal{S}_\mathcal{T} = (\rho + L_M)\,\frac{\dot{f}_2}{f_2}\,,$$

which vanishes when we adopt the Lagrangian $L_M = -\rho$, and the hypersource is given as

$$\mathcal{S}_\mathcal{H} = -\frac{2}{\sqrt{-g}} u^\nu \Big[\nabla_\alpha \nabla_\beta H_\nu{}^{\alpha\beta} + \frac{1}{f_2}\,(\nabla_\alpha \nabla_\beta f_2)\, H_\nu{}^{\alpha\beta} + \frac{1}{f_2} f_{2,(\alpha} \nabla_{\beta)} H_\nu{}^{\alpha\beta}\Big]\,.$$

Note that the non-conservation of the energy-momentum tensor implies non-geodesic motion, where an extra-force $\mathcal{F}^\lambda$ arises due to the $Q$-matter coupling

$$\frac{d^2 x^\lambda}{ds^2} + \left\{{}^{\lambda}_{\mu\nu}\right\} u^\mu u^\nu = \mathcal{F}^\lambda. \tag{14}$$

The extra-force can be decomposed as

$$\mathcal{F}^\lambda = -\frac{h^{\alpha\lambda}\nabla_\alpha p}{\rho + p} + \mathcal{F}^\lambda_\mathcal{T} + \mathcal{F}^\lambda_\mathcal{H}\,, \tag{15}$$

where the first term on the right-hand-side is the usual general relativistic contribution of the pressure gradient. The extra force $\mathcal{F}^\lambda$ consists of the term $\mathcal{F}^\lambda_\mathcal{T}$, defined as

$$\mathcal{F}^\lambda_\mathcal{T} = (-p + L_M)\, h^\lambda_\nu \nabla^\nu \log f_2\,, \tag{16}$$

and the hyperforce $\mathcal{F}^\lambda_\mathcal{H}$ is given by

$$\mathcal{S}_\mathcal{H} = -\frac{2}{\sqrt{-g}} u^\nu \Big[\nabla_\alpha \nabla_\beta H_\nu{}^{\alpha\beta} \frac{1}{f_2}\,(\nabla_\alpha \nabla_\beta f_2)\, H_\nu{}^{\alpha\beta} + \frac{1}{f_2} f_{2,(\alpha} \nabla_{\beta)} H_\nu{}^{\alpha\beta}\Big]\,. \tag{17}$$

It is interesting that the extra force (16) vanishes identically for a perfect fluid if we adopt the Lagrangian prescription $L_M = p$.

## 3. Cosmological applications

In the framework of standard Friedman-Robertson-Walker (FRW) geometry, we have $Q = 6H^2$. Introducing the effective energy density $\rho_{\rm eff}$ and effective pressure $p_{\rm eff}$ of the cosmological fluid, defined as

$$\rho_{\rm eff} = -\frac{f_2}{2F}\left(\rho - \frac{f_1}{2f_2}\right), \qquad p_{\rm eff} = \frac{2\dot{F}}{F} H - \frac{f_2}{2F}\left(\rho + 2p + \frac{f_1}{2f_2}\right), \tag{18}$$

we can write the gravitational field equations in a form similar to the Friedmann equations of GR as

$$3H^2 = \rho_{\rm eff}, \qquad 2\dot{H} + 3H^2 = -p_{\rm eff}. \tag{19}$$

To describe cosmological evolution, and the possible transition to an accelerated phase, we also introduce the parameter $w$ of the dark energy equation of state, defined as

$$w=\frac{p_{\rm eff}}{\rho_{\rm eff}}=\frac{-4\dot{F}H+f_2\left(\rho+2p+\frac{f_1}{2f_2}\right)}{f_2\left(\rho-\frac{f_1}{2f_2}\right)}. \tag{20}$$

The deceleration parameter can be written as

$$q=\frac{1}{2}\left(1+3w\right)=2+\frac{3\left(4\dot{F}H-f_1-2f_2p\right)}{f_1-2f_2\rho}\,. \tag{21}$$

Thus, as a first step in this direction we have obtained the generalized Friedmann equations describing the cosmological evolution in flat FRW type geometry. The coupling between matter and the $Q$ field introduces two types of corrections. The first is the presence of a term of the form $f_2/2F$ multiplying the components of the energy-momentum tensor (energy density and pressure) in both Friedmann equations. Secondly, an additive term of the form $f_1/4F$ also appears in the generalized Friedmann equations. The basic equations describing the cosmological dynamics can then be reformulated in terms of an effective energy density and pressure, which both depend on the standard components of the energy-momentum tensor, and on the functions $f_i(Q)$, $i=1,2$, and on $F$. In the vacuum case $\rho=p=0$, the deceleration parameter takes the form $q=-1+12\dot{F}H/f_1$, showing that, depending on the mathematical forms of the coupling functions, a large number of cosmological evolutionary scenarios can be obtained. Generally, we have shown explicitly that for late times, the Universe attains an exponentially accelerating de Sitter phase[22].

## 4. Conclusions

In this work, we have explored an extension of the symmetric teleparallel gravity, by considering a new class of theories where the nonmetricity $Q$ is coupled nonminimally to the matter Lagrangian, in the framework of the metric-affine formalism. As in the standard curvature-matter couplings, this nonminimal $Q$-matter coupling entails the nonconservation of the energy-momentum tensor, and consequently the appearance of an extra force. Thus, in summary, we have established the theoretical consistency and motivations on these extensions of $f(Q)$ family of theories. Furthermore, we considered cosmological applications, in which the presented approach provides gravitational alternatives to dark energy. As future avenues of research, one should aim in characterizing the phenomenology predicted by these theories with a nonmetricity-matter coupling, in order to find constraints arising from observations.

## Acknowledgments

FSNL acknowledges funding by the Fundação para a Ciência e a Tecnologia (FCT, Portugal) through an Investigador FCT Research contract No. IF/00859/2012.

GJO is funded by the Ramon y Cajal contract RYC-2013-13019 (Spain). DRG is funded by the FCT postdoctoral fellowship No. SFRH/BPD/102958/2014. FSNL and DRG also acknowledge funding from the research grants UID/FIS/04434/2013, No. PEst-OE/FIS/UI2751/2014 and No. PTDC/FIS-OUT/29048/2017. This work is supported by the Spanish projects FIS2014-57387-C3-1-P, FIS2017-84440-C2-1-P (AEI/FEDER, EU), the project H2020-MSCA-RISE-2017 Grant FunFiCO-777740, the project SEJI/2017/042 (Generalitat Valenciana), the Consolider Program CPANPHY-1205388, and the Severo Ochoa grant SEV-2014-0398 (Spain).

## References

1. T. P. Sotiriou and V. Faraoni, *Rev. Mod. Phys.* **82**, 451 (2010).
2. A. De Felice and S. Tsujikawa, *Living Rev. Rel.* **13**, 3 (2010).
3. S. Nojiri and S. D. Odintsov, *Phys. Rept.* **505**, 59 (2011).
4. S. Capozziello and M. De Laurentis, *Phys. Rept.* **509**, 167 (2011).
5. F. S. N. Lobo, in *Dark Energy-Current Advances and Ideas*, 173-204 (2009), Research Signpost, ISBN 978-81-308-0341-8.
6. H. F. M. Goenner, *Foundations of Physics* **14**, 9 (1984).
7. T. Koivisto, *Class. Quant. Grav.* **23**, 4289 (2006).
8. O. Bertolami, C. G. Boehmer, T. Harko and F. S. N. Lobo, *Phys. Rev. D* **75**, 104016 (2007).
9. T. Harko and F. S. N. Lobo, Extensions of $f(R)$ Gravity: Curvature-Matter Couplings and Hybrid Metric-Palatini Theory, (Cambridge Monographs on Mathematical Physics). Cambridge: Cambridge University Press.
10. G. Allemandi, A. Borowiec, M. Francaviglia and S. D. Odintsov, *Phys. Rev. D* **72**, 063505 (2005).
11. S. Nojiri and S. D. Odintsov, *Phys. Lett. B* **599**, 137 (2004).
12. T. Harko and F. S. N. Lobo, *Galaxies* **2**, no. 3, 410 (2014)
13. G. J. Olmo, *Int. J. Mod. Phys. D* **20**, 413 (2011).
14. T. Harko, T. S. Koivisto, F. S. N. Lobo and G. J. Olmo, *Phys. Rev. D* **85**, 084016 (2012).
15. S. Capozziello, T. Harko, T. S. Koivisto, F. S. N. Lobo and G. J. Olmo, *Universe* **1** no.2, 199 (2015).
16. L. Järv, M. Rünkla, M. Saal and O. Vilson, *Phys. Rev. D* **97**, no. 12, 124025 (2018).
17. J. W. Maluf, *Annalen Phys.* **525**, 339 (2013).
18. J. B. Jimenez, L. Heisenberg and T. Koivisto, *Phys. Rev. D* **98**, 044048 (2018).
19. A. Conroy and T. Koivisto, *Eur. Phys. J. C* **78**, 923 (2018).
20. T. Koivisto, *Int. J. Geom. Meth. Mod. Phys.* **15**, 1840006 (2018).
21. J. Beltrn Jimnez, L. Heisenberg and T. S. Koivisto, *JCAP* **1808**, 039 (2018).
22. T. Harko, T. S. Koivisto, F. S. N. Lobo, G. J. Olmo and D. Rubiera-Garcia, *Phys. Rev. D* **98**, 084043 (2018).

# Some classical features of polynomial higher derivative gravities

Breno Loureiro Giacchini*

*Department of Physics, Southern University of Science and Technology, Shenzhen, Guangdong, China and Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, Rio de Janeiro, Brazil*
**E-mail: breno@sustech.edu.cn*

Tibério de Paula Netto

*Department of Physics, Southern University of Science and Technology, Shenzhen, Guangdong, China and Departamento de Física, ICE, Universidade Federal de Juiz de Fora, Juiz de Fora, Minas Gerais, Brazil*
*E-mail: tiberio@sustech.edu.cn*

In this talk we discuss some classical aspects of general polynomial higher-derivative gravity. In particular, we describe the behaviour of the weak-field solutions associated to a point-like mass at small distances and provide necessary and sufficient conditions for the metric to be regular. We also consider the metric for a collapsing thick null shell, and verify that it is regular if the aforementioned conditions are valid.

*Keywords*: Higher-derivative gravity; Lee-Wick gravity; Spacetime singularities.

## 1. Introduction

Higher-derivative theories of gravity possess interesting features from both quantum and classical point of views. Indeed, the inclusion of higher-derivative terms in the Lagrangian can make the theory perturbatively renormalizable,[1,2] while they can also smooth out classical singularities.[3–6] Owed to these good properties, higher derivatives are often considered in the search for a fundamental theory of gravity. One of the main difficulties of this approach to quantum gravity, nonetheless, is the presence of ghost-like degrees of freedom, which classically generate instabilities in the solutions and, in the quantum perspective, violate unitarity. For example, consider the model defined by the action[2]

$$S = \frac{1}{4\kappa}\int d^4x\sqrt{-g}\left(2R + RF_1(\Box)R + R_{\mu\nu}F_2(\Box)R^{\mu\nu}\right), \tag{1}$$

where $F_j(\Box)$ is a real polynomial of degree $\delta_j$ of the d'Alembert operator. The propagator associated this model, in the Landau gauge, is given by

$$G_{\mu\nu,\alpha\beta}(k) = \frac{P^{(2)}_{\mu\nu,\alpha\beta}}{k^2 f_2(-k^2)} - \frac{P^{(0-s)}_{\mu\nu,\alpha\beta}}{2k^2 f_0(-k^2)}, \tag{2}$$

where $f_0$ and $f_2$ are polynomial functions of degree $d_0 = \max\{\delta_1, \delta_2\} + 1$ and $d_2 = \delta_2 + 1$, respectively, defined as

$$f_0(\Box) = 1 - [F_2(\Box) + 3F_1(\Box)]\,\Box\,, \quad f_2(\Box) = 1 + \frac{1}{2}F_2(\Box)\,\Box\,. \tag{3}$$

Hence, if the polynomial $f_s(z)$ has $N_s$ distinct real roots $z = -m^2_{(s)i}$ with multiplicity $n_{(s)i}$, then the propagator has $N_0 + N_2$ massive poles. In Ref. 2 it was shown that half of these excitations correspond to ghost modes.

Recently there have been some proposals for dealing with these ghosts still in the framework of higher-derivative gravity (HDG). For example, in the Lee-Wick HDG[7,8] the ghost-like poles of the propagator are associated to complex masses. Such modes can then appear as virtual states only, yielding an unitary scattering matrix.[7] Another approach to avoid ghosts is to make a non-local extension of the polynomial HDG by using non-polynomial functions $F_j$ such that the propagator has no other degrees of freedom besides the graviton.[9–13]

In this talk we present some classical aspects of the general polynomial-derivative gravity model (1), namely, we discuss the avoidance of Newtonian singularities. We remark that, in the weak field approximation, the local model (1) is the most general one with higher derivatives, and it contains the case of Lee-Wick gravity as particular case. The original results presented here were published in Ref. 14. Using similar arguments it is possible to extend considerations to some classes of non-local gravity theories, as carried out in a more general manner in Ref. 15.

## 2. Singularities in the Newtonian limit

In order to evaluate the field generated by a point-like source in the static non-relativistic weak-field approximation we consider metric fluctuations around the Minkowski spacetime, $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$, sourced by the energy-momentum tensor $T_{\mu\nu} = M\,\delta^3(\mathbf{r})\,\delta^0_\mu\,\delta^0_\nu$. In this case the metric can be written in the isotropic form

$$ds^2 = -\left[1 + \frac{2}{3}(2\chi_2 + \chi_0)\right] dt^2 + \left[1 - \frac{2}{3}(\chi_2 - \chi_0)\right](dx^2 + dy^2 + dz^2) \qquad (4)$$

and one can show that the (linearised) equations of motion for the potentials $\chi_s$ $(s = 0, 2)$ are equivalent to solving[14]

$$f_s(\Delta)\Delta\chi_s = \kappa_s\, M\, \delta^3(\mathbf{r})\,, \qquad (5)$$

with $\kappa_0 = -\kappa/2$, $\kappa_2 = \kappa$. We remark that the decomposition of the usual Newtonian potentials into $\varphi = \frac{1}{3}(2\chi_2 + \chi_0)$ and $\psi = \frac{1}{3}(\chi_2 - \chi_0)$ allows a great simplification in the notation and considerations, as it splits the contribution of the scalar and spin-2 modes, through $\chi_0$ and $\chi_2$, respectively.

The solution of (5) can be obtained by the Laplace[14,16] or Fourier[15] transform technique and reads

$$\chi_s(r) = -\frac{\kappa_s M}{4\pi r} + \frac{\kappa_s M}{4\pi^{3/2}} \sum_{i=1}^{N_s} \sum_{j=1}^{n_{(s)i}} \frac{a_{(s)i,j}}{(j-1)!} \left(\frac{r}{2m_{(s)i}}\right)^{j-\frac{3}{2}} K_{j-\frac{3}{2}}(m_{(s)i} r)\,, \qquad (6)$$

where $K_\nu$ is the modified Bessel function of the second kind. Also, $z = -m^2_{(s)i}$ is one of the $N_s$ distinct roots of the equation $f_s(-z) = 0$ and $n_{(s)i}$ is its multiplicity. Of course, if $d_s$ is the degree of $f_s(z)$, then $\sum_i n_{(s)i} = d_s$. The coefficient $a_{(s)i,j}$ can

be obtained by the Heaviside residue method, and follows from the partial fraction decomposition of $[zf_s(-z)]^{-1}$.

The solution above can be expanded in power series around $r = 0$, which gives

$$\chi_s(r) = -\frac{\kappa_s M}{4\pi r}\left(1 - S_s^{(0)}\right) + c_s + \frac{\kappa_s M}{8\pi}\sum_{i=1}^{N_s}\left(S_s^{(1)} - S_s^{(2)}\right) r + O(r^2)\,, \qquad (7)$$

where $c_s$ is a constant and

$$S_s^{(0)} = \sum_{i=1}^{N_s} a_{(s)i,1}\,, \qquad S_s^{(1)} = \sum_{i=1}^{N_s} m_{(s)i}^2 a_{(s)i,1}\,, \qquad S_s^{(2)} = \sum_{i=1}^{N_s} a_{(s)i,2}. \qquad (8)$$

It is possible to show[5,14] that $S_s^{(0)} = 1$ for any $d_s \geqslant 1$, which means that all the theories which have at least four derivatives in the spin-$s$ sector have a finite potential $\chi_s(r)$ at $r = 0$. Moreover, one can prove[14,15] that $S_s^{(1)} = S_s^{(2)}$ if and only if $d_s \geqslant 2$ — in other words, $\chi_s'(0) = 0$ in all theories with more than four derivatives in the spin-$s$ sector. It can be shown that, for these models, the condition $\chi_0'(0) = \chi_2'(0) = 0$ is necessary and sufficient to have regular curvature invariants.[14,16] Thus, it follows that the metric (4) is regular in all theories which have at least six derivatives in both spin-0 and spin-2 sectors.[14,15]

## 3. Singularities in the ultrarelativistic limit

The static Newtonian solution presented in the previous section can be used to construct the metric of a non-spinning gyraton (see, *e.g.*, Refs. 17, 18). The general idea is to apply a boost to the non-relativistic metric and then take the Penrose limit. With this solution one can consider a homogeneous spherical shell distribution of gyratons with total mass $M$ imploding towards its centre, and analyse the occurrence of singularities and the formation of mini black holes.[16,17,19]

In Ref. 14 it was shown that the Kretschmann scalar associated to the $\mathcal{I}$ domain[a] of the collapsing thick null shell in a general polynomial theory is

$$R^2_{\mu\nu\alpha\beta} \;=\; \frac{32G^2M^2}{3\tau^2}\left[4\Delta^2(\ln r)^2 + c'\Delta\ln r \;+\; c''\Delta + \left(S_2^{(0)} - S_2^{(1)}\right)^2\right] + O(r^2)\,, \qquad (9)$$

where $\tau > 0$ is the thickness of the shell, $\Delta \equiv S_2^{(1)} - S_2^{(2)}$ and $c'$ and $c''$ are constants which depend on the massive parameters of the theory. Inasmuch as $\Delta = 0$ for the theories with more than four derivatives in the spin-2 sector, the collapse of the thick null shell does not generate a singularity. On the other hand, the Kretschmann invariant diverges for models with less than six derivatives.

[a]The $\mathcal{I}$ domain, defined by the locus of the spacetime points for which $r+|t| < \tau/2$, is characterized by the intersection of the in-coming and the out-coming fluxes of null fluid. In this domain the shell assumes its highest density, favouring the mini black hole formation and the emergence of singularities.

The solution for the collapsing shell can also be used to verify the existence of mass gap for the formation of mini black holes. This is related to the occurrence of apparent horizons in the solution, *i.e.*, regions such that $g \equiv (\nabla\varrho)^2 = 0$, where $\varrho$ is the component $g_{\theta\theta}$ of the spherically symmetric metric. In fact, one can show that the invariant $g$ reads

$$g(r) \,=\, 1 + \frac{2GM(S_2^{(0)} - S_2^{(2)})r^2}{3\tau} \,+\, O(r^4) \tag{10}$$

for the collapsing thick null shell in any polynomial gravity theory.[14] Since $r < \tau$ on the domain $\mathcal{I}$, it follows that

$$\frac{2GM|S_2^{(0)} - S_2^{(2)}|r^2}{3\tau} < \frac{2GM|S_2^{(0)} - S_2^{(2)}|\tau}{3}\,. \tag{11}$$

In other words, given any $\tau$ it is possible to avoid the existence of an apparent horizon on $\mathcal{I}$ provided that the total mass $M$ of the shell is sufficiently small. This result was obtained for the first time in Ref. 19 for the theory with only four derivatives; in Refs. 17 and 16 it was extended for the ghost-free gravity and polynomial gravity with only real and non-degenerate masses, while in Ref. 14 general polynomial theories were considered, including the case of Lee-Wick gravity.

## 4. Conclusions

The results presented here show that there is a significant difference of the HDG theories with four and more derivatives. Even though the modified Newtonian potential is finite in both cases (if there are at least four derivatives in the spin-2 and spin-0 sectors), the former always contains curvature singularities in the solution associated to a point-like mass. On the other hand, these singularities are regularized in the models with more than four derivatives. The same situation occurs for the collapsing thick null shell, for which curvature singularities are avoided only in the models with six or more derivatives.

These results bring more motivations for further studies on the occurrence of singularities in the spherically symmetric static solutions in the full non-linear regime of HDG models. For example, the numerical searches for solutions in theories with 6, 8 and 10 derivatives reported in Ref. 20 only found regular solutions. In what concerns the fourth-derivative gravity, it is known that singularities are present in both regimes.[3,21,22] This subject certainly deserves more investigation.

## Acknowledgments

B.L.G. is grateful to CNPq for supporting his Ph.D. studies, and to ICTP for partial financial support to his attendance at the 15$^{\text{th}}$ Marcel Grossmann Meeting. The work of T.P.N. was partially funded by the PNPD program by CAPES.

## References


1. K.S. Stelle, *Phys. Rev. D* **16**, 953 (1977).
2. M. Asorey, J.L. López and I.L. Shapiro, *Int. J. Mod. Phys. A* **12**, 5711 (1997), arXiv:hep-th/9610006.
3. K.S. Stelle, *Gen. Rel. Grav.* **9**, 353 (1978).
4. L. Modesto, T. de Paula Netto and I.L. Shapiro, *J. High Energy Phys.* **1504**, 098 (2015), arXiv:1412.0740.
5. B.L. Giacchini, *Phys. Lett. B* **766**, 306 (2017), arXiv:1609.05432.
6. A. Accioly, B.L. Giacchini and I.L. Shapiro, *Phys. Rev. D* **96**, 104004 (2017), arXiv:1610.05260.
7. L. Modesto and I.L. Shapiro, *Phys. Lett. B* **755**, 279 (2016), arXiv:1512.07600.
8. L. Modesto, *Nucl. Phys. B* **909**, 584 (2016), arXiv:1602.02421.
9. N.V. Krasnikov, *Theor. Math. Phys.* **73**, 1184 (1987) [*Teor. Mat. Fiz.* **73**, 235 (1987)].
10. Yu.V. Kuz'min, *Sov. J. Nucl. Phys.* **50**, 1011 (1989) [*Yad. Fiz.* **50**, 1630 (1989)].
11. E.T. Tomboulis, arXiv:hep-th/9702146, *Superrenormalizable gauge and gravitational theories*, 1997.
12. L. Modesto, *Phys. Rev. D* **86**, 044005 (2012), arXiv:1107.2403.
13. T. Biswas, E. Gerwick, T. Koivisto and A. Mazumdar, *Phys. Rev. Lett.* **108** 031101 (2012), arXiv:1110.5249.
14. B.L. Giacchini and T. de Paula Netto, *Eur. Phys. J. C* **79**, 217 (2019), arXiv:1806.05664.
15. B.L. Giacchini and T. de Paula Netto, *J. Cosmol. Astropart. Phys.* **1907** 013 (2019), arXiv:1809.05907.
16. V.P. Frolov, *Phys. Rev. Lett.* **115**, 051102 (2015), arXiv:1505.00492.
17. V.P. Frolov, A. Zelnikov and T. de Paula Netto, *J. High Energy Phys.* **1506**, 107 (2015), arXiv:1504.00412.
18. V.P. Frolov and A. Zelnikov, *Introduction to Black Hole Physics* (Oxford University Press, New York, 2011).
19. V.P. Frolov and G.A. Vilkovisky, *Phys. Lett. B* **106**, 307 (1981).
20. B. Holdom, *Phys. Rev. D* **66**, 084010 (2002), arXiv:hep-th/0206219.
21. H. Lü, A. Perkins, C.N. Pope and K.S. Stelle, *Phys. Rev. Lett.* **114**, 171601 (2015), arXiv:1502.01028.
22. H. Lü, A. Perkins, C.N. Pope and K.S. Stelle, *Phys. Rev. D* **92**, 124019 (2015), arXiv:1508.00010.

# New 2D dilaton gravity for nonsingular black holes

Hideki Maeda

*Department of Electronics and Information Engineering, Hokkai-Gakuen University, Sapporo 062-8605, Japan*
*E-mail: h-maeda@hgu.jp*

Gabor Kunstatter

*Department of Physics, University of Winnipeg and Winnipeg Institute for Theoretical Physics, Winnipeg, Manitoba, Canada R3B 2E9*
*E-mail: g.kunstatter@uwinnipeg.ca*

Tim Taves

*Centro de Estudios Científicos (CECs), Casilla 1469, Valdivia, Chile*
*E-mail: timtaves@gmail.com*

We construct a two-dimensional action that is an extension of spherically symmetric Lovelock gravity. In spite that the action contains arbitrary functions of the areal radius and the norm squared of its gradient, the field equations are second order and obey the Birkhoff's theorem. Similar to the spherically symmetric Lovelock gravity, the field equations admit the generalized Misner-Sharp mass that determines the form of the vacuum solution. The arbitrary functions in the action allow for vacuum solutions that describe a larger class of nonsingular black-hole spacetimes than previously available.

*Keywords*: Nonsingular black hole; Two-dimensional effective theory.

## 1. 2D effective actions for spherically symmetric spacetimes

The metric for $n(\geq 3)$-dimensional spherically symmetric spacetimes is given by

$$\begin{aligned} \mathrm{d}s^2_{(n)} &= g_{\mu\nu}(x)\mathrm{d}x^\mu \mathrm{d}x^\nu \\ &= \bar{g}_{AB}(y)\mathrm{d}y^A \mathrm{d}y^B + R(y)^2 \mathrm{d}\Omega^2_{(n-2)}, \end{aligned} \tag{1}$$

where $\bar{g}_{AB}(y)$ $(A, B = 0, 1)$ is the general two-dimensional (2D) Lorentzian metric, $\mathrm{d}\Omega^2_{(n-2)}$ is the line-element on the unit $(n-2)$-sphere, and $R$ is the areal radius.

After imposing spherical symmetry and integrating out the angular variables, the general $n(\geq 3)$-dimensional gravitational action,

$$I_n = \frac{1}{16\pi G_{(n)}} \int \mathrm{d}^n x \sqrt{-g}\mathcal{L}(\mathcal{R}, \mathcal{R}_{\mu\nu}, \mathcal{R}_{\mu\nu\rho\sigma}), \tag{2}$$

reduces to a 2D effective action. The variation of this effective 2D action will give the same equations of motion for the original action (2)[1–3].

We adopt units such that $c = \hbar = 1$ and $G_{(n)}$ denotes the $n$-dimensional gravitational constant. In the following, $D_A$ and $\mathcal{R}[\bar{g}]$ denote the covariant derivative and the Ricci scalar with respect to $\bar{g}_{AB}$, respectively. We also define $(DR)^2 := (D_A R)(D^A R)$ and a length parameter $l$ proportional to the Planck length as $l^{n-2} := 16\pi G_{(n)}/\mathcal{A}_{(n-2)}$, where $\mathcal{A}_{(n-2)}$ is the volume of a unit $(n-2)$-sphere. The complete analysis for the results presented here is available in Ref. 4.

### 1.1. *Effective 2D action for Einstein gravity and its generalization*

The Einstein-Hilbert action for general relativity corresponds to $\mathcal{L} = \mathcal{R}$ in the action (2) and its effective 2D action takes the form

$$I_{\mathrm{EH}(2)} = \frac{1}{l^{n-2}} \int \mathrm{d}^2y\sqrt{-\bar{g}}\Big\{ R^{n-2}\mathcal{R}[\bar{g}] + (n-2)(n-3)R^{n-4}(DR)^2 + (n-2)(n-3)R^{n-4}\Big\}. \tag{3}$$

By the Birkhoff's theorem, the unique vacuum solution with spherical symmetry is the well-known Schwarzschild-Tangherlini solution.

A natural way to generalize the spherically symmetric action (3) in Einstein gravity is the following 2D dilaton gravity:

$$I_{(2)} = \frac{1}{l^{n-2}} \int \mathrm{d}^2y\sqrt{-\bar{g}}\Big\{ \phi(R)\mathcal{R}[\bar{g}] + h(R)(DR)^2 + V(R)\Big\}, \tag{4}$$

where $\phi(R)$, $h(R)$, and $V(R)$ are arbitrary functions of a scalar field $R$. (See Ref. 5 for a review on this class of 2D gravity.) This standard 2D dilaton gravity theory (4) obeys the Birkhoff's theorem[6]. Namely, the general vacuum solution has the following form:

$$\mathrm{d}s^2 = -f(R)\mathrm{d}t^2 + f(R)^{-1}\mathrm{d}R^2. \tag{5}$$

If one chooses $h(R) = V(R) = \phi_{,RR}(R)$, the metric function is given by

$$f(R) = 1 - \frac{l^{n-2}M}{j(R)}, \quad \Big(j(R) := \int V(R)\mathrm{d}R\Big). \tag{6}$$

Especially, in the case with $n = 4$ and $\phi_{,R} = j(R) = (R^2+l^2)^{3/2}/R^2$, the general vacuum solution is the well-known Bardeen metric[7]:

$$\mathrm{d}s^2_{(4)} = -\left(1 - \frac{l^2MR^2}{(R^2+l^2)^{3/2}}\right)\mathrm{d}t^2 + \left(1 - \frac{l^2MR^2}{(R^2+l^2)^{3/2}}\right)^{-1}\mathrm{d}R^2 + R^2\mathrm{d}\Omega^2_{(2)}. \tag{7}$$

The Bardeen spacetime (7) is certainly nonsingular everywhere, however, this class of nonsingular black holes are considered to be unphysical. This is because the metric (7) violates the *limiting curvature conjecture*, which asserts that the curvature invariants are bounded by some fundamental value in a viable fundamental theory[8]. In fact, to the best of our knowledge, this limiting curvature condition cannot be fulfilled within the framework of the action for pure 2D dilaton gravity (4). This is the main reason why we consider a more general class of 2D dilaton gravity.

### 1.2. *Effective 2D action for Lovelock gravity and its generalization*

Lovelock gravity is a natural generalization of general relativity in arbitrary dimensions as a second-order quasilinear theory of gravity[9]. The second-order field equations ensure the ghost-free nature of the theory and Lovelock gravity reduces to general relativity with a cosmological constant in four dimensions.

In the action (2), Lovelock gravity in vacuum corresponds to

$$\mathcal{L}=\sum_{p=0}^{[n/2]}2^{-p}\alpha_{(p)}\delta^{\mu_1\cdots\mu_p\nu_1\cdots\nu_p}_{\rho_1\cdots\rho_p\sigma_1\cdots\sigma_p}\mathcal{R}_{\mu_1\nu_1}{}^{\rho_1\sigma_1}\cdots\mathcal{R}_{\mu_p\nu_p}{}^{\rho_p\sigma_p}, \tag{8}$$

where $\delta^{\mu_1\cdots\mu_p}_{\rho_1\cdots\rho_p}:=p!\delta^{\mu_1}_{[\rho_1}\cdots\delta^{\mu_p}_{\rho_p]}$, and its effective 2D action was obtained[10,11] as

$$\begin{aligned}I_{\mathrm{L}(2)}=&\frac{1}{l^{n-2}}\int \mathrm{d}^2y\sqrt{-\bar{g}}R^{n-2}\sum_{p=0}^{[n/2]}\frac{(n-2)!}{(n-2p)!}\alpha_{(p)}\\&\times\Bigg[p\mathcal{R}[\bar{g}]R^{2-2p}+(n-2p)(n-2p-1)\Big\{(1-Z)^p+2pZ\Big\}R^{-2p}\\&+p(n-2p)R^{1-2p}\Big\{1-(1-Z)^{p-1}\Big\}(D_AR)\frac{(D^AZ)}{Z}\Bigg],\end{aligned} \tag{9}$$

where we have defined $Z:=(DR)^2$. The Birkhoff's theorem in Lovelock gravity shows that, under several technical assumptions, the unique vacuum solution is given by the Schwarzschild-Tangherlini-type solution[12,13].

In analogy with the action (4), we now propose[4] the following natural extension of the spherically symmetric Lovelock action (9)

$$I_{\mathrm{XL}}=\frac{1}{l^{n-2}}\int \mathrm{d}^2y\sqrt{-\bar{g}}\Big\{\phi(R)\mathcal{R}[\bar{g}]+\eta(R,Z)+\chi(R,Z)(D_AR)\frac{(D^AZ)}{Z}\Big\}, \tag{10}$$

where $\eta(R,Z)$ and $\chi(R,Z)$ are as yet arbitrary functions of a scalar field $R$ and $Z$. For any given $\phi(R)$ and $\chi(R,Z)$, one can choose the function $\eta(R,Z)$ as

$$\phi_{,RR}=\eta_{,Z}-\chi_{,R}, \tag{11}$$

where a comma denotes the partial derivative, so that the field equations obey the Birkhoff's theorem for $Z=(DR)^2\neq 0$ and $\chi-\phi_{,R}\neq 0$. Then the resulting general vacuum solution has the following form:

$$\mathrm{d}s^2=-f(R)\mathrm{d}t^2+f(R)^{-1}\mathrm{d}R^2. \tag{12}$$

Actually, the condition (11) ensures the existence of the generalized Misner-Sharp mass which satisfies the unified first law. Under the condition (11), the existence of Minkowski vacuum requires

$$\eta(R,1)=2\phi_{,RR}. \tag{13}$$

## 2. Designing nonsingular black holes

Now we show how to construct specific nonsingular black holes as exact solutions by making appropriate choices for the functions in the action (10). We are interested in constructing nonsingular black holes that satisfy the limiting curvature condition, namely curvature invariants are everywhere bounded for arbitrarily large $M$.

Such an example is the following Hayward nonsingular black hole[14]:

$$f(R) = 1 - \frac{l^2 M R^2}{R^3 + l^4 M}, \tag{14}$$

of which generalization in $n$ dimensions is given by[4]

$$f(R) = 1 - \frac{l^{n-2} M R^2}{R^{n-1} + l^n M}. \tag{15}$$

This $n$-dimensional Hayward black hole (15) is the unique vacuum solution in the theory with

$$\eta(R, Z) = 2\phi_{,RR} Z + \frac{(n-3) l^{n-2} R^n (1-Z) - (n-1) l^n R^{n-2} (1-Z)^2}{\{l^{n-2} R^2 - l^n (1-Z)\}^2},$$
$$\chi(R, Z) = \phi_{,R} - \frac{l^{n-2} R^{n+1}}{\{l^{n-2} R^2 - l^n (1-Z)\}^2}. \tag{16}$$

Also, the following Bardeen-type nonsingular black hole[4]

$$f(R) = 1 - \frac{l^{n-2} M R^2}{(R^2 + M^{2/(n-1)} l^{2n/(n-1)})^{(n-1)/2}} \tag{17}$$

or new nonsingular black hole[4]

$$f(R) = 1 + \frac{R^{n+1}}{2 l^{n+2} M} \left( 1 - \sqrt{1 + \frac{4 l^{2n} M^2}{R^{2(n-1)}}} \right) \tag{18}$$

can be the unique vacuum solution in the theory with a suitable choice of $\eta(R, Z)$ and $\chi(R, Z)$.

## References


1. R.S. Palais, *Comm. Math. Phys.* **69**, 19 (1979).
2. M.E. Fels and C.G. Torre, *Class. Quant. Grav.* **19**, 641 (2002).
3. S. Deser and B. Tekin, *Class. Quant. Grav.* **20**, 4877 (2003).
4. G. Kunstatter, H. Maeda, and T. Taves, *Class. Quant. Grav.* **33**, 105005 (2016).
5. D. Grumiller, W. Kummer, and D.V. Vassilevich, *Physics Reports* **369**, 327 (2002).
6. D. Louis-Martinez and G. Kunstatter, *Phys. Rev. D* **49**, 5227 (1994).
7. J. Bardeen. *Non-Singular General-Relativistic Gravitational Collapse.* Presented at GR5, Tiblisi, U.S.S.R., and published in the conference proceedings in the U.S.S.R., 1968.
8. V.P. Frolov, *Phys. Rev. D* **94**, 104056 (2016).
9. D. Lovelock, *J. Math. Phys.* **12**, 498 (1971).
10. G. Kunstatter, T. Taves, and H. Maeda, *Class. Quant. Grav.* **29**, 092001 (2012).
11. G. Kunstatter, H. Maeda, and T. Taves, *Class. Quant. Grav.* **30**, 065002 (2013).
12. R. Zegers, *J. Math. Phys.* **46**, 072502 (2005).
13. H. Maeda, S. Willison, and S. Ray, *Class. Quant. Grav.* **28**, 165005 (2011).
14. S.A. Hayward, *Phys. Rev. Lett.* **96**, 031103 (2006).

# Gravitational waves in a molecular medium: dispersion, extra polarizations and quantitative estimates

Giovanni Montani

*ENEA, Fusion and Nuclear Safety Department, C. R. Frascati, Via E. Fermi 45*
*00044 Frascati (Roma), Italy*
*Physics Department, Sapienza University of Rome, P.le Aldo Moro 5*
*00185 Roma, Italy*

Fabio Moretti

*Physics Department, Sapienza University of Rome, P.le Aldo Moro 5*
*00185 Roma, Italy*
*E-mail: fabio.moretti@uniroma1.it*

We analyze the propagation of gravitational waves in a molecular matter medium: our findings demonstrate that dispersion only is expected, together with the emergence of three extra polarizations, able to induce longitudinal stresses. We also give quantitative estimates of the predicted effects.

*Keywords*: Macroscopic gravity, gravitational wave dispersion, extra polarizations, longitudinal stress, multi-messenger astronomy.

## 1. Introduction

The Theory of General Relativity provides us a very accurate and elegant theoretical framework from which forecast on all natural phenomena dominated by gravity. One of the most fascinating features of the theory is the prediction of the fact that the gravitational field is radiative: curvature can propagate through spacetime with the speed of light. With the first direct measurement of a gravitational wave, in 2015, we are witnessing the birth of a new era in the fields of astronomy, astrophysics and cosmology. The major part of theoretical studies have dealt with the asymptotic properties of gravitational waves, far from bounded sources, Refs. 1, 2, 3, 4. Many other authors have dealt with the problem of the propagation of gravitational waves in a matter medium: the case of a dissipative fluid is considered in Refs. 5, 6, whilst a collisionless kinetic gas is studied in Refs. 7, 8, 9, 10, 11; the interaction between gravitational waves and neutrinos is analyzed in Refs. 12, 13, 14 whereas the case of spinning particles is investigated in Ref. 15.

Our aim is to analyze the propagation of gravitational wave in a matter medium described as a set of point-like masses, grouped into molecules: it has been shown, Ref. 16 that such a medium can be modelled, through the application of Kaufman molecular moments averaging method, Ref. 17 (whose application is denoted by $\langle\cdot\rangle$), as

$$\langle T_{\mu\nu}\rangle = T^{(f)}_{\mu\nu} + \frac{c^2}{2}\,\partial^\rho\partial^\sigma Q_{\mu\rho\nu\sigma}, \tag{1}$$

where $T^{(f)}_{\mu\nu}$ is the stress energy tensor that describes a set of free particles, *i.e.* the centers of mass of the molecules, and $Q_{\mu\rho\nu\sigma}$ is the quadrupole polarization tensor, that gives account of the molecular structure, characterized by the same set of symmetries of Riemann tensor.

We find, through a direct calculation performed via the geodesic deviation equation, that the quadrupole tensor is generated, on a static level, by the tensor of the second spatial derivatives of the Newtonian potential of the molecule. Moreover, if the molecule is perturbed with a vacuum gravitational wave we find that the latter is source of an induced contribution to the quadrupole moment of the molecule: the alteration in the structure of the molecules causes the fact that the medium reacts to the gravitational wave with a mean field that modifies the propagation of the wave inside the medium itself.

We show that only dispersion is expected, along with the appearance of three extra degrees of freedom able to induce longitudinal stresses on test particles. We also give some quantitative estimates of the predicted effects, by assuming the medium to be either similar to our own galaxy or to a more exotic object.

## 2. Constitutive relation and modified wave equation

The derivation of the constitutive relation is made under the hypothesis of slow motion of the point-like masses composing the molecule and small departure from Minkowski background metric. Moreover, the molecule is imagined as a sphere of constant mass density, characterized by a definite macroscopic radius $L$: the reduced wavelength $\lambda\!\!\!^{-}$ of the gravitational radiation interacting with the molecule must be greater than the molecular size. We perturb the molecule with a vacuum, plus polarized gravitational wave and we express the quadrupole tensor in terms of the amplitudes of the wave. It can be checked that the following constitutive relation comes out:

$$Q_{i0j0} = \epsilon_g \left( \frac{2}{c^2}\, \phi_{,ij} + \frac{1}{2}\, h_{ij,00} \right) \qquad , \qquad Q_{0ijk} = Q_{ijkl} = 0. \tag{2}$$

Here $\phi$ is the Newtonian potential of the molecule and $h_{\mu\nu}$ is the small departure from Minkowski background, the dynamical metric perturbation called gravitational wave. The constant $\epsilon_g$ is a real gravitational dielectric constant and can be calculated as

$$\epsilon_g = \frac{N L^5 c^2}{4G}, \tag{3}$$

being $N$ the density of molecules inside the medium. We write the macroscopic form of Einstein field equation

$$G_{\mu\nu} = \chi \left( T^{(f)}_{\mu\nu} + \frac{c^2}{2}\, \partial^\rho \partial^\sigma Q_{\mu\rho\nu\sigma} \right) \qquad , \qquad \chi = \frac{8\pi G}{c^4} \tag{4}$$

at first order in $h_{\mu\nu}$, implementing the constitutive relation (2): we find a set of fourth order partial differential equation for the wave inside the medium. We write

the problem in terms of the trace reversed matrix $\bar{h}_{\mu\nu}$ on which we are allowed to activate Hilbert gauge fixing, given the diffeomorphism invariance of the theory:

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - \frac{1}{2}\eta_{\mu\nu}h \qquad , \qquad \partial^\mu \bar{h}_{\mu\nu} = 0. \tag{5}$$

We find that the trace $\bar{h}$ solves

$$\Box\left(\partial_0^2 + m^2\right)\bar{h} = 0, \tag{6}$$

where $\Box$ is d'Alembert wave operator and $m^2 = \dfrac{c^2}{4\pi G\epsilon_g}$: hence $\bar{h}$ is a superposition of an harmonic part, solution of d'Alembert equation, that can be canceled out with a further gauge transformation that preserves Hilbert gauge, plus a solution of the harmonic oscillator $\partial_0^2 + m^2$ that does not propagate. This means that we can set $\bar{h}$ to zero: the remaining five degrees of freedom solve the fourth order PDE

$$\partial_0^4 \bar{h}_{\mu\nu} + m^2\Box\bar{h}_{\mu\nu} = 0. \tag{7}$$

We search for plane wave solutions of (7), choosing the $z$ axis to be coincident with the direction of propagation of the wave, $\bar{h}_{\mu\nu} \propto e^{i(\omega(k)t - kz)}$ and we calculate the dispersion relation:

$$\omega_\pm^2(k) = c^2\left(-\frac{m^2}{2} \pm \sqrt{\frac{m^4}{4} + m^2k^2}\right). \tag{8}$$

It can be shown that only the plus signed solution has a physical significance, because it goes back to be the standard vacuum dispersion relation $\omega(k) = ck$ when the material medium is removed, in the limit $\epsilon_g \to 0$, or $m^2 \to \infty$, in contrast to the minus signed, that shows a divergence in the same limit. It is also easy to show that $\omega_+(k)$ is real for any value of the wavenumber: dispersion only is expected. We calculate the group velocity

$$v_g(k) = \frac{m^2kc}{2\sqrt{\left(\sqrt{m^2k^2 + \dfrac{m^4}{4}} - \dfrac{m^2}{2}\right)\left(m^2k^2 + \dfrac{m^4}{4}\right)}} \xrightarrow[k \ll m]{} c\left(1 - \frac{3k^2}{2m^2}\right). \tag{9}$$

We will show that the condition $\lambda\!\!\!^{-} \gg L$ always implies $k \ll m$.

## 3. Polarizations

We exploit the gauge freedom, fixing

$$\bar{h} = \bar{h}_{11} + \bar{h}_{22} + \left(1 - \frac{c^2k^2}{\omega^2}\right)\bar{h}_{33} = 0, \tag{10}$$

via the following equations:

$$\begin{aligned} \bar{h}_{11} &= \bar{h}_+ + \bar{h}^* \\ \bar{h}_{22} &= -\bar{h}_+ + \bar{h}^* \\ \bar{h}_{33} &= \frac{2\omega^2}{c^2k^2 - \omega^2}\bar{h}^*. \end{aligned} \tag{11}$$

In order to analyze which kind of deformation is induced by each component on a sphere of test particles, we calculate the geodesic deviation equation, taken in the comoving frame

$$\frac{1}{c^2}\frac{d^2\xi^\mu}{dt^2} = R^\mu_{\nu\rho\sigma}u^\nu u^\rho \xi^\sigma = R^\mu_{00i}\xi^i, \tag{12}$$

where $\xi^\mu = (0, \xi^x, \xi^y, \xi^z)$ is a vector denoting the separation between two nearby geodesics and the Riemann tensor is constructed up to the first order in $\bar{h}_{\mu\nu}$. We fix the vector $\xi^\mu$ as

$$\xi^\mu = \left(0, \xi^x_{(0)} + \delta\xi^x, \xi^y_{(0)} + \delta\xi^y, \xi^z_{(0)} + \delta\xi^z\right), \tag{13}$$

being $\xi^i_{(0)}$ the initial positions and $\delta\xi^i$ the displacements of order $O(h)$ induced by the gravitational wave. We compute (12) up to order $h$ separately for each component of $\bar{h}_{\mu\nu}$, finding that:

(1) $\bar{h}_+$ is a standard plus mode in the $xy$ plane,
(2) $\bar{h}_{12}$ is a standard cross mode in the $xy$ plane,
(3) $\bar{h}_{13}$ is a cross mode in the $xz$ plane,
(4) $\bar{h}_{23}$ is a cross mode in the $yz$ plane,
(5) $\bar{h}^*$ is a superposition of two pure modes: a breathing mode (Refs. 18, 19, 20) in the $xy$ plane and a longitudinal mode along the $z$ direction.

The ratio between the amplitudes of the anomalous polarizations and the standard ones (plus and cross polarization in the transverse plane) are found to be of order $k^2/m^2$.

## 4. Models of macroscopic media

Now we will give some quantitative estimates of $\epsilon_g$, without any intention of being too accurate, but merely realistic, in the characterization of the involved parameters. Let us begin with the case of a medium composed by molecules whose size is roughly the typical size of a binary system, Refs. 21, 22, 23: we will set $L = 30\,AU = 4.4 \cdot 10^{12}\,m$. In terms of the wavenumber, the condition $\lambda \gg L$ reads as $kL \ll 1$: we will fix $kL = 0.05$, *i.e.* $k = 1.14 \cdot 10^{-14}\,m^{-1}$. The parameter $N$ changes considerably in different regions of the Galaxy, therefore we calculate $\epsilon_g$ for three different values of $N$: in $(i)$ we set $N = 1\,pc^{-3}$ (our neighborhood), in $(ii)$ we choose the value $N = 100\,pc^{-3}$ (inner region) and in $(iii)$ we fix $N = 10^5\,pc^{-3}$ (Galaxy core). We report the results obtained for $\epsilon_g$ and for the ratio $k^2/m^2$, that gives the entity of the deviation from $c$ in the group velocity and the ratio between the amplitudes of the anomalous polarizations and the standard ones:

$(i)$ $\epsilon_g = 1.89 \cdot 10^{40}\,kg\,m$ $\qquad$ $k^2/m^2 = 2.27 \cdot 10^{-14}$
$(ii)$ $\epsilon_g = 1.89 \cdot 10^{42}\,kg\,m$ $\qquad$ $k^2/m^2 = 2.27 \cdot 10^{-12}$
$(iii)$ $\epsilon_g = 1.89 \cdot 10^{45}\,kg\,m$ $\qquad$ $k^2/m^2 = 2.27 \cdot 10^{-9}$.

Now we consider the case of a material medium composed by molecules whose size is roughly comparable with the size of an open cluster, Refs. 24, 25: we set for $L$ the value $L = 3\,pc = 9.26{\cdot}10^{16}\,m$. As before, we fix $kL = 0.05$, *i.e.* $k = 5.40{\cdot}10^{-19}\,m^{-1}$. Our galaxy is estimated to contain about $100,000$ open clusters, mainly located in the central disc. We will calculate the density of molecules as the number of molecules ($10^5$) divided by the volume of the Galaxy, assuming it to be a disc with diameter of $45\,kpc$ and thickness of the disc of about $0.6\,kpc$ (roughly the size of the Milky Way, Ref. 26). We obtain the value $N = 3.55 \cdot 10^{-57}\,m^{-3}$. We calculate $\epsilon_g$ and $k^2/m^2$:

$$\epsilon_g = 8.13 \cdot 10^{54}\,kg\,m \qquad\qquad k^2/m^2 = 2.21 \cdot 10^{-8}.$$

One can imagine a medium in which the density of clusters is 5 or 50 times greater than in our Galaxy: being the ratio $k^2/m^2$ linear in $N$, we expect for such a medium a magnitude of Macroscopic Gravity effects of order $10^{-7} \div 10^{-6}$.

## 5. Conclusions

In this work we have demonstrated that the propagation of a gravitational wave in a molecular medium shows remarkable differences with respect to the vacuum case. By establishing a constitutive relation between the quadrupole tensor and the perturbative fields acting on the molecule, namely a vacuum gravitational wave $h_{\mu\nu}$ and the Newtonian potential of the molecule itself $\phi$, we have derived the law of motion for the wave inside the medium, a fourth order PDE: we have ruled out the possibility of damped and growing modes and determined that dispersion only occurs. We have outlined the fact that only the trace of the gravitational wave is still solution of d'Alembert equation, hence erasable with a gauge transformation. The remaining five degrees of freedom are not harmonic: they cannot be canceled out by a gauge transformation that preserves Hilbert gauge. This finding demonstrates that a gravitational wave in a matter medium possesses five degrees of freedom and it is able to induce longitudinal stresses in a sphere of test particles. We have showed that the quantity $k^2/m^2$ acts as a quantifier of two different observables: the deviation from $c$ in the group velocity and the ratio between the amplitudes of the anomalous polarizations and the standard ones. Lastly we have established some models of molecular medium, varying the values of the parameters $L$ and $N$: we have estimated the entity of Macroscopic Gravity effects, obtaining for $k^2/m^2$ values in the range $10^{-14} \div 10^{-6}$. We focused our analysis on two specific regions of wavelengths. In the first case (binaries) the signal is characterized by $\lambda \simeq 10^{13}\,m$, that is comparable with the scale of lengths to which the space interferometer LISA will be sensitive, Refs. 27, 28. In the second case (clusters) the wavelength of the signal is in the region $\lambda \simeq 10^{18}\,m$ and it is, in principle, detectable by experiments like IPTA, Refs. 29, 30. Despite the modest entity of the expected deviation from vacuum General Relativity, the integrated effect over very large distances can become significant and the propagation of gravitational waves can

be considerably different from that of an electromagnetic counterpart, in a typical "follow up" scenario of multi-messenger astronomy.

## References


1. Landau L.D., Lifshitz E.M., *The Classical Theory of Fields* (Pergamon, Oxford, 1971).
2. Misner C.W., Thorne K.S., Wheeler J.A., *Gravitation* (Freeman, San Francisco CA, 1973).
3. Wald R., *General Relativity* (The University of Chicago Press, Chicago and London, 1984).
4. Maggiore M., *Gravitational Waves* (Oxford University Press, Oxford, 2008).
5. Hawking S.W. *Astrophys. J.* **145**, 544 (1966).
6. Prasanna A.R. *Phys. Lett. A* **257**, 120 (1999).
7. Ignatyev Yu. G., Shulikovsky V. Yu. *Gravitation & Cosmology* **12**, 1 (2006).
8. Polnarev A.G. *Sov. Phys. JETP* **35**, 834 (1972).
9. Chesters D. *Phys. Rev. D* **7**, 2863 (1973).
10. Weinberg S. *Phys. Rev. D* **69**, 023503 (2004).
11. Flauger R., Weinberg S. *Phys. Rev. D* **97**, 123506 (2018).
12. Lattanzi M., Montani G. *Mod. Phys. Lett. A* **20**, 2607 (2005).
13. Lattanzi M., Benini R., Montani G. *Class. Quantum Grav.* **27**, 194008 (2010).
14. Benini R., Lattanzi M., Montani G. *Gen. Rel. Grav.* **43**, 945 (2011).
15. Milillo I., Lattanzi M., Montani G. *Int. J. Mod. Phys. A* **23**, 1278 (2008).
16. Szekeres P., *Ann. Phys.* **64**, 599 (1971).
17. Kaufman A.N., *Ann. Phys.* **18**, 261 (1962).
18. Bombacigno F., Montani G., *Phys. Rev. D* **99**, 064016 (2019).
19. Gong Y., Hou S., *EPJ Web of Conferences* **168**, 01003 (2018).
20. Philippoz L., Jetzer P., *J. Phys.: Conf. Ser.* **840**, 012057 (2017).
21. Abt H.A., *ARA&A* **21**, 343 (1983).
22. Duquennoy A., Mayor M., *A&A* **248**, 485 (1991).
23. Lada C.J., *Astrophys.J.* **640**, L63-L66 (2006).
24. Cathie J.C., Robert D.M., Reid I.N., *Dynamics of Young Star Clusters and Associations* (Springer 2015).
25. Allison M., *Star Clusters and How to Observe Them* (Springer 2006).
26. Rix H.W., Bovy J., *The Astronomy and Astrophysics Review* **21**, 61 (2013).
27. Moore C.J., Cole R.H., Berry C.P.L., *Class. Quantum Grav.* **32**, 015014 (2015).
28. Larson S.L., Hiscock W.A., Hellings R.W., *Phys. Rev. D* **62**, 062001 (2000).
29. Hobbs G. et al., *Class. Quantum Grav.* **27**, 084013 (2010).
30. Verbiest J.P.W. et al., *MNRAS* **458**, 1267 (2016).

# Ring paradigm as quantum gravity

Jan Novák

*Department of physics, Technical University in Liberec,*
*Liberec, Czech Republic*
** E-mail: jan.novak@johnynewman.com*
*www.tul.cz*

Construction of a model of Quantum Gravity, which will be some day in concordance with experiments, is one of the most fascinating tasks we have in modern theoretical physics. There are a plethora of common problems, which must be solved to find a viable candidate for Quantum Gravity. We introduce the concept of the nonlinear graviton and we end with possible experimental evidence for our approach.

*Keywords*: String Theory; Loop Quantum Gravity; Causal Set Approach; Causal Dynamical Triangulation;non-locality; background independence; dimensional reduction; arrow of time; determinism, Wheeler delayed choice experiment; EPR paradox.

## 1. Introduction

We have two main branches in theoretical physics, which describe the Universe at very small and very large scales. These are called General Relativity (GR) and Quantum Mechanics (QM). QM and GR passed many experimental tests and they seem to be fundamental theories of Nature except very high energies.

There are a plethora of reasons, why we should try to build a theory, which would reduce to QM and GR in certain limits. One of the reasons is the unification and the other argument is the existence of singularities. There are many common issues, which must be solved, when we want to unify QM with GR,[1]. We need to make the theory non-local[2] and background independent[3]. We must also deal with dimensional reduction at high energies[4]. We disscused other issues, which are connected to this problematics in[1]. Let us mention, for example, problem of arrow of time[5], determinism or indeterminism of the theory[6], connections with particle physics[7], dark matter[8], dark energy[9], Wheeler delayed choice experiment and EPR paradox[10,11].

Already many attempts have been made during the 20th century to reconcile QM and GR. Let us mention String Theory (ST[3]), Loop Quantum Gravity (LQG[3]), Causal Dynamical Triangulation (CDT[3]), Causal Set Approach (CSA[3]), Group Field Theory,[3] and many others. We present a distinct approach in this article. We claim that, the mathematical apparatus for this theory is hidden in parts of algebraic geometry[12], and algebraic topology.

## 2. First remarks about a new paradigm

We want to show that there is hidden a new philosophical concept that could lead to a new theory in the physical foundations of previously mentioned theories. We will use the knowledge that particles are not point-like objects. But the spacetime will be not continuous for us. So, ST will be for us just a toy model.

Fig. 1. The Universe could be modeled as a finite structure of intersection of Hopf-linked rings, on which are traveling other rings. We suppose that there was one ring in the Big Bang and further rings were created after, which are the trajectories on which could travel particles and fields. Particle and fields could be described as vibrating rings traveling around other rings.

We will not use deep knowledge from LQG. But the investigations are connected to studies of the following set, where the basic objects are topologically circles, and they are Hopf-linked, Figure 1. The formulation of the paradigm, which will follow leads to a mathematical investigation of this set.

Space (not spacetime!) will be discretized to objects, which are called rings. These will be deformable circles, which could be Hopf-linked. We want to argue that points are not the mathematical abstraction, which we will need. When we would discretize to points, the singularities of the type of black holes and initial singularities would be unavoidable. A motivation to this discretization comes from a philosophical branch called finitism, which avoids completed infinities altogether. And this effort to discretize to these rings is well motivated by discrete causal set theory, where we want to put particles, fields, and spacetime on equal footing. Let's stop for a moment on this key point of this paradigm.

We all know very well that GR is perturbatively non-renormalizable. This means that when we try to construct Feynman diagrams and deal with gravitons similarly as in QM the theory diverges.

Another fact from a different area of physics is that the notion of particles is non-unique in quantum field theory in curved background. This serves us as an inspiration for our construction of graviton on the fundamentally nonlinear level.

Our paradigm will come with the existence of a new object, which will be the ring, as we already said. It is a finite deformable circle, which has a finite circumference of 1 Planck length and could stretch from 2 Planck lengths to possibly macroscopic distances. Why we need this object and why it is useful to have particles, fields, and rings? The answer is that we would be possible to put particles, fields, and spacetime on the equal level. Now we will always keep any particle and field together with some trajectory, which will be part of the ring. This will enable us at the end to model particle and fields in a unique description.

The first thing what we need to do is to prove that our discretization is correct according to a deep principle, the so-called holographical principle. It states that the area of any surface $S$ enclosing a volume $V$ measures the information content of the underlying theory in the volume $V$. Our discretization is in concordance with this principle. The number of rings scales as the area of the enclosed volume.

An urgent question comes to our mind. Is this discretization just a mathematical tool or a real physical object? We claim that it is not a mathematical abstraction.

**A keynote about nonlinear graviton**

What we have done so far is an artificial discretization. So what does it have to do with gravity? We need to mention the standard picture of GR and QM. According to GR is a matter moving in 4- dimensional continuous spacetime $(M, g)$ described by topological variety $M$ with a metric $g$. As was quoted by J.A.Wheeler, matter says spacetime how to curve, and spacetime says to matter how to move.

On the other hand, particles jump up on this manifold according to the quantum field theory and are interacting with each other by exchanging the elementary particles of fields. They locally curve the spacetime. These two pictures are incompatible.

What we have in our minds as a picture for a description of gravity are some elementary particles exchanging gravitons. Do we have any experimental evidence that gravitons really exist? It is interesting that detection of the inflationary gravitational wave backgroun, would provide us with direct observational confirmation of the fact that gravity must be quantized. While on the purely theoretical level there is no doubt that it should be so, in particular, if one treats gravity as a low-energy effective QFT, there is so far no evidence for the quantum nature of gravity. Discovery of the inflationary gravitational wave background would be revolutionary. We could always pose a question, is it possible to observe just one graviton in some apparatus as we could observe just one photon? It is only an academic question because one graviton has the too tiny effect to be observable in any possible experiment with interferometers. But we want to indicate that the answer is no. It is not possible to observe the effect of just one graviton, because the picture we have in our mind from the standard QM, is incorrect.

Let's consider Figure 2. We have two particles (we don't say now, how we want to model them) and we switch on the gravity. Then a ring (trajectory or element of space) is created between these two particles, which links it. This link will create an attractive force between these two particles. We could think about it as a spring between them, although it does not have properties of a linear spring. It simultaneously creates a trajectory for exchanging other fields and particles, so a creation of other three interactions.

One question immediately comes to our minds. Is it not a contradiction with the basic principles of QM? We do not say that a particle has a trajectory in QM. We speak about probabilities. But this trajectory creates in Planck time from 2 Planck

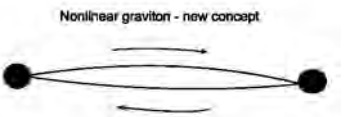



Fig. 2. The old picture of particles exchanging gravitons is incorrect according to RP (Ring Paradigm). When we switch on the extragravity (which is the unification of all 4 interations ), an element of space - ring - is created between these two particles. Other particles could then travel on this trajectories. The important thing is that these trajectories are created in Planck time.

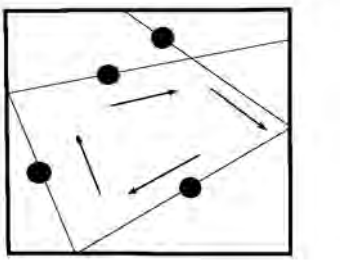


Fig. 3. There is a particle knocked in a box at the figure. It jumps from one trajectory to other trajectory. The trajectories are parts of rings, which are created between the walls of the box and they are changing in Planck time.

lengths to possibly arbitrary distances, and it is changing in Planck time. So, let's consider some particle knocked in the box as in Figure 3.

It is a well-known fact that we could not localize it. There are randomly created trajectories in the box, and the particle jumps from one trajectory to other as we depicted it.

We now just move forward. What we want to draw ultimately is a more advanced version of a Feynman diagram as in Figure 4.

There are two trajectories with opposite orientation, then a trajectory is created in Planck time between these two trajectories. Particles (vibrating rings) are coming on the first ring, then a vibrating ring is traveling in-between, and two particles originate on the second ring.

It is immediately clear that we will need to make this picture of vibrating rings traveling around other rings dynamical. We will need the process of creating a ring, the process of absorption of rings and breaking of rings, Figure 5. We wish to model the false vacuum (creation and absorption) at the end and also accelerated expansion (breaking). We said that we could talk about the ring as a very strange spring between the two particles. We can stretch it, the force is increasing first, and then it is decreasing as $\frac{1}{r^2}$ with the distance $r$ between these two rings.

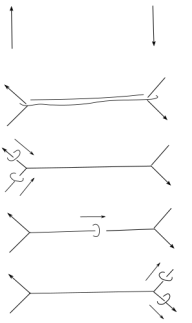


Fig. 4. We need to draw an advanced version of the Feynman diagram in RP. We switch on extragravity at the beginning, then a ring is created in Planck time between these two rings. Particles collide on these rings on the left, then a particle is propagating on the ring between these two rings, and finally, two new particles are created on these rings on the right.

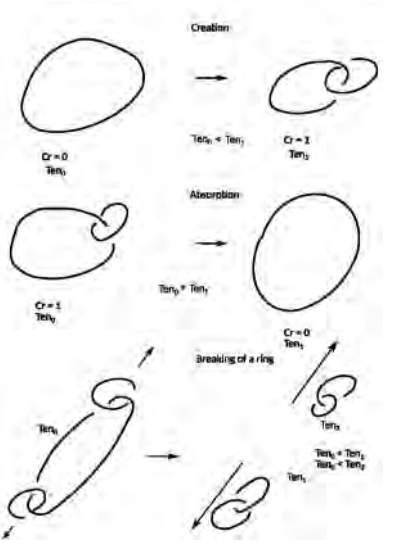



Fig. 5. The basic processes with rings are creation, absorption and breaking of a ring.

## Conclusion

We introduced a concept of nonlinear graviton in this short article. We claim that graviton is not a particle. It is a ring (element of space, trajectory), which connects other rings, and which produces force between them.

One immediate result is the limit velocity, by which we are able to send information in our Universe. Particles and fields of the standard model could travel only with maximal velocity of light in vacuum. If we will find other particles outside

standard model, they will still maximally travel with velocity of light. But the ring is created in Planck time on possibly arbitrary distances in the Universe. Could we send some information by it? Could it have some technical applications? These are the important questions, on which we need to find answers in future research.

## References


1. J. Novák , RT-paradigm as Quantum Gravity, 2019, sent to Foundations of physics
2. J. Henson, Constructing an interval of Minkowski space from a causal set Class.Quant.Grav.**23**: L29-L35, 2006 arXiv: 0601069 [gr-qc]
3. D. Oriti 2009 Approaches to Quantum Gravity (Cambridge University Press), 2009
4. S. Carlip, Dimensional reduction in causal set gravity Class. Quantum Grav. **32**, 232001, 2015, arXiv: 1506.08775[gr-qc]
5. D. Lazarovici, P. Reichert, Arrow of time without past hypothesis, History and Philosophy of Physics, arXiv: 1809.04646v1
6. G. 't Hooft, Determinism and dissipation in quantum gravity, SPIN-2000/07, arXiv: 0003005v2 [hep-th]
7. S. O. Bilson-Thompson, F. Markopoulou, L. Smolin, Quantum gravity and the standard model, Class.Quant.Grav. 24:3975-3994, 2007, arXiv: 0603022 [hep-th]
8. K. Garrett and G. Duda, Dark matter: a primer, Adv. Astron.: 968283, 2011, arXiv: 1006.2483 [hep-ph]
9. T. Abbott et al., The dark energy survey: more than dark energy - an overview, Mon. Not. Roy. Astron. Soc. 460, no.2, 1270 - 1299, 2016
10. H. D. Zeh, On the intepretation of measurement in quantum theory, Foundations of Physics, Vol 1, No. 1, 1970
11. V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch, Experimental realization of Wheeler's delayed choice GedankenExperiment, Science 315, 5814, 966-8, 2007
12. A. Postnikov, Total positivity, Grassmannians, and networks. Preprint, http://math.mit.edu/ apost/papers/tpgrass.pdf, arXiv: 0609764v1 [math.CO]

# Perfect fluid geometries in Rastall's cosmology

M. Palese* and M. Capone

*Department of Mathematics, University of Torino,*
*via C. Alberto 10 Torino, I-10123 Italy*
*marcella.palese@unito.it*

The equation of state of an ultrarelativistic perfect fluid is obtained as a necessary condition for a perfect fluid space-time in Rastall's cosmology.

*Keywords*: perfect fluid space-time; Rastall's cosmology; ultrarelativistic fluid

## 1. Introduction

It is well known the importance of a perfect fluid description of gravitational sources in General Relativity[3]. It has been also stressed the relevance of compatibility of perfect-fluid solutions with modified or extended theories of gravity[1].

In this note we point out some cosmological features emerging from the request that the Rastall model[11] describe a perfect fluid geometry of space-time.

We shortly recall some features of the model introduced by Rastall in 1972; we refer to this paper for the relevant issues and stress that the chosen signature for the underling Lorentzian manifold will be $(-,+,+,+)$, in accordance with such a source.

Let then $Lor\boldsymbol{X}$ be the open submanifold of the subbundle of symmetric tensors constituted by regular metrics with the above choice of the Lorentzian signature. An induced fibered chart on $J^2Lor\boldsymbol{X}$ has local coordinates $(x^\alpha, g_{\alpha\beta}, g_{\alpha\beta,\gamma}, g_{\alpha\beta,\gamma\nu})$ and we can define other local coordinates $g^{\mu\nu}$ and so on by the relation $g_{\alpha\mu}g^{\mu\nu} = \delta^\nu_\alpha$.

As well known Einstein equations read

$$G_{\mu\nu} \doteq R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa_{GR}T_{\mu\nu}\,,$$

where $\kappa_{GR} = \frac{8\pi G}{c^4}$; the Ricci tensor is obtained from a metric connection, so that $R_{\mu\nu} = R_{\mu\nu}(j^2g)$ and the scalar curvature $R$ has to be intended as $R(j^2g) = g^{\alpha\beta}R_{\alpha\beta}(j^2g)$. These equations naturally imply the conservation of energy-momentum tensor as a consequence of the Bianchi identities.

The Rastall model for gravity coupled with matter governed by an energy-momentum tensor satisfying (with $\lambda$ a suitable non-null dimensional constant)

$$\hat{T}^\mu_{\nu;\mu} = \lambda R_{,\nu}$$

where the comma denotes the partial derivative, and by field equations

$$R_{\mu\nu} - \frac{1}{2}\left(1 - 2\kappa_r\lambda\right)Rg_{\mu\nu} = \kappa_r\hat{T}_{\mu\nu}\,, \tag{1}$$

*Corresponding Author

where $\kappa_r$ is a dimensional constant to be determined in order to give the right Poisson equation in the static weak-field limit. This model introduces an energy-momentum tensor fulfilling the requirement $\nabla_\mu \hat{T}^\mu_\nu \doteq \hat{T}^\mu_{\nu;\mu} \neq 0$ without violating the Bianchi identities; for a 4-vector, say $a_\nu$, $\hat{T}^\mu_{\nu;\mu} = a_\nu$ and $a_\nu \neq 0$ on curved spacetime, but $a_\mu = 0$ on flat spacetime (in agreement with Special Relativity). Of course, in general, $\kappa_r \neq \kappa_{GR}$. Taking the trace of Eqs.(1) gives us the so-called structural or master equation $(4\kappa_r\lambda - 1)\, R = \kappa_r \hat{T}$ $(\kappa_r\lambda \neq \frac{1}{4})$.

By this relation the Rastall equations (1) can be recast in the form of Einstein equations. Indeed, one can immediately write $G_{\mu\nu} = \kappa_r S_{\mu\nu}$, where $S_{\mu\nu} \doteq \hat{T}_{\mu\nu} - \frac{\kappa_r\lambda}{4\kappa_r\lambda - 1} g_{\mu\nu}\hat{T}$. By construction, this new energy-momentum tensor is conserved, *i.e.* $S^\mu_{\nu;\mu} = 0$. Furthermore, (if $u_\mu$ is a 4-velocity with $g^{\mu\nu}u_\mu u_\nu = -1$) by assuming $\hat{T}_{\mu\nu} = -(\rho + p)u_\mu u_\nu - p g_{\mu\nu}$, *i.e.* the source is a perfect isentropic fluid with energy density $\rho$ and pressure $p$, we can explicitly work out an expression for $S_{\mu\nu}$. The source turns out to be still a perfect fluid, provided we redefine its energy density and pressure.[11]

## 2. Perfect fluid space-times in Rastall's cosmology

Rastall's cosmology has been recently object of attention in scientific community; some papers are concerned with the querelle whether Rastall's stress-energy tensor corresponds to an artificially isolated part of the physical conserved stress-energy or not[2,7,10,15], some others are concerned with the derivation of a related variational principle (*i.e.* with the question how to get a - eventually modified - Rastall model which could be derived by a Lagrangian and physically different from Einstein gravity), see *e.g.* Refs. 12, 14, 13.

We study this model from the point of view of the compatibility of its geometry with a perfect fluid description of space-time. Indeed, as a gravity model it provides a non minimal coupling between metric and matter and we try to understand how this is related with this compatibility requirement.

Indeed it is well known that by suitably introducing a matter density $\mu$, the energy density of space-time $\bar{\rho}$ can be defined as a function of $\mu$ satisfying $\bar{p} = \mu\bar{\rho}' - \bar{\rho}$ where the prime denotes derivation with respect to $\mu$; (this is equivalent with asking whether we can derive the tensor $R_{\mu\nu} \equiv \tilde{T}_{\mu\nu}$ from a variational principle, *i.e.* with looking for a Lagrangian density depending on $\rho(\mu)$ from which $\tilde{T}_{\mu\nu}$ can be derived variationally; such kind of space-times appear *e.g.* in the dual Lagrangian description of the Ricci tensor for barotropic perfect relativistic fluids[6], see also Refs. 4, 5, 8).

We therefore have

$$\tilde{T}_{\alpha\beta} = \bar{\rho}(\mu) g_{\alpha\beta} - \mu\bar{\rho}'(\mu)[g_{\alpha\beta} + u_\alpha u_\beta]\,,$$

with trace $\tilde{T} = 4\bar{\rho} - 3\mu\bar{\rho}'$; moreover $\bar{p} = \mu^2 \frac{d}{d\mu}(\frac{\bar{\rho}}{\mu}) = \mu\bar{\rho}' - \bar{\rho}$.

### 2.1. *Ultrarelativistic equation of state from a perfect fluid geometry*

By requiring the perfect fluid Ricci tensor to be *also* a Rastall's geometry the following identity should hold true.

$$\begin{aligned}\tilde{T}_{\alpha\beta} &\equiv \hat{T}_{\alpha\beta} - \frac{1}{2}\frac{2\tilde{\lambda}-1}{4\tilde{\lambda}-1} g_{\alpha\beta}\hat{T} \\ &= [\rho - \mu\rho' - \frac{1}{2}\frac{2\tilde{\lambda}-1}{4\tilde{\lambda}-1}(4\rho - 3\mu\rho')]g_{\alpha\beta} - \mu\rho' u_\alpha u_\beta \\ &\equiv (\bar{\rho} - \mu\bar{\rho}')g_{\alpha\beta} - \mu\bar{\rho}' u_\alpha u_\beta = -\bar{p}g_{\alpha\beta} - (\bar{\rho}+\bar{p})u_\alpha u_\beta \,. \end{aligned} \tag{2}$$

The above identity leads to

$$\rho' = \bar{\rho}' \,.$$

By inserting this and deriving with respect to $\mu$ we get

$$\frac{2\tilde{\lambda}-1}{4\tilde{\lambda}-1}(4\rho' - 3(\mu\rho')') = 0 \,.$$

Let us now exclude the case $2\tilde{\lambda} - 1 = 0$. We see that a necessary condition for the Ricci tensor of a nontrivial Rastall's geometry of matter to be a perfect fluid is

$$4\rho' - 3(\mu\rho')' = 0 \,.$$

*i.e.*

$$\rho' - 3p' = 0 \,.$$

the equation of state is therefore

$$p = \frac{1}{3}\rho + const.(\mu) \,,$$

*i.e.* if we denote $\alpha$ the integration constant (a constant function of $\mu$) (uniquely determined by initial data)

$$p - \alpha = \frac{1}{3}\rho$$

In particular it is easy to get that $-3\alpha = \rho - 3p = \tilde{T}$, *i.e.* $\alpha = -\frac{1}{3}\tilde{T}$

By redefining the pressure $\pi = p - \alpha = p + \frac{1}{3}\tilde{T}$ the equation of state is the one of an ultrarelativistic perfect fluid with pressure $\pi$ and energy density $\rho$.

$$\pi = \frac{1}{3}\rho \iff p = \frac{1}{3}(\rho - \tilde{T})$$

*i.e.* with a rescaled pressure or equivalently a *energy density rescaled by the trace of the energy-momentum tensor*. This relation gives us a model with constant (with respect to the matter density $\mu$) energy-tensor trace $\tilde{T}$ and consequently constant (with respect to $\mu$) scalar curvature.

**Acknowledgements**

The first author (M.P.) is supported by the Department of Mathematics - University of Torino's project *Strutture geometriche e algebriche in Fisica Matematica e applicazioni*, GNSAGA of INdAM, and the Cost Action *CA17139 - European Topology Interdisciplinary Action*. The second author (M.C.) has been supported by the Regione Piemonte research grant *Aspetti Geometrici delle Teorie Alternative della Gravitazione*. Thanks are due to M. Ferraris for useful discussions.

**References**

1. S. Capozziello, C.A. Mantica, L.G. Molinari: Cosmological perfect-fluids in f(R) gravity, *Int. J. Geom. Methods Mod. Phys.* **16** (1) (2019) 1950008 .
2. F. Darabi, H. Moradpour, I. Licata, Y. Heydarzade, C. Corda: Einstein and Rastall theories of gravitation in comparison, *Eur. Phys. J. C* (2018) 78–25.
3. L.P. Eisenhart: Space-time Continua of Perfect Fluids in General Relativity, *Trans. American Math. Soc.* **26** (1924) 205–220.
4. M. Ferraris, M. Francaviglia: Energy-momentum tensors and stress tensors in geometric field theories, *J. Math. Phys.* **26** (6) (1985)1243–1252.
5. M. Ferraris, M. Francaviglia, and I. Volovich: The universality of vacuum Einstein equations with cosmological constant, *Class. Quant. Grav.* **11** (6) (1994) 1505.
6. M. Ferraris, J. Kijowski: On the equivalence between the purely metric and the purely affine theories of gravitation, *Rend. Sem. Mat. Univ. Politec. Torino* **41** (3) (1984) 169–201.
7. S. Hansraj, A. Banerjee, P. Channuie: Impact of the Rastall parameter on perfect fluid spheres, *Annals of Physics* **400** (2019) 320–345.
8. D. Krupka: Variational principles for energy-momentum tensors, *Rep. Math. Phys.***49** (2/3) (2002) 259–268.
9. H Moradpour, Y. Heydarzade, F. Darabi, I.G. Salaco: A generalization to the Rastall theory and cosmic eras, *Eur. Phys. J. C* (2017) 77: 259.
10. A. M. Oliveira, H. E. S. Velten, J. C. Fabris, L. Casarini: Neutron Stars in Rastall Gravity, *Phys. Rev. D* **92** (2015) 044020.
11. P. Rastall: Generalization of the Einstein Theory, *Phys. Rev. D* **6** (12) (1972) 3357.
12. L.L. Smalley: Rastall's and related theories are conservative gravitational theories although physically inequivalent to general relativity, *J. Phys. A: Math. Gen* **16** (1983) 2179–2185.
13. L.L. Smalley: Variational Principle for a Prototype Rastall Theory of Gravitation, *Il Nuovo Cimento B* **80** (1) (1984) 42–47.
14. L.L. Smalley: A variational principle for a spacetime-dependent cosmological function, *Class. Quant. Grav.* **10** (1993) 1179–1184.
15. M. Visser: Rastall gravity is equivalent to Einstein gravity, *Phys. Lett.* **B782** (2018) 83–86.

# How to use unimodular quantum cosmology for the prediction of a late-time classical universe?

Natascha Riahi

*University of Vienna, Faculty of Physics, Gravitational Physics*
*Boltzmanng. 5, 1090 Vienna, Austria*
*E-mail: natascha.riahi@gmx.at*

Unimodular quantum cosmology admits wavepacket solutions that evolve according to a kind of Schrödinger equation. Though this theory is equivalent to general relativity on the classical level, its canonical structure is different and the problem of time does not occur. We present an Ehrenfest theorem for the long term evolution of the expectation value of the scale factor for a spatially flat Friedmann universe with a scalar field. We find that the classical and the quantum behavior in the asymptotic future coincide for the special case of a massless scalar field. We examine the general behavior of uncertainties in order to single out models that can lead to a classical universe.

## 1. Introduction

The canonical quantization of general relativity leads to the so-called problem of time (see[1] and references therein). In most non-perturbative approaches of quantum gravity time has disappeared from the theory and is seen as an artifact of the classical limit. Here we investigate quantum cosmology in the framework of unimodular gravity. This theory is practically equivalent to general relativity at the classical level, but since it has a different canonical structure time does not disappear from the quantum theory[2] and the quantization of unimodular theory yields different results compared to the quantization of general relativity[3]. In[4] we constructed a class of unitarily evolving solutions with a negative expectation value of the Hamiltonian for the special case of a spatially flat universe with a massless scalar field. Investigating a special example, we found that the classical and quantum dynamics of the scale factor coincide for the asymptotic future (though with significant spread). Here we compare the expectation value of the scale factor to the evolution of its classical counterpart for solutions of the spatially flat Friedmann universe with an arbitrary scalar field which yields an Ehrenfest theorem for the late time behavior. We examine the evolution of the uncertainties in order to single out models that can lead to a classical universe in the asymptotic future.

## 2. About Unimodular Gravity

We start with the Einstein Hilbert action (1)

$$S_{EH} = \frac{1}{2\kappa}\int_{\mathcal{M}} d^4x\,\sqrt{-g}\,(R-2\Lambda) - \frac{1}{\kappa}\int_{\partial\mathcal{M}} d^3x\,\sqrt{h}\,K\,, \tag{1}$$

where

$$\kappa = \frac{8\pi G}{c^4}$$

contains the velocity of light $c$ and the gravitational constant $G$. We also take into account the matter action $S_m$ that describes the fields. If we vary the action $S = S_m + S_{EH}$ with respect to the metric $g_{\mu\nu}$ under the restriction $-g = 1$, we obtain Einsteins equations with an arbitrary additional constant $\Lambda$ , that can be identified with the cosmological constant of general relativity ([2]).

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = \kappa\, T_{\mu\nu} - \Lambda\, g_{\mu\nu}, \quad \sqrt{-g} - 1 = 0\,. \tag{2}$$

This theory is called unimodular gravity. Any solution of unimodular gravity (2) is also a solution of general relativity for a specific cosmological constant and vice versa. The only difference between the two theories is, that $\Lambda$ is a natural constant in general relativity while it is a conserved quantity in unimodular gravity. But since in both theories the cosmological constant can not vary over the whole universe, we would have to investigate different universes to determine if solutions with different $\Lambda$ exist (unimodular theory) or if $\Lambda$ is a "true" natural constant. So the two theories are practically indistinguishable on the classical level.

## 3. The Unimodular Hamiltonian of a spatially flat Friedmann Universe

The metric of a homogeneous and isotropic spacetime (Friedmann universe)

$$ds^2 = -N^2(t) c^2 dt^2 + a^2(t) d\Omega_3^2 \tag{3}$$

is characterized by the lapse function N(t) and the scale factor $a(t)$. If the spatial curvature is zero, $d\Omega_3^2$ is the line element of three-dimensional flat space.

Inserting the metric into the Einstein-Hilbert action (1) with $\Lambda = 0$ yields ([1])

$$S_{EH} = \frac{3}{\kappa} \int dt\, N \left( -\frac{\dot{a}^2 a}{c^2 N^2} \right) v_0\,,$$

where $v_0$ is is the volume of the spacelike slices according to (3).

The action of a scalar field in a Friedmann universe (3) reads

$$S_m = \int dt\, N a^3 \left( \frac{\dot{\phi}^2}{2N^2 c^2} - V(\phi) \right) v_0\,. \tag{4}$$

Using the unimodular condition for the lapse function $N = a^{-3}$, we find for the Hamiltonian ([4]) of the unimodular theory

$$H_{uni} = \frac{c^2}{2} \frac{p_\phi^2}{a^6} - \frac{c^2}{4\epsilon} \frac{p_a^2}{a^4} + V(\phi) \tag{5}$$

The Hamiltonian is a conserved quantity and not a constraint as in general relativity

The canonical quantization of this Hamiltonian yields

$$\hat{p}_a = -i\hbar \frac{\partial}{\partial a}\,, \quad \hat{p}_\phi = -i\hbar \frac{\partial}{\partial \phi}\,, \tag{6}$$

$$\widehat{H} = \frac{\hbar^2 c^2}{4\epsilon}\frac{1}{a^5}\frac{\partial}{\partial a}a\frac{\partial}{\partial a} - \frac{\hbar^2 c^2}{2}\frac{1}{a^6}\frac{\partial^2}{\partial \phi^2} + V(\phi)\,. \tag{7}$$

Here we have chosen the factor ordering that gives the part of the Hamiltonian that is quadratic in the momenta the form of a Laplace Beltrami operator ([1]).

The evolution of the wavefunction $\psi(a,\phi,t)$ is determined by

$$\widehat{H}\psi = i\hbar\frac{\partial}{\partial t}\psi\,. \tag{8}$$

The Hamiltonian is symmetric with respect to the inner product defined by the measure $a^5 da d\phi$, where $a \in (0,\infty)$ and $\phi \in (-\infty,\infty)$.

Applying the coordinate transformations

$$A = a^3/3 \qquad B = \frac{3}{\sqrt{2\epsilon}}\phi \tag{9a}$$

and

$$u = Ae^{-B} \qquad v = Ae^{B}\,, \tag{9b}$$

we obtain for the Hamilton operator

$$\widehat{H} = \frac{\hbar^2 c^2}{\epsilon}\frac{\partial^2}{\partial u \partial v} + V\left(\frac{u}{v}\right) \tag{10}$$

The volume element is given by $dudv$ and $u \in (0,\infty)$, $v \in (0,\infty)$.

The classical unimodular Hamiltonian then reads

$$H = -\frac{c^2}{\epsilon}p_u\, p_v + V\left(\frac{u}{v}\right)\,.$$

The Laplace-Beltrami factor ordering ensures that the quantization of the Hamiltonian commutes with the coordinate transformation if we understand classical transformations as canonical point transformations (see also[5]).

## 4. General properties of the time evolution

In[4] we have derived a class of unitarily evolving wavepacket solutions of (8) for the special case of a massless scalar field ($V = 0$). We found that these solutions fulfill in the late phase of time evolution

$$\lim_{t\to\infty}\,\psi(0,v,t) = \lim_{t\to\infty}\,\psi(u,0,t) = 0\,. \tag{11}$$

Now we assume that a wavepacket solution for a general scalar field can be found and that it also fulfills (11). We investigate the physical behaviour of these solutions compared to the evolution of the classical quantities.

As in ordinary quantum mechanics, the evolution of the expectation values of an observable $\widehat{O}$ with respect to a solution $\psi(u,v,t)$ of the Schrödinger equation (8) is given by

$$\frac{d}{dt}\langle\psi|\widehat{O}|\psi\rangle = -\frac{i}{\hbar}\left(\langle\psi|\widehat{O}\widehat{H}|\psi\rangle - \langle\widehat{H}\psi|\widehat{O}\psi\rangle\right)\,. \tag{12}$$

This implies the equation

$$\frac{d}{dt}\langle\psi|\widehat{O}|\psi\rangle = -\frac{i}{\hbar}\left\langle\psi\left[\widehat{O},\widehat{H}\right]\psi\right\rangle\,, \tag{13}$$

only if $\widehat{O}\psi$ obeys

$$\langle\widehat{H}\psi|\widehat{O}\psi\rangle \stackrel{!}{=} \langle\psi|\widehat{H}\widehat{O}|\psi\rangle\,. \tag{14}$$

In the case of the Hamiltonian(10), this condition reads

$$\int_0^\infty \psi^*(u,v,t)\frac{\partial}{\partial v}\widehat{O}\psi(u,v,t)dv\,\Big|_{u=0} - \int_0^\infty\left(\frac{\partial}{\partial u}\psi^*(u,v,t)\right)\widehat{O}\psi(u,v,t)du\,\Big|_{v=0} = 0\,. \tag{15}$$

We find that this condition is fulfilled in the limit $t\to\infty$ for wavepackets with the property (11). This means that the time evolution that we derive by the application of (13), gives the correct result for the limit $t\to\infty$, and is approximately valid in the late phase of time evolution.

We find for the variable $uv$, related to the scalefactor by $A^2 = uv = \frac{a^6}{9}$

$$\lim_{t\to\infty}\frac{d^2}{dt^2}\langle A^2\rangle = \frac{c^2}{\epsilon}\left\langle -2\,\widehat{H} + V + u\frac{\partial V}{\partial u} + v\frac{\partial V}{\partial v}\right\rangle\,, \tag{16a}$$

whereas the classical time evolution reads

$$\lim_{t\to\infty}\frac{d^2}{dt^2}A^2 = \frac{c^2}{\epsilon}\left(-2\,H + V + u\frac{\partial V}{\partial u} + v\frac{\partial V}{\partial v}\right)\,. \tag{16b}$$

So we see that in the special case of a massless scalar field the late time behavior of the classical scalefactor and its expectation value according to unimodular quantum cosmology are the same. In general (16) represents the Ehrenfest theorem for unimodular quantum cosmology.

The result for the uncertainty $\Delta(A^2)$ in the special case of a massless scalar field reads

$$\lim_{t\to\infty}\frac{d^4}{dt^4}(\Delta(A^2))^2 = \frac{24c^2}{\epsilon^2}(\Delta H)^2\,,$$

which implies that the uncertainty $\Delta(A^2)$ is monotonically growing with $t^2$ in the late phase of time evolution.

## 5. Evolution of uncertainties for general matter models

All dynamical equations in this section are derived under the assumption (11) using (13) and are therefore valid for late times. We find for the dynamics of the expectation values

$$\frac{d}{dt}\langle v\rangle = -\mu\,\langle\widehat{p_u}\rangle\,,\ \frac{d}{dt}\langle u\rangle = -\mu\,\langle\widehat{p_v}\rangle\,,\ \frac{d}{dt}\langle\widehat{p_u}\rangle = -\left\langle\frac{\partial V}{\partial v}\right\rangle\,,\ \frac{d}{dt}\langle\widehat{p_v}\rangle = -\left\langle\frac{\partial V}{\partial u}\right\rangle\,, \tag{17}$$

where $\mu = c^2/\epsilon = 2\pi G/(3c^2)$.

If $\frac{\partial V}{\partial v}, \frac{\partial V}{\partial u}$ fulfill

$$\langle F(u,v)\rangle \approx F(\langle u\rangle, \langle v\rangle) \tag{18}$$

the equations constitute a closed system of ordinary differential equations for the expectation values, which coincide with the classical evolution equations according to (7).

The four observables $\hat{u}, \hat{v}, \widehat{p_v}, \widehat{p_u}$ give rise to ten uncertainties $(\Delta u)^2, (\Delta v)^2$, $(\Delta p_u)^2, \Delta(p_v)^2$, $\Delta(u,v), \Delta(u,p_u)$, $\Delta(v,p_v)$, $\Delta(u,p_v)$, $\Delta(v,p_u)$, $\Delta(p_u,p_v)$, where the mixed uncertainties of two observables are defined by

$$\Delta(v_1,v_2) = \frac{1}{2}\left\langle \widehat{v_1}\widehat{v_2} + \widehat{v_2}\widehat{v_1} - 2\left\langle \widehat{v_1}\right\rangle \left\langle \widehat{v_1}\right\rangle \right\rangle$$

The application of (13) yields the evolution equations that contain expectation values of partial derivatives of $V$ as for instance

$$\frac{d}{dt}\Delta(u,p_v) = -\mu\Delta(p_v)^2 - \left\langle \frac{\partial V}{\partial v}u \right\rangle + \left\langle \frac{\partial V}{\partial v} \right\rangle \langle u\rangle \ . \tag{19}$$

Expanding the occurring derivatives of $V$ around the expectation values given by (17)and neglecting all stochastic moments of higher than second order yields a closed system of linear differential equations for the variances (quadratic uncertainties):

$$\begin{aligned}
&\frac{d}{dt}(\Delta u)^2 = -2\mu(\Delta(u,p_v)\,, \quad \frac{d}{dt}(\Delta v)^2 = -2\mu(\Delta(v,p_u))\,, \\
&\frac{d}{dt}\Delta(u,p_v) = -\mu\Delta(p_v)^2 - V_{22}\Delta(u,v) - V_{12}(\Delta u)^2\,, \\
&\frac{d}{dt}\Delta(v,p_u) = -\mu\Delta(p_u)^2 - V_{11}\Delta(u,v) - V_{12}(\Delta v)^2\,, \\
&\frac{d}{dt}\Delta(u,v) = -\mu\Delta(v,p_v) - \mu\Delta(u,p_u)\,, \\
&\frac{d}{dt}(p_u)^2 = -2V_{11}(\Delta(u,p_u) - 2V_{12}(\Delta(v,p_u)\,, \\
&\frac{d}{dt}(p_v)^2 = -2V_{22}(\Delta(v,p_v) - 2V_{12}(\Delta(u,p_v)\,, \\
&\frac{d}{dt}\Delta(p_u,p_v) = -V_{11}\Delta(u,p_v) - V_{12}\Delta(p_v,v) - V_{22}\Delta(v,p_u) - V_{12}\Delta(p_u,u)\,, \\
&\frac{d}{dt}\Delta(u,p_u) = -\mu\Delta(p_u,p_v) - V_{11}(\Delta u)^2 - V_{12}\Delta(u,v)\,, \\
&\frac{d}{dt}\Delta(v,p_v) = -\mu\Delta(p_u,p_v) - V_{22}(\Delta v)^2 - V_{12}\Delta(u,v)\,,
\end{aligned} \tag{20}$$

where we have chosen the abbreviations

$$V_{11} = \frac{\partial^2 V}{\partial v^2}\,, \quad V_{22} = \frac{\partial^2 V}{\partial u^2} \quad V_{12} = \frac{\partial^2 V}{\partial v \partial u}\,.$$

Given that all contributions of higher moments can be neglected, all uncertainties remain bounded if the system (20) is stable. This is especially the case if $V_{11}, V_{22}, V_{12}$

converge to fixed values for late times and fulfill

$$|V_{12}| < \sqrt{V_{11}V_{22}}\,, \quad V_{12} < 0$$

since then all characteristic roots of (20) have zero real part and are of simple type (see for instance[6]).

If we neglect deviations from the classical dynamics which would be of the order of magnitudes of variances (see (18)) we can evaluate (20) inserting the classical solutions for $\langle u\rangle\,, \langle v\rangle\,, \langle\widehat{p_u}\rangle\,, \langle\widehat{p_v}\rangle$ and verify if the initial assumption of small uncertainties (variances) is justified.

## References


1. Kiefer, C.: Quantum gravity, Oxford university press (2012).
2. Henneaux, M., Teitelboim,C. (1989): The cosmological constant and general covariance, Phys.Lett.B 222/2, 195.
3. Unruh, W.G. (1989): Unimodular theory of canonical quantum gravity, Phys.Rev.D 40/4, 1048–1051.
4. Riahi, N. (2017): Wavepacket evolution in unimodular quantum cosmology, Galaxies 2018, 6(1), 8.
5. Kuchar, K. (1986): Hamiltonian dynamics of gauge systems, Phys. Rev. D 34/10 3031-3043, 3043–3057.
6. Coppel, W.A.: Stability and Asymptotic Behaviour of Differential Equations, Heath and Company, Boston(1965), p. 56.

# Newton gauge cosmological perturbations for static spherically symmetric modifications of the de Sitter metric

Antonio Enea Romano[1,2],Camilo Santa Vélez[2]

[1]*King's College London, Strand, London, WC2R 2LS, United Kingdom*
[2]*Instituto de Fisica, Universidad de Antioquia, A.A.1226, Medellin, Colombia*

Static coordinates can be convenient to solve the vacuum Einstein's equations in presence of spherical symmetry, but for cosmological applications comoving coordinates are more suitable to describe an expanding Universe, especially in the framework of cosmological perturbation theory (CPT). Using CPT we develop a general method to transform in the weak field limit any static spherically symmetric (SSS) metric from static coordinates to the Newton gauge. We apply the method to the Schwarzschild de Sitter (SDS) metric to a SSS solution of the Brans-Dicke theory, confirming the results obtained independently by solving the perturbation equations in the Newton gauge. We then consider different classes of SSS metrics obtained by modifying the SDS metric and compute the corresponding Bardeen potentials. Using the gauge invariance of the Bardeen potentials we obtain a gauge invariant definition of the turn around radius, checking it is consistent with the result obtained in static coordinates for the SDS metric and for other SSS metrics.

## 1. Introduction

Spherically symmetric solutions are important because they allow to compute different observables within the solar system or to study the stability of large scale structure[1,2]. They can consequently provide a very important tool to test modified gravity theories[3]. In general relativity the most general spherically symmetric vacuum solution of the Einstein's field equations with a cosmological constant, the Schwarzschild-De-Sitter solution, can be written in static coordinates[4]. This is a consequence of the Birkhoff's theorem, but for a general modified theory of gravity such a coordinate system may not be possible. For a modified gravity theory in fact the most general spherically symmetric solution may not be expressed in static coordinates, but only a subclass of the general solutions. This is the consequence of the fact that the Birkhoff's theorem may not hold in modified gravity theories[5].

In cosmology comoving coordinates are more convenient to describe an expanding Universe, especially in the framework of cosmological perturbation theory (CPT), and it is therefore useful to find the coordinate transformation from static to comoving coordinates. In this paper we develop a general method to transform in the weak field limit any spherically symmetric metric from static coordinates to the Newton gauge and apply it to different known and new cases. We first apply the method to the Schwarzschild de Sitter (SDS) metric, confirming it does indeed allow to compute correctly the Newtonian potential due to a point mass in an expanding Universe. We then compute the Newtonian potentials for a spherically symmetric solution of the Brans-Dicke theory in static coordinates, confirming the results obtained independently by solving the Brans-Dicke field equations in the Newton gauge[3]. Using the gauge invariance of the Bardeen potentials we obtain a gauge

invariant definition of the turn around radius, checking it is consistent with the result obtained in static coordinates for the SDS metric and for other static spherically symmetric (SSS) metrics. We finally consider different generalized classes of SSS metrics and using the same method we transform them to the Newton gauge.

## 2. Static Coordinates

Assuming a general spherically symmetric metric ansatz of the type

$$ds^2 = F(t,R)dt^2 - H(t,R)dR^2 - R^2 d\Omega^2, \tag{1}$$

one of the Einstein's equations gives $\partial_t f = 0$, implying the existence of the well known static coordinates solution

$$F(t,R) = H(t,R)^{-1} = \left(1 - \frac{2m}{R} - \frac{R^2}{l^2}\right), \tag{2}$$

where have defined $l^2 = 3/\Lambda$. Substituting the same general ansatz in the field equation of a different gravity theory the equation $\partial_t f = 0$ may not hold anymore, and a general spherically symmetric solution may not be written in static coordinates anymore. The de-Sitter metric, i.e. the $m = 0$ limit of the SDS metric, can also be written in isotropic coordinates

$$R = e^{t/l} r \tag{3}$$

$$t = \tilde{t} - \frac{l}{2}\log(r^2 e^{2t/l} - l^2) \tag{4}$$

$$ds^2 = d\tilde{t}^2 - a^2(t)(dr^2 + r^2 d\Omega^2), \tag{5}$$

where $a(\tilde{t}) = e^{\tilde{t}/l}$. These coordinates are called comoving coordinates in cosmology because they are interpreted as the coordinates of the observer comoving with the Hubble flow, and for this reason we will use this terminology in the rest of this paper. We will use a similar coordinate transformation from static to comoving coordinates to re-write the SDS metric far from the Schwarzschild radius, i.e. for $m \ll r$ in terms of cosmological perturbations respect to the FRW background.

## 3. Cosmological perturbations in a spherically symmetric space and their relation with static coordinates

For cosmological applications it is useful to re-write the SSS metrics in comoving coordinates. The the most general scalar perturbations respect to the flat FRW background can be written as[6]

$$ds^2 = a^2\Big\{(1+2\psi)d\tau^2 - 2\partial_i\omega d\tau dx^i - [(1-2\phi)\delta_{ij} + D_{ij}\chi]\, dx^i dx^j\Big\}, \tag{6}$$

where $D_{ij} = \partial_i\partial_j - \frac{1}{3}\delta_{ij}\nabla^2$.

Assuming spherical symmetry the metric can be written as[7]

$$ds^2 = a^2\Big[(1+2\psi)d\tau^2 - \left(1-2\phi+\frac{2}{3}\mathcal{E}\right)dr^2 - 2\omega' d\tau dr - \left(1-2\phi-\frac{1}{3}\mathcal{E}\right)r^2 d\Omega^2\Big], \tag{7}$$

where

$$\mathcal{E} = \chi'' - \frac{\chi'}{r}\,, \tag{8}$$

and the prime $'$ denotes derivative respect to $r$.

Static coordinates are not commonly used in cosmology, where the comoving coordinate are normally preferred, but can be useful to solve different problems such as the estimation of the maximum size of gravitationally bounded structures. This is due to the fact that in static coordinates it is more convenient to define the turn around radius[3]. For this reason it can be interesting to understand what cosmological perturbations metrics can be written in static coordinates. In general it is difficult to establish if static coordinates exist for a generic metric written in the form in eq.(7), so we will to approach the problem from the opposite direction and look for an answer to this question : what form does it take in terms of cosmological perturbations theory a metric which can be written in static coordinates?

## 4. SDS metric in the Newton gauge

After substituting $R = e^{t/l}r$ in the SDS metric in static coordinates, far from the cosmological and Schwarzschild horizon , i.e. for $m \ll r \ll l$, at leading order we get

$$ds^2 = \left(1-\frac{2m}{ar}\right)dt^2 - 4m\dot{a}\,dtdr - a^2\left(1+\frac{2m}{ar}\right)dr^2 - a^2r^2d\Omega^2\,, \tag{9}$$

where $\dot{a} = da/dt$. Introducing conformal time $d\tau = dt/a(t)$ we then get

$$ds^2 = a^2\left[\left(1-\frac{2m}{ar}\right)d\tau^2 - 4m\frac{a_\tau}{a^2}d\tau dr - \left(1+\frac{2m}{ar}\right)dr^2 - r^2d\Omega^2\right], \tag{10}$$

and comparing with equation (7) we get

$$\psi = -\frac{m}{ar}\,, \tag{11}$$

$$\phi = -\frac{m}{3ar}\,, \tag{12}$$

$$\omega' = \frac{2ma_\tau}{a^2} \tag{13}$$

$$\mathcal{E} = \frac{2m}{ar}. \tag{14}$$

After integrating eq.(8) and $\omega'$ we finally get

$$\omega = \frac{2ma_\tau}{a^2}r \tag{15}$$

$$\chi = -\frac{2mr}{a} + \frac{1}{2}r^2C(\tau) + D(\tau)\,, \tag{16}$$

where $C$ and $D$ are functional constants of integration. Since we are only interested in perturbations which should vanish in a limit in which the mass vanishes, the physically interesting solutions correspond to $C = D = 0$.

Using cosmological perturbation theory (CPT) we can derive explicitly the gauge transformation between the static coordinates and the Newton gauge. Under an infinitesimal space-time translation of the form

$$\tilde{x}^0 = x^0 + \zeta, \tag{17}$$
$$\tilde{x}^i = x^i + \partial^i \beta, \tag{18}$$

the gauge transformations are

$$\tilde{\phi} = \phi - \frac{1}{3}\nabla^2\beta + \frac{a_\tau}{a}\zeta, \tag{19}$$
$$\tilde{\omega} = \omega + \zeta + \beta_\tau, \tag{20}$$
$$\tilde{\psi} = \psi - \zeta_\tau - \frac{a_\tau}{a}\zeta, \tag{21}$$
$$\tilde{\chi} = \chi + 2\beta. \tag{22}$$

Imposing the Newton gauge condition

$$\omega_N = \chi_N = 0, \tag{23}$$

after solving the differential equations (19) we get

$$\beta = \frac{mr}{a}, \tag{24}$$
$$\zeta = -\frac{mra_\tau}{a^2}. \tag{25}$$

We can then use the gauge transformations to calculate the perturbations in the Newton gauge up to leading order

$$\Psi_N = -\frac{m}{ar}, \tag{26}$$
$$\Phi_N = -\frac{m}{ar}. \tag{27}$$

Instead of finding the transformation taking to the Newton gauge we could have also computed the Bardeen potentials[8] directly from eq.(11) up to leading order:

$$\Psi_B = \psi - \frac{1}{a}\left[a\left(\frac{\chi_\tau}{2} - \omega\right)\right]_\tau = -\tfrac{m}{ar}, \tag{28}$$
$$\Phi_B = \phi + \frac{1}{6}\nabla^2\chi - \frac{a_\tau}{a}\left(\omega - \frac{\chi_\tau}{2}\right) = -\tfrac{m}{ar}\,. \tag{29}$$

As expected the the Bardeen potentials reduce to the Newton gauge potentials obtained in eq.(27) and eq.(26), and the metric takes the form

$$ds^2 = a^2\left[\left(1 - \frac{2m}{ar}\right)d\tau^2 - \left(1 + \frac{2m}{ar}\right)\left(dr^2 + r^2 d\Omega^2\right)\right]. \tag{30}$$

## 5. Spherically symmetric solutions in Brans-Dicke theory in static coordinates and in the Newton gauge

In Brans-Dicke (BD) theory the the Jebsen-Birkhoff theorem[5] is valid if the scalar field is time independent. As a consequence under the the assumption of a static scalar field also the static ansatz for the metric adopted in[3] should give the most general spherically symmetric solution. Applying a perturbative approach the solution of the field equations can be written as[9]

$$ds^2 = \left[1-(1+\epsilon)\frac{2m}{R}-(1-2\epsilon)\frac{\Lambda R^2}{3}\right]dt^2 - \left[1-(1-\epsilon)\frac{2m}{R}-(1-4\epsilon)\frac{\Lambda R^2}{3}\right]dR^2 - R^2 d\Omega^2, \tag{31}$$

where $\epsilon = \frac{1}{2\omega+3}$. The SDS solution is recovered in the limit $\epsilon \to 0$, which is also the limit in which the BD theory reduced to GR.

Adopting comoving coordinate defined as $R = ar$ the metric takes the form

$$ds^2 = \left[1-(1+\epsilon)\frac{2m}{ar}-(1-2\epsilon)\frac{\Lambda (ar)^2}{3}\right]dt^2 - \left[1-(1-\epsilon)\frac{2m}{ar}-(1-4\epsilon)\frac{\Lambda (ar)^2}{3}\right](adr+a'rdt)^2 - (ar)^2 d\Omega^2, \tag{32}$$

and after introducing conformal time $dt = ad\tau$

$$ds^2 = a^2\left\{\left[1-(1+\epsilon)\frac{2m}{ar}-(1-2\epsilon)\frac{\Lambda a^2 r^2}{3}\right]d\tau^2 - \left[1-(1-\epsilon)\frac{2m}{ar}-(1-4\epsilon)\frac{\Lambda a^2 r^2}{3}\right]\left(dr^2+2\frac{a'r}{a}drd\tau+\frac{r^2a'^2}{a^2}d\tau^2\right) - r^2 d\Omega^2\right\}, \tag{33}$$

we obtain the Bardeen potentials from the perturbations in the gauge given

$$\Psi_B = -\frac{m}{ar}(1+\epsilon), \tag{34}$$

$$\Phi_B = -\frac{m}{ar}(1-\epsilon), \tag{35}$$

which coincide with the Newton gauge result as expected, due to the gauge invariance of $\Psi_B$ and $\Phi_B$. These potentials reduce to the GR result $\Phi_N = \Psi_N = -\frac{m}{ar}$ in eq.(26,27) when $\epsilon \to 0$.

In general relativity the absence of anisotropic pressure perturbations in the vacuum implies that $\Phi_N = \Psi_N$ while in BD theory the field equations do imply this anymore, and they can be different. Note we have recovered, far form the cosmological horizon, the metric computed in[3] solving the perturbation equations in the Newton gauge. This shows explicitly what is the coordinate transformation between the solution in static and comoving coordinates, and that the solutions are indeed the same. The advantage of this approach is that it allows to derive the metric as a Newton gauge perturbation of the FRW solution directly from the

metric in static coordinates, without the need to solve again the perturbed field equations as it was done in[3].

**Acknowledgments**

We thank Constantinos Skordis for useful discussions.

**References**

1. V. Pavlidou and T. N. Tomaras, Where the world stands still: turnaround as a strong test of Λ CDM cosmology, *JCAP* **1409**, p. 020 (2014).
2. D. Tanoglidis, V. Pavlidou and T. Tomaras, Testing CDM cosmology at turnaround: where to look for violations of the bound?, *JCAP* **1512**, p. 060 (2015).
3. S. Bhattacharya, K. F. Dialektopoulos, A. E. Romano, C. Skordis and T. N. Tomaras, The maximum sizes of large scale structures in alternative theories of gravity (2016).
4. K. Schleich and D. M. Witt, A simple proof of Birkhoff's theorem for cosmological constant, *J. Math. Phys.* **51**, p. 112502 (2010).
5. V. Faraoni, The Jebsen-Birkhoff theorem in alternative gravity, *Phys. Rev.* **D81**, p. 044002 (2010).
6. V. F. Mukhanov, H. A. Feldman and R. H. Brandenberger, Theory of cosmological perturbations. Part 1. Classical perturbations. Part 2. Quantum theory of perturbations. Part 3. Extensions, *Phys. Rept.* **215**, 203 (1992).
7. T. Biswas and A. Notari, Swiss-Cheese Inhomogeneous Cosmology and the Dark Energy Problem, *JCAP* **0806**, p. 021 (2008).
8. J. M. Bardeen, Gauge Invariant Cosmological Perturbations, *Phys. Rev.* **D22**, 1882 (1980).
9. S. Bhattacharya, K. F. Dialektopoulos, A. E. Romano and T. N. Tomaras, Brans-Dicke Theory with $\Lambda > 0$: Black Holes and Large Scale Structures, *Phys. Rev. Lett.* **115**, p. 181104 (2015).

# Scalar-nonmetricity theory

Margus Saal*, Laur Järv, Mihkel Rünkla and Ott Vilson

*Institute of Physics, University of Tartu,*
*W. Ostwaldi 1, 50411 Tartu, Estonia*
**E-mail: margus.saal@ut.ee*

Motivated by the symmetric teleparallel equivalent of general relativity we present a class of extended gravity theories where a scalar field is nonminimally coupled to nonmetricity $Q$. The Lagrange multipliers ensure that the connection is torsion free and flat. As a consequence, there exists a coordinate system where the connection coefficients vanish globally. We give the field equations for the metric, connection, and scalar field, and then briefly discuss cosmology, conformal invariance, and the relation with $f(Q)$ theory.

*Keywords*: symmetric teleparallel gravity, scalar-nonmetricity gravity

## 1. Introduction

General relativity (GR) as a geometric theory of gravity is based on metric Levi-Civita connection endowed with curvature $R$. However, by invoking alternative assumptions about the connection, one may reformulate the theory as teleparallel equivalent of general relativity (TEGR) based on torsion $T$, or symmetric teleparallel equivalent of general relativity (STEGR) based on nonmetricity $Q$.[1] Very recently the latter was generalized to $f(Q)$ gravity[2] in analogy with $f(R)$ theories. In the present note we review our work on constructing a scalar-nonmetricity theory,[3,4] in analogy with the scalar-tensor gravity based on curvature.

## 2. Triality of gravitational theories

### 2.1. *Connections*

On metric-affine spacetimes no constraints are imposed on the metric $g_{\mu\nu}$ and on the connection $\Gamma^{\lambda}{}_{\mu\nu}$ so that they are completely independent variables. The most general affine connection can be decomposed into three parts

$$\Gamma^{\lambda}{}_{\mu\nu} = \left\{{}^{\lambda}{}_{\mu\nu}\right\} + K^{\lambda}{}_{\mu\nu} + L^{\lambda}{}_{\mu\nu} \tag{1}$$

given by the Levi-Civita connection $\left\{{}^{\lambda}{}_{\mu\nu}\right\}$ of the metric $g_{\mu\nu}$, contortion

$$K^{\lambda}{}_{\mu\nu} \equiv \frac{1}{2} g^{\lambda\beta} \left(T_{\mu\beta\nu} + T_{\nu\beta\mu} + T_{\beta\mu\nu}\right) = -K_{\nu\mu}{}^{\lambda}\,, \tag{2}$$

and disformation

$$L^{\lambda}{}_{\mu\nu} \equiv \frac{1}{2} g^{\lambda\beta} \left(-Q_{\mu\beta\nu} - Q_{\nu\beta\mu} + Q_{\beta\mu\nu}\right) = L^{\lambda}{}_{\nu\mu}\,. \tag{3}$$

The contortion tensor can be defined via torsion,

$$T^{\lambda}{}_{\mu\nu} \equiv 2\Gamma^{\lambda}{}_{[\mu\nu]} = \Gamma^{\lambda}{}_{\mu\nu} - \Gamma^{\lambda}{}_{\nu\mu}\,, \tag{4}$$

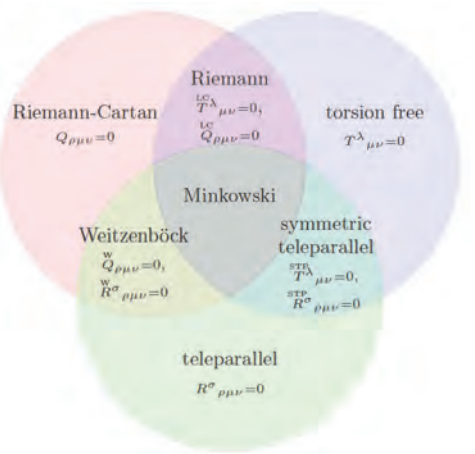



Fig. 1. Restrictions to connection defines different geometries.[3]

and the disformation tensor via nonmetricity,

$$Q_{\rho\mu\nu} \equiv \nabla_\rho g_{\mu\nu} = \partial_\rho g_{\mu\nu} - \Gamma^\beta{}_{\rho\mu} g_{\beta\nu} - \Gamma^\beta{}_{\rho\nu} g_{\mu\beta} \,. \tag{5}$$

Let us recall, that the torsion measures the non-closure of the parallelogram if two vectors are parallel transported along to each other, and the nonmetricity is associated with the change in length of a vector upon parallel transport.

Strictly speaking, nonmetricity, torsion, as well as curvature

$$R^\sigma{}_{\rho\mu\nu} \equiv \partial_\mu \Gamma^\sigma{}_{\nu\rho} - \partial_\nu \Gamma^\sigma{}_{\mu\rho} + \Gamma^\alpha{}_{\nu\rho}\Gamma^\sigma{}_{\mu\alpha} - \Gamma^\alpha{}_{\mu\rho}\Gamma^\sigma{}_{\nu\alpha} \tag{6}$$

are all properties of the connection. Restricting the connection we can define different geometries, see Fig. 1. In the present note we focus our attention to the case, when torsion and curvature are zero, i.e. the symmetric teleparallel (STP) connection. To denote a situation where a particular property is imposed on the connection, and consequently on the covariant derivative, curvature, etc, we use overset labels, e.g., $\overset{\text{STP}}{\Gamma}{}^{\lambda}{}_{\mu\nu}$, $\overset{\text{W}}{\nabla}_\mu$, $\overset{\text{LC}}{R}{}^{\sigma}{}_{\rho\mu\nu}$.

In what follows we shall utilize the symmetric teleparallel connection, $\overset{\text{STP}}{\Gamma}{}^{\lambda}{}_{\mu\nu} = \overset{\text{STP}}{\Gamma}{}^{\lambda}{}_{(\mu\nu)}$, which is torsion free, $\overset{\text{STP}}{T}{}^{\sigma}{}_{\mu\nu} = 0$, and flat, $\overset{\text{STP}}{R}{}^{\sigma}{}_{\rho\mu\nu} = 0$. In that case there exists a coordinate system $\{\xi^\sigma\}$ where the connection coefficients $\overset{\text{STP}}{\Gamma}{}^{\lambda}{}_{\mu\nu}$ vanish globally, $\overset{\text{STP}}{\Gamma}{}^{\lambda}{}_{\mu\nu}(\xi^\sigma) = 0$.

The result leads to corollaries. First, the covariant derivatives commute[1] $\overset{\text{STP}}{\nabla}_\mu \overset{\text{STP}}{\nabla}_\nu \mathbb{T}\Big|_{\{\xi^\sigma\}} = \partial_\mu\partial_\nu \mathbb{T} = \partial_\nu\partial_\mu \mathbb{T} = \overset{\text{STP}}{\nabla}_\nu \overset{\text{STP}}{\nabla}_\mu \mathbb{T}\Big|_{\{\xi^\sigma\}}$, where $\mathbb{T}$ is a tensor (density) of arbitrary rank (and weight). Secondly, in an arbitrary coordinate system $\{x^\mu\}$, the connection coefficients read[2]

$$\overset{\text{STP}}{\Gamma}{}^{\lambda}{}_{\mu\nu} = \frac{\partial x^\lambda}{\partial \xi^\sigma} \frac{\partial}{\partial x^\mu} \left( \frac{\partial \xi^\sigma}{\partial x^\nu} \right) , \tag{7}$$

where $\{\xi^\sigma\}$ are the coordinates for which the connection vanishes.

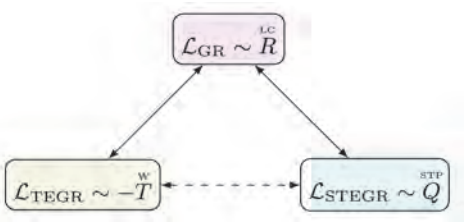



Fig. 2. Triple equivalence of gravitational theories: general relativity (GR) based on connection with vanishing nonmetricity and torsion, teleparallel equivalent of general relativity (TEGR) based on connection with vanishing nonmetricity and curvature, and symmetric teleparallel equivalent of general relativity (STEGR) based on connection with vanishing curvature and torsion.[3]

### 2.2. *Equivalent formulations of general relativity*

The curvature scalar for general connection reads

$$R = \overset{\rm LC}{R} + M^{\lambda}{}_{\nu\rho} M^{\mu}{}_{\mu\lambda} g^{\nu\rho} - M^{\lambda}{}_{\mu\rho} M^{\mu}{}_{\nu\lambda} g^{\nu\rho} + \overset{\rm LC}{\nabla}_{\mu} \left( M^{\mu}{}_{\nu\rho} g^{\nu\rho} - M^{\nu}{}_{\nu\rho} g^{\mu\rho} \right) , \tag{8}$$

where we have used the notation $M^{\lambda}{}_{\mu\nu} = K^{\lambda}{}_{\mu\nu} + L^{\lambda}{}_{\mu\nu}$. If we restrict the geometry to have vanishing torsion and nonmetricity, hence $M^{\lambda}{}_{\nu\rho} = 0$, and the curvature scalar (8) is simply $R = \overset{\rm LC}{R}$ which is the case of GR. If we choose a Weitzenbck connection whereby the curvature and the nonmetricity are zero, thus $\overset{\rm W}{R} = 0$, $M^{\lambda}{}_{\nu\rho} = \overset{\rm W}{K}{}^{\lambda}{}_{\nu\rho}$, then Eq. (8) yields $\overset{\rm LC}{R} = -\overset{\rm W}{T} - 2\overset{\rm LC}{\nabla}_{\rho} \overset{\rm W}{T}{}^{\rho}$, where we introduced the torsion scalar $T$ and its trace $T^{\rho}$. The action given by $\overset{\rm W}{T}$ differs from the one of GR by just a total divergence term and thus provides an equivalent theory (TEGR).

A third possibility is to impose vanishing curvature and torsion, which is the case of symmetric teleparallel connection. Plugging $\overset{\rm STP}{R} = 0$ and $M^{\lambda}{}_{\nu\rho} = \overset{\rm STP}{L}{}^{\lambda}{}_{\nu\rho}$ into (8) now yields

$$\overset{\rm LC}{R} = \overset{\rm STP}{Q} - \overset{\rm LC}{\nabla}_{\rho} (\overset{\rm STP}{Q}{}^{\rho} - \overset{\rm STP}{\tilde{Q}}{}^{\rho}) , \tag{9}$$

where we introduced the nonmetricity scalar $Q$ as

$$Q = -\frac{1}{4} Q_{\alpha\beta\gamma} Q^{\alpha\beta\gamma} + \frac{1}{2} Q_{\alpha\beta\gamma} Q^{\beta\gamma\alpha} + \frac{1}{4} Q_{\alpha} Q^{\alpha} - \frac{1}{2} Q_{\alpha} \tilde{Q}^{\alpha} , \tag{10}$$

and two independent contractions,

$$Q_{\alpha} \equiv Q_{\alpha}{}^{\beta}{}_{\beta} , \qquad \tilde{Q}^{\alpha} \equiv Q_{\beta}{}^{\beta\alpha} . \tag{11}$$

As the action constructed with the nonmetricity scalar (10) restricted to symmetric teleparallel connection would differ from the GR action only by a total divergence term, and the latter does not contribute to the equations of motion, we get another formulation of Einstein's gravity, i.e. STEGR. The three equivalent formulations are summarized on Fig. 2.

## 3. Scalar-nonmetricity theory

Let us introduce the action

$$S = \frac{1}{2}\int \mathrm{d}^4x\sqrt{-g}\,(\mathcal{L}_\mathrm{g} + \mathcal{L}_\ell) + S_\mathrm{m}\,, \tag{12}$$

with gravitational Lagrangian of a scalar $\Phi$ coupled to nonmetricity $Q$,

$$\mathcal{L}_\mathrm{g} = \mathcal{A}(\Phi)Q - \mathcal{B}(\Phi)g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi - 2\mathcal{V}(\Phi)\,, \tag{13}$$

and Lagrange multipliers terms that impose vanishing curvature and torsion,

$$\mathcal{L}_\ell = 2\lambda_\sigma{}^{\nu\mu\rho}R^\sigma{}_{\nu\mu\rho} + 2\lambda_\sigma{}^{\mu\nu}T^\sigma{}_{\mu\nu}\,. \tag{14}$$

Here $S_\mathrm{m}$ denotes the action of matter fields (not directly coupled to the scalar field $\Phi$). The general action contains three functions describing the nonminimal coupling, kinetic term, and potential of the scalar field $\mathcal{A}(\Phi), \mathcal{B}(\Phi), \mathcal{V}(\Phi)$. The nonmetricity scalar $Q$ is given by (10). In the case when $\mathcal{A} = 1$ and $\mathcal{B} = \mathcal{V} = 0$ the theory reduces to plain STEGR.

Variation of the action (12) with respect to the metric gives symmetric field equations

$$\begin{aligned}&\frac{2}{\sqrt{-g}}\overset{\mathrm{STP}}{\nabla}_\sigma\left(\sqrt{-g}\mathcal{A}\overset{\mathrm{STP}}{P}{}^\sigma{}_{\mu\nu}\right) - \frac{1}{2}g_{\mu\nu}\mathcal{A}\overset{\mathrm{STP}}{Q} + \mathcal{A}\left(\overset{\mathrm{STP}}{P}{}_{\mu\sigma\rho}\overset{\mathrm{STP}}{Q}{}_\nu{}^{\sigma\rho} - 2\overset{\mathrm{STP}}{Q}{}_{\sigma\rho\mu}\overset{\mathrm{STP}}{P}{}^{\sigma\rho}{}_\nu\right) + \\ &+\frac{1}{2}g_{\mu\nu}\left(\mathcal{B}g^{\sigma\rho}\partial_\sigma\Phi\partial_\rho\Phi + 2\mathcal{V}\right) - \mathcal{B}\partial_\mu\Phi\partial_\nu\Phi = \mathcal{T}_{\mu\nu}\,.\end{aligned} \tag{15}$$

Here $\overset{\mathrm{STP}}{P}{}_{\mu\sigma\rho}$ is the nonmetricity conjugate (or superpotential)

$$\overset{\mathrm{STP}}{P}{}^\sigma{}_{\mu\nu} = -\frac{1}{2}\overset{\mathrm{STP}}{L}{}^\sigma{}_{\mu\nu} + \frac{1}{4}\overset{\mathrm{STP}}{Q}{}^\sigma g_{\mu\nu} - \frac{1}{4}\left(\overset{\mathrm{STP}}{\tilde{Q}}{}^\sigma g_{\mu\nu} + \delta^\sigma_{(\mu}\overset{\mathrm{STP}}{Q}{}_{\nu)}\right) \tag{16}$$

which satisfies $\overset{\mathrm{STP}}{Q} = \overset{\mathrm{STP}}{Q}{}_\sigma{}^{\mu\nu}\overset{\mathrm{STP}}{P}{}^\sigma{}_{\mu\nu}$ and the matter energy-momentum tensor $\mathcal{T}_{\mu\nu}$ is defined as usual.

Variation with respect to the scalar field $\Phi$ gives

$$2\mathcal{B}\overset{\mathrm{LC}}{\nabla}_\alpha\overset{\mathrm{LC}}{\nabla}{}^\alpha\Phi + \mathcal{B}'\,g^{\mu\nu}\partial_\mu\Phi\partial_\nu\Phi + \mathcal{A}'\overset{\mathrm{STP}}{Q} - 2\mathcal{V}' = 0\,, \tag{17}$$

where the primes mean derivative with respect to the scalar field. The equation for the metric does not contain second derivatives of the scalar field, and analogously the equation for the scalar field does not contain second derivatives of the metric. Hence the equations are naturally "debraided". On the other hand, the scalar field is not sourced by ordinary matter, as it would be in the scalar-curvature-tensor case.

Variation of the action (12) with respect to connection yields an equation containing Lagrange multipliers

$$\overset{\mathrm{STP}}{\nabla}_\rho\left(\sqrt{-g}\lambda_\sigma{}^{\nu\mu\rho}\right) + \sqrt{-g}\lambda_\sigma{}^{\mu\nu} = \sqrt{-g}\mathcal{A}\overset{\mathrm{STP}}{P}{}^{\mu\nu}{}_\sigma\,. \tag{18}$$

Due to the symmetric teleparallel connection the covariant derivatives commute and it is possible to eliminate the Lagrange multipliers, which are antisymmetric with

respect to their last indices, from (18) by acting on it with $\overset{\text{STP}}{\nabla}_\mu \overset{\text{STP}}{\nabla}_\nu$. The result can be simplified further by using the equivalent of Bianchi identity, $\overset{\text{STP}}{\nabla}_\alpha \overset{\text{STP}}{\nabla}_\beta (\sqrt{-g}\overset{\text{STP}}{P}{}^{(\alpha\beta)}{}_\mu) = 0$, to give

$$\overset{\text{STP}}{\nabla}_\mu \left[ \left(\partial_{[\nu}\mathcal{A}\right) \overset{\text{STP}}{\nabla}_{\lambda]} \left(\sqrt{-g}g^{\mu\nu}\right)\right] = 0\,. \tag{19}$$

The equation for the connection (19) is not trivially satisfied and in order to find a solution we must solve the three equations simultaneously. In addition, we have the continuity equation for matter, which follows from the three above mentioned equations and can be used as a separate consistency check.

## 4. Remarks

### 4.1. *Example: flat FRLW cosmology*

Let us consider the spatially (Levi-Civita) flat line element $\mathrm{d}s^2 = -\mathrm{d}t^2 + a(t)^2\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j$. From the equation for the connection (19) we observe that if the metric and the scalar field depend on one and the same variable $t$, then the coincident gauge $\overset{\text{STP}}{\Gamma}{}^\lambda{}_{\mu\nu} = 0$ is a consistent choice, because the anti-symmetrization vanishes for $[0,0]$ case. Calculating the Lagrangian $\overset{\text{STP}}{Q}$ in these coordinates yields $\overset{\text{STP}}{Q} = -6(\dot{a}/a)^2 = -6H^2$ and the corresponding Friedmann equations read

$$H^2 = \frac{1}{3\mathcal{A}}\Big(\rho + \frac{1}{2}\mathcal{B}\dot{\Phi}^2 + \mathcal{V}\Big)\,, \tag{20}$$

$$2\dot{H} + 3H^2 = \frac{1}{\mathcal{A}}\Big(-2\mathcal{A}'H\dot{\Phi} - \frac{1}{2}\mathcal{B}\dot{\Phi}^2 + \mathcal{V} - p\Big)\,, \tag{21}$$

$$\mathcal{B}\ddot{\Phi} + (3\mathcal{B}H + \frac{1}{2}\dot{\mathcal{B}})\dot{\Phi} + \mathcal{V}' + 3\mathcal{A}'H^2 = 0\,. \tag{22}$$

Here $\rho$ is the energy density and $p$ is the pressure of the matter fluid. By combining the equations we arrive at $\dot{\rho} + 3H(\rho + p) = 0$, thus checking once more that the coincident gauge $\overset{\text{STP}}{\Gamma}{}^\lambda{}_{\mu\nu} = 0$ is a suitable choice. These equations coincide with those obtained in the scalar-torsion case, and hence describe both early and late time accelerating phase equally well.

### 4.2. *Conformal transformations*

The nonmetricity scalar transforms under the conformal transformation $\bar{g}_{\mu\nu} = \mathrm{e}^{\Omega(\Phi)}g_{\mu\nu}$ as follows:

$$\bar{Q} = \mathrm{e}^{-\Omega}\Big(Q + \frac{3}{2}g^{\mu\nu}\partial_\mu\Omega\partial_\nu\Omega + (Q^\alpha - \tilde{Q}^\alpha)\partial_\alpha\Omega\Big). \tag{23}$$

The term $g^{\mu\nu}\partial_\mu\Omega\partial_\nu\Omega$ can be absorbed into the redefinition of the kinetic term of the scalar field, however the term $(Q^\alpha - \tilde{Q}^\alpha)\partial_\alpha\Omega$ does not appear in the original action and causes the action (12) not to preserve its structure under conformal transformations. However, if we add a term $\frac{1}{2}(Q^\alpha - \tilde{Q}^\alpha)\partial_\alpha\mathcal{A}(\Phi)$ to the original

Lagrangian, we obtain the exact equivalent to the familiar scalar-curvature theory, which is covariant under the conformal transformations and scalar field redefinitions.

### 4.3. *Equivalence with $f(Q)$*

Following the standard procedure it is possible to show that $f(Q)$ theories[2] where $\mathcal{L}_{\mathrm{g},f(Q)} = f(Q)$ form a subclass of scalar-nonmetricity theories. Utilizing an auxiliary field $\Phi$ we can write

$$\mathcal{L}_{\mathrm{g},aux} = f'(\Phi)\,Q - (f'(\Phi)\,\Phi - f(\Phi))\ . \tag{24}$$

By varying the action (24) with respect to $\Phi$ yields $f''(\Phi)(\Phi - Q) = 0$. By varying the action (24) with respect to $\Phi$ one gets $f''(\Phi)(\Phi - Q) = 0$. As $f''(\Phi) \neq 0$, the equation of motion gives $Q = \Phi$ and we restore the original action. After identifications $\mathcal{A}(\Phi) = f'(\Phi)$ and $2\mathcal{V}(\Phi) = f'(\Phi)\,\Phi - f(\Phi)$ we get scalar-nonmetricity action without a kinetic term for the scalar field (i.e $\mathcal{B}(\Phi) = 0$).

## 5. Summary

Of the three alternative geometric reformulations of general relativity, the symmetric teleparallel gravity and its extensions have received the least attention so far. In our work we endeavor to explore the ground by considering a generic setting where a scalar field is nonminimally coupled to the nonmetricity scalar in the symmetric teleparallel framework. A lot of research waits ahead, most obviously constructing solutions and clarifying their features, but also understanding the relations between the theories established in different geometric settings.

## Acknowledgments

The work was supported by the Estonian Ministry for Education and Science through the projects IUT02-27 and PUT790, as well as the European Regional Development Fund through the Center of Excellence TK133 “The Dark Side of the Universe”.

## References

1. J. M. Nester and H. J. Yo, Chin. J. Phys. **37**, 113 (1999).
2. J. Beltrán Jiménez, L. Heisenberg and T. Koivisto, Phys. Rev. D **98**, no. 4 044048 (2018).
3. L. Järv, M. Rünkla, M. Saal and O. Vilson, Phys. Rev. D **97**, no. 12 124025 (2018).
4. M. Rünkla and O. Vilson, Phys. Rev. D **98**, no. 8 084034 (2018).

# Anisotropic universe with power law $f(R)$ gravity

M. Farasat Shamir*

*Department of Humanities and Sciences,*
*National University of Computer and Emerging Sciences, Lahore Campus, Pakistan*
**E-mail: farasat.shamir@nu.edu.pk*

This study highlights the dynamics of anisotropic universe in modified gravity. To meet this aim, locally rotationally symmetric Bianchi type $I$ spacetime is studied in the metric theory of $f(R)$ gravity. Anisotropic fluid is considered to study the exact solutions of modified field equations. In particular, a general solution metric is reported using the well-known power law $f(R)$ gravity model. The graphical analysis of equation of state parameter is given which includes the corresponding values predicted for cosmic expansion. The energy conditions are also discussed for a range of specific model parameter. It is shown that anisotropic universe in modified gravity anticipate some interesting solutions and viable power law $f(R)$ gravity models can be reconstructed using suitable values of model parameter.

*Keywords*: Modified Gravity; $f(R)$ Gravity; Expansion of Universe.

## 1. Introduction

Sufficient experimental evidence in the recent decade have unveiled a mysterious picture of accelerated expansion of universe[1]. It is also believed that universe is currently exhibiting a transient phase of cosmic expansion with $\omega = -1$ and ultimately it will collapse sometime in the future[2]. This interesting phenomenon of late time acceleration and dark energy is somehow justified by modification of gravity. The $f(R)$ theory of gravity has been debated seriously with some fruitful results[3]. Interesting reviews[4] may be helpful for a better comprehension of the theory.

The involvement of a general function of Ricci scalar in the theory predicts many solutions of modified field equations as compared to ordinary general relativity. However, one can also recover already known solutions. The simplest example is the existence of de Sitter solution when dealing with the vacuum case and power law $f(R)$ gravity model[5]. Similarly, Einstein static universe does exist in $f(R)$ theory with barotropic perfect fluid[6]. Thus investigating modified gravity, in particular, using power law model with anisotropic background seems interesting. In this study, locally rotationally symmetric (LRS) Bianchi type $I$ spacetime is investigated by considering anisotropic fluid and power law $f(R)$ gravity model.

## 2. Field Equations in $f(R)$ Gravity

The $f(R)$ gravity field equations are

$$f_R(R)R_{\mu\nu} - \frac{1}{2}f(R)g_{\mu\nu} - \nabla_\mu\nabla_\nu f_R(R) + g_{\mu\nu}\Box f_R(R) = \kappa T_{\mu\nu}, \tag{1}$$

where $f_R(R)$ is the derivative of $f(R)$ with respect to $R$ and the other notations have their usual meanings. Moreover, we consider LRS Bianchi type $I$ spacetime

$$ds^2 = dt^2 - X^2(t)dx^2 - Y^2(t)[dy^2 + dz^2], \tag{2}$$

and assume that the universe is filled with anisotropic fluid

$$T^{\mu}_{\nu} = diag[\rho, -p_x, -p_y, -p_z]. \tag{3}$$

We further characterize the anisotropic fluid as

$$T^{\mu}_{\nu} = diag[1, -\omega, -(\omega+\delta), -(\omega+\delta)]\rho, \tag{4}$$

where $\omega$ and $\delta$ are equation of state and skewness parameters respectively. Since $f(R)$ gravity field equations are highly non-linear in nature, so we adopt a conventional physical assumption of proportionality of shear scalar and expansion scalar which gives

$$X = Y^s, \qquad s \in \mathbb{R}. \tag{5}$$

Using, Eqs. (2), (4) and (5), the modified field equations simplify to

$$(2s+1)\frac{\dot{Y}^2}{Y^2}f_R - \frac{f}{2} - \tilde{Y}f_R + (s+2)\frac{\dot{Y}}{Y}\dot{f}_R = \kappa\rho, \tag{6}$$

$$-\left(\frac{2\ddot{Y}}{Y} + \frac{\dot{Y}^2}{Y^2}\right)f_R + \frac{f}{2} + \tilde{Y}f_R - \frac{2\dot{Y}}{Y}\dot{f}_R - \ddot{f}_R = \kappa\omega\rho, \tag{7}$$

$$-\left((s+1)\frac{\ddot{Y}}{Y} + s^2\frac{\dot{Y}^2}{Y^2}\right)f_R + \frac{f}{2} + \tilde{Y}f_R - (s+1)\frac{\dot{Y}}{Y}\dot{f}_R - \ddot{f}_R = \kappa(\omega+\delta)\rho, \tag{8}$$

where $\tilde{Y} \equiv (s+2)\frac{\ddot{Y}}{Y} + (s^2+s+1)\frac{\dot{Y}^2}{Y^2}$ and the over dot is to denote time derivative.

## 3. Dynamics of Anisotropic Universe with Power Law $f(R)$ Model

Now we consider the well known $f(R)$ gravity model as

$$f(R) = \zeta R^{r+1}, \qquad \zeta,\ r \in \mathbb{R} - \{0\}. \tag{9}$$

Differentiating Eq.(9) with respect to $R$, it follows that

$$f_R(R) = \zeta(r+1)R^r, \tag{10}$$

Without loss of any generality, we may choose $\zeta = \frac{1}{r+1}$. This simplifies the work for further analysis. Since the field equations (6)-(8) are still complicated even with the choice of power law model. Thus, a power law form of metric coefficient is chosen

$$Y(t) = \xi t^{\frac{1}{s+2}}, \quad \xi \in \mathbb{R} - \{0\}. \tag{11}$$

Using these settings, Eqs.(6)-(8) provide the corresponding expressions for energy density and pressure components

$$\rho = \frac{-1}{t^2(r+1)(s+2)^2}\Bigg[\big(2rs^2+2r^2s^2-2s+4rs+8sr^2+ 8r^2+6r-1\big)\left(\frac{2(2s+1)}{t^2(s+2)^2}\right)^r\Bigg], \tag{12}$$

$$p_x = \frac{-1}{t^2(r+1)(s+2)^2}\Bigg[\big(2rs^2-2s+4r^3s^2+6r^2s^2+4rs+ 20sr^2+16r^3s+16r^2+16r^3-1\big)\left(\frac{2(2s+1)}{t^2(s+2)^2}\right)^r\Bigg], \tag{13}$$

$$p_y = \frac{-1}{t^2(r+1)(s+2)^2}\Bigg[\big(2rs-2s+4r^2s^2+4r^3s^2+16r^3s+ 18sr^2+4r+16r^3+20r^2-1\big)\left(\frac{2(2s+1)}{t^2(s+2)^2}\right)^r\Bigg] = p_z. \tag{14}$$

Manipulating Eqs.(12)-(14), EoS and skewness parameters turn out to be

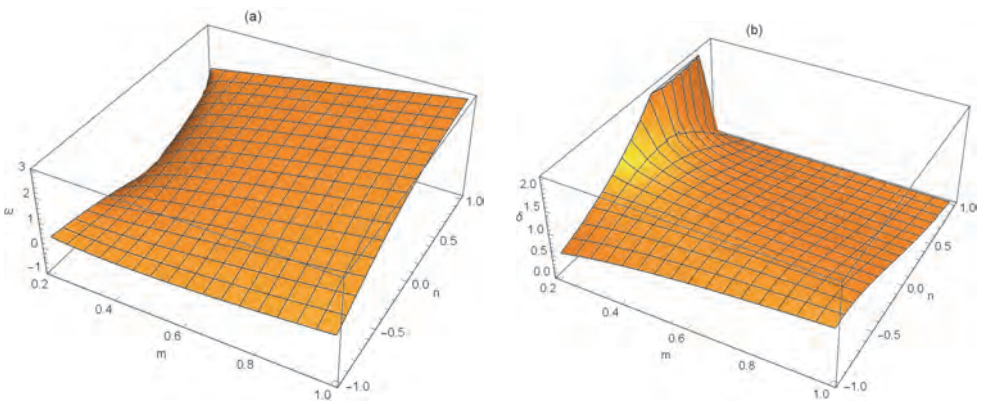



Fig. 1. Behavior of equation of state and skewness parameters

$$\omega = \frac{4r^3s^2+2s^2r+6r^2s^2-2s+20sr^2+16r^3s+4sr+16r^3+16r^2-1}{2r^2s^2+2s^2r+8sr^2+4sr-2s+8r^2+6r-1}, \tag{15}$$

$$\delta = -\frac{2r(r+1)(s+2)(s-1)}{2r^2s^2+2s^2r+8sr^2+4sr-2s+8r^2+6r-1}. \tag{16}$$

The behavior of these parameters is reflected in Fig. $(1a)$ and $(1b)$ respectively. It is worthwhile to notice that many possibilities of equation of state parameter exist. For example, an interesting case arise when $\omega = -1$ describing cosmic expansion[2,7,8].

In this case, explicit solutions of Eq.(15) turn out to be

$$r = \frac{-1 - s \pm \sqrt{3 + 6s + s^2}}{2(s+2)}, \tag{17}$$

and

$$s = \frac{-4r^2 - 3r + 1 \pm \sqrt{-5r^2 - 4r + 1}}{2r(r+1)}. \tag{18}$$

So in the light of above constraints, one can choose suitable power law $f(R)$ gravity model and the corresponding solution metric. In this case, a generic solution metric turns out to be

$$ds^2 = dt^2 - \xi^{2s} t^{\frac{2s}{s+2}} dx^2 - \xi^2 t^{\frac{2}{s+2}} (dy^2 + dz^2). \tag{19}$$

It is important to notice here that the anisotropy parameter involved in the above spacetime may also be used to reconstruct some well known solutions for physical relevance. For example, $s = -1/2$ gives the well known Kasner's universe and obviously this value corresponds to vacuum case (see Eqs.(12)-(14)).

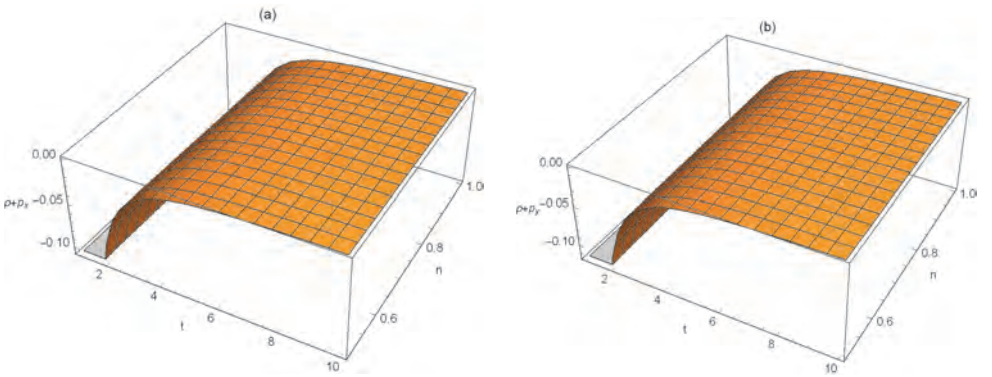



Fig. 2. Plots of NEC and WEC

### 3.1. *Power Law model and Energy Bounds*

The energy bounds have gained popularity in the modern day cosmology. It is important to mentation here that viability of cosmological models and some important theorems about black holes are associated with the energy bounds. The usual energy conditions namely null energy conditions (NEC), weak energy conditions (WEC), strong energy conditions (SEC) and dominant energy conditions (DEC)

are defined as

$$\begin{aligned}
&\text{NEC:} \quad \rho + p_x \geq 0, \quad \rho + p_y \geq 0 \\
&\text{WEC:} \quad \rho \geq 0, \quad \rho + p_x \geq 0, \quad \rho + p_y \geq 0, \\
&\text{SEC:} \quad \rho + 3p_x \geq 0, \quad \rho + 3p_y \geq 0, \quad \rho + p_x \geq 0, \quad \rho + p_y \geq 0, \\
&\text{DEC:} \quad \rho \geq 0, \quad \rho \pm p_x \geq 0, \quad \rho \pm p_y \geq 0.
\end{aligned} \tag{20}$$

The graphical behavior of these conditions is reflected in Figs. (2)-(4). It is mentioned here that we have fixed $\kappa = 1$ and the model parameter $r = 1/3$. Moreover, we choose the same range of anisotropy parameter $s$ as in the case of EoS parameter plot (see Fig. (1)). Since these values corresponds to negative $\omega$, so it is already anticipated that most of the energy bounds will not be satisfied. In particular, Fig. (3) depicts that SEC is violated. It is interesting as this violation seems to support an accelerating universe.

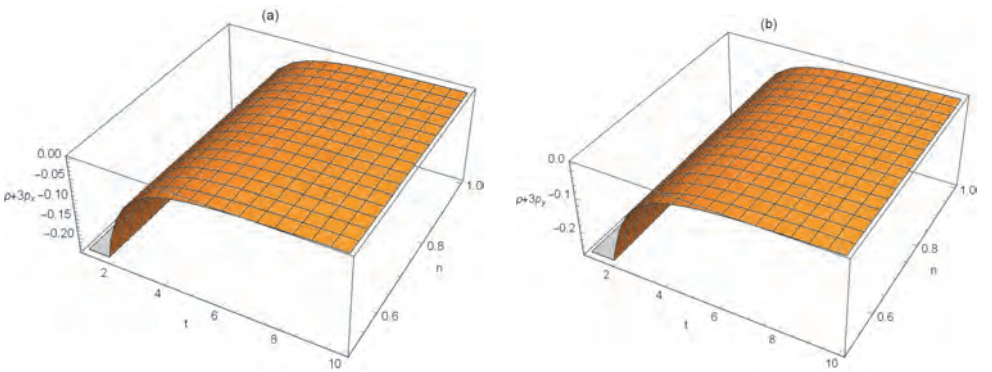



Fig. 3. Plots of SEC

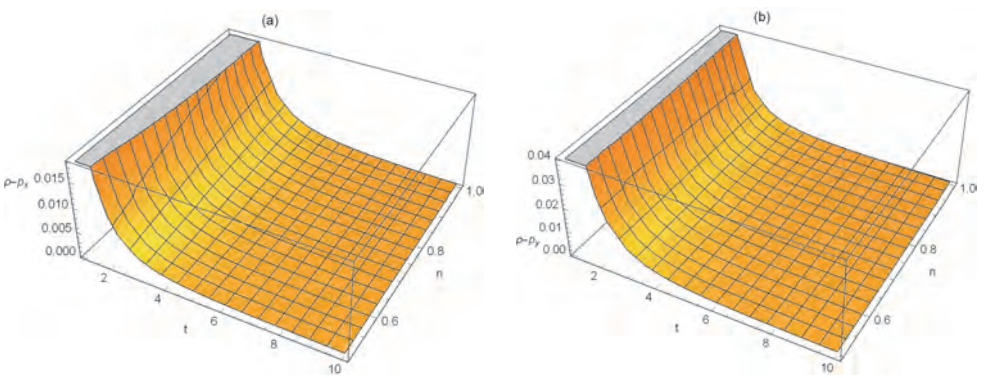



Fig. 4. Plots of DEC

## 4. Outlook

In the present work, we study the dynamics of anisotropic universe in modified gravity. LRS Bianchi type $I$ spacetime is investigated in the metric theory of $f(R)$ gravity to meet this aim. Moreover, an anisotropic fluid is considered with deviation parameter to discuss the dynamics of modified field equations. A general solution metric is reported using the well-known power law $f(R)$ gravity model. In particular, explicit equation for EoS parameter is reported in terms of anisotropy and $f(R)$ model parameters. The graphical analysis of equation of state parameter is also given which includes $\omega = -1$ predicted for cosmic expansion. The energy conditions are also discussed for a range of specific model parameter. We have only discussed some limited choices and many other possibilities can also be explored. It is shown that anisotropic universe in modified gravity anticipate some interesting solutions and viable power law $f(R)$ gravity models can be reconstructed using suitable values of model parameter.

## References

1. Riess, A.G. et al.: Astrophys. J. **607**(2004)665; Tegmartk, M. et al.: Phys. Rev. **D69**(2004)103501; Spergel, D.N. et al.: Astrophys. J. Suppl. **148**(2003)175.
2. Kaloper, N. and Padilla, A.: Phys. Rev. **D90**(2014)084023.
3. Nojiri, S. and Odintsov, S.D.: Int. J. Geom. Meth. Mod. Phys. **115**(2007)4; Nojiri, S. and Odintsov, S.D.: *Problems of Modern Theoretical Physics*, A Volume in honour of Prof. Buchbinder, I.L. in the occasion of his 60th birthday, p.266-285, (TSPU Publishing, Tomsk), arXiv:0807.0685.
4. Felice, A.D and Tsujikawa, S.: Living Rev. Rel. **13**(2010)3; Bamba, K., Capozziello, S., Nojiri, S. and Odintsov, S.D.: Astrophys. Space Sci. **342**(2012)155; Sotiriou, T.P. and Faraoni, V.: Rev. Mod. Phys. **82**(2010)451.
5. Starobinsky, A.A.: Phys. Lett. **B91**(1980)99.
6. Seahra, S.S., and Boehmer, C.G.: Phys. Rev. **D79**(2009)064009.
7. Corasaniti, P. S. et al: Phys. Rev. **D70**(2004)083006.
8. Hogan, J.: Nature **448**(2007)240.

# Stability of Einstein universe in matter-curvature coupling gravity

Sharif M.* and Waseem A.

*Department of Mathematics, University of the Punjab,*
*Lahore, 54590, Pakistan*
**E-mail: msharif.math@pu.edu.pk*
*www.pu.edu.pk*

We investigate the existence and stability of the Einstein universe in the context of $f(R,T,Q)$ gravity, where $Q = R_{\mu\nu}T^{\mu\nu}$. Considering linear homogeneous perturbations around scale factor and energy density, we formulate static as well as perturbed field equations. We parameterize the stability regions corresponding to conserved as well as non-conserved energy-momentum tensor using linear equation of state parameter for particular models of this gravity. The graphical analysis concludes that for a suitable choice of parameters, the stable regions of the Einstein universe are obtained which indicate that the big-bang singularity can be avoided successfully by the emergent mechanism in non-minimal matter-curvature coupled gravity.

*Keywords*: Stability analysis; Einstein universe; Modified gravity.

## 1. Introduction

The strong non-minimal coupling between gravity and matter fields present in $f(R,T,Q)$ gravity has gained much attention to study the issue of current accelerated cosmic expansion[1]. Odintsov and Sáez-Gómes[2] explored matter instability, $\Lambda$CDM model and de Sitter solutions in this modified theory. Sharif and Zubair investigated the validity of thermodynamical laws[3] and derived the energy conditions[4] for two different models of this gravity. The physical behavior of compact objects is also discussed in this gravity[5]. We have analyzed the dust spherical solution in the same gravity[6].

The emergent universe scenario to resolve the issue of big-bang singularity is based on the stability of Einstein universe (EU) against all kinds of perturbations. In general relativity, the idea of this emergent universe is not proved successful due to unstable EU against homogenous perturbations[7]. Seahra and Böhmer[8] showed that stable EU solutions exist in $f(R)$ models for perfect fluid with linear equation of state (EoS) and remain unstable against inhomogeneous perturbations. Shabani and Ziaie[9] discussed the stable solutions of EU in $f(R,T)$ gravity which were unstable in $f(R)$ gravity. Sharif and Ikram[10] investigated the stability of EU against linear homogeneous perturbations in $f(\mathcal{G},T)$ gravity and found stable solutions.

This paper explores the effects of strong non-minimal coupling on the stability of EU using homogeneous linear perturbations in $f(R,T,Q)$ gravity. In the next section, we formulate the field equations of this theory and investigate the stability of EU for both conserved and non-conserved energy-momentum tensor (EMT). In the last section, we concluding our results.

## 2. Stability Analysis of Einstein Universe in $f(R,T,Q)$ Gravity

The line element for closed FRW universe model is given by

$$ds^2 = dt^2 - a^2(t)\left(\frac{1}{1-r^2}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right), \tag{1}$$

where $a(t)$ represents the scale factor. The EMT for perfect fluid is

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - pg_{\mu\nu}, \tag{2}$$

where $\rho,\ p$ and $u_\mu$ indicate energy density, pressure and four velocity, respectively. In closed FRW universe background, the field equations of $f(R,T,Q)$ gravity corresponding to $\mathcal{L}_m = -p$ are obtained as

$$\begin{aligned}\frac{3}{a^2}(1+\dot{a}^2) = \frac{1}{f_R + pf_Q}\Bigg[&\kappa^2\rho + (\rho+p)f_T + 3\left(\frac{a\ddot{a}+\dot{a}^2+1}{a^2}\right)f_R + \frac{1}{2}\\ &\times\ f(R,T,Q) - 3\frac{\dot{a}}{a}\partial_t f_R - \frac{3}{2}\left(\frac{4\dot{a}^2 - a\ddot{a}+2}{a^2}\right)\rho f_Q - \frac{3}{2}\\ &\times\ \left(\frac{a\ddot{a}+2\dot{a}^2}{a^2}\right)pf_Q + \frac{3\dot{a}}{2a}\partial_t[(p-\rho)f_Q]\Bigg],\end{aligned} \tag{3}$$

$$\begin{aligned}-2a\ddot{a} - (1+\dot{a}^2) = \frac{1}{f_R + pf_Q}\Bigg[&\kappa^2a^2p - \frac{a^2}{2}f(R,T,Q) - 3(a\ddot{a}+\dot{a}^2+1)f_R\\ &+\ 2a\dot{a}\partial_t f_R + a^2\partial_{tt}f_R + \frac{1}{2}\left(4\dot{a}^2 - a\ddot{a}+2\right)pf_Q + \frac{1}{2}(a\ddot{a}\\ &+\ 2\dot{a}^2)\rho f_Q + 2a\dot{a}\partial_t[(p+\rho)f_Q] + \frac{a^2}{2}\partial_{tt}[(\rho-p)f_Q]\Bigg],\end{aligned} \tag{4}$$

where dot shows derivative with respect to time. The conservation equation is

$$\begin{aligned}\dot{\rho} + 3\frac{\dot{a}}{a}(\rho+p) = \frac{1}{2\kappa^2 + 3(f_T - \frac{2\dot{a}^2+a\ddot{a}}{a^2}f_Q)}\Bigg[&\Bigg\{3\frac{\dot{a}}{a^3}\left(a\ddot{a} - 4\dot{a}^2\right)(\rho+p)\\ &-\ 3(\frac{a\ddot{a}-\dot{a}^2}{a^2})\partial_t p\Bigg\}f_Q - 2(\rho+p)\partial_t f_T + 6(\frac{\dot{a}}{a})^2(\rho+p)\\ &\times\ \partial_t f_Q + \partial_t p f_T\Bigg].\end{aligned} \tag{5}$$

We consider linear homogeneous perturbations to investigate the stability of EU in $f(R,T,Q)$ gravity. For this purpose, we take $a(t) = a_0 = constant$ for EU and the corresponding field equations (3) and (4) reduce to

$$\frac{3}{a_0^2} = \frac{1}{f_R + p_0 f_Q}\left[\kappa^2\rho_0 + (\rho_0+p_0)f_T + \frac{f(R_0,T_0,Q_0)}{2} + \frac{3}{a_0^2}f_R - \frac{3}{a_0^2}\rho_0 f_Q\right], \tag{6}$$

$$-\frac{1}{a_0^2} = \frac{1}{f_R + p_0 f_Q}\left[\kappa^2 p_0 - \frac{1}{2}f(R_0,T_0,Q_0) - \frac{3}{a_0^2}f_R + \partial_{tt}f_R + \frac{p_0}{a_0^2}f_Q\right], \tag{7}$$

where $R_0 = R(a_0) = -\frac{6}{a_0^2}$, $T_0 = \rho_0 - 3p_0$ and $Q_0 = 6\frac{p_0}{a_0^2}$. Here $\rho_0$ and $p_0$ denote the unperturbed energy density and pressure, respectively. In order to examine the stability regions, we consider linear EoS defined as $p(t) = \omega\rho(t)$ ($\omega$ is the EoS

parameter) and introduce the expressions for linear perturbations in scale factor and energy density depending only on time as follows

$$a(t) = a_0 + a_0\delta a(t), \quad \rho(t) = \rho_0 + \rho_0\delta\rho(t), \tag{8}$$

where $\delta a(t)$ and $\delta\rho(t)$ express the perturbed scale factor and energy density, respectively. We assume that $f(R,T,Q)$ is analytic and by applying Taylor series expansion for three variables upto first order, this turns out to be

$$f(R,T,Q) = f(R_0,T_0,Q_0) + f_R\delta R + f_T\delta T + f_Q\delta Q. \tag{9}$$

Using linear EoS, $\delta R$, $\delta T$ and $\delta Q$ become

$$\delta R = -6(\delta\ddot{a} - 2\frac{\delta a}{a_0^2}), \quad \delta T = T_0\delta\rho_0, \quad \delta Q = 3\rho_0[(\omega-1)\delta\ddot{a} - 4\frac{\omega}{a_0^2}\delta a + 2\frac{\omega}{a_0^2}\delta\rho]. \tag{10}$$

Substituting Eqs.(6)-(10) into the field equations (3) and (4), we obtain the linearized perturbed equations as

$$6(f_R+\rho_0 f_Q)\delta a + a_0^2\rho_0[\kappa^2 + (1+\omega)f_T + \frac{1}{2}(1-3\omega)f_T - \frac{3}{a_0^2}f_Q]\delta\rho = 0, \tag{11}$$

$$2(f_R+\rho_0 f_Q)\delta\ddot{a} + \frac{2}{a_0^2}(\rho_0\omega f_Q - f_R)\delta a + \rho_0[\kappa^2\omega - \frac{1}{2}(1-3\omega)f_T - \frac{\omega}{a_0^2}f_Q]\delta\rho$$
$$+\frac{1}{2}\rho_0(1-\omega)f_Q\delta\ddot{\rho} = 0. \tag{12}$$

These express a direct relation between perturbed scale factor and energy density.

### 2.1. *Stability for conserved case*

The conservation law does not hold in $f(R,T,Q)$ gravity but we assume that this law holds in this gravity for which the right hand side of Eq.(5) becomes zero and we obtain

$$\dot{p}f_T - 2(\rho+p)\partial_t f_T + \frac{3}{a^2}\left[\frac{\dot{a}}{a}(a\ddot{a} - 4\dot{a}^2)(\rho+p) - (a\ddot{a}-\dot{a}^2)\dot{p}\right]f_Q$$
$$+6(\frac{\dot{a}}{a})^2(\rho+p)\partial_t f_Q = 0. \tag{13}$$

From the standard conservation equation, we obtain the relation defined by $\delta\dot{\rho}(t) = -3(1+\omega)\delta\dot{a}(t)$. To construct the perturbed field equation in the form of perturbed scale factor, we eliminate $\delta\rho$ from Eqs.(11) and (12), obtain $a_0^2$ from Eqs.(6) and (7) and then substitute the expression of $\delta\dot{\rho}(t)$ in the resulting equation which yields

$$\left[2(f_R-\omega\rho_0 f_Q)\left[\rho_0\left\{\kappa^2 + (1+\omega)f_T - \left(\frac{3\rho_0(1+\omega)(\kappa^2+f_T)}{2f_R+\rho_0(3+\omega)f_Q}\right)f_Q + \frac{1}{2}(1-3\omega)\right.\right.\right.$$
$$\left.\left.\times\ f_T\}\right] + 6\rho_0(f_R+\rho_0 f_Q)\left\{\kappa^2\omega - \left(\frac{\omega\rho_0(1+\omega)(\kappa^2+f_T)}{2f_R+\rho_0(3+\omega)f_Q}\right)f_Q - \frac{1}{2}(1-3\omega)f_T\right\}\right]\delta a$$
$$+\left[\rho_0\left(\frac{2f_R+\rho_0(3+\omega)f_Q}{\rho_0(1+\omega)(\kappa^2+f_T)}\right)\left\{\kappa^2 + (1+\omega)f_T - \left(\frac{3\rho_0(1+\omega)(\kappa^2+f_T)}{2f_R+\rho_0(3+\omega)f_Q}\right)f_Q + \frac{1}{2}\right.\right.$$
$$\left.\left.\times\ (1-3\omega)f_T\right\}\left\{\frac{3}{2}\rho_0(1-\omega^2)f_Q - 2(f_R+\rho_0 f_Q)\right\}\right]\delta\ddot{a} = 0. \tag{14}$$

We consider a specific form of $f(R,T,Q)$ gravity given by $f(R,T,Q) = R + f(T)+g(Q)$, where $f(T)$ and $g(Q)$ are the generic functions of $T$ and $Q$, respectively. We assume that the conservation law holds for this model. Using this assumption along with $a(t) = a_0 = constant$, Eq.(13) leads to $\dot{p}f'(T) - 2(\rho + p)f''(T)\dot{T} = 0$, where prime shows derivative with respect to $T$. Consequently, using linear EoS ($p = \omega\rho$) and perturbation in energy density given in Eq.(8), the resulting second-order differential equation is obtained as $\omega f'(T) - 2(1 + \omega)Tf''(T) = 0$, whose solution is

$$f(T) = \frac{c_1 T(1+\omega)(2T(1+\omega))^{\frac{\omega}{2(1+\omega)}}}{2+3\omega} + c_2, \tag{15}$$

where $c_1$ and $c_2$ are integration constants.

It is mentioned here that the conservation law holds only for this unique expression of $f(T)$ in the proposed model. For the considered model, substitution of Eq.(15) into Eq.(14) leads to the following resulting differential equation

$$\left[6\{\Delta_1 + \Delta_2\Delta_3 + 2\Delta_3^2\Delta_4\} - 3g'(Q)\{2\Delta_5 + \Delta_3\Delta_6 - \Delta_3^2\Delta_7\}\right]\delta a - [(\Delta_8 + \Delta_9) \times g'(Q) + 2\Delta_3\Delta_{10}]\delta\ddot{a} = 0, \tag{16}$$

where $\Delta_i$'s $(i = 1, 2, 3, ..., 10)$ are

$$\Delta_1 = \frac{\rho_0\kappa^4}{3}(1 + 4\omega + 3\omega^2), \quad \Delta_2 = \rho_0\kappa^2\left(1 + \omega(10 + 7\omega) + 2\rho_0\omega(1 + 3\omega + \omega^2)\right),$$
$$\Delta_3 = \frac{c_1}{2}\left(2\rho_0(1+\omega)(1-3\omega)\right)^{\frac{\omega}{2(1+\omega)}}, \quad \Delta_4 = \rho_0\omega(2 + \omega + \rho_0(1 + \omega + \omega^2)),$$
$$\Delta_5 = \rho_0^2\omega\kappa^4(1 + 4\omega + 3\omega^2), \quad \Delta_7 = \rho_0^2(\kappa^2\omega^2(1 + 3\omega) - 3(1 + 2\omega) + 3\omega^3(2 + \omega)),$$
$$\Delta_6 = \rho_0^2\kappa^2(8\omega(1 + \omega) + 3(1 - \omega^4)), \quad \Delta_8 = \kappa^2\left(2(3 + \omega) + \rho_0(1 - 4\omega + 3\omega^2)\right)$$
$$\Delta_9 = 9 - \omega^2 + \frac{\rho_0}{2}(9 - 13\omega + 7\omega^2 - 3\omega^3), \quad \Delta_{10} = \frac{2\kappa^2}{\Delta_3} + (3 - \omega).$$

The solution of Eq.(16) is $\delta a(t) = b_1 e^{\Omega t} + b_2 e^{-\Omega t}$, where $b_1$ and $b_2$ are integration constants and $\Omega$ represents the frequency of small perturbation which is of the form

$$\Omega^2 = \frac{6(\Delta_1 + \Delta_2\Delta_3 + 2\Delta_3^2\Delta_4) - 3g'(Q)(2\Delta_5 + \Delta_3\Delta_6 - \Delta_3^2\Delta_7)}{(\Delta_8 + \Delta_9)g'(Q) + 2\Delta_3\Delta_{10}}. \tag{17}$$

### 2.2. *Stability for non-conserved case*

Here, we discuss the stability when EMT is not conserved. We consider another specific model $f(R,T,Q) = \alpha R + f(Q)$, where $\alpha$ is an arbitrary constant. For this model, the perturbed field equations (11) and (12) lead to

$$6\left[\alpha + \rho_0 f'(Q)\right]\delta a + a_0^2\rho_0\left[\kappa^2 - \frac{3}{a_0^2}f'(Q)\right]\delta\rho = 0, \tag{18}$$

$$\left[2\alpha + \rho_0(\frac{1+3\omega^2}{2})f'(Q)\right]\delta\ddot{a} + \frac{2}{a_0^2}\left[\rho_0\omega f'(Q) - \alpha\right]\delta a + \rho_0\omega[\kappa^2 - \frac{1}{a_0^2}f'(Q)]\delta\rho = 0. \tag{19}$$

The differential equation in perturbed scale factor is obtained by eliminating $\delta\rho$ and by substituting the value of $a_0^2$ of the form

$$2\kappa^2\rho_0(1+\omega)\left[\alpha(1+3\omega)+4\omega\rho_0 f'(Q)\right]\delta a-\left[4\alpha^2+\alpha\rho_0(3\omega^2-2\omega+7)f'(Q)\right]\delta\ddot{a}=0,$$

whose solution is $\delta a(t) = a_1 e^{\bar{\Omega}t} + a_2 e^{-\bar{\Omega}t}$, where $a_j$'s $(j = 1, 2)$ are integration constants and frequency of small perturbation $(\bar{\Omega})$ with $\xi_2 = \rho_0 f'(Q)$ is given by

$$\bar{\Omega}^2 = \frac{2\kappa^2\rho_0(1+\omega)\left(\alpha(1+3\omega)+4\omega\xi_2\right)}{4\alpha^2+\alpha(3\omega^2-2\omega+7)\xi_2}. \tag{20}$$

For graphical analysis, we take $\kappa^2 = 1$, $\rho_0 = 0.3$ and $g'(Q) = \xi_1$ as a new parameter. Fig. **1** shows the existence of stable EU for $\Omega^2$ corresponding to different values of $c_1$. It is observed that the stable EU exists for all values of $\omega \geq -1$ and these regions are becoming more smooth with increasing value of $c_1$. For negative values of $c_1$, the stable modes exist in the range $-1 < \omega < \frac{1}{3}$ and with decreasing value of $c_1$, the graphs show more stable regions towards positive values of EoS parameter. Figure **2** indicates the stable EU for $\bar{\Omega}^2$ and it is found that for $\alpha = 1$, the stable regions appear only for negative values of $\omega$ but for greater values of $\alpha$, we obtain stable region also for positive values of $\omega$. The stable EU for negative values of $\alpha$ also exists and this stability increases with decreasing value of $\alpha$.

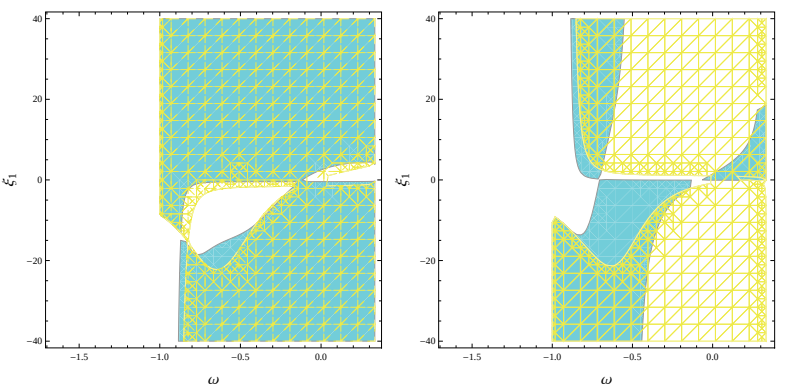



Fig. 1. Plots of stable regions in $(\omega, \xi_1)$ space for $\Omega^2$ with $c_1 = 1$ (cyan), $c_1 = 7$ (yellow) for left plot, $c_1 = -1$ (cyan) and $c_1 = -7$ (yellow) for right plot.

## 3. Concluding Remarks

This paper studies the stability of EU with closed FRW universe model and perfect fluid in $f(R, T, Q)$ gravity. We have formulated static and perturbed field equations using scalar homogeneous perturbations about energy density and scale factor. These equations are parameterized by linear EoS parameter. The second order perturbed differential equations are constructed whose solutions provide the

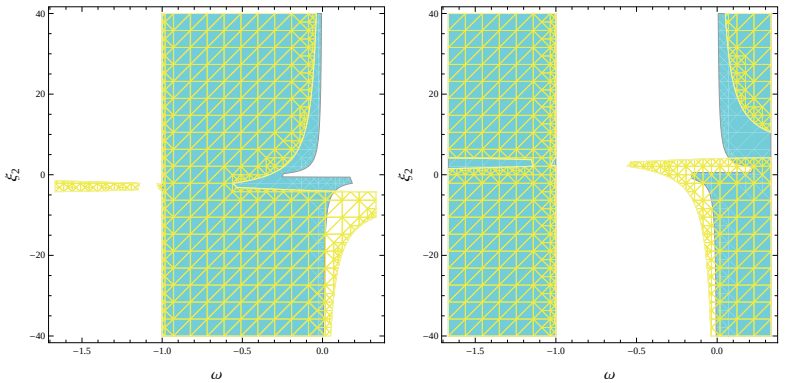



Fig. 2. Plots of stable regions in $(\omega, \xi_2)$ space for $\bar{\Omega}^2$ with $\alpha = 1$ (cyan), $\alpha = 7$ (yellow) for right plot, $\alpha = -1$ (cyan) and $\alpha = -7$ (yellow) for left plot.

existence and stability regions of EU for particular $f(R, T, Q)$ models. We have analyzed both conserved as well as non-conserved EMT cases against perturbations scheme. We conclude that stable modes of EU exist against homogeneous scalar perturbations for all values of EoS parameter if the model constraints are chosen appropriately in this gravity. The stable EU against vector perturbations also exist because any initial vector perturbations remain frozen. It is worthwhile to mention here that the range of EoS parameter is greatly enhanced as compared to that of $f(R)$ gravity and more stable regions are found as compared to $f(R, T)$ theory due to the presence of generic function of $Q$. Like all other modified theories, our results also reduce to GR in the absence of dark source terms.

## Acknowledgments

We would like to thank the Higher Education Commission, Islamabad, Pakistan for its financial support through the *Indigenous Ph.D. Fellowship, Phase-II, Batch-III.*

## References

1. Z. Haghani, et al., *Phys. Rev. D* **88**, 044023 (2013).
2. S.D. Odintsov and D. Sáez-Gómes, *Phys. Lett. B* **725**, 437 (2013).
3. M. Sharif and M. Zubair, *J. High Energy Phys.* **12**, 079 (2013).
4. M. Sharif and M. Zubair, *J. Cosmol. Astropart. Phys.* **11**, 042 (2013).
5. M. Sharif and A. Waseem, *Eur. Phys. J. Plus* **131**, 190 (2016).
6. M. Sharif and A. Waseem, *Eur. Phys. J. Plus* **133**, 136 (2018).
7. A.S. Eddington, *Mon. Not. R. Astron. Soc.* **90**, 668 (1930).
8. S.S. Seahra and C.G. Böhmer, *Phys. Rev. D* **79**, 064009 (2009).
9. H. Shabani and A.H. Ziaie, *Eur. Phys. J. C* **77**, 31 (2017).
10. M. Sharif and A. Ikram, *Int. J. Mod. Phys. D* **26**, 1750084 (2017).

# Collapsing and expanding solutions in $f(R,T)$ gravity

Sharif M. * and Siddiqa A.

*Department of Mathematics, University of the Punjab,*
*Lahore, 54590, Pakistan*
**E-mail: msharif.math@pu.edu.pk*

We discuss anisotropic source describing the phenomena of collapse and expansion in the context of $f(R,T)$ theory. For this purpose, we take an auxiliary solution of the Einstein field equations and evaluate expansion scalar whose negative values lead to collapse and positive values give expansion. For both cases, the behavior of density, pressures as well as anisotropic parameter is explored and the effects of model parameter on these quantities are examined. We also check the energy conditions for physical viability of these generating solutions.

*Keywords*: $f(R,T)$ gravity; Self-gravitating objects.

## 1. Introduction

The force of gravity is responsible for maintaining the structure of celestial bodies, restricting them in their respective tracks as well as controlling different astrophysical phenomena. Stars are one of these heavenly objects that are in equilibrium state due to balance of inward gravity and outward thermal pressure. When this balance is disturbed, star first experiences expansion and becomes red giant or massive stars lead to red super giants. After this expansion phase, a supernova explosion occurs and star collapses resulting into compact objects like white dwarfs, neutron stars and black holes. Oppenheimer and Snyder[1] are the pioneers to explore the gravitational collapse of dust matter. Glass[2] investigated the collapsing and expanding solutions for a spherical system within the framework of general relativity (GR). Abbas explored these solutions for plane symmetric stellar structure[3], spherical symmetric charged system[4] as well as cylindrical charged fluid configurations[5] in GR. These collapsing and expanding solutions are also generated in the background of $f(R,T)$ gravity for a sphere in the absence[6] and presence of electromagnetic field.

To discuss the mysteries of dark energy and dark matter, researchers proposed different extensions or alternatives to GR. Here, we consider the $f(R,T)$ theory which couples curvature scalar and trace of energy momentum tensor. This manuscript generates the collapsing and expanding solutions for a plane symmetric anisotropic source in $f(R,T)$ context. In next section, we define the basic equations for a plane symmetric stellar structure in $f(R,T)$ theory and its subsections contain the generating solution for collapse as well as expansion. Finally, we summarize our results.

## 2. Basic Equations in $f(R,T)$ Theory

We consider the following metric and energy-momentum tensor to describe the geometry and matter distribution inside the anisotropic plane symmetric source,

respectively

$$ds^2 = -A^2(t,z)dt^2 + B^2(t,z)(dx^2+dy^2) + C^2(t,z)dz^2, \tag{1}$$

$$T_{\mu\nu} = (\rho + p_\perp)V_\mu V_\nu + p_\perp g_{\mu\nu} + (p_z - p_\perp)\chi_\mu\chi_\nu, \tag{2}$$

where $V_\mu$, $\chi_\mu$, $\rho$, $p_z$ and $p_\perp$ denote the four velocity, unit four-vector, density as well as pressures in $z$ and its perpendicular direction, respectively. Here, the four-vectors $V_\mu$ and $\chi_\mu$ have the following expressions

$$V^\mu = (A^{-1},0,0,0), \quad K^\mu = (0,0,0,C^{-1}),$$
$$\text{and} \quad V^\mu V_\nu = -1,\ \chi^\mu\chi_\nu = 1,\ V^\mu\chi_\nu = 0.$$

The $f(R,T)$ gravity action

$$S = \int d^4x\sqrt{-g}\left[\frac{1}{16\pi}f(R,T) + \mathcal{L}_m\right], \tag{3}$$

produces the following field equations

$$f_R R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}f + (g_{\mu\nu}\Box - \nabla_\mu\nabla_\nu)f_R = 8\pi T_{\mu\nu} - f_T(T_{\mu\nu}+\Theta_{\mu\nu}) + 8\pi E_{\mu\nu}, \tag{4}$$

where $f_R$ and $f_T$ denote the derivatives of $f(R,T)$ with respect to $R$ and $T$, respectively, while $\Theta_{\mu\nu}$ is defined as

$$\Theta_{\mu\nu} = g^{\gamma\alpha}\frac{\delta T_{\gamma\alpha}}{\delta g^{\mu\nu}}, \quad T_{\mu\nu} = g_{\mu\nu}\mathcal{L}_m - \frac{\partial\mathcal{L}_m}{\partial g^{\mu\nu}}, \tag{5}$$

We consider $f(R,T) = R + 2\lambda T$ suggested by Harko *et al.*[7] to investigate how curvature-matter coupling effects the generating solutions for collapse and expansion. This $f(R,T)$ model and $\mathcal{L}_m = -\rho$ yield the following field equations

$$G_{\mu\nu} = (8\pi + 2\lambda)T_{\mu\nu} + 2\lambda\rho g_{\mu\nu} + \lambda T g_{\mu\nu}, \tag{6}$$

which produces the following set of equations

$$\frac{\dot{B}^2}{B^2A^2} - \frac{B'^2}{C^2B^2} - 2\frac{B''}{BC^2} + 2\frac{\dot{C}}{C}\frac{B}{BA^2} + 2\frac{B'}{B}\frac{C'}{C^3} = (8\pi+\lambda)\rho - 2\lambda p_\perp - \lambda p_z, \tag{7}$$

$$\frac{-\ddot{B}}{BA^2} + \frac{\dot{A}\dot{B}}{A^3B^2} + \frac{A'B'}{C^2AB^2} + \frac{B''}{B^2C^2} - \frac{\dot{C}\dot{B}}{CB^2A^2} - \frac{B'C'}{B^2C^3} + \frac{A''}{ABC^2} - \frac{\ddot{C}}{CA^2B}$$
$$+\frac{\dot{A}\dot{C}}{A^2CB} - \frac{A'C'}{ABC^2} = (8\pi+4\lambda)p_\perp + \lambda\rho + \lambda p_z, \tag{8}$$

$$\frac{-2\ddot{B}}{BA^2} + \frac{2\dot{A}\dot{B}}{A^3B} + 2\frac{A'B'}{ABC^2} - \frac{\dot{B}^2}{A^2B^2} + \frac{B'^2}{B^2C^2} = (8\pi+3\lambda)p_z + \lambda\rho + 2\lambda p_\perp, \tag{9}$$

where an over dot and prime correspond to the derivative with respect to $t$ and $z$, respectively. Simultaneous solution of field equations gives the following values of

density and pressures

$$\rho = \frac{A^{-3}B^{-2}C^{-3}}{8(2\pi+\lambda)(4\pi+\lambda)}\Big[-AC\Big\{(8\pi+5\lambda)(A^2C^2\dot{B}^2+A^4B'^2)+\lambda C^4(\dot{B}^2+B'^2)\Big\}$$
$$+\ 2\lambda B^4\Big\{C^2(\dot{A}\dot{C}-A\ddot{C})-A^2A'C'+A^2CA''\Big\}+2B\Big\{(8\pi+\lambda)A^3C^2\dot{B}\dot{C}+\lambda C^5$$
$$+\ \Big(\dot{A}\dot{B}-A\ddot{B}-A'B')+(8\pi+5\lambda)(A^5B'C'A^5CB'')\Big\}+2\lambda B^3\Big\{-AC^2\dot{B}\dot{C}$$
$$+\ +C^3(\dot{A}\dot{B}-A\ddot{B})-A^3B'C'+A^2C(A'B'+AB'')\Big\}\Big], \tag{10}$$

$$p_z = \frac{A^{-3}B^{-2}C^{-3}}{8(2\pi+\lambda)(4\pi+\lambda)}[AC\{-\lambda A^2C^2\dot{B}^2+\lambda A^4B'^2-(8\pi+3\lambda)C^4(\dot{B}^2+B'^2)\}$$
$$-\ 2\lambda B^4\{C^2(\dot{A}\dot{C}-A\ddot{C})-A^2A'C'+A^2CA''\}+2B\{-\lambda A^3C^2\dot{B}\dot{C}+(8\pi+3\lambda)$$
$$\times\ C^5(\dot{A}\dot{B}-A\ddot{B}-A'B')-\lambda A^5B'C'+\lambda A^5CB''-2\lambda B^3\{-AC^2\dot{B}\dot{C}+C^3$$
$$\times\ (\dot{A}\dot{B}-A\ddot{B})-A^3B'C'+A^2C(A'B'+AB'')\}\}], \tag{11}$$

$$p_\perp = \frac{A^{-3}B^{-2}C^{-3}}{8(2\pi+\lambda)(4\pi+\lambda)}[AC\{-\lambda A^2C^2\dot{B}^2+A^4B'^2+C^4(\dot{B}^2+B'^2)\}+(4\pi+\lambda)$$
$$\times\ B^4\{C^2(\dot{A}\dot{C}-A\ddot{C})-A^2A'C'+A^2CA''\}-2\lambda B\{A^3C^2\dot{B}\dot{C}+C^5(\dot{A}\dot{B}$$
$$-\ A\ddot{B}-A'B')+A^5B'C'-A^5C\ddot{B}\}+2(4\pi+\lambda)B^3\{-AC^2\dot{B}\dot{C}+C^3(\dot{A}\dot{B}$$
$$-\ A\ddot{B})-A^3\dot{B}\dot{C}+A^2C(A'B'+AB'')\}]. \tag{12}$$

The pressure anisotropy is defined by the relation

$$\triangle = p_z - p_\perp. \tag{13}$$

The expansion scalar is obtained as

$$\Theta = \frac{1}{A}\left(\frac{2\dot{B}}{B}+\frac{\dot{C}}{C}\right), \tag{14}$$

and we consider an auxiliary solution of Eq.(8) as

$$A = \frac{\alpha\dot{B}}{B^\gamma}, \quad C = B^\gamma, \tag{15}$$

where $\gamma$ and $\alpha > 0$ are arbitrary constants. The above solution leads to

$$\Theta = (2+\gamma)B^{1-\gamma}. \tag{16}$$

The positive values of $\Theta$ provide expansion and its negative values correspond to collapse. We discuss these cases in the following subsections.

### 2.1. *Generating solution for collapse*

For collapsing solution, we find the metric coefficient $B$ using the Taub's mass[8] for plane symmetric configuration and then using the concept of trapped surfaces (that

whether the collapse lead to the formation of trapped surfaces or not). The mass function is evaluated as

$$m(t,z)=\frac{(g_{11})^{\frac{3}{2}}}{2}R_{12}^{12}=\frac{B}{2}\left(B^{2\gamma}-\frac{B'^2}{B^{2\gamma}}\right). \tag{17}$$

For $B'=B^{2\gamma}$, it has been discussed that no trapping surfaces are found for a plane symmetric object experiencing collapse[3]. Hence we have

$$B_{col}=[(1-2\gamma)z+g(t)]^{\frac{1}{1-2\gamma}}, \tag{18}$$

with $g(t)$ an integration function. This expression of $B_{col}$ in turn yields the following values of $A$ and $C$ (we take $\alpha=1$)

$$A_{col}=\frac{\dot{g(t)}}{1-2\gamma}[(1-2\gamma)z+g(t)]^{\frac{\gamma}{1-2\gamma}}, \tag{19}$$

$$C_{col}=[(1-2\gamma)z+g(t)]^{\frac{\gamma}{1-2\gamma}}. \tag{20}$$

The corresponding values of density and pressures are obtained by substituting this solution in Eqs.(10)-(12). For graphical analysis, we chose $g(t)=t$ and the analysis is summarized in Table **1**. The graphs of $\rho$ and $p_z$, for different values of $\lambda$ are shown in Figure **1**. For collapse solution, there is no change in different parameters with respect to time.

Table **1**: Change in parameters with respect to $r$, $t$ and $\lambda$ for the collapsing solution.

| **Parameter** | **As $z$ increases** | **As $\lambda$ increases** |
|---|---|---|
| $\rho$ | positively decreases | positively decreases |
| $p_z$ | positively decreases | positively increases |
| $p_\perp$ | negatively increases | negatively increases |
| $\triangle$ | positively decreases | positively decreases |

### 2.2. *Generating solution for expansion*

In order to generate a solution for an expanding plane symmetric configuration, we need the value of $B$ that leads to expansion. We assume that it is a sum of temporal and radial components and choose the values of free parameters such that $\Theta$ remains positive for these values. Thus we have the following expanding solution

$$A_{exp}=\frac{\alpha}{(r+t)^\gamma},\quad B_{exp}=\alpha(r+t),\quad C_{exp}=\alpha(r+t)^\gamma. \tag{21}$$

After replacing these values in Eqs.(10)-(12), we find the expressions of density and pressures. The plots of $\rho$ and $p_z$ are given in Figure **2** as well as the graphical analysis is summarized in Table **2**. For the expanding solution, the physical parameters also vary with varying time.

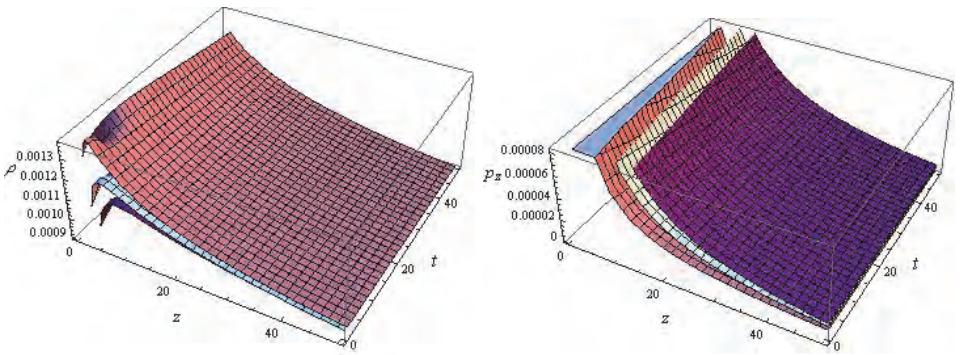



Fig. 1. Plots of $\rho$ and $p_z$ versus $r$ and $t$ for $\gamma = -5$, $\alpha = 1$, $\lambda = 0.3$ (pink), $\lambda = 0.6$ (blue), $\lambda = 0.9$ (purple).

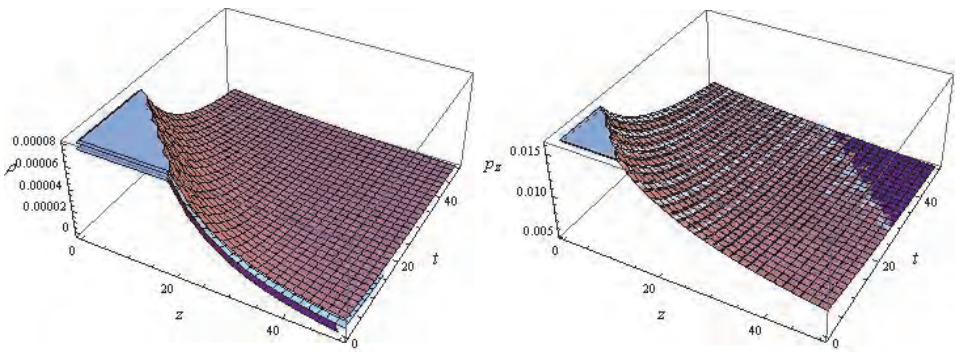



Fig. 2. Plots of $\rho$ and $p_z$ versus $r$ and $t$ for $\gamma = 0.08$, $\alpha = 1$, $\lambda = 0.01$ (pink), $\lambda = 0.02$ (blue), $\lambda = 0.03$ (purple).

Table **2**: Change in parameters with respect to $r$, $t$ and $\lambda$ for the expanding solution.

| **Parameter** | **As $z$ increases** | **As $t$ increases** | **As $\lambda$ increases** |
|---|---|---|---|
| $\rho$ | positively decreases | positively decreases | positively decreases |
| $p_z$ | positively decreases | positively decreases | positively decreases |
| $p_\perp$ | negatively decreases | negatively increases | negatively increases |
| $\triangle$ | positively decreases | positively decreases | positively decreases |

## 3. Summary

The expanding universe can affect different phenomena happening in the universe one of which is the life of a star. In this paper, we investigate the generating solutions for collapsing and expanding phases during stellar evolution and consider

an anisotropic plane symmetric source in the context of $f(R,T)$ theory. Firstly, we assume a solution in the form of metric coefficient $B$ and then find it to generate a solution for collapse and expansion. We also investigate the role of model parameter $\lambda$ on different quantities. It is found that the energy conditions are satisfied in both solutions for chosen values of parameters implying the validity of solutions.

**Acknowledgments**

We would like to thank the Higher Education Commission, Islamabad, Pakistan for its financial support through the *Indigenous Ph.D. 5000 Fellowship Program Phase-II, Batch-III.*

**References**

1. J.R. Oppenheimer and H. Snyder, *Phys. Rev.* **56**, 455 (1939).
2. E.N. Glass, Gen. Relativ. Gravit. **45**, 266 (2013).
3. G. Abbas, *Astrophys. Space Sci.* **350**, 307 (2014).
4. G. Abbas, *Astrophys. Space Sci.* **352**, 955 (2014).
5. G. Abbas, *Astrophys. Space Sci.* **357**, 56 (2015).
6. G. Abbas, and R. Ahmed, *Eur. Phys. J. C* **77**, 441 (2017).
7. T. Harko, F.S.N. Lobo, S. Nojiri, and S.D. Odintsov, *Phys. Rev. D* **84** 024020 (2011).
8. T. Zannias, *Phys. Rev. D* **41**, 3252 (1990).

# A classical and quantum tomographic analysis of the de Sitter universe

C. Stornaiolo*

*INFN-Sezione di Napoli, Complesso Universitario di Monte S. Angelo, Ed. 6, via Cinzia 80126 Napoli, Italy*
*E-mail: * cosmo@na.infn.it*

In this work we reformulate classical and quantum cosmology in terms of a tomographic description. This approach enables us to describe the quantum and classical states of the universe in the same way and allows us to examine some aspects of the transition from the initial quantum stage to the next classical one, which are not evident in the other formulations of quantum cosmology. For reasons of simplicity and because they allow us to very effectively illustrate some aspects of the quantum-classical transition, we have considered the Hartle-Hawking model for a de Sitter universe and the corresponding classical model. The analysis obtained can also be extended to the tunneling from nothing of Vilenkin-LInde, provided that some restrictions are made in the models. We demonstrate that the Hartle-Hawking model does not have a classic limit.

*Keywords*: Quantum cosmology; Quantum tomography.

*1. Introduction.* Tomography was originally introduced for clinical purposes as an imaging technique to create three-dimensional images of the interior of the human body from two-dimensional sections obtained by X-rays. Its application was subsequently extended to many other fields of scientific research. In particular the same principles were applied to quantum tomography in order to reconstruct the state of a quantum system. In fact the tomographic analysis of a system either in quantum optics as in quantum mechanics was used to reconstruct the Wigner function, which is a distribution function on the phase space. Similarly, a classical probability function in the phase space can be obtained through a tomographic analysis.[1, 2] This shows that classical and quantum states can both be described by a tomographic approach. This motivates the extension of the tomographic approach to quantum cosmology.

Introduction of quantum cosmology (see e.g.[3][4]) is motivated by the necessity to address the conceptual problems introduced by the prediction in classical general relativity of an initial singularity which is unacceptable from a physical point of view because the knowledge of an initial state enables us to predict the future evolution of a classical system, otherwise we can establish the probabilities of the evolution of a quantum system. With quantum cosmology we substitute this indefinite classical state with a quantum state. This is a plausible working hypothesis because distance scales are extremely small and the energy densities are extremely high at times close to the initial singularity. But the theoretical approach to quantum cosmology is quite different from that of ordinary quantum systems, because for the latter the initial conditions are imposed with respect to an external environment that does not exist in cosmology and they can be varied at will, while in cosmology the initial conditions become a physical law to all effects, because in this context we are considering just one and only one "experiment". Quantum cosmology is formulated as

the restriction of quantum gravity to homogeneous space-times; with this restriction a theory very close to quantum mechanics emerges, where the Wheeler-de Witt equation takes the place of the Schrödinger equation.

The importance of applying quantum tomography to quantum cosmology is that through the tomographic representation both classical and quantum states can be represented in the same way allowing us to compare them and eventually to study the quantum-classical transition for a system like the universe. In this paper we determine the tomogram[5] of the Hartle-Hawking initial state and compare its classical limit with the classical tomogram.

*2. Tomograms* In quantum mechanics where, due to the uncertainty principle, the state of a system is specified in terms of probability distributions by introducing a wave function or the density matrix. A quantum state can be also represented in the phase space, in analogy with classical mechanics, by the Wigner function[6] defined by the non linear transform of the wave function $\psi(x)$,

$$W(x,p) = \frac{1}{2\pi}\int \psi\left(x+\frac{u}{2}\right)\psi^*\left(x-\frac{u}{2}\right)\mathrm{e}^{-i\frac{up}{\hbar}}du\,. \tag{1}$$

and if non pure states are considered, its definition can be extended to any matrix density by

$$W(x,p) = \frac{1}{2\pi}\int \rho\left(x+\frac{u}{2},x-\frac{u}{2}\right)\mathrm{e}^{-i\frac{up}{\hbar}}du\,. \tag{2}$$

The Wigner function has to satisfy the normalization condition,

$$\int W(q,p)dq\,dp = 1. \tag{3}$$

Similarly in quantum cosmology alongside the concepts of wave function and density matrix the Wigner function was introduced with the purpose to better understand the problem of the classical transition of a quantum universe.[7, 8] Unfortunately, these representations can hardly be related directly to the observational data.

Recently the notion of standard positive probability distribution function (tomogram) was introduced in cosmology to describe the quantum state of universe as an alternative to wave function or to density matrix.[5, 9] The tomogram can be defined as the modified Radon transform, of the Wigner function

$$\mathcal{W}(X,\mu,\nu) = \int W(q,p)\delta(X-\mu x-\nu p)\frac{dxdp}{\hbar}\,. \tag{4}$$

It depends on the random variable $X$ given by the linear combination of position and momentum:

$$X = \mu x + \nu p \tag{5}$$

where $\mu = s\cos\theta$ and $\nu = s^{-1}\sin\theta$ parameterize the rotation and squeezing of the reference frames of the phase space through a set of linear canonical transformations. So, we say that the quantum state is given if the position probability distribution $\mathcal{W}(X,\mu,\nu)$ in an ensemble of rotated and squeezed reference frames in the classical

phase space is given. From Eqs. (1) and (4) it follows that we can derive the tomogram from the wave function $\psi(x)$ which is

$$\mathcal{W}(X,\mu,\nu) = \frac{1}{2\pi|\nu|\hbar}\left|\int \psi(y)\exp\left[i\left(\frac{\mu}{2\hbar\nu}y^2 - \frac{X}{\hbar\nu}y\right)\right]dy\right|^2 . \tag{6}$$

Finally when $\mu = 0$ and $\nu = 1$ the tomographic amplitude reduces to a Fourier transform from $x$ to $X \equiv p$, and is proportional to the density $|\psi(p)|^2$ in the $p$ representation.

The symplectic tomogram $\mathcal{W}(X,\mu,\nu)$ has the following properties of a) non negativity, i.e. $\mathcal{W}(X,\mu,\nu) \geq 0$ and b) normalization $\int \mathcal{W}(X,\mu,\nu)dX = 1.$ , which characterizes it as a marginal probability function. For more details see e.g. ref. [5].

*3. Classical symplectic tomography* The definition of tomogram (4) can be extended to any distribution on the phase space. Therefore we apply this definition to a classical distribution $f(q,p)$, for instance a solution of the Boltzmann equation,

$$\mathcal{W}(X,\mu,\nu) = \int f(x,p)\delta(X - \mu x - \nu p)dxdp\,. \tag{7}$$

If the classical probability distribution $f(q,p)$ is normalized also the tomogram is normalized and satisfies the following conditions,

$$\int \mathcal{W}(X,\mu,\nu)\,dX = 1 \tag{8}$$

The (initial) state of a simple particle is described the distribution function

$$f(q,p) = \delta(x - x_0)\delta(p - p_0) \tag{9}$$

we obtain by eq. (7) the tomogram of a free particle

$$\mathcal{W}(X,\mu,\nu) = \delta(X - \mu x_0 - \nu p_0) \tag{10}$$

where for a moving particle the time dependent with distribution function

$$f(q,p,t) = \delta(x - x(t))\delta(p - p(t)) \tag{11}$$

the corresponding tomogram is

$$\mathcal{W}(X,\mu,\nu) = \delta\left(X - \mu x(t) - \nu p(t)\right)\,. \tag{12}$$

*3. The de Sitter model in General Relativity.* Let us determine the tomogram that describes the classical state of the de Sitter universe where the cosmological constant is responsible for the universe expansion even in the absence of material fields. We consider a closed homogeneous and isotropic model with a cosmological constant $\Lambda$. Using the notations of [7] [8], it is described by the metric

$$ds^2 = l_p^2\left[-N^2(t)dt^2 + a^2(t)d\Omega_3^2\right] \tag{13}$$

where $a(t)$ is the expansion factor of the universe, $d\Omega_3^2$ is the metric of the unit three-sphere $N(t)$ is the lapse function, $l_p = 2/3\ell_p$, with $\ell_p$ the Planck length. Introducing the phase space variable $q = a^2$ and Hamiltonian

$$\mathcal{H} = \frac{1}{2}\left(-4p^2 + \lambda q - 1\right) \tag{14}$$

where $\lambda$ is the cosmological constant in Planck units. The constraint $\mathcal{H} = 0$ implies that position and momentum are not independent. It corresponds to a trajectory on the phase space. At any fixed time the distribution on the phase space reduces to just one point of this trajectory. Therefore the distribution function is the Dirac delta of the constraint,

$$f(q,p) = \delta\left(-4p^2 + \lambda q - 1\right) \tag{15}$$

From Eqs. (7) and (15), the classical tomogram is

$$\mathcal{W}(X,\mu,\nu) = \frac{1}{2|\mu|}\frac{1}{\left|\sqrt{\frac{\lambda^2\nu^2}{16\mu^2} + \frac{\lambda X}{\mu}} - 1\right|}\,. \tag{16}$$

As the function in Eq. (16 )is not integrable on $-\infty$ to $+\infty$, we cannot normalize the tomogram on this interval. The tomogram is a probability function and it has to satisfy the normalization condition. Therefore we impose that function (16) is different from zero only on the interval $C = \left[\frac{\mu}{\lambda}\left(1 - \frac{\lambda^2\nu^2}{16\mu^2}\right), \lambda\mu + \frac{\mu}{\lambda}\left(1 - \frac{\lambda^2\nu^2}{16\mu^2}\right)\right]$. Finally we see that if the cosmological constant is of the order of magnitude of $10^{-122}$, which is in Planck units the value nowadays accepted, the interval $C$ becomes very narrow. in other words the very high degree of homogeneity of the classical (de Sitter) universe is a consequence of the extreme smallness of the cosmological constant.

*4. The de Sitter quantum cosmological models.* In quantum cosmology the Hamiltonian constraint becomes an equation, in the de Sitter it is[10]

$$\left(4\hbar^2\frac{d^2}{dq^2} + \lambda q - 1\right)\psi(q) = 0\,. \tag{17}$$

Equation (17) becomes the Airy equation[13, 14]

All the solutions of the Airy equation are linear combinations of $\mathrm{Ai}(x)$ and $Bi(x)$. The most interesting are those which can be evolved to a classical universe with the properties actually observed in the present time. In this work we recall the Hartle and Hawking proposal,[11] we mention that there are also the models related to the "no-boundary condition" and the Vilenkin-Linde "tunneling of the universe from nothing beginning ".[12] The wave functions corresponding to all these initial conditions in a de Sitter (quantum) universe were derived originally by Halliwell and Louko[10] . This models where analyzed in[10] and extend the discussion of these models to the Wigner functions, see[7, 8] . Here we will consider just the Hartle and Hawking wave function,

$$\psi_{HH} = A\,\mathrm{Ai}\left(\frac{1-\lambda q}{(2\hbar\lambda)^{2/3}}\right), \tag{18}$$

because only with this wave function we can express the corresponding s tomogram in an analytic form allowing us to compare it with the classical tomogram (16).

To derive the tomogram the Hartle-Hawking tomogram we insert the wave function (18) in (6) and obtain

$$\mathcal{W}(X,\mu,\nu)=\frac{A^2}{|\mu|}\left|\mathrm{Ai}\left(\frac{1}{(2\hbar\lambda)^{2/3}}\left(1-\frac{\lambda X}{\mu}-\frac{\lambda^2}{16}\frac{\nu^2}{\mu^2}\right)\right)\right|^2. \tag{19}$$

Then we calculate its classical limit by taking the limit $\hbar\to 0$; indeed we consider the limit $(2\hbar\lambda)^{2/3}\to 0$ by keeping $\lambda$ fixed, as $(2\hbar\lambda)^{2/3}$ approaches zero the argument goes to infinite and using the Airy function can be substituted by its asymptotic expressions[13, 14] , in particular we notice that the argument of the Airy function in (18) is the negative of the argument of (16), so we are interested in

$$\mathcal{W}(X,\mu,\nu)\approx\frac{A^2}{8\pi^2\hbar|\mu|}\frac{(2\hbar\lambda)^{4/3}}{\left|1-\frac{\lambda X}{\mu}-\frac{\lambda^2}{16}\frac{\nu^2}{\mu^2}\right|^{1/2}}\times\left|\cos\left(\frac{2}{3}S^{3/2}-\frac{\pi}{4}\right)\right|^2, \tag{20}$$

$$\tag{21}$$

when its argument is negative.[13, 14] When its argument is positive the asymptotic expression of (19) vanishes in this limit giving no contribution to the tomogram.

The normalization constant $A$ is fixed by imposing that the quantum and classical tomograms have the same coefficient, so $A=\frac{2^{5/6}\pi}{\hbar^{1/6}\lambda^{2/3}}$.

We notice that the one over square root factor of (20) coincides with the classical tomogram, but the square modulus of the cosine does not converge when its argument goes to infinity. We conclude that the Hartle and Hawking model does not have a classical limit. Similarly we cannot find a classical limit for the Vilenkin-Linde solutions because these functions are not integrable on the real line, so that we cannot obtain a tomogram at least we do not try to modify them in some way. Finally we notice that the classical limit of the tomogram, whose expression incidentally is similar to Vilenkin's wave function,

$$\mathcal{W}(X,\mu,\nu)=\frac{A^2}{|\mu|}\left|\frac{1}{2}\left[Ai\left(\frac{1}{(2\hbar\lambda)^{2/3}}\left(1-\frac{\lambda X}{\mu}-\frac{\lambda^2}{16}\frac{\nu^2}{\mu^2}\right)\right)\right.\right. \tag{22}$$

$$\left.\left.+i\mathrm{Bi}\left(\frac{1}{(2\hbar\lambda)^{2/3}}\left(1-\frac{\lambda X}{\mu}-\frac{\lambda^2}{16}\frac{\nu^2}{\mu^2}\right)\right)\right]\right|^2. \tag{23}$$

coincides with (16) therefore It represents a quantum universe that converges to the classical de Sitter universe.

## References


1. V. I. Man'ko, "Conventional quantum mechanics without wave function and density matrix," AIP Conf. Proc. **464** (1999) no.1, 191 [quant-ph/9902079].
2. A. Ibort, V. I. Man'ko, G. Marmo, A. Simoni and F. Ventriglia, "An Introduction to the tomographic picture of quantum mechanics," Phys. Scripta **79** (2009) 065013

3. D. L. Wiltshire, "An Introduction to quantum cosmology," gr-qc/0101003.
4. J. J. Halliwell, "Introductory Lectures On Quantum Cosmology," In *Jerusalem 1989, Proceedings, Quantum cosmology and baby universes* 159-243 and MIT Cambridge - CTP-1845 (90,rec.May) 108 p
5. C. Stornaiolo, "Tomographic analysis of the De Sitter model in quantum and classical cosmology," Int. J. Geom. Meth. Mod. Phys. **16** (2018) no.01, 1950012
6. E.P. Wigner, "On the quantum correction for thermodynamic equilibrium," Phys. Rev. **40** (1932) 749.
7. S. Habib, "The classical limit in quantum cosmology. 1 Quantum mechanics and the Wigner function," Phys. Rev. D **42** (1990) 2566; S. Habib and R. Laflamme, "Wigner function and decoherence in quantum cosmology," Phys. Rev. D **42** (1990) 4056.
8. R. Cordero, H. Garcia-Compean and F. J. Turrubiates, "Deformation quantization of cosmological models," Phys. Rev. D **83** (2011) 125030
9. V. I. Manko, G. Marmo and C. Stornaiolo, "Radon transform of Wheeler-de Witt equation and tomography of quantum states of the universe," Gen. Rel. Grav. **37** (2005) 99; V. I. Man'ko, G. Marmo and C. Stornaiolo, "Cosmological dynamics in tomographic probability representation," Gen. Rel. Grav. **37** (2005) 2003; V. I. Man'ko, G. Marmo and C. Stornaiolo, "Tomographic entropy and cosmology," Gen. Rel. Grav. **40** (2008) 1449; S. Capozziello, V. I. Man'ko, G. Marmo and C. Stornaiolo, "Tomographic Representation of Minisuperspace Quantum Cosmology and Noether Symmetries," Gen. Rel. Grav. **40** (2008) 2627; S. Capozziello, V. I. Man'ko, G. Marmo and C. Stornaiolo, "A Tomographic description for classical and quantum cosmological perturbations," Phys. Scripta **80** (2009) 045901; C. Stornaiolo, "Tomographic Represention of Quantum and Classical Cosmology," Astrophys. Space Sci. Proc. **38** (2014) 3, 211; C. Stornaiolo, "Tomographic cosmology," Phys. Scripta **90** (2015) 7, 074032.
10. J. J. Halliwell and J. Louko, "Steepest Descent Contours in the Path Integral Approach to Quantum Cosmology. 1. The De Sitter Minisuperspace Model," Phys. Rev. D **39** (1989) 2206.
11. J. B. Hartle and S. W. Hawking, "Wave Function of the Universe," Phys. Rev. D **28** (1983) 2960 [Adv. Ser. Astrophys. Cosmol. **3** (1987) 174].
12. A. Vilenkin, "Boundary Conditions in Quantum Cosmology," Phys. Rev. D **33** (1986) 3560; A. Vilenkin, "Quantum Cosmology and the Initial State of the Universe," Phys. Rev. D **37** (1988) 888; A. D. Linde, "Quantum Creation of the Inflationary Universe," Lett. Nuovo Cim. **39** (1984) 401. doi:10.1007/BF02790571
13. Abramowitz, M. and Stegun, I. (1965). Handbook of Mathematical Functions, Dover Publications, New York.
14. O. Vallee, M. Soares, Airy Functions and Applications to Physics, Imperial College Press, (2004)

# Spherically symmetric solutions with any cosmological constant in the Einstein-Weyl gravity

Robert Švarc* and Jiří Podolský†

*Institute of Theoretical Physics, Faculty of Mathematics and Physics, Charles University, Prague, V Holešovičkách 2, 180 00 Praha 8, Czech Republic*
*E-mail: *robert.svarc@mff.cuni.cz, †podolsky@mbox.troja.mff.cuni.cz*
*www.mff.cuni.cz*

Vojtěch Pravda‡ and Alena Pravdová◇

*Institute of Mathematics of the Czech Academy of Sciences, Žitná 25, 115 67 Prague 1, Czech Republic*
*E-mail: ‡pravda@math.cas.cz, ◇pravdova@math.cas.cz*
*www.math.cas.cz*

In our contribution we present spherically symmetric solutions in four dimensions to the Einstein-Weyl (and also general quadratic) gravity admitting arbitrary value of the cosmological constant. The main attention is paid to the black hole spacetimes representing one parameter extension of the well-known Schwarzschild-(anti-)de Sitter geometry of Einstein's general relativity. This additional parameter corresponds to the non-vanishing value of the Bach tensor at the black hole horizon. Our work thus generalises recent paper by Lü et al. (Phys. Rev. Lett. 114 (2015) 171) to any value of the cosmological constant. Moreover, using more convenient metric ansatz, the field equations form an autonomous system and their solution is explicitly obtained in an exact form of power series. The physical interpretation of these geometries is also discussed, namely we investigate specific tidal effects on free test particles or their basic thermodynamic properties.

*Keywords*: Einstein-Weyl theory; Quadratic gravity; Black holes; Bach tensor.

## 1. Introduction

Black holes represent one of the most important theoretical predictions of the Einstein general relativity. Almost immediately after its formulation Karl Schwarzschild[1] found a static spherically symmetric solution to the field equations. However, its 'black-hole nature' was understood many years later. In our contribution, we are interested in the extension of such a fundamental solution in the framework of the so called quadratic gravity[2,3] (including the Einstein–Weyl theory). This paper is based on our recent letters[4,5], where the main aim was to simplify the field equations, to find their explicit solution in a closed form, and to admit any value of the cosmological constant.

The most general four-dimensional quadratic gravity in vacuum can be introduced via the Einstein-Hilbert action containing the Ricci scalar $R$, the cosmological constant $\Lambda$, and a contraction of the Weyl tensor $C_{abcd}$,

$$S = \int \mathrm{d}^4x\,\sqrt{-g}\,\Big(\gamma\,(R-2\Lambda) + \beta\,R^2 - \alpha\,C_{abcd}\,C^{abcd}\Big)\,, \tag{1}$$

where $\alpha$, $\beta$, and $\gamma = G^{-1}$ are constants. Here we employ a natural simplifying assumption of the constant scalar curvature, namely

$$R = \text{const.}\,, \qquad \Rightarrow \qquad R = 4\Lambda\,, \tag{2}$$

under which the field equations become

$$R_{ab} - \Lambda\, g_{ab} = 4k\, B_{ab}\,, \qquad \text{with} \qquad k \equiv \frac{\alpha}{\gamma + 8\beta\Lambda}\,, \tag{3}$$

where $B_{ab} \equiv \big(\nabla^c\nabla^d + \frac{1}{2}R^{cd}\big)C_{acbd}$ is the traceless, symmetric and conserved *Bach tensor* ($g^{ab}B_{ab} = 0$, $B_{ab} = B_{ba}$, $B_{ab}{}^{;b} = 0$).

## 2. The spacetime geometry

A spherically symmetric spacetime is usually presented as

$$\mathrm{d}s^2 = -h(\bar{r})\,\mathrm{d}t^2 + \frac{\mathrm{d}\bar{r}^2}{f(\bar{r})} + \bar{r}^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\,. \tag{4}$$

In Refs. 4,5 it was shown that its *more suitable form* for analysing the quadratic gravity field equation is

$$\mathrm{d}s^2 = \Omega^2(r)\big[\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2 - 2\,\mathrm{d}u\,\mathrm{d}r + \mathcal{H}(r)\,\mathrm{d}u^2\,\big]\,, \tag{5}$$

which is related to the classic form (4) by the transformation

$$\bar{r} = \Omega(r)\,, \qquad t = u - \int \mathcal{H}(r)^{-1}\mathrm{d}r\,, \tag{6}$$

and

$$h(\bar{r}) = -\Omega^2\,\mathcal{H}\,, \qquad f(\bar{r}) = -\left(\frac{\Omega'}{\Omega}\right)^2\mathcal{H}\,, \tag{7}$$

with prime denoting the derivative with respect to the coordinate $r$. The metric (5) is *conformal* to a direct-product Kundt 'seed' of the algebraic type D, see Refs. 6–8. The *Killing horizons* corresponding to $\partial_u = \partial_t$ are located at $r_h$ introduced via condition

$$\mathcal{H}\big|_{r=r_h} = 0\,. \tag{8}$$

The geometries (5) can be characterized using the *scalar curvature invariants*,

$$C_{abcd}\,C^{abcd} = \tfrac{1}{3}\,\Omega^{-4}\,(\mathcal{H}'' + 2)^2\,, \tag{9}$$

$$B_{ab}\,B^{ab} = \tfrac{1}{72}\,\Omega^{-8}\big[(\mathcal{B}_1)^2 + 2(\mathcal{B}_1 + \mathcal{B}_2)^2\big]\,, \tag{10}$$

where we have introduced *two independent Bach components*, namely

$$\mathcal{B}_1 \equiv \mathcal{H}\mathcal{H}''''\,, \qquad \mathcal{B}_2 \equiv \mathcal{H}'\mathcal{H}''' - \tfrac{1}{2}\mathcal{H}''^2 + 2\,. \tag{11}$$

Obviously, $\mathcal{B}_1 = 0 = \mathcal{B}_2 \;\Leftrightarrow\; B_{ab} = 0 \;\Leftrightarrow\; B_{ab}\,B^{ab} = 0$ and thus we may invariantly distinguish two geometrically different classes of solutions corresponding to $B_{ab} = 0$ and $B_{ab} \neq 0$, respectively.

## 3. The quadratic gravity field equations

Using the ansatz (5) together with the conformal behaviour of the Bach tensor ($B_{ab} = \Omega^{-2}\, B_{ab}^{\rm Kundt}$) and the contracted Bianchi identities, we obtain the field equations (3) in the form of two simple ODEs for the functions $\Omega(r)$ and $\mathcal{H}(r)$, namely

$$\Omega\Omega'' - 2\Omega'^2 = \tfrac{1}{3} k\, \mathcal{B}_1 \mathcal{H}^{-1}\,, \tag{12}$$

$$\Omega\Omega'\mathcal{H}' + 3\Omega'^2\mathcal{H} + \Omega^2 - \Lambda\Omega^4 = \tfrac{1}{3} k\, \mathcal{B}_2\,. \tag{13}$$

Moreover, the trace of (3), i.e., $R = 4\Lambda$, becomes

$$\mathcal{H}\Omega'' + \mathcal{H}'\Omega' + \tfrac{1}{6}(\mathcal{H}'' + 2)\Omega = \tfrac{2}{3}\Lambda\,\Omega^3\,. \tag{14}$$

These equations do not explicitly depend on the variable $r$ (an autonomous system) and their solutions can be found as *power series* around *any* point $r_0$,

$$\Omega(r) = \Delta^n \sum_{i=0}^{\infty} a_i\, \Delta^i\,, \qquad \mathcal{H}(r) = \Delta^p \sum_{i=0}^{\infty} c_i\, \Delta^i\,, \tag{15}$$

with $\Delta \equiv r - r_0$, and $n, p \in \mathbb{R}$.

### 3.1. *Vanishing Bach tensor*

In fact, for $B_{ab} = 0 \;\Leftrightarrow\; \mathcal{B}_1 = 0 = \mathcal{B}_2$ Einstein's theory is restored, and using the gauge freedom $r \to \lambda\, r + \nu$, $u \to \lambda^{-1} u$ of the metric (5), field equations (12) and (13) yield the classic solution

$$\Omega(r) = \bar{r} = -\frac{1}{r}\,, \qquad \mathcal{H}(r) = \frac{\Lambda}{3} - r^2 - 2m\, r^3\,. \tag{16}$$

Employing the explicit relations (7) it is straightforward to show that these functions really represent the *Schwarzschild-(anti–)de Sitter* spacetime[6,7] corresponding to $f = h = 1 - 2m\,\bar{r}^{-1} - \frac{1}{3}\Lambda\,\bar{r}^2$ in the metric (4).

### 3.2. *Non-vanishing Bach tensor*

A generic case $B_{ab} \neq 0$ leads to the non-trivially coupled system of the field equations (12), (13) which solutions can be found in the form (15). Analysing dominant powers of $\Delta$ in the equations we may immediately restrict the admitted values of parameters $[n,\, p]$ and the cosmological constant $\Lambda$, see Tab. 1. The most interesting black hole case $[n,\, p] = [0,\, 1]$ will be discussed in the following section.

Table 1. The only allowed values of $[n,\, p]$ and $\Lambda$ in (15), which are restricted by dominant powers of $\Delta$ in (12), (13), and (14).

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $n$ | 0 | 0 | 1 | $-1$ | $-1$ | 0 | 0 | $< 0$ |
| $p$ | 1 | 0 | 0 | 2 | 0 | 2 | $\geq 2$ | $2n+2$ |
| $\Lambda$ | any | any | any | 0 | $\neq 0$ | $\neq 0$ | $\frac{3}{8k}$ | $\frac{11n^2+6n+1}{1-4n^2}\,\frac{3}{8k}$ |

## 4. Explicit black hole solution expanded around horizon

In the case [0, 1], the single root of $\mathcal{H}$ given by $r_0 \equiv r_h$ represents the non-degenerate Killing horizon, see equation (8). Subsequently, the field equations (12), (13), together with the trace (14), specifically restrict the coefficients in (15) with free parameters $a_0,\ c_0,\ c_1$.

To obtain physical meaning of these free parameters we identify the Schwarzschild-(anti-)de Sitter 'background' spacetime (16). We begin with evaluation the Bach tensor (11) on the horizon, i.e., $\mathcal{B}_1(r_h) = 0$, $\mathcal{B}_2(r_h) = -\frac{3}{k} a_0^2\, b$. Here we have introduced the dimensionless parameter $b \equiv \frac{1}{3}(c_1 - 2 + \Lambda a_0^2)$. Surprisingly, setting $b = 0$ makes the Bach tensor vanishes *everywhere* and effectively leads to the Einstein case, where using the gauge freedom we may set

$$a_0 = -\frac{1}{r_h}\,, \qquad c_0 = r_h - \frac{\Lambda}{r_h}\,, \tag{17}$$

and sum up the series to obtain exactly the *Schwarzschild-(anti-)de Sitter black hole spacetime* (16) with the identification $\frac{\Lambda}{3} - r_h^2 = 2m\, r_h^3$.

In the generic case $b \neq 0$, we separate the 'Bach contribution' proportional to the parameter $b$ in the original coefficients $a_i$, $c_i$ of (15) by introducing their parts $\alpha_i$, $\gamma_i$ (and factorizing some constants). With the 'background' gauge choice (17), we get a one-parameter *extension* of the Schwarzschild-(A)dS spacetime,

$$\Omega(r) = -\frac{1}{r} - \frac{b}{r_h} \sum_{i=1}^{\infty} \alpha_i \Big( \frac{r_h - r}{\rho\, r_h} \Big)^i\,, \tag{18}$$

$$\mathcal{H}(r) = (r - r_h) \bigg[ \frac{r^2}{r_h} - \frac{\Lambda}{3r_h^3} \big( r^2 + r r_h + r_h^2 \big) + 3b\, \rho\, r_h \sum_{i=1}^{\infty} \gamma_i \Big( \frac{r - r_h}{\rho\, r_h} \Big)^i \bigg]\,, \tag{19}$$

where

$$\rho \equiv 1 - \frac{\Lambda}{r_h^2}\,, \qquad \alpha_1 \equiv 1\,, \qquad \gamma_1 = 1\,, \qquad \gamma_2 = \frac{1}{3} \Big[ 4 - \frac{1}{r_h^2} \Big( 2\Lambda + \frac{1}{2k} \Big) + 3b \Big]\,, \tag{20}$$

and $\alpha_l$, $\gamma_{l+1}$ for $l \geq 2$ are (with $\alpha_0 \equiv 0$) given by

$$\begin{aligned} \alpha_l = \frac{1}{l^2} \Big[ & - \frac{2\Lambda}{3r_h^2} \sum_{j=0}^{l-1} \sum_{i=0}^{j} \big[ \alpha_{l-1-j}\, \rho^j + \big( \rho^{l-1-j} + b\, \alpha_{l-1-j} \big) \big( \alpha_i\, \rho^{j-i} + \alpha_{j-i} (\rho^i + b\, \alpha_i) \big) \big] \\ & - \tfrac{1}{3} \alpha_{l-2} (2 + \rho) \rho (l-1)^2 + \alpha_{l-1} \big[ \tfrac{1}{3} + (1 + \rho) \big( l(l-1) + \tfrac{1}{3} \big) \big] \\ & - 3 \sum_{i=1}^{l} (-1)^i\, \gamma_i\, (\rho^{l-i} + b\, \alpha_{l-i}) \big( l(l-i) + \tfrac{1}{6} i(i+1) \big) \Big]\,, \end{aligned} \tag{21}$$

$$\gamma_{l+1} = \frac{(-1)^l}{k r_h^2\, (l+2)(l+1) l (l-1)} \sum_{i=0}^{l-1} \big[ \alpha_i\, \rho^{l-i} + \alpha_{l-i} \big( \rho^i + b\, \alpha_i \big) \big] (l-i)(l-1-3i)\,. \tag{22}$$

This solution represents a *three-parameter family of spherically symmetric static Schwarzschild-Bach-(anti-)de Sitter black holes* in the quadratic gravity, where

- the vaule $r = r_h$ corresponds to the *Killing horizon* since $\mathcal{H}(r_h) = 0$,
- the parameter $\Lambda$ representing the *cosmological constant* can be set to zero which recovers the results of Ref. 4,
- the *parameter* $b$ determines contribution of the non-trivial Bach tensor.

These physical parameters determine the scalar invariants (9), (10) on the horizon,

$$C_{abcd}\, C^{abcd}(r_h) = 12\left((1+b)r_h^2 - \tfrac{1}{3}\Lambda\right)^2, \qquad B_{ab}\, B^{ab}(r_h) = \frac{r_h^4}{4k^2}\, b^2\,. \tag{23}$$

To examine convergence of the series in (18), (19) we use the d'Alembert ratio test. With $n$ growing, the ratio between two subsequent terms approaches a specific constant and the series *asymptotically behave as geometric series*, see Fig. 1. The metric function $\mathcal{H}(r)$ outside the black-hole horizon is plotted in Fig. 2. We may observe a qualitative difference between $\Lambda < 0$ and $\Lambda > 0$, respectively, corresponding to the presence of an outer boundary of the static region ($r > r_h$) for $\Lambda > 0$.

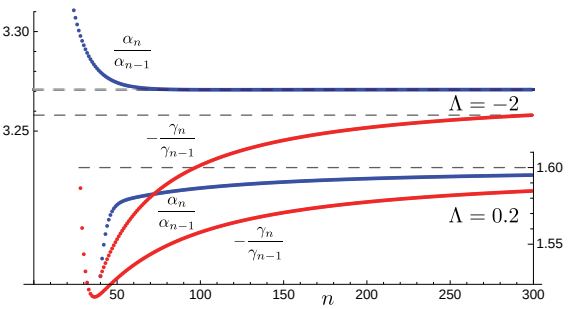



Fig. 1. The ratio convergence test for solutions (18), (19), here given by $r_h = -1$, $k = 0.5$ with $b = 0.2$, $\Lambda = -2$ (top) and $b = 0.3$, $\Lambda = 0.2$ (bottom).

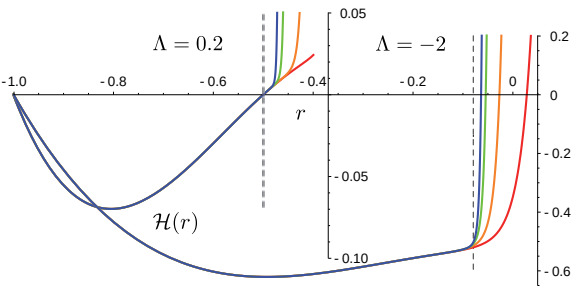



Fig. 2. Two plots of the metric function $\mathcal{H}(r)$ given by (19) with the same parameters as in Fig. 1. Vertical dashed lines indicate the radii of the convergence. For positive $\Lambda$ the function $\mathcal{H}(r)$ seems to have another root corresponding to the cosmological horizon. First 50 (red), 100 (orange), 200 (green), 300 (blue) terms in the expansions are visualized.

## 5. Physical properties

Interestingly, for the solutions introduced above the two independent parts (11) of the Bach tensor $\mathcal{B}_1, \mathcal{B}_2$ can be observed via a *specific relative motion of free test particles* described by the equation of geodesic deviation. Technical detail and explicit expressions can be found in Refs. 4,5.

Finally, let us mention the expression for the black hole *entropy* which also contain the remnant of the Bach tensor. Applying the generalized definition of entropy[9] we obtain

$$S = \tfrac{1}{4}\mathcal{A}\left(1 + \tfrac{4}{3}k\Lambda - 4k\, r_h^2\, b\right) = \tfrac{1}{4}\mathcal{A}\left(1 + \tfrac{4}{3}k\Lambda - 4k\, \bar{r}_h^{-2}\, b\right), \tag{24}$$

where $\mathcal{A} = 4\pi\, r_h^{-2} = 4\pi\, \bar{r}_h^2$ is the horizon area. More detailed discussion as well as other thermodynamic quantities are presented in our original letters.[4,5]

## Acknowledgments

This work was supported by the Czech Science Foundation under projects GAČR 17-01625S (RŠ and JP) and GAČR 19-09659S (VP and AP), and the Research Plan RVO: 67985840 (VP and AP).

## References

1. K. Schwarzschild, Über das Gravitationsfeld eines Massenpunktes nach der Einsteinschen Theorie, *Sitz. Preuss. Akad. Wiss. Berlin* **7** 189 (1916).
2. H. Lü, A. Perkins, C. N. Pope and K. S. Stelle, Black holes in higher derivative gravity, *Phys. Rev. Lett.* **114** 171601 (2015).
3. H. Lü, A. Perkins, C. N. Pope and K. S. Stelle, Spherically symmetric solutions in higher-derivative gravity, *Phys. Rev. D* **92** 124019 (2015).
4. J. Podolský, R. Švarc, V. Pravda and A. Pravdová, Explicit black hole solutions in higher-derivative gravity, *Phys. Rev. D* **98** 021502(R) (2018).
5. R. Švarc, J. Podolský, V. Pravda and A. Pravdová, Exact Black Holes in Quadratic Gravity with any Cosmological Constant, *Phys. Rev. Lett.* **121** 231104 (2018).
6. H. Stephani et al., *Exact Solutions of Einstein's Field Equations* (Cambridge University Press, Cambridge, 2003).
7. J. Griffiths and J. Podolský, *Exact Space-Times in Einstein's General Relativity* (Cambridge University Press, Cambridge, 2009).
8. V. Pravda, A. Pravdová, J. Podolský and R. Švarc, Exact solutions to quadratic gravity, *Phys. Rev. D* **95** 084025 (2017).
9. R. M. Wald, Black hole entropy is the Noether charge, *Phys. Rev. D* **48** R3427(R) (1993).

# Energy nonconservation: $f(R,T)$ gravity and noncommutative quantum theory

Ronaldo V. Lobato[†] and G. A. Carvalho[*]

*Departamento de Física, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, 12228-900, Brazil.*
*ICRANet, P.zza della Repubblica 10, I–65122 Pescara, Italy.*
[†]*E-mail: rvlobato@pm.me*
[*]*E-mail: araujogc@ita.br*

A. G. Martins

*Departamento de Ciências Naturais, Universidade do Estado do Pará, Belém, Pará, 66050-540, Brazil.*
*E-mail: andrey_martins@uepa.br*

P. H. R. S. Moraes

*Departamento de Física, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, 12228-900, Brazil.*
*E-mail: moraes.phrs@gmail.com*

The $f(R,T)$ gravity theory was proposed as an extension of the $f(R)$ theories, for which besides geometrical correction terms, proportional to the Ricci scalar $R$, one has also material correction terms, proportional to the trace of the energy-momentum tensor $T$. Those material extra terms prevent the energy-momentum tensor of the theory to be conserved. On the other hand, in the context of noncommutative quantum mechanics, the presence of compact dimensions whose coordinates do not commute with time imply that time evolution is discretized, a feature which induces violations of energy conservation. In the present work we propose a connection between these two effects, so that the energy nonconservation observed in the 4-dimensional $f(R,T)$ gravity can be understood as a macroscopic effect of nonconservative quantum transitions involving the compact extra dimension. It turns out that the energy flows between the ordinary (commutative) 4-dimensional spacetime and the compact extra dimension.

*Keywords*: Modified gravity; $f(R,T)$ gravity; noncommutative quantum mechanics; discrete time; violation of energy conservation.

## 1. Introduction

The $f(R,T)$ gravity[1] is a generalization of the $f(R)$ theories. Its gravitational action depends on an arbitrary function of both the Ricci scalar $R$ and the trace of the energy-momentum tensor $T$. The dependence on $T$ was inspired initially by the consideration of existence of exotic imperfect fluids and quantum effects, the later sometimes associated with conformal anomaly[2,3].

The $f(R,T)$ action reads[1]

$$\mathcal{S}=\int d^4x\sqrt{-g}\left[\frac{f(R,T)}{16\pi}+\mathcal{L}_m\right]. \tag{1}$$

In (1), $f(R,T)$ is the general function of $R$ and $T$, $\mathcal{L}_m$ is the matter Lagrangian and $g$ is the determinant of the metric tensor $g_{\mu\nu}$.

Field equations of $f(R,T)$ gravity present extra terms, and due to those new terms the covariant derivative of the energy-momentum tensor is shown to be non-null, that is,

$$\nabla^{\mu}T_{\mu\nu} = \mathtt{F}\left[(\mathcal{L}_m g_{\mu\nu} - T_{\mu\nu})\nabla^{\mu}\ln f_T(R,T) + \nabla^{\mu}\left(\mathcal{L}_m - \frac{1}{2}T\right)g_{\mu\nu}\right], \tag{2}$$

where we introduce the notation

$$\mathtt{F} = \frac{f_T(R,T)}{8\pi + f_T(R,T)}$$

and $f_T(R,T) = \partial f(R,T)/\partial T$.

Considering the energy-momentum tensor $T_{\mu\nu}$ to describe a perfect fluid we can assume the matter Lagrangian to be $\mathcal{L}_m = \rho$.

In General Relativity the energy-momentum tensor is covariantly conserved, so that a flat space-time background leads to well-defined energy and momentum conservation laws, in particular described by the continuity and Navier-Stokes equations. Therefore, we consider the Minkowski metric as the background space-time metric, such that $\nabla_{\mu}T^{\mu\nu} = \partial_{\mu}T^{\mu\nu}$. Then, from the covariant derivative of the energy-momentum tensor (2), for $\nu = 0$, one can obtain the modified continuity equation[4]

$$\frac{\partial\rho}{\partial t} + \nabla\vec{j} = \frac{f_T}{8\pi + f_T}\left[\frac{1}{2}\frac{\partial\rho}{\partial t} - \vec{j}\cdot\vec{\nabla}(\ln f_T)\right], \tag{3}$$

where $\rho$ is the energy density of the system and $\vec{j}$ represents the energy current density.

From (3), for $f_T = 0$, the standard continuity equation is recovered. It is worth to note that equation (3) implies that in $f(R,T)$ gravity the energy conservation law is not respected, even in a flat spacetime background. This fact can be interpreted as energy flux that is ejected (injected) from (in) the system (see Ref.[4] for a detailed account). The loss or gain of energy will depend on the function $f_T(R,T)$.

## 2. Noncommutative quantum theory

### 2.1. *The algebra of noncommutative spacetime*

The canonical $D$-dimensional noncommutative spacetime is defined by replacing the coordinates $x_i$ by Hermitian operators $\hat{x}_i$ which obey the commutation relations[5,6]

$$[\hat{x}_i, \hat{x}_j] = i\theta_{ij}\mathbb{1}, \tag{4}$$

where $\theta_{ij}$ are the components of an antisymmetric $D\times D$ constant tensor and $\mathbb{1}$ is the unity element of the spacetime noncommutative algebra[5]. The components of $\theta$ determine the spacetime discretization, in the same manner as the Planck constant $\hbar$ discretizes the phase space in quantum mechanics. The geometric notion of point is

no longer meaningful, because the coordinates cannot be simultaneously measured, leading to a spacetime uncertainty relation,

$$\Delta x_i \Delta x_j \geq |\theta_{ij}|/2. \tag{5}$$

It was argued[7] that one can reach a lower bound limit on the measurements of lengths, $\Delta x_{\min} = l_{\mathrm{P}}$. In the low-energy limit, $l_{\mathrm{P}} \to 0$, one recovers the usual classical spacetime with commutative coordinates.

### 2.1.1. *The noncommutative plane*

In the special case of $D = 2$, one obtains the noncommutative plane, also called Moyal plane[8], obeying the commutation relation ($i, j \in \{0, 1\}$)

$$[\hat{x}_i, \hat{x}_j] = i\theta\epsilon_{ij}, \tag{6}$$

where $\epsilon_{ij}$ is the 2-dimensional Levi-Civita pseudo-tensor and $\theta > 0$. Equation (6) leads to the Heisenberg-like uncertainty relation

$$\Delta x_0 \Delta x_1 \geq \frac{\theta}{2}. \tag{7}$$

### 2.1.2. *Noncommutative cylinder*

Another noncommutative spacetime of interest is the noncommutative cylinder, which can be seen as resulting from the compactification of the spatial coordinate of the Moyal plane. As a consequence, the time evolution gets quantized, as shown in Ref.[9].

The classical (commutative) cylinder is topologically equivalent to $\mathbb{R} \times S^1$, where $S^1$ is a circle. In the present work we consider the compact dimension as a hidden warped space-like extra dimension, analogously to the Kaluza-Klein theory, but with the hidden dimension replaced by a noncommutative circle. Quantum Mechanics in the noncommutative cylinder was considered in Ref.[10], where it was shown that energy is no longer conserved in scattering processes.

Taking into account the extra dimension (from now on denoted by $\hat{x}_4$) together with the commutative coordinates, $x_1$, $x_2$ and $x_3$, we consider a Hamiltonian of the form $\hat{H}_{\text{ total}} = \hat{H}_{\text{commut}}(x_1, x_2, x_3) + \hat{H}_{\text{extra}}(\hat{x}_4)$, so that the physical states can be expressed as

$$\Psi_{\text{total}}(\hat{x}_0, x_1, x_2, x_3, \hat{x}_4) \equiv \psi(x_1, x_2, x_3)\, \hat{\psi}(\hat{x}_0, \hat{x}_4). \tag{8}$$

As a consequence of the topology of the extra dimension, the functions $\hat{\psi}(\hat{x}_0, \hat{x}_4)$ are periodic in $\hat{x}_4$. The spatial period is given by $2\pi r$, where $r$ is the circle radius. It follows that $\hat{\psi} = \hat{\psi}(\hat{x}_0, e^{i\frac{\hat{x}_4}{r}})$. In general the states can be expressed as linear combinations of the form

$$\hat{\psi} = \sum_{m=-\infty}^{\infty} C_m e^{im\frac{\hat{x}_4}{r}}, \tag{9}$$

and one can consider that the algebra of the noncommutative cylinder is generated by $\hat{x}_0$ and $e^{i\frac{\hat{x}_4}{r}}$. From (6) we obtain the commutation relation

$$[\hat{x}_0, e^{i\frac{\hat{x}_4}{r}}] = -\frac{\theta}{r} e^{i\frac{\hat{x}_4}{r}}, \tag{10}$$

and using the Baker-Campbell-Hausdorff formula[11] one can show that

$$e^{i\frac{\hat{x}_4}{r}} e^{i\frac{2\pi r}{\theta}\hat{x}_0} = e^{i\frac{2\pi r}{\theta}\hat{x}_0} e^{i\frac{\hat{x}_4}{r}}. \tag{11}$$

Therefore, the operator $e^{i\frac{2\pi r}{\theta}\hat{x}_0}$ commutes with all elements of the algebra of the noncommutative cylinder. According to the Schur's lemma[11], its action on the quantum states is given by $e^{i\varphi}\mathbb{1}$, where the phase $e^{i\varphi}$ classifies the possible irreducible representations of the noncommutative cylinder algebra. In what follows we choose $\varphi = 0$, so that $e^{i\frac{2\pi r}{\theta}\hat{x}_0} = \mathbb{1}$.

Now, from the eigenvalue equation $\hat{x}_0 |x_0\rangle = x_0 |x_0\rangle$ we get

$$e^{i\frac{2\pi r}{\theta}\hat{x}_0} |x_0\rangle = e^{i\frac{2\pi r}{\theta}x_0} |x_0\rangle . \tag{12}$$

Since $e^{i\frac{2\pi r}{\theta}\hat{x}_0} = \mathbb{1}$, one obtains $e^{\frac{i2\pi x_0}{\theta}} = 1$, which leads to $\frac{2\pi r x_0}{\theta} = 2\pi n$ and $x_0 = n\frac{\theta}{r}$, with $n \in \mathbb{Z}$. Hence, the spectrum of $\hat{x}_0$ is

$$\text{spec } \hat{x}_0 \equiv \left\{ n\frac{\theta}{r}, \quad \text{with} \quad n \in \mathbb{N} \right\}. \tag{13}$$

The values of the time variable correspond to the spectrum of $\hat{x}_0$, i.e., time is discretized, with a minimum interval $\frac{\theta}{r} \neq 0$ and generic time interval $\Delta t = n\frac{\theta}{r}$, with $n \in \mathbb{N}$.

According to[9,10], time discretization implies that time evolution is given by integral powers of the minimum time evolution operator $U\left(\frac{\theta}{r}\right)$, that is,

$$U\left(n\frac{\theta}{r}\right) = e^{-in\frac{\theta}{r}\hat{H}}. \tag{14}$$

Equation (14) indicates that the time evolution allows quantum transitions from eigenstates of initial energy $E_i$ to any final energy of the form $E_f = E_i + \frac{2\pi r m}{\theta}$, thus leading to the possibility of energy nonconservation.

In Ref.[8] it was shown that the energy nonconservation should occur in scattering and decay processes. Here we argue that this phenomenon can be verified in $f(R,T)$ gravity, considered as an effective theory defined on a 4-dimensional sub-space of a 5-dimensional space-time, where the time coordinate does not commute with the compactified extra coordinate. The main effect of the nonconservative quantum transitions is an energy flow between the extra dimension $S^1$ and the usual 4-dimensional spacetime. This macroscopic effect is captured by (3), where the extra terms in the continuity equation leads to energy nonconservation. In Ref.[4] this link between noncommutative quantum mechanics and the nonconservation of energy of $f(R,T)$ gravity is explained in detail.

## 3. Conclusions

Energy and momentum conservation laws are well-stablished in flat spacetime backgrounds, in the context of General Relativity. By its turn, $f(R,T)$ gravity does not present this characteristic. Since the energy-momentum tensor is not covariantly conserved, it follows that the field equations lead to extra terms in the continuity equation, even in a flat spacetime background, thus showing that the energy is not conserved.

Within the context of noncommutative theory we can also have energy nonconservation due to the time discretization induced by the presence of a noncommutative compact extra dimension.

We argue that the energy nonconservation in a 4-dimensional $f(R,T)$ theory can be understood as an effect of quantum transitions in a larger spacetime containing a fifth compact dimension which does not commute with the time operator. As a consequence, there is a flow of energy between the usual spacetime and this extra dimension. Even though the conservation laws are preserved in this larger scenario, in the 4-dimensional effective theory energy is no longer conserved.

## Acknowledgements

RVL thanks CNPq (Conselho Nacional de Desenvolvimeno Científico e Tecnológico) process 141157/2015-1 and CAPES/PDSE/88881.134089/2016-01 for financial support. GAC thanks to CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) for financial support process 88881.188302/2018-01.

## References

1. T. Harko, F. S. N. Lobo, S. Nojiri and S. D. Odintsov, F(R,T) gravity, *Physical Review D* **84** (July 2011), 00776.
2. T. Harko, Thermodynamic interpretation of the generalized gravity models with geometry-matter coupling, *Physical Review D* **90**, p. 044067 (August 2014).
3. T. Harko, F. S. N. Lobo, J. P. Mimoso and D. Pavón, Gravitational induced particle production through a nonminimal curvature–matter coupling, *The European Physical Journal C* **75**, p. 386 (August 2015), 00024.
4. R. V. Lobato, G. A. Carvalho, A. G. Martins and P. H. R. S. Moraes, Energy nonconservation as a link between f(R,T) gravity and noncommutative quantum theory, *The European Physical Journal Plus* **134**, p. 132 (April 2019), 00000.
5. M. R. Douglas and N. A. Nekrasov, Noncommutative Field Theory, *Reviews of Modern Physics* **73**, 977 (November 2001).
6. R. J. Szabo, Quantum Field Theory on Noncommutative Spaces, *Physics Reports* **378**, 207 (May 2003), 00000.
7. S. Doplicher, K. Fredenhagen and J. E. Roberts, Spacetime quantization induced by classical gravity, *Physics Letters B* **331**, 39 (1994).

8. A. P. Balachandran, T. R. Govindarajan, C. Molina and P. Teotonio-Sobrinho, Unitary Quantum Physics with Time-Space Noncommutativity, *Journal of High Energy Physics* **2004**, 072 (October 2004), 00000.
9. A. P. Balachandran, T. R. Govindarajan, A. G. Martins and P. Teotonio-Sobrinho, Time-Space Noncommutativity: Quantised Evolutions, *Journal of High Energy Physics* **2004**, 068 (November 2004).
10. A. P. Balachandran, A. G. Martins and P. Teotonio-Sobrinho, Discrete time evolution and energy nonconservation in noncommutative physics, *Journal of High Energy Physics* **2007**, p. 066 (May 2007), 00010.
11. B. C. Hall, An Elementary Introduction to Groups and Representations, **222** (May 2000).

# Quantization of the Szekeres spacetime through generalized symmetries

Andronikos Paliathanasis

*Instituto de Ciencias Físicas y Matemáticas, Universidad Austral de Chile, Valdivia, Chile*

*Institute of Systems Science, Durban University of Technology, POB 1334 Durban 4000, South Africa*

Adamantia Zampeli*

*Institute of Theoretical Physics, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, 18000 Prague 8, Czech Republic*

Theodosios Christodoulakis

*Nuclear and Particle Physics section, Physics Department, University of Athens, 15771 Athens, Greece*

M.T. Mustafa

*Department of Mathematics, Statistics and Physics, College of Arts and Sciences, Qatar University, Doha 2713, Qatar*

We present the effect of the quantum corrections on the Szekeres spacetime, a system important for the study of the inhomogeneities of the pre-inflationary era of the universe. The study is performed in the context of canonical quantisation in the presence of symmetries. We construct an effective classical Lagrangian and impose the quantum version of its classical integrals of motion on the wave function. The interpretational scheme of the quantum solution is that of Bohmian mechanics, in which one can avoid the unitarity problem of quantum cosmology. We discuss our results in this context.

*Keywords*: Szekeres system; Silent universe; Quantisation; Semiclassical approach

## 1. Introduction

We focus on the quantisation of the Szekeres spacetime metric with the aim to study the effect of possible quantum corrections in the dynamics. This metric has the form [1]

$$ds^2 = -dt^2 + e^{2\alpha}dr^2 + e^{2\beta}\left(dy^2 + dz^2\right) \tag{1}$$

where $\alpha \equiv \alpha(t, r, y, z)$ and $\beta \equiv \beta(t, r, y, z)$ and represents an irrotational perfect fluid with vanishing pressure and magnetic Weyl tensor, $p = \omega_{ab} = H_{ab} = 0$. The interest in the silent universe lies on the fact that it can be seen as inhomogeneous solutions of Einstein equations with no symmetries which generalise Kantowski-Sachs, FRW and Tolman-Bondi spacetimes. Thus it is proper for the description of perturbations on these spacetimes [2–5].

We start by writing the field equations on the covariant variables $(\rho, \theta, \sigma, E)$, where $\rho = T^{\mu\nu}u_\mu u_\nu$, with $T^{\mu\nu}$ being the energy-momentum tensor of the matter,

*azampeli@phys.uoa.gr

$\theta = (\nabla_\nu u_\mu)\, h^{\mu\nu}$ is the expansion rate of the observer, while $\sigma$ and $E$ are the shear and electric component of the Weyl tensor, $E^\mu_\nu = Ee^\mu_\nu,\ \sigma^\mu_\nu = \sigma e^\mu_\nu$, in which the set of $\{u^\mu, e^\mu_\nu\}$ defines an orthogonal tetrad. The field equations then become

$$\dot{\rho} + \theta\rho = 0, \tag{2a}$$

$$\dot{\theta} + \frac{\theta^2}{3} + 6\sigma^2 + \frac{1}{2}\rho = 0, \tag{2b}$$

$$\dot{\sigma} - \sigma^2 + \frac{2}{3}\theta\sigma + E = 0, \tag{2c}$$

$$\dot{E} + 3E\sigma + \theta E + \frac{1}{2}\rho\sigma = 0, \tag{2d}$$

together with the algebraic equation which is the Hamiltonian constraint

$$\frac{\theta^2}{3} - 3\sigma^2 + \frac{{}^{(3)}R}{2} = \rho \tag{2e}$$

where $\dot{}$ denotes the directional derivative along $u^\mu$, the energy density is $\rho = T^{\mu\nu}u_\mu u_\nu$, with $T^{\mu\nu}$ being the energy-momentum tensor of the matter, the parameter $\theta = (\nabla_\nu u_\mu)\, h^{\mu\nu}$ is the expansion rate of the observer, while $\sigma$ and $E$ are the shear and electric component of the Weyl tensor, $E^\mu_\nu = Ee^\mu_\nu,\ \sigma^\mu_\nu = \sigma e^\mu_\nu$, in which the set of $\{u^\mu, e^\mu_\nu\}$ defines an orthogonal tetrad. We note that in addition to these equations, the spatial constraints ensure the integrability of the system.

In the following sections we present the quantization of this system in terms of canonical quantization in the presence of symmetries[6]. The starting point is the effective Lagrangian obtained in[7] and we adopt the Bohmian approach for our analysis[8,9] following e.g.[10], since its causal character suits the context of quantum cosmology, where the notion of an external observer cannot be justified.

## 2. Classical and Quantum Dynamics

In[7] the Szekeres system (2) was written in an equivalent form of a two second-order differential equations system which are equations of motion of a Lagrangian of the form

$$L = \frac{1}{2}G_{\alpha\beta}(q(t))\dot{q}^\alpha(t)\dot{q}^\beta(t) - V(q(t)), \quad \alpha, \beta = 0, \ldots n-1 \tag{3}$$

where $q(t)$ denote the degrees of freedom of the system and $G_{\alpha\beta}$ the metric on the configuration space of variables. Adopting proper coordinates for our case, $\rho = \frac{6}{(1-v)u^2},\ E = \frac{v}{u^3(v-1)}$, the system of the two second order equations becomes

$$\ddot{v} - \frac{2v}{u^3} = 0, \tag{4a}$$

$$\ddot{u} + \frac{1}{u^2} = 0 \tag{4b}$$

derivable from $L = \dot{u}\dot{v} - \frac{\dot{v}}{u^2}$. The system (4) admits two integrals of motion, quadratic in the velocities; the first is the Hamiltonian function since the system

is autonomous, while the second one is the quadratic function $I_0$ which can be constructed by the application of Noether's theorem for contact symmetries[7], which in the phase space become

$$p_u p_v + \frac{v}{u^2} = h, \tag{5a}$$

$$p_v^2 - 2u^{-1} = I_0 \tag{5b}$$

When turned to quantum operators and imposed on the wave function, according to the rules $\hat{p}_\alpha = -i\frac{\partial}{\partial q^\alpha}$, $\{.,.\} \to -\frac{i}{\hbar}[.,.]$, with operator-ordering respecting the general covariance and hermiticity ensured under the inner product $\int d^n q \mu \psi_1^* \psi_2$, they lead to two eigenvalue equations

$$\left(-\partial_{uv} + \frac{v}{u^2}\right)\Psi = h\Psi, \tag{6a}$$

$$\left(\partial_{vv} + \frac{2}{u}\right)\Psi = -I_0\Psi, \tag{6b}$$

In the first one we can recognise the time-independent Schroödinger equation and their solution is

$$\Psi\left(I_0, u, v\right) = \frac{\sqrt{u}}{\sqrt{2 + I_0 u}}\left(\Psi_1 \cos f\left(u, v\right) + \Psi_2 \sin f\left(u, v\right)\right) \tag{7}$$

where

$$f\left(u, v\right) = \frac{\left(hu + I_0 v\right)\sqrt{2I_0 + I_0^2 u} - 2h\sqrt{u}\,\mathrm{arcsinh}\,\sqrt{\frac{I_0 u}{2}}}{I_0^{3/2}\sqrt{u}}, \quad \text{for } I_0 \neq 0, \tag{8}$$

$$f\left(u, v\right) = \frac{\sqrt{2}\left(hu^2 + 3v\right)}{3\sqrt{u}}, \quad \text{for } I_0 = 0. \tag{9}$$

and $\Psi_1$, $\Psi_2$ denote constants of integration. Due to the linearity of (6a), the general solution is $\Psi_{Sol}\left(u, v\right) = \sum_{I_0} \Psi\left(I_0, u, v\right)$.

## 3. Semiclassical analysis and probability

In the context of the Bohmian approach, the departure from the classical theory is determined by an additional term in the classical Hamilton-Jacobi equation, known as quantum potential $Q_V = -\frac{\Box\Omega}{2\Omega}$, where $\Omega$ denotes the amplitude of the wave function in polar form, $\Psi(u, v) = \Omega(u, v)e^{iS(u,v)}$. When the quantum potential is zero, the identification

$$\frac{\partial S}{\partial q_i} = p_i = \frac{\partial}{\partial \dot{q}_i} \tag{10}$$

is possible. If this classical definition for the momenta is retained even when $Q \neq 0$, the semiclassical solutions will differ from the classical ones.

Under the assumption that the quantum corrections in the general solution (7) follow from the "frequency $I_0$" with the highest peak in the wave function, which is

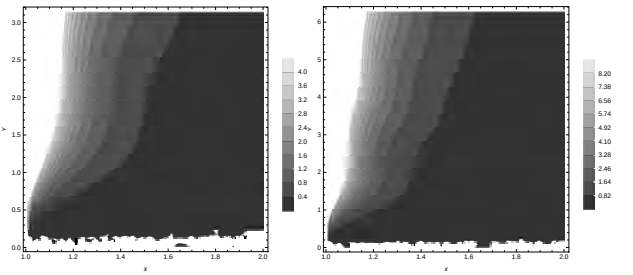


Fig. 1. Contour plot for the probability function in the space of variables $x, v$. We observe that as $x \to 0$ and $v$ is small, that is, $I_0 \to 0$, the function $P(x, v)$ reaches to a minimum extreme.

in agreement with the so-called Hartle criterion[11], the quantum potential vanishes. This provides no quantum corrections and the semiclassical equations (10) give the classical solution, i.e. the Szekeres universe remains "silent", even at the quantum level. In the particular case $h = 0$ and $\Psi_1 \to 0$, the wave function is well behaved at $u \to 0$ and $u \to \infty$. We can thus define a probability which, after a change of coordinates to $u \to \frac{2}{x^2 - I_0}$ becomes

$$P = \int_{\sqrt{I_0}+\epsilon}^{\lambda} dx \int_0^{2k\pi} dv \; \frac{4c_3^2 \sin(xv)}{x(x^2 - I_0)^2} \; , \; k \in \mathbb{N}. \tag{11}$$

where the cut-off constant $\lambda$ is introduced to exclude the case $E = 0, \rho = 0$. The qualitative behaviour of the probability function is given in the contour plot in Fig. 1. The plots show that for $I_0 \to 0$ the probability function reaches its minimum.

## 4. Conclusions

Our quantum analysis of the Szekeres system was based on the canonical quantization in the presence of symmetries and the results were interpreted by adopting the Bohmian mechanics approach. The starting point was an effective classical point-like Lagrangian which can reproduce the two dimensional system of second-order differential equations resulted from the initial field equations. This Lagrangian is autonomous, thus there exists a conservation law of "energy" corresponding to the Hamiltonian function. As for the extra contact symmetry, it leads to a quadratic in the momenta conserved quantity attributed to a Killing tensor of the second-rank. The two conserved quantities give two eigenequations at the quantum level, the Hamiltonian function being the Schrödinger equation.

The assumption that the wave function is peaked around its classical value leads to the lack of quantum corrections and the recovery of the classical solutions, thus leading to the conclusion that the Szekeres universe remains silent at the quantum level. Finally, for the particular case $h = 0$ it was shown that the probability function and relate one (unstable) exact solution with the existence of a minimum of this probability.

The classical exact solution which corresponds at the values $h = I_0 = 0$ of the integration constants is $u_A(t) = \frac{6^{\frac{2}{3}}}{2} t^{\frac{2}{3}}, v_A(t) = v_0 t^{-\frac{1}{3}}$ [7] and corresponds to an unstable critical point for the dynamical system (2). It is interesting the fact that these conditions also correspond to the extremum of the probability function, something which might be related to the existence and stability of the exact solution. This result is also in accordance with the analysis of the probability extrema in [12] where it was shown that the extrema of the probability lie on the classical values.

## References


1. P. Szekeres, A Class of Inhomogeneous Cosmological Models, *Commun. Math. Phys.* **41**, p. 55 (1975).
2. M. Ishak and A. Peel, The growth of structure in the Szekeres inhomogeneous cosmological models and the matter-dominated era, *Phys. Rev.* **D85**, p. 083502 (2012).
3. K. Bolejko and M.-N. Celerier, Szekeres Swiss-Cheese model and supernova observations, *Phys. Rev.* **D82**, p. 103510 (2010).
4. D. Vrba and O. Svitek, Modelling inhomogeneity in Szekeres spacetime, *Gen. Rel. Grav.* **46**, p. 1808 (2014).
5. M. Bruni, S. Matarrese and O. Pantano, Dynamics of silent universes, *Astrophys. J.* **445**, 958 (1995).
6. T. Christodoulakis, N. Dimakis, P. A. Terzis, G. Doulis, T. Grammenos *et al.*, Conditional Symmetries and the Canonical Quantization of Constrained Minisuperspace Actions: the Schwarzschild case, *J.Geom.Phys.* **71**, 127 (2013).
7. A. Paliathanasis and P. G. L. Leach, Symmetries and Singularities of the Szekeres System, *Phys. Lett.* **A381**, 1277 (2017).
8. D. Bohm, A Suggested interpretation of the quantum theory in terms of hidden variables. 1., *Phys. Rev.* **85**, 166 (1952).
9. D. Bohm, A Suggested interpretation of the quantum theory in terms of hidden variables. 2., *Phys. Rev.* **85**, 180 (1952).
10. A. Zampeli, T. Pailas, P. A. Terzis and T. Christodoulakis, Conditional symmetries in axisymmetric quantum cosmologies with scalar fields and the fate of the classical singularities, *JCAP* **1605**, p. 066 (2016).
11. J. B. Hartle, *Prediction in Quantum Cosmology*, in *Gravitation in Astrophysics: Cargèse 1986*, eds. B. Carter and J. B. Hartle (Springer US, Boston, MA, 1987), Boston, MA, pp. 329–360.
12. N. Dimakis, P. A. Terzis, A. Zampeli and T. Christodoulakis, Decoupling of the reparametrization degree of freedom and a generalized probability in quantum cosmology, *Phys. Rev.* **D94**, p. 064013 (2016).

# Investigation of the stability of orbits by using the adiabatic theory of motion in general relativity

S. Toktarbay*, M.E. Abishev[1], A.Z. Talkhat [1], A. Muratkhan [1],
S. S. Kenzhebayeva[1] and A.Zh. Abylayeva[1]

[1] *Department of Theoretical and Nuclear Physics, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan*
* *E-mail: saken.yan@yandex.com, https://www.kaznu.kz/en*

We investigate the problem of the orbital stability of the motion of a test body in the restricted three-body problem, where all bodies have their own rotation. The stability of the orbits is investigated by using vector elements of orbits such as orbital moment and its time derivative. We show that it is possible to get some insight into the stability properties of the motion of test bodies.

*Keywords*: Restricted three-body problem; orbital stability; vector elements

## 1. Introduction

The problem of the motion of bodies in general relativity (GR) can be split into two correlated problems, namely, the problem of motion of point like masses and the problem of motion of extended bodies. There are several methods for obtaining the equations of motion from the equations of the gravitational field, such as the method of Einstein-Infeld-Hoffmann (EIH)[1], the first approximate method of Fock[2], the second approximate method of Fock, etc. the EIH and Infeld methods allow us to find the equations of motion of point masses. There are several works devoted to investigate the restricted three body problem in GR. In[3] for three finite masses, it was showed that a triangular configuration satisfies the post-Newtonian equation of motion in general relativity, if and only if it has the relativistic corrections to each side length. Also the relativistic effect of restricted three body problem was investigated in[4] and it was shown that the relativistic effects become important when the parameter $M/R$ is increased. The motion of an extended body in GR mechanics eas first formulated by the V. Fok and after that developed by Petrova, Brumberg and Abdilldin[2 5 6 7]. Fock's methods make it possible to derive the equations of motion of extended bodies, allow us to take into account the internal structure and shape of bodies. In[8 9], we investigated the problem of the orbital stability of a circular motion of a test body in the restricted three body problem, when perturbations from the second body are of the order of relativistic corrections to the motion of a test body from the central body. It was shown than in this case the orbit of the test body is stable.

This work is organized as follows. In Sec. 2, we present the relativistic Lagrangian and derive the equation of motion. In Sec. 3, we present the evolution equation of motion of a test body. Finally, Sec. 4 contains discussions of our results.

## 2. The Lagrangian and equation of motion

The problem of the orbital stability of circular motion of a test body in the restricted three-body problem has been investigated in[8]. The resting position of the central body coincides with the reference point of coordinates, the second body moves along the circle around the central body and is not subject to the disturbance. The test body moves in a perturbed circular orbit and the relativistic corrections are:

$$U_2 \ll U_1; U_1/U_2 \approx v^2/c^2, \tag{1}$$

The relativistic Lagrange function can be represented as:

$$L = L^{(0)} + L^{(rot)}. \tag{2}$$

where $L^{(0)}$ is the relativistic Lagrange function for three point masses[10], and the second term $L^{(rot)}$ is responsible for corrections containing rotational terms. Regarding the $L^{(0)}$ term, in[8] it is shown that the motion of the test body in the plane of the orbit of the two body is stable when all bodies have no proper rotation. In this work, we consider the case when all bodies have their own rotation. Our aim to find the evolution equation of motion for the test (third) body, which describes the average change of its orbital momentum. To do that, we will study the evolution equations of motion by using the asymptotic methods of adiabatic theory, through the process of averaging of the corresponding equations using the vector elements **M** (the orbital moment) and **A** (the Laplace vector). M.M.Abdildin[6] and Burmberg[5] devoted an entire book to the motion of extended body with the internal structure and rotation in GR, where the Lagrange derived for N body using the Fok's approach. We derived the second term of Eq.(2) for the three masses without consideration of internal structure. The details of this derivation will be presented elsewhere. According to the adiabatic theory, the evolutionary motion of the test(third)body describes the average change of its orbital momentum. Therefore, we write down the orbital angular momentum of the test body:

$$\vec{M} = [\vec{r}, \vec{p}] \tag{3}$$

where the square bracket means the vector cross product. The evolutionary motion of the test(third)body is:

$$\dot{\vec{M}} = \left[\dot{\vec{r}}, \vec{p}\right] + \left[\vec{r}, \dot{\vec{p}}\right] \tag{4}$$

where the dot represents time derivative.

According to the Eqs.(2) the time derivative of orbital angular momentum reduce to $\dot{\vec{M}}^{(rot)}$

$$\dot{\vec{M}} = \dot{\vec{M}}^{(0)} + \dot{\vec{M}}^{(rot)} \tag{5}$$

The quantities $\dot{\vec{r}}$ and $\dot{\vec{p}}$ are found by means of the Hamilton equations from the Hamilton function of the system and which satisfies

$$H = \vec{\vartheta}_i \frac{\partial L}{\partial \vec{\vartheta}_i} - L \tag{6}$$

After lengthy computations, it is possible to isolate the derivatives of the canonical variables of the test body:

$$\dot{\vec{r}}_3^{(rot)} = \frac{1}{c^2} I_3 \left(\vec{\omega}_3 \cdot \vec{p}_3\right) \vec{\omega}_3 - \frac{\gamma}{2c^2} \cdot \frac{m_2 I_2 + m_3 I_3}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \times \left[ -\vec{p}_2 + 3 \frac{\left(\left(\vec{r}_2 - \vec{r}_3\right) \cdot \vec{p}_2\right) \cdot \left(\vec{r}_2 - \vec{r}_3\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \right], \tag{7}$$

$$\begin{aligned}
\dot{\vec{p}}_3^{(*)} = & \left\{ -\frac{3\gamma}{2c^2} \left(m_2 I_2 + m_3 I_3\right) \frac{\vec{p}_2 \vec{p}_3}{\left|\vec{r}_2 - \vec{r}_3\right|^5} + \frac{15\gamma}{2c^2} \frac{\left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{p}_2\right) \cdot \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{p}_3\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^7} \right. \\
& - \frac{3\gamma}{c^2} \frac{1}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \left(m_2 \omega_3^2 I_3 + m_3 \omega_2^2 I_2\right) + \frac{3\gamma^2}{2c^2} \frac{1}{\left|\vec{r}_2 - \vec{r}_3\right|^5} \left(m_2 I_3 + m_3 I_2\right) \\
& + \frac{\gamma^2}{2c^2} m_1 m_2 I_3 \frac{1}{\left|\vec{r}_3\right|^3 \left|\vec{r}_2 - \vec{r}_3\right|^3} - \frac{3\gamma^2}{2c^2} m_1 m_2 I_3 \frac{\left(\vec{r}_3 \left(\vec{r}_2 - \vec{r}_3\right)\right)}{\left|\vec{r}_3\right|^3 \left|\vec{r}_2 - \vec{r}_3\right|^3} \\
& \left. - \frac{\gamma^2}{2c^2} m_2 m_3 I_1 \frac{1}{\left|\vec{r}_2\right|^3 \left|\vec{r}_3\right|^3} + - \frac{9 \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{\omega}_2\right) \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{\omega}_3\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^5} \right\} \vec{r}_2 \\
& + \left[ \vec{r}_3, \left\{ -\frac{9\gamma}{2c^2} \left( m_2 I_2 + m_3 I_3 \frac{\vec{p}_2 \vec{p}_3}{\left|\vec{r}_2 - \vec{r}_3\right|^5} \right) \right. \right. \\
& - \frac{15\gamma}{2c^2} \frac{\left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{p}_2\right) \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{p}_3\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^7} + \frac{3\gamma}{c^2} \frac{1}{\left|\vec{r}_3\right|^3} \left(m_1 \omega_3^2 I_3 + m_3 \omega_1^2 I_1\right) \\
& + \frac{3\gamma}{c^2} \frac{1}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \left(m_2 \omega_3^2 I_3 + m_3 \omega_2^2 I_2\right) - \frac{3\gamma^2}{2c^2} \frac{\left(m_1 I_3 + m_3 I_1\right)}{\left|\vec{r}_3\right|^5} \\
& - \frac{3\gamma^2}{2c^2} \frac{\left(m_2 I_3 + m_3 I_2\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^5} - \frac{\gamma^2}{c^2} m_1 m_2 I_3 \frac{1}{\left|\vec{r}_3\right|^3 \left|\vec{r}_2 - \vec{r}_3\right|^3} \\
& - \frac{3\gamma^2}{2c^2} m_1 m_2 I_3 \frac{\left(\vec{r}_3 \left(\vec{r}_2 - \vec{r}_3\right)\right)}{\left|\vec{r}_3\right|^5 \left|\vec{r}_2 - \vec{r}_3\right|^3} - \frac{3\gamma^2}{2c^2} m_1 m_2 I_3 \frac{\left(\vec{r}_3 \left(\vec{r}_2 - \vec{r}_3\right)\right)}{\left|\vec{r}_3\right|^3 \left|\vec{r}_2 - \vec{r}_3\right|^3} \\
& + \frac{3\gamma^2}{2c^2} m_2 m_3 I_1 \frac{\left(\vec{r}_2 \vec{r}_3\right)}{\left|\vec{r}_2\right|^3 \left|\vec{r}_3\right|^5} + \frac{3\gamma^2}{2c^2} m_1 m_3 I_2 \frac{\left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{r}_2\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^5 \left|\vec{r}_2\right|^3} \\
& \left. \left. + \frac{12\gamma}{c^2} \frac{I_1 I_3}{\left|\vec{r}_3\right|^5} + \frac{12\gamma}{c^2} \frac{I_2 I_3}{\left|\vec{r}_2 - \vec{r}_3\right|^5} + \frac{9\gamma^2}{2c^2} \frac{\left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{\omega}_2\right) \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{\omega}_3\right) \vec{r}_3}{\left|\vec{r}_2 - \vec{r}_3\right|^5} \right\} \vec{r}_3 \right] \\
& + \left[ \vec{r}_3, \frac{\gamma}{c^2} \left\{ \frac{9 \left(\vec{r}_3 \vec{\omega}_2\right) \left(\vec{r}_3 \vec{\omega}_3\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \vec{r}_3 - \frac{3 \left(\vec{r}_3 \vec{\omega}_3\right)}{\left|\vec{r}_3\right|^3} \vec{\omega}_1 + \frac{3 \left(\vec{r}_3 \vec{\omega}_1\right)}{\left|\vec{r}_3\right|^3} \vec{\omega}_3 + \right. \right. \\
& \left. \left. \frac{3 \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{\omega}_3\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \vec{\omega}_2 + \frac{3 \left(\left(\vec{r}_2 - \vec{r}_3\right) \vec{\omega}_2\right)}{\left|\vec{r}_2 - \vec{r}_3\right|^3} \vec{\omega}_3 \right\} \right]
\end{aligned} \tag{8}$$

where $\vartheta_i$, $r_i$, $p_i$, $\omega_i$ and $I_i$ are the velocity, radius, momentum, angular velocity and inertia momentum of $i$-th mass, respectively [9].

## 3. Equations of motion

Our approach consists in finding the evolutionary equation of motion the test body (third), which describes the average change of its orbital momentum.

$$\dot{\vec{M}}^{(rot)} = -\frac{\gamma}{2c^2}\cdot\frac{m_2I_2+m_3I_3}{\left|\vec{r}_2-\vec{r}_3\right|^3}\left[[-\vec{p}_2,\vec{p}_3]+3\frac{\left(\left(\vec{r}_2-\vec{r}_3\right)\vec{p}_2\right)}{\left|\vec{r}_2-\vec{r}_3\right|^2}\left[\left(\vec{r}_2-\vec{r}_3\right),\vec{p}_3\right]\right]$$
$$+\left\{-\frac{3\gamma}{2c^2}(m_2I_2+m_3I_3)\frac{\vec{p}_2\vec{p}_3}{\left|\vec{r}_2-\vec{r}_3\right|^5}+\frac{15\gamma}{2c^2}\frac{\left(\left(\vec{r}_2-\vec{r}_3\right)\vec{p}_2\right)\cdot\left(\left(\vec{r}_2-\vec{r}_3\right)\vec{p}_3\right)}{\left|\vec{r}_2-\vec{r}_3\right|^7}\right.$$
$$-\frac{3\gamma}{c^2}\frac{1}{\left|\vec{r}_2-\vec{r}_3\right|^3}\left(m_2\omega_3^2I_3+m_3\omega_2^2I_2\right)+\frac{3\gamma^2}{2c^2}\frac{1}{\left|\vec{r}_2-\vec{r}_3\right|^5}\left(m_2I_3+m_3I_2\right)$$
$$+\frac{\gamma^2}{2c^2}m_1m_2I_3\frac{1}{\left|\vec{r}_3\right|^3\left|\vec{r}_2-\vec{r}_3\right|^3}-\frac{3\gamma^2}{2c^2}m_1m_2I_3\frac{\left(\vec{r}_3\left(\vec{r}_2-\vec{r}_3\right)\right)}{\left|\vec{r}_3\right|^3\left|\vec{r}_2-\vec{r}_3\right|^3}-\frac{\gamma^2}{2c^2}m_2m_3I_1\frac{1}{\left|\vec{r}_2\right|^3\left|\vec{r}_3\right|^3}$$
$$\left.+\frac{\gamma}{c^2}\frac{9\left(\left(\vec{r}_2-\vec{r}_3\right)\vec{\omega}_2\right)\left(\left(\vec{r}_2-\vec{r}_3\right)\vec{\omega}_3\right)}{\left|\vec{r}_2-\vec{r}_3\right|^5}\right\}[\vec{r}_3,\vec{r}_2] \tag{9}$$

According to the Eq.(5), the second term of the time derivative of the orbital momentum is:

$$\overline{\dot{\vec{M}}^{(rot)}} = \frac{1}{T}\int_0^T \dot{\vec{M}}^{(rot)}dt. \tag{10}$$

In order to obtain the equations of motion one needs to integrate the Eq.(10) for the repetition period of the system configurations T (synodic period of the test body).

In general, it is a complicated task to integrate the above differential equation. Suppose that the test body is moving along a circular orbit

$$\vec{r}_3 = r_{kep}\left(\vec{i}\cos\omega_3 t+\vec{j}\sin\omega_3 t\right) \tag{11}$$

and the second body

$$\vec{r}_2 = r_2\left(\vec{i}\cos\omega_2 t+\vec{j}\sin\omega_2 t\right) \tag{12}$$

Substituting the radius-vector of the test body in Eq.(10) and integrating the relativistic rotational component we can find that

$$\overline{\dot{\vec{M}}^{(rot)}} = 0 \tag{13}$$

The orbital stability of the test body by definition means the equality to zero of the average change of the angular momentum. As we can see from this expression, in GR, in the first approximation, the circular motion of a test body in the plane of the orbit of the second body in the restricted circular three-body problem is stable

## 4. Conclusions

In this work, we considered the orbital stability problem for the circular motion of a test body in the restricted three-body problem within the framework of GR, when all bodies have their own rotation. We derived the equation of motion from the Lagrange function with the rotational terms and the stability problem of the orbit is investigated by using the adiabatic theory of motion in the GR. Moreover, we derived the explicit form of the corresponding motion equation and the time derivative of momentum. The resulting equation was integrated (averaged) by the sinodic period of the test body, which eliminates perturbations,leaving only evolutionary effects.

The result of the integral is confirmed that the circular motion will be stable if the spins of all bodies are collinear within a plane. The orbital stability of quasicircular orbits of the test body is a task of further investigations.

## Acknowledgments

This work was partially supported by UNAM-DGAPA-PAPIIT, Grant No. 111617, and by the Ministry of Education and Science of RK, Grant No. AP05133630.

## References

1. A. Einstein and L. Infeld, *Can. J. Math.* **1** p.209. (1949).
2. V. Fok. *The theory of space, time and gravitation (2nd English edn.).* (Pergamon, Oxford 1964)
3. Kei Yamada and Hideki Asada *Phys. Rev.* **D 82** 104019 (2010).
4. T. I. Maindl,R. Dvorak. *J.Math. Phys.* **7**, 1137-1143(1966).
5. V.A. Brumberg, *Relativistic Celestial Mechanics, Moscow, Nauka* (1972).
6. M. M. Abdil'din,*Mechanins of Einstein theory of gravity, Alma-Ata: Nauka*, (1988), 198 p.
7. M. M. Abdil'din , M. E. Abishev , N. A. Beissen *Gravit. Cosmol.***15** 141(2009).
8. M.E. Abishev, S. Toktarbay, B. A. Zhami, *Gravit. Cosmol.* **20** 252-254(2014).
9. M.E. Abishev, S. Toktarbay, A.Zh. Abylayeva, A.Z. Talkhat, *EPJ Web of Conferences,*. **168** 04001 (2018).
10. L.D. Landau , E.M. Lifshitz *The Classical Theory of Fields, ( Volume 2 of A Course of Theoretical Physics ) Pergamon Press (1971)*

# On self-force for particles coupled to higher-order electrodynamics and scalar fields

Alan Baza, Angel Harb, Vu Hoang* and Maria Radosz

*Department of Mathematics, University of Texas at San Antonio,*
*One UTSA Circle,*
*San Antonio, TX-78249, United States*
**E-mail: duynguyenvu.hoang@utsa.edu*
*www.math.utsa.edu*

Aaron DeLeon

*Department of Physics and Astronomy, University of Texas at San Antonio,*
*One UTSA Circle,*
*San Antonio, TX-78249, United States*
*www.utsa.edu/physics*

We address the question of point particle motion coupled to classical fields, in the context of scalar fields derived from higher-order Lagrangians and BLTP electrodynamics.

*Keywords*: Point particles, fields, higher-order Lagrangian

## 1. Introduction

Finding a consistent and well-posed dynamical equations of motion for a system of point particles interacting with classical fields is an important problem of relativistic physics. The field could be scalar, gravitational or electromagnetic. For the gravitational field (Einstein-Infeld-Hoffmann problem), the subject's history is reviewed in [7], so we will restrict ourselves to brief remarks concerning electromagnetic interaction. The problem of point particle motion famously started with attempts to model the electron as point particles in classical electrodynamics. The field has been plagued ever since by the appearance of *infinities* for which renormalization procedures are required (see e.g. [6,16]). An approach free of infinities started in 1933 with a paper by Born [4], which was continued by Born and Infeld in [5]. This time, the field equations of electrodynamics were altered to the nonlinear Maxwell–Born–Infeld field equations. The MBI field equations are nonlinear and hence difficult to handle.

Bopp [2,3], Landé–Thomas [10,11], and Podolsky [13,14] went into a different direction by proposing linear, but *higher-order derivative* field equations to remove the infinite field energy problems. Recently the self-force problem in BLTP electrodynamics was studied in quite some generality by J. Gratus, W. Tucker and V. Perlick [9], after more limited earlier studies in [11] and [18]. Recently M. Kiessling and A.S. Tahvildar-Zadeh considered a system of $N$ charged point particles together with their associated BLTP self-fields. Their paper [8] contains the first well-posedness result for the combined Cauchy problem of particles and electrodynamic fields.

In this paper, we report on recent results [1,12] to provide a rigorous framework for point particles interacting via *retarded fields*. As we explain in section 2, the

usual Lagrangian framework does not lead to viable particle equations of motion. An additional principle is needed, which we take to be the conservation of energy and momentum (see also[8]). It turns out that for the higher-order field equations we describe here, a rigorous derivation of the particle equations of motion can be given, which is free of ad-hoc renormalization procedures.

## 2. Scalar Fields

The fundamental issues can be most readily illustrated in the context of scalar fields. We therefore first consider a scalar field $\phi(x^\alpha)$ defined on Minkowski space-time (with Minkowski metric $g_{\alpha\beta}$), interacting with a particle described by its world-line $\tau \mapsto q^\alpha(\tau)$. The particle world-line is associated with a scalar charge density

$$\rho(x^\alpha) = \kappa \int \frac{\delta^{(4)}(x - q(\tau))}{\sqrt{-g}} d\tau \tag{1}$$

where $\kappa$ is coupling constant between particle and field. Note that the scalar current $\rho\dot{q}^\alpha$ is conserved, i.e. $\partial_\alpha(\rho\dot{q}^\alpha) = 0$. The action $S$ of the system is *formally* given by

$$S[\phi, q] = S_{\text{field}} + S_{\text{int}} + S_{\text{particle}}$$

where

$$S_{\text{field}} = -\frac{1}{2}\int d^4x\sqrt{-g}\; g^{\alpha\beta}\partial_\alpha\phi\partial_\beta\phi, \qquad S_{\text{int}} = -\int d^4x\sqrt{-g}\;\phi\rho$$

$$S_{\text{particle}} = \int m(\tau)d\tau$$

Here, $m(\tau)$ is the bare mass of the particle, which for scalar particles has to be taken as time-dependent (see[16]). Formal variation of the action leads to the equations

$$\frac{d}{d\tau}\left(m\dot{q}^\alpha\right) = -\kappa\partial^\alpha\phi\big|_{x^\beta = q^\beta(\tau)} \tag{2}$$

$$\Box\,\phi = -\rho. \tag{3}$$

Now the retarded solution of the wave equation (3), for all $x^\alpha$ not on the world-line, is given by

$$\phi(x^\beta) = \left.\frac{\kappa}{4\pi\dot{q}_\gamma(x^\gamma - q^\gamma)}\right|_{\text{ret}} \tag{4}$$

where "ret" indicates evaluation at the retarded time on the particle world-line, i.e. $\dot{q}_\gamma(x^\gamma - q^\gamma)$ is a light-like Minkowski vector. But combining the equations (2), (3) and (4) makes no apparent sense, since (2) requires evaluation of $\phi, \partial^\alpha\phi$ at the particle's position and (4) is not well-defined there. To remedy this situation, authors have proposed renormalization procedures, avereraging axioms and the use of a combination of retarded and advanced solutions of the field equation (3) (see e.g.[16]). Clearly the Lagrangian formulation does not determine the particle equations of motion uniquely and an additional principle is needed. Preferably, the additional principle should be as fundamental as possible. Here, we consider the

possibility of using *conservation of energy and momentum* to derive the equations of motion. The most convenient, relativistic covariant way to implement this is to define the energy-momentum tensor of the particle by

$$T^{\mu\nu}_{\text{particle}}(x^\alpha) = m_0 \int \frac{u^\mu u^\nu \delta^{(4)}(x - q(\tau))}{\sqrt{-g}}\, d\tau \tag{5}$$

The conservation of energy and momentum is then simply

$$\partial_\mu \left( T^{\mu\nu}_{\text{particle}} + T^{\mu\nu}_{\text{field}} \right) = 0 \tag{6}$$

where $T^{\mu\nu}_{\text{field}}$ is the canonical energy-momentum tensor of the field. In view of (5), we need to be precise on how to mathematically interpret (6). The most straightforward way is to interpret (6) in a distributional sense. So as the definition of the action of $T^{\mu\nu}_{\text{field}}$ on any smooth tensor field $\varphi_{\mu\nu}$ we might take

$$T^{\mu\nu}_{\text{field}}(\varphi_{\mu\nu}) = \int \left[ -\frac{1}{2} g^{\mu\nu} \partial_\alpha \phi \partial^\alpha \phi + \partial^\mu \phi \partial^\nu \phi \right] \varphi_{\mu\nu}\, d^4x \tag{7}$$

where $\phi \in C_0^\infty(\mathbb{R}^4)$ is a smooth test function. The problem, of course, is that the integrand is *not locally integrable* if we plug in the retarded field (4). This is most easily seen for a charge at rest, for which $T^{00} \sim r^{-4}$ with $r$ being the spatial distance between the charge and the point $(x^\alpha)$.

## 3. Higher-Order Scalar Fields

A possible way out without resorting to renormalization is to modify the field equations. It should be kept in mind that other choices of the action are conceivable and a modification of the Lagrangian is a direction worth pursuing. As an example, we would like to mention that there is a large body of literature on modifications of Einstein's general relativity (see e.g.[17]). For now, we consider the modified action

$$S_{\text{field}} = \int d^4x \sqrt{-g}\ \left[ -g^{\alpha\beta} \partial_\alpha \phi \partial_\beta \phi + a(\Box\, \phi)^2 \right], \tag{8}$$

by the higher-order term $(\Box\, \phi)^2 = \left[ \frac{1}{\sqrt{-g}} \partial_\alpha \left( \sqrt{-g} g^{\alpha\beta} \partial_\beta \phi \right) \right]^2$. This modification keeps the field equations linear, but changes the *short-distance physics* in a way explained below. Observe also that (7) is the unique action functional involving second derivatives of $\phi$ that leads to linear field equations. The only other covariant quadratic contribution involving second order derivatives would be

$$g^{\alpha\gamma} g^{\beta\rho} (\partial_\alpha \partial_\beta \phi)(\partial_\gamma \partial_\rho \phi)$$

but this would reduce to $(\Box\, \phi)^2$ after an integration by parts. The resulting field equation for $\phi$ is

$$\Box\, (\phi + a \Box\, \phi) = -\rho. \tag{9}$$

where $\rho$ is given by (1). It turns out that the higher-order equation (9) has a regularizing effect on fields with point sources. The retarded solution $\phi$ of (9) can

be determined by Green's functions techniques and has the property that both $\phi$ and $\partial_\beta\phi$ are now bounded in the vicinity of the particle, although the value of $\partial_\beta\phi$ on the particle world-line is still ill-defined.

Recall that the canonical energy-momentum tensor is given by the variation of $S_{\text{field}}$ with respect to $g^{\mu\nu}$:

$$\delta S_{\text{field}} = -\int d^4x\sqrt{-g}\ (T_{\text{field}})_{\mu\nu}g^{\mu\nu} \tag{10}$$

A lengthy calculation gives (for $g_{\alpha\beta}$ being the Minkowski metric):

$$\begin{aligned} T^{\alpha\beta}_{\text{field}} = &-\frac{1}{2}g^{\alpha\beta}\partial_\mu\phi\partial^\mu\phi + \partial^\alpha\phi\partial^\beta\phi \\ &+ 2a\partial^\alpha\Box\,\phi\partial^\beta\phi - \frac{a}{2}(\Box\,\phi)^2 - a\partial_\mu(\Box\,\phi)(\partial^\mu\phi)g^{\alpha\beta} \end{aligned} \tag{11}$$

The energy-momentum tensor of the field can now be defined as a distribution, since it contains singularities of order at most $r^{-2}$. A thorough analysis of (6) now yields a well-defined equation of motion (see[1]), from which the "self-force" contribution can be read off. The form of the self-force can depend on the type of problem considered (e.g. a scattering problem or initial-value problem for particles and fields).

## 4. BLTP Electrodynamics

Higher-order field linear equations for electrodynamics were proposed by Bopp[2,3], Landé–Thomas[10,11], and Podolsky[13,14]. Starting point is the usual Maxwell action with an additional term containing second-order derivatives of the field tensor $F_{\mu\nu}$:

$$S_{\text{BLTP}} = \int d^4x\sqrt{-g}\ \left[-\frac{1}{16\pi}F_{\alpha\beta}F^{\alpha\beta} - \frac{1}{8\pi\kappa^2}\partial_\gamma F^{\gamma\alpha}\partial^\lambda F_{\lambda\alpha}\right] \tag{12}$$

This is the unique action giving linear field equations of order less or equal to four. $\kappa > 0$ is a fixed parameter throughout. Variation of the action with respect to the vector potential $A_\alpha$ gives

$$\begin{aligned} (I - \kappa^{-2}\Box\,)\partial^\alpha F_{\alpha\beta} &= -4\pi j_\beta \\ \partial_\alpha F_{\beta\gamma} + \partial_\beta F_{\gamma\alpha} + \partial_\gamma F_{\alpha\beta} &= 0 \end{aligned} \tag{13}$$

where $j_\alpha$ is usual particle 4-current vector. We will refer to a theory based on (13) as BLTP electrodynamics. As before, the Lagrangian formulation does not determine the particle equations of motion and we additionally impose (6) in a distributional sense. The canonical energy-momentum tensor of the field is given by

$$\begin{aligned} -4\pi T^{\alpha\gamma}_{\text{BLTP}} := &g^{\alpha\rho}F_{\rho\mu}F^{\mu\gamma} + \frac{1}{4}g^{\alpha\gamma}F_{\rho\theta}F^{\rho\theta} - \kappa^{-2}(g^{\alpha\rho}F_{\rho\mu}\Box\,F^{\mu\gamma} + g^{\alpha\rho}F^{\gamma\mu}\Box\,F_{\mu\rho} \\ &+ g^{\alpha\rho}\partial_\mu F^{\mu\gamma}\partial^\beta F_{\rho\beta}) - \frac{1}{2}\kappa^{-2}(F_{\rho\theta}\Box\,F^{\rho\theta} + \partial_\rho F^{\rho\theta}\partial^\beta F_{\beta\theta})g^{\alpha\gamma} \end{aligned} \tag{14}$$

and it can be shown that

$$\varphi_{\mu\nu} \mapsto \int T^{\mu\nu}_{\text{BLTP}} \varphi_{\mu\nu} \; d^4x$$

defines a well-defined distribution. In [12], we prove the following Theorem:

**Theorem 4.1.** *Suppose* $\{q_n^\alpha(\tau_n)\}_{n=1}^N$ *are a collection of non-intersecting world-lines satisfying certain conditions and let the field associated to the n-th particle be defined by*

$$(F_n)_{\alpha\beta}(x^\alpha) = e_n\kappa^2 \left.\frac{R_{[\alpha}u_{\beta]}}{2u^\gamma R_\gamma}\right|_{ret} + e_n\kappa^2 \int_{-\infty}^{\tau_{ret}(x^\alpha)} \frac{J_2(\kappa D) R_{[\alpha}u_{\beta]}}{D^2} d\tau \tag{15}$$

*where* $e_n$ *is the charge of the n-th particle and the integral extends over the world-line* $q_n(\tau)$*. R is defined by* $R^\alpha(x^\beta, \tau) = x^\alpha - q^\alpha(\tau)$ *and* $D = (q_n^\gamma(\tau) - q_n^\gamma(\tau'))((q_n)_\gamma(\tau) - (q_n)_\gamma(\tau'))$*. Suppose that* (6) *holds. Then*

$$m_n \frac{du_n^\alpha}{d\tau_n} = e_n \left( \mathcal{F}_n^{\alpha\beta} + \sum_{m\neq n} F_m^{\alpha\beta} \right) u_\beta \tag{16}$$

*holds. Note in particular that the force on the n-th particle splits into a self-force* $\mathcal{F}^{\alpha\beta}u_\beta$ *and the Lorentz-force exerted by all other particles. The self-field* $\mathcal{F}^{\alpha\beta}$ *is given by the formula*

$$(\mathcal{F}_n)_{\alpha\beta}(\tau) = e_n\kappa^2 \int_{-\infty}^{\tau} \frac{J_2(\kappa D) R_{[\alpha}u_{\beta]}}{D^2} d\tau'. \tag{17}$$

More details can be found in [12]. The class of world-lines covered by the Theorem is extremely broad, comprising essentially arbitrary subluminal world-lines that are twice differentiable. Hence in the context of BLTP electrodynamics, the equations of motion rigorously follow from (6).

## Acknowledgments

The authors wish to thank M. Kiessling and A. S. Tahvildar-Zadeh for a great number of stimulating discussions and for inviting VH to participate in the session "The Einstein-Infeld-Hoffmann Legacy in Mathematical Relativity" at MG'15. VH acknowledges National Science Foundation support under grants DMS-1614797 and NSF DMS-1810687.

## References

1. Baza, A., DeLeon, A., Harb, A., Hoang, V. , Particle motion coupled to higher-order scalar fields. *In preparation.*
2. Bopp, F., Eine lineare Theorie des Elektrons, *Annalen Phys.* **430**, 345–384 (1940).

3. Bopp, F., Lineare Theorie des Elektrons. II, *Annalen Phys.* **434**, 573–608 (1943).
4. Born, M., Modified field equations with a finite radius of the electron, *Nature 132*, 282 (1933).
5. Born, M. and Infeld, L, Foundations of the new field theory. *Proc. R. Soc. London A* 144, 425-451 (1934)
6. Dirac, P.A.M. *Proc. R. Soc. London A* 167, 148 (1938).
7. Kiessling, M., The Einstein-Infeld-Hoffmann Legacy in Mathematical Relativity I: The Classical Motion of Charged Point Particles, *Proceedings of the 15th Marcel Grossmann meeting.*
8. Kiessling, M., Tahvildar-Zadeh, A.S., BLTP electrodynamics as an initial-value problem. *In preparation.*
9. Gratus, J., Perlick, V., and Tucker, R.W., On the self-force in Bopp-Podolsky electrodynamics, *J. Phys. A, Math. Theor.* 48, 401-435 (28pp.) (2015).
10. Landé, A., Finite Self-Energies in Radiation Theory. Part I, *Phys. Rev.* 60, 121-126 (1941).
11. Landé, A., and Thomas, L.H., Finite Self-Energies in Radiation Theory. Part II, *Phys. Rev.* 60, 514-523 (1941).
12. Hoang, V. and Radosz, M. , On self-force in higher-order electrodynamics. Preprint `https,//arxiv.org/abs/1902.06386`.
13. Podolsky, B., A generalized electrodynamics. Part I, Non-quantum, *Phys. Rev.* 62, 68-71 (1942).
14. Podolsky, B., and Schweb, P., A review of generalized electrodynamics, *Rev. Mod. Phys.* 20, 40-50 (1948)
15. Poisson, E., A Relativist's Toolkit, the Mathematics of Black-Hole Mechanics. Cambridge University Press (2004).
16. Poisson, E., Pound, A., and Vega, I., The motion of point particles in curved spacetime, *Living Rev. Rel.* **14**,7(190) (2011).
17. Sotiriou, T.P, f(R) theories of gravity. *Rev.Mod.Phys.* 82 (2010) 451-497.
18. Zayats, A.E., Self-interaction in the Bopp–Podolsky electrodynamics, Can the observable mass of a charged particle depend on its acceleration?, *Annals Phys. (NY)* ***342****, 11–20 (2014).*

# The Einstein-Infeld-Hoffmann legacy in mathematical relativity I: The classical motion of charged point particles

Michael K.-H. Kiessling* and A. Shadi Tahvildar-Zadeh

*Department of Mathematics, Rutgers University*
*110 Frelinghuysen Rd., Piscataway, NJ 08854, USA*
**E-mail: miki@math.rutgers.edu*

Einstein, Infeld, and Hoffmann (EIH) claimed that the field equations of general relativity theory alone imply the equations of motion of neutral matter particles, viewed as point singularities in space-like slices of spacetime; they also claimed that they had generalized their results to charged point singularities. While their analysis falls apart upon closer scrutiny, the key idea merits our attention. This rapport identifies necessary conditions for a well-defined general-relativistic joint initial value problem of $N$ classical point charges and their electromagnetic and gravitational fields. Among them, in particular, is the requirement that the electromagnetic vacuum law guarantees a finite field energy-momentum of a point charge. This disqualifies the Maxwell(–Lorentz) law used by EIH. On the positive side, if the electromagnetic vacuum law of Bopp, Landé–Thomas, and Podolsky (BLTP) is used, and the singularities equipped with a non-zero bare rest mass, then a joint initial value problem can be formulated in the spirit of the EIH proposal, and shown to be locally well-posed — *in the special-relativistic zero-G limit*. With gravitational coupling (i.e. $G > 0$), though, changing Maxwell's into the BLTP law and assigning a bare rest mass to the singularities is by itself not sufficient to obtain even a merely well-defined joint initial value problem: the gravitational coupling also needs to be changed, conceivably in the manner of Jordan and Brans–Dicke.

## 1. Brief History and State of Affairs

> "I am plaguing myself with the derivation of the equations of motion of material points, conceived of as singularities [in the gravitational field], from the equations of general relativity." **Albert Einstein**, in a letter to **Max Born** on Dec. 4, 1926.

We don't know when **Einstein** first conceived of the notion of point particles as singularities in relativistic fields,[a] but his letter to Max Born makes it plain that by the end of 1926 his ideas had matured to the point where he pursued a dynamical theory for such point singularities, expecting that their law of motion could be extracted from his gravitational field equations. Already a month later **Einstein** & **Grommer** announced that "the law of motion is completely determined by the field equations, though shown in this work only for the case of equilibrium." In that paper[7] the case of a static, spherically symmetric spacetime with a single time-like singularity was studied. The truly dynamical many-body problem was treated a decade later by **Einstein, Infeld**, and **Hoffmann** in their famous paper Ref. 8, with follow-ups in Refs. 9, 10. They argued explicitly that the field equations of general relativity theory alone determine the equations of motion of neutral matter

[a] In 1909 he remarked that "light quanta" might be point singularities in "a field," their motion being guided by the electromagnetic field. See part II of our rapport.

particles, viewed as point singularities in space-like slices of spacetime. They also claimed that they had generalized their results to charged point-singularities, with the details written up in a set of notes deposited with the secretary of the IAS. In 1941 the motion of charged point-singularities was revisited by Infeld's student **P. R. Wallace**, who presented the details of the calculations in Ref. 28.

Here is the gist of the **Einstein-Infeld-Hoffmann** argument (modern terminology):

- Suppose you have a four-dimensional, time-oriented, asymptotically flat electromagnetic **Lorentz** spacetime $\mathcal{M}^{1,3}$ with $N$ charged, time-like singularities of infinite extent, presentable as a graph over $\mathbb{R}^{1,3} \setminus \{N \text{ time-like world-lines}\}$.
- Away from the singularities the spacetime structure obeys **Einstein**'s equations

$$\mathbf{R} - \tfrac{1}{2}R\mathbf{g} = \tfrac{8\pi G}{c^4}\mathbf{T}[\mathbf{F}, \mathbf{g}], \tag{1}$$

  where $\mathbf{T}[\mathbf{F}, \mathbf{g}]$ is the energy-momentum-stress tensor of the electromagnetic vacuum field $\mathbf{F}$, satisfying **Maxwell**'s field equations in vacuum,

$$\boldsymbol{d}\mathbf{F} = \mathbf{0} \qquad \& \qquad \boldsymbol{d}{*}\mathbf{F} = \mathbf{0}. \tag{2}$$

- The twice contracted second **Bianchi** identity implies energy-momentum conservation:

$$\boldsymbol{\nabla}\cdot\left(\mathbf{R} - \tfrac{1}{2}R\mathbf{g}\right) = \mathbf{0} \implies \boldsymbol{\nabla}\cdot\mathbf{T}[\mathbf{F}, \mathbf{g}] = \mathbf{0}. \tag{3}$$

- "Massive, charged" singularities are associated with field "fluxes," and thus

$$\boxed{\boldsymbol{\nabla}\cdot\mathbf{T}[\mathbf{F}, \mathbf{g}] = \mathbf{0} \quad \& \quad \text{flux conditions}} \implies \text{law of the time-like singularities.} \tag{4}$$

The main bullet point is of course the last one. Here are **EIH** in their own words (p.66): "It is shown that for two-dimensional [closed] spatial surfaces containing singularities , certain surface integral conditions are valid which determine the motion ." [Emphasis ours.]

Unfortunately, despite its publication in the Annals of Mathematics, the 1938 **EIH** paper is not only not rigorous, it contains questionable technical assumptions and serious blunders. Some were addressed in Ref. 10, yet *their main conclusions turn out to be false*. All the same, the core idea merits a deeper inquiry.

The issue is how to correctly handle singularities. **EIH** state (p.66): "By means of a new method of approximation, specially suited to the treatment of quasi-stationary fields, the gravitational field due to moving particles is determined." This approximation method assumes that the particles are moving *slowly* and the field strengths (as seen in a **Lorentz** frame of **Minkowski** space, in which the particles move slowly) are *weak*. As a consequence, one has to choose the radii of the closed surfaces sufficiently large to satisfy the weak-field assumption needed for the

convergence of the expansion (which was not shown). Yet, on p. 92 one reads: "It is most convenient to take definite, infinitesimally small spheres whose centers are at the singularities, ..." which clearly violates their weak-field condition. Indeed, **EIH** realize that: "... in this case infinities of the types

$$\lim const./r^n, \qquad n \text{ a positive integer}, \qquad r \to 0 \tag{5}$$

can occur in the values of the partial integrals," but then commit a major blunder by stating (**EIH**, p. 92): "Since these must cancel, however, in the final result, we may merely ignore them throughout the calculation of the surface integrals."

*Alas, the infinities do not cancel!*

In the same year **P. A. M. Dirac** invented *negative infinite bare mass renormalization* to handle those infinities, in the simpler special-relativistic purely electrodynamical setting.[6] For an electron with positive "observable mass" $m_{\rm obs}$ and charge $-e$ he obtained the equation of motion (in **Misner**–**Thorne**–**Wheeler** notation)

$$m_{\rm obs}\tfrac{\mathrm{d}^2}{\mathrm{d}\tau^2}\mathbf{q} = \mathbf{f}^{\rm ext} + \mathbf{f}^{\rm LAUE}, \tag{6}$$

where

$$\mathbf{f}^{\rm ext} = -\tfrac{e}{c}\mathbf{F}^{\rm ext}(\mathbf{q}) \cdot \tfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{q} \tag{7}$$

is a **Lorentz Minkowski**-force due to "external sources,"

$$\mathbf{f}^{\rm LAUE} = \tfrac{2e^2}{3c^3}\left(\mathbf{g} + \tfrac{1}{c^2}\tfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{q} \otimes \tfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{q}\right) \cdot \tfrac{\mathrm{d}^3}{\mathrm{d}\tau^3}\mathbf{q} \tag{8}$$

is **von Laue**'s radiation-reaction **Minkowski**-force of the electron, and

$$m_{\rm obs} = \lim_{r\downarrow 0}\left(m_{\rm b}(r) + \tfrac{e^2}{2c^2}\tfrac{1}{r}\right) \tag{9}$$

defines $m_{\rm b}(r)$. [N.B.: $m_{\rm b}(r) \downarrow -\infty$ as $r \downarrow 0$]. Here, $r$ is the radius of a sphere in the instantaneous rest-frame of the electron, centered at the electron, which plays the role of the surfaces containing singularities invoked by **EIH**.

As is well-known, such mass-renormalization computations have become the template for the much more elaborate — and quite successful — renormalization group computations in quantum electrodynamics (and more generally, quantum field theory). Be that as it may, **Dirac** himself later in life became very dissatisfied with this approach, and so are we.

First of all, supposing a point electron has a bare mass, then how could it possibly depend on the radius $r$ of a sphere over which a theoretical physicist averages the fields?

Second, the third proper time derivative featuring in the **von Laue Minkowski**-force means that (6) is a third-order ODE for the position of the particle as a function of proper time, requiring vector initial data for position, velocity, and acceleration. Yet a classical theory of point particle motion should only involve initial data for position and velocity.

In 1951 **Lev Landau** & **Eugenii Lifshitz** addressed the $\dddot{\mathbf{q}}$ problem as follows:

- Test particle theory works well for many practical purposes.
- In such situations **von Laue**'s $\ddot{\mathbf{q}}$ force term must be a small perturbation of $\mathbf{f}^{\text{ext}}$.
- Compute $\dddot{\mathbf{q}}$ perturbatively: take the proper time derivative of the test-particle law,

$$\tfrac{\mathrm{d}^3}{\mathrm{d}\tau^3}\mathbf{q} \approx -\tfrac{e}{m_{\text{obs}}c}\tfrac{\mathrm{d}}{\mathrm{d}\tau}\left(\mathbf{F}^{\text{ext}}(\mathbf{q})\cdot\tfrac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{q}\right). \tag{10}$$

- The right-hand side depends only on $\mathbf{q}$, $\dot{\mathbf{q}}$, $\ddot{\mathbf{q}}$. Substitute it for $\frac{\mathrm{d}^3}{\mathrm{d}\tau^3}\mathbf{q}$ in (8).

R.h.s.(8) with r.h.s.(10) substituted for $\frac{\mathrm{d}^3}{\mathrm{d}\tau^3}\mathbf{q}$ will be called the **Landau–Lifshitz Minkowski**-force of radiation-reaction, denoted $\mathbf{f}^{\text{LL}}$. Equation (6) with $\mathbf{f}^{\text{Laue}}$ replaced by $\mathbf{f}^{\text{LL}}$ is known as the **Landau–Lifshitz** equation of motion for the electron. It seems to work quite well for practical purposes in which $\mathbf{F}^{\text{ext}}$ can be approximated by some smooth field tensor, on time scales beyond the one where test particle theory works well, but not arbitrarily far beyond.[25] However[4], as soon as $\mathbf{F}^{\text{ext}}$ is taken to be the field generated by all other particles the **Landau–Lifschitz** equation of motion is typically well-posed only until the moment that a point charge meets the forward initial light cone of another point charge, a ludicrously short time span!

**Dirac**'s idea of infinite negative bare mass renormalization and **Landau–Lifshitz**'s perturbative treatment of the **von Laue** radiation-reaction **Minkowski**-force have become standard ingredients also in general-relativistic treatments of charged point particle motion. Thus, for a point electron moving in a given curved background, **Eric Poisson, Adam Pound**, & **Ian Vega** in their review Ref. 23 present the following equations of motion:

$$m_{\text{obs}}\tfrac{\mathrm{D}}{\mathrm{d}\tau}\mathbf{u} = \mathbf{f}^{\text{ext}} + \mathbf{f}^{\text{Laue}} + \mathbf{f}^{\text{tail}}, \tag{11}$$

where $\mathbf{u} := \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{q}$ and $\frac{\mathrm{D}}{\mathrm{d}\tau}\mathbf{u} := \frac{\mathrm{d}}{\mathrm{d}\tau}\mathbf{u} + \mathbf{\Gamma}^{\text{ext}}(\mathbf{u},\mathbf{u})$, and $\mathbf{f}^{\text{ext}} = -\frac{e}{c}\mathbf{F}^{\text{ext}}(\mathbf{q})\cdot\mathbf{u}$ as before, but now

$$\mathbf{f}^{\text{Laue}} = \tfrac{2}{3}e^2\left(\mathbf{g} + \tfrac{1}{c^2}\mathbf{u}\otimes\mathbf{u}\right)\cdot\left(\tfrac{1}{6}\mathbf{R}^{\text{ext}}\cdot\tfrac{1}{c}\mathbf{u} + \tfrac{1}{c^3}\tfrac{\mathrm{D}^2}{\mathrm{d}\tau^2}\mathbf{u}\right) \tag{12}$$

with

$$\tfrac{\mathrm{D}^2}{\mathrm{d}\tau^2}\mathbf{u} \approx -\tfrac{e}{m_{\text{obs}}c}\tfrac{\mathrm{D}}{\mathrm{d}\tau}\left(\mathbf{F}^{\text{ext}}(\mathbf{q})\cdot\mathbf{u}\right), \tag{13}$$

and

$$\mathbf{f}^{\text{tail}} = 2e^2\int_{-\infty}^{\tau}\boldsymbol{H}^{\text{ret}}(\mathbf{q}(\tau),\mathbf{q}(\tau'))\cdot\mathbf{u}(\tau')\mathrm{d}\tau'\cdot\mathbf{u}(\tau), \tag{14}$$

where $\boldsymbol{H}^{\text{ret}}(\mathbf{q}(\tau),\mathbf{q}(\tau'))$ is a retarded type of Green function for the electromagnetic field tensor in curved spacetime. Equation (11) does not yet include gravitational radiation-reaction, which **Poisson, Pound**, & **Vega** discuss also, in particular the approaches of **Quinn** & **Wald** and **Detweiler** & **Whiting** (see Ref. 23), but we don't need to go there because (11) already displays a major problem due to the so-called tail force term:

*Equation (11), even with r.h.s.(13) substituted for* $\frac{\mathrm{D}^2}{\mathrm{d}\tau^2}\mathbf{u}$ *at r.h.s.(12), does not pose a second-order initial value problem for the position of the point electron but instead requires the input of the entire past history of the motion!*

One can try to extricate oneself from this dilemma by once again having recourse to a **Landau**–**Lifshitz**-type perturbation argument: the tail force, also a radiation-reaction term, must be small in situations where test particle theory works well. In this case, backward from the initial instant (say at $\tau = 0$) one can approximately replace $\mathbf{q}(\tau')$ and $\mathbf{u}(\tau')$ in the integrand by the pertinent expressions computed from test particle theory, with particle data for position and velocity given, and then treat this so-modified equation as a second-order integro-differential equation from the initial instant on forward, with $\mathbf{q}(\tau')$ and $\mathbf{u}(\tau')$ in the integrand for $\tau' > 0$ no longer test-particle expressions. This set of "effective equations of motion" may work well in many practical situations.

Yet from a mathematical relativity point of view this state of affairs is very unsatisfactory, both technically (being non-rigorous) and conceptually (involving heuristic but arbitrary arguments).

## 2. Rigorous Approach

In the following we report on recent rigorous advances in formulating a joint initial value problem for classical charged point particles and the electromagnetic and gravitational fields they generate, with the key idea of the 1938 **EIH** paper, as outlined on p.1 of this rapport, serving as our point of departure. To avoid the mistakes made by **EIH**, we inquire into necessary conditions on the energy-momentum-stress tensor which allow one to extract a law of motion associated with the time-like singularities from an equation like (4) without invoking infinite mass renormalization, nor arbitrary averaging over some neighborhood of an a-priori ill-defined force field, as done in Ref. 23. For the simpler special-relativistic zero-gravity limit (cf. Ref. 16, 17) we even state a well-posedness theorem, so this case is treated first.

### 2.1. *The zero-G Limit*

#### 2.1.1. *Time-like particle world-lines in Minkowski spacetime*

In the limit $G \downarrow 0$, (1) is solved by $\mathcal{M}^{1,3} = \mathbb{R}^{1,3} \setminus \{N \text{ time-like world-lines}\}$, with metric $\mathbf{g} = \boldsymbol{\eta}$ away from the world-lines. The question is which conditions on $\mathsf{T}$ lead, in a mathematically clean way, to the **EIH**-type conclusion

$$\boxed{\boldsymbol{\nabla}\cdot\mathsf{T}[\mathsf{F},\boldsymbol{\eta}] = \mathbf{0} \quad \& \quad \text{flux conditions}} \implies \text{law of the time-like world-lines.} \qquad (1)$$

To answer this question, we extend $\mathcal{M}^{1,3}$ continuously to $\mathbb{R}^{1,3}$ (by adding the time-like world-lines), and switch to a distributional formulation. We can formulate everything for $N$ charged, massive, time-like world-lines, but for simplicity we set $N = 1$ in the following.

Since the issue is the formulation of a well-posed initial value problem, we also choose an arbitrary Lorentz frame, with space vector $\boldsymbol{s} \in \mathbb{R}^3$ and time $t \in \mathbb{R}$. Then the space part of l.h.s.(1) becomes the *local conservation law for the total momentum*,

$$\tfrac{\partial}{\partial t}\boldsymbol{\Pi}(t,\boldsymbol{s}) + \nabla \cdot T(t,\boldsymbol{s}) = \boldsymbol{0}, \tag{2}$$

where it is *postulated* (equivalent to "minimal coupling") that

$$\boldsymbol{\Pi}(t,\boldsymbol{s}) := \boldsymbol{\Pi}^{\text{field}}(t,\boldsymbol{s}) + \boldsymbol{\Pi}^{\text{charge}}(t,\boldsymbol{s}) \tag{3}$$

is the total momentum vector-density, with $\boldsymbol{\Pi}^{\text{field}}(t,\boldsymbol{s})$ the contribution from the field and $\boldsymbol{\Pi}^{\text{charge}}(t,\boldsymbol{s})$ (a distribution) the usual contribution from the point charge, which must be assigned a non-vanishing bare rest mass $m_{\text{b}}$. Similarly, it is *postulated* that

$$T(t,\boldsymbol{s}) := T^{\text{field}}(t,\boldsymbol{s}) + T^{\text{charge}}(t,\boldsymbol{s}) \tag{4}$$

is the symmetric total stress tensor, with $T^{\text{charge}}(t,\boldsymbol{s})$ the usual stress tensor of the point particle, and $T^{\text{field}}(t,\boldsymbol{s})$ that of the field (except for our unconventional choice of sign!)

Incidentally, energy conservation follows as a corollary from momentum conservation.

Next, we integrate (2) over all $\boldsymbol{s} \in \mathbb{R}^3$, at $t \in \mathbb{R}$. This yields total momentum conservation

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) = -\tfrac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3} \boldsymbol{\Pi}^{\text{field}}(t,\boldsymbol{s})d^3s. \tag{5}$$

Clearly, for (5) to make sense, the field momentum vector-density $\boldsymbol{\Pi}^{\text{field}}(t,\boldsymbol{s})$ has to be integrable over $\mathbb{R}^3$, and this integral differentiable in time. This rules out the **Maxwell**–**Lorentz** field equations, but leaves other options, notably the **M**$_{\text{axwell}}$–**B**$_{\text{orn}}$–**I**$_{\text{nfeld}}$ (**MBI**) and **M**$_{\text{axwell}}$–**B**$_{\text{opp}}$–**L**$_{\text{andé}}$–**T**$_{\text{homas}}$–**P**$_{\text{odolsky}}$ (**MBLTP**) field equations (see below).

Comparing equation (5) with **Newton**'s law for the rate of change of momentum,

$$\tfrac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{p}(t) = \boldsymbol{f}(t), \tag{6}$$

where $\boldsymbol{f}(t)$ is the force acting on the particle at time $t$, it is clear that the force $\boldsymbol{f}(t)$ needs to be extracted from r.h.s.(5). Since the particle's bare momentum $\boldsymbol{p}(t)$ is given in terms of its bare mass $m_{\text{b}}$ and velocity $\boldsymbol{v}(t) = \frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{q}(t)$ by the **Einstein-Lorentz-Poincaré** law

$$\boldsymbol{p}(t) := m_{\text{b}}\frac{\boldsymbol{v}(t)}{\sqrt{1-\frac{1}{c^2}|\boldsymbol{v}(t)|^2}}, \tag{7}$$

the expression for the force also has to be compatible with the requirements of a second-order initial value problem for the position of the point particle! Thus, beside the existence of the time-derivative of the space integral over the field momentum vector-density, it is important that the result involves, initially, only the initial

electromagnetic fields and the initial data for position and velocity of the point particle and, at later times $t > 0$, at most the history of position, velocity, and acceleration of the particle, and of the fields, from the initial instant on, yet not beyond $t$. Whenever this is possible we obtain a well-defined joint initial value problem for field and particle, which may or may not be well-posed.

We were able to explicitly extract a well-defined force on the point charge from the **MBLTP** field equations, and to prove that the resulting joint initial value problem for charge and field is well-posed. We expect this to be feasible also for the **MBI** field equations, but their formidable nonlinearity makes rigorous progress a slow process.

Common to all these classical systems of electromagnetic field equations are the *pre-metric* ***Maxwell–Lorentz*** *field equations*, viz. the evolution equations

$$\partial_t \mathbf{B}(t,\boldsymbol{s}) = -c\boldsymbol{\nabla} \times \mathbf{E}(t,\boldsymbol{s}) \tag{8}$$

$$\partial_t \mathbf{D}(t,\boldsymbol{s}) = +c\boldsymbol{\nabla} \times \mathbf{H}(t,\boldsymbol{s}) + 4\pi e\dot{\boldsymbol{q}}(t)\delta_{\boldsymbol{q}(t)}(s) \tag{9}$$

and the constraint equations

$$\boldsymbol{\nabla} \cdot \mathbf{B}(t,\boldsymbol{s}) = 0 \tag{10}$$

$$\boldsymbol{\nabla} \cdot \mathbf{D}(t,\boldsymbol{s}) = -4\pi e\delta_{\boldsymbol{q}(t)}(s) \tag{11}$$

for the **B**, **D** fields. They differ in the *Electromagnetic Vacuum Law*: $(\mathbf{B},\mathbf{D}) \leftrightarrow (\mathbf{H},\mathbf{E})$. The **Born-Infeld** law[2] reads

$$\mathbf{H} = \frac{\mathbf{B} - \frac{1}{b^2}\mathbf{D}\times(\mathbf{D}\times\mathbf{B})}{\sqrt{1+\frac{1}{b^2}(|\mathbf{B}|^2+|\mathbf{D}|^2)+\frac{1}{b^4}|\mathbf{B}\times\mathbf{D}|^2}} \tag{12}$$

$$\mathbf{E} = \frac{\mathbf{D} - \frac{1}{b^2}\mathbf{B}\times(\mathbf{B}\times\mathbf{D})}{\sqrt{1+\frac{1}{b^2}(|\mathbf{B}|^2+|\mathbf{D}|^2)+\frac{1}{b^4}|\mathbf{B}\times\mathbf{D}|^2}} \tag{13}$$

The **Bopp–Landé–Thomas–Podolsky** law[1,19,20,22] reads

$$\mathbf{H}(t,\boldsymbol{s}) = \left(1+\varkappa^{-2}\Box\right)\mathbf{B}(t,\boldsymbol{s}) \tag{14}$$

$$\mathbf{D}(t,\boldsymbol{s}) = \left(1+\varkappa^{-2}\Box\right)\mathbf{E}(t,\boldsymbol{s})\,. \tag{15}$$

(N.B.: $\Box := c^{-2}\partial_t^2 - \Delta$.) When $b \to \infty$, respectively when $\varkappa \to \infty$, both these vacuum laws reduce to the **Maxwell-Lorentz** law $\mathbf{H} = \mathbf{B} \quad \& \quad \mathbf{E} = \mathbf{D}$.

For given particle motions with subluminal velocity $|\dot{\boldsymbol{q}}(t)| < c$, the field **Cauchy** problems are globally well-posed in the sense of distributions for both the **ML** and **MBLTP** field equations. The **MBI** field Cauchy problem, unfortunately, has not yet been conquered to the extent needed. Global well-posedness of the classical initial value problem has only been shown with small data (no charges!) in Ref. 24; **F. Pasqualotto** presented an extension of **Speck**'s result to **MBI** field evolutions on the **Schwarzschild** background.[21] A local well-posedness result for **MBI** field evolutions with subluminal point sources (and inevitably large data) should be possible, but so far only the special case of electrostatic solutions with $N$ point

charge sources has been conquered in Ref. 15. There it was shown that a unique finite-energy electrostatic weak solution of the **MBI** field equations with $N$ point charges placed anywhere in $\mathbb{R}^3$ exists, and that the solution is real analytic away from the point charges for any choices of their signs and magnitudes.

For the field momentum densities $\mathbf{\Pi}$ (dropping the superscript "field") one has the following expressions. For the **ML** and for **MBI** field equations,

$$4\pi c\mathbf{\Pi} = \mathbf{D} \times \mathbf{B}, \tag{16}$$

whereas for the **MBLTP** field equations,

$$4\pi c\mathbf{\Pi} = \mathbf{D} \times \mathbf{B} + \mathbf{E} \times \mathbf{H} - \mathbf{E} \times \mathbf{B} - \varkappa^{-2}\big(\nabla \cdot \mathbf{E}\big)\big(\nabla \times \mathbf{B} - \varkappa \dot{\mathbf{E}}\big). \tag{17}$$

For **MBLTP** field evolutions with point sources we showed that $\mathbf{\Pi}(t, \boldsymbol{s})$ is in $L^1_{loc}(\mathbb{R}^3)$ about each $\boldsymbol{q}(t)$, see Ref. 17. We expect such a result also for **MBI** fields. It is surely false for **ML** fields!

With appropriate decay rates at spatial infinity imposed on the field initial data, the **MBLTP** field momentum $\int_{\mathbb{R}^3} \mathbf{\Pi}(t, \boldsymbol{s})\mathrm{d}^3 s$ exists for all $t$. Moreover, given **Lipschitz** maps $t \mapsto \boldsymbol{q}(t)$, $t \mapsto \boldsymbol{v}(t)$ and bounded $t \mapsto \boldsymbol{a}(t)$, we showed that $\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3} \mathbf{\Pi}(t, \boldsymbol{s})\mathrm{d}^3 s$ exists for all $t$.

The crucial step in showing that (5), with (7), yields an equation of motion is now the following. The fields $\mathbf{B}, \mathbf{D}, \mathbf{E}, \dot{\mathbf{E}}$ (and $\mathbf{H}$) at $(t, \boldsymbol{s})$ are given by explicit functionals of the vector functions $\boldsymbol{q}(\cdot)$ and $\boldsymbol{v}(\cdot)$, and $\mathbf{D}$ & $\mathbf{H}$ also involve $\mathbf{a}(\cdot)$; their dependence on $\mathbf{a}(\cdot)$ is linear. For $t < 0$ we set $\boldsymbol{q}(t) = \boldsymbol{q}(0) + \boldsymbol{v}(0)t$ and $\boldsymbol{v}(t) = \boldsymbol{v}(0)$, and $\boldsymbol{a}(t) = \mathbf{0}$. Treating $\boldsymbol{q}(\cdot)$ and $\boldsymbol{v}(\cdot)$ as given, and $\mathbf{a}(\cdot)$ as independent vector function variable for $t > 0$, (5) together with (7) is equivalent to a **Volterra** integral equation for $\mathbf{a} = \mathbf{a}[\boldsymbol{q}, \boldsymbol{p}]$, viz.

$$\mathbf{a} = W[\boldsymbol{p}] \cdot \Big(\boldsymbol{f}^{\mathrm{vac}}[\boldsymbol{q}, \boldsymbol{v}] + \boldsymbol{f}^{\mathrm{source}}[\boldsymbol{q}, \boldsymbol{v}; \mathbf{a}]\Big) \tag{18}$$

where

$$\boldsymbol{v} = \frac{1}{m_{\mathrm{b}}} \frac{\boldsymbol{p}}{\sqrt{1 + \frac{|\boldsymbol{p}|^2}{m_{\mathrm{b}}^2 c^2}}}; \quad m_{\mathrm{b}} \neq 0 \tag{19}$$

and

$$W[\boldsymbol{p}] := \mathrm{sgn}(m_{\mathrm{b}}) \frac{c}{\sqrt{m_{\mathrm{b}}^2 c^2 + |\boldsymbol{p}|^2}} \left[\boldsymbol{I}_{3\times 3} - \frac{\boldsymbol{p} \otimes \boldsymbol{p}}{m_{\mathrm{b}}^2 c^2 + |\boldsymbol{p}|^2}\right], \tag{20}$$

and where we have written the field as a sum of a source-free (vacuum) field and a field having the point charge as source, resulting in a **Lorentz** force due to that vacuum field,

$$\boldsymbol{f}^{\mathrm{vac}}[\boldsymbol{q}, \boldsymbol{v}](t) \equiv -e\left[\mathbf{E}^{\mathrm{vac}}(t, \boldsymbol{q}(t)) + \tfrac{1}{c}\boldsymbol{v}(t) \times \mathbf{B}^{\mathrm{vac}}(t, \boldsymbol{q}(t))\right], \tag{21}$$

and a "self"-type force $\boldsymbol{f}^{\text{source}}[\boldsymbol{q},\boldsymbol{v};\mathbf{a}]$, in ***BLTP*** *electrodynamics*[b] given by

$$\boldsymbol{f}^{\text{source}}[\boldsymbol{q},\boldsymbol{v};\mathbf{a}](t) = -\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^3}\boldsymbol{\Pi}^{\text{source}}(t,\boldsymbol{s})d^3s \tag{22}$$

$$= -\frac{\mathrm{d}}{\mathrm{d}t}\int_{B_{ct}(\boldsymbol{q}_0)}\left(\boldsymbol{\Pi}^{\text{source}}(t,\boldsymbol{s}) - \boldsymbol{\Pi}^{\text{source}}(0,\boldsymbol{s}-\boldsymbol{q}_0-\boldsymbol{v}_0t)\right)d^3s \tag{23}$$

$$= \tfrac{e^2}{4\pi}\Big[ - \mathbf{Z}^{[2]}_{\boldsymbol{\xi}}(t,t) + \mathbf{Z}^{[2]}_{\boldsymbol{\xi}^\circ}(t,t) \tag{24}$$

$$- \sum_{0\leqslant k\leqslant 1} c^{2-k}(2-k)\int_0^t\Big[\mathbf{Z}^{[k]}_{\boldsymbol{\xi}}\big(t,t^{\mathrm{r}}\big) - \mathbf{Z}^{[k]}_{\boldsymbol{\xi}^\circ}\big(t,t^{\mathrm{r}}\big)\Big](t-t^{\mathrm{r}})^{1-k}\mathrm{d}t^{\mathrm{r}} \tag{25}$$

$$- \sum_{0\leqslant k\leqslant 2} c^{2-k}\int_0^t\Big[\tfrac{\partial}{\partial t}\mathbf{Z}^{[k]}_{\boldsymbol{\xi}}\big(t,t^{\mathrm{r}}\big) - \tfrac{\partial}{\partial t}\mathbf{Z}^{[k]}_{\boldsymbol{\xi}^\circ}\big(t,t^{\mathrm{r}}\big)\Big](t-t^{\mathrm{r}})^{2-k}\mathrm{d}t^{\mathrm{r}}\Big] \tag{26}$$

where $\boldsymbol{\xi}(t) \equiv (\boldsymbol{q},\boldsymbol{v},\mathbf{a})(t)$, and $\boldsymbol{\xi}^\circ(t) \equiv (\boldsymbol{q}_0+\boldsymbol{v}_0t, \boldsymbol{v}_0, \mathbf{0})$, and where

$$\mathbf{Z}^{[k]}_{\boldsymbol{\xi}}\big(t,t^{\mathrm{r}}\big) = \int_0^{2\pi}\int_0^{\pi}\big(1-\tfrac{1}{c}\big|\boldsymbol{v}(t^{\mathrm{r}})\big|\cos\vartheta\big)\,\boldsymbol{\pi}^{[k]}_{\boldsymbol{\xi}}\big(t,\boldsymbol{q}(t^{\mathrm{r}})+c(t-t^{\mathrm{r}})\boldsymbol{n}\big)\sin\vartheta\mathrm{d}\vartheta\mathrm{d}\varphi\,, \tag{27}$$

with $\boldsymbol{n} = (\sin\vartheta\cos\varphi\ ,\ \sin\vartheta\sin\varphi\ ,\ \cos\vartheta)$, and where, with the kernels

$$\mathrm{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s}) := \tfrac{J_1\left(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\right)}{\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}}, \tag{28}$$

$$\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s}) := \tfrac{J_2\left(\varkappa\sqrt{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\right)}{c^2(t-t')^2-|\boldsymbol{s}-\boldsymbol{q}(t')|^2}\left(\boldsymbol{s}-\boldsymbol{q}(t')-\boldsymbol{v}(t')(t-t')\right), \tag{29}$$

we have

$$\boldsymbol{\pi}^{[0]}_{\boldsymbol{\xi}}(t,\boldsymbol{s}) = -\varkappa^4\frac{1}{4}\left[\frac{\left(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}\right)\times\left(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{1}{c}\boldsymbol{v}\right)}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^2}\right]_{\text{ret}} \tag{30}$$

$$+\varkappa^4\frac{1}{2}\left[\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\text{ret}}\times\int_{-\infty}^{t^{\text{ret}}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \tag{31}$$

$$-\varkappa^4\frac{1}{2}\left[\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{1}{c}\boldsymbol{v}}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\text{ret}}\times\int_{-\infty}^{t^{\text{ret}}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}c\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \tag{32}$$

$$-\varkappa^4\int_{-\infty}^{t^{\text{ret}}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}c\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t'\times\int_{-\infty}^{t^{\text{ret}}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \tag{33}$$

$$-\varkappa^4c\int_{-\infty}^{t^{\text{ret}}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\mathrm{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t'\int_{-\infty}^{t^{\text{ret}}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\mathrm{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\boldsymbol{v}(t')\mathrm{d}t', \tag{34}$$

[b]None of the four original contributors formulated a well-defined expression for the force, yet we believe that our formulation accomplishes what they had intended; hence the name of the theory.

$$\boldsymbol{\pi}^{[1]}_{\boldsymbol{\xi}}(t,\boldsymbol{s}) = -\varkappa^2\frac{1}{2}\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\left(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}\right)\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]_{\rm ret} \tag{35}$$

$$-\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\left(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}\right)\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]_{\rm ret}\times\int_{-\infty}^{t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \tag{36}$$

$$+\varkappa^2\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\left[\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\left(\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}\right)\times\boldsymbol{a}}{c^2\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]\right]_{\rm ret}\times\int_{-\infty}^{t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}c\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \tag{37}$$

$$-\varkappa^3\left[\frac{1}{1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})}\right]_{\rm ret}\int_{-\infty}^{t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\mathrm{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\left[\boldsymbol{v}(t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s}))+\boldsymbol{v}(t')\right]\mathrm{d}t', \tag{38}$$

$$\boldsymbol{\pi}^{[2]}_{\boldsymbol{\xi}}(t,s) = -\varkappa^2\left[\frac{1}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^2}\frac{1}{c}\boldsymbol{v}\right]_{\rm ret} \tag{39}$$

$$+\varkappa^2\left[\left[1-\tfrac{1}{c^2}\left|\boldsymbol{v}\right|^2\right]\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\times\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]_{\rm ret}\times\int_{-\infty}^{t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}c\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t' \tag{40}$$

$$-\varkappa^2\left[\left[1-\tfrac{1}{c^2}\left|\boldsymbol{v}\right|^2\right]\frac{\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})-\frac{1}{c}\boldsymbol{v}}{\left(1-\frac{1}{c}\boldsymbol{v}\cdot\boldsymbol{n}(\boldsymbol{q},\boldsymbol{s})\right)^3}\right]_{\rm ret}\times\int_{-\infty}^{t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s})}\boldsymbol{v}(t')\times\mathbf{K}_{\boldsymbol{\xi}}(t',t,\boldsymbol{s})\mathrm{d}t', \tag{41}$$

and $\big|_{\rm ret}$ means that $\boldsymbol{q}(\tilde{t})$, $\boldsymbol{v}(\tilde{t})$, $\boldsymbol{a}(\tilde{t})$ are evaluated at $\tilde{t}=t^{\rm ret}_{\boldsymbol{\xi}}(t,\boldsymbol{s})$.

**Remark 2.1.** The decomposition of the electromagnetic fields into a sum of two types of fields, one with the point charge as source, the other source-free, is to some extent arbitrary. For this reason it is futile to try to identify *the self-field force* of a point charge. At best one can speak of a "self"-field force, the scare quotes referring to the ambiguity in identifying how much of the field is "self"-generated by the charge and how much is not.

The following *key proposition* about the Volterra equation is proved in Ref. 17. Its proof takes several dozen pages of careful estimates.

**Proposition 2.2.** *Given $C^{0,1}$ maps $t\mapsto\boldsymbol{q}(t)$ and $t\mapsto\boldsymbol{p}(t)$, with Lip$(\boldsymbol{q})=v$, Lip$(\boldsymbol{v})=$ a big enough, and $|\boldsymbol{v}(t)|\leqslant v<c$, the Volterra equation (18) as a fixed point map has a unique $C^0$ solution $t\mapsto\mathbf{a}(t)=\boldsymbol{\alpha}[\boldsymbol{q}(\,\cdot\,),\boldsymbol{p}(\,\cdot\,)](t)$ for $t\geqslant 0$. Moreover, the solution depends **Lipschitz** continuously on the maps $t\mapsto\boldsymbol{q}(t)$ and $t\mapsto\boldsymbol{p}(t)$ (treated as independent).*

The well-posedness result of the joint initial value problem for **MBLTP** fields and their point charge sources is a corollary of Proposition 2.2. Namely, now setting $\boldsymbol{a}(t):=\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v}(t)$ and $\boldsymbol{v}(t):=\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{q}(t)$, and recalling (7), the solution to the **Volterra** integral equation for the acceleration poses a **Newton**-type second-order initial value problem for the position of the point charge with a complicated yet **Lipschitz**-continuous force. Once the motion is computed, inserting the vector functions $\boldsymbol{q}(\cdot)$,

$\boldsymbol{v}(\cdot)$, and $\mathbf{a}(\cdot)$ of the solution into the functionals of the fields yields $\mathbf{B}, \mathbf{D}, \mathbf{E}, \dot{\mathbf{E}}$ at $(t, \boldsymbol{s})$ for $t > 0$, too. This is a Theorem in Ref. 17, summarized informally as follows.

**Theorem 2.3.** *As a consequence of Proposition 2.2, the joint initial value problem for **MBLTP** fields and their point charge source is equivalent to the fixed point equations*

$$\boldsymbol{q}(t) = \boldsymbol{q}(0) + \frac{1}{m_b}\int_0^t \frac{\boldsymbol{p}}{\sqrt{1+|\boldsymbol{p}|^2/m_b^2}}(\tilde{t})d\tilde{t} =: Q_t(\boldsymbol{q}(\cdot), \boldsymbol{p}(\cdot)) \tag{42}$$

$$\boldsymbol{p}(t) = \boldsymbol{p}(0) - \int_{\mathbb{R}^3} (\boldsymbol{\Pi}(t, \boldsymbol{s}) - \boldsymbol{\Pi}(0, \boldsymbol{s}))\, d^3s =: P_t(\boldsymbol{q}(\cdot), \boldsymbol{p}(\cdot)), \tag{43}$$

*where $Q_t$ and $P_t$ are **Lipschitz** maps. Thus, **BLTP** electrodynamics is locally well-posed.*

In fact, in Ref. 17 the **Cauchy** problem for the **MBLTP** field with $N$ point charges is treated. *Local well-posedness* is proved for *admissible initial data* (see below) & $m_\mathrm{b} \neq 0$, and *global well-posedness* shown to hold if in finite time: (a) no particle reaches the speed of light; (b) no particle is infinitely accelerated; (c) no two particles reach the same location.
By "admissible" initial data we mean the following: the initial particle velocities are subluminal ($|\boldsymbol{v}(0)| < c$) and no two particles occupy the same location; the electromagnetic initial fields are the sum of a spatially sufficiently rapidly decaying vacuum field plus $N$ fields each with a single point charge source — the sourced fields are boosted electrostatic fields with boost velocity equal to the initial velocity of the source.

Although our result seems to be the first formulation of a well-posed joint initial value problem for classical electromagnetic fields and their point charge sources, and this endows **BLTP** electrodynamics with a mathematically superior status compared to the ill-defined **Lorentz** electrodynamics, we do not claim that **BLTP** electrodynamics is the correct classical limit of the illusive mathematically well-formulated quantum theory of electromagnetism. In particular, the **MBLTP** field equations feature "physical" oddities: (a) a field energy functional which is unbounded below; (b) subluminal transversal electromagnetic wave modes; (c) longitudinal electrical wave modes.
Moreover, since the **MBLTP** field equations demand initial data for $\mathbf{B}, \mathbf{D}, \mathbf{E}, \dot{\mathbf{E}}$ at $t = 0$ while physically we only can prescribe $\mathbf{B}$, $\mathbf{D}$ (N.B.: in **Lorentz** electrodynamics, $\mathbf{D} = \mathbf{E}$; not so in **BLTP** electrodynamics), one needs to find a prescription which expresses $\mathbf{E}, \dot{\mathbf{E}}$ at $t = 0$ in terms of $\mathbf{B}, \mathbf{D}$ at $t = 0$. In Ref. 17 we show that a reasonable choice is the map $(\mathbf{B}, \mathbf{D})_0 \mapsto (\mathbf{E}, \dot{\mathbf{E}})_0$ obtained by *maximizing the field energy* w.r.t. $\mathbf{E}_0$ and $\dot{\mathbf{E}}_0$ (treated independently). This can be made co-variant by stipulating that the maximization is carried out in the **Lorentz** frame in which total momentum vanishes initially.

Back to the mathematically superior status of **BLTP** electrodynamics, one can now apply rigorous analysis, and controlled numerical techniques, to study

the theory. For instance, a rigorous comparison of our expression (22) for the "self"-force with a differently defined "self"-field force which was studied in Ref. 12 was carried out recently by Hoang & Radosz and their students, see Ref. 5, 13. One of our next projects is to rigorously extract effective equations of motion with more user-friendly expressions for the "self"-force. In particular, whether or to which extent the **Landau–Lifshitz** equation approximately governs the motion is an interesting question.

### 2.1.2. *Topologically non-trivial flat spacetimes with time-like singularities*

The *zero-gravity limit* of singular spacetimes does not automatically yield **Minkowski** spacetime minus a number of world-lines. In Ref. 27 the zero-$G$ limit of **Carter**'s maximal analytical extension of the electromagnetic **Kerr–Newmann** spacetime was analyzed. The limiting spacetime is axially symmetric and static. It is locally isometric to **Minkowski** spacetime, but is topologically non-trivial, featuring **Zipoy** topology. Its constant-time slices are double-sheeted, and have the topology of $\mathbb{R}^3$ branched over the un-knot. The spacetime is singular on a 2-dimensional time-like cylinder $\mathbb{S}^1 \times \mathbb{R}$, which is the world-tube of a space-like ring-type singularity. The most intriguing aspect of this solution is that at any instant of time the ring singularity, when viewed from one of the two sheets of space, appears to be positively charged, and from the other sheet, negatively charged, as already noticed by **Carter**. The electromagnetic fields it supports where discovered in the 19th century by **P. Appell** as "multi-valued electromagnetic fields," while **A. Sommerfeld** soon realized that they represent single-valued electromagnetic fields on a topologically non-trivial multi-sheeted space.

A natural next step would be to formulate the corresponding zero-gravity two-body problem for two space-like ring-type singularities of z$G$**KN**-type, evolving in time jointly with the electromagnetic fields they generate. We know for example that the space should be four-sheeted (in general, $2^N$ sheets are needed for $N$ ring-type singularities). However, as the z$G$**KN** fields solve the **Maxwell–Lorentz** vacuum field equations away from the singularities, it is not surprising that one again encounters the infinite field energy-momentum problems which plague **Lorentz** electrodynamics. Thus one would first need to find either **MBI** or **MBLTP** analogues of the **Appell–Sommerfeld** fields. The nonlinearity of the **MBI** field equations makes this a daunting task, but we are confident that the feat can be accomplished with the **MBLTP** field equations. In that case the zero-$G$ **EIH**-type approach explained in the previous subsection should allow the formulation of a well-posed initial value problem for ring-type singularities and their electromagnetic **MBLTP** fields. Note that the law of motion would most likely be a system of partial differential equations, as the ring-type singularity has to be allowed to bend, twist, warp, stretch, and such.

### 2.2. *Turning on gravity: $G > 0$*

#### 2.2.1. *The neighborhood of Minkowski spacetime*

By rigorously establishing well-posedness of **BLTP** electrodynamics with point charges as a consequence of postulating the conservation law (2) for the total momentum vector-density (3), with the expressions for the particles given be the usual special-relativistic ones and those for the fields determined by the field equations,Ref. 17 demonstrates that a key idea of the 1938 **EIH** paper is viable in the zero-$G$ limit when applied with a suitable set of electromagnetic field equations, and with non-zero bare rest mass assigned to the point charges. We now address the question whether this result extends continuously to a $G > 0$ neighborhood of special relativity, free of black holes. By a result of **Geroch**–**Traschen**[11] a no-Black-Holes spacetime with a one-dimensional time-like singularity cannot exist if the singularity has positive bare mass, different from the special-relativistic case, where the bare mass merely had to be non-zero.

One of the main ingredients of the **EIH** argument is of course, that the conservation law of energy-momentum is implied by the twice contracted $2^{nd}$ **Bianchi** identity in concert with **Einstein**'s general-relativistic field equations. However, for the spacetimes with time-like singularities envisaged by **EIH** (and **Weyl**) this is not automatically true. An important step, therefore, is to determine the **Lorentz** spacetimes with time-like singularities on which the twice contracted $2^{nd}$ **Bianchi** identity holds in a weak form.

Together with **A. Y. Burtscher** we have begun a systematic study of the favorable conditions. In Ref. 3 we study the simplest non-trivial case: static spherically symmetric spacetimes with a single time-like singularity. We obtain some necessary and, for certain electromagnetic vacuum laws, also sufficient conditions that the twice contracted $2^{nd}$ **Bianchi** identity holds in a weak form. Interestingly, in the naked singularity regime (no Black Hole!), the answer is *negative* for the electromagnetic **Reissner**–**Weyl**–**Nordström** spacetime, but *positive* for the **Hoffmann** spacetime[c] in which electromagnetic **MBI** fields are coupled with **Einstein**'s gravity; see Ref. 26 for a rigorous discussion. Our goal is to generalize our study, one step at a time, to dynamical spacetimes without symmetry.

The crucial question, then, is whether the weak twice contracted $2^{nd}$ **Bianchi** identity implies the law of the electromagnetic spacetime's time-like singularities with bare energy-momentum assigned to them. Interestingly, it seems that merely changing **Maxwell**'s into the **BLTP** or **MBI** vacuum law is by itself not sufficient to allow the formulation of a well-defined joint initial value problem for the massive point charges and the electromagnetic and gravitational fields they generate: the gravitational coupling also needs to be changed, for instance in the manner of

[c]It is curious that although **Hoffmann**[14] had worked out his spacetime solution by 1935, the fact that his spacetime with non-positive bare mass is less singular than the **RWN** spacetime did not in 1938 compell **EIH** to use the **MBI** instead of **Maxwell**–**Lorentz** field theory.

**Jordan, Brans–Dicke**, or $f(\mathbf{R})$ gravity, to obtain a well-defined joint **Cauchy** problem. Put differently, the gravitational coupling of spacetime structure with bare matter and the electromagnetic fields needs to be "mediated" by a certain type of scalar field which moderates the strength of the spacetime singularities enough so that the strategy explained in the previous subsection can be applied. This scalar field in the classical theory would play a role vaguely reminiscent of the role played by the scalar Higgs field in the quantum field-theoretical standard model of elementary particle physics.

2.2.2. *The neighborhood of zGKN-type spacetimes*

Everything stated in the previous subsection about the neighborhood of the **Minkowski** spacetime has an analogue problem for the gravitational neighborhood of topologically non-trivial flat spacetimes of (generalized) z*G***KN** type.

**Acknowledgments**

We gratefully acknowledge interesting discussions with: E. Amorim, A. Burtscher, H. Carley, D. Deckert, V. Hartenstein, V. Hoang, M. Kunze, V. Perlick, M. Radosz, J. Speck, H. Spohn. We thank the organizers, R. Ruffini and R. Jantzen, for inviting us to MG15.

**References**

1. Bopp, F., *Eine lineare Theorie des Elektrons*, Annalen Phys. **430**, 345–384 (1940).
2. Born, M., and Infeld, L., *Foundation of the new field theory*, Proc. Roy. Soc. London **A 144**, 425–451 (1934).
3. Burtscher, A., Kiessling, M.K.-H., and Tahvildar-Zadeh, A.S., *Weak second Bianchi identity for spacetimes with time-like singularities*, arXiv:1901.00813 (2019).
4. Deckert, D.-A., and Hartenstein, V., *On the initial value formulation of classical electrodynamics*, J. Phys. A: Math. Theor. **49**, 445202 (19pp.) (2016).
5. DeLeon, A., Baza, A., Harb, A., Hoang, V. and Radosz, M., *On the Self-force for Particles Coupled to Higher-order Electrodynamics and Scalar Fields*, these proceed.
6. Dirac, P. A. M., *Classical theory of radiating electrons*, Proc. Roy. Soc. A **167**, 148–169 (1938).
7. Einstein, A., and Grommer, J., *General Relativity and the Law of Motion*, Sitzungsber. Preuss. Akad., Jan. 6, pp.2-13 (1927).
8. Einstein, A., Infeld, L., and Hoffmann, B., *The gravitational equations and the problem of motion*, Annals Math. **39**, 65-100 (1938).
9. Einstein, A., and Infeld, L., *The gravitational equations and the problem of motion, II*, Annals Math. **41**, 455–464 (1940).

10. Einstein, A., and Infeld, L., *On the motion of particles in general relativity theory*, Canad. J. Phys. **1**, 209–241 (1949).
11. Geroch, R., and Traschen, J., *Strings and other distributional sources in general relativity*, Phys. Rev. D, **36**, 1017ff (1987).
12. Gratus, J., Perlick, V., and Tucker, R.W., *On the self-force in Bopp–Podolsky electrodynamics*, J. Phys. A: Math. Theor. **48**, 435401 (28pp.) (2015).
13. Hoang, V., and Radosz, M., *On the self-force in higher-order electrodynamics*, U. Texas San Antonio Preprint, 41pp., arXiv:1902.06386 (2019).
14. Hoffmann, B., *Gravitational and electromagnetic mass in the Born–Infeld electrodynamics*, Phys. Rev. **47**, 877–880 (1935).
15. Kiessling, M.K.-H., *On the quasi-linear elliptic PDE* $-\nabla\cdot(\nabla u/\sqrt{1-|\nabla u|^2}) = 4\pi\sum_k a_k\delta_{s_k}$ *in physics and geometry*, Commun. Math. Phys. **314**, 509–523 (2013); *Correction:* Commun. Math. Phys. **364**, 825–833 (2018).
16. Kiessling, M.K.-H., *On the force which acts on a point charge source of the classical electromagnetic field*, (in preparation; 2019).
17. Kiessling, M.K.-H., and Tahvildar-Zadeh, A. S., *BLTP electrodynamics as initial value problem* (in preparation; 2019).
18. Landau, L. D., and Lifshitz, E., *The Classical Theory of Fields*, Pergamon Press (First English edition) (1951).
19. Landé, A., *Finite Self-Energies in Radiation Theory. Part I*, Phys. Rev. **60**, 121–126 (1941).
20. Landé, A., and Thomas, L.H., *Finite Self-Energies in Radiation Theory. Part II*, Phys. Rev. **60**, 514–523 (1941).
21. Pasqualotto, F., *Nonlinear stability for the Maxwell–Born–Infeld system on a Schwarzschild background*, arXiv:1706.07764, 101pp. (2017).
22. Podolsky, B., *A generalized electrodynamics. Part I: Non-quantum*, Phys. Rev. **62**, 68–71 (1942).
23. Poisson, E., Pound, A., and Vega, I., *The motion of point particles in curved spacetime*, Living Rev. Rel. **14**,7(190) (2011).
24. Speck, J. R., *The nonlinear stability of the trivial solution to the Maxwell-Born-Infeld system*, J. Math. Phys. **53**, 083703 (2012)
25. Spohn, H., *Dynamics of charged particles and their radiation fields*, Cambridge UP (2004).
26. Tahvildar-Zadeh, A. S., *On the Static Spacetime of a single Point Charge*, Rev. Mod Phys. **23**, 309–346 (2011).
27. Tahvildar-Zadeh, A. S., *On a zero-gravity limit of Kerr–Newman spacetimes and their electromagnetic fields, Journal of Mathematical Physics.* **56** 042501 (2015).
28. Wallace, P. R., *Relativistic equations of motion in electromagnetic theory*, Am. J. Math. **63**, 729–749 (1941).

# The Einstein-Infeld-Hoffmann legacy in mathematical relativity II: Quantum laws of motion for singularities of spacetime

A. Shadi Tahvildar-Zadeh† and Michael K. H. Kiessling

*Dept. of Mathematics, Rutgers–New Brunswick*
*E-mail: †shadit@math.rutgers.edu, miki@math.rutgers.edu*

We report on recent developments towards a relativistic quantum mechanical theory of motion for a fixed, finite number of electrons, photons, and their anti-particles, as well as its possible generalizations to other particles and interactions.

## 1. Introduction

As reported in Part I, the goal of the classical part of our program is to formulate and analyze a generally covariant joint initial value problem for the motion of massive charged particles, together with the electromagnetic and gravitational fields that they generate. We believe that proving local well-posedness, i.e. existence and uniqueness of a solution to this initial value problem, and continuous dependence of the solution on initial data, is a first necessary step towards the formulation of a deeper quantum theory of motion, for which the above could serve as the classical limit.

In the second part of our report, we expand on what such a relativistic quantum theory could look like, and present the results of our preliminary efforts towards achieving our ultimate goal, which is twofold: (1) to provide an accurate account of empirical electromagnetism in terms of a special-relativistic $N$-body quantum theory of electrons, photons, and their anti-particles; and (2) to extend that effort to general relativity by taking gravitational effects into account.

At the same time, inspired by Einstein's[12,13] and Bell's[3] writings, we want to find out whether it is possible to formulate such a theory as a generalization of the non-relativistic theory of de Broglie[6] and Bohm[4] in which the quantum-mechanical wave function $\psi$ guides the actual motion of all the particles involved.

Our approach to a relativistic quantum theory is based on a number of working hypotheses. Before listing them, we summarize the assumptions that formed the basis of our approach to the *classical* theory of motion of charged particles, as was outlined in Part I of this report:

1. All forms of matter are of a discrete (particulate) nature.
2. Particles, whether feature-less point-particles or structured ones such as ring-particles, have well-defined world-lines or world-tubes that form the *singular boundary* of spacetime, in the sense that the spacetime metric and/or classical fields permeating the spacetime develop a singularity as a point on this boundary is approached. Moreover, these singularities must be weak enough to give rise to locally integrable field energy-momentum

densities, but not so weak as to be confined to characteristic hypersurfaces (which are completely determined by the field equations).
3. The metric of the spacetime satisfies the weak version of the twice-contracted second Bianchi identity everywhere, including in a neighborhood of any singularity of spacetime (of the type discussed in item 2 above), thus allowing the usual conservation laws of total (i.e. field+particle) energy, momentum, and angular momentum to hold in a weak sense.

In Part I, we showed how these principles may be used to derive a classical equation of motion for world-lines of matter particles (which are identified with time-like singularities of spacetime) based entirely on momentum balance laws.

Our approach to a quantum law of motion begins by replacing the third item above with another set of hypotheses[a]:

3. The quantum-mechanical wave function of a single particle is a section of a Clifford-algebra-valued spin bundle over configuration spacetime. Different species of particles are distinguished by their wave functions at a point in spacetime belonging to different *generalized ideals* (stable subspaces) of that Clifford algebra[b].
4. The wave function of a system of $N$ relativistic particles is defined on the $N$-body configuration spacetime, and takes its values in a tensor space generated by the tensor product of bases for the constituent 1-body wave function spaces.
5. The wave function of a single particle satisfies a relativistic wave equation that is equivariant with respect to the action of the Poincaré group on the bundle of frames for the tangent space at any point in spacetime.
6. As a consequence of satisfying that wave equation, the wave function of a single particle possesses a conserved, future-directed, timelike (or causal, if the particle is massless), and normalizable current, whose component in the normal direction to a given spacelike foliation of spacetime, when evaluated at a point on a leaf of that foliation gives the probability density of the particle trajectory crossing the leaf at that point (by the Born rule.) The same holds for the $N$-particle version of the wave function and an $N$-particle tensor current that is jointly conserved.
7. The motion of a system of $N$ particles is guided by a multi-time wave function defined on the $N$-particle configuration spacetime, in accordance with a relativistic generalization of the de Brogile-Bohm guiding law, such as the Hypersurface Bohm-Dirac Law, in which a conserved $N$-particle tensor current evaluated at the actual positions of those $N$-particles determines their actual velocities.

[a] We emphasize that this is a preliminary list based on what we know so far, and is subject to modification and expansion as our investigations continue.
[b] See Furey[14,15] for details on this point of view.

8. In the classical limit, the quantum motions converge to those described in part I.

## 2. Least Invasive Quantization

In this section we demonstrate by way of a very simple example, how it may be possible to continuously deform a classical law into a quantum law of motion for the same particle.

Consider a particle of (bare) mass $m$ in an external potential $V = V(t, q)$, $q \in \mathbb{R}^d$. The relativistic Lagrangian for the particle's motion is

$$\ell(t, q, \dot{q}) = m(1 - \sqrt{1 - \dot{q}^2}) - V(t, q). \tag{1}$$

The corresponding action is

$$\mathcal{A}[q] := \int_{t_1}^{t_2} \ell(t, q, \dot{q}) dt. \tag{2}$$

We define the canonical momentum to be

$$p := \frac{\partial \ell}{\partial \dot{q}} = \frac{m\dot{q}}{\sqrt{1 - \dot{q}^2}}, \tag{3}$$

so the Hamiltonian is

$$H(t, q, p) = p\dot{q} - \ell = \sqrt{m^2 + p^2} - m + V(t, q). \tag{4}$$

The Hamilton equations are therefore

$$\dot{q} = \frac{\partial H}{\partial p} = \frac{p}{\sqrt{m^2 + p^2}}, \qquad \dot{p} = -\frac{\partial H}{\partial q} = -\partial_q V(t, q). \tag{5}$$

The *Hamilton-Jacobi equation*

$$\frac{\partial S}{\partial t} + H(q, \frac{\partial S}{\partial q}, t) = 0, \tag{6}$$

is a first-order nonlinear PDE for the *Hamilton-Jacobi phase function* $S = S(t, q)$. It arises in the context of solving Hamilton's equations by finding a *canonical transformation*, i.e. a change of variable in phase space $(q, p)$ such that the transformed Hamiltonian is constant. We note that $S$ is defined on the space of generic positions of the particle, i.e. it's a function defined on the configuration space (or configuration spacetime, in the relativistic setting.)

In the case of our particular example, it's easy to see that the Hamilton-Jacobi equation has the form of an eikonal equation:

$$|\partial_t \tilde{S} + V|^2 - |\partial_q \tilde{S}|^2 = m^2, \qquad \tilde{S} := S - mt. \tag{7}$$

Moreover, this equation is derivable from an action principle: Let $\varrho \geqslant 0$ be another function defined on the configuration spacetime, set

$$\mathcal{L} := \varrho \left( |\partial_t \tilde{S} + V|^2 - |\partial_q \tilde{S}|^2 - m^2 \right), \tag{8}$$

and let $\mathcal{S} := \iint \mathcal{L} dqdt$ be the corresponding action. It is then obvious that the stationary points of $\mathcal{S}$ with respect to compactly supported variations in $\varrho$ satisfy (7). If we now let $\varrho$ be an active variable, and set the variation of $\mathcal{S}$ with respect to $\tilde{S}$ also equal to zero, we obtain the following continuity-type equation for $\varrho$:

$$\partial^\mu(\varrho u_\mu) = 0. \tag{9}$$

Here $\mu = 0, \ldots, d$, indices are raised and lowered using the Minkowski metric $\eta = \mathrm{diag}(1, -1, \ldots, -1)$ on the configuration spacetime, and we are employing the Einstein summation convention. The covectorfield $u$ is by definition

$$u_0 := \partial_t \tilde{S} + V, \qquad u_i = \partial_{q^i} \tilde{S}, \tag{10}$$

so that by (7),

$$u_\mu u^\mu = m^2. \tag{11}$$

Let us now define a complex-valued field on the configuration spacetime

$$\psi := \sqrt{\varrho} e^{i\tilde{S}/\hbar}. \tag{12}$$

Evidently, using that $(\varrho, \tilde{S})$ satisfy the Euler-Lagrange equations for the stationary points of the action $\mathcal{S}$, one can derive the equation that $\psi$ must satisfy, which will be a nonlinear PDE.

The goal of a *least invasive quantization* procedure is to find a continuous deformation of the Lagrangian $\mathcal{L}$, now thought of as $\mathcal{L}[\psi]$, to a *nearby* Lagrangian $\tilde{\mathcal{L}}[\psi] = \mathcal{L}[\psi] + \hbar^2 \mathcal{L}'[\psi]$ in such a way that the Euler-Lagrange equation of the new action corresponding to $\tilde{\mathcal{L}}$ is a *linear* PDE in $\psi$. This is easily accomplished by adding a *Fisher-like* term to $\mathcal{L}$: Let

$$\tilde{\mathcal{L}} := \mathcal{L} + \hbar^2 \frac{\partial_\mu \varrho \partial^\mu \varrho}{4\varrho}. \tag{13}$$

A computation then shows that

$$\tilde{\mathcal{L}}[\psi] = \hbar^2 |(\partial_t + \frac{i}{\hbar} V)\psi|^2 - \hbar^2 |\partial_q \psi|^2 + m^2 |\psi|^2 \tag{14}$$

so that a stationary point $\psi$ will satisfy the linear Klein-Gordon equation

$$\left(D^\mu D_\mu + m^2\right) \psi = 0, \qquad D_\mu := -i\hbar\partial_\mu + A_\mu, \tag{15}$$

which (under appropriate assumptions relating $A_\mu$ and $V$), is a Lorentz-covariant equation for a massive spin-0 particle.

It should be noted that in the above admittedly over-simplified setting, $V$ was taken to be an "external" field, but in order to connect with our program of finding equations of motions for singularities of a dynamic field, $V$ needs to be determined by the field for which the particle in question is a singularity. For example, if we are to identify the electron with certain singularities of the electromagnetic field, then the corresponding $V$ in the wave equation for the electron would have to be determined in terms of that electromagnetic field, evaluated at generic positions of

the electron. For a detailed discussion of a proper least invasive quantization in the electromagnetic context see Ref. 21. In any case, since the electron is a spin-1/2 particle, the appropriate wave equation is not Klein-Gordon's but Dirac's, so the above process needs to be suitably modified.

Finally, we should note that even though in (9) one may formally set $\rho := \varrho u^0$ and $v^k := u^k/u^0$ to obtain what looks like a continuity equation for a probability density

$$\partial_t \rho + \partial_k(\rho v^k) = 0 \tag{16}$$

and proceed to derive a Born Rule and guiding law for the particle from it, there is a-priori no reason why $\rho$ would be non-negative. This suggests that even for spin-0 particles, the Klein-Gordon may not be the correct equation. Indeed, the lack of a future-directed timelike conserved current for solutions of the Klein-Gordon equation is a main reason for researchers doubting whether bosons can be treated as particles. We believe they can, and have already addressed this issue for photons[19]. The question of what the correct wave equation is for spin-0 bosons will be addressed in a forthcoming publication of ours.

## 3. Quantum Law of Motion for Point Test Particles

### 3.1. *Point particle in Hoffmann-like spacetimes*

We now move on to a more realistic scenario, of a single point charge, thought of as the only time-like singularity of a static, spherically symmetric 4-dimensional electrovacuum spacetime, i.e. a solution of the Einstein-Maxwell equations. As was shown in Ref. 7, the case where the electromagnetic vacuum law is the standard, linear law of Maxwell ($D = E, H = B$) yields a spacetime that is too singular for weak Bianchi identity to be satisfied at the singularity, so that our third working hypothesis is not satisfied. We showed however, that there are *nonlinear* vacuum laws, including the one proposed by Born & Infeld[5], for which the corresponding static spherically symmetric solution of the Einstein-Maxwell equations (which was fully analyzed in Ref. 23) *does* satisfy the weak Bianchi identity. The spacetime that corresponds to the Born-Infeld vacuum law was discovered by Hoffmann[16] in 1935.

As a first step towards a full-blown two-body quantum-mechanical problem, one may study the dynamics of a single electron, thought of as a *test* point charge, placed in the electrostatic and gravitational field of another, much more massive point charge such as a nucleus, the latter thought of as the central singularity of a static spherically symmetric spacetime, while the contribution of the former to the geometry of the spacetime is ignored. At the same time, the Born-Oppenheimer approximation allows one to reduce the 2-body quantum mechanical situation to a single body one. The dynamics of the electron is thus determined by its wave function that, due to the electron being a massive spin-1/2 particle, satisfies the

Dirac equation on the background spacetime of the nucleus, with the electrostatic field of the nucleus appearing in the Dirac operator as a minimal coupling term.

This program was carried out successfully in Ref. 1, where it was shown that the Dirac Hamiltonian on Hoffmann-like spacetimes is essentially self-adjoint regardless of the nuclear charge, its essential spectrum is the same as that of the free Dirac operator on Minkowski space, its point spectrum is non-empty and consists of infinitely many eigenvalues in the gap, accumulating at the right endpoint. Numerical investigations of this problem are in the works, including comparison of the eigenvalues with those of the Dirac+Coulomb potential and the Dirac+Born electrostatic potential Hamiltonians on a Minkowski background.

### 3.2. *Point particle in zero-gravity Kerr-Newmann spacetime*

As mentioned above, the main obstacle on the path to a well-defined quantum theory that could encompass both gravity and electromagnetism is that when Maxwell-Lorentz electrodynamics is coupled to Einstein's gravity, the infinities inherent in ML cause the spacetime to have curvature singularities that are too strong for even a weak notion of energy-momentum conservation to hold. In our quest to find a remedy for the strong curvature singularities of well-known solutions of Einstein-Maxwell equations such as the Kerr-Newmann solution, another avenue that we have pursued is to work with the *zero-gravity limit* of such singular spacetimes. This is a (geometrically well-defined) limit when Newton's gravitational constant $G$, which appears as a parameter in these metrics, is set to zero. The limiting spacetime and the electromagnetic field on it were analyzed in Ref. 22. It is axially symmetric, static, and locally isometric to Minkowksi space, but is topologically non-trivial. These facts were already known to Carter[8], who discovered the maximal analytical extension of the Kerr-Newmann solution. It was also known that the constant time slices of this manifold are double-sheeted, and have the topology of $\mathbb{R}^3$ branched over the un-knot (See Fig. 1.)

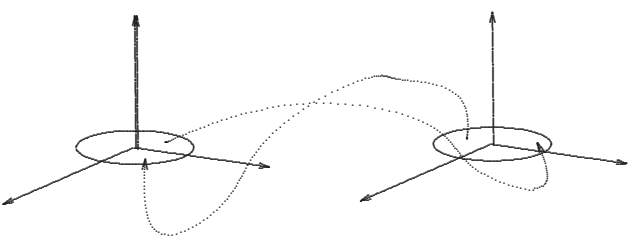


Fig. 1. Visualizing the double-sheeted spacelike slices of zero-$G$ Kerr-Newmann spacetime as a branched cover of the un-knot. Arrows indicate that the top of the disk in one copy of $\mathbb{R}^3$ is to be identified with the bottom of the disk in the other copy.

The spacetime is indeed singular on a 2-dim. timelike cylinder $\mathbb{S}^1 \times \mathbb{R}$, which is the world-tube of a ring-like singularity. The most intriguing aspect of this solution

is that at any instant of time the ring singularity, when viewed from one of the two sheets of space, appears to be positively charged, and from the other sheet, negatively charged.

As a first step in studying the quantum problem, one may identify the zero-gravity Kerr-Newmann (zGKN) solution with the spacetime outside a ring-like singularity representing a positively charged particle such as a proton. Since the metric of zGKN and the electromagnetic field corresponding to it are well known, it is easy to formulate the quantum dynamics of a test electron placed in the vicinity of the ring singularity by studying the Dirac equation on the zGKN background. This task was carried out in Ref. 17, where it was shown, using techniques developed in Refs. 2 & 26, that the pertinent Dirac Hamiltonian is essentially self-adjoint and its essential spectrum is the same as the standard Dirac operator on Minkowski space. We further showed that its discrete spectrum is nonempty, provided the radius of the ring is small enough. The ground state of the Dirac Hamiltonian for the single test electron has support in both sheets of the spacetime, but is mainly concentrated in the sheet where the ring singularity appears to be positively charged, see Fig. 2, where the horizontal axis is a radial coordinate that runs from $-\infty$ to 0 in one sheet and 0 to $\infty$ in the other sheet, while the vertical axis is $|\psi|^2$.

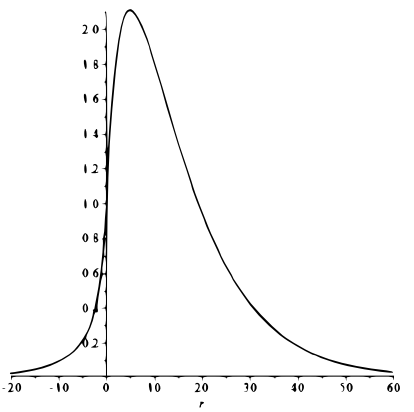



Fig. 2. Ground state of single-electron Dirac Hamiltonian on a zGKN background

Further analysis showed that the radius of the ring should be of the order of the *anomalous magnetic moment of the electron*, and that given this choice, the full spectrum will be close to the standard Dirac+Coulomb problem for Hydrogenic atoms. Furthermore, the broken symmetry due to nonzero ring radius causes the well-known degeneracy in the Dirac spectrum of Hydrogen to be broken, resulting in effects that are qualitatively similar to Lamb shift and hyperfine splitting, without the need to appeal to QED methods, see Fig. 3 (axes have the same meaning as in Fig. 2.)

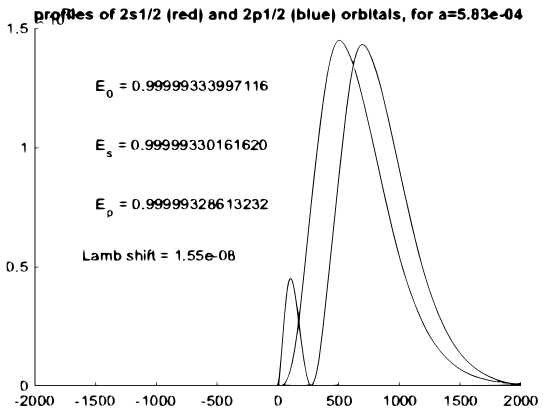



Fig. 3. Numerical evidence of a qualitative Lamb shift for Dirac's Hamiltonian on zGKN. The prameter $a$ is the ring radius, which here is set equal to electron's anomalous magnetic moment, measured in units of the Compton wavelength of electron.

## 4. Towards a Quantum Law of Motion for Ring-Like Singularities of Spacetime

Going back to the intriguing feature of the ring singularity of zGKN, i.e. that the same ring is found to have a positive charge when seen from one of the two sheets of that spacetime, and a negative charge when seen from the other sheet, one may wonder whether this ring singularity is representing both a particle and its anti-particle. In Ref. 18, extrapolating from the existence of such "bi-particle" structures in General Relativity, we proposed a novel interpretation of Dirac's "wave equation for the relativistic electron" in which the electron and the positron are merely the two different "topological spin" states of a single more fundamental particle, not distinct particles in their own right. This novel interpretation resolves the dilemma that Dirac's wave equation seems to be capable of describing both the electron and the positron in 'external' fields in many relevant situations, while the bi-spinorial wave function has only a single position variable in its argument, not two –as it should if it were a quantum-mechanical two-particle wave equation. We formulated a Dirac equation for such a ring-like bi-particle which interacts with a static point charge located elsewhere in the topologically non-trivial physical space associated with the moving ring particle. Indeed, for quasi-static motions of the ring, this equation is nothing but the one we have been discussing in the previous section, namely Dirac's equation on the zGKN background! The difference is in the interpretation, since the wave function in the present case is defined on the configuration space of the center of the ring, which is now moving with respect to a fixed point charge located on one of the two sheets. See Fig. 4 for a schematic illustration of the two interpretations. Furthermore, we showed that the motion of the ring can be governed by a de Broglie-Bohm type law extracted from the Dirac equation.

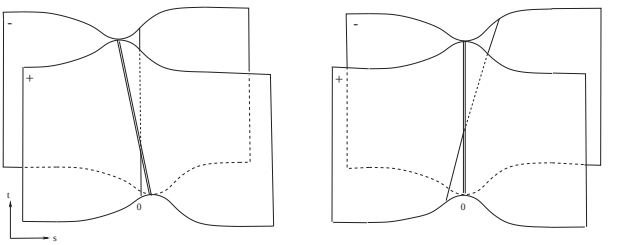



Fig. 4. Left: The physical spacetime with a single ring singularity in straight line motion relative to the rest frame of a designated origin, which in the depicted scenario gets "swept over" by the ring singularity; Right: Trajectory of a test particle in straight line motion in the static z$G$KN spacetime having the center of its ring singularity as the origin (marked 0). The test particle transits through the ring from one sheet to the other.

Our work in this direction is so far limited to quasi-static motions of the ring, which result in the ring remaining circular to the leading order, so that its motion can be specified solely in terms of the trajectory of the center of the circle, together with the evolving unit normal to the plane of the circle. Clearly, much work remains to be done in order to extend these results to the general case, which would presumably involve deformations of the ring as it moves.

The next natural step in this direction would be to study the full 2-body quantum problem, either in the context of nonlinear electrodynamics, or in the zero-gravity context by studying the two-rings analog of zGKN. Due to the strong nonlinearity of the BI vacuum law however, there is currently no explicit solution known to the *classical* 2-body problem, even in the absence of gravity, thus making any progress in that direction quite difficult. As to the two rings, that analysis is already under way. We know for example that the space should be four-sheeted (in general, $2^N$ sheets are needed for $N$ ring singularities), and that it is possible to come up with a Dirac equation for the 2-body wave function (with synchronized times), equipped with a minimal coupling term that models the electromagnetic interaction of the two rings. It is important to note however, that there is no chance of "turning gravity back on" in these models, since doing so will bring back the strong singularities of the type present in the Kerr-Newman spacetime, together with all the causal pathologies that they entail.

As described in Part I of this report, there are other electromagnetic theories, such as the Bopp-Lande-Thomas-Podolsky (BLTP) theory, which may allow us to find electromagnetic spacetimes whose singularities are mild enough so that they can represent charged particles. Once such solutions are known, a least invasive quantization procedure as outlined in Section 2 could conceivably lead to a quantum law of motion for those particles. Investigations are currently under way on various aspects of this program.

## 5. Quantum Laws of Motion for Interacting Point-Particle Systems

As mentioned at the outset, our goal is the eventual formulation of a relativistic quantum-mechanical theory of motion for a fixed number $N$ of electrons, photons, and their anti-particles in 3+1-dimensional Minkowski spacetime, and its generalization to other particles and interactions. We are certainly not the first researchers to state this as our goal, as can be attested to by the following passage in Ref. 9:

> "The Compton effect, at its discovery, was regarded as a simple collision of two bodies, and yet the detailed discussion at the present time involves the idea of the annihilation of one photon and the simultaneous creation of one among an infinity of other possible ones. We would like to be able to treat the effect as a two-body problem, with the scattered photon regarded as the same individual as the incident, in just the way we treat the collisions of electrons."

Such a goal however, of treating electrons and photons on an equal footing, within a quantum-*mechanical* framework of a fixed number of particles, has so far remained elusive. Progress has been obstructed in particular by the lack of a viable candidate for the quantum-mechanical photon wave function and its pertinent relativistic wave equation which furnishes a conserved probability current for the photon position, obeying Born's rule. Recently such a photon wave function and wave equation have been constructed in Ref. 19. In a forthcoming joint work with M. Lienert[20], we show how the photon wave equation of Ref. 19 can be coupled with Dirac's well-known relativistic wave equation for the electron in a Lorentz-covariant manner to accomplish what Darwin has asked us to do: to "treat the effect as a two-body problem, with the scattered photon regarded as the same individual as the incident, in just the way we treat the collisions of electrons."

We work with Dirac's manifestly Lorentz-covariant formalism of multi-time wave functions[10]. For our $N = 2$ body problem the wave function $\Psi(\mathbf{x}_{\rm ph}, \mathbf{x}_{\rm el})$ depends on the two generic spacetime events $\mathbf{x}_{\rm el}$ and $\mathbf{x}_{\rm ph}$ of the electron and the photon, respectively, which must be space-like separated. Both the Dirac operator of a free electron[25], and the Dirac-type operator of a free photon constructed in Ref. 19, act on $\Psi(\mathbf{x}_{\rm ph}, \mathbf{x}_{\rm el})$. Unique solvability of this system of evolution equations requires imposing a suitable *boundary condition* at the subset of co-incident events, $\{\mathbf{x}_{\rm el} = \mathbf{x}_{\rm ph}\}$. Conservation of the particle currents dictates the boundary condition, up to a choice in a phase. We prove that the resulting initial-boundary-value problem is well-posed.

It is intuitively obvious that a boundary condition at co-incident events $\{\mathbf{x}_{\rm el} = \mathbf{x}_{\rm ph}\}$ amounts to a local pair interaction between electron and photon in a Lorentz-covariant manner. Our boundary condition is compatible with the kind of interaction expected for an electron and a photon in Compton scattering. In particular, when we plot the joint 2-body probability density $\rho$ as a function of generic positions of the two particles, we observe that, as it evolves in time, the density forms four peaks, corresponding to the wave function being mostly supported near four distinct configurations: (1) both particles going to the left, (2) both particles going

to the right, (3) particles going away from each other for all time; and (4) particles initially moving towards each other and "bouncing off" of one another. Indeed the fourth peek appears to "hit" the boundary and get reflected, see Fig. 5 for snapshots of the evolving density[c].

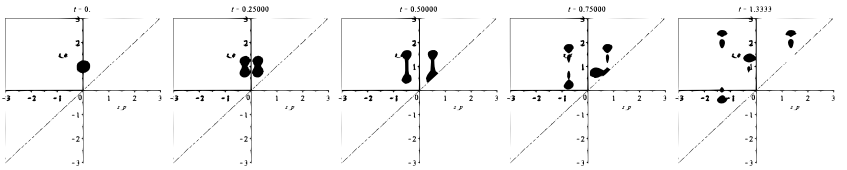


Fig. 5. Contour plot of the density $\rho$ in photon-electron configuration space at six consecutive snapshots of common time $t = t_{\mathrm{el}} = t_{\mathrm{ph}}$. The photon axis is horizontal, the electron axis vertical.

We have also carried out a limited number of numerical experiments with our system of equations, which indeed demonstrate the process of Compton scattering, but which also have revealed an unexpected novel phenomenon: photon capture and subsequent release by the electron. For all practical purposes this scenario seems indistinguishable from the scenario of an annihilation of a photon, followed by a subsequent emission of another one. In our quantum-mechanical $N = 2$-body model the photon of course never gets destroyed or created, precisely as envisioned by Darwin.

Beyond demonstrating that relativistic quantum mechanics with a fixed finite number $N$ of interacting particles is feasible, we also formulate the de Broglie–Bohm-type foundations of this Lorentz-covariant quantum model[d]. This is accomplished by adapting to our interacting 2-body model the so-called "hypersurface Bohm–Dirac"-type formulation for non-interacting particles[11]. This formulation requires one to specify a foliation of spacetime by spacelike hypersurfaces. In our model these hypersurfaces are given by a time-like Killing vector field that is determined self-consistently by the initial data of the wave function. The guiding law for the particles is furnished by the conserved current of our quantum-mechanical multi-time wave function. We extend a theorem of Teufel–Tumulka[24], which implies that unique particle motions typically exist globally in time.

The numerically computed trajectories confirm that our boundary condition causes the particles to scatter off of one another. See Fig. 6 for example of a two pairs of particles with identical initial conditions, one evolved according to the free evolution (i.e. without any boundary condition on the coincidence set) and the other using our boundary condition.

[c]You can see an animation of the density plot of $\rho$ as a function of the two positions by going to `http://sites.math.rutgers.edu/~shadi/phel1dint_slow.gif`

[d]This should put to rest the often-voiced claim that the de Broglie–Bohm theory could not be made relativistic.

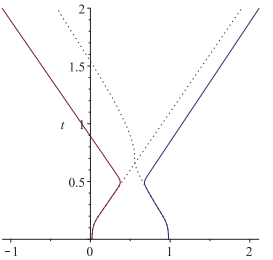



Fig. 6. Interacting (solid) versus non-interacting (dotted) electron-photon trajectories.

The intriguing phenomenon of photon "capture and release" by the electron that we alluded to in the above can also be illustrated by the numerically computed electron and photon trajectories, see Fig. 7.

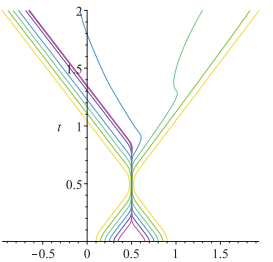



Fig. 7. World-lines corresponding to possible capture-and-release phenomena (Different colors correspond to the photon-electron pair starting at different initial positions that are successively closer to each other.)

Finally, by sampling a large ensemble of random positions distributed according to the initial wave function, we can also illustrate that the empirical statistics over the possible actual trajectories reproduces Born's rule in our model, a consequence of the equivariance of the evolution of the probability densities. See Fig. 8 and note the four clusters of trajectories that correspond to the four peaks of the joint probability density that was mentioned in the above.

More numerical investigation is currently under way, to explore various regions in the parameter space for the dynamics (e.g. mass and energy of incident electron, frequency profile of incident photon, their initial distance, etc.) and to quantify the energy and momentum transfer between the two particles.

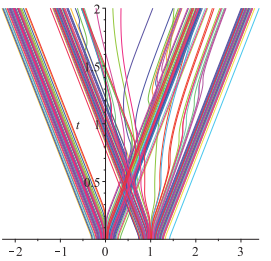



Fig. 8. Electron & photon world-lines corresponding to 100 pairs of typical initial positions.

**Acknowledgements**

We gratefully acknowledge helpful discussions with: M. K. Balasubramanian, E. Carlen, O. Darrigol, C. Furey, S. Goldstein, J.-M. Graf, F. Hehl, M. Jansen, H. Jauslin, J. Lebowitz, M. Lein, N. Leopold, M. Lienert, T. Newman, B. Simon, A. Soffer, H. Spohn, W. Struyve, R. Tumulka, and M. Winklmeier. We thank the organizers, R. Ruffini and R. Jantzen, for inviting us to MG15.

**References**

1. Balasubramanian, M. K. , "Scalar fields and spin-half fields on mildly singular spacetimes", Thesis (Ph.D.)–Rutgers, The State University of New Jersey, 77 pages, (2015).
2. D. Batic, H. Schmid, and M. Winklmeier. "On the eigenvalues of the Chandrasekhar-Page angular equation," *J. Math. Phys.*, **46**(1):012504, 35, (2005).
3. Bell, J.S., "Speakable and Unspeakable in Quantum Mechanics," $2^{nd}$ ed., Cambridge Univ. Press, Cambridge (2004).
4. Bohm, D., "A suggested interpretation of the quantum theory in terms of 'hidden' variables. Part I," *Phys. Rev.* **85**:166–179 (1952); "Part II," *ibid.*, 180–193 (1952).
5. Born, M., and Infeld, L., "Foundations of the New Field Theory", *Nature*, **132**, p. 1004, (1933).
6. de Broglie, L.V.P.R., *La nouvelle dynamique des quanta,* in "Cinquième Conseil de Physique Solvay" (Bruxelles 1927), ed. J. Bordet, (Gauthier-Villars, Paris, 1928); English transl.: "The new dynamics of quanta", p.374-406 in: G. Bacciagaluppi and A. Valentini, "Quantum Theory at the Crossroads," (Cambridge Univ. Press, 2009).
7. Burtscher, A. Y., Kiessling, M. K.-H., and Tahvildar-Zadeh, A. S., "Weak second Bianchi identity for spacetimes with timelike singularities," *preprint* (20 pages), `[arXiv:1901.00813]`, (2019).
8. Carter, B., "Global structure of the Kerr family of gravitational fields," *Phys. Rev.*, **174**:1559–1571, (1968).

9. Darwin, C. G., "Notes on the Theory of Radiation," *Proc. Roy. Soc. London A,* **136** (829), pp. 36-52, (1932).
10. Dirac, P. A. M., "Relativistic Quantum Mechanics," *Proc. R. Soc. Lond. A*, **136**:453–464 (1932).
11. Dürr, D., Goldstein, S., Münch-Berndl, K., and Zanghì, N., "Hypersurface Bohm–Dirac models" *Phys. Rev. A*, **60**: 2729–2736 (1999).
12. Einstein, A., "Zum gegenwärtigen Stand des Strahlungsproblems," *Phys. Zeitschr.* **10**, 185–193 (1909).
13. Einstein, A., "Über die Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung," *Verh. Deutsch. Phys. Ges.* **7**, 482–500 (1909).
14. Furey, C., "Charge quantization from a number operator," *Physics Letters B,* **742**, pp. 195-199, (2015).
15. Furey, C., "$SU(3)_C \times SU(2)_L \times U(1)_Y$ ($\times U(1)_X$) as a symmetry of division algebraic ladder operators", *The European Physical Journal C* **78**(5), p. 375, (2018).
16. Hoffmann, B., "On the new field theory," *Proc. R. Soc. Lond. A*, **148**, pp. 353–364, (1935)
17. Kiessling, M. K.-H. & Tahvildar-Zadeh, A. S.,"The Dirac point electron in zero-gravity Kerr–Newman spacetime," *Journal of Mathematical Physics* **56** 042303 (2015).
18. Kiessling, M. K.-H. & Tahvildar-Zadeh, A. S., "A novel quantum-mechanical interpretation of the Dirac equation," *Journal of Physics A* **49** 135301 (2016).
19. Kiessling, M. K.-H. & Tahvildar-Zadeh, A. S., "On the Quantum Mechanics of a Single Photon," *Jour. Math Phys.* 59, 112302, 34 pages, (2018).
20. Kiessling, M. K.-H., Lienert, M., & Tahvildar-Zadeh, A. S., "A Lorentz-Covariant Interacting Electron-Photon System in One Space Dimension," *preprint*, 34 pp., [ arXiv:1906.03632], (2019).
21. Kiessling, M. K.-H., "Electromagnetic field theory without divergence problems. II. A least invasively quantized theory", *J. Statist. Phys.*, **116**, 1-4, pp. 1123–1159, (2004).
22. Tahvildar-Zadeh, A. S., "On a zero-gravity limit of Kerr–Newman spacetimes and their electromagnetic fields," *Journal of Mathematical Physics.* **56** 042501 (2015).
23. Tahvildar-Zadeh, A. S., "On the static spacetime of a single point charge," *Rev. Math. Phys.*, **23** (3), pp. 309–346, (2011).
24. Teufel, S. and Tumulka, R., "Simple Proof for Global Existence of Bohmian Trajectories," *Commun. Math. Phys.*, **258**:349–365 (2005).
25. Thaller, B., *The Dirac Equation*, Springer Verlag, Berlin (1992).
26. Winklmeier, M., and Yamada, O., "Spectral analysis of radial Dirac operators in the Kerr-Newman metric and its applications to time-periodic solutions," *J. Math. Phys.*, **47**(10):102503, 17, (2006.)

# The problem of Boltzmann brains and how Bohmian mechanics helps solve it

Roderich Tumulka

*Mathematisches Institut, Eberhard-Karls-Universität,*
*Auf der Morgenstelle 10, 72076 Tübingen, Germany*
*E-mail: roderich.tumulka@uni-tuebingen.de*

Most versions of classical physics imply that if the 4-volume of the entire space-time is infinite or at least extremely large, then random fluctuations in the matter will by coincidence create copies of us in remote places, so called "Boltzmann brains." That is a problem because it leads to the wrong prediction that *we* should be Boltzmann brains. The question arises, how can any theory avoid making this wrong prediction? In quantum physics, it turns out that the discussion requires a formulation of quantum theory that is more precise than the orthodox interpretation. Using Bohmian mechanics for this purpose, we point out a possible solution to the problem based on the phenomenon of "freezing" of configurations.

*Keywords*: quantum fluctuation; Bunch-Davies vacuum; late universe; scalar fields in cosmology; de Sitter space-time

## 1. The problem of Boltzmann brains

A "Boltzmann brain" [1,2] is this: Let $M$ be the present macro-state of your brain. For a classical gas in thermal equilibrium, it has probability 1 that after sufficient waiting time, some atoms will "by coincidence" ("by fluctuation," "by ergodicity") come together in such a way as to form a subsystem in a micro-state belonging to $M$. That is, this brain comes into existence not by childhood and evolution of life forms, but by coincidence; this brain has memories (duplicates of your present memories), but they are false memories: the events described in the memories never happened to this brain! Boltzmann brains are, of course, very unlikely. They are fluctuations in which entropy spontaneously decreases. But they will happen if the waiting time is long enough, and they will happen more frequently if the system is larger (bigger volume, higher number of particles).

The problem with Boltzmann brains is this: If the universe continues to exist forever, and if it reaches universal thermal equilibrium at some point, then the overwhelming majority of brains in the history of the universe will be Boltzmann brains. According to the "Copernican principle," we should see what a typical observer sees. But we do not, as most Boltzmann brains find themselves surrounded by thermal equilibrium, not by other intelligent beings on a planet. In fact, if a theory predicts that the great majority of brains are Boltzmann brains, then one should conclude that we are Boltzmann brains. But we are not. We had childhoods. We know because we trust out memories more than any physics theory.

How can any of our serious theories avoid making this incorrect prediction? One possibility is that there will be a Big Crunch, and that the total 4-volume of the universe will be finite and not too large. Then the probability that an

unlikely fluctuation ever happens can be small. However, it is expected (e.g., from Λ-cold-dark-matter, often called the standard model of cosmology) that the late universe will be close to de Sitter space-time and thus has infinite lifetime and infinite 4-volume. We will describe another way out of the problem.[5,9]

## 2. Concrete version of the Boltzmann brain problem

It is further expected from Λ-cold-dark-matter that the state of matter in the late universe will be close (in terms of local observables) to the Bunch-Davies vacuum, a quantum state invariant under the isometries of de Sitter space-time (in particular, a stationary state). The $|\psi|^2$ probability distribution it defines on configuration space gives $> 99\%$ weight to thermal equilibrium configurations, but positive probability to brain configurations, in fact $> 99\%$ probability to configurations containing brains if 3-space is large enough (in particular if infinite).

Does this mean there are Boltzmann brains in the Bunch-Davies vacuum? Actually, does a stationary state mean that nothing happens? Or does the factual situation visit different configurations over time according to $|\psi|^2$? This is a point at which foundational questions of quantum theory become relevant to cosmology. We need to ask, what is the significance of this particular wave function for reality?

We will analyze a Bohm-type model of the situation and conclude that there are no Boltzmann brains in the Bunch-Davies state. For comparison, for Everett's many-worlds interpretation of quantum theory, Boddy, Carroll, and Pollack[2] have argued that in a stationary state nothing moves, whereas David Wallace, a leading advocate of the many-worlds view, disagrees (personal communication).

## 3. Bohmian mechanics

Bohmian mechanics[3,4] is a version of quantum mechanics with trajectories. For non-relativistic quantum mechanics, it can be defined as follows. The theory takes "particles" literally and proposes that electrons are material points moving in 3-space that have a definite position $\boldsymbol{Q}_k(t)$ at every time $t$. There is a wave-particle duality in the sense that there are particles and there is a wave (the usual wave function $\psi_t$ of quantum mechanics on the configuration space $\mathbb{R}^{3N}$ of $N$ particles).

The theory has two dynamical laws: for $\psi_t$ the usual Schrödinger equation

$$i\hbar \frac{\partial \psi_t}{\partial t} = H\psi\,, \tag{1}$$

and for $\boldsymbol{Q}_k(t)$ Bohm's equation of motion

$$\frac{d\boldsymbol{Q}_k(t)}{dt} = \frac{\hbar}{m_k} \nabla_k \operatorname{Im} \log \psi_t\big(\boldsymbol{Q}_1(t), \ldots, \boldsymbol{Q}_N(t)\big)\,. \tag{2}$$

Figure 1 shows a sample of Bohmian trajectories after passing through a double slit. The law of motion (2) is equivalent to

$$\frac{dQ}{dt} = \frac{j(Q)}{\rho(Q)}\,, \tag{3}$$

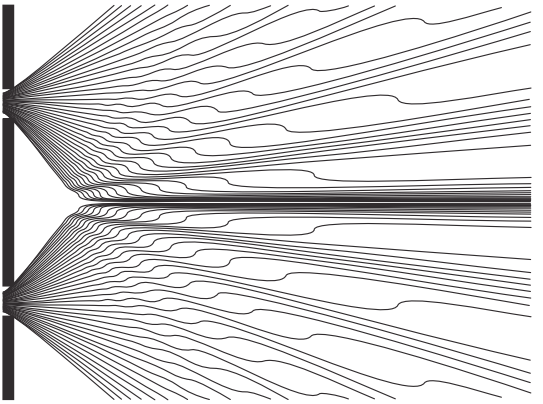


Fig. 1. Several alternative Bohmian trajectories in a double slit experiment, each coming from the left and passing through one of the slits. (Reprinted from Ref. 4)

where $Q = (\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_N)$ is the configuration, $\rho = |\psi|^2$ is the standard probability density, and $j$ is the standard probability current vector field in configuration space. As a consequence of the continuity equation for $\rho$ and $j$, one obtains the *equivariance theorem* that if $Q(t = 0)$ is random with distribution density $|\psi(t = 0)|^2$, then at any time $t$, $Q(t)$ has distribution $|\psi(t)|^2$. As a part of the theory, we assume that the initial configuration is in fact so chosen. Figure 1 illustrates that the arrival points on the screen are $|\psi|^2$ distributed. The central fact about the empirical predictions of Bohmian mechanics is that the inhabitants of a Bohmian world would observe outcomes in agreement with the predictions of quantum mechanics.

## 4. How Bohmian mechanics helps against Boltzmann brains

### 4.1. *Freezing in Bohmian mechanics*

In Bohmian mechanics, there is the phenomenon of "freezing," whose basic form is as follows.

**Theorem 4.1.** *If $\psi$ is a non-degenerate eigenstate of a Hamiltonian that is a non-relativistic Schrödinger operator, then the Bohmian configuration does not move.*

That is because the conjugate of $\psi$ must be another eigenstate with the same eigenvalue, so $\psi$ must be real up to a global phase. As a consequence, according to (2), particle velocities vanish. Freezing may be surprising because the momenum distribution, $|\hat{\psi}(k)|^2$ with $\hat{\psi}$ the Fourier transform of $\psi$, is not concentrated on the origin. In Bohmian mechanics, momentum corresponds not to the instantaneous velocity but to the asymptotic velocity that the particle would reach if the potential were turned off. [4]

It follows that, if non-relativistic Bohmian mechanics were true, and if the late universe were in a non-degenerate eigenstate, then the configuration would

be frozen. Arguably, the Boltzmann brain problem is absent then, as any such brain, even if it existed, would not be functioning. Turning now to a more realistic scenario, we will see that although freezing as such does not occur in the late universe, asymptotic freezing does.

### 4.2. *Concrete model of the late universe*

Let us take for granted that the late universe is asymptotically a de Sitter space-time with metric

$$ds^2 = dt^2 - e^{2Ht}\delta_{ij}dx^i dx^j\,, \tag{4}$$

where $H$ is the Hubble parameter (expansion speed of the universe—not the Hamiltonian!) and $i, j$ are summed over the values $1, 2, 3$. Actually, this coordinate patch covers only half of de Sitter space-time, but that is alright because we are considering only the late universe. As a model of the matter, we consider a simple quantum field theory with a single Hermitian scalar quantum field $\varphi(\mathbf{x}, t)$. The quantum state can then be written as a wave functional $\Psi(\varphi, t)$ on the space of field configurations $\varphi$, which are real-valued functions of $\mathbf{x} = (x^1, x^2, x^3)$. It will be convenient to rescale time according to $dt = e^{Ht}\,d\eta$, where the new coordinate $\eta$ is called the conformal time. As $t$ varies from $-\infty$ to $+\infty$, $\eta$ runs from $-\infty$ only up to 0; in particular, the long-time limit $t \to \infty$ that is relevant to us corresponds to $\eta \to 0$ from the left. We also rescale the field variable according to $y = e^{Ht}\varphi$. Moreover, it will be useful to express the field configuration in terms of its Fourier modes $y_{\mathbf{k}}$ (note that $y_{-\mathbf{k}} = y_{\mathbf{k}}^*$ since $y(\mathbf{x})$ is real). The Schrödinger equation for $\Psi$ in the Schrödinger picture then reads[5,7]

$$i\frac{\partial\Psi}{\partial\eta} = \int_{\mathbb{R}^{3+}} d^3\mathbf{k}\left[-\frac{\delta^2}{\delta y_{\mathbf{k}}^*\delta y_{\mathbf{k}}} + k^2\, y_{\mathbf{k}}^*\, y_{\mathbf{k}} + \frac{i}{\eta}\left(\frac{\delta}{\delta y_{\mathbf{k}}^*}y_{\mathbf{k}}^* + y_{\mathbf{k}}\frac{\delta}{\delta y_{\mathbf{k}}}\right)\right]\Psi\,, \tag{5}$$

where $k = |\mathbf{k}|$, and $\mathbb{R}^{3+}$ means an arbitrary choice of half space (containing exactly one of $\mathbf{k}$ and $-\mathbf{k}$ for almost every $\mathbf{k}$).

We consider a Bohmian model with a field ontology,[6,7] that is, we assume that there is an actual field configuration $\varphi(\mathbf{x}, t)$ (the analog of $Q(t)$), equivalently expressed as $y_{\mathbf{k}}(\eta)$, which is "guided" by $\Psi$ according to the equation

$$\frac{dy_{\mathbf{k}}}{d\eta} = \frac{\delta \mathrm{Im}\log\Psi}{\delta y_{\mathbf{k}}^*} - \frac{1}{\eta}y_{\mathbf{k}}\,, \tag{6}$$

which is the natural analog in this setting of Bohm's equation of motion (2).[7]

### 4.3. *Asymptotic freezing in the Bunch-Davies state*

In this representation, the Bunch-Davies state reads

$$\Psi = \prod_{\mathbf{k}\in\mathbb{R}^{3+}} \frac{1}{\sqrt{2\pi}f}\exp\left\{-\frac{1}{2f^2}|y_{\mathbf{k}}|^2 + i\left[\left(\frac{f'}{f} + \frac{1}{\eta}\right)|y_{\mathbf{k}}|^2 - \mathrm{phase}(k,\eta)\right]\right\} \tag{7}$$

with $f = f_k(\eta) = \sqrt{1+1/k^2\eta^2}/\sqrt{2k}$. In particular, the different field modes $\mathbf{k}$ are disentangled. The phase of $\Psi$ is not constant, so $y_\mathbf{k}$ is not frozen, but $\Psi$ is simple enough so we can explicitly solve[5] the Bohmian equation of motion (6):

$$y_\mathbf{k}(\eta) = \tilde{c}_\mathbf{k}\, f_k(\eta) \tag{8}$$

or, expressed in the original variables,

$$\varphi_\mathbf{k}(t) = c_\mathbf{k}\sqrt{1+k^2\exp(-2Ht)/H^2}\,. \tag{9}$$

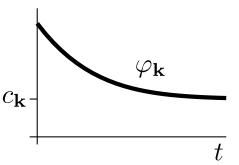



Fig. 2. Graph of $\varphi_\mathbf{k}$ as a function of $t$

Note that $\lim_{t\to\infty}\varphi_\mathbf{k}(t)$ exists and is equal to $c_\mathbf{k}$, which means the Bohmian field is asymptotically freezing. This suggests that the configuration becomes constant for large $t$, precluding a functioning brain. In fact, the situation is a little more subtle because the $x^j$ coordinates are moving apart, so that at any time only the modes with wave lengths large compared to the Hubble distance $1/H$ are frozen. But even before (9) becomes constant, its simple behavior is as good as freezing, as it is too simple to support the complex behavior of a functioning brain. Hence, in Bohmian mechanics there are no Boltzmann brains in the Bunch-Davies vacuum.

### 4.4. *Asymptotic freezing in a generic state*

However, $\Psi$ will not be close to the Bunch-Davies state in Hilbert space. It will look locally similar, but Bohmian mechanics depends nonlocally on the wave function. This issue is taken care of by the following

**Theorem 4.2 (Ryssens[8] 2012, Tumulka[9] 2015).** *For a large class of wave functions $\Psi$ and most initial field configurations, the asymptotic long-time behavior of $\varphi$ is*

$$\varphi_\mathbf{k}(t) \approx c_\mathbf{k}\sqrt{1+k^2\exp(-2Ht)/H^2} \quad \text{for } t > t_0\,, \tag{10}$$

*where $t_0$ is independent of $\mathbf{k}$ (but depends on $\Psi$). In particular, $\lim_{t\to\infty}\varphi_\mathbf{k}(t)$ exists.*

### 4.5. *Idea of proof*

Consider a single mode $\mathbf{k}$. Rescale the field variable, setting $z = \gamma(\eta)^{-1}y$, with the function $\gamma$ to be chosen later. Rescale and phase-transform wave function,

$$\Phi(z,\eta) = e^{\alpha(\eta)+i\beta(\eta)z^*z}\Psi(\gamma(\eta)\,z,\eta) \tag{11}$$

with functions $\alpha$ and $\beta$ to be chosen later. Also rescale time according to $d\tau = \gamma^{-2}\, d\eta$. For a suitable (explicitly known[9]) choice of the scaling functions $\alpha, \beta, \gamma$, the evolution of $\Phi$ reduces to a non-relativistic Schrödinger equation in a 2d harmonic oscillator potential,

$$i\frac{\partial \Phi}{\partial \tau} = -\frac{\partial^2 \Phi}{\partial z^* \partial z} + \omega^2 z^* z \Phi \tag{12}$$

and the non-relativistic Bohmian equation of motion

$$\frac{dz}{d\tau} = \frac{\partial \mathrm{Im} \log \Phi(z,\tau)}{\partial z^*}\,. \tag{13}$$

These evolutions do not become singular as $\tau \to 0-$ ($\Leftrightarrow$ $\eta \to 0-$ $\Leftrightarrow$ $t \to \infty$). Thus, $\lim_{\tau \to 0-} z(\tau)$ exists, so, for $\tau$ sufficiently close to 0, $z(\tau) \approx$ const. $\Leftrightarrow$ $y(\eta) \approx \gamma(\eta) \times$ const. $\Leftrightarrow$ $\varphi(t) \approx c_{\mathbf{k}} \sqrt{1 + k^2 \exp(-2Ht)/H^2}$.

### 4.6. *Upshot*

For almost any $\Psi$, the Bohmian motion becomes very simple as $t \to \infty$. While there is a positive probability for a brain configuration to occur, this subsystem would not function as a brain because it cannot move in a complex way. (In fact, the probability of a brain configuration in a Hubble volume is tiny.) Thus, Boltzmann brains do not occur in Bohmian mechanics according to this particular model.

## References


1. A. Albrecht, in J.D. Barrow, P.C.W. Davies, C.L. Harper (editors), *Science and Ultimate Reality* (Cambridge University Press 2004), `http://arxiv.org/abs/astro-ph/0210527`
2. K.K. Boddy, S.M. Carroll, and J. Pollack, in K. Chamcham, J. Barrow, J. Silk, and S. Saunders (editors), *The Philosophy of Cosmology* (Cambridge University Press 2016), `http://arxiv.org/abs/1505.02780`
3. D. Bohm, *Physical Review* **85**, 166 (1952)
4. D. Dürr and S. Teufel, *Bohmian mechanics* (Springer, Heidelberg, 2009)
5. S. Goldstein, W. Struyve, and R. Tumulka, to appear in A. Ijjas and B. Loewer (editors), *Guide to the Philosophy of Cosmology* (Oxford University Press, 2019), `http://arxiv.org/abs/1508.01017`
6. B.J. Hiley and A.H. Aziz Mufti, p. 141 in M. Ferrero and A. van der Merwe (editors), *Fundamental Theories of Physics* **73** (Kluwer, Dordrecht, 1995)
7. N. Pinto-Neto, G. Santos, and W. Struyve, *Physical Review D* **85**, 083506 (2012), `http://arxiv.org/abs/1110.1339`
8. W. Ryssens, *On the Quantum-to-Classical Transition of Primordial Perturbations*, Master thesis, Physics & Astron. Dept., Katholieke Univ. Leuven (2012)
9. R. Tumulka, *General Relativity and Gravitation* **48**, 2 (2016), `http://arxiv.org/abs/1507.08542`

# Traversable wormholes and Yukawa potentials

Remo Garattini

*Università degli Studi di Bergamo,*
*Dipartimento di Ingegneria e scienze applicate,*
*Viale Marconi,5 24044 Dalmine (Bergamo) ITALY*
*I.N.F.N. - sezione di Milano, Milan, Italy*
*E-mail:remo.garattini@unibg.it

Traversable Wormhole are amazing astrophysical objects predicted by General Relativity which are able to connect remote region of space-time. Even if their existence has not been proved yet they are object of continuous investigation. From the theoretically point of view, to exist, traversable wormholes need a special form of energy density termed "*exotic*". Since this exotic source must be concentrated on the throat of the wormhole, we discuss the implications of assuming Yukawa-like profiles which could be realize such a configuration.

## 1. Introduction

Traversable wormholes are solutions of the Einstein's Field Equations (EFE) which have the property of connecting remote space-time regions[1,2]. In Schwarzschild coordinates, the traversable wormhole metric can be cast into the form

$$ds^2 = -\exp\left(-2\phi\left(r\right)\right)dt^2 + \frac{dr^2}{1-\frac{b(r)}{r}} + r^2 d\Omega^2. \tag{1}$$

where $\phi\left(r\right)$ is called the redshift function, while $b\left(r\right)$ is called the shape function and where $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$ is the line element of the unit sphere. $\phi(r)$ and $b(r)$ are arbitrary functions of the radial coordinate $r \in [r_0, +\infty)$. A fundamental property of a wormhole is that a flaring out condition of the throat, given by $(b - b'r)/b^2 > 0$, must be satisfied as well as the request that $1 - b(r)/r > 0$. Furthermore, at the throat $b(r_0) = r_0$ and the condition $b'(r_0) < 1$ is imposed to have wormhole solutions. It is also fundamental that there are no horizons present, which are identified as the surfaces with $e^{2\phi} \to 0$, so that $\phi(r)$ must be finite everywhere. The first step to establish if a traversable wormhole exists is given by solving the EFE in an orthonormal reference frame

$$\rho\left(r\right) = \frac{1}{8\pi G}\frac{b'}{r^2}, \tag{2}$$

$$p_r\left(r\right) = \frac{1}{8\pi G}\left[\frac{2}{r}\left(1-\frac{b\left(r\right)}{r}\right)\phi' - \frac{b}{r^3}\right], \tag{3}$$

$$p_t\left(r\right) = \frac{1}{8\pi G}\left(1-\frac{b\left(r\right)}{r}\right)\left[\phi'' + \phi'\left(\phi' + \frac{1}{r}\right)\right] - \frac{b'r - b}{2r^2}\left(\phi' + \frac{1}{r}\right), \tag{4}$$

where $\rho\left(r\right)$ is the energy density, $p_r\left(r\right)$ is the radial pressure, and $p_t\left(r\right)$ is the lateral pressure and using the conservation of the stress-energy tensor, in the same

orthonormal reference frame, one gets

$$p_r' = \frac{2}{r}\left(p_t - p_r\right) - \left(\rho + p_r\right)\phi'. \tag{5}$$

One strategy to obtain solutions for the EFE is represented by imposing an Equation of State (EoS) of the following forms:

(1) $p_r(r) = \omega\rho(r)$, with $\omega$ constant,
(2) $p_r(r) = \omega\left(r\right)\rho(r)$, with $\omega$ function of the radial coordinate $r$,
(3) $p_r(r) = \omega\rho^{\gamma}(r)$, with $\gamma \in \mathbb{R}$.

Of course, this list does not exhaust the possibilities of constraining the relationship between $p_r(r)$ and $\rho(r)$. For instance, by imposing the first EoS, one finds

$$\phi' = \frac{b\left(r\right) + \omega b\left(r\right)' r}{2r^2\left(1 - \frac{b(r)}{r}\right)} \tag{6}$$

and if we also assume Zero Tidal Forces (ZTF), one gets

$$\phi\left(r\right) = C \qquad \text{and} \qquad b\left(r\right) = r_0\left(\frac{r_0}{r}\right)^{\frac{1}{\omega}}, \tag{7}$$

where the condition $b\left(r_0\right) = r_0$ has been used. The parameter $\omega$ is restricted by the following conditions

$$b'\left(r_0\right) < 1; \qquad \frac{b\left(r\right)}{r} \underset{r\to+\infty}{\to} 0 \qquad \Longrightarrow \qquad \begin{cases} \omega > 0 \\ \omega < -1 \end{cases} \tag{8}$$

and the metric (1) assumes the particular simple expression

$$ds^2 = -dt^2 + \frac{dr^2}{1 - \left(\frac{r_0}{r}\right)^{\frac{1}{\omega}+1}} + r^2 d\Omega^2. \tag{9}$$

It is interesting to note that when $\omega \to \infty$, one finds

$$ds^2 = -dt^2 + \frac{dr^2}{1 - \frac{r_0}{r}} + r^2 d\Omega^2, \tag{10}$$

namely a traversable wormhole with zero energy density and ZTF. The corresponding Stress-Energy Tensor (SET) is given by

$$T_{\mu\nu} = \left(\rho\left(r\right), p_r\left(r\right), p_t\left(r\right), p_t\left(r\right)\right) = \frac{r_0}{2\kappa r^3}\left(0, 0, 1, 1\right), \tag{11}$$

that is it is formed by pure transverse pressure. To establish how much negative energy density is necessary, it is useful the computation of the "volume integral quantifier", which provides information about the "total amount" of averaged null energy condition (ANEC) violating matter in the space-time[3]. This is defined by

$$I_V = \int \left[\rho(r) + p_r(r)\right] dV \tag{12}$$

and for the line element (1), one can write

$$I_V = \frac{1}{\kappa}\int_{r_0}^{+\infty}(r-b(r))\left[\ln\left(\frac{e^{2\phi(r)}}{1-\frac{b(r)}{r}}\right)\right]' dr. \tag{13}$$

For instance, for the metric (10), one finds

$$I_V = -\frac{1}{\kappa}\int_{r_0}^{+\infty}(r-r_0)\left[\ln\left(1-\frac{r_0}{r}\right)\right]' dr = -\frac{r_0}{\kappa}\left[\ln(r)\right]_{r_0}^{+\infty} \to -\infty \tag{14}$$

which means that an infinite amount of negative energy is necessary to build such a wormhole. This was also confirmed in Ref.[1]. For this reason, we are going to explore the possibilities offered by the second EoS. The third EoS requires a careful analysis and it will not presented here. In particular, we would like to consider Yukawa-type profiles for $b(r)$ and $\omega(r)$. The hope is that such a profile concentrates more the energy density to the wormhole throat and thus minimizes the usage of exotic matter.

## 2. The Inhomogeneous Equation of State and the Volume Integral Quantifier

When we apply the inhomogeneous EoS $p_r = \omega(r)\rho$ to find the corresponding shape function, one finds

$$b(r) = r_0 \exp\left[-\int_{r_0}^{r}\frac{d\bar{r}}{\omega(\bar{r})\bar{r}}\right]. \tag{15}$$

The shape function (15) is obtained by imposing $\phi'(r) = 0$. Since we know the form of the redshift function and of the shape function, the SET can also be easily computed

$$T_{\mu\nu} = \frac{r_0}{\kappa r^3}\left(-\frac{1}{\omega(r)}, -1, \frac{1}{2\omega(r)}+\frac{1}{2}, \frac{1}{2\omega(r)}+\frac{1}{2}\right)\exp\left[-\int_{r_0}^{r}\frac{d\bar{r}}{\omega(\bar{r})\bar{r}}\right]. \tag{16}$$

Note that the SET (16) is traceless. We will examine two specific choices for $\omega(r)$.

### 2.1. *Two examples for $\omega(r)$*

We are going to focus our attention on two examples. The first one is represented by

$$\omega(r) = \frac{1}{\mu r} \qquad \Longrightarrow \qquad \omega(r_0) = \frac{1}{\mu r_0}, \tag{17}$$

leading to

$$b(r) = r_0\exp\left[-\mu(r-r_0)\right] \qquad and \qquad b'(r) = -\mu r_0\exp\left[-\mu(r-r_0)\right] \tag{18}$$

satisfying therefore the flare-out condition since $b'(r_0) = -\mu r_0$. The form of the metric (1) therefore becomes

$$ds^2 = -dt^2 + \frac{dr^2}{1-\frac{r_0\exp[-\mu(r-r_0)]}{r}} + r^2 d\Omega^2, \tag{19}$$

which looks like a Yukawa profile, at least on the radial part. From the line element (19), the SET is easily computed and assumes the following form

$$T_{\mu\nu} = \frac{r_0}{2\kappa r^3}\left[diag\left(-2\mu r, -2, \mu r+1, \mu r+1\right)\right]\exp\left[-\mu\left(r-r_0\right)\right]. \tag{20}$$

Note that the SET is also traceless as it should be, because it is a particular case of the SET (16). On the throat the SET becomes

$$T_{\mu\nu} = \frac{1}{2\kappa r_0^2}\left[diag\left(-2\mu r_0, -2, \mu r_0+1, \mu r_0+1\right)\right] \tag{21}$$

and the following property is satisfied

$$\lim_{\mu\to 0}\lim_{r\to r_0} T_{\mu\nu} = \lim_{r\to r_0}\lim_{\mu\to 0} T_{\mu\nu}. \tag{22}$$

On the other hand, we find

$$\lim_{\mu\to\infty}\lim_{r\to r_0} T_{\mu\nu} \neq \lim_{r\to r_0}\lim_{\mu\to\infty} T_{\mu\nu}. \tag{23}$$

In particular, $\lim\limits_{\mu\to\infty}\lim\limits_{r-r_0} T_{\mu\nu}$ is not defined, while (27)

$$\lim_{r\to r_0}\lim_{\mu\to\infty} T_{\mu\nu} = \left[diag\left(0,0,0,0\right)\right] \tag{24}$$

corresponding to the Minkowski space written in spherical coordinates. This can be confirmed also by looking at the line element (19). The corresponding $I_V$ for the line element (1) becomes

$$\begin{aligned} I_V &= \frac{1}{\kappa}\int_{r_0}^{+\infty}\left(r - r_0 e^{-\mu(r-r_0)}\right)\left[\ln\left(1-\frac{r_0}{r}e^{-\mu(r-r_0)}\right)\right]' dr \\ &= \frac{r_0}{\kappa}\left(1+e^{\mu r_0}\,\mathrm{Ei}_1\left(\mu r_0\right)\right), \end{aligned} \tag{25}$$

where $\mathrm{Ei}_1\left(x\right)$ is the exponential integral. Since the integrand is finite, we can easily evaluate the behavior close to the throat, whose result is

$$\begin{aligned} I_V \simeq \frac{r_0}{\kappa}\,&\left(1-\exp\left(\mu r_0\right)\mathrm{Ei}_1\left(\mu\left(r_0+\varepsilon\right)\right) + \exp\left(\mu r_0\right)\mathrm{Ei}_1\left(\mu r_0\right) - \exp\left(-\mu\varepsilon\right)\right) \to 0 \\ &\text{when } \varepsilon \to 0, \end{aligned} \tag{26}$$

namely $I_V$ can be arbitrarily small. The second example we are going to discuss is the following

$$\omega\left(r\right) = \frac{\exp\left[-\mu\left(r-r_0\right)\right]}{\mu r}, \tag{27}$$

which has the following properties

$$\omega\left(r\right) \underset{r\to\infty}{\to} 0 \qquad and \qquad \omega\left(r_0\right) = \frac{1}{\mu r_0}. \tag{28}$$

Plugging (27) into the shape function (15), one finds

$$\begin{aligned} b(r) &= r_0\exp\left[1-\exp\left(\mu\left(r-r_0\right)\right)\right] \\ b'(r) &= -\mu r_0\,\exp\left[\mu\,\left(r-r_0\right)+1-\exp\left(\mu\,\left(r-r_0\right)\right)\right], \end{aligned} \tag{29}$$

satisfying the flare-out condition since

$$b'(r_0) = -\mu r_0\,. \tag{30}$$

Even in this case, the SET is easily computed and assumes the explicit form

$$T_{\mu\nu} = \frac{r_0}{2\kappa r^3}\left[diag\left(-\frac{2}{\omega\,(r)}, -2, \frac{1}{\omega\,(r)}+1, \frac{1}{\omega\,(r)}+1\right)\right]\exp\left[1-\exp\left(\mu\,(r-r_0)\right)\right], \tag{31}$$

where $\omega\,(r)$ is given by Eq.(27). On the throat, one gets

$$T_{\mu\nu} = \frac{1}{2\kappa r_0^2}\left[diag\left(-2\mu r_0, -2, \mu r_0+1, \mu r_0+1\right)\right] \tag{32}$$

and also in this case we obtain the same behavior of the SET (20). It is interesting to note that, on the throat, the behavior of the SET in (20) and in (31) are equal. The corresponding $I_V$ for the line element (1) becomes

$$I_V = -\frac{1}{\kappa}\int_{r_0}^{+\infty}\left(r - r_0\exp\left[1-\exp\left(\mu\,(r-r_0)\right)\right]\right) \times\left[\ln\left(1-\frac{r_0\exp\left[1-\exp\left(\mu\,(r-r_0)\right)\right]}{r}\right)\right]'dr = -\frac{1}{\kappa}\int_{r_0}^{+\infty} f\,(r)\,dr, \tag{33}$$

where

$$f\,(r) = \frac{r_0\exp\left[1-\exp\left(\mu\,(r-r_0)\right)\right]\left(\exp\left(\mu\,(r-r_0)\right)\mu\,r+1\right)}{r}. \tag{34}$$

The exact evaluation of $I_V$ is quite complicated but, an estimate when $r \gg r_0$ is possible to see if the integral is finite for large $r$. Indeed, one finds that the integrand becomes

$$f\,(r) \underset{r\to\infty}{\simeq} \frac{\mu r_0\exp\left(\mu\,(r-r_0)+1\right)}{\exp\left(\exp\left(\mu\,(r-r_0)\right)\right)}$$
$$\Longrightarrow I_V \simeq -r_0\exp\left(1-\exp\left(\mu\,(r-r_0)\right)\right) \to 0 \qquad \text{when } r\to\infty \tag{35}$$

and therefore $I_V$ will be finite close to infinity. On the other hand, close to the throat, we can write

$$I_V \simeq \frac{1}{\kappa}\int_{r_0}^{r_0+\varepsilon}\left(\mu r_0+1\right)dr \to 0 \qquad \text{when } \varepsilon\to 0. \tag{36}$$

Note that we might have chosen the following form

$$\omega\,(r) = \frac{\exp\left(-\mu r\right)}{\mu r}, \tag{37}$$

instead of (27). However, while the properties at infinity are equal to the profile (27), close to the throat one finds

$$\omega\,(r_0) = \frac{\exp\left(-\mu r_0\right)}{\mu r_0}. \tag{38}$$

As a consequence, the shape function becomes

$$b\left(r\right)=r_{0}\exp\left(\exp\left(\mu r_{0}\right)-\exp\left(\mu r\right)\right) \tag{39}$$

while the energy density is

$$\rho(r)=-\frac{\mu r_{0}}{\kappa r^{2}}\exp\left(\mu r+\exp\left(\mu r_{0}\right)-\exp\left(\mu r\right)\right). \tag{40}$$

If we compare the energy density obtained with (27) with the one computed with (37), we can see that, on the throat, the amount of negative energy density is larger for the choice (37) because

$$\rho(r)=-\frac{\mu}{\kappa r_{0}}\exp\left(\mu r_{0}\right). \tag{41}$$

This implies that the choice (27) is favored from an energetically point of view. The same conclusion can be reached comparing the energy density obtained with (37) with the one obtained with (17). On the other hand, when we compare the energy density obtained with (17) and the one obtained with (27), we find

$$\frac{\rho_{\omega_{2}}(r)}{\rho_{\omega_{1}}(r)}=\frac{\exp\left[\mu\ \left(r-r_{0}\right)+1-\exp\left(\mu\ \left(r-r_{0}\right)\right)\right]}{\exp\left[-\mu\left(r-r_{0}\right)\right]}\underset{r\rightarrow\infty}{\longrightarrow}0, \tag{42}$$

where $\rho_{\omega_{1}}(r)$ is the energy density obtained with the help of (17) and represented by the first component of the SET (20), while $\rho_{\omega_{2}}(r)$ is the energy density obtained with the help of (27) and represented by the first component of the SET (31). The behavior shown in (42) reveals that the negativity of the energy density is really concentrated in the proximity of the throat. A further investigation inspired by Yukawa profiles could be related to the Self Sustained Traversable Wormholes, i.e., traversable wormholes which are sustained by their own quantum fluctuations[4–7].

## References


1. M. S. Morris and K. S. Thorne, *Am. J. Phys.* **56**, 395 (1988).
2. Visser M 1995 *Lorentzian Wormholes: From Einstein to Hawking* (American Institute of Physics, New York).
3. M. Visser, S. Kar and N. Dadhich, Phys. Rev. Lett. 90, 201102 (2003).
4. R. Garattini, *Class. Quant. Grav.* **22**, 1105 (2005); arXiv:gr-qc/0501105.
5. R. Garattini, *Class. Quant. Grav.* **24**, 1189 (2007); arXiv: gr-qc/0701019.
6. R. Garattini and F. S. N. Lobo, *Class. Quant. Grav.* **24**, 2401 (2007); arXiv:gr-qc/0701020.
7. R. Garattini and F. S. N. Lobo, *Phys. Rev.* **D 85** (2012) 024043; ArXiv: 1111.5729 [gr-qc].

# Black holes and wormholes in light of Weyl transformations*

Fayçal Hammad

*Department of Physics and Astronomy & STAR Research Cluster, Bishop's University*
*2600 College Street, Sherbrooke, QC, J1M 1Z7 Canada*
*Physics Department, Champlain College-Lennoxville*
*2580 College Street, Sherbrooke, QC, J1M 0C8 Canada*
*E-mail: fhammad@ubishops.ca*

A lot can be learned about black holes and wormholes by re-scaling spacetime itself without changing the coordinates used to describe it. Such a conformal transformation is called a Weyl transformation. It takes spacetime from a given frame — called Einstein frame — to a conformal frame — called Jordan frame. Such a transformation reveals that horizons and wormholes might appear/disappear in the conformal frame even if they were absent/present in the original frame. It arises both from the simple prescription for defining black holes and wormholes, as well as from the more sophisticated definitions. In addition, some definitions might be transformed into one another under Weyl transformations.

*Keywords*: Black holes; Wormholes; Weyl transformation; Cosmology.

## 1. Introduction

A Weyl transformation consists in re-scaling spacetime without changing its coordinates. It is described by a simple transformation that takes an original spacetime metric $g_{\mu\nu}$ into a new metric $\tilde{g}_{\mu\nu}$ such that,

$$\tilde{g}_{\mu\nu} = e^{2\Omega} g_{\mu\nu}, \tag{1}$$

where $\Omega(x)$ is spacetime-dependent function, everywhere regular and non-vanishing.

On the other hand, various definitions of black holes and wormholes have been given in the literature. They range from what could be considered as a simple "prescription" to what could be called "sophisticated" definitions.

## 2. The simple "prescription"

This prescription works only for spherically symmetric metrics that depend only on the time coordinate $t$ and a radial coordinate. The latter could be chosen to be the physical areal radius $R$ that multiplies $\mathrm{d}\Theta^2 = \mathrm{d}\vartheta^2 + \sin^2\vartheta \mathrm{d}\varphi^2$. The radial coordinate could also be chosen to be any other real parameter $r$ on which all the components of the metric would depend.

Using $R$ as the radial coordinate, a spherically symmetric metric is written as,

$$\mathrm{d}s^2 = -A(t,R)\mathrm{d}t^2 + B(t,R)\mathrm{d}R^2 + R^2\mathrm{d}\Theta^2, \tag{2}$$

where, $A(t,R)$ and $B(t,R)$ are functions of the time coordinate $t$ and the areal radius $R$. The simple prescription for defining black holes and wormholes consists then

*This talk is based on the published paper 1

(see Ref. 2 and the references therein) in checking whether the following algebraic equation,

$$g^{\mu\nu}\nabla_\mu R\nabla_\mu R = 0, \tag{3}$$

has a single- or a double-root. The single-root solution is identified with the location of the closed 2-surface of a black hole horizon, whereas the double-root is identified with the location of the 2-surface of a wormhole throat. Although this procedure is incomplete and imprecise, it is useful in practice since it gives a quick way of localizing the radial coordinate of an apparent horizon and/or a wormhole throat.

### 2.1. *Weyl transformation and the simple prescription*

The Weyl conformal transformation (1) changes the prescription (3) and the metric (2) into the following ones, respectively, to be used in the conformal frame,[2]

$$\mathrm{d}\tilde{s}^2 = -e^{2\Omega}A(t,R)\mathrm{d}t^2 + e^{2\Omega}B(t,R)\mathrm{d}R^2 + \tilde{R}(t)^2\mathrm{d}\Theta^2, \tag{4}$$

$$\tilde{g}^{\mu\nu}\tilde{\nabla}_\mu\tilde{R}\tilde{\nabla}_\mu\tilde{R} = 0. \tag{5}$$

Here, $\tilde{R}(t) = e^{2\Omega}R(t)$ is the areal radius in te conformal frame. Using the metric (4), after having rewritten it entirely in terms of the pair $(\tilde{t},\tilde{R})$ — which is in itself a very tedious task 2 — and then applying the conformal prescription (5), one finds,

$$\frac{1}{B}(\Omega_{,R}R+1)^2 - \frac{1}{A}\Omega_{,t}^2R^2 = 0. \tag{6}$$

While this condition does give the locations of the various possible black hole horizons and wormhole throats in the static case, it does not describe in detail their real origin in the dynamical case. As shown in Ref. 1, if one uses instead an arbitrary parameter $r$ as the radial coordinate and writes the metric (2) as,

$$\mathrm{d}s^2 = -A(t,r)\mathrm{d}t^2 + B(t,r)\mathrm{d}r^2 + R(t,r)^2\mathrm{d}\Theta^2, \tag{7}$$

Weyl's conformal transformation (1) turns the latter into the following form,

$$\mathrm{d}\tilde{s}^2 = -e^{2\Omega}A(t,r)\mathrm{d}t^2 + e^{2\Omega}B(t,r)\mathrm{d}r^2 + \tilde{R}(t,r)^2\mathrm{d}\Theta^2. \tag{8}$$

Applying the conformal prescription (5) on this last form of the metric gives after a pretty short and a simple calculation the following condition for detecting black holes and/or wormholes,[1]

$$\frac{1}{B}(\Omega_{,r}R+R_{,r})^2 - \frac{1}{A}(\Omega_{,t}R+R_{,t})^2 = 0. \tag{9}$$

We clearly see form this condition how and when a black hole and/or a wormhole might arise in the conformal frame. In fact, contrary to Eq. (6), the condition (9) contains all the necessary information on the time and space dependences of both the areal radius $R$ and the conformal exponent $\Omega$.

## 3. "Sophisticated" definitions for general spacetimes

Besides the simple prescription (3) for detecting black hole horizons and wormholes, various other definitions are also given in the literature. Since these other definitions are rather valid for any spacetime and involve the concept of null vectors, trapped surfaces, and geodesic expansions, we call them the sophisticated definitions. An important fundamental feature of the sophisticated definitions is that, in contrast to the simple prescription (3), these more rigorous definitions do not allow a black hole horizon to coincide with a wormhole throat. These definitions make a clear distinction between the two concepts.

### 3.1. *Black hole horizon*

A generic black hole horizon is defined as being the future outer trapping horizon.[3] Such a statement is translated into the following three conditions to be satisfied on the 2-surface $H$ of the horizon,[3]

$$\theta_+|_H = 0, \qquad \theta_-|_H < 0, \qquad \partial_-\theta_+|_H < 0. \tag{10}$$

The quantities $\theta_\pm$ represent the expansion of the outgoing (ingoing) null geodesics, the tangent vectors of which are denoted $l^\mu_\pm$, respectively. The partial derivatives, $\partial_\pm$, stand for derivative with respect to an affine parameter $u^\pm$ along the geodesic $l^\mu_\pm$.

### 3.2. *Hochberg-Visser wormhole*

A simple, covariant, and quasilocal definition of a wormhole throat, as defined in Ref. 4, and which we shall call a Hochberg-Visser wormhole, is that of a marginally anti-trapped surface. This definition does not involve any information about the faraway region outside the throat. It simply consists of the hypersurface foliated by compact spatial 2-surfaces $S$ on which the following conditions are satisfied,[4]

$$\theta_\pm|_S = 0 \qquad \text{and} \qquad \partial_\pm\theta_\pm|_S > 0. \tag{11}$$

### 3.3. *Hayward wormhole*

Another simple, covariant, and quasilocal definition for wormholes is the one given in Ref. 5, and which we shall call here a Hayward wormhole. According to this definition, a wormhole throat is a *timelike* hypersurface foliated by a non-vanishing minimal spatial 2-surface $S$ on a null hypersurface, *i.e.*, a timelike trapping horizon. This statement is formally expressed by the following three conditions,[5]

$$\theta_\pm|_S = 0, \qquad \partial_\pm\theta_\pm|_S > 0, \qquad \partial_\mp\theta_\pm|_S < 0. \tag{12}$$

In other words, Hochberg-Visser wormholes include spacelike hypersurfaces and therefore are not necessarily Hayward wormholes.

### 3.4. *Maeda-Harada-Carr wormhole*

This wormhole is a 2-surface required to be extremal on a spacelike hypersurface. Using a null coordinate foliation, a spherically symmetric metric takes the form $\mathrm{d}s^2 = -2e^{2f}\mathrm{d}u\mathrm{d}v + R^2\mathrm{d}\Theta^2$, where $u$ and $v$ are the null coordinates and $f = f(u, v)$ is a function of these. The Maeda-Harada-Carr wormhole then consists of the 2-sphere $S$, of radius $R = R(u, v)$, which is extremal and minimal in a spacelike radial direction $\zeta^\mu$. Formally, this translate into the following two conditions,[6]

$$R_{|A}\zeta^A\big|_S = 0 \qquad \text{and} \qquad R_{|AB}\zeta^A\zeta^B\big|_S > 0. \tag{13}$$

A vertical bar with the subscript, $|_A$, stands for a covariant derivative with respect to the two-metric $g_{AB}$ of the two-dimensional spacetime spanned by the null vectors $\partial_u$ and $\partial_v$.

### 3.5. *Tomikawa-Izumi-Shiromizu wormhole*

As in the case of the Maeda-Harada-Carr wormhole, the minimality of the 2-surface $S$ representing a Tomikawa-Izumi-Shiromizu wormhole is imposed on a spacelike hypersurface. In contrast to all the above definitions, however, what is required for this wormhole is the vanishing, not of the expansions themselves, but of the difference $\theta_+ - \theta_-$ between the outgoing and ingoing expansions. This translates into the following two conditions,[7]

$$\theta_+ - \theta_-|_S = 0 \qquad \text{and} \qquad (\partial_+ - \partial_-)(\theta_+ - \theta_-)|_S > 0. \tag{14}$$

## 4. Weyl transformation and the sophisticated definitions

In order to find the affect of a Weyl transformation on each of the above more rigorous definitions, one only needs to figure out how the null tangents $l^\mu_\pm$, the transverse metric $h_{\mu\nu}$, and the null expansions $\theta_\pm$ transform under (1). It is straightforward to show that $\tilde{h}_{\mu\nu} = e^{2\Omega}h_{\mu\nu}$, $\tilde{l}^\mu_\pm = e^{-\Omega}l^\mu_\pm$ and that $\tilde{\theta}_\pm = e^{-\Omega}\left(\theta_\pm + 2\partial_\pm\Omega\right)$.[1] With these simple transformations at hand, it is easy to find the behavior under Weyl's conformal transformation of all the above sophisticated black hole and wormhole definitions.

### 4.1. *Conformal black holes*

The definition (10) conformally transforms into the following three conditions,[1]

$$\begin{gathered}\theta_+ + 2\partial_+\Omega\big|_H = 0, \qquad \theta_- + 2\partial_-\Omega\big|_H < 0, \\ \partial_-\theta_+ + 2\partial_-\partial_+\Omega\big|_H < 0.\end{gathered} \tag{15}$$

From the first equality, we learn that in the absence of a black hole in the original frame, it is possible for a black hole horizon to arise in the new frame provided only that the conformal exponent $\Omega$ does vary with $u^+$. On the other hand, to have a black hole in the conformal frame if one already exists in the original frame, $\Omega$ must be independent of $u^+$.

### 4.2. *Conformal Hochberg-Visser wormhole*

The definition (11) of such a wormhole conformally transforms into the following two conditions,[1]

$$\theta_{\pm} + 2\partial_{\pm}\Omega\big|_S = 0 \quad \text{and} \quad \partial_{\pm}\theta_{\pm} + 2\partial_{\pm}\partial_{\pm}\Omega\big|_S > 0. \tag{16}$$

From the first equality we see that to have a Hochberg-Visser wormhole in the new frame if one already exists in the old frame, the conformal exponent must again be independent of the parameter $u^{\pm}$.

### 4.3. *Conformal Hayward wormhole*

The definition (12) of such a wormhole conformally transforms into the following three conditions,[1]

$$\begin{gathered} \theta_{\pm} + 2\partial_{\pm}\Omega\big|_S = 0, \qquad \partial_{\pm}\theta_{\pm} + 2\partial_{\pm}\partial_{\pm}\Omega\big|_S > 0, \\ \partial_{\mp}\theta_{\pm} + 2\partial_{\mp}\partial_{\pm}\Omega\big|_S < 0. \end{gathered} \tag{17}$$

Like for the Hochberg-Visser wormhole, if a Hayward wormhole already exists in the old frame another one might arise in the new frame if $\Omega$ is indepednet of $u^{+}$. Recall that a Hayward wormhole is necessarily a Hochberg-Visser wormhole but the converse is not true. However, if the conformal factor is chosen such tha the last inequality in (17) is not satisfied but the second inequality is, then a Hayward wormhole transforms into a pure Hochberg-Visser wormhole.

### 4.4. *Conformal Maeda-Harada-Carr wormhole*

The definition (13) of such a wormhole conformally transforms into the following three conditions,[1]

$$\begin{aligned} &\zeta^A \left(R_{|A} + R\Omega_{|A}\right)\Big|_S = 0 \\ &\zeta^A\zeta^B \left(R_{|AB} + R\Omega_{|AB} - R\Omega_{|A}\Omega_{|B}\right) \\ &\qquad + \zeta^A\zeta^B g_{AB}\left(R_{|C}\Omega^{|C} + R\Omega_{|C}\Omega^{|C}\right)\Big|_S > 0. \end{aligned} \tag{18}$$

Although these conditions are more complicated, it is clear that a Maeda-Harada-Carr wormhole can arise or disappear in the conformal frame depending on one's choice of $\Omega$. For a worked out example, though, see Ref. 1.

### 4.5. *Conformal Tomikawa-Izumi-Shiromizu wormhole*

The definition (14) of such a wormhole conformally transforms into the following two conditions,[1]

$$\begin{gathered} \theta_+ - \theta_- + 2\left(\partial_+\Omega - \partial_-\Omega\right)\Big|_S = 0, \\ \left(\partial_+ - \partial_-\right)\left[\theta_+ - \theta_- + 2\left(\partial_+\Omega - \partial_-\Omega\right)\right]\Big|_S > 0. \end{gathered} \tag{19}$$

Again, we see that a Tomikawa-Izumi-Shiromizu wormhole can arise in the conformal frame, even if in the original frame there was none, provided that the conformal factor is chosen to satisfy both conditions (19). Worked out examples are found in Ref. 1.

## 5. Conclusion

A conformal transformation can make black holes and wormholes appear or disappear just by judiciously choosing the conformal factor $\Omega$. A conformal transformation thus does have a non-trivial effect on a given physical concept just as it was already shown in Refs. 8, 9 regarding the fate of the concept of a quasilocal mass.

## Acknowledgements

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2017-05388), as well as by the STAR Research Cluster of Bishop's University.

## References

1. F. Hammad, "Revisiting black holes and wormholes under Weyl transformations", *Phys. Rev.* D **97**, 124015 (2018) [arXiv:1806.01388].
2. V. Faraoni, A. Prain and A. F. Zambrano Moreno, "Black holes and wormholes subject to conformal mappings", *Phys. Rev.* D **93**, 024005 (2016) [arXiv:1509.04129].
3. S. A. Hayward, "General laws of black-hole dynamics", *Phys. Rev.* D **49**, 6467 (1994) [arXiv:gr-qc/9303006].
4. D. Hochberg and M. Visser, "Dynamic wormholes, anti-trapped surfaces, and energy conditions", *Phys. Rev.* D **58**, 044021 (1998) [arXiv:gr-qc/9802046].
5. S. A. Hayward, "Wormhole dynamics in spherical symmetry", *Phys. Rev.* D **79**, 124001 (2009) [arXiv:0903.5438].
6. H. Maeda, T. Harada and B. J. Carr, "Cosmological wormholes", *Phys. Rev.* D **79**, 044034 (2009) [arXiv:0901.1153].
7. Y. Tomikawa, K. Izumi and T. Shiromizu, "New definition of a wormhole throat", *Phys. Rev.* D **91**, 104008 (2015) [arXiv:1503.01926].
8. F. Hammad, "Conformal mapping of the Misner-Sharp mass from gravitational collapse", *Int. J. Mod. Phys.* D **25**, 1650081 (2016) [arXiv:1610.02951].
9. F. Hammad, "More on the conformal mapping of quasi-local masses: The Hawking-Hayward case", *Class. Quantum Grav.* **33**, 235016 (2016) [arXiv:1611.03484].

# Towards constraining realistic Lorentzian wormholes through observations

Sayan Kar[a], Sukanta Bose[b], S. Aneesh[a]

[a] *Department of Physics and CTS, Indian Institute of Technology Kharagpur, 721302, India*

[b] *Inter University Centre for Astronomy and Astrophysics, Post Bag 4, Ganeshkhind, Pune, 410007, India*
*&*
*Department of Physics and Astronomy, Washington State University, 1245 Webster, Pullman, WA 99164-2614, USA*

Ever since their revival three decades ago, in the seminal work of Morris and Thorne, Lorentzian wormholes in General Relativity have led an uncomfortable existence because they require matter which violates the well-known energy conditions. However, in scalar-tensor and other theories of gravity, realistic wormholes can indeed exist with 'normal matter'. We illustrate this with some known examples and also by explicitly constructing a zero Ricci scalar wormhole in a four dimensional scalar-tensor, on-brane gravity theory arising from the two-brane Randall-Sundrum model with one extra dimension. If such a wormhole could arise as the end-state of some astrophysical process, its ringdown may be studied using gravitational waves. With this aim, we obtain the scalar quasinormal modes in this class of wormholes and choose to identify them as for the 'breathing mode' associated with gravitational waves in scalar-tensor theories. Finally, if a breathing mode is indeed observed in LIGO-like detectors with design sensitivity, and has a maximum amplitude equal to that of the tensor mode that was observed of GW150914, then for a range of values of the wormhole parameters we will be able to discern it from a black hole. If in future observations we are able to confirm the existence of such wormholes, we would, at one go, have some indirect evidence of a modified theory of gravity as well as extra spatial dimensions.

## 1. Lorentzian wormholes: a quick introduction

The earliest example of a spacetime resembling the Lorentzian wormhole of today is the well-known Einstein-Rosen bridge[1], which, though a two-sheeted geometry, is however non-traversable and has a degenerate metric (the determinant of $g_{ij}$ vanishes at the throat). Subsequent to Einstein-Rosen, wormholes appear in the Misner-Wheeler study of *geons*[2]. Later, in the early 1970s, Ellis and Bronnikov[3] constructed a wormhole solution of the Einstein-scalar equations (the 'phantom' scalar had a wrong-sign kinetic energy). The modern-day Lorentzian wormhole was defined largely in the 1988 paper of Morris and Thorne[4] who also constructed a wormhole time machine[5]. Various other methods, such as *Schwarzschild surgery* were developed by Visser in the 1990s (see[6] for further details and an overview of early work).

Lorentzian wormholes have an existence problem in General Relativity(GR), largely due to the fact that they require matter to violate the *Weak or Null Energy Condition*. The cause behind this is the nature of the geometry which acts as a *defocusing lens* for a congruence of null or timelike geodesics[4]. Such defocusing entails a violation of the *convergence conditions, null or timelike*. Since the convergence

conditions are directly related to the energy conditions via Einstein's field equations of GR, we end up violating the energy conditions obn matter as well[4]. Various ways have been suggested in order to resolve this issue[7] within GR, but success, at best, is very limited.

A way out of this impasse is to consider modified theories of gravity. Here we will focus on one class of such theories which are born out of the effective on-brane scalar-tensor gravity resulting from the two-brane Randall-Sundrum model with one orbifolded extra dimension[8]. In such an effective theory one can avoid energy-condition violation, though violation of the convergence condition (as expected for wormholes) remains upheld. The above possibility emerges because we have a *scalar field* contribution on the R. H. S. of the field equations, apart from the energy-momentum of matter. The scalar field piece is not quite 'matter', because its origin lies in the *distance between the branes embedded in the bulk*. We can therefore have *realistic Lorentzian wormholes* in this theory, a fact we demostrate below, in brief.

Given such a possible wormhole with 'good' matter, one is therefore encouraged to see if there are chances of detecting it. If we imagine some astrophysical collision process creating such a wormhole, then, its quasinormal (QNM) ringing will indeed be a signature. It is also important to note that among the allowed gravitational wave polarisations in a scalar-tensor theory there is a *scalar breathing mode*, apart from the standard + and × polarisations which exist in GR. Hence, a study of the scalar wave equation and a determination of the scalar quasinormal modes could be a way of detecting the wormhole. After deriving the frequencies and time-constants of the QNMs as functions of the metric parameters, we use them to estimate the percentage errors in the parameters via the Fisher matrix formalism. We follow this line of investigation in our work. Our findings suggest that observations with multiple detectors can be used to put interesting limits on scalar-mode emissions and, thereby, constrain parameters of such wormholes.

## 2. A WEC satisfying realistic wormhole in a two-brane RS model

Let us first write down the wormhole spacetime which is a solution in a effective on-brane scalar-tensor theory and is threaded by matter satisfying the WEC. The Einstein field equations for such a scalar-tensor theory are given as[8]:

$$\Phi G_{\mu\nu} = \bar{\kappa} T^b_{\mu\nu} + (\nabla_\mu \nabla_\nu \Phi - g_{\mu\nu} \nabla^\alpha \nabla_\alpha \Phi) - \frac{3}{2(1+\Phi)} \left( \nabla_\mu \Phi \nabla_\nu \Phi - \frac{1}{2} g_{\mu\nu} \nabla^\alpha \Phi \nabla_\alpha \Phi \right) \tag{1}$$

The scalar field equation turns out to be:

$$\Box \Phi = \frac{\bar{\kappa}}{l} \frac{T^b}{2\omega + 3} - \frac{1}{2\omega + 3} \frac{d\omega}{d\Phi} \nabla^\alpha \Phi \nabla_\alpha \Phi \tag{2}$$

Note that the R. H. S of the Einstein equations has a 'matter stress energy' (the $T^b_{\mu\nu}$) and a contribution from the scalar field $\phi$. If one uses the $G_{\mu\nu}$ in the R. H. S. of,

say, the null Raychaudhuri equation, one notices that *the convergence condition can be violated, without violating the WEC for matter*. In other words, the negativity required for the violation of the convergence condition can be supplied via the $\phi$ dependent pieces on the R. H. S. in the Einstein equation, keeping $T^b_{\mu\nu}k^\mu k^\nu$ always positive. The $R = 0$ line element which achieves this, with an everywhere finite and non-zero scalar field is[9–11]:

$$ds^2 = -\left(\kappa + \lambda\sqrt{1 - \frac{2M}{r}}\right)^2 dt^2 + \frac{dr^2}{1 - \frac{2M}{r}} + r^2 d\Omega_2^2 \tag{3}$$

where $\kappa, \lambda > 0$, $r \geq 2M$. The wormhole throat is at $r = 2M$. It has a spatial slice identical to that for the Schwarzschild black hole.

The WEC inequalities for the required matter ($\rho \geq 0, \rho + \tau \geq 0, \rho + p \geq 0$, where $\rho$, $\tau$, $p$ are the non-zero, diagonal components of the energy momentum tensor), have been checked and are satisfied[11] as shown in Figs. 1,2,3. In these plots $x = \frac{M}{2r'}$ where $r'$ is the isotropic coordinate.

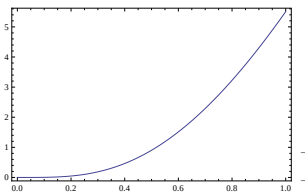



Fig. 1. $\rho$ vs. $x$

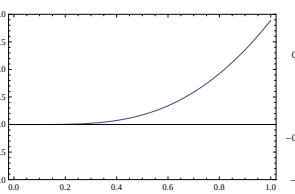



Fig. 2. $\rho + \tau$ vs. $x$

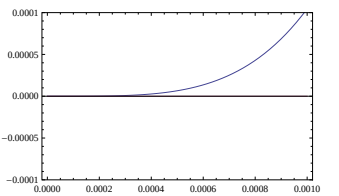



Fig. 3. $10^6(\rho + p)$ vs. $x$

## 3. How to detect: observational possibities using GWs

As mentioned earlier, we imagine the end-state of a collision of astrophysical objects (black holes, neutron stars, wormholes?) as a two-parameter ($M$ and $\frac{\kappa}{\lambda}$) wormhole of the type discussed in the previous section. Being a solution of a scalar-tensor theory, where the *breathing mode* polarisation exists apart from the standard GR + and × polarisations[12], it is plausible that scalar perturbations produce it. It can be shown that the scalar wave equation which we wrote earlier, when perturbed (i.e. $\phi \to \phi + \delta\phi$) yields, with a suitable gauge choice, an equation of the form $\Box\delta\phi = 0$. Therefore, by studying the scalar QNMs one can obtain information about the frequencies and damping rates of the ring-down process. Of course, gravitational perturbations must also be present, but here we focus only on the scalar breathing mode. The effective potential as a function of the tortoise coordinate $r_*$ is shown in Fig. 4.[13] Note the double barrier structure of the potential. The time-domain profiles for the ring-down are shown in Figs. 5, 6.[13]

### 3.1. *Scalar quasinormal modes–breathing mode*

We find the quasinormal modes (fundamentals) by using Prony Fit as well as by using direct integration[13]. The values found using the two methods do agree with each other. In Figs.7,8 the real and imaginary parts of the QNMs are shown as functions of $\lambda$, where we have chosen $\kappa + \lambda = 1$. In order to use the QNMs for evaluating the percentage errors we fit the QNM profiles using *NonlinearLeastSquareFit* in *Mathematica 10.0*. The fitted curves are shown as the continuous lines in Figs. 7, 8.

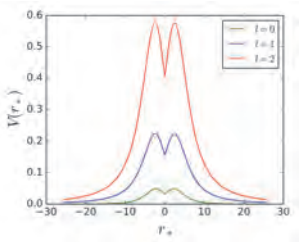



Fig. 4. The effective potential $V_{eff}(r_*)$.

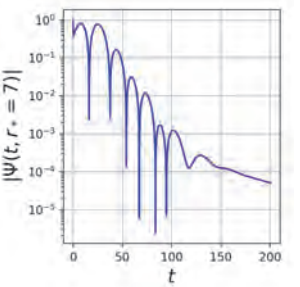



Fig. 5. Time domain profile

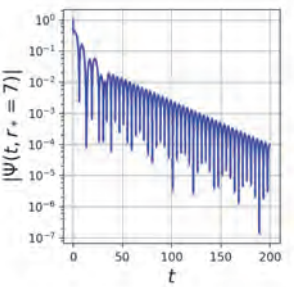



Fig. 6. Time domain profile

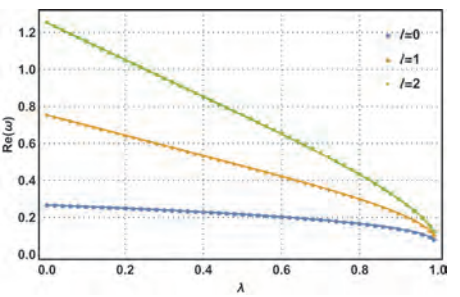



Fig. 7. Real part of QNM vs. $\lambda$.

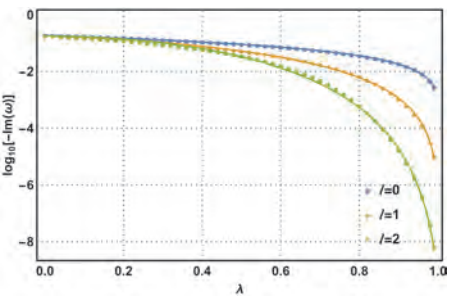



Fig. 8. Imaginary part of QNM vs $\lambda$.

An example of the fitted profile function is shown below. For $l = 0$, we have,

$$\nu = \frac{8628.13}{M}\left(1 - 0.29\frac{\lambda}{\lambda+\kappa}\right)\left(1 - \left(\frac{\lambda}{\lambda+\kappa}\right)^{2.36}\right)^{0.24} \tag{4}$$

$$\tau = \frac{M}{38908.58}\left(1 - 0.99\left(\frac{\lambda}{\lambda+\kappa}\right)^{0.94}\right)^{-1.02} \tag{5}$$

For $M = 68M_{sun}$ and $\frac{\kappa}{\lambda} = 0.1$ we obtain $\nu = 64$ Hz.

The breathing mode signal may be written as: $h(t) = A\sin(2\pi\nu t)e^{-t/\tau}$ where the strain amplitude $A$ contains the breathing-mode antenna pattern. Thus, knowing $\nu$ and $\tau$ from the extracted signal, one can find the $\frac{\kappa}{\lambda}$ and $M$, in principle. However, one must extract the signal from detector noise for which it is crucial to estimate errors.

### 3.2. *Error estimates*

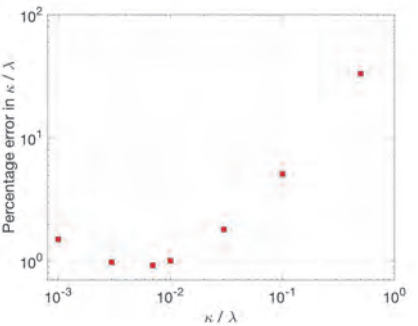



Fig. 9. Percentage error vs. $\frac{\kappa}{\lambda}$

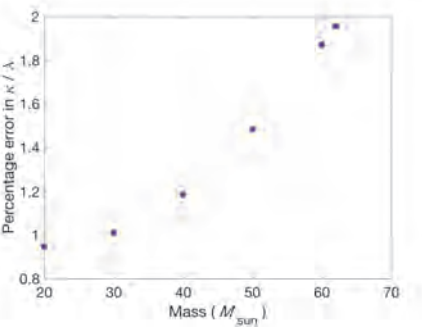



Fig. 10. Percentage error vs. $M$.

The percentage errors in $\frac{\kappa}{\lambda}$ can be calculated via the Fisher information matrix formalism. We estimate how accurately the wormhole parameters will be measurable using future interferometric detectors like aLIGO. To estimate the error in $\kappa/\lambda$, we compute that matrix for the damped-sinusoid signal in a single aLIGO detector at design sensitivity, for that parameter alone. The matrix is determined by the derivative of the signal $h$ with respect to $\kappa/\lambda$, which influences both the frequency and the damping time-constant of the signal.

We take $M$ to be known. For reference, the maximum QNM strain amplitude is assumed as $10^{-21}$, which is approximately the maximum amplitude (quadrupolar) of the GW150914 signal. Finally, we invert the information matrix to derive the estimated variance in the measured values of $\kappa/\lambda$. Its square-root gives the lower bound on the statistical error in $\kappa/\lambda$.

Fig. 9 shows the percentage errors in $\frac{\kappa}{\lambda}$ for $M = 30M_{sun}$. The error intially reduces as $\frac{\kappa}{\lambda}$ increases. The mode frequency gradually shifts to the more sensitive parts of the detector band. For higher values of $\frac{\kappa}{\lambda}$ (less than 1) the error increases due to decreasing time constant (heavy damping). Errors for $\frac{\kappa}{\lambda} > 1$ upto 10 have also been found using Einstein Telescope design sensitivities[13]. In Fig . 10, we have fixed $\frac{\kappa}{\lambda} = 0.01$. The error increases with increasing $M$ because the mode frequency decreases, placing the signal in the less sensitive part of the detector band. More details and other plots are available in[13].

### 3.3. *Summary and conclusions*

In the above, we have outlined a way of knowing, through GW observations, whether wormholes could exist. Earlier attempts towards detecting wormholes were largely using gravitational lensing. With new GW data, our proposal seems a better prospect. However, one really needs an astrophysical merger model. It is also necessary to study gravitational perturbations in order to obtain a more complete picture.

If we ever see the specific Lorentzian wormhole we proposed here, it will also be indirect evidence for extra dimensions and modified gravity. We hope that observations in the near future would be able to put better constraints and help decide the viability of Lorentzian wormholes in nature.

## Acknowledgments

One of the authors (SK) thanks Francisco Lobo for inviting him to give this talk. He (SK) also thanks the organisers of MG15 for their warm hospitality during the conference.

## References

1. A. Einstein and N. Rosen, Phys. Rev. **48**, 73 (1935).
2. C. W. Misner and J. A. Wheeler, Ann. Phys. **2**, 525(1957)
3. H. Ellis, J. Math. Phys. **14**, 104 (1973); K. A. Bronnikov, Acta Physica Polonica **B4**, 251 (1973).
4. M. S. Morris and K. S. Thorne, Am. J. Phys. **56**, 395 (1988).
5. M. S. Morris, K. S. Thorne and U. Yurtsever, Phys. Rev. Letts. **61**, 1446 (1988)
6. M. Visser, *Lorentzian wormholes: from Einstein to Hawking*, AIP Series in Computational and Applied Mathematical Physics, 1996.
7. S. Kar, Phys. Rev. **D49**, 862 (1994); M. Visser, S. Kar and N. Dadhich, Phys. Rev. Letts. **90**, 201102 (2003); H. Maeda and M. Nozawa, Phys. Rev. **D 78**, 024005 (2008); N. M. Garcia and F. S. N. Lobo, Class.Quant.Grav. **28**, 085018 (2011); R. Shaikh, Phys. Rev. **D 92**, 024015 (2015).
8. S. Kanno and J. Soda, Phys. Rev. **D 66**, 083506 (2002); L. Randall and R. Sundrum, Phys. Rev. Letts. **83**, 3370 (1999).
9. N. Dadhich, S. Kar, S. Mukherjee, and M. Visser, Phys. Rev. D **65**, 064004 (2002).
10. R. Casadio, A. Fabbri, and L. Mazzacurati, Phys. Rev. D **65**, 084040 (2001).
11. S. Kar, S. Lahiri, and S. SenGupta, Phys. Lett. B **750**, 319-324 (2015); R. Shaikh and S. Kar, Phys. Rev. **D94**,024011 (2016).
12. C. M. Will, *Theory and Experiment in Gravitational Physics*, Cambridge University Press (1993).
13. S. Aneesh, S. Bose and S. Kar, Phys. Rev. **D97**, 124004 (2018).

# Classical and quantum strong energy inequalities and the Hawking singularity theorem

P. J. Brown, C. J. Fewster and E.-A. Kontou*

*Department of Mathematics, University of York, Heslington, York, YO10 5DD, UK*
**E-mail: eleni.kontou@york.ac.uk*

Hawking's singularity theorem concerns matter obeying the strong energy condition (SEC), which means that all observers experience a non-negative effective energy density (EED). The SEC ensures the timelike convergence property. However, for both classical and quantum fields, violations of the SEC can be observed even in the simplest of cases, like the Klein-Gordon field. Therefore there is a need to develop theorems with weaker restrictions, namely energy conditions averaged over an entire geodesic and weighted local averages of energy densities such as quantum energy inequalities (QEIs). We present lower bounds of the EED for both classical and quantum scalar fields allowing nonzero mass and nonminimal coupling to the scalar curvature. In the quantum case these bounds take the form of a set of state-dependent QEIs valid for the class of Hadamard states. We also discuss how these lower bounds are applied to prove Hawking-type singularity theorems asserting that, along with sufficient initial contraction, the spacetime is future timelike geodesically incomplete.

*Keywords*: quantum fields, gravity, energy conditions, quantum inequalities, singularities

## 1. Introduction

A spacetime is defined to be singular if it possesses at least one incomplete geodesic. The question of whether or not cosmological models either originate or terminate in singularities has been an active subject of research since the formulation of the general theory of relativity. Initial efforts focused on models with high levels of symmetry, until Raychaudhuri's 1955 paper[1] paved the way to more general results. The Raychaudhuri equations in their modern form[2] present the evolution of geodesic congruences, and are the heart of most singularity theorems. For a timelike irrotational congruence with velocity field $U^\mu$, the expansion $\theta = \nabla \cdot U$ satisfies

$$\nabla_U \theta = R_{\mu\nu} U^\mu U^\nu - 2\sigma^2 - \theta^2/(n-1)\,, \tag{1}$$

where $n$ is the spacetime dimension, $\sigma$ is the shear scalar and $R_{\mu\nu}$ is the Ricci tensor. Senovilla[3] has described the singularity theorems in terms of a 'pattern theorem' with three ingredients: an energy condition establishes a focussing effect for geodesics, a causality condition removes the possibility of closed timelike curves, and a boundary or initial condition establishes the existence of some trapped region of spacetime. We divide singularity theorems into 'Hawking-type' after Hawking's original theorem[4] and 'Penrose-type' after Penrose's[5], depending on whether they demonstrate timelike or null geodesic incompleteness respectively. Hawking-type results, which concern us here, are based on the strong energy condition (SEC), which requires that the effective energy density (EED) is non-negative. The EED is defined as $\rho_U := T_{\mu\nu} U^\mu U^\nu - T/(n-2)$ and is easily made negative at individ-

ual points even for the classical minimally coupled scalar field. This situation is exacerbated in quantum field theory (QFT), in which none of the pointwise energy conditions can hold[6].

For these reasons there has long been interest in establishing singularity theorems under weakened energy assumptions, such as requiring the averaged energy along the geodesic to be non-negative. Examples of singularity theorems with such conditions include Refs. 7–9 but none of them address the case of a condition obeyed by quantum fields. Within a QFT the weakened energy conditions take the form of quantum energy inequalities (QEIs). These were introduced by Ford[10] and express a restriction on the possible magnitude and duration of any negative energy densities or fluxes. Our approach follows Ref. 11 which proved singularity theorems with energy conditions inspired by QEIs. For the Hawking-type theorem, it is assumed that

$$\int R_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu f^2(\tau)d\tau \geq -|||f|||^2\,, \tag{2}$$

on any timelike geodesic. Here $\dot{\gamma}^\mu$ is the tangent to the timelike geodesic $\gamma$, $|||\cdot|||^2$ is a Sobolev norm of the form $|||f|||^2 = \sum_{\ell=0}^{L} Q_\ell ||f^{(\ell)}||^2$ and $Q_\ell$ are non-negative constants. It is also assumed that there is a compact Cauchy hypersurface intersected by $\gamma$ at $\tau = 0$. Then the spacetime is future timelike geodesic incomplete provided that the required initial contraction obeys

$$\theta(0) \leq -\frac{c}{2} - \int_{-\tau_0}^{\infty} R_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu d\tau - |||f|||^2\,, \tag{3}$$

for some $c > 0$ and $f \in C^\infty(\mathbb{R})$ with $\operatorname{supp} f \subset [-\tau_0, \infty)$, $f(\tau) = e^{-c\tau/(n-1)}$ on $[0, \infty)$. It might seem strange at first that large positive values of $R_{\mu\nu}\dot{\gamma}^\mu\dot{\gamma}^\nu$ require larger initial contraction. The reason is that the average energy condition of Eq. (2) means that large positive energies in the past (before the $\theta(0)$ is measured) allow large negative energies in the future that can lead to a divergence of the congruence.

The purpose of this contribution is to describe bounds on the EED for classical and quantum fields, show that they are of the form of Eq. (2) and estimate the contraction needed to prove singularity theorems. In Sec. 2 we treat the classical Einstein-Klein-Gordon field while in Sec. 3 we present a bound on the renormalised expectation value of the EED for the non-minimally coupled quantum scalar field. In Sec. 4 we conclude with a summary and discussion of future work. Our geometrical conventions are those of Ref. 11 and we work on a globally hyperbolic manifold $M$.

## 2. A singularity theorem for the Einstein-Klein-Gordon field

The non-minimally coupled scalar field obeys the field equation $P_\xi\phi = 0$ with $P_\xi := \Box_g + m^2 + \xi R$, where $\xi$ is the coupling constant and

$$T_{\mu\nu} = (\nabla_\mu\phi)(\nabla_\nu\phi) + \frac{1}{2}g_{\mu\nu}(m^2\phi^2 - (\nabla\phi)^2) + \xi(g_{\mu\nu}\Box_g - \nabla_\mu\nabla_\nu - G_{\mu\nu})\phi^2\,. \tag{4}$$

For this stress-energy tensor, after some manipulation, the averaged EED on a timelike geodesic $\gamma$ is

$$\int_\gamma d\tau\, \rho_U\, f^2(\tau) = \int_\gamma d\tau \Bigg\{ -\frac{1-2\xi}{n-2} m^2 f^2(\tau) + \left(1 - 2\xi\frac{n-1}{n-2}\right)(\nabla_{\dot\gamma}\phi)^2 f^2(\tau) + \frac{2\xi}{n-2} h^{\mu\nu}(\nabla_\mu\phi)(\nabla_\nu\phi) f^2(\tau) + 2\xi[\nabla_{\dot\gamma}(f(\tau))\phi]^2 - 2\xi\phi^2(f'(\tau))^2 - \xi\mathcal{R}_\xi\phi^2 \Bigg\}, \quad (5)$$

where $\mathcal{R}_\xi = R_{\mu\nu}\dot\gamma^\mu\dot\gamma^\nu - 2\xi/(n-2)R$, and $h^{\mu\nu} = \dot\gamma^\mu\dot\gamma^\nu - g^{\mu\nu}$ is a positive definite metric. For $\xi \in [0, \xi_c]$, where $\xi_c$ is the conformal coupling constant, all the curvature independent terms have a definite sign, so

$$\int_\gamma d\tau\, \rho_U\, f^2(\tau) \geq -\int_\gamma d\tau \left\{ \frac{1-2\xi}{n-2} m^2 f^2(\tau) + \xi\left(2(f'(\tau))^2 + \mathcal{R}_\xi f^2(\tau)\right)\right\}\phi^2\,. \quad (6)$$

For a field obeying the Einstein's equation, we have that $8\pi\rho_U = R_{\mu\nu}\dot\gamma^\mu\dot\gamma^\nu$ and $(n/2-1)R = 8\pi T$. Moving the curvature terms to the left hand side of Eq. (6)

$$\int R_{\mu\nu}\dot\gamma^\mu\dot\gamma^\nu f(\tau)^2\, d\tau \geq -Q(\|f'\|^2 + \tilde{Q}^2\|f\|^2), \quad (7)$$

with $Q$ and $\tilde{Q}$ positive constants that depend on $\phi_{\max} \geq |\phi|$, $m$, $\xi$ and $n$. We can show that if

$$\theta(0) < -\tilde{Q}\sqrt{Q(n-1) + \frac{Q^2}{2}} - \frac{1}{2}QK\coth(K\tau_0)\,, \quad \text{with } K^2 \geq \tilde{Q}^2 + Q^{-1}R_{\mu\nu}\dot\gamma^\mu\dot\gamma^\nu \text{ on } (-\tau_0, 0]\,, \quad (8)$$

on a compact Cauchy surface, the spacetime is future timelike geodesic incomplete.

Finally, we want to estimate the necessary contraction for realistic scalar fields, using a hybrid model: a quantised scalar field in Minkowski spacetime of dimension 4, in a thermal state of temperature $T < T_m$, $T_m = mc^2/k$ with the maximum field value set by the expectation value of the Wick square at that temperature, $\phi^2_{\max} \sim \langle{:}\phi^2{:}\rangle_T$. As an example we consider the Higgs field with $m = 125\text{GeV}/c^2$, $\theta(0) \sim 10^{-14}\text{s}^{-1}$ for temperature up to $T = 10^{13}$K. Similar results are obtained for other examples. We conclude that when the field mass is taken equal to that of an elementary particle, and provided that the temperature stays below early universe levels we need very little initial contraction for geodesic incompleteness [a].

## 3. Strong quantum energy inequality

To quantise the EED we will follow the algebraic approach. We only consider quasifree, Hadamard states $\omega$, for which the two-point function $W(x,y) = \langle\Phi(x)\Phi(y)\rangle_\omega$ has a prescribed singularity structure so that the difference between two states is smooth.

[a] A detailed proof of the theorem and estimation of the initial contraction is given in Ref. 12.

We follow the methods of Hollands and Wald[13,14] to quantize the stress-energy tensor in a systematic and locally covariant way, and then define the EED as a quantum field by $\rho_U(f) = T_{\mu\nu}(U^\mu U^\nu - g^{\mu\nu}/(n-2)f)$. We are interested in expectation values of the quantized EED in state $\omega$, normal ordered relative to a reference Hadamard state $\omega_0$, $\langle{:}\rho_U{:}(f)\rangle_\omega = \langle\rho_U(f)\rangle_\omega - \langle\rho_U(f)\rangle_{\omega_0}$.

We now turn to the derivation of a quantum strong energy inequality (QSEI), a bound on the renormalised expectation value of EED averaged along a timelike geodesic $\gamma$. Choose any smooth $n$-bein $e_a$ $(a = 0, \ldots, n-1)$ on a tubular neighbourhood of $\gamma$, so that $U^\mu = e_0^\mu$ is everywhere timelike and agrees with $\dot\gamma^\mu$ on $\gamma$. The expectation values of the EED in Hadamard state $\omega$ can be written in terms of the coincidence limits, $[\![\cdot]\!]$ acting on $:W: = W - W_0$. Then we have

$$\langle{:}\rho_U{:}\rangle_\omega = [\![\hat\rho_1{:}W{:}]\!] + [\![\hat\rho_2{:}W{:}]\!] + \left(\xi\mathcal{R}_\xi - \frac{1-2\xi}{n-2}m^2\right)[\![{:}W{:}]\!] \text{ with} \tag{9}$$

$$\hat\rho_1 = \left(1 - 2\xi\frac{n-1}{n-2}\right)(\nabla_U \otimes \nabla_U) + \frac{2\xi}{n-2}\sum_{a=1}^{n-1}(\nabla_{e_a} \otimes \nabla_{e_a})\,, \tag{10a}$$

$$\hat\rho_2 = -2\xi(\mathbb{1} \otimes_{\mathfrak{s}} U^\mu U^\nu \nabla_\mu \nabla_\nu)\,, \tag{10b}$$

where $\otimes_{\mathfrak{s}}$ is the symmetrised tensor product $P \otimes_{\mathfrak{s}} P' = [(P \otimes P') + (P' \otimes P)]/2$. The contribution of the terms deriving from $\hat\rho_1$ to the averaged EED can be bounded from below, uniformly in $\omega$, using the methods of Ref. 15. By contrast, the mass term is negative definite for $\xi < 1/2$, while the geometric term $\mathcal{R}_\xi$ has no definite sign in general. This leaves $\hat\rho_2$, the contribution of which can be manipulated to a more convenient form. With these considerations we can prove

$$\langle{:}\rho_U{:} \circ \gamma\rangle_\omega(f^2) \geq -\left(\mathfrak{Q}_1[f] + \langle{:}\Phi^2{:} \circ \gamma\rangle_\omega(\mathfrak{Q}_2[f] + \mathfrak{Q}_3[f])\right) \text{ where} \tag{11}$$

$$\mathfrak{Q}_1[f] = \int_0^\infty \frac{d\alpha}{\pi}\left(\phi^*(\hat\rho_1 W_0)(\bar f_\alpha, f_\alpha) + 2\xi\alpha^2\phi^* W_0(\bar f_\alpha, f_\alpha)\right)\,, \tag{12a}$$

$$\mathfrak{Q}_2[f] = \frac{1-2\xi}{n-2}m^2 f^2(\tau) + 2\xi(f'(\tau))^2 \text{ and } \mathfrak{Q}_3[f] = \xi\mathcal{R}_\xi(\gamma(\tau))f(\tau)^2\,. \tag{12b}$$

Here $\phi^*$ is the distributional pull-back on the geodesic. An important feature of the QEI (11), is that the lower bound depends on the state of interest $\omega$. We should note however, that the only nontrivial quantum field appearing in the bound is the Wick square $:\Phi^2:$ which enables us to show that the QEI derived is nontrivial (see Ref. 16 for more details).

The establishment of a QSEI is a first step towards a Hawking-type singularity theorem result employing QEI hypotheses. Here we outline a method of obtaining such a result, part of an ongoing work to appear elsewhere. The singularity theorems require a geometric assumption, a condition on the curvature rather than the stress-energy tensor. In the case of classical fields we can use the Einstein equation while when we are treating quantum fields on a classical curved background we can instead use the semiclassical Einstein equation (SEE), $\langle T_{\mu\nu}\rangle_\omega = 8\pi G_{\mu\nu}$. In order to use the

QSEI and the SEE for general curved spacetimes the EED needs to be renormalised by subtracting the Hadamard parametrix instead of a reference state. To overcome this problem we note there is evidence (see e.g. Ref. 17) that in situations where the curvature is bounded we can find a uniform length $\tau_0$ which is small compared to local curvature. Then the Hadamard parametrix on $\tau_0$ approximates that of flat spacetime. Following, we can use Eq. (11) for $f \in C_0^\infty$ supported only to intervals with lengths at most $2\tau_0$. Choosing a state $\omega$ and a metric $g_{\mu\nu}$ that satisfy the EED and the stress energy tensor renormalised with respect to the Minkowski vacuum,

$$\int d\tau f^2(\tau) R_{\mu\nu} U^\mu U^\nu \geq -8\pi \left[ \int_0^\infty \frac{d\alpha}{\pi} \phi^* ((\nabla_U \otimes \nabla_U) W_0)(\bar{f_\alpha}, f_\alpha) + \frac{\mu^2 \phi_{\max}^2 ||f||^2}{n-2} \right], \tag{13}$$

where we also set $\xi = 0$, $\mu$ is the mass of the field and we restrict to a class of Hadamard states for which $|(:\Phi^2:\gamma)_\omega| \leq \phi_{\max}^2$.

To discuss averages over long timescales we will use a partition of unity. We define bump functions each supported on $(-\tau_0, \tau_0)$ and thus obtain a sum of integrals, each of which can be bounded by Eq. (13). Then for even numbers of spacetime dimensions ($m = n/2$), we can show that for $f$ supported on $(-\infty, \infty)$ we have

$$\int_{-\infty}^{\infty} d\tau f^2(\tau) R_{\mu\nu} U^\mu U^\nu \geq -Q_m (||f^{(m)}||^2 + \tilde{Q}_m^2 ||f||^2), \tag{14}$$

where the $Q_m$ and $\tilde{Q}_m$ depend on $\phi_{\max}$, $\mu$, and the maximum values of the chosen bump function and its derivatives. This is an inequality of the form of Eq. (2) so it can be used as a condition to a Hawking-type singularity theorem.

## 4. Conclusions

In this work we presented the derivation of lower bounds on EED for the classical and quantum non-minimally coupled scalar field. In the classical case we presented the proof of a Hawking-type singularity theorem for the Einstein-Klein-Gordon field and estimated the required initial contraction to have future timelike geodesic incompleteness. In the quantum case, using the SEE we showed that we can obtain a condition on the curvature to be used as a hypothesis for a singularity theorem.

Apart from estimating the required initial contraction for the quantum case, the obvious extension of this work is the derivation of an absolute (Hadamard renormalised) QSEI for spacetimes with curvature. Another direction is the derivation of a Penrose-type singularity theorem with a condition obeyed by quantum fields. Finally we should note that similar proofs with weakened energy conditions seem to be possible for other classical relativity theorems.

## Acknowledgements

This work is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 744037 "QuEST".

## References


1. A. Raychaudhuri, Relativistic cosmology. 1., *Phys. Rev.* **98**, 1123 (1955).
2. J. Ehlers, Contributions to the relativistic mechanics of continuous media, *Gen. Relativity Gravitation* **25**, 1225 (1993), Translation from German of an article originally published in Akad. Wiss. Lit. Mainz Abh. Math.-Nat. Kl., Nr. 11, 792–837 (1961).
3. J. M. M. Senovilla, Singularity theorems and their consequences, *Gen. Relativity Gravitation* **30**, 701 (1998).
4. S. W. Hawking, The Occurrence of singularities in cosmology, *Proc. Roy. Soc. Lond.* **A294**, 511 (1966).
5. R. Penrose, Gravitational collapse and space-time singularities, *Phys. Rev. Lett.* **14**, 57 (1965).
6. H. Epstein, V. Glaser and A. Jaffe, Nonpositivity of energy density in Quantized field theories, *Nuovo Cim.* **36**, p. 1016 (1965).
7. F. J. Tipler, Energy conditions and spacetime singularities, *Phys. Rev.* **D17**, 2521 (1978).
8. C. Chicone and P. Ehrlich, Line integration of Ricci curvature and conjugate points in Lorentzian and Riemannian manifolds, *Manuscripta Math.* **31**, 297 (1980).
9. A. Borde, Geodesic focusing, energy conditions and singularities, *Classical Quantum Gravity* **4**, 343 (1987).
10. L. H. Ford, Quantum Coherence Effects and the Second Law of Thermodynamics, *Proc. Roy. Soc. Lond.* **A364**, 227 (1978).
11. C. J. Fewster and G. J. Galloway, Singularity theorems from weakened energy conditions, *Class. Quant. Grav.* **28**, p. 125009 (2011).
12. P. J. Brown, C. J. Fewster and E.-A. Kontou, A singularity theorem for einstein–klein–gordon theory, *General Relativity and Gravitation* **50**, p. 121 (2018).
13. S. Hollands and R. M. Wald, Local Wick polynomials and time ordered products of quantum fields in curved space-time, *Commun. Math. Phys.* **223**, 289 (2001).
14. S. Hollands and R. M. Wald, Conservation of the stress tensor in interacting quantum field theory in curved spacetimes, *Rev. Math. Phys.* **17**, 227 (2005).
15. C. J. Fewster, A General worldline quantum inequality, *Class. Quant. Grav.* **17**, 1897 (2000).
16. C. J. Fewster and E.-A. Kontou, Quantum strong energy inequalities, *Phys. Rev.* **D99**, p. 045001 (2019).
17. E.-A. Kontou and K. D. Olum, Quantum inequality in spacetimes with small curvature, *Phys. Rev.* **D91**, p. 104005 (2015).

# Wormholes immersed in rotating matter

Christian Hoffmann[1], Theodora Ioannidou[2], Sarah Kahlen[1], Burkhard Kleihaus[1], Jutta Kunz[1]

[1] *Institut für Physik, Universität Oldenburg, Postfach 2503 D-26111 Oldenburg, Germany*
*E-mail: christian.hoffmann@uni-oldenburg.de; sarah.kahlen@uni-oldenburg.de; b.kleihaus@uni-oldenburg.de; jutta.kunz@uni-oldenburg.de*
[2] *Faculty of Civil Engineering, School of Engineering Aristotle University of Thessaloniki, 54249, Thessaloniki, Greece*
*E-mail: ti3@auth.gr*

We consider Ellis wormholes immersed in rotating matter in the form of an ordinary complex boson field. The resulting wormholes may possess full reflection symmetry with respect to the two asymptotically flat spacetime regions. However, there arise also wormhole solutions where the reflection symmetry is broken. The latter always appear in pairs. We analyse the properties of these rotating wormholes and show that their geometry may feature single throats or double throats. We also discuss the ergoregions and the lightring structure of these wormholes.

*Keywords*: wormholes, boson stars, lightrings

## 1. Introduction

The non-trivial topology of wormholes requires the presence of exotic matter in Einstein's General Relativity (see e.g. the recent review[1] of the field). Choosing a massless phantom (scalar) field for the exotic matter, Ellis[2,3] and Bronnikov[4] found static spherically symmetric wormhole solutions, which connect two asymptotically flat regions of space-time.

Their rotating generalizations were first constructed perturbatively for slow rotation[5,6] and later numerically for rapid rotation[7]. In these wormhole solutions the rotation of the throat and thus the spacetime is achieved by an appropriate choice of the boundary conditions. However, this results in the fact that the two asymptotic regions are rotating with respect to one another. Thus while both asymptotic regions are asymptotically flat, the spacetime is not symmetric with respect to reflection at the throat.

In order to obtain rotating wormholes that exhibit a reflection symmetry at the throat, one can immerse the throat inside rotating matter[8,9]. Then the rotation of the matter will drag the spacetime and thus the throat of the wormhole.

## 2. Wormholes Immersed in Rotating Matter

A nice model to study such wormholes immersed in rotating matter is obtained by adding an ordinary massive complex scalar field to the real phantom scalar field into the action, coupling both to gravity. Without the phantom scalar field the model would yield non-rotating and rotating boson stars. The combination of both scalar fields then allows for rotating wormholes immersed in scalar matter, that are

reflection symmetric. Moreover, a new type of non-symmetric wormholes emerges. In the following we will briefly discuss the model and present its rotating wormhole solutions and analyze their properties[8–11].

### 2.1. *Theoretical Setting*

We consider the action $S$

$$S = \int \left[\frac{1}{2\kappa}\mathcal{R} + \mathcal{L}_{\rm bs} + \mathcal{L}_{\rm ph}\right] \sqrt{-g}\; d^4x \tag{1}$$

with the Einstein-Hilbert term, the Lagrangian $\mathcal{L}_{\rm bs}$ of the complex scalar field $\Phi$

$$\mathcal{L}_{\rm bs} = -\frac{1}{2} g^{\mu\nu} \left(\partial_\mu \Phi^* \partial_\nu \Phi + \partial_\nu \Phi^* \partial_\mu \Phi\right) - m_{\rm bs}^2 |\Phi|^2 \ , \tag{2}$$

and the Lagrangian $\mathcal{L}_{\rm ph}$ of the phantom field $\Psi$,

$$\mathcal{L}_{\rm ph} = \frac{1}{2} \partial_\mu \Psi \partial^\mu \Psi \ . \tag{3}$$

The field equations then consist of the Einstein equations

$$G_{\mu\nu} = \mathcal{R}_{\mu\nu} - \frac{1}{2} g_{\mu\nu} \mathcal{R} = \kappa T_{\mu\nu} \tag{4}$$

and the matter field equations

$$\nabla^\mu \nabla_\mu \Psi = 0 \tag{5}$$

and

$$\nabla^\mu \nabla_\mu \Phi = m_{\rm bs}^2 \Phi \ . \tag{6}$$

An appropriate Ansatz for the metric is given by

$$ds^2 = -e^f dt^2 + e^{q-f} \left[e^b (d\eta^2 + h d\theta^2) + h \sin^2\theta (d\varphi - \omega dt)^2\right] \ , \tag{7}$$

where $f$, $q$, $b$ and $\omega$ are functions of $\eta$ and $\theta$, $h = \eta^2 + \eta_0^2$ with throat parameter $\eta_0$, and $\eta$ takes positive and negative values, $-\infty < \eta < \infty$. The ansatz for the rotating bosonic matter is taken as for boson stars

$$\Phi(t, \eta, \theta, \varphi) = \phi(\eta, \theta)\; e^{i\omega_s t + i n \varphi} \ , \tag{8}$$

where $\phi(\eta, \theta)$ is a real function, $\omega_s$ is the boson frequency, $n$ is a rotational quantum number, and the ansatz for the phantom field $\Psi$ is simply

$$\Psi(t, \eta, \theta, \varphi) = \psi(\eta, \theta) \ . \tag{9}$$

The resulting set of six coupled non-linear partial differential equations is then solved numerically subject to an appropriate set of boundary conditions in the two asymptotic regions $\eta \to \pm\infty$, on the axis of rotation $\theta = 0$, and in the equatorial plane $\theta = \pi/2$.

### 2.2. *Symmetric Wormholes*

Let us now discuss the properties of the symmetric wormholes immersed in bosonic matter. In Fig.1a and b we show their mass $M$ and particle number $Q$ versus the boson frequency $\omega_s$ for a typical set of such wormholes ($\eta_0 = 1$, $n = 0, 1, 2$). Their angular momentum $J$ is given by $J = nQ$.

Clearly, the domain of existence is limited by a maximal value $\omega_{\max} = m_b$, where a vacuum configuration with $M = 0 = Q$ is reached, analogous to boson stars. For large values of $\omega_s$ the global charges of the wormholes follow those of boson stars (see the thin black lines in the figures). However, for small $\omega_s$ the spiralling behaviour present in boson stars is basically lost. In fact, the would-be spirals unwind with respect to the frequency and continue to lower frequencies (possibly reaching a singular configuration[10]).

We emphasize that these solutions satisfy the same boundary conditions in both asymptotic regions. Thus in the case of rotation, it is the complex scalar field with its finite rotational quantum number $n$ that imposes the rotation on the configuration. Then the rotation of the scalar field is reflected in the rotation of the

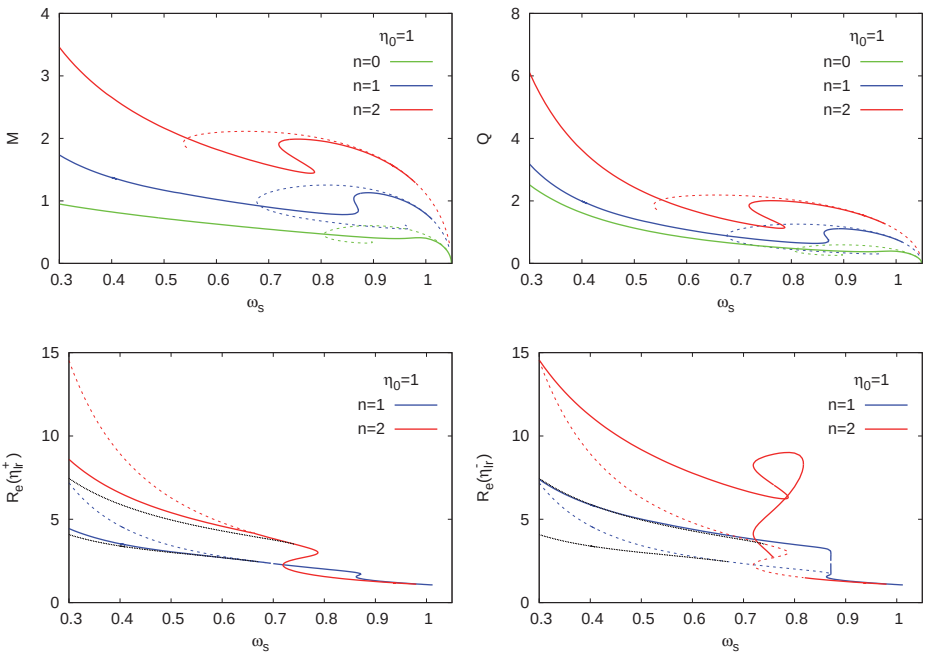



Fig. 1. Properties of symmetric wormhole solutions (throat parameter $\eta_0 = 1$, rotational quantum numbers $n = 0, 1, 2$) versus the boson frequency $\omega_s$: (a) the mass $M$; (b) the particle number $Q$; the dashed lines indicate the respective boson star solutions; (c) the corotation lightring circumferential radius $R_e(\eta_{lr}^+)$; (d) the counterrotation lightring circumferential radius $R_e(\eta_{lr}^-)$; also shown are the circumferential radii of the ergosurfaces (black lines).

spacetime, leading to a rotating throat and frame dragging. Not too surprisingly therefore a sufficiently fast rotation will lead to ergoregions in the wormhole spacetimes. The circumferential radii of the ergoregions are exhibited by the black lines in Fig. 1c and d.

When we consider the geometry of the wormhole solutions we realize that there arises a transition from ordinary single throat wormholes to double throat wormholes with an equator in between, as the frequency $\omega_s$ is decreased. At the transition value the throat degenerates to an inflection point, i.e., the circumferential radius at the center $\eta = 0$ has vanishing first and second derivative.

Of interest are also the lightrings of these spacetimes, as exhibited in Fig. 1c and d for corotating and counterrotating photon orbits, respectively. In particular, we show their circumferential radii $R_e(\eta_{lr}^+)$ and $R_e(\eta_{lr}^-)$. We note that a single lightring exists for large $\omega_s$. For smaller $\omega_s$ two more lightrings emerge. One lightring is always located at $\eta = 0$ (i.e., at the throat or equator), and the additional ones are located symmetrically w.r.t. $\eta = 0$. In the $n = 2$ case up to five lightrings of counterrotating massless particles exist.

### 2.3. *Asymmetric Wormholes*

Let us now focus on another new aspect, namely the presence of symmetric and asymmetric wormhole solutions immersed in bosonic matter. Whereas the field equations are symmetric with respect to reflection of the radial coordinate at the center, and the same boundary conditions are employed in both asymptotically flat regions, the solutions may, however, still be either symmetric or asymmetric with respect to such a reflection. It is the non-trivial topology, which allows for asymmetric solutions, as well.

Starting again the discussion with the global charges, we exhibit the mass and the particle number of these asymmetric solutions in Fig. 2, and compare to the

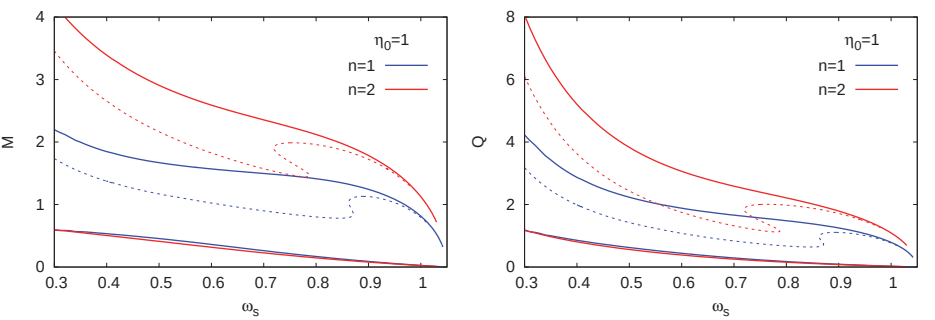



Fig. 2. Properties of asymmetric wormhole solutions (throat parameter $\eta_0 = 1$, rotational quantum numbers $n = 1, 2$) versus the boson frequency $\omega_s$: (a) the mass $M$; (b) the particle number $Q$; the dashed lines indicate the respective symmetric solutions.

corresponding symmetric solutions. Because of the asymmetry the boson field is different in both regions of the spacetime, resulting in different global charges, read off asymptotically. Thus for a given $n$ there is one curve for the symmetric wormholes, but there are two curves for the asymmetric wormholes. Note, that for these asymmetric solutions the angular momentum no longer satisfies $J = nQ$.

As in the case of the symmetric wormholes, there arise double throat wormholes, but now the equator will not reside at $\eta = 0$, but instead arise somewhere in one of the regions. Of course, for each asymmetric solution there exists a second solution, obtained for $\eta \to -\eta$. The asymmetric wormholes may also possess ergoregions and multiple lightrings.

## 3. Conclusions and Outlook

We have obtained a new type of rotating wormhole by immersing the throat in rotating bosonic matter, adopting some features from boson stars. We have studied various physical properties of these solutions, like their global charges, their ergoregions, and their lightrings. To conclude, let us illustrate these new wormhole solutions immersed in rotating matter, via embeddings of a symmetric and an asymmetric double throat wormhole in Fig. 3.

Fig. 3. Embeddings of symmetric (a) and asymmetric (b) wormhole solutions.

All these wormholes are based on General Relativity and therefore need exotic matter for their existence. It will be interesting to consider this new type of wormhole solutions in generalized theories of gravity, which allow for wormholes without the need for exotic matter. Another point of interest will be the study of the stability of these solutions.

## References


1. F. S. N. Lobo, Fundam. Theor. Phys. **189**, pp. (2017).
2. H. G. Ellis, J. Math. Phys. **14**, 104 (1973).
3. H. G. Ellis, Gen. Rel. Grav. **10**, 105 (1979).
4. K. A. Bronnikov, Acta Phys. Polon. **B4**, 251 (1973).
5. P. E. Kashargin and S. V. Sushkov, Grav. Cosmol. **14**, 80 (2008).
6. P. E. Kashargin and S. V. Sushkov, Phys. Rev. D **78**, 064071 (2008).
7. B. Kleihaus and J. Kunz, Phys. Rev. D **90**, 121503 (2014)
8. C. Hoffmann, T. Ioannidou, S. Kahlen, B. Kleihaus and J. Kunz, Phys. Lett. B **778**, 161 (2018)
9. C. Hoffmann, T. Ioannidou, S. Kahlen, B. Kleihaus and J. Kunz, Phys. Rev. D **97**, no. 12, 124019 (2018)
10. V. Dzhunushaliev, V. Folomeev, C. Hoffmann, B. Kleihaus and J. Kunz, Phys. Rev. D **90**, no. 12, 124038 (2014)
11. C. Hoffmann, T. Ioannidou, S. Kahlen, B. Kleihaus and J. Kunz, Phys. Rev. D **95**, no. 8, 084010 (2017)

# Wormholes, energy conditions and time machines

Francisco S. N. Lobo[1,*] and Diego Rubiera-Garcia[2,**]

[1] *Instituto de Astrofísica e Ciências do Espaço, Faculdade de Ciências da Universidade de Lisboa, Edifício C8, Campo Grande, P-1749-016 Lisbon, Portugal*
[2] *Departamento de Física Teórica and IPARCOS, Universidad Complutense de Madrid, E-28040 Madrid, Spain*
[*] *E-mail: fslobo@fc.ul.pt,* [**] *E-mail: drubiera@ucm.es*

This report is based on the Parallel Session AT3 "Wormholes, Energy Conditions and Time Machines" of the Fifteenth Marcel Grossmann Meeting - MG15, held at the University of Rome "La Sapienza" – Rome, in 2018.

*Keywords*: Wormholes; energy conditions; time machines.

## 1. Introduction

Traditionally, the solutions of General Relativity (GR) are obtained by first considering a plausible distribution of mass and energy, described by a stress-energy tensor, and through the Einstein field equations (EFE) one obtains the spacetime metric. However, one may adopt the reverse philosophy and engineer an interesting metric and then solve the EFE to find the matter sources that sustain the corresponding geometry. In this manner, in wormhole physics, the tunnel-like structure imposes the so-called flaring-out condition at the wormhole throat, the latter being the location of the minimum radius of the geometry. Thus, from this flaring-out condition, and using the EFE, in GR one discovers that the null energy condition (NEC) is violated[1]. In fact, all of the pointwise energy conditions are violated. Another interesting feature of these spacetimes is that they allow "effective" superluminal travel, and consequently generate closed timelike curves[2], although locally the speed of light is not surpassed. One may mention that, in fact, these solutions are primarily useful as "gedanken-experiments" and as a theoretician's probe of the foundations of GR. We refer the reader to the books[3,4] for an extensive review on the subject.

The AT3 Parallel Session, "Wormholes, Energy Conditions and Time Machines", was dedicated to the state-of-art contributions on this fascinating branch of research. In this report, we find it useful to organise the following sections, which reflect the internal schedule that was adopted during the session.

## 2. Traversable wormholes: Mathematical structure and applications

Ever since their revival, three decades ago, in the seminal work of Morris and Thorne[1,2], Lorentzian wormholes in GR have led an uncomfortable existence as they require matter which violates the well known energy conditions. However, in other theories of gravity beyond GR wormholes can indeed exist supported by "standard matter" satisfying the energy conditions[5].

Sayan Kar illustrated this fact with some known examples and also by explicitly constructing a zero Ricci scalar wormhole in a four dimensional scalar-tensor, on-brane gravity theory arising from the two-brane Randall-Sundrum model with one extra dimension[6]. If such a wormhole could arise as the end-state of some astrophysical process, its ringdown may be studied using gravitational waves. With this aim, the scalar quasinormal modes in this class of wormholes were obtained and identified as the 'breathing mode' associated with gravitational waves in scalar-tensor theories. Moreover, using the frequencies and time constants of the breathing mode as well as the results arising out of the GW150914 event, the size of the wormhole parameters using standard tools were estimated. The analysis suggests that it may be possible to constrain wormholes and distinguish them from black holes using this kind of observations.

Jutta Kunz considered Ellis-Bronnikov wormholes immersed in rotating matter in the form of an ordinary complex boson field[7,8]. The resulting wormholes may possess full reflection symmetry with respect to the two asymptotically flat spacetime regions. However, wormhole solutions also arise where the reflection symmetry is broken, and in such a case they always appear in pairs. The properties of these rotating wormholes were analysed and it was shown that their geometry may feature single throats or double throats. The ergoregions and the light ring structure of these wormholes were also discussed.

Kirill Bronnikov analysed spherically symmetric configurations in GR, supported by nonlinear electromagnetic fields with gauge-invariant Lagrangians[9] depending on the single invariant $f = F_{\mu\nu}F^{\mu\nu}$. Considering metrics with two independent functions of time, a natural generalization of the class of wormholes previously introduced by Arellano and Lobo with a time-dependent conformal factor was found[10]. Such wormholes are shown to be only possible for some particular choices of the Lagrangian function $L(f)$, having no Maxwell weak-field limit. Their time evolution contains cosmological-type singularities. For general non-Maxwell $L(f)$, instead of the usual electric-magnetic duality there exists the so-called FP duality, connecting theories with different $L(f)$; accordingly, for some of the wormhole solutions naturally emerging with magnetic fields, the electric counterparts with the same metric are ill-defined from the viewpoint of the Lagrangian formalism.

Gonzalo Olmo showed that the field equations of metric-affine theories of gravity whose Lagrangian is a nonlinear function of the Ricci tensor can be written in an Einstein form introducing both an auxiliary metric and auxiliary matter fields[11,12]. In this way the resolution of the equations can be done directly from within the framework of GR coupled to nonlinear matter fields. In other words, the nonlinearities of the gravity sector can be transferred to the matter sector. The correspondence with the original nonlinear gravity theory only involves algebraic manipulations. Standard analytical and numerical methods can thus be used to find solutions of these theories starting from GR ones.

Rajibul Shaikh studied shadows cast, under some circumstances, by a certain class of rotating wormholes, and explored the dependence of the shadows on the wormhole spin[13]. The results were then compared with that of a Kerr black hole. For small spin, the shapes of the shadows cast by a wormhole and a black hole are nearly identical. However, with increasing values of the spin, the shape of a wormhole shadow starts deviating considerably from that of a Kerr black hole. Detection of such a considerable deviation in future observations may possibly indicate the presence of a wormhole. In other words, the results indicate that the wormholes considered in this work, and that have reasonable spin, can be distinguished from a black hole through the observation of their shadows.

João Rosa presented wormhole solutions in the scalar-tensor representation of the generalized hybrid metric-Palatini matter theory, given by a gravitational Lagrangian density $f(R, \mathcal{R})$, where $R$ is the metric Ricci scalar, and $\mathcal{R}$ is a Palatini scalar curvature defined in terms of an independent connection[14]. The main interest in the solutions found is that the matter field obeys the NEC everywhere, including the throat and up to infinity, so that there is no need for exotic matter. The wormhole geometry with its flaring out at the throat is supported by the higher-order curvature terms, which can be interpreted as a gravitational fluid. Thus, in this theory, in building a wormhole, it is possible to exchange the exoticity of matter by the exoticity of the gravitational sector.

Sung-Won Kim presented recent work on the exact solution of the cosmological model, in an expanding Friedmann-Lemaitre-Robertson-Walker universe, from the isotropic form of the Morris-Thorne type wormhole[15]. Matter, radiation, and the cosmological constant $\Lambda$ were considered as the single component of the universe. For the multi-component universe case, the $\Lambda$CDM universe with matter and $\Lambda$ was adopted. In the case of the single-component universe, the apparent horizons coincide at a certain time. However, they have two coincidence times in the multi-component universe, which means that they start at early times and finish at late times like the combination of two single-component universes.

Mikhail Volkov applied duality rotations and complexifications to the vacuum Weyl metrics generated by massive rods or by point masses[16]. As a first step this gives families of prolate and oblate vacuum metrics. Further duality transformations produce a scalar field from the vacuum, which can be either the conventional scalar or the phantom field with negative kinetic energy. This gives rise to large classes of axially symmetric solutions, presumably including all previously known solutions for gravity-coupled massless scalar fields. Particularly interesting are the oblate solutions which, irrespectively of whether they are coupled to a scalar field or not, describe wormholes connecting several asymptotic regions. In the one-wormhole sector they reduce to the ring wormholes in the vacuum case and to the Bronnikov-Ellis wormhole in the phantom one. The two-wormhole solutions were also studied and it was found that two of their four asymptotic regions are completely regular while two others contain an infinitely long strut along the symmetry axis.

Remo Garattini argued that, since Yukawa corrections to Newtonian potentials appear in some modified theories of gravity, it is worth exploring some aspects of traversable wormholes with a shape function having a Yukawa-like profile and related generalizations[17]. These aspects also include the introduction of some equations of state. More specifically, an equation of state was considered in which the sum of the energy density and radial pressure is proportional to a Yukawa profile. In some cases, it was shown that the specific wormhole solution has an asymptotic behaviour corresponding to a global monopole.

Fayçal Hammad presented the behavior of black hole horizons and wormholes under Weyl conformal transformations[18]. First, a shorter, but more general, derivation of the Weyl transformation of the simple prescription for detecting horizons and wormholes given recently in the literature for spherically symmetric spacetimes was provided. Then, the conformal behavior of black hole horizons and wormholes in more general spacetimes, based on more "sophisticated" definitions, was presented. The study showed that black holes and wormholes might always arise in the new frame even if they were absent in the original frame. Moreover, it was shown that the various definitions found in the literature might be transformed into one another under such transformations. Finally, the conformal behavior of the required energy conditions for wormholes was discussed.

Jin Young Kim considered a new approach to construct wormholes without introducing exotic matter in Einstein-Born-Infeld gravity with a cosmological constant[19]. It was argued that the exoticity of the energy-momentum tensor is not essential to sustain the wormhole. The stability of the new wormholes with a scalar perturbation was also investigated, and it was confirmed that the Breitenlohner-Freedman bound holds in Einstein-Born-Infeld gravity.

Elisa Maggio discussed that gravitational wave astronomy can give us access to the structure of black holes, and to potentially probing microscopic corrections at the horizon scale[20]. A general model of a exotic compact object consisting of a Kerr geometry with a reflective surface near the horizon was investigated. This framework can be applied to thin-shell wormholes when the surface is perfectly reflecting. The stability of these geometries under scalar and electromagnetic perturbations was analysed. It was shown that exotic compact objects with a perfectly reflecting surface are affected by the an ergoregion instability when spinning sufficiently fast. This instability might have a crucial impact on their phenomenology. On the other hand, it was found that a partial absorption at the surface is sufficient to quench the ergoregion instability completely. This finding has important consequences for the viability of exotic compact objects and it suggests that they are not necessarily ruled out by the ergoregion instability.

## 3. Energy conditions

Eleni-Alexandra Kontou presented a theme related to the strong quantum energy inequality and the Hawking singularity theorem[21]. The latter concerns matter

obeying the strong energy condition (SEC), which means that all observers experience a nonnegative effective energy density thereby guaranteeing the timelike convergence property. However, for both classical and quantum fields, violations of the SEC can be observed in some of the simplest cases, for example, in the massive Klein-Gordon field. Therefore, there is a need to develop theorems with weaker restrictions, namely, energy conditions averaged over an entire geodesic and quantum inequalities, weighted local averages of energy densities. In fact, lower bounds of the effective energy density were derived in the presence of both classical and quantum scalar fields allowing a nonzero mass and nonminimal coupling to the scalar curvature. In the quantum case these bounds take the form of a set of state-dependent quantum energy inequalities valid for the class of Hadamard states. Finally, it was discussed how the classical and quantum inequalities derived can be used as an assumption to a modified Hawking-type singularity theorem.

## 4. Causal structure of spacetime

Sandipan Sengupta presented, in classical gravity theory, explicit examples of vacuum solutions that admit the possibility of time travel (to the past) through their geodesics[22]. These geometries are built upon metrics whose determinant can continuously go to zero over some extended region of the spacetime. These solutions to the first order field equations satisfy the energy conditions. One may see the existence of these solutions as a motivation to revisit the status of causality in the formulation of classical gravity.

Vaishak Prasad presented the possibility of resolving the chronology protection problem due to the existence of closed time-like curves in Kerr-Newman black holes using modified gravity[23]. First, the details of the causality violation in the Kerr-Newman spacetime were revisited and quantified. It was then shown that the issue also extends onto two of the modified Kerr-Newman solutions: non-commutative geometry inspired and $f(R)$ gravity solutions. The geodesic connectivity of the causality violating region was discussed in both scenarios and the existence of null geodesics was proved. The possibility of preventing causality violation was also explored. It was shown that although in both models the parameters can be chosen such that the causality violating region is removed, the $f(R)$ case is not consistent with cosmological observations. In the case of the non-commutative geometry inspired solution, it was shown that the chronology protection can be ensured by choosing suitable values for the non-commutative parameter, thereby eliminating the causality violating region and the Cauchy horizon.

Ilia Musco showed that in the gravitational collapse to form black holes, trapping horizons (foliated by marginally trapped surfaces) make their first appearance either within the collapsing matter or where it joins on to a vacuum exterior. Those which then move outwards with respect to the matter have been proposed for use in defining black holes, replacing the global concept of an event horizon, which has some serious drawbacks for practical applications. Results were presented from a

study of the properties of both outgoing and ingoing trapping horizons, assuming strict spherical symmetry throughout [24]. The causal nature was investigated following two different approaches, namely, one using a geometrical quantity related to expansions of null geodesic congruences, and the other using the horizon velocity measured with respect to the collapsing matter. The models treated are simplified, but do include pressure effects in a meaningful way, and the behavior of the horizon evolution was analysed depending on the initial conditions of energy density and pressure of the collapse.

## References


1. M. S. Morris and K. S. Thorne, Am. J. Phys. **56**, 395 (1988).
2. M. S. Morris, K. S. Thorne and U. Yurtsever, Phys. Rev. Lett. **61**, 1446 (1988).
3. M. Visser, "Lorentzian wormholes: From Einstein to Hawking", AIP press, New York (1995).
4. F. S. N. Lobo, "Wormholes, Warp Drives and Energy Conditions", Fundam. Theor. Phys. **189**, Springer, Switzerland (2017)
5. T. Harko, F. S. N. Lobo, M. K. Mak and S. V. Sushkov, Phys. Rev. D **87**, 067504 (2013).
6. S. Aneesh, S. Bose and S. Kar, Phys. Rev. D **97**, 124004 (2018).
7. C. Hoffmann, T. Ioannidou, S. Kahlen, B. Kleihaus and J. Kunz, Phys. Lett. B **778**, 161 (2018).
8. C. Hoffmann, T. Ioannidou, S. Kahlen, B. Kleihaus and J. Kunz, Phys. Rev. D **97**, 124019 (2018).
9. K. A. Bronnikov, Int. J. Mod. Phys. D **27**, 1841005 (2018).
10. A. V. B. Arellano and F. S. N. Lobo, Class. Quant. Grav. **23**, 5811 (2006).
11. V. I. Afonso, G. J. Olmo, E. Orazi and D. Rubiera-Garcia, Eur. Phys. J. C **78**, 866 (2018).
12. V. I. Afonso, G. J. Olmo, E. Orazi and D. Rubiera-Garcia, Phys. Rev. D **99**, 044040 (2019).
13. R. Shaikh, Phys. Rev. D **98**, 024044 (2018).
14. J. L. Rosa, J. P. S. Lemos and F. S. N. Lobo, Phys. Rev. D **98**, 064054 (2018).
15. S. W. Kim, arXiv:1811.07164 [gr-qc].
16. G. W. Gibbons and M. S. Volkov, JCAP **1705**, 039 (2017).
17. R. Garattini, arXiv:1907.03622 [gr-qc].
18. F. Hammad, Phys. Rev. D **97**, 124015 (2018).
19. J. Y. Kim and M. I. Park, Eur. Phys. J. C **76**, 621 (2016).
20. E. Maggio, V. Cardoso, S. R. Dolan and P. Pani, Phys. Rev. D **99**, 064007 (2019).
21. P. J. Brown, C. J. Fewster and E. A. Kontou, Gen. Rel. Grav. **50**, 121 (2018).
22. S. Sengupta, Phys. Rev. D **97**, 124038 (2018).
23. V. Prasad, R. Srinivasan and S. Gutti, Phys. Rev. D **99**, 024023 (2019).
24. A. Helou, I. Musco and J. C. Miller, Class. Quant. Grav. **34**, 135012 (2017).

# Ergoregion instability of exotic compact objects

Elisa Maggio[1,2,*]

[1] *Dipartimento di Fisica, "Sapienza" Università di Roma & Sezione INFN Roma1, Piazzale Aldo Moro 5, 00185, Roma, Italy*

[2] *Consortium for Fundamental Physics, School of Mathematics and Statistics, University of Sheffield, Hicks Building, Hounsfield Road, Sheffield S3 7RH, United Kingdom*

[*] *E-mail: elisa.maggio@roma1.infn.it*

Gravitational waves are probes of gravity in the strong-field regime that allow to understand the nature of compact objects. Several families of exotic compact objects have been conceived in order to overcome some paradoxes associated to black holes, particularly the existence of the event horizon. Some models of exotic compact objects are characterized by microscopic or even Planckian corrections at the horizon scale. In particular, the wormhole solution has a perfectly reflecting surface instead of the event horizon. Spinning horizonless compact objects may be unstable against an "ergoregion instability" when spinning sufficiently fast. We analyse the instability and its astrophysical impact for the viability of exotic compact objects.

## 1. Introduction

Recently, there has been a growing interest in investigating the nature of compact objects and testing the existence of horizons.[1,2] Following the recent detections of gravitational waves (GWs) from compact binary coalescences,[3] it is reasonable to enquire whether the compact remnant is compatible with the predictions of General Relativity. Several quantum-gravity models of exotic compact objects (ECOs) have been proposed as alternatives to black holes (BHs).[4–6] ECOs are horizonless compact objects which can overcome some paradoxes associated to BHs, such as their thermodynamical instability and their entropy which is far in eccess of a typical stellar progenitor.[7] ECO models have a unique fingerprint in the postmerger ringdown phase of a GW signal that allow to distinguish them from canonical BH solutions.[1,8]

In the following, we analyse the wormhole model as a horizonless compact object supported by an exotic distribution of matter.[9] Spinning wormhole solutions are prone to the so-called *ergoregion instability* when spinning sufficiently fast.[10] The latter is an instability that develops in any spacetime featuring an ergoregion but without an event horizon in the absence of dissipation mechanisms. The ergoregion instability could have a crucial impact on the dynamics of the ECO and provide a strong theoretical argument in favour of the BH paradigm.

## 2. Gravitational waves as probes of strong gravity

The recent detections of GWs provide the opportunity to analyse the two-body motion in the strong-field and large-velocity regime of gravity.[11] The detected GW signals are characterized by three phases: (i) the inspiral, when the two bodies have large separations and the GW emission decreases the binary's orbital period; (ii) the

merger, when the two bodies coalesce; (iii) the ringdown, when the final remnant relaxates to a stationary and equilibrium solution.

The ringdown stage is dominated by a superposition of the characteristic frequencies of the remnant, the so-called *quasi-normal modes* (QNMs). The latter are complex frequencies, $\omega = \omega_R + i\omega_I$, which describe the response of the compact object to a perturbation.[12] The emitted ringdown signal is a sum of exponentially damped sinusoids whose frequencies are proportional to the real part of the QNMs ($f_{\rm GW} = \omega_R/2\pi$) and whose damping times are proportional to the imaginary part of the QNMs ($\tau_{\rm damp} = -1/\omega_I$.)[13] According to our convention, a mode with $\omega_I < 0$ describes a stable remnant, whereas a mode with $\omega_I > 0$ describes an unstable one.

From the detection of the ringdown stage of a binary coalescence, the QNMs of the remnant can be inferred, leading to an understanding of the nature of the remnant. Solely in the first event GW150914, the least-damped QNM of the remnant has been observed.[14] The latter is consistent with a Kerr BH of mass $62^{+4}_{-4}\ M_\odot$ and spin $0.67^{+0.05}_{-0.07}$ at 90% confidence.[15] However alternatives to BHs are not excluded and the characterization of the remnant is still an open problem.[16–18] A precise modeling of the gravitational waveform in a variety of strong gravity processes is necessary, including the signal emitted by alternatives to BHs.

## 3. Exotic compact objects

Several arguments have recently been put forward, suggesting that new physics at the horizon scale might alter the formation of BHs during the gravitational collapse.[5,19,20] In this context, several families of ECOs have been proposed as near-horizon quantum structures.[4,7,21] ECOs are theoretical compact objects that can mimic the features of BHs by means of electromagnetic observations since they can be as compact as BHs. However, ECOs are not endowed with an event horizon nor a curvature singularity.

In the following, we consider rotating solutions described by the Kerr metric and characterized by microscopic corrections at the horizon scale.[22,23] The surface of the compact object is located at $r_0 = r_+(1+\epsilon)$, where $r_+$ is the location of the would-be horizon and $0 < \epsilon \ll 1$ is a parameter related to the compactness of the object. The properties of the object's interior are parametrized in terms of a surface reflection coefficient $\mathcal{R}$.[24] A perfectly reflecting surface is described by $|\mathcal{R}|^2 = 1$, whereas a partially absorbing surface (e.g. due to dissipation, viscosity, fluid mode excitations) is described by $|\mathcal{R}|^2 < 1$.

### 3.1. *Wormholes*

In the static limit, the wormhole solution can be constructed by taking two copies of the Schwarzschild metric with $r > r_0$ and by pasting them at $r = r_0$, which is the location of the wormhole throat. This procedure is called *Schwarzschild surgery* and is analogous to considering one Schwarzschild metric with $r > r_0$ and a perfectly reflecting surface at $r = r_0$.[9]

According to Birkhoff's theorem, the Schwarzschild metric is the only solution of Einstein's field equations in vacuum which describes a spherically symmetric spacetime. As a consequence, the wormhole solution must be sustained by a distribution of matter. In this case, a thin shell of matter is located at the wormhole throat.[25] Furthermore, the distribution is exotic since it violates the weak energy condition according to which the surface energy density is negative.[25]

## 4. Ergoregion instability

Spinning wormhole solutions might turn unstable above a critical threshold of the spin due to the *ergoregion instability*.[10] Since physical negative-energy states can exist inside the ergoregion – which is key ingredient of the Penrose's process–[26] it is energetically favorable to cascade toward more negative states. The only way to prevent such an infinite cascade is by absorbing negative-energy states. For this reason, Kerr BHs are stable against the ergoregion instability. Whereas, ECOs suffer from the instability in the absence of dissipation mechanisms.

### 4.1. *Linear perturbations*

In order to analyse the stability of ECOs, let us consider a massless scalar field governed by the Klein-Gordon equation in Kerr geometry. By decomposing the scalar field in Boyer-Lindquist coordinates, the Klein-Gordon equation can be separated in terms of Teukolsky's master equations.[27]

By imposing physically motivated boundary conditions at infinity and at the surface of the ECO, the equations define an eigenvalue problem whose (complex) eigenvalues are the QNMs of the system. In particular, at infinity the proper vibrations of the object require only outgoing waves.[28] In the the BH case, at $r = r_+$ the presence of the horizon would require purely ingoing waves. Whereas, in the ECO case, at $r = r_0$ there is a superposition of ingoing and outgoing waves such that the flux of energy is conserved.[22,23]

### 4.2. *Quasi-normal mode spectrum*

In the static case, the QNMs of the wormhole are stable since $\omega_I < 0$. As discussed in Ref. 8, in the BH limit (as $\epsilon \to 0$) the QNM frequencies approaches zero and the damping time becomes infinite, thus the modes become extremely long lived. The fact that the imaginary part of the modes vanishes as $\epsilon \to 0$ suggests that the modes might turn unstable in the spinning case, due to the Zeeman splitting of the frequencies with different azimuthal numbers. This expectation is confirmed by our results for an ECO with a perfectly reflecting surface. In particular, the QNM frequencies have a zero crossing at a critical value of the spin where the ECO turns from stable into unstable against the ergoregion instability.[22]

### 4.3. *Astrophysical impact of the instability*

The onset of the ergoregion instability depends on the compactness of the ECO. In particular, in the BH limit ($\epsilon \to 0$) the critical value of the spin approaches zero. Thus, even slowly spinning ECOs suffer from the ergoregion instability.

The instability time scale is extremely short compared to typical astrophysical time scales. For example, for a stellar mass ECO the instability occurs after few seconds the formation of the object.[22] Thus, for a perfectly reflecting surface the ergoregion instability plays an important role in the dynamics of the ECO.

### 4.4. *The role of absorption*

The case of a perfectly reflecting surface is an idealization and it is expected that the compact object would absorb part of the ingoing radiation, e.g., through viscosity, dissipation. In this case, at $r = r_0$ there is a superposition of ingoing and outgoing waves such that a certain fraction of the incident flux of energy is absorbed at the surface of the ECO.

By deriving the QNMs of the compact object, we obtain that an absorption rate of 0.4% is sufficient to quench the ergoregion instability completely.[22] Thus, any ECO with at least this level of absorption is stable under scalar field perturbations.

An indicative estimate shows that this level of absorption can be naturally achieved by a viscous fluid in the ECO. The dissipation can occur through sound waves which propagate into the interior of the fluid and through shear waves which heat the surface of the ECO.[29,30]

## 5. Conclusions and future prospectives

In the newly-born era of GWs, we have the first opportunity to look for new physics at the horizon scale. We analysed the stability of an ECO described by the Kerr metric whose surface is at a Planckian distance from the would-be event horizon. We analysed a test scalar field perturbation and we showed two interesting properties:

i. An ECO with a perfectly reflecting surface suffers from the ergoregion instability above a critical value of the spin. The instability time scale is short compared to astrophysical time scales.
ii. An absorption rate of 0.4% at the surface of the ECO is sufficient to quench the ergoregion instability completely.

An extension of this work is the analysis of gravitational perturbations. It has been recently computed in Ref. 23 that an absorption rate of 0.3% (6%) is necessary in order to have a stable ECO with spin $\chi \lesssim 0.7$ ($\chi \lesssim 0.9$) for any perturbation.

Finally, an understanding of the evolution and the end point of the instability – including possible nonlinearities – remain interesting open problems.

**Acknowledgments**

E.M. acknowledge networking support by the COST Action CA16104. E.M. acknowledge the financial support provided under the European Union's H2020 ERC, Starting Grant agreement no. DarkGRA–757480. E.M. acknowledge support from the Amaldi Research Center funded by the MIUR program "Dipartimento di Eccellenza" (CUP: B81I18001170001).

**References**

1. V. Cardoso, S. Hopper, C. F. B. Macedo, C. Palenzuela, and P. Pani, "Gravitational-wave signatures of exotic compact objects and of quantum corrections at the horizon scale," *Phys. Rev.*, vol. D94, no. 8, p. 084031, 2016.
2. V. Cardoso and P. Pani, "Tests for the existence of horizons through gravitational wave echoes," *Nat. Astron.*, vol. 1, pp. 586–591, 2017.
3. B. P. Abbott *et al.*, "GWTC-1: A Gravitational-Wave Transient Catalog of Compact Binary Mergers Observed by LIGO and Virgo during the First and Second Observing Runs," 2018.
4. S. D. Mathur, "Fuzzballs and the information paradox: A Summary and conjectures," 2008.
5. S. B. Giddings, "Possible observational windows for quantum effects from black holes," *Phys. Rev.*, vol. D90, no. 12, p. 124033, 2014.
6. B. Holdom and J. Ren, "Not quite a black hole," *Phys. Rev.*, vol. D95, no. 8, p. 084034, 2017.
7. P. O. Mazur and E. Mottola, "Gravitational vacuum condensate stars," *Proc. Nat. Acad. Sci.*, vol. 101, pp. 9545–9550, 2004.
8. V. Cardoso, E. Franzin, and P. Pani, "Is the gravitational-wave ringdown a probe of the event horizon?," *Phys. Rev. Lett.*, vol. 116, no. 17, p. 171101, 2016. [Erratum: Phys. Rev. Lett.117,no.8,089902(2016)].
9. M. Visser, *Lorentzian wormholes: From Einstein to Hawking*. Woodbury, USA: AIP, 1996.
10. J. L. Friedman, "Ergosphere instability," *Communications in Mathematical Physics*, vol. 63, pp. 243–255, Oct. 1978.
11. B. P. Abbott *et al.*, "Binary Black Hole Mergers in the first Advanced LIGO Observing Run," *Phys. Rev.*, vol. X6, no. 4, p. 041015, 2016.
12. E. Berti, V. Cardoso, and A. O. Starinets, "Quasinormal modes of black holes and black branes," *Class. Quantum Grav.*, vol. 26, p. 163001, 2009.
13. B. P. Abbott *et al.*, "The basic physics of the binary black hole merger GW150914," *Annalen Phys.*, vol. 529, no. 1-2, p. 1600209, 2017.
14. B. P. Abbott *et al.*, "Tests of general relativity with GW150914," *Phys. Rev. Lett.*, vol. 116, no. 22, p. 221101, 2016.
15. B. P. Abbott *et al.*, "Observation of gravitational waves from a binary black hole merger," *Phys. Rev. Lett.*, vol. 116, p. 061102, Feb 2016.

16. J. Abedi, H. Dykaar, and N. Afshordi, "Echoes from the Abyss: Tentative evidence for Planck-scale structure at black hole horizons," *Phys. Rev.*, vol. D96, no. 8, p. 082004, 2017.
17. R. S. Conklin, B. Holdom, and J. Ren, "Gravitational wave echoes through new windows," *Phys. Rev.*, vol. D98, no. 4, p. 044021, 2018.
18. J. Westerweck, A. Nielsen, O. Fischer-Birnholtz, M. Cabero, C. Capano, T. Dent, B. Krishnan, G. Meadors, and A. H. Nitz, "Low significance of evidence for black hole echoes in gravitational wave data," *Phys. Rev.*, vol. D97, no. 12, p. 124037, 2018.
19. R. Brustein, A. J. M. Medved, and K. Yagi, "Discovering the interior of black holes," *Phys. Rev.*, vol. D96, no. 12, p. 124021, 2017.
20. G. F. Giudice, M. McCullough, and A. Urbano, "Hunting for Dark Particles with Gravitational Waves," *JCAP*, vol. 1610, no. 10, p. 001, 2016.
21. K. Skenderis and M. Taylor, "The fuzzball proposal for black holes," *Phys. Rept.*, vol. 467, pp. 117–171, 2008.
22. E. Maggio, P. Pani, and V. Ferrari, "Exotic Compact Objects and How to Quench their Ergoregion Instability," *Phys. Rev.*, vol. D96, no. 10, p. 104047, 2017.
23. E. Maggio, V. Cardoso, S. R. Dolan, and P. Pani, "Ergoregion instability of exotic compact objects: electromagnetic and gravitational perturbations and the role of absorption," *Phys. Rev.*, vol. D99, no. 6, p. 064007, 2019.
24. Z. Mark, A. Zimmerman, S. M. Du, and Y. Chen, "A recipe for echoes from exotic compact objects," *Phys. Rev.*, vol. D96, no. 8, p. 084002, 2017.
25. W. Israel, "Singular hypersurfaces and thin shells in general relativity," *Nuovo Cim.*, vol. B44S10, p. 1, 1966. [Nuovo Cim.B44,1(1966)].
26. R. Penrose, "Gravitational collapse: The role of general relativity," *Riv. Nuovo Cim.*, vol. 1, pp. 252–276, 1969. [Gen. Rel. Grav.34,1141(2002)].
27. S. A. Teukolsky, "Rotating black holes - separable wave equations for gravitational and electromagnetic perturbations," *Phys. Rev. Lett.*, vol. 29, pp. 1114–1118, 1972.
28. S. A. Teukolsky and W. H. Press, "Perturbations of a rotating black hole. III - Interaction of the hole with gravitational and electromagnet ic radiation," *Astrophys. J.*, vol. 193, pp. 443–461, 1974.
29. F. P. Esposito, "Absorption of Gravitational Energy by a Viscous Compressible Fluid," *The Astrophys. J.*, vol. 165, p. 165, apr 1971.
30. D. Papadopoulos and F. P. Esposito, "Absorption of gravitational energy by a viscous compressible fluid in a curved spacetime," *The Astrophys. J.*, vol. 292, pp. 330–338, May 1985.

# Wormhole solutions in generalized hybrid metric-Palatini gravity

João Luís Rosa* and José P. S. Lemos**

*Centro de Astrofísica e Gravitação - CENTRA, Departamento de Física, Instituto Superior Técnico - IST, Universidade de Lisboa - UL, Avenida Rovisco Pais 1, 1049-001, Portugal*
**E-mail: joaoluis92@gmail.com*
***E-mail: joselemos@ist.utl.pt*

Francisco S.N. Lobo

*Instituto de Astrofísica e Ciências do Espaço,*
*Faculdade de Ciências da Universidade de Lisboa,*
*Edifício C8, Campo Grande, P-1749-016 Lisbon, Portugal*
*E-mail: fslobo@fc.ul.pt*

Wormhole solutions in a generalized hybrid metric-Palatini matter theory, given by a gravitational Lagrangian density $f(R, \mathcal{R})$, where $R$ is the metric Ricci scalar, and $\mathcal{R}$ is a Palatini scalar curvature defined in terms of an independent connection, and a matter Lagrangian, are found. The solutions are worked out in the scalar-tensor representation of the theory, where the Palatini field is traded for two scalars, $\varphi$ and $\psi$, and the gravitational term $R$ is maintained. The main interest in the solutions found is that the matter field obeys the null energy conditions everywhere, including the throat and up to infinity, so that there is no need for exotic matter. The wormhole geometry with its flaring out at the throat is supported by the higher-order curvature terms, or equivalently, by the two fundamental scalar fields, which either way can be interpreted as a gravitational fluid. Thus, in this theory, in building a wormhole, it is possible to exchange the exoticity of matter by the exoticity of the gravitational sector.

*Keywords*: Wormholes; modified gravity; hybrid metric-Palatini gravity.

## 1. Introduction

The flaring-out condition is the fundamental ingredient in wormhole physics[1–5], and in general relativity this geometric condition entails the violation of the null energy condition (NEC). The latter states that $T_{ab}\, k^a k^b \geq 0$, where $T_{ab}$ is the matter stress-energy tensor and $k^a$ is any null vector. Matter that violates the NEC is denoted as exotic matter. In fact, in the context of modified theories of gravity[6], it has been shown in principle that normal matter may thread the wormhole throat[7], and it is the higher-order curvature terms, which may be interpreted as a gravitational fluid, that support these nonstandard wormhole geometries. Indeed, it was shown explicitly that in $f(R)$ theories[8] and extensions[9–11], wormhole throats can be theoretically constructed without the presence of exotic matter, where the throat is solely sustained by the geometric fields presented in modified gravity.

The main aim of this work is to theoretically construct wormhole solutions in generalized hybrid metric-Palatini gravity[12–14], where the matter obeys the NEC throughout the geometry. In the scalar representation of generalized hybrid metric-Palatini gravity we are studying, the gravitational field is complemented by the two other fundamental scalar fields[14,15], whereas the matter fields in this theory are still encoded in the stress-energy tensor $T_{ab}$. Thus, to allow for a mandatory

flaring-out condition, the matter $T_{ab}$ can obey the NEC, as long as some appropriate combination of $T_{ab}$ with the other fundamental fields violates the NEC. We will make use of the junction conditions for the generalized hybrid theory[16].

## 2. Scalar-tensor representation of the generalized hybrid metric-Palatini gravity

Consider the general hybrid action $S$ given by

$$S=\frac{1}{2\kappa^2}\int\sqrt{-g}f\left(R,\mathcal{R}\right)d^4x+\int\sqrt{-g}\,\mathcal{L}_m d^4x, \tag{1}$$

where $\kappa^2\equiv 8\pi G$, $g$ is the determinant of the spacetime metric $g_{ab}$, $R$ is the metric Ricci scalar, $\mathcal{R}\equiv\mathcal{R}^{ab}g_{ab}$ is the Palatini Ricci scalar, which is defined in terms of an independent connection $\hat{\Gamma}^c_{ab}$ as, $\mathcal{R}_{ab}=\partial_c\hat{\Gamma}^c_{ab}-\partial_b\hat{\Gamma}^c_{ac}+\hat{\Gamma}^c_{cd}\hat{\Gamma}^d_{ab}-\hat{\Gamma}^c_{ad}\hat{\Gamma}^d_{cb}$, $\partial_a$ denotes partial derivative; $f\left(R,\mathcal{R}\right)$ is a well-behaved function of $R$ and $\mathcal{R}$, and $\mathcal{L}_m$ is the matter Lagrangian density minimally coupled to the metric $g_{ab}$.

It is useful to express the action (1) in a scalar-tensor representation, given by[14]

$$S=\frac{1}{2\kappa^2}\int\sqrt{-g}\left[\left(\varphi-\psi\right)R-\frac{3}{2\psi}\partial^a\psi\partial_a\psi-V\left(\varphi,\psi\right)\right]d^4x+\int\sqrt{-g}\,\mathcal{L}_m d^4x\,, \tag{2}$$

where $\varphi$ and $\psi$ are two scalar fields and $V=V(\varphi,\psi)$ is the potential. Varying the action (2) with respect to the metric $g_{ab}$ gives[15]

$$\begin{aligned}\left(\varphi-\psi\right)G_{ab}=&\,\kappa^2T_{ab}+\nabla_a\nabla_b\varphi-\nabla_a\nabla_b\psi+\frac{3}{2\psi}\partial_a\psi\partial_b\psi\\&-\left(\Box\varphi-\Box\psi+\frac{1}{2}V+\frac{3}{4\psi}\partial^c\psi\partial_c\psi\right)g_{ab}\,,\end{aligned} \tag{3}$$

where $\nabla_a$ is the covariant derivative and $\Box=\nabla^a\nabla_a$ the d'Alembertian. Varying the action with respect to the field $\varphi$ and to the field $\psi$ yields[15],

$$\Box\varphi+\frac{1}{3}\left(2V-\psi V_\psi-\varphi V_\varphi\right)=\frac{\kappa^2T}{3}\,, \tag{4}$$

$$\Box\psi-\frac{1}{2\psi}\partial^a\psi\partial_a\psi-\frac{\psi}{3}\left(V_\varphi+V_\psi\right)=0\,, \tag{5}$$

respectively. In this work, we are interested in finding wormhole solutions that satisfy the NEC throughout the spacetime geometry.

## 3. Wormhole geometries in the generalized hybrid metric-Palatini gravity

### 3.1. *Gravitational field equations*

We impose that the wormhole solutions are described by a static and spherically symmetric metric where the line element is given by[1]

$$ds^2=-e^{\zeta(r)}dt^2+\left[1-\frac{b\left(r\right)}{r}\right]^{-1}dr^2+r^2\left(d\theta^2+\sin^2\theta d\phi^2\right)\,, \tag{6}$$

where $\zeta(r)$ is the redshift function and $b(r)$ is the shape function. The latter obeys the boundary condition $b(r_0) = r_0$, where $r_0$ is the radius of the wormhole throat. The flaring-out condition at the throat vicinity is given by $(b - b'r)/b^2 > 0$. This imposes that at the throat $b'(r_0) < 1$. The redshift function must be finite everywhere to avoid the presence of horizons. We also assume that $\varphi = \varphi(r)$ and $\psi = \psi(r)$.

The three independent components of Eq. (3) are given by

$$(\varphi-\psi)\frac{b'}{r^2} - \frac{(\varphi-\psi)'}{2r}(1-b') - \frac{V}{2} - \left(1-\frac{b}{r}\right)\left[(\varphi-\psi)'' + \frac{3(\varphi-\psi)'}{2r} + \frac{3\psi'^2}{4\psi}\right] = \kappa^2\rho\,, \tag{7}$$

$$(\varphi-\psi)\left[\frac{\zeta'}{r}\left(1-\frac{b}{r}\right) - \frac{b}{r^3}\right] + \left(1-\frac{b}{r}\right)\times \times\left[-\frac{3\psi'^2}{4\psi} + \frac{2(\varphi-\psi)'}{r} + \frac{\zeta'(\varphi-\psi)'}{2}\right] + \frac{V}{2} = \kappa^2 p_r, \tag{8}$$

$$(\varphi-\psi)\left[\left(1-\frac{b}{r}\right)\left(\frac{\zeta''}{2} + \frac{\zeta'^2}{4} + \frac{\zeta'}{2r}\right) + \frac{b-rb'}{2r^3}\left(1+\frac{\zeta' r}{2}\right)\right] + \frac{\phi'}{2r}(1-b') + \frac{V}{2} + \left(1-\frac{b}{r}\right)\left[(\varphi-\psi)'' + \frac{\zeta'(\varphi-\psi)'}{2} + \frac{3\psi'^2}{4\psi} + \frac{(\varphi-\psi)'}{2r}\right] = \kappa^2 p_t\,, \tag{9}$$

where $\rho = \rho(r)$ is the energy density, $p_r = p_r(r)$ is the radial pressure, and $p_t = p_t(r)$ is the transverse pressure. The equation for $\varphi(r)$, Eq. (4), can be written as

$$\left(1-\frac{b}{r}\right)\left(\zeta'' + \frac{\zeta'^2}{2} + \frac{2\zeta'}{r}\right) - \frac{b+rb'}{r^3} + \frac{b-rb'}{r^3}\left(1+\frac{r\zeta'}{2}\right) + V_\varphi = 0\,, \tag{10}$$

and the equation for $\psi(r)$, Eq. (5), yields

$$\left(1-\frac{b}{r}\right)\left(\psi'' + \frac{\zeta'\psi'}{2} + \frac{3\psi'}{2r} - \frac{\psi'^2}{2\psi}\right) + \frac{\psi'}{2r}(1-b') - \frac{\psi}{3}(V_\varphi + V_\psi) = 0\,. \tag{11}$$

The fundamental gravitational fields are the metric fields, $\zeta(r)$ and $b(r)$, and the scalar fields $\varphi(r)$ and $\psi(r)$ that form the potential $V$. The scalar fields $\varphi$ and $\psi$ should be well-behaved throughout $r$.

### 3.2. *Interior solution*

We have the freedom to choose three functions out of the eight variables, $\zeta(r)$, $b(r)$, $\varphi(r)$, $\psi(r)$, $V(r)$, $\rho(r)$, $p_r(r)$, and $p_t(r)$, since there are only five independent equations (7)-(10). The three functions that we choose are $\zeta(r)$, $b(r)$, and $V(\varphi, \psi)$.

Consider, for simplicity, the following metric functions

$$\zeta(r) = \zeta_0\,, \qquad b(r) = \frac{r_0^2}{r}\,. \tag{12}$$

Let $V_0$ be a constant and consider further a power-law potential of the form

$$V(\varphi,\psi)=V_0\left(\varphi-\psi\right)^2 . \tag{13}$$

The gravitational field equations (7)–(9), yield the following solutions

$$\rho=\frac{r_0^2}{2r^6V_0\kappa^2}\left(24-31\frac{r_0^2}{r^2}\right), \qquad p_r=\frac{r_0^2}{2r^6V_0\kappa^2}\left(16-13\frac{r_0^2}{r^2}\right), \tag{14}$$

$$p_t=\frac{r_0^2}{2r^6V_0\kappa^2}\left(-32+39\frac{r_0^2}{r^2}\right), \tag{15}$$

respectively. The solutions for the scalar fields, $\varphi(r)$ and $\psi(r)$, are given by

$$\varphi(r)=\psi_0-\frac{r_0^2}{r^4V_0}, \qquad \psi(r)=\psi_0 . \tag{16}$$

From Eqs. (16), we see that both scalar fields are finite for $r$ at the throat $r_0$, and for values of $r$ in its vicinity, and indeed for all $r$.

The expressions for the NEC are given by

$$\rho+p_r=\frac{2r_0^2}{r^6V_0\kappa^2}\left(10-11\frac{r_0^2}{r^2}\right), \qquad \rho+p_t=\frac{4r_0^2}{r^6V_0\kappa^2}\left(-1+\frac{r_0^2}{r^2}\right). \tag{17}$$

We impose that $V_0$ negative, i.e., $V_0<0$, so that the NEC is valid in the throat's neighborhood up to to the radius $r=r_0\sqrt{11/10}$.

### 3.3. *Exterior vacuum solution*

Now, in order to guarantee that the complete solution obeys the NEC for all values of $r$, we match the interior solution at some $r_\Sigma$, to an external vacuum and asymptotically flat spherically symmetric solution, i.e., $T_{ab}=0$. Consider that the scalar fields are constant in the exterior solution, i.e., $\varphi\left(r\right)=\varphi_e$ and $\psi\left(r\right)=\psi_e$ with $\varphi_e$ and $\psi_e$ constants, $\varphi_e\neq\psi_e$, where the subscript $e$ stands for exterior.

It was shown that the Schwarzschild solution with a cosmological constant[14] of general relativity is a vacuum solution of the generalized hybrid theory we are studying[14]. These solutions are known as Schwarzschild-de Sitter, Schwarzschild-dS for short, if the constant cosmological term is positive, and Schwarzschild-anti-de Sitter, Schwarzschild-AdS for short, if the constant cosmological term is negative. The metric fields $\zeta(r)$ and $b(r)$ for the exterior region outside some radius $r_\Sigma$, i.e., $r>r_\Sigma$, are then

$$e^{\zeta(r)}=\left(1-\frac{2M}{r}-\frac{V_0\left(\varphi_e-\psi_e\right)r^2}{6}\right)e^{\zeta_e}, \qquad b\left(r\right)=2M+\frac{V_0\left(\varphi_e-\psi_e\right)r^3}{6}, \tag{18}$$

respectively, where $M$ is a constant of integration and represents the mass, and the factor $e^{\zeta_e}$ is a useful constant.

The sign of the term $V_0\left(\varphi_e-\psi_e\right)/2$ will determine whether the solution is Schwarzschild-dS or Schwarzschild-AdS[3]. This sign will be determined by the

matching surface and the imposition that the NEC holds everywhere. Thus, exterior solution, in addition to Eqs. (18), has the following expressions, $\varphi(r) = \varphi_e$, $\psi(r) = \psi_e$, $V = V_0(\varphi_e - \psi_e)^2$, and $\rho(r) = p_r(r) = p_t(r) = 0$.

### 3.4. *Surface stresses at the junction interface*

To match the interior to the exterior solution we need the junction conditions at the hypersurface $\Sigma$, which is located at the junction radius $r_\Sigma$, for the generalized hybrid metric-Palatini gravity[16]. These conditions are

$$[h_{\alpha\beta}] = 0\,, \qquad [\varphi] = 0\,, \qquad [\psi] = 0\,, \qquad n^a\,[\partial_a \psi] = 0\,, \tag{19}$$

$$n^a\,[\partial_a \varphi] = \frac{\kappa^2}{3} S\,, \qquad S^\beta_\alpha - \frac{1}{3}\delta^\beta_\alpha S = -\frac{(\varphi_\Sigma - \psi_\Sigma)}{\kappa^2}\left[K^\beta_\alpha\right]\,, \qquad [K] = 0\,, \tag{20}$$

where $h_{\alpha\beta}$ is the induced metric at the junction hypersurface $\Sigma$, with Greek indexes standing for $(0, 2, 3)$, the brackets $[X]$ denote the discontinuity of $X$ across $\Sigma$, $n^a$ is the unit normal vector to $\Sigma$, $K = K^\alpha_\alpha$ is the trace of the extrinsic curvature $K_{\alpha\beta}$ of the surface $\Sigma$, $S = S^\alpha_\alpha$ is the trace of the stress-energy tensor $S_{\alpha\beta}$ of the thin shell, and the subscripts $\Sigma$ indicate the value computed at the junction interface.

The second condition in Eq. (20) yield the field equations at the hypersurface

$$\sigma = \frac{4r_0^2}{\kappa^2 V_0 r_\Sigma^5}\left(1 - \frac{1}{4}\, r_\Sigma \left[K_0^0\right]\right)\,, \qquad p = -\frac{4r_0^2}{\kappa^2 V_0 r_\Sigma^5}\left(1 + \frac{1}{8}\, r_\Sigma \left[K_0^0\right]\right)\,, \tag{21}$$

respectively, where $\sigma = -S_0^0$ the surface energy density and $p = S_1^1$ and the surface transverse pressure. with

$$\left[K_0^0\right] = \frac{r_0 \zeta_0}{2r_\Sigma^2}\sqrt{1 - \frac{r_0^2}{r_\Sigma^2}} + \frac{\frac{r_0^2}{r_\Sigma} - 6M}{6r_\Sigma^2\sqrt{1 - \frac{2M}{r_\Sigma} + \frac{r_0^2}{6r_\Sigma^2}}}\,. \tag{22}$$

We also impose that the matter NEC should be satisfied at the thin shell[4,5], so that $\sigma + p \geq 0$ and the formulas given in Eq. (21) imply

$$\sigma + p = -\frac{3r_0^2}{2\kappa^2 V_0 r_\Sigma^4}\left[K_0^0\right]\,. \tag{23}$$

Finding a combination of parameters that fulfills all the junction conditions, plus the matter NEC everywhere, is a problem that requires some fine-tuning, and was analysed in detail[14].

## 4. Conclusions

In this work, we found traversable asymptotically AdS wormhole solutions that obey the NEC everywhere in the generalized hybrid metric-Palatini gravity theory, so there is no need for exotic matter. The interior wormhole solution obtained verifies the NEC near and at the throat of the wormhole. The matching of the interior solution to an exterior vacuum solution yields a thin shell respecting also

the NEC. We showed that it is possible to build the full wormhole solution, where the matter satisfies the NEC throughout the wormhole geometry, and found that it is asymptotically AdS.

## Acknowledgments

JLR acknowledges financial support of FCT-IDPASC through grant no. PD/BD/114072/2015. JPSL thanks Fundação para a Ciência e Tecnologia (FCT), Portugal, for financial support through Grant No. UID/FIS/00099/2013 and Grant No. SFRH/BSAB/128455/2017. FSNL acknowledges financial support of the FCT through an Investigador FCT Research contract, with reference IF/00859/2012, and the grants UID/FIS/04434/2013 and No. PTDC/FIS-OUT/29048/2017.

## References

1. M. S. Morris and K. S. Thorne, *Am. J. Phys.* **56**, 395 (1988).
2. M. Visser, *Lorentzian wormholes: From Einstein to Hawking* (Springer-Verlag, New York, 1996).
3. J. P. S. Lemos, F. S. N. Lobo, and S. Q. Oliveira, *Phys. Rev. D* **68**, 064004 (2003).
4. F. S. N. Lobo, *Class. Quant. Grav.* **21**, 4811 (2004).
5. F. S. N. Lobo, *Gen. Rel. Grav.* **37**, 2023 (2005).
6. T. Harko and F. S. N. Lobo, *Extensions of $f(R)$ Gravity: Curvature-Matter Couplings and Hybrid Metric-Palatini Theory* (Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, 2019).
7. T. Harko, F. S. N. Lobo, M. K. Mak, and S. V. Sushkov, *Phys. Rev. D* **87**, 067504 (2013).
8. F. S. N. Lobo and M. A. Oliveira, *Phys. Rev. D* **80**, 104012 (2009).
9. N. M. Garcia and F. S. N. Lobo, *Phys. Rev. D* **82**, 104018 (2010).
10. N. Montelongo Garcia and F. S. N. Lobo, *Class. Quant. Grav.* **28**, 085018 (2011).
11. S. Capozziello, T. Harko, T. S. Koivisto, F. S. N. Lobo, and G. J. Olmo, *Phys. Rev. D* **86**, 127504 (2012).
12. T. Harko, T. S. Koivisto, F. S. N. Lobo, and G. J. Olmo, *Phys. Rev. D* **85**, 084016 (2012).
13. S. Capozziello, T. Harko, T. S. Koivisto, F. S. N. Lobo, and G. J. Olmo, *Universe* **1**, 199 (2015).
14. J. L. Rosa, J. P. S. Lemos, and F. S. N. Lobo, *Phys. Rev. D* **98**, 064054 (2018).
15. J. L. Rosa, S. Carloni, J. P. S. Lemos, and F. S. N. Lobo, *Phys. Rev. D* **95**, 124035 (2017).
16. J. L. Rosa and J. P. S. Lemos, "Junction conditions and thin shells in the generalized hybrid metric-Palatini gravity", to be submitted (2019).

# Causality in first order gravity

Sandipan Sengupta*

*Department of Physics and Centre for Theoretical Studies,*
*Indian Institute of Technology Kharagpur, Kharagpur-721302, India*
**E-mail: sandipan@phy.iitkgp.ac.in*

We explore the possibility of 'time travel' through geodesics in the first order formulation of classical gravity theory. The analysis is developed around a few vacuum solutions to the field equations, constructed such that the proper time exhibits a non-monotonic flow along the geodesic. These geometries are built upon metrics whose determinant vanish over some region. The associated spacetimes are smooth and the curvature two-form fields are finite everywhere.

*Keywords*: Time machine, Degenerate metric, Causality violation; MG15 Proceedings; World Scientific Publishing.

## 1. Introduction

The possibility of obtaining theoretical models of time machine spacetimes within Einstein's theory of general relativity, as well as the practical feasibility of such a potential construction, have been and still is a matter of open debate. Whereas Einstein's equations (with or without matter) have been shown to admit spacetime solutions that exhibit closed timelike curves (and hence implying a possible realization of time travel), they seem to violate at least one or more of the stability criteria motivated from classical and quantum physics. To emphasize, while some of them fail to satisfy the energy conditions, others are either unstable against quantum fluctuations or are supported by matter fields which have no realization in nature[1–8].

Here we discuss vacuum spacetime solutions of gravity theory, with the special property that the associated metric is degenerate over a region (but not everywhere) and the (non-spacelike) geodesics are characterized by nonmonotonic proper time[9]. This could possibly imply 'time travel', provided there is a way to go across the turning points of the affine parameter along the geodesic.

These configurations satisfy the energy conditions at the invertible phase. At the noninvertible phase, there is no exotic matter. Although it is beyond the scope of this article to provide a precise formulation of the energy conditions for the degenerate metric phase, there are reasons to believe that the energy conditions are satisfied there as well[10]. In addition, these solutions are free of any divergence either in the curvature two-form fields or in the lower dimensional curvature scalars that could be associated with the zero-determinant phase. Even though it is not clear whether such solutions could be realized in nature, their very existence as classical solutions is intriguing.

## 2. Solutions of first order gravity

Spacetimes with degenerate metrics[11] appear naturally in the context of first order formulation of gravity theory, described by the Hilbert-Palatini action[12–14]. The corresponding equations of motion are obtained through the variation with respect to the tetrad and connection fields:

$$e^{[K}_{[\mu} D_\nu(\omega) e^{L]}_{\alpha]} = 0, e^{[J}_{[\nu} R_{\alpha\beta]}^{\ \ KL]}(\omega) = 0 \ . \tag{1}$$

Here we shall present three solutions to these equations with the properties mentioned earlier. Whereas the connection fields are flat for the first two, the last example corresponds to a nontrivial curvature two-form.

### 2.1. *Solution-I: Extension of Rindler metric*

Let us introduce a metric which divides the spacetime into three regions; The determinant being zero at the intermediate region and non-zero everywhere else:

$$\begin{aligned} ds^2 &= -f^2(x)dt^2 + f'^2(x)dx^2 + dy^2 + dz^2 \ \text{ at } |x| > x_0, \\ &= 0 - F^2(x)dx^2 + dy^2 + dz^2 \ \text{ at } |x| \leq x_0 \end{aligned} \tag{2}$$

where, each of the coordinates $(t, x, y, z)$ spans the whole real line $(-\infty, \infty)$ and the functions $f(x)$ & $F(x)$ satisfy:

$$f(\pm x_0) = 0 = f'(\pm x_0), \ F(x_0) = 0. \tag{3}$$

The internal metric used to raise or lower the $SO(3,1)$ indices is given by $\eta_{IJ} = diag[-1,1,1,1]$ everywhere. At the two regions $|x| > x_0$, the spacetime metric becomes equivalent to the Rindler metric upon a reparametrization $x \to u = f(x)$. The only nonvanishing component of the torsionless spin connection is $\omega_t^{\ 01} = 1$. Within the patch $|x| \leq x_0$, the spin-connection fields are chosen to be:

$$\hat{\omega}_\mu^{\ IJ} = 0, \tag{4}$$

leading to trivial $\hat{R}_{\mu\nu}^{\ \ IJ}$. Note that although the spin-connection fields as given above are not manifestly continuous across $x = x_0$, they may be made to be so through a $SO(3,1)$ gauge transformation on the fields:

$$\hat{e}'^I_\mu = \Lambda^I_{\ L} \hat{e}^L_\mu, \ \hat{\omega}'^{\ IJ}_\mu = \Lambda^I_{\ K} \hat{\omega}_\mu^{\ KL} \left(\Lambda^{-1}\right)^{\ J}_L + \Lambda^I_{\ K} \partial_\mu \left(\Lambda^{-1}\right)^{KJ}.$$

where transformation $\Lambda$ is a boost:

$$\Lambda^I_{\ J} = \begin{pmatrix} \cosh t & \sinh t & 0 & 0 \\ \sinh t & \cosh t & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

As is straightforward to verify, both the configurations $(e^I_\mu, \omega_\mu^{\ IJ})$ and $(\hat{e}^I_\mu, \hat{\omega}_\mu^{\ IJ})$ are solutions to the first order equations of motion (1). Also, the gauge-covariant

fields (i.e. tetrad, torsion and field-strength) are continuous at the phase boundaries $x = \pm x_0$.

For the invertible and noninvertible metric phases, the (torsionless) affine connection components are defined as, respectively:

$$\Gamma_{\alpha\beta\rho} = g_{\rho\sigma}\Gamma_{\alpha\beta}^{\ \ \sigma} = \frac{1}{2}\left[\partial_\alpha g_{\beta\rho} + \partial_\beta g_{\alpha\rho} - \partial_\rho g_{\alpha\beta}\right],$$
$$\hat{\Gamma}_{\alpha\beta\rho} = \hat{g}_{\rho\sigma}\hat{\Gamma}_{\alpha\beta}^{\ \ \sigma} = \frac{1}{2}\left[\partial_\alpha \hat{g}_{\beta\rho} + \partial_\beta \hat{g}_{\alpha\rho} - \partial_\rho \hat{g}_{\alpha\beta}\right] \tag{5}$$

The noninvertibility of the metric with $\hat{g}_{t\mu} = 0$ gets reflected through the indeterminacy of the affine connection components $\hat{\Gamma}_{\alpha\beta}^{\ \ t}$. This essentially implies that within the degenerate region, the null coordinate (t) becomes nondynamical and any evolution along that is equivalent to a gauge motion.

The geodesic equations at the region $x > x_0$ are defined as $u^\alpha \mathcal{D}_\alpha u^\beta := u^\alpha \partial_\alpha u^\beta + \Gamma_{\alpha\rho}^{\ \ \beta} u^\alpha u^\rho = 0$, leading to:

$$f^2(x)\dot{t}^2 - f'^2(x)\dot{x}^2 - \dot{y}^2 - \dot{z}^2 = k,\ f^2(x)\dot{t} = E,\ \dot{y} = k_y,\ \dot{z} = k_z \tag{6}$$

where $E, k_y, k_z$ are constants of motion and $k = 1, 0, -1$ characterise the timelike, null and spacelike geodesics, respectively. The affine parameter is solved as:

$$\lambda = \pm\frac{1}{(k + k_y^2 + k_z^2)}\left[E^2 - (k + k_y^2 + k_z^2)f^2(x)\right]^{\frac{1}{2}} + const. \tag{7}$$

Within the degenerate region $x \leq x_0$, the equations for the physical degrees of freedom read (the equation for the null coordinate $t$ being redundant):

$$F^2(x)\dot{x}^2 - \dot{y}^2 - \dot{z}^2 = k,\ \dot{y} = k_y,\ \dot{z} = k_z \tag{8}$$

with the solution:

$$\lambda = \pm\frac{1}{(k + k_y^2 + k_z^2)^{\frac{1}{2}}}\int dx\ F(x) + const. \tag{9}$$

In the above expressions, one could choose the smooth function $F(x)$ such that the proper time is monotonic at the degenerate region. However, it is also possible to do otherwise, leading to a turning point of the proper time. For instance, a possible set of such functions is given by:

$$f(x) = \alpha x^2 exp\left[-\frac{x_0^2}{x^2 - x_0^2}\right],\ F(x) = f'(x),$$

where $\alpha$ is a constant. This corresponds to a motion along the geodesic where a massive particle apparently can come back to the same value of proper time after a trip within the degenerate core.

### 2.2. *Solution-II: Extension of Kasner metric*

In our next example, the invertible phase of the spacetime corresponds to a particular cosmological solution of the Kasner family. Here, the null eigenvalue at the noninvertible phase lies along the $x$-direction, unlike in the previous example:

$$\begin{aligned} ds^2 &= -dt^2 + t^2 dx^2 + dy^2 + dz^2 \ \text{ at } t > 0, \\ &= -F^2(t)dt^2 + 0 + dy^2 + dz^2 \ \text{ at } t \leq 0, \end{aligned} \tag{10}$$

The function $F(t)$ obeys the boundary conditions: $F(0) = 1,\ \dot{F}(0) = 0$. The internal metric is given by $\eta_{IJ} = diag[-1, 1, 1, 1]$.

Using the affine connection components, the geodesic equations for timelike and null geodesics at $t > 0$ become:

$$\dot{t}^2 - t^2\dot{x}^2 - \dot{y}^2 - \dot{z}^2 = k,\ t^2\dot{x} = k_x,\ \dot{y} = k_y,\ \dot{z} = k_z; \tag{11}$$

Which are solved as:

$$\lambda = \pm\frac{1}{(k + k_y^2 + k_z^2)}[(k + k_y^2 + k_z^2)t^2 + k_x^2]^{\frac{1}{2}} + const. \tag{12}$$

At the degenerate phase $t \leq 0$, the physical motion is described by the remaining three coordinates through the respective geodesic equations (while the dynamics happens to be trivial along the $x$-direction as expected):

$$F^2(t)\dot{t}^2 - \dot{y}^2 - \dot{z}^2 = k,\ \dot{y} = k_y,\ \dot{z} = k_z, \tag{13}$$

These imply:

$$\lambda = \pm\frac{1}{(k + k_y^2 + k_z^2)^{\frac{1}{2}}}\int dt\ F(t) + const. \tag{14}$$

As earlier, a suitable choice of the smooth function $F(t)$ may now be made so that the proper time is nonmonotonic at $t \leq 0$.

### 2.3. *Solution III: Extension of Schwarzschild metric*

Our final example is based on a degenerate extension of the Schwarzschild exterior geometry. In terms of the chart ($t \in (-\infty, \infty),\ u \in (-\infty, \infty),\ \theta \in [0, \pi]$, $\phi \in [0, 2\pi]$), the spacetime metric at the two phases are given by:

$$\begin{aligned} ds^2 &= -\left[1 - \frac{2M}{f(u)}\right]dt^2 + \frac{f'^2(u)}{\left[1 - \frac{2M}{f(u)}\right]}du^2 + f^2(u)\left[d\theta^2 + \sin^2\theta d\phi^2\right]\ (u > u_0), \\ &= 0 - F^2(u)du^2 + H^2(u)\left[d\theta^2 + \sin^2\theta d\phi^2\right]\ (u \leq u_0), \end{aligned} \tag{15}$$

where we impose the boundary conditions:

$$f(u_0) = 2M,\ f'(u_0) = 0,\ F(u_0) = 0 \tag{16}$$

and also the constraint:

$$\left(1+\frac{H'^2(u)}{F^2(u)}\right)F(u)+2H(u)\left(\frac{H'(u)}{F(u)}\right)'=0 \tag{17}$$

The field-strength at $u > u_0$ is completely determined by the associated (torsionless) spin-connection fields $\omega_\alpha{}^{IJ}(e)$. At the region $u \leq u_0$ with a degenerate phase, we assume the nonvanishing spin-connection components to be:

$$\hat{\omega}_\theta{}^{12}=i\frac{H'(u)}{F(u)},\ \hat{\omega}_\phi{}^{23}=-\cos\theta,\ \hat{\omega}_\phi{}^{31}=-i\frac{H'(u)}{F(u)}\sin\theta \tag{18}$$

The corresponding $SO(3,1)$ field strength components read:

$$\hat{R}_{u\theta}{}^{12}=i\left[\frac{H'(u)}{F(u)}\right]',\ \hat{R}_{\theta\phi}{}^{23}=\left[1+\left(\frac{H'(u)}{F(u)}\right)^2\right]\sin\theta,\ \hat{R}_{\phi u}{}^{31}=i\left[\frac{H'(u)}{F(u)}\right]'\sin\theta, \tag{19}$$

all other components being zero. Note that even though some of the field-strength components are imaginary, the physical fields, given by their $SO(3,1)$ gauge-invariant counterparts $\hat{R}_{\mu\nu\alpha\beta}=\hat{R}_{\mu\nu}{}^{IJ}\hat{e}_{\alpha I}\hat{e}_{\beta J}$, are all real.

It is easy to verify that the configuration above satisfies the vacuum equations of motion of first order gravity everywhere. For definiteness, we shall consider the following set of functions $F(u)$ and $H(u)$ which solve the boundary conditions and the constraint:

$$F(u)=-\frac{f'(u)}{\left(\frac{2M}{f(u)}-1\right)^{\frac{1}{2}}},\ H(u)=f(u) \tag{20}$$

With this, one may now solve the geodesic equations explicitly, the procedure being exactly similar to the earlier cases. Without going into any further detail of these solutions [a], it is sufficient to note that the affine parameter $\lambda$ is characterized completely by the (smooth) function $f(u)$. An appropriate choice of $f(u)$ can be made so that the spacetime is smooth and the proper time $\lambda(u)$ exhibits turning point(s) within the degenerate region. If an observer is able to cross any such point along the geodesic, the proper clock would start running backwards.

## 3. Conclusions

Here we have constructed smooth vacuum solutions of first-order gravity where the geodesics correspond to a non-monotonic motion of the proper time. While this feature is intriguing, it is certainly not clear if this could imply a 'time travel' in reality. One should note that the turning points of the proper time precisely corresponds to those points where there exists no orientability of time. Thus, even though there is no divergence in the associated curvature two-form fields, these degenerate metric solutions are not completely satisfactory in a strict sense, at least

[a] The reader is referred to ref. [9] for more details in this regard.

classically. Rather, the appropriate habitat of such solutions should be the quantum theory, which could provide a meaningful interpretation of this non-orientability, and of the acausal features associated with these in general. A deeper understanding of such issues is essential in order to answer the question as to whether these could be seen as vacuum time machine solutions, without having to bring in additional features such as exotic matter or else.

## Acknowledgments

The author gratefully acknowledges the support (in part) of the Grant No. ECR/2016/000027 (ECR award), SERB, Department of Science and Technology, Government of India and also (in part) of the ISIRD grant (RAQ).

## References

1. W. J. van Stockum, *Proc. R Soc. Edinb.* **57**, 135 (1937).
2. K. Godel, *Rev. Mod. Phys.* **21**, 447 (1949).
3. C. W. Misner, *Relativity Theory and Astrophysics I: Relativity and Cosmology, edited by J. Ehlers (AmericanMathematical Society), Providence* (1967).
4. F. J. Tipler, *Phys. Rev. D* **9**, 2203 (1974);
   F. J. Tipler, *Phys. Rev. Lett* **37**, 879 (1976);
   F. J. Tipler, *Ann. Phys. (N.Y.)* **108**, 1 (1977).
5. J. R. Gott, *Phys. Rev. Lett.* **66**, 1126 (1991).
6. M. S. Morris, K. S. Thorne and U. Yurtsever, *Phys. Rev. Lett.* **61**, 1446 (1988).
7. A. Ori, *Phys. Rev. D* **76**, 044002 (2007);
   A. Ori, *Phys. Rev. Lett.* **95**, 021101 (2005);
   S. Krasnikov, *Phys. Rev. D* **65**, 064013 (2002).
8. S. W. Hawking, *Phys. Rev. D* **46**, 603 (1992).
9. S. Sengupta, *Phys. Rev. D* **97**, 124038 (2018).
10. J. Samuel, *Proc. Conf. of Physics at the Planck Scale (Puri), ed. J. Maharana* (1994).
11. A. Einstein and N. Rosen, *Phys. Rev.* **48**, 73 (1935).
12. A. A. Tseytlin, *J. Phys. A: Math. Gen.* **15**, L105 (1982).
13. R.K. Kaul and S. Sengupta, *Phys. Rev. D* **93**, 084026 (2016).
14. R.K. Kaul and S. Sengupta, *Phys. Rev. D* **94**, 104047 (2016).
15. R.K. Kaul and S. Sengupta, *Phys. Rev. D* **96**, 104011 (2017).
16. S. Sengupta, `arXiv:gr-qc/1803.08618` , *Arxiv Preprint*, 2018.

# Shadows cast by rotating wormholes

Rajibul Shaikh

*Department of Astronomy & Astrophysics, Tata Institute of Fundamental Research, Mumbai, Maharashtra 400005, India*
** E-mail: rajibul.shaikh@tifr.res.in, rshaikh@iitk.ac.in*

We study shadows cast, under some circumstances, by a certain class of rotating wormholes and explore the dependence of the shadows on the wormhole spin. We compare our results with that of a Kerr black hole. For small spin, the shapes of the shadows cast by a wormhole and a black hole are nearly identical to each other. However, with increasing values of the spin, the shape of a wormhole shadow start deviating considerably from that of a Kerr black hole. Detection of such considerable deviation in future observations may possibly indicate the presence of a wormhole. In other words, our results indicate that the wormholes which are considered in our work and have reasonable spin can be distinguished from a black hole, through the observation of their shadows.

*Keywords*: General Relativity; Shadow; Wormhole; Black Hole; Gravitational Lensing.

## 1. Introduction

It is commonly believed that the central suppermassive compact region of our Galaxy as well as those of many other galaxies contain suppermassive black holes. Typically, the event horizon of a black hole, together with a set of unstable light rings (or a photon sphere in the case of a static, spherically symmetric black hole) present in its exterior geometry, creates a characteristic shadow-like image (a darker region over a brighter background) of the photons emitted by nearby photon sources or of the radiation emitted by an accretion flow around it. With the purpose of detecting this shadow in the image of the suppermassive compact object Sagittarius A$^*$ (Sgr A$^*$) present at our Galactic Center and in the image of that present at the center of the nearby galaxy M87, the event horizon telescope (EHT)[1–3] is being constructed and has begun collecting data.

While the intensity map of an image depends on the details of the emission mechanisms of photons, the boundary (silhouette) of the shadow is determined only by the spacetime metric itself, since it corresponds to the apparent shape of the photon capture orbits as seen by a distant observer. Shadows and images cast by different black holes have been studied by several researchers in the past. See Ref. 4 for a recent brief review on this subject. Some recent studies, however, suggest that the presence of a shadow does not by itself prove that it is being necessarily cast by a black hole. Other horizonless compact objects[5–12] can also cast similar shadows. In case of a wormhole, the throat plays very crucial role in casting a shadow[9]. This crucial role was overlooked in Ref. 6, resulting in incomplete results. Later, this problem was revisited in Ref. 9 where the correct and complete shapes of rotating wormhole shadows were obtained by incorporating the crucial role of a wormhole throat. In what follows, we summarize the main results obtained in Ref. 9.

## 2. The rotating wormhole spacetime and null geodesics

The spacetime geometry of a stationary, axisymmetric spacetime metric describing a rotating traversable wormhole of Teo class is given by[13]

$$ds^2 = -N^2dt^2 + \frac{dr^2}{1-\frac{b}{r}} + r^2K^2\left[d\theta^2 + \sin^2\theta(d\varphi - \omega dt)^2\right], \tag{1}$$

where $-\infty < t < \infty$, and $r_0 \leq r < \infty$, $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$ are spherical coordinates. The functions $N$, $b$, $K$, and $\omega$ depend on $r$ and $\theta$. $N$ and $b$ are, respectively, known as analog redshift function and shape function. $K$ determines the area radius through $R = rK$ and $\omega$ measures the angular velocity of the wormhole. To ensure that there are no curvature singularities or event horizons, the above metric functions have to satisfy certain conditions (see Ref. 13 for details). The spacetime describes two identical, asymptotically flat regions connected together at the throat $r = r_0 = b > 0$.

We choose the following form of the metric functions which are frequently used in the literature[6,14,15]:

$$N = \exp\left[-\frac{r_0}{r} - \gamma\left(\frac{r_0}{r}\right)^{\delta}\right], \quad b(r) = r_0, \quad K = 1, \quad \omega = \frac{2J}{r^3}, \tag{2}$$

where $J$ is the angular momentum of the wormhole and $\gamma$ and $\delta$ ($> 1$) are real constants. Also, we can write $r_0 = 2M$, $M$ being the mass of the wormhole[9].

Null geodesics in the spacetime of the wormholes given in Eq. (1) have been studied in Ref. 6. The radial part of the geodesic equation for a photon is given by[6,9]

$$\left(\frac{dr}{d\tilde{\lambda}}\right)^2 + V_{eff} = 0, \quad V_{eff} = -\frac{1}{N^2}\left(1 - \frac{b}{r}\right)R(r), \tag{3}$$

$$R(r) = (1-\omega\xi)^2 - \eta\frac{N^2}{r^2K^2}, \tag{4}$$

where $V_{eff}$ is the effective potential, and $\xi$ and $\eta$ are impact parameters defined by

$$\xi = \frac{L}{E}, \quad \eta = \frac{Q}{E^2}. \tag{5}$$

Here, $E$, $L$ and $Q$ are three constants of motion and represent, respectively, the energy, the angular momentum (about the axis of symmetry) of the photon and the Carter constant.

## 3. Shadows

A wormhole connects two regions of spacetime through its throat. To obtain the shadow cast by a wormhole, we consider that the wormhole is illuminated by photons coming from an extended light source situated in one of the regions (first region), and no light sources are present in the vicinity of the throat in the other region (second region). Depending on the impact parameters, some of the photons from

the light source get captured by the wormhole, and some get scattered away to a faraway observer situated in the first region. In the observer's sky, the scattered photons which are received by him/her form bright spots. On the other hand, the captured photons plunge through the wormhole and go to the other region, thereby forming dark spots in the observer's sky. The union of the dark spots in the observer's sky constitute what we call shadow.

A set of unstable circular orbits of photons, which are also known as unstable light rings, separate the escaping and plunging photons and are characterized by certain critical values of the impact parameters $\xi$ and $\eta$. These critical impact parameters define the boundary of the shadow. The unstable circular photon orbits correspond to the maximum of the effective potential $V_{eff}$ and are determined by the standard conditions

$$V_{eff} = 0, \quad \frac{dV_{eff}}{dr} = 0, \quad \frac{d^2V_{eff}}{dr^2} \leq 0. \tag{6}$$

The unstable circular orbits may lie outside as well as at the wormhole throat[12]. For the unstable circular orbits whose radii do not coincide with the throat radius $r_0$ (i.e. for the unstable circular orbits which lie outside the throat such that $(1-\frac{b}{r}) \neq 0$ on the orbits), in terms of $R(r)$, the above set of conditions can be rewritten as

$$R(r) = 0, \quad \frac{dR}{dr} = 0, \quad \frac{d^2R}{dr^2} \geq 0. \tag{7}$$

Using $R = 0$ and $dR/dr = 0$, we obtain[6]

$$\eta = \left[\frac{r^2K^2}{N^2}(1-\omega\xi)^2\right]_{r=r_{ph}}, \quad \xi = \frac{\Sigma}{\Sigma\omega - \omega'}\Big|_{r=r_{ph}}, \quad \Sigma = \frac{1}{2}\frac{d}{dr}\ln\left(\frac{N^2}{r^2K^2}\right), \tag{8}$$

where $r_{ph}$ is the radius of a circular photon orbit. On the other hand, since $(1-\frac{b}{r})$ vanishes at the throat $r = r_0$, for the unstable circular orbits which are located at the throat (i.e., for unstable orbits for which $r_{ph} = r_0$), it can be seen from Eq. (3) that the conditions in Eq. (6) are satisfied at the throat when[9]

$$R(r_0) = 0, \quad \frac{dR}{dr}\Big|_{r=r_0} \geq 0, \tag{9}$$

which are different from those in Eq. (7). Note that, in Ref. 9, the second condition on the right hand side of Eq. (9) is written as $\frac{d^2R}{dr^2}\Big|_{r=r_0} \geq 0$ instead of $\frac{dR}{dr}\Big|_{r=r_0} \geq 0$. This was a typographic error in that paper. For nonrotating wormholes, a set of conditions similar to those in Eq. (9) have been obtained in Ref. 12. For the circular photon orbits at the throat, $R(r_0) = 0$ gives

$$(1-\omega_0\xi)^2 - \eta\frac{N_0^2}{r_0^2K_0^2} = 0, \tag{10}$$

where the subscript '0' implies that the quantities are evaluated at the throat.

In realistic observation, the apparent shape of a shadow is measured in the observer's sky, a plane which passes through the center of the wormhole geometry

and is normal to the line of sight joining the observer and the center of the wormhole geometry. The coordinates on this plane are denoted by $\alpha$ and $\beta$ and are known as celestial coordinates. The celestial coordinates are related to the impact parameters by[6]

$$\alpha = -\frac{\xi}{\sin\theta_{obs}}, \quad \beta = \left(\eta - \frac{\xi^2}{\sin^2\theta_{obs}}\right)^{1/2}, \tag{11}$$

where $\theta_{obs}$ is the inclination angle of the observer, i.e., the angle between the rotation axis of the wormhole and the line of sight of the observer. In the $(\alpha, \beta)$-plane, Eqs. (8) and (11) define the part of the boundary of the shadow formed by the unstable circular photon orbits which lie outside the throat. However, for the part of the boundary of the shadow which is due to the unstable circular orbits at the throat, we have, from Eqs. (10) and (11)[9],

$$(N_0^2 - \omega_0^2 r_0^2 K_0^2 \sin^2\theta_{obs})\alpha^2 - 2\omega_0 r_0^2 K_0^2 \sin\theta_{obs}\alpha - r_0^2 K_0^2 + N_0^2\beta^2 = 0. \tag{12}$$

The last equation, which incorporates the effect of a wormhole throat and was not considered in Ref. 6, plays very crucial role in determining the shape of the shadow of a wormhole. The complete apparent shape of the shadow is given by the common region enclosed by the curve obtained from Eqs. (8) and (11) and that obtained from Eq. (12) (see Ref. 9 for details).

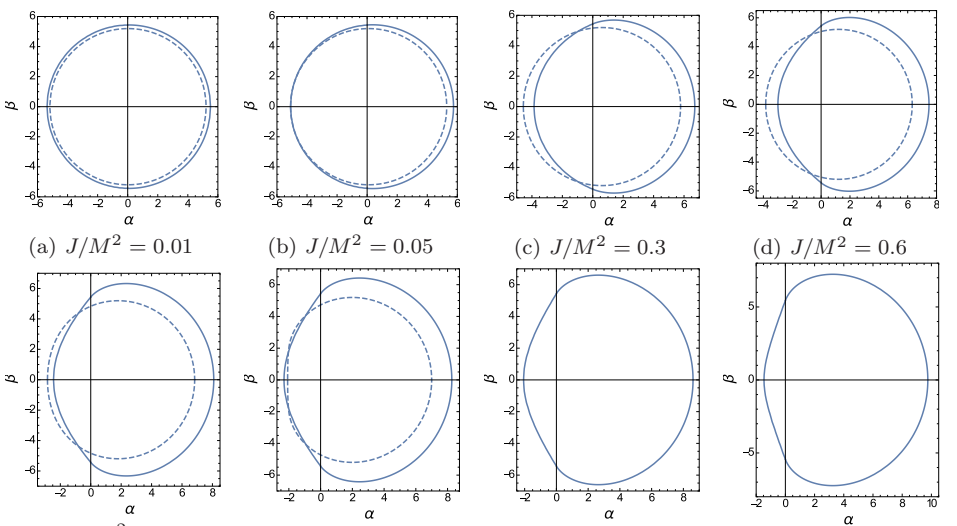



Fig. 1. The shape of the shadow of a rotating wormhole (solid curve) and the Kerr black hole (dashed curve) for different spin values [(a)-(h)]. The metric functions of the wormhole are given by Eq. (2) with $\gamma = 0$ and $\delta = 0$. Here, $\theta_{obs} = 90°$. The axes are in units of the mass $M$. The figures are extracted from Ref. 9.

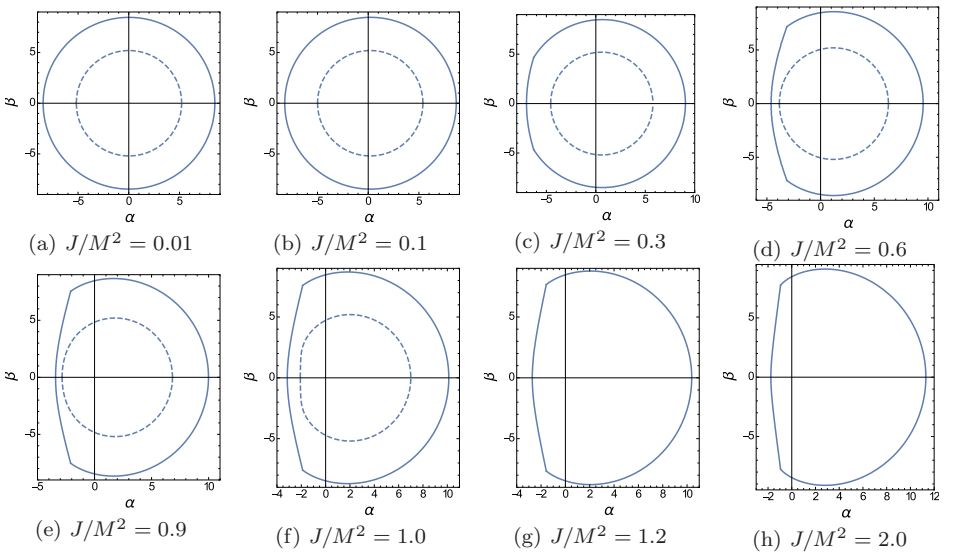



Fig. 2. The shape of the shadow of a rotating wormhole (solid curve) and the Kerr black hole (dashed curve) for different spin values [(a)-(h)]. The metric functions of the wormhole are given by Eq. (2) with $\gamma = 1$ and $\delta = 2$. Here, $\theta_{obs} = 90°$. The axes are in units of the mass $M$. The figures are extracted from Ref. 9.

Figures 1 and 2 show the shadow cast by a rotating wormhole whose metric functions are given by Eq. (2). For comparison, we have also plotted the shadow cast by a Kerr black hole for each set of the parameter values. Note that, for small spin $a$ $(= J/M^2)$, the shape of a wormhole shadow is qualitatively similar to that of the Kerr black hole as they both are almost circular. However, with increasing spin, the characteristic deformation (from the circular shape) in the wormhole shadow is more and more prominent than that in the black hole shadow, resulting in considerable deviation between the two. Such considerable deviation of the wormhole shadow from that of the Kerr black hole may be relevant to discriminate between the two in future observations.

## 4. Conclusion

We have studied shadows cast by a class of rotating Teo wormhole by incorporating the the crucial role that a wormhole throat can play. We have compared our results with those of a Kerr black hole. We find that, for small spin, the shapes of the both black hole and wormhole shadow are qualitatively identical to each other, i.e., both are nearly circular. However, with increasing spin, the characteristic deformation (from the circular shape) in the wormhole shadow is more and more prominent than that in the black hole shadow, resulting in considerable deviation between the two. Such considerable deviation, if detected in future observations, may possibly indicate the presence of a wormhole. In other words, our results indicate that,

through the observations of the shapes of their shadows, a wormhole which have reasonable spin can be distinguished from a black hole.

However, our conclusions are largely based on the specific types of wormholes we have chosen to work with. It will be worthwhile to see whether or to what extent the conclusions drawn here carry over to a broader class of rotating wormholes.

**Acknowledgments**

The author acknowledges Tata Institute of Fundamental Research, Mumbai, India, for providing financial assistance to attend the conference.

**References**

1. S. S. Doeleman et al., *Nature* **455**, 78 (2008).
2. S. S. Doeleman et al., *Science* **338**, 355 (2012).
3. `http://www.eventhorizontelescope.org.`
4. P. V. P. Cunha and C. A. R. Herdeiro, *Gen. Relativ. Gravit.* **50**, 42 (2018).
5. A. E. Broderick and R. Narayan, *Astrophys. J.* **638**, L21 (2006).
6. P. G. Nedkova, V. Tinchev, and S. S. Yazadjiev, *Phys. Rev. D* **88**, 124019 (2013).
7. C. Bambi, *Phys. Rev. D* **87**, 107501 (2013).
8. T. Ohgami and N. Sakai, *Phys. Rev. D* **91**, 124020 (2015).
9. R. Shaikh, *Phys. Rev. D* **98**, 024044 (2018).
10. P. V. P. Cunha, C. A. R. Herdeiro, and M. J. Rodriguez, *Phys. Rev. D* **97**, 084020 (2018).
11. R. Shaikh, P. Kocherlakota, R. Narayan, and P. S. Joshi, *Mon. Not. Roy. Astron. Soc.* **482** (2019) 52.
12. R. Shaikh, P. Banerjee, S. Paul, and T. Sarkar, *Phys. Lett. B* **789**, 270 (2019).
13. E. Teo, *Phys. Rev. D* **58**, 024014 (1998).
14. T. Harko, Z. Kovacs, and F. Lobo, *Phys. Rev. D* **79**, 064001 (2009).
15. C. Bambi, *Phys. Rev. D* **87**, 084039 (2013).

# Deformations of massive field theories

Sebastian Garcia-Saenz,[1,†] Nicolas Boulanger,[2] Cédric Deffayet[1,3] and Lucas Traina[2]

[1] *Sorbonne Université, UPMC Paris 6 and CNRS, UMR 7095, Institut d'Astrophysique de Paris, GReCO, Paris 75014, France*
[2] *Groupe de Mécanique et Gravitation, Physique théorique et mathématique, Université de Mons – UMONS, 20 place du Parc, 7000 Mons, Belgium*
[3] *IHES, Le Bois-Marie, 35 route de Chartres, Bures-sur-Yvette 91440, France*
[†] *E-mail: sebastian.garcia-saenz@iap.fr*

The BRST-BV formalism provides a general and systematic framework to classify consistent deformations of gauge theories. This note is a brief review of the generalization of the method to accommodate the treatment of massive field theories in the Stueckelberg formulation. Applications include the classification of the cubic and quartic interaction vertices for a multiplet of massive spin-1 fields, and of the cubic deformations of the theory of a single massive spin-2 field. The results are shown to correctly reproduce the known vertices of massive Yang–Mills theory and massive gravity, respectively. The formalism also sheds light on the characterization of Stueckelberg gauge theories, by demonstrating in particular that they can be abelianized by means of local field redefinitions.

## 1. Introduction

The problem of *deformations* of field theories concerns the following basic question: given a free theory for a some set of fields, what are the most general interactions that one can add in a *consistent* way?[a] The requirement of consistency is an essential assumption, and is the statement that any deformation must preserve the number and nature of the degrees of freedom. In a gauge theory, in particular, this implies that the number of gauge symmetries must stay the same — since a gauge symmetry is nothing but a constraint on the number of propagating fields — although the form of the symmetries may in general change as one deforms the initial theory.

The solution to the problem of consistent deformations is well known in some relatively simple cases. For instance in the case of the free Maxwell theory, it is easy to prove that the abelian $U(1)$ gauge symmetry admits no deformations, meaning that any interaction vertices must be invariant under this symmetry. In four dimensions one can also show that all such invariants must of the Born–Infeld type, i.e. they can be written solely in terms of a basic set of invariant tensors. In the context of Maxwell theory this implies that any deformation of the action of a massless spin-1 particle must be given by Lorentz-invariant contractions of $F_{\mu\nu}$ and its derivatives. The story is more interesting for a collection of vector fields $A^a_\mu$, with $a = 1, \ldots, n$. The free action may always be diagonalized (assuming no negative-norm states) so as to take the form $S_0 = -\frac{1}{4}\int d^D x\, F^{\mu\nu}_a F^a_{\mu\nu}$, where again

[a] Although in this work we have focused on deformations of free theories, the general formalism is not restricted to this situation. For instance, one could consider higher order deformations of models that are already interacting, or the question of how to couple two sets of degrees of freedom which are independently described by non-trivial theories.

$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu$ and color indices are raised and lowered with $\delta_{ab}$. As in the single-field case, there is always the possibility of constructing interactions in the form of Born–Infeld terms, which are however less interesting as they correspond to higher-dimension operators. The more compelling direction is instead to allow for deformations of the $U(1)^n$ gauge symmetry, which are no longer trivial if $n \geq 3$. This procedure leads, under some assumptions on the number of derivatives, to Yang–Mills theory as the most general interacting theory for a multiplet of massless spin-1 fields. In other words, the problem of consistent deformations allows one to *derive* Yang–Mills theory purely from the knowledge of the particle content and the global symmetries (in this case Poincaré invariance) that define the model. In particular, the Lie algebraic nature of the deformed gauge transformations is not an added assumption but follows as a consequence of the formalism. More recent application of this method to models of relevance in gravity include the treatment of massless spin-2 fields in [1] and of partially massless spin-2 fields in [2].

These examples were concerned with massless fields for which gauge invariance plays a defining role. Indeed, the BRST-BV method that we will describe below is tailored to the analysis of gauge theories. In this regard, it may seem that deformations of *massive* theories would be much simpler to study given that the consistency of the gauge symmetries is not an issue anymore in this setting. Nevertheless, even in the absence of constraints imposed by gauge invariance, there is still the concern that interactions may introduce extra degrees of freedom. A famous example is the case of a massive spin-2 field, for which generic interactions lead to the presence of an additional scalar mode in the spectrum, the so-called Boulware–Deser ghost.[b] This is our main motivation to generalize the BRST-BV deformation procedure to include the treatment of massive field theories.

Before we proceed with a description of the formalism, it is instructive to comment on the obstacles that one would encounter by following the more direct approach of the Noether procedure. In this set-up, one seeks to construct interactions perturbatively, by expanding the action as $S = S^{(2)} + S^{(3)} + S^{(4)} + \cdots$, while at the same time allowing for field-dependent deformations of the gauge symmetry, i.e. $\delta_\epsilon \varphi^i = \delta^{(0)}_\epsilon \varphi^i + \delta^{(1)}_\epsilon \varphi^i + \delta^{(2)}_\epsilon \varphi^i + \cdots$, where the superscript $(k)$ means $k$ powers of the fields $\varphi^i$, and $\epsilon$ is a gauge parameter. The consistency condition is given by the gauge invariance of the full action, $\delta_\epsilon S = 0$, which is to be solved iteratively for the deformations $S^{(k)}$ and $\delta^{(k)}_\epsilon$. Although straightforward, this method suffers from several drawbacks. First and foremost is the fact that interaction vertices are in general ambiguous due to the existence of trivial deformations, which can arise from total derivatives, from non-linear redefinitions of the fields, and from

[b] From the point of view of effective field theory (EFT), one may argue that additional degrees of freedom arising from the interactions should not be considered an issue, as they will necessarily lie at or above the cutoff of the EFT. Nevertheless, the search for interactions for which such extra fields are absent is still of interest, as it may lead to a higher cutoff compared to what a generic derivative expansion would imply, often making the model more compelling for physical applications. The case of massive gravity is a classic example.

field-dependent redefinitions of the gauge parameters. Next there is the issue of accounting for the existence of gauge symmetries that vanish on-shell, which are likewise trivial and should be discarded. Lastly, although the Noether procedure is in principle perfectly general, in practice it can be quite cumbersome to apply, especially for theories in curved spacetimes where some simplifying techniques cannot be easily applied. We will see that these obstacles can be addressed in a very systematic way within the BRST-BV method.

## 2. BRST-BV formalism

The origin of the BRST-BV method goes back to the antifield formalism of Batalin and Vilkovisky (BV),[3] whose purpose is to provide a fully general approach to the quantization of non-abelian gauge theories, generalizing the Faddeev–Popov procedure as well as the original BRST method.

Consider an action $S[\varphi]$ for a set of fields $\varphi^i$ and with gauge symmetries $\delta_\epsilon \varphi^i = R^i{}_\alpha[\varphi]\epsilon^\alpha$. As in the BRST method, we introduce a ghost $C^\alpha$ for each gauge symmetry, and note that $C^\alpha$ is fermionic if $\epsilon^\alpha$ is bosonic. In addition, we introduce a set of antifields $\varphi^*_i$ and $C^*_\alpha$, respectively for each field and for each ghost. By definition, the antifields have opposite statistics to the fields they match to. The fundamental object that defines the dynamics of the theory is the *BV functional*

$$W = S + \int d^D x \Big[ R^i{}_\alpha[\varphi]\varphi^*_i C^\alpha + f^\gamma{}_{\alpha\beta}[\varphi] C^*_\gamma C^\alpha C^\beta + M^{ij}_{\alpha\beta}[\varphi]\varphi^*_i \varphi^*_j C^\alpha C^\beta + \cdots \Big] . \quad (1)$$

The tensors $R^i{}_\alpha$, $f^\gamma{}_{\alpha\beta}$ and $M^{ij}_{\alpha\beta}$ are in general non-linear functionals of the fields, and the ellipsis stands for an a priori infinite series in antifields, each term multiplied by a "structure functional" of the fields as well as powers of the ghosts. The form of the structure functionals is determined by the *master equation*,

$$(W, W) = 0 , \quad (2)$$

where $(A, B)$ stands for the antibracket of any two functionals $A$ and $B$ (we will not need the explicit definition; see for details). The extension of the action given by the BV functional implies that the original gauge invariance of the theory is removed. However, the information about the gauge symmetry is still encoded in full within the master equation, in the following sense. By expanding (2) in antifields one finds at lowest order the relation $R^i{}_\alpha \frac{\delta S}{\delta \varphi^i} = 0$, which is nothing but the Noether identity associated to the gauge invariance. At the next order one has $R^j{}_\alpha \frac{\delta R^i{}_\beta}{\delta \varphi^j} - R^j{}_\beta \frac{\delta R^i{}_\alpha}{\delta \varphi^j} = f^\gamma{}_{\alpha\beta} R^i{}_\gamma + M^{ij}_{\alpha\beta} \frac{\delta S}{\delta \varphi^j}$, which is the statement that gauge transformations form an algebra. Notice that in general the gauge algebra is not of the Lie type due to the fact that the $f^\gamma{}_{\alpha\beta}$ are not constant. In this way one may continue to higher orders in antifields to derive an infinite tower of consistency relations for the functionals that define the structure of the gauge algebra. For instance at the following order one encounters the Jacobi identity that encodes the consistency of the functionals $f^\gamma{}_{\alpha\beta}$, just like the gauge algebra encodes the consistency of the operators $R^i{}_\alpha$ that define the original gauge transformations.

The main idea of the deformation analysis is to revert this logic.[4] We are now given only a free theory, with some fields and gauge symmetries, and we seek to determine the most general set of interactions by solving the master equation. In practice we do this perturbatively, expanding the BV action as $W = W_0 + gW_1 + g^2W_2 + \cdots$, where $g$ is a formal deformation parameter that will play the role of a coupling constant. At first order in $g$ eq. (2) reads $(W_0, W_1) = 0$, which is to be solved for $W_1$; at second order we have to solve for $W_2$ from the equation $(W_1, W_1) + 2(W_0, W_2) = 0$, and so on. At this stage it may seem that we have not simplified the problem by much in comparison to the Noether procedure — instead of solving for the action and gauge transformations, we now have to solve for the BV functional order by order. The real technical advantage of the BRST-BV formalism may be seen by rephrasing the analysis as a cohomological problem, as we now explain.

We have remarked that the BV action has no gauge invariance, which is precisely the reason for introducing the ghost fields in the first place. But a remnant of the gauge symmetry is still present in $W_0$ in the form of a *BRST symmetry*, $sW_0 = 0$, where $s$ denotes the BRST differential. One can then show that $W_0$ acts as a generator of BRST transformations in the sense that $sA = (W_0, A)$ for any functional $A$. The order-$g$ part of the master equation then states that $W_1$ is BRST-closed, $sW_1 = 0$. Since $s^2 = 0$, given any solution of this equation we can always add a BRST-exact term $W_1' = sB_1$. But it can be proved that such terms are precisely the expressions that encode the possible trivial deformations alluded to above. In other words, the non-trivial solutions of the master equation that we are after are defined by the cohomology of the BRST differential. The method thus provides a very clean mathematical characterization of trivial deformations of field theories.

## 3. Application to massive field theories

It is clear from the construction of the BV functional that gauge symmetries play a crucial role in the formalism. This may seem to preclude the application of the method to treat massive theories, which in their usual parametrization possess no gauge invariance. However, as is well known, gauge symmetries may be straightforwardly introduced simply by adding some extra fields in such a way that the action regains the invariance of the massless theory. This is known as the *Stueckelberg procedure*. In the following we summarize our results of the application of the BRST-BV formalism to analyze interactions for a multiplet of massive spin-1 particles and for a single massive spin-2 particle.[5]

### 3.1. *Multiple massive spin-1 fields*

The starting point is the free action for a collection of massive spin-1 fields $A^a_\mu$ $(a = 1, \ldots, n)$, and is given by a sum of $n$ Proca Lagrangians, and we assume all the fields to have the same mass $m$. To this we add a set of Stueckelberg scalars

$\pi^a$ responsible for restoring the local $U(1)^n$ symmetry of the massless theory, with gauge functions $\epsilon^a$ and corresponding ghosts $C^a$. The antifields are denoted by $A^{*\mu}_a$, $\pi^*_a$ and $C^*_a$.

The solution of the first order deformation of the BV functional may be written as $W_1 = \int d^D x(a_0 + a_1 + a_2)$, where the subscript in $a_k$ denotes the antifield number. From the master equation we obtain $a_2 = \frac{g}{2} f^a{}_{bc} C^*_a C^b C^c$, $a_1 = g f^a{}_{bc} A^{*\mu}_a A^b_\mu C^c + \frac{g}{2} f^a{}_{bc} \pi^*_a \pi^b C^c$, and $a_0 = \frac{g}{2} f^a{}_{bc} \left(-F^{\mu\nu}_a A^b_\mu A^c_\nu + A^\mu_a \partial_\mu \pi^b \pi^c\right)$. Each term has a clear interpretation: $a_2$ encodes the deformation of the gauge algebra, so that the coefficients $f^a{}_{bc}$ play the role of structure constants; $a_1$ encodes the deformation of the gauge transformations at linear order in the fields; and $a_0$ encodes the cubic vertices of the deformed theory. Notice that the existence of a non-trivial $a_0$ enforces the constraint $f_{abc} = f_{[abc]}$ on the structure constants, whereas initially we only had $f^a{}_{bc} = f^a{}_{[bc]}$.

Continuing to the next order, one finds an obstruction when solving for $W_2$ unless one imposes the constraint $f^a{}_{b[c} f^b{}_{de]} = 0$. This is nothing but the Jacobi identity for the structure constants, and therefore the gauge algebra of the theory must be a Lie algebra for the deformations to be consistent. We omit the explicit result for $W_2$, but we remark that it does correctly reproduce, together with $W_1$ as given above, the cubic and quartic non-abelian interactions for a multiplet of massive spin-1 fields, which have been derived previously via other methods.[6]

### 3.2. *Massive spin-2 field*

The free theory of a spin-2 field $h_{\mu\nu}$ of mass $m$ is given by the Fierz–Pauli action. Linear diffeomorphism invariance is restored by the inclusion of a Stueckelberg vector $B_\mu$, and the corresponding ghost is denoted by $C_\mu$. In order to have a well behaved high energy limit, it is useful to further add a Stueckelberg scalar $\varphi$ in order to restore a local $U(1)$ symmetry, whose ghost we denote by $C$.

The calculations in this case are more involved and so we limit ourselves to present only the final result for the cubic vertices (we refer the reader to[5] for details), omitting contributions from the Stuckelberg fields:

$$a_0\Big|_{B_\mu=0=\varphi} = \kappa \mathcal{L}^{(3)}_{\mathrm{EH}}[h] + \frac{\kappa m^2}{2}\left(h^{\mu\nu} h_\mu{}^\rho h_{\rho\nu} - \frac{5}{4} h h^{\mu\nu} h_{\mu\nu} + \frac{1}{4} h^3\right) + \alpha_1 \bar{a}^{(\mathrm{dRGT})}_0 + \alpha_2 \bar{a}^{(\mathrm{PL}_1)}_0 + \alpha_3 \bar{a}^{(\mathrm{PL}_2)}_0 \,. \tag{3}$$

Here $\mathcal{L}^{(3)}_{\mathrm{EH}}$ is the cubic part of the Einstein–Hilbert Lagrangian, so that the deformation parameter $\kappa$ may be identified with $1/M_P^{(D-2)/2}$, with $M_P$ the Planck mass. There are also three arbitrary constants $\alpha_{1,2,3}$ multiplying the homogeneous solutions $\bar{a}_0$ of the master equation. These are all terms of the Born–Infeld type, constructed out of a basic invariant tensor $H_{\mu\nu}$. In particular, $\bar{a}^{(\mathrm{dRGT})}_0$ contains no derivatives of $H_{\mu\nu}$, and when added to the first line in (3) it produces the cubic interactions of dRGT massive gravity.[7] Lastly, the operators $\bar{a}^{(\mathrm{PL})}_0$ involve second

derivatives of $H_{\mu\nu}$ and correspond to the cubic pseudo-linear interaction terms for a massive spin-2 field.[8] Our result in eq. (3) thus provides a full classification of the possible cubic vertices for a single massive graviton.

### 3.3. *Abelianization*

We conclude by commenting on an interesting property of Stueckelberg gauge theories, which has been made manifest thanks to our application of the BRST-BV formalism. This is the property that interacting Stueckelberg theories may always be *abelianized*, that is there always exists a choice of field variables for which the gauge symmetries become abelian, reducing to the symmetries of the free theory. This follows from the fact that the ghost associated to a Stueckelberg symmetry is necessarily BRST-exact, which in turn implies that all deformations of the gauge algebra and the gauge symmetries are trivial. A further interesting consequence of this abelianization property is that the action is of the Born–Infeld type, which makes it a simple matter to construct the most general set of interactions for a given set of fields. That being said, the use of variables that makes the abelian nature of the theory manifest is not always the smartest choice, as this formulation may obscure the correct power counting of operators in the derivative expansion.

### Acknowledgments

The work of SGS and his attendance to the MG15 Meeting were supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 307934, NIRG project).

### References

1. N. Boulanger, T. Damour, L. Gualtieri and M. Henneaux, Inconsistency of interacting, multigraviton theories, *Nucl. Phys. B* **597** (2001), 127-171.
2. N. Boulanger, C. Deffayet, S. Garcia-Saenz and L. Traina, Theory for multiple partially massless spin-2 fields, *Phys. Rev. D* **100** (2019) no.10, 101701.
3. I. A. Batalin and G. A. Vilkovisky, Gauge Algebra and Quantization, *Phys. Lett.* **102B**, 27 (1981), 463.
4. G. Barnich and M. Henneaux, Consistent couplings between fields with a gauge freedom and deformations of the master equation, *Phys. Lett.* **B311**, 123 (1993).
5. N. Boulanger, C. Deffayet, S. Garcia-Saenz and L. Traina, Consistent deformations of free massive field theories in the Stueckelberg formulation, *JHEP* **07**, p. 021 (2018).
6. Yu. M. Zinoviev, On massive spin 2 interactions, *Nucl. Phys.* **B770**, 83 (2007).
7. C. de Rham, G. Gabadadze and A. J. Tolley, Resummation of Massive Gravity, *Phys. Rev. Lett.* **106**, p. 231101 (2011).
8. K. Hinterbichler, Ghost-Free Derivative Interactions for a Massive Graviton, *JHEP* **10**, p. 102 (2013).

# Scalar-Gauss-Bonnet theories: Evasion of no-hair theorems and novel black-hole solutions

Panagiota Kanti

*Division of Theoretical Physics, Department of Physics,University of Ioannina, Ioannina GR-45110, Greece*
**E-mail: pkanti@cc.uoi.gr*

Athanasios Bakopoulos

*Division of Theoretical Physics, Department of Physics,University of Ioannina, Ioannina GR-45110, Greece*
*E-mail: abakop@cc.uoi.gr*

Nikolaos Pappas

*Nuclear and Particle Physics Section, Physics Department, National and Kapodistrian University of Athens, Athens GR-15771, Greece*
*E-mail: npappas@cc.uoi.gr*

We consider a general Einstein–scalar–Gauss-Bonnet theory with a coupling function $f(\phi)$ between the scalar field and the quadratic gravitational Gauss-Bonnet term. We show that the existing no-hair theorems are easily evaded, and therefore black holes may emerge in the context of this theory. Indeed, we demonstrate that, under mild only assumptions for $f(\phi)$, asymptotic solutions describing either a regular black-hole horizon or an asymptotically-flat solution always emerge. We then show, through numerical integration, that the field equations allow for the smooth connection of these asymptotic solutions, and thus for the construction of a complete, regular black-hole solution with non-trivial scalar hair. We present and discuss the physical characteristics of a large number of such solutions for a plethora of coupling functions $f(\phi)$. Finally, we investigate whether pure scalar-Gauss-Bonnet black holes may arise in the context of our theory when the Ricci scalar may be altogether ignored.

*Keywords*: Generalised Gravitational Theories, Gauss-Bonnet term, no-hair theorems, black-hole solutions, scalar hair

## 1. Introduction

The General Theory of Relativity is a beautiful mathematical theory that predicts a variety of gravitational solutions, with the black holes being the most fascinating example. In the context of General Relativity, the black-hole solutions have been uniquely determined and classified according to their properties (mass, charge and angular-momentum). No-Hair theorems, that forbid the association of a black hole with any other "charge" or field, were formulated quite early on. The existence of black-hole solutions associated with a non-trivial scalar field in the region outside the black-hole horizon has also been intensively studied. The *old no-hair theorem*[1] was formulated in the seventies, and excluded static black holes with a scalar field. However, this was outdated by the discovery of black holes with Yang-Mills[2], Skyrme fields[3] or conformally-coupled scalar fields[4]. Twenty years later, the *novel no-hair theorem*[5] was formulated (for more recent analyses, see[6–8]) but this was also

shown to be evaded in the context of the Einstein-Dilaton-Gauss-Bonnet theory[9] and in shift-symmetric Galileon theories[10,11].

In fact, the black-hole solutions[9–11] were derived in the context of the so-called generalised gravitational theories, where additional fields and higher gravitational terms may be present. These theories comprise a popular test-bed for the formulation of the ultimate theory of gravity beyond Einstein's General Theory of Relativity, and are under intense research activity. In this work, we will consider a wide class of gravitational theories where a scalar field $\phi$ has a general coupling function $f(\phi)$ to the quadratic gravitational Gauss-Bonnet (GB) term. Choosing the coupling function to be of an exponential or a linear form, one recovers the two novel black-hole solutions with non-trivial scalar hair[9,11], respectively. In[12] we demonstrated that, in fact, this class of theories with an arbitrary $f(\phi)$ always evades the existing no-hair theorems and allow for the emergence of novel black-hole solutions, with a regular horizon and an asymptotically-flat limit. Here, we review these results and discuss the characteristics of these solutions. We also investigate whether solutions arise in the context of the pure scalar-Gauss-Bonnet theory where the Ricci scalar may be ignored.

## 2. The Einstein-Scalar-Gauss-Bonnet theory

We will therefore consider the following generalised gravitational theory

$$S = \frac{1}{16\pi}\int d^4x\sqrt{-g}\left[R - \frac{1}{2}\,\partial_\mu\phi\,\partial^\mu\phi + f(\phi)R^2_{GB}\right], \tag{1}$$

where the GB term is defined as $R^2_{GB} = R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma} - 4R_{\mu\nu}R^{\mu\nu} + R^2$. By varying the above action with respect to the metric tensor and scalar field, we obtain the following gravitational field equations and the equation for the scalar field:

$$G_{\mu\nu} = T_{\mu\nu}\,, \qquad \nabla^2\phi + \dot{f}(\phi)R^2_{GB} = 0\,, \tag{2}$$

respectively, where a dot denotes the derivative with respect to the scalar field. The energy-momentum tensor has the form

$$T_{\mu\nu} = -\frac{1}{4}g_{\mu\nu}\partial_\rho\phi\partial^\rho\phi + \frac{1}{2}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}(g_{\rho\mu}g_{\lambda\nu} + g_{\lambda\mu}g_{\rho\nu})\eta^{\kappa\lambda\alpha\beta}\tilde{R}^{\rho\gamma}{}_{\alpha\beta}\nabla_\gamma\partial_\kappa f. \tag{3}$$

In the above, $\tilde{R}^{\rho\gamma}{}_{\alpha\beta} = \eta^{\rho\gamma\sigma\tau}R_{\sigma\tau\alpha\beta} = \epsilon^{\rho\gamma\sigma\tau}R_{\sigma\tau\alpha\beta}/\sqrt{-g}$. In the context of the above theory, we will look for regular, static, spherically-symmetric and asymptotically-flat black-hole solutions described by the line-element

$$ds^2 = -e^{A(r)}dt^2 + e^{B(r)}dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\varphi^2)\,. \tag{4}$$

Using the above expression, the Einstein's equations take the following explicit form

$$4e^B(e^B + rB' - 1) = \phi'^2\left[r^2e^B + 16\ddot{f}(e^B - 1)\right] - 8\dot{f}\left[B'\phi'(e^B - 3) - 2\phi''(e^B - 1)\right], \tag{5}$$

$$4e^B(e^B - rA' - 1) = -\phi'^2r^2e^B + 8\left(e^B - 3\right)\dot{f}A'\phi', \tag{6}$$

$$e^B\left[rA'^2 - 2B' + A'(2 - rB') + 2rA''\right] = -\phi'^2 re^B \\ +8\phi'^2\ddot{f}A' + 4\dot{f}[\phi'(A'^2 + 2A'') + A'(2\phi'' - 3B'\phi')], \qquad (7)$$

while the scalar equation reads

$$2r\phi'' + (4 + rA' - rB')\,\phi' + \frac{4\dot{f}e^{-B}}{r}\left[(e^B - 3)A'B' - (e^B - 1)(2A'' + A'^2)\right] = 0. \qquad (8)$$

In the above, we have assumed that the scalar field depends only on the radial coordinate, and thus the prime denotes differentiation with respect to $r$.

The unknown quantities, that we seek to determine through the solution of the system of Eqs. (5)-(8), are the scalar field $\phi$ and the metric functions $A$ and $B$. Of these, the metric function $B$ may be easily determined in terms of $(\phi, A)$ through Eq. (6). Then, the remaining field equations lead to a system of two independent, ordinary differential equations of second order for the functions $A$ and $\phi$:

$$A'' = \frac{P}{S}\,, \qquad \phi'' = \frac{Q}{S}\,. \qquad (9)$$

The functions $P$, $Q$ and $S$ are rather complicated expressions of $(r, \phi', A', \dot{f}, \ddot{f})$ and may be found in[12].

For a regular horizon to form, we demand that $e^A \to 0$ in Eq. (4), while $\phi$, $\phi'$ and $\phi''$ remain finite, as $r \to r_h$. Then, the 2nd of Eqs. (9) yields the constraint

$$\phi'_h = \frac{r_h}{4\dot{f}_h}\left(-1 \pm \sqrt{1 - \frac{96\dot{f}_h^2}{r_h^4}}\right). \qquad (10)$$

The quantity under the square-root should be positive which results in the additional bound $\dot{f}_h^2 < r_h^4/96$. Using the above in the 1st of Eqs. (9), we may uniquely determine the form of $A'$ near the horizon. Putting everything together, the near-horizon solution reads

$$e^A = a_1(r - r_h) + \dots\,, \qquad e^{-B} = b_1(r - r_h) + \dots\,, \\ \phi = \phi_h + \phi'_h(r - r_h) + \phi''_h(r - r_h)^2 + \dots\,. \qquad (11)$$

On the other hand, at asymptotic infinity, we assume power-law expressions for the metric functions and scalar field as customary. Substituting these expressions into the field equations, we obtain

$$e^A = 1 - \frac{2M}{r} + \frac{MD^2}{12r^3} + \dots\,, \quad e^B = 1 + \frac{2M}{r} + \frac{16M^2 - D^2}{4r^2} + \dots\,, \\ \phi = \phi_\infty + \frac{D}{r} + \frac{MD}{r^2} + \frac{32M^2D - D^3}{24r^3} + \dots\,. \qquad (12)$$

The above asymptotic behaviour is characterised by the ADM mass $M$ and scalar charge $D$ of the black hole. We may therefore conclude that the scalar-tensor theory

(1) with a general coupling function $f(\phi)$ is always compatible with either a regular horizon or an asymptotically-flat limit.

However, no complete black-hole solution may be constructed unless the aforementioned asymptotic solutions are smoothly matched. To investigate whether this is in principle possible, we turn to the novel no-hair theorem[5] and examine its requirements under which it may forbid the existence of such a solution. This theorem assumes first that, at asymptotic infinity, the $T^r{}_r$ component of the energy-momentum tensor is positive and decreasing. Indeed, we find that this has the form

$$T^r{}_r = \frac{e^{-B}\phi'}{4}\left[\phi' - \frac{8e^{-B}\left(e^B-3\right)\dot{f}A'}{r^2}\right] \simeq \frac{\phi'^2}{4} \sim \mathcal{O}(\frac{1}{r^4}). \tag{13}$$

In the near-horizon regime, $T^r{}_r$ should be negative and increasing according to[5]; however, employing the asymptotic solution (11), we find that in our case

$$T^r{}_r = -\frac{2e^{-B}}{r^2}A'\phi'\dot{f} + \mathcal{O}(r-r_h)\,. \tag{14}$$

This expression is always positive-definite since, close to the horizon, $A' > 0$, and $\dot{f}\,\phi' < 0$ according to Eq. (10) for a regular horizon. Also, we find that $T^r{}_r$ is always decreasing close to $r_h$ and as a result, the novel no-hair theorem is non-applicable in our theory.

The above result opens the way for the construction of novel black-hole solutions in the context of the general theory (1). We have therefore numerically solved the system of equations (9), and determined a large number of black-hole solutions with scalar hair for a variety of forms of the coupling function $f(\phi)$: exponential, odd and even power-law, odd and even inverse-power-law. Once the form of $f(\phi)$ was chosen, the input values $(\phi_h, \phi'_h)$, with $\phi'_h$ being given by Eq. (10), always led to a regular black-hole solution with scalar hair. The scalar field and profile of $T^r{}_r$ for those solutions are depicted in Figs. 1(a,b).

Some of the characteristics of the black-hole solutions we found[12] are represented in Figs. 2(a,b), where we depict the indicative case of $f(\phi) = \alpha/\phi$. The scalar charge $D$ is a function of the black-hole mass and thus a dependent quantity; this renders

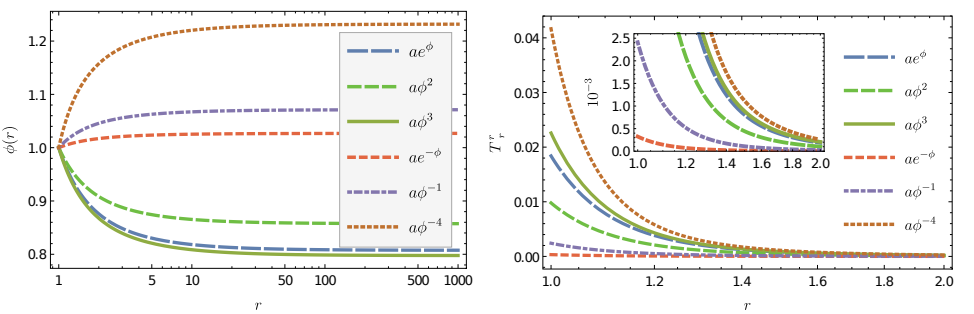



Fig. 1. The scalar field $\phi$ (left plot) and the $T^r{}_r$ component (right plot) for different coupling functions $f(\phi)$, for $a = 0.01$ and $\phi_h = 1$.

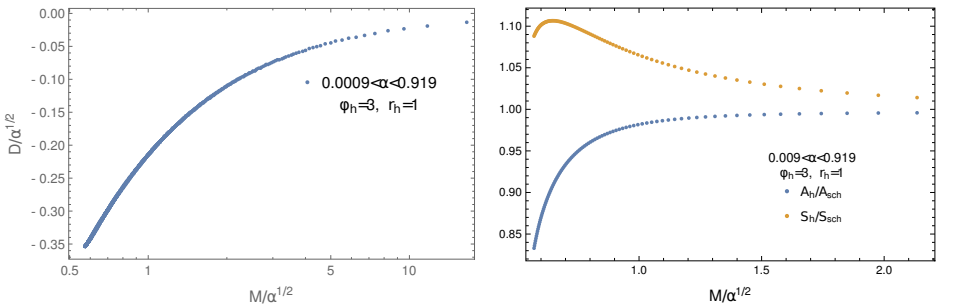



Fig. 2. The scalar charge $D$ (left plot), and the ratios $A_h/A_{Sch}$ and $S_h/S_{Sch}$ (right plot, lower and upper curve respectively) in terms of the mass $M$, for $f(\phi) = \alpha/\phi$.

the scalar hair secondary. For a large mass, the scalar charge vanishes and our black-hole solutions match the Schwarzschild solution. The horizon area is always smaller than the one of the Schwarzschild solution exhibiting also a lower value beyond which the black hole ceases to exist — the latter feature is due to the additional bound emerging from the positivity of the quantity under the square-root in Eq. (10). Its entropy is larger than that of the Schwarzschild case and thus thermodynamically more stable.

## 3. The pure scalar-Gauss-Bonnet theory

We will now investigate whether a regular black-hole solution can arise in the context of a pure scalar-GB theory, i.e. in the absence of the linear Ricci term. By ignoring all terms in the field equations related to the Ricci term, these are simplified — but can we construct again a regular horizon? If we assume as before that, as $r \to r_h$, $\phi'$ remains finite while $A'$ diverges, Eq. (6) now yields: $e^B \simeq 3 + \mathcal{O}(1/A')$; but this does not describe a black hole. We may alternatively demand that $e^B \to \infty$ instead, as $r \to r_h$; then, Eq. (6) gives: $A' \simeq r^2\phi'/8\dot{f} + \mathcal{O}(e^{-B})$. In this case, $A(r)$ is the dependent quantity, and Eqs. (5) and (7) form a system of two differential equations for $B$ and $\phi$. In the limit $r \to r_h$, we find the results[12]

$$B' = -\frac{2}{r}\, e^B + \mathcal{O}\left(e^{-B}\right), \qquad \phi'' = -\frac{e^B}{r}\,\phi' + \mathcal{O}\left(e^{-B}\right). \tag{15}$$

Upon integration, the first equation leads to the solution $e^{-B} = 2\ln(r/r_h)$, which does resemble a horizon, but the second one reveals that this horizon is not regular unless $\phi'(r_h) = 0$, an assumption that trivialises the contribution of the GB term. Alternative ansatzes for the form of the spacetime around the sought-for black hole have also failed to lead to a regular horizon in the absence of the Ricci scalar.

## 4. Conclusions

In the context of a general Einstein-scalar-GB theory, we have demonstrated that the emergence of regular black-hole solutions is a generic feature. For an arbitrary

coupling function $f(\phi)$, we were always able to construct a regular black-hole horizon as well as an asymptotically-flat solution at infinity, and to explicitly show that the novel no-hair theorem is then easily evaded. Our numerical analysis has subsequently led to a large number of regular black-hole solutions for different choices of $f(\phi)$, all characterised by a non-trivial scalar hair (for similar black-hole solutions, see also[13,14]). The study of the pure scalar-GB theory, and the failure to obtain a regular horizon, clearly demonstrates that the presence of the GB term in the theory is a necessary condition for the emergence of novel black holes but not a sufficient one as it must be supplemented by the presence of the linear Ricci term.

## Acknowledgments

This research is implemented through the Operational Program "Human Resources Development, Education and Lifelong Learning" and is co-financed by the European Union (European Social Fund) and Greek national funds.

## References

1. J. D. Bekenstein, Phys. Rev. Lett. **28** (1972) 452; C. Teitelboim, Lett. Nuovo Cim. **3S2** (1972) 397.
2. M. S. Volkov and D. V. Galtsov, JETP Lett. **50** (1989) 346; P. Bizon, Phys. Rev. Lett. **64** (1990) 2844; B. R. Greene, S. D. Mathur and C. M. O'Neill, Phys. Rev. D **47** (1993) 2242; K. I. Maeda, T. Tachizawa, T. Torii and T. Maki, Phys. Rev. Lett. **72** (1994) 450.
3. H. Luckock and I. Moss, Phys. Lett. B **176** (1986) 341; S. Droz, M. Heusler and N. Straumann, Phys. Lett. B **268** (1991) 371.
4. J. D. Bekenstein, Annals Phys. **82** (1974) 535; Annals Phys. **91** (1975) 75.
5. J. D. Bekenstein, Phys. Rev. D **51** (1995) no.12, R6608.
6. C. A. R. Herdeiro and E. Radu, Int. J. Mod. Phys. D **24** (2015) no.09, 1542014.
7. T. P. Sotiriou and V. Faraoni, Phys. Rev. Lett. **108** (2012) 081103.
8. L. Hui and A. Nicolis, Phys. Rev. Lett. **110** (2013) 241104.
9. P. Kanti, N. E. Mavromatos, J. Rizos, K. Tamvakis and E. Winstanley, Phys. Rev. D **54** (1996) 5049; Phys. Rev. D **57** (1998) 6255.
10. E. Babichev and C. Charmousis, JHEP **1408** (2014) 106.
11. T. P. Sotiriou and S. Y. Zhou, Phys. Rev. Lett. **112** (2014) 251102; Phys. Rev. D **90** (2014) 124063.
12. G. Antoniou, A. Bakopoulos and P. Kanti, Phys. Rev. Lett. **120** (2018) no.13, 131102; Phys. Rev. D **97** (2018) no.8, 084037.
13. D. D. Doneva and S. S. Yazadjiev, Phys. Rev. Lett. **120** (2018) no.13, 131103.
14. H. O. Silva *et al.*, Phys. Rev. Lett. **120** (2018) no.13, 131104.

# $f(R)$ Hořava-Lifshitz cosmologies via Noether's symmetries

Myrzakul Shynaray* and Myrzakul Tolkynay

[1] *Department of Theoretical and Nuclear Physics, al-Farabi Kazakh National University Almaty, 050040, Kazakhstan*

[2] *Department of General Theoretical Physics, L.N.Gumilyov Eurasian National University, Nur-Sultan, 010008, Kazakhstan*

** E-mail: srmyrzakul@gmail.com*

*www.kaznu.kz_myrzakul.shynaray@kaznu.kz*

Hernando Quevedo

*Department of Physics, University of Rome "La Sapienza", Rome, I-00185, Italy*

*E-mail: quevedo@nucleares.unam.mx*

We investigate the general approach to finding exact cosmological solutions in $f(R)$ Hořava-Lifshitz gravity, based on Noether's theorem. A feature of this approach is that it uses the behavior of an effective Lagrangian under infinitesimal transformations of the desired symmetry, explicitly determining the form $f(R)$ for which such symmetries exist. It is shown that the dynamics of the scale factor changes according to either a exponential function of time.

*Keywords*: Hořava-Lifshitz gravity; $f(R)$ gravity; Noether symmetry.

## 1. Introduction

At the end of the past millennium, it became clear from the observation of type Ia supernovae and the cosmic microwave background that our universe expands with acceleration [1, 2]. The explanation of this phenomenon within the framework of the general theory of relativity led to the formulation of a large number of models, one of which consists in introducing a mysterious substance, the so-called dark energy (see for example Ref. [3] and references therein for some reviews). The nature of dark energy is still not clear, but mathematically it fits well with the ΛCDM model a wide range of data [4]. Nevertheless, this model has strong theoretical flaws [5], which motivated the search for alternative models [6, 7]. One of the alternatives is the modification of the Einstein-Hilbert term, replacing the Ricci scalar $R$ in the action with some common functions $f(R)$ of the Ricci scalar (see for example Ref. [8]).

In parallel, some models for quantum gravity were developed. In 2009, Hořava [9], based on an idea proposed by Lifshitz [10], formulated a model for a theory of quantum gravity, which takes into account ultraviolet mode renormalizability, due to an anisotropic scaling between space and time, so that Lorentz invariance is violated on ultraviolet scales. However, when choosing the parameter $\lambda = 1$, the infrared limit of the theory reproduces the general theory of relativity. This modification of the general theory of relativity consists in introducing high-order terms into the Einstein-Hilbert action, which lead to different scalings and divide the coordinates into space and time. In theory, there are no ghosts, since there are only second-order

time derivatives in the action, but a clear violation of general covariance introduces a new scalar degree of freedom, due to which pathologies appear [11]. This model developed further and is known as the Hořava-Lifshitz theory (see for example Ref. [12]). From the above considerations, it is clear that the quantum theory of Hořava-Lifshitz, combined with the alternative theory $f(R)$, is a promising candidate for completing the general theory of relativity in the ultraviolet range [13, 14].

In this paper, we consider the Friedman-Lemaître-Robertson-Walker (FLRW) flat spacetime in the framework of the metric formalism $f(R)$ of gravity. We consider a general approach to constructing modified gravity, which is invariant with respect to diffeomorphisms and preserve foliation. The approach was proposed in [14], where special attention was paid to the formulation of modified $f(R)$ Hořava–Lifshitz gravity and its Hamiltonian structure. Following [15], we calculated an effective Lagrangian in which the scale factor $a$ and the Ricci scalar $R$ play the role of independent dynamic variables. This Lagrangian is constructed in such a way that its variation with respect to $a$ and $R$ gives the equations of motion of the Hořava–Lifshitz theory. The form of the function $f(R)$, appearing in the modified action, is then determined by the requirement that the Lagrangian admits the required Noether symmetry [16]. Under Noether symmetry of this cosmological model, we understand that there is a vector field $X$, which is an infinitesimal generator of a symmetry in the tangent space of the configuration space, such that the derivative of the Lagrangian along this vector field vanishes. We will see that, by requiring the Noether symmetry as a feature of the Lagrangian of the model under consideration, we can obtain the explicit form of the function $f(R)$. Since the existence of symmetry leads to constants of motion, we can integrate the field equations, which then lead to a exponential expansion for the universe.

## 2. Modified $f(R)$ Hořava-Lifshitz Gravity

In this work, we consider a more general model of the Hořava-Lifshitz gravity proposed in [14]. The action of such a model has the form

$$S_{f(R_{GHL})} = \int \mathrm{d}^4x\sqrt{g^{(3)}}Nf(R_{GHL}). \tag{1}$$

Here $g^{(3)}$ is determinant of the three-dimensional metric tensor $g_{ij}^{(3)}$ for the ADM metric given in the following form

$$ds^2 = -N^2\mathrm{d}t^2 + g_{ij}^{(3)}\left(\mathrm{d}x^i + N^i\mathrm{d}t\right)\left(\mathrm{d}x^j + N^j\mathrm{d}t\right), \tag{2}$$

where $i, j = 1, 2, 3$, $N$ is the so-called lapse variable and $N^i$ is the shift 3-vector. In the action, we use the function $f(R_{GHL})$, which denotes the generalized curvature of the Hořava-Lifshitz gravity $R_{GHL}$ and is defined as

$$R_{GHL} \equiv K^{ij}K_{ij} - \lambda K^2 + 2\mu\nabla_\mu\left(n^\mu\nabla_\nu n^\nu - n^\nu\nabla_\nu n^\mu\right) - E^{ij}\mathcal{G}_{ijkl}E^{kl}, \tag{3}$$

where $K_{ij}$ is the extrinsic curvature

$$K_{ij} = \frac{1}{2N}\left(\dot{g}^{(3)}_{ij} - \nabla^{(3)}_i N_j - \nabla^{(3)}_j N_i\right), \quad K = K^i{}_i\,, \tag{4}$$

$n^\mu$ is a unit vector perpendicular to the three-dimensional hypersurface $\Sigma_t$ defined by $t =$ constant, and $\nabla^{(3)}_i$ expresses the covariant derivative on the hypersurface $\Sigma_t$. In the last term of Eq. (3), $\mathcal{G}_{ijkl}$ is the inverse of the generalized De Witt metric

$$\mathcal{G}^{ijkl} = \frac{1}{2}(g^{(3)ik}g^{(3)jl} + g^{(3)il}g^{(3)jk}) - \lambda g^{(3)ij}g^{(3)kl}. \tag{5}$$

Here it is important to note that $\mathcal{G}^{ijkl}$ is singular for $\lambda = 1/3$ and $G_{ijkl}$ exist if $\lambda \neq 1/3$. The expression for $E_{ij}$ is constructed to satisfy the "detailed balance principle" [17] and is defined as

$$\sqrt{g^{(3)}}E^{ij} = \frac{\delta W[g^{(3)}_{kl}]}{\delta g^{(3)}_{ij}}, \tag{6}$$

where the form of $W[g^{(3)}_{kl}]$ is given in [18] for $z = 2$ and $z = 3$.

Consider a spatially flat FLRW universe

$$ds^2 = -N^2\mathrm{d}t^2 + a(t)^2 \sum_{i=1,2,3}\left(\mathrm{d}x^i\right)^2, \tag{7}$$

where $N$ can be considered as time-independent and we will fix as $N = 1$. The scalar curvature (3) can be written as

$$R_{GHL} = 3(1 - 3\lambda + 4\mu)\frac{\dot{a}^2}{a^2} + 6\mu\frac{\ddot{a}}{a}. \tag{8}$$

Following [16], in order to investigate Noether's symmetry of the model, it is necessary to determine the effective Lagrangian of the action (1) of the minisuperspace under consideration, in which the scale factor $a$ and the scalar curvature $R_{GHL}$ play the role of independent dynamical variables

$$S = \int dt\mathcal{L}(a, \dot{a}, R_{GHL}, \dot{R}_{GHL}) = \\ \int dt\Big[a^3 f(R_{GHL}) - \nu\Big\{R_{GHL} - \Big(3(1 - 3\lambda + 4\mu)\frac{\dot{a}^2}{a^2} + 6\mu\frac{\ddot{a}}{a}\Big)\Big\}\Big], \tag{9}$$

where $\nu = a^3 df(R_{GHL})/dR_{GHL}$ is a Lagrange multiplier. Then, the effective Lagrangian will have the form

$$\mathcal{L}(a, \dot{a}, R_{GHL}, \dot{R}_{GHL}) = (9\lambda - 3)\dot{a}^2 a f' + 6\mu\dot{a}\dot{R}_{GHL}a^2 f'' + a^3(f'R_{GHL} - f). \tag{10}$$

The equations of motion will have then the following form

$$3H^2 + 2\dot{H} = -\frac{2}{3\lambda - 1}\frac{1}{f'}\Big[\mu f'''\dot{R}^2_{GHL} + \mu f''\ddot{R}_{GHL} + \\ +(3\lambda - 1)f''H\dot{R}_{GHL} + \frac{1}{2}\left(f - R_{GHL}f'\right)\Big]. \tag{11}$$

Also, we have the zero energy condition associated with the above Lagrangian as

$$H^2 = \frac{1}{3(3\lambda-1)}\frac{1}{f'}\left[f'R_{GHL} - f - 6\mu\dot{R}_{GHL}Hf''\right]. \tag{12}$$

Knowing the Lagrangian of the generalized Hořava-Lifshitz gravity, it is possible to determine the Noether symmetry.

## 3. The Noether symmetries in $f(R_{GHL})$ theory of gravity

Here, our aim is to find the function $f(R_{GHL})$ such that the corresponding Lagrangian exhibits the desired symmetry. Following [16], we define the Noether symmetry induced on the model by a vector field $X$ on the tangent space $TQ = \left(a, \dot{a}, R, \dot{R}_{GHL}\right)$ of the configuration space $Q = (a, R_{GHL})$ of the Lagrangian (10) through

$$X = \alpha\frac{\partial}{\partial a} + \beta\frac{\partial}{\partial R_{GHL}} + \frac{d\alpha}{dt}\frac{\partial}{\partial \dot{a}} + \frac{d\beta}{dt}\frac{\partial}{\partial \dot{R}_{GHL}}, \tag{13}$$

such that the Lie derivative of the Lagrangian with respect to this vector field vanishes

$$L_X\mathcal{L} = 0. \tag{14}$$

In Eq.(13), $\alpha$ and $\beta$ are functions of $a$ and $R_{GHL}$ and $\frac{d}{dt}$ represents the Lie derivative along the dynamical vector field, that is,

$$\frac{d}{dt} = \dot{a}\frac{\partial}{\partial a} + \dot{R}_{GHL}\frac{\partial}{\partial R_{GHL}}. \tag{15}$$

Here we substitute the expressions for the effective Lagrangian (10) and combine the coefficients in front of $\dot{a}^2, \dot{R}^2_{GHL}, \dot{a}\dot{R}_{GHL}$. Equating to zero the resulting expression, we obtain the following equations

$$3(3\lambda-1)\left(\alpha + 2a\alpha_a\right)f' + \left[3(3\lambda-1)\beta a + 6\mu a^2\beta_a\right]f'' = 0, \tag{16}$$

$$6\mu a^2\alpha_{R_{GHL}}f'' = 0, \tag{17}$$

$$6\mu\left(2a\alpha + a^2\alpha_a\right)f'' + 6(3\lambda-1)a\alpha_{R_{GHL}}f' + 6\mu a^2\left(\beta f''' + \beta_{R_{GHL}}f''\right) = 0. \tag{18}$$

Then, we collect the remaining free member

$$3\alpha a^2(f'R - f) + \beta a^3 Rf'' = 0. \tag{19}$$

Now, our task is to solve the system of equations (16) - (19) in order to find the cosmological parameters that describe the dynamics of the universe in the framework of the $f(R_{GHL})$ Hořava-Lifshitz gravity. From Eq.(17), it is clear that two cases must be considered: $f'' = 0$ and $\frac{d\alpha}{dR_{GHL}} = 0$, but the solution for $f'' = 0$ has no physical meaning; therefore, we consider only $\frac{d\alpha}{dR_{GHL}} = 0$, then

$$\alpha(a) = \alpha_0 a^{\frac{\beta_0}{\alpha_0}+1}, \beta(a, R) = \beta_0 a^{\frac{\beta_0}{\alpha_0}}R, \tag{20}$$

and

$$f = f_0 R^{-3\frac{\alpha_0}{\beta_0}}, \tag{21}$$

where $\alpha_0 = \frac{2\mu}{3\lambda-6\mu-1}\beta_0$, $f_0, \beta_0$ are integral constants.

## 4. Cosmological solutions

In this section, to describe the dynamics of the universe, we solve analytically the field equations (11) - (12). To this end, we need to find the explicit dependence of the scale factor $a$ in terms of the time $t$. We rewrite Eqs. (11)–(12) as follows

$$3H^2 + 2\dot{H} = -\frac{4\mu(3\lambda-3\mu-1)}{(3\lambda-6\mu-1)^2}\frac{\dot{R}^2_{GHL}}{R^2_{GHL}} + \frac{2\mu}{3\lambda-6\mu-1}\frac{\ddot{R}_{GHL}}{R_{GHL}} + \frac{2(3\lambda-1)}{3\lambda-6\mu-1}\frac{\dot{R}_{GHL}}{R_{GHL}}H + \frac{1}{6\mu}R_{GHL}, \tag{22}$$

$$H^2 = \frac{R_{GHL}}{18\mu} - \frac{2\mu}{1-3\lambda+6\mu}\frac{\dot{R}_{GHL}}{R_{GHL}}H\ . \tag{23}$$

To solve this system we obtain

$$H = \sqrt{\frac{C_1}{C_2}}\tanh(\sqrt{C_1C_2}t), \tag{24}$$

where $C_1 = \frac{Z}{3\mu+3\lambda-1}$, $C_2 = \frac{(3\lambda-1)(3\mu-3\lambda+1)}{2\mu(3\mu+3\lambda-1)}$, or, equivalently,

$$a = a_0\left(e^{\sqrt{C_1C_2}t} + e^{-\sqrt{C_1C_2}t}\right)^{\frac{1}{C_2}}\ . \tag{25}$$

Thus, we have found a general solution to the modified $f(R)$ Hořava-Lifshitz gravity theory. In general, it represents an expanding cosmological model with a scale factor with grows exponentially.

## 5. Conclusion

In this work, we analyzed the $f(R)$ Hořava-Lifhitz gravity model, which is a modification of the original Hořava-Lifhitz model for quantum gravity. We consider the spatially flat FLRW line element to find out what kind of cosmological scenarios are allowed in this modified gravity. Since usually it is very difficult to find analytic solutions to the field equations, we use here the alternative method, which is based on the analysis of the Noether symmetries of a particular effective Lagrangian. In this work, we derive the effective Lagrangian, from the corresponding field equations for the $f(R)$ Hořava-Lifhitz theory in the presence of spatially flat, isotropic and homogeneous line element.

It turns out that the field equations can be represented as a set of two second-order ordinary differential equations, which can be expressed as one differential equation for different variables so that it can be solved by using the standard method

of separation of variables. This simple representation allows us to integrate explicitly the scale factor as well as the Ricci scalar. Both quantities are then explicit functions of time.

The results presented in this work show that the method of Noether symmetries can be applied also in the case of the modified $f(R)$ Hořava-Lifhitz gravity model to derive cosmological solutions. The resulting functions for the scale factor show that the corresponding universe expands either exponentially. This scenario is possible in relativistic cosmology so that, in principle, we could compare our results with observational data from different epochs of the universe evolution. This could be used to set limits on the values of the parameters that enter the Hořava-Lifhitz action.

## Acknowledgments

This work was partially supported by UNAM-DGAPA-PAPIIT, Grant No. 111617, and by the Ministry of Education and Science of RK, Grant No. BR05236322 and AP05133630.

## References

1. A. G. Riess, *et. al.*, *Astron. J.* **116** 1009 (1998).
   S. Perlmutter, *et. al.*, *Astrophys. J.* **517** 565 (1999).
2. C. L. Bennett, *et. al.*, *Astrophys. J. Suppl.* **148** 1 (2003).
3. S. Nojiri and S. D. Odintsov, *Int. J. Geom. Meth. Mod. Phys.* **4** 115 (2007).
4. U. Seljak, *et.al.*, *Phys. Rev. D* **71** 103515 (2005).
5. S.M. Carroll , W.H. Press , E.L. Turner, *J.Ann. Rev. Astron. Astroph.* **30** 499 (1992).
6. P.J.E.Peebles, B.Rathra , *Rev. Mod. Phys.* **75**, 559 (2003).
7. E.J. Copeland, M. Sami, S. Tsujikawa, *Int. J. Mod. Phys.* **D 15**, 1753 (2006).
8. M. Calzà, M.Rinaldi, L.Sebastiani *L. Eur. Phys. J. C* **78**178 (2018).
9. P. Horava, *Phys. Rev. D* **79**, 084008 (2009).
10. P. Horava, *JHEP* **0903**, 020 (2009).
    P. Horava, *Phys. Lett. B* **694**, 172-176 (2010).
11. R. G. Cai, B. Hu and H. B. Zhang, *Phys. Rev. D* **80**, 041501 (2009).
12. G.E. Volovik,*JETP Lett.* **89**, 525-528 (2009).
13. J. Kluson, *Phys. Rev. D* **81**, 064028 (2010).
14. M. Chaichian, S. Nojiri, S. D. Odintsov, M. Oksanen and A. Tureanu, *Class. Quant. Grav.* **27**, 185021 (2010).
15. S. Capozziello and A. De Felice, *JCAP* **0808**, 016 (2008).
16. B. Vakili, N.Khosravi and H. R. Sepangi, *Class. Quantum Grav.* **24** 931 (2007).
17. A. De Felice and S. Tsujikawa, *Living Rev.Rel.* **13**, 3 (2010).
18. G.J. Olmo, *Phys. Rev. D* **75**, 023511 (2007).
19. Fourteenth Marcel Grossmann Meeting (2015), `http://www.icra.it/MG/mg14/`.

# Extended mimetic gravity: Hamiltonian analysis and gradient instabilities

Kazufumi Takahashi* and Tsutomu Kobayashi

*Department of Physics, Rikkyo University, Toshima, Tokyo 171-8501, Japan*
**E-mail: takahashik@rikkyo.ac.jp*

We propose a novel class of degenerate higher-order scalar-tensor theories as an extension of mimetic gravity. By performing a noninvertible conformal transformation on "seed" scalar-tensor theories which may be nondegenerate, we can generate a large class of theories with at most three physical degrees of freedom. We identify a general seed theory for which this is possible. Cosmological perturbations in these extended mimetic theories are also studied. It is shown that either of tensor or scalar perturbations is generically plagued with ghost/gradient instabilities. See Ref. 1 for more details.

*Keywords*: Scalar-tensor theories; mimetic gravity; cosmological perturbations.

## 1. Introduction

When one constructs a field theory with higher derivatives, a guiding principle comes from the theorem of Ostrogradsky[2], which states that any theory described by a nondegenerate higher derivative Lagrangian has unstable extra degrees of freedom (DOFs), i.e., Ostrogradsky ghost. Therefore, a theory without Ostrogradsky ghost, often referred to as a *healthy* theory, must have a degenerate Lagrangian.

Within (single-field) scalar-tensor theories in four dimensions, the Horndeski theory[3] (or its equivalent formulation known as generalized Galileons[4,5]) provides a basic ground for studying a wide class of such healthy theories having three DOFs, since it is the most general scalar-tensor theory that yields second-order Euler-Lagrange equations. There are further possibilities of healthy theories beyond the Horndeski class, such as Gleyzes-Langlois-Piazza-Vernizzi (GLPV) theories[6] and quadratic/cubic degenerate higher-order scalar-tensor (DHOST) theories[7–9]. Those quadratic/cubic DHOST theories form the broadest class of healthy scalar-tensor theories known so far. However, these theories are obtained under the assumption that the Lagrangian depends on up to quadratic/cubic order in $\nabla_\mu\nabla_\nu\phi$ (hence the name "quadratic/cubic DHOST"), and thus the very boundary of healthy scalar-tensor theories remains unclear.

In light of this situation, we explore the possibility to construct a new class of DHOST theories by use of conformal/disformal transformation of the metric, which has been employed for investigating the relation between the known scalar-tensor models. To this end, we perform a *noninvertible* conformal transformation on generic scalar-tensor theories that could possess an unwanted extra DOF. Here, the noninvertibility of the transformation is crucial because otherwise the resultant theory could also have an extra DOF[10]. Since the formulation of the new theory

can be regarded as an extension of the mimetic gravity model introduced in Ref. 11, we call the theory *extended mimetic gravity*. We show that these models have at most three DOFs based on Hamiltonian analysis and thus form a broad class of DHOST theories, most of which lie outside the quadratic/cubic DHOST class. We also study the linear stability of cosmological perturbations in our extended mimetic gravity and show that the models obtained in the above explained way generically exhibit the problem of ghost/gradient instabilities.

## 2. Extended Mimetic Gravity

We start from the following general scalar-tensor theory:

$$S_{\rm seed} = \int d^4x\sqrt{-g}\left[f_2(\phi, X)\mathcal{R} + f_3(\phi, X)\mathcal{G}^{\mu\nu}\phi_{\mu\nu} + F(g_{\mu\nu}, \phi, \phi_\mu, \phi_{\mu\nu})\right], \quad (1)$$

with $\mathcal{R}$ and $\mathcal{G}_{\mu\nu}$ being the four-dimensional Ricci scalar and Einstein tensor, $\phi_\mu \equiv \nabla_\mu\phi$, $\phi_{\mu\nu} \equiv \nabla_\mu\nabla_\nu\phi$, and $X \equiv \phi^\mu\phi_\mu$. Note that, for a generic choice of the functions $f_2, f_3$, and $F$, the theory yields Ostrogradsky ghost. Using the action of the form (1) as a seed, we perform the following conformal transformation:

$$g_{\mu\nu} \;\rightarrow\; \tilde{g}_{\mu\nu} = -Xg_{\mu\nu}, \quad (2)$$

Here, $\tilde{g}_{\mu\nu}$ is identified as the metric in the original frame (1), while $g_{\mu\nu}$ is now the metric of the new theory, which we call *extended mimetic gravity*. The transformation (2) is noninvertible as the right-hand side is invariant under conformal transformation of $g_{\mu\nu}$. As a result, the new theory acquires a local conformal symmetry.

To study whether the so-obtained extended mimetic gravity models possess a problematic extra DOF or not, one needs to perform a Hamiltonian analysis. In doing so, we write the action in the ADM language, which becomes of the form

$$S_{\rm EMG} = \int dtd^3x\left[N\sqrt{\gamma}L_{\rm EMG} + \Lambda(NA_* + N^iD_i\phi - \dot{\phi})\right], \quad (3)$$

where $L_{\rm EMG}$ is some function of $(\gamma_{ij}, R_{ij}, \phi, A_*; V_{ij}; D_i)$. Here, we have defined

$$V_{ij} \equiv K_{ij} + \frac{\dot{A}_* - D^i\phi D_iN - N^iD_iA_*}{NA_*}\gamma_{ij}, \quad (4)$$

and introduced an auxiliary variable $A_*$ with a Lagrange multiplier $\Lambda$ so that second-order time derivatives do not appear explicitly in the action. Reflecting the aforementioned local conformal symmetry, $\dot{\gamma}_{ij}$ and $\dot{A}_*$ (namely, $\ddot{\phi}$) appear only in a special combination $V_{ij}$. When one proceeds to a Hamiltonian analysis, this relation leads to an additional primary constraint, which turns out to be first class. Thus, it kills the otherwise existing unwanted DOF, leaving only three DOFs. For a more rigorous and detailed analysis, see Ref. 1.

## 3. Cosmological Perturbations

To see whether the general mimetic theories are phenomenologically viable or not, let us analyze perturbations around the Friedmann-Lemaître-Robertson-Walker (FLRW) background in the mimetic gravity models.

For simplicity, we take the unitary gauge to write $\phi = t$ and impose the constraint $X = -1$ which fixes the conformal gauge DOF. As a consequence, any function of $(\phi, X)$ can be regarded as a function of $t$ only. We also have $N = 1$ since $X$ is written in terms of $N$ as $X = -1/N^2$ in the unitary gauge. Under this setup, the action can be recast in the following form:

$$S_{\rm EMG} = \int dtd^3x\sqrt{\gamma}\left[\left(f_2 - \frac{1}{2}\dot{f}_3\right)R + \mathcal{F}(t, \mathcal{K}_1, \mathcal{K}_2, \mathcal{K}_3, \cdots, \mathcal{K}_\ell)\right], \tag{5}$$

where $\mathcal{F}$ is some function of $t$ and $\mathcal{K}_n \equiv K^{i_1}_{i_2}K^{i_2}_{i_3}\cdots K^{i_n}_{i_1}$ $(n = 1, 2, \cdots, \ell)$. It is useful to define the first and second derivatives of $\mathcal{F}$ respectively as

$$\mathcal{F}_n \equiv \frac{\partial\mathcal{F}}{\partial\mathcal{K}_n}, \quad \mathcal{F}_{mn} \equiv \frac{\partial^2\mathcal{F}}{\partial\mathcal{K}_m\partial\mathcal{K}_n}. \tag{6}$$

Now we substitute the following metric ansatz to the action (5),

$$N = 1, \quad N_i = \partial_i\chi, \quad \gamma_{ij} = a^2e^{2\zeta}\left(\delta_{ij} + h_{ij} + \frac{1}{2}h_{ik}h_{jk} + \cdots\right), \tag{7}$$

where $\chi$ and $\zeta$ are scalar perturbations and $h_{ij}$ denotes a transverse-traceless tensor perturbation. The background EOM is given by

$$\dot{\mathcal{P}} + 3H\mathcal{P} - \mathcal{F} = 0, \quad \mathcal{P} \equiv \sum_{n=1}^{\ell} n\,H^{n-1}\mathcal{F}_n, \tag{8}$$

where $H \equiv \dot{a}/a$ is the Hubble parameter and $\mathcal{F}_n$ are evaluated at the background, $\mathcal{K}_n = 3H^n$. This equation will be used to simplify the expressions of the quadratic actions for the tensor and scalar perturbations below.

The quadratic action for the tensor perturbation $h_{ij}$ is given by

$$S^{(2)}_{\rm T} = \int dtd^3x\frac{a^3}{4}\left[\mathcal{B}\dot{h}^2_{ij} - \mathcal{E}\frac{(\partial_k h_{ij})^2}{a^2}\right], \tag{9}$$

where

$$\mathcal{B} \equiv \sum_{n=2}^{\ell}\frac{n(n-1)}{2}H^{n-2}\mathcal{F}_n, \quad \mathcal{E} = f_2 - \frac{1}{2}\dot{f}_3. \tag{10}$$

The tensor perturbations are stable if $\mathcal{B} > 0$ and $\mathcal{E} > 0$.

The quadratic action for the scalar perturbations $\zeta$ and $\chi$ is

$$S^{(2)}_{\rm S} = \int dtd^3xa^3\left[\frac{3}{2}(3\mathcal{A} + 2\mathcal{B})\dot{\zeta}^2 + 2\mathcal{E}\frac{(\partial_k\zeta)^2}{a^2}\right.$$
$$\left. + \frac{1}{2}(\mathcal{A} + 2\mathcal{B})\left(\frac{\partial^2\chi}{a^2}\right)^2 - (3\mathcal{A} + 2\mathcal{B})\dot{\zeta}\frac{\partial^2\chi}{a^2}\right], \tag{11}$$

where

$$\mathcal{A} \equiv \sum_{m=1}^{\ell} \sum_{n=1}^{\ell} m\, n\, H^{m+n-2} \mathcal{F}_{mn}. \tag{12}$$

Provided that $\mathcal{A} + 2\mathcal{B} \neq 0$, $\chi$ can be eliminated from the action by use of its EOM. Then, we are left with the following quadratic action for curvature perturbation:

$$S_{\rm S}^{(2)} = 2 \int dt d^3x\, a^3 \left[ \frac{\mathcal{B}(3\mathcal{A} + 2\mathcal{B})}{\mathcal{A} + 2\mathcal{B}} \dot{\zeta}^2 + \mathcal{E} \frac{(\partial_k \zeta)^2}{a^2} \right]. \tag{13}$$

Written in this form, one notices that the stability condition for the tensor perturbations, $\mathcal{E} > 0$, is not compatible with the stability for the scalar perturbation, $\mathcal{E} < 0$. This indicates that either of the tensor or scalar perturbations exhibits gradient instabilities, even if one circumvents ghosts by choosing the coefficients in front of the time derivative terms in Eqs. (9) and (13) to be positive.

One would notice that if $\mathcal{B}(\mathcal{A}+2\mathcal{B})(3\mathcal{A}+2\mathcal{B}) = 0$ then the scalar perturbations appear to be nondynamical. This is indeed the case if we choose a Horndeski (or GLPV) theory as seed, where $\mathcal{A} + 2\mathcal{B} = 0$. In this case, it is important to take into account the presence of matter fields other than $\phi$ to discuss the viability of mimetic cosmology. If another scalar field $\psi$ is added to the seed Lagrangian, one can show that the scalar perturbations revive, i.e., there are now two scalar DOFs, and one of them is a ghost.

Thus, we have established that all the mimetic gravity models with three DOFs obtained so far are plagued with ghost/gradient instabilities on a cosmological background (except for the special case of strongly-coupled scalar perturbations). More detailed discussion is found in Ref. 1 (see also Ref. 12).

## References


1. K. Takahashi and T. Kobayashi, *JCAP* **1711**, 038 (2017).
2. R. P. Woodard, *Scholarpedia* **10**, 32243 (2015).
3. G. W. Horndeski, *Int. J. Theor. Phys.* **10**, 363 (1974).
4. C. Deffayet, X. Gao, D. A. Steer, and G. Zahariade, *Phys. Rev. D* **84**, 064039 (2011).
5. T. Kobayashi, M. Yamaguchi, and J. Yokoyama, *Prog. Theor. Phys.* **126**, 511 (2011).
6. J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, *Phys. Rev. Lett.* **114**, 211101 (2015).
7. D. Langlois and K. Noui, *JCAP* **1602**, 034 (2016).
8. M. Crisostomi, K. Koyama, and G. Tasinato, *JCAP* **1604**, 044 (2016).
9. J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, *JHEP* **12**, 100 (2016).

10. K. Takahashi, H. Motohashi, T. Suyama, and T. Kobayashi, *Phys. Rev. D* **95**, 084053 (2017).
11. A. H. Chamseddine and V. Mukhanov, *JHEP* **11**, 135 (2013).
12. D. Langlois, M. Mancarella, K. Noui, and F. Vernizzi, `arXiv:1802.03394 [gr-qc]`.

# Solutions of gravitational field equations for weakly birefringent spacetimes

Nils Alex

*Department of Physics, Friedrich-Alexander University Erlangen-Nuremberg, Staudtstr. 7, 91058 Erlangen, Germany*
*E-mail: nils.alex@fau.de*

The constructive gravity programme applied to electrodynamics with vacuum birefringence yields the — up to unknown gravitational constants — unique compatible gravity theory for the underlying non-metric geometry. Starting from a perturbative variant of this procedure, we solve the resulting gravitational field equations for the geometry perturbation and point out the non-metric refinements to the linear Schwarzschild and the gravitational wave solution.

*Keywords*: Constructive gravity; pre-metric electrodynamics; modified gravity; gravitational waves; massive gravity

## 1. Introduction

Equations of motion for a matter field (e.g. an electromagnetic potential) are always *incomplete* in the sense that the coefficients in these equations, which are functions of the geometric field to which the matter field couples, are not determined. General relativity, in its essence, *closes* Maxwell's equations by providing dynamic equations for the metric tensor. The gravitational closure programme provides a prescription for performing this closure for any kind of matter theory subject to a few conditions.[1] Applying a perturbative variant of this procedure to a certain generalization of Maxwell electrodynamics which does not exclude vacuum birefringence *a priori* has been shown to yield gravitational field equations that predict the dynamics of these birefringent perturbations.[2]

## 2. Birefringent electrodynamics

The most general theory for an electrodynamic potential has been derived by Hehl and Obukhov in Ref. 3 from the principles of conservation of electric charge and magnetic flux and a linear constitutive law. Instead of the usual metric tensor $g$ from Maxwell electrodynamics, the potential couples to a rank 4 tensor field $G$ with symmetries

$$0 = G^{(ab)cd} = G^{ab(cd)} = G^{abcd} - G^{cdab} \tag{1}$$

via the action

$$S[A, G] = \int \mathrm{d}^4x\, \omega_G\, G^{abcd} F_{ab} F_{cd}, \tag{2}$$

where $F = \mathrm{d}A$ denotes the electromagnetic field strength and $\omega_G$ a 1-density built from $G$. This theory is called *General Linear Electrodynamics (GLED)*. Maxwell

electrodynamics is recovered by setting $G^{abcd} = g^{ac}g^{bd} - g^{ad}g^{bc} - \sqrt{-g}\,\epsilon^{abcd}$ and $\omega_G = \frac{1}{24}\epsilon_{abcd}G^{abcd} = \sqrt{-g}$.

A very distinctive property of GLED is the causality of the field equations: While wave covectors in the WKB limit of Maxwell electrodynamics are selected by the causality condition $0 = P(k) = g(k,k)$, which is a homogeneous polynomial of degree two, the causality condition of GLED is of degree four.[3] This allows for more roots of $P$ which introduces the possibility of a polarization-dependent speed of light — birefringence in vacuo.

## 3. Gravitational closure of weakly birefringent electrodynamics

While the gravitational closure procedure is fully laid out, it remains very difficult to follow through because the problem boils down to a set of infinitely many homogeneous first-order PDEs. There is, however, a way around this obstacle: *Perturbative* gravitational closure, which is only concerned with gravitational dynamics of small perturbations of the geometric field. This is also the sector of great phenomenological relevance, because corrections to Einstein gravity and field theories coupling to the geometry should — if present at all — be rather minute.

In Ref. 2, the perturbative gravitational closure procedure has been applied to perturbations $H$ of the GLED geometry $G$,

$$G^{abcd} = \eta^{ac}\eta^{bd} - \eta^{ad}\eta^{bc} - \epsilon^{abcd} + H^{abcd}, \tag{3}$$

which results in a 11-parameter family of field equations for the perturbation $H$, where $H$ has been gauge-fixed and decomposed according to a $3+1$ split into a set of spatial fields (greek indices being spatial indices from 1 to 3, raised and lowered at will with the spatial part $\gamma$ of $\eta$):

**transverse traceless (tt) tensor modes** $U_{\alpha\beta}, V_{\alpha\beta}, W_{\alpha\beta}$ satisfying

$$\begin{aligned} 0 = U_{[\alpha\beta]}, \quad 0 = \partial_\alpha U^{\alpha\beta} \quad \text{and} \quad 0 = \gamma_{\alpha\beta}U^{\alpha\beta}, \\ V_{\alpha\beta}, W_{\alpha\beta} \quad \text{similar}, \end{aligned} \tag{4}$$

**solenoidal vector modes** $U_\alpha, W_\alpha, B_\alpha$ satisfying

$$0 = \partial_\alpha U^\alpha, \quad W_\alpha, B_\alpha \quad \text{similar}, \tag{5}$$

**scalar modes** $V, W, \widetilde{U}, \widetilde{V}, A$.

Einstein gravity corresponds to the sector where only $U_{\alpha\beta}, B_\alpha, \widetilde{U}$, and $A$ are non-zero. In the following sections, we will solve the field equations for two prototypical matter distributions: The inertial point mass, which yields corrections to

the linear Schwarzschild solution of Einstein gravity, and the vacuum, which reveals the causality of the theory.

## 4. Inertial point mass source

An inertial point mass resting at the spatial origin of a coordinate chart is described by the world line

$$\gamma^a(\tau) = \tau\delta^a_0 \tag{6}$$

and produces a contribution of $-M\delta^{(3)}(x)$ in the equation of motion which corresponds to the variation of the action with respect to $A$. Since the matter distribution (6) is stationary, we may restrict our attention to stationary solutions for the gravitational field by setting all time derivatives to zero. In this setup, the field equations for tt and solenoidal vector modes are homogeneous and yield

$$0 = T_{\alpha\beta} \quad \text{for all tt modes,} \tag{7}$$

$$0 = V_\alpha \quad \text{for all solenoidal vector modes.} \tag{8}$$

The field equations for the scalar modes read

$$\begin{aligned}
0 &= \sigma_1\Delta V + \sigma_2\widetilde{U} + \sigma_4\widetilde{V} - 4\sigma_2 A,\\
0 &= \sigma_4\Delta V + \sigma_5 V + \sigma_6\Delta W + \sigma_7 W + (-3\sigma_1+\sigma_2)\widetilde{U} + (-9\sigma_1+3\sigma_2)\widetilde{V}\\
&\quad + (-12\sigma_1+\sigma_2)A,\\
0 &= \sigma_6\Delta V + \sigma_7 V + \sigma_8\Delta W - \sigma_5 W,\\
0 &= (-2\sigma_1 + \frac{2}{3}\sigma_2 - \frac{2}{9}\sigma_3)\Delta\Delta V - \frac{2}{3}\sigma_3\Delta\widetilde{U} + \sigma_9\Delta\widetilde{V} + \sigma_{10}\widetilde{V}\\
&\quad + (4\sigma_2 + \frac{4}{3}\sigma_3)\Delta A,\\
-M &= (-4\sigma_1 + \frac{8}{3}\sigma_2)\Delta\Delta V + 4\sigma_2\Delta\widetilde{U} + (6\sigma_2+2\sigma_3)\Delta\widetilde{V},
\end{aligned} \tag{9}$$

where the constants $\sigma_i$ denote combinations of the 11 independent gravitational constants. The general solution of this system admits the form

$$\begin{aligned}
V(x) &= \frac{M}{4\pi r}[G_1 + G_2 f_\nu(r) + G_3 f_\rho(r)]\\
W(x) &= \frac{M}{4\pi r}[G_4 + G_5 f_\nu(r) + G_6 f_\rho(r)]\\
\widetilde{U}(x) &= \frac{M}{4\pi r}[G_7 + G_8 f_\mu(r) + G_9 f_\nu(r) + G_{10} f_\rho(r)]\\
\widetilde{V}(x) &= \frac{M}{4\pi r}[G_{11} f_\mu(r)]\\
A(x) &= \frac{M}{4\pi r}[G_{12} + G_{13} f_\mu(r) + G_{14} f_\nu(r) + G_{15} f_\rho(r)]
\end{aligned} \tag{10}$$

with functions

$$f_\alpha(r) = \begin{cases} \cos(\sqrt{-\alpha}r) & \alpha < 0 \\ \mathrm{e}^{-\sqrt{\alpha}r} & \alpha \geq 0 \end{cases} \tag{11}$$

and constants $G_i, \mu, \nu, \rho$ depending on the 11 gravitational constants.[a]

The solutions (10) can be regarded as corrections to the linearized Schwarzschild solution of Einstein gravity, which in the context of GLED reads

$$V = 0, \quad W = 0, \quad \widetilde{V} = 0, \quad \widetilde{U}(x) = 4\kappa\frac{M}{4\pi r}, \quad \text{and} \quad A(x) = \kappa\frac{M}{4\pi r}. \tag{12}$$

If (10) shall only introduce *short-ranging* Yukawa corrections — instead of *long-ranging* Coulomb corrections —, the constants $G_1$ and $G_4$ must vanish and the ratio $G_7/G_{12}$ must be equal to 4. These conditions are satisfied by only one condition on the 11 independent gravitational constants and greatly simplify the form of (10) to

$$\begin{aligned} V(x) &= 0, \\ W(x) &= 0, \\ \widetilde{U}(x) &= \frac{M}{4\pi r}[\alpha + \beta\mathrm{e}^{-\sqrt{\mu}r}], \\ \widetilde{V}(x) &= \frac{M}{4\pi r}\gamma\mathrm{e}^{-\sqrt{\mu}r}, \\ A(x) &= \frac{M}{4\pi r}[\frac{1}{4}\alpha - \frac{1}{4}(\beta + 3\gamma)\mathrm{e}^{-\sqrt{\mu}r}]. \end{aligned} \tag{13}$$

Arriving at (13) concludes — perturbatively and only for this special matter distribution — the gravitational closure programme for a theory of electrodynamics with vacuum birefringence. The result is a *prediction* of where and how non-metric refinements are produced by an inertial point mass, depending on 4 unknown gravitational constants. The ramifications of such refinements have been investigated both on the astronomic scale for e.g. galaxy rotation curves[4] as well as on the quantum scale for quantum electrodynamics[5]. Experimental tests based upon these studies could now, in principle, provide approximations or bounds for the unknown parameters in (13).

## 5. Vacuum solutions

In the case of a spacetime with no matter content, the gravitational field equations are a system of coupled homogeneous wave equations with mass and damping terms. A very critical property of the solutions of this system can be examined by calculating the complex eigenvalues of the time evolution: the stability of solutions

[a] This is the result for one of two cases which have to be considered, depending on the exact values of the gravitational constants. The solution in the other case contains products of exponential and trigonometric functions of $r$, but the discussion below and especially the result (13) remain the same.

with respect to small variations of the initial data. Stability in this sense is guaranteed if and only if the real part of each eigenvalue is non-positive. Because we are interested in investigating the sector of the theory which produces only small phenomenological refinements to Einstein gravity, we implement this stability conditions and find that the remaining 10 gravitational constants must satisfy two additional conditions.

Furthermore, for propagating gravitational fields it is desirable that the causality condition in the WKB limit coincides with the Minkowski causality condition $\eta(k,k)=0$. Otherwise, consistent co-evolution of gravitational and electrodynamic fields as proposed by the gravitational closure programme would not be possible. This imposes one further condition on the gravitational constants. With this final condition implemented, we arrive at a 7-parameter family of gravitational field equations wich in vacuo simplify to

**one massless wave equation**

$$0=\Box U_{\alpha\beta}, \tag{14}$$

**massive wave equations**

$$\begin{aligned}
0&=\Box V_{\alpha\beta}+mV_{\alpha\beta}=\Box W_{\alpha\beta}+mW_{\alpha\beta},\\
0&=\Box U_{\alpha}+mU_{\alpha}=\Box W_{\alpha}+mW_{\alpha},\\
0&=\Box V+mV=\Box W+mW,\\
0&=\Box\widetilde{V}+m_{\mathrm{S}}\widetilde{V},
\end{aligned} \tag{15}$$

**constraint equations**

$$\begin{aligned}
B_{\alpha}&=-\frac{1}{2}\dot{U}_{\alpha},\\
\widetilde{U}&=-\frac{1}{3}\Delta V-\frac{\sigma+3}{2}\widetilde{V},\\
A&=\frac{1}{4}\ddot{V}+\frac{1}{4}\frac{\sigma-3}{2}\widetilde{V},
\end{aligned} \tag{16}$$

with three combinations of gravitational constants: a mass $m$, a second mass $m_{\mathrm{S}}$ for the scalar mode $\widetilde{V}$, and $\sigma$. Comparing this result with the vacuum equation for Einstein gravity,

$$0=\Box U_{\alpha\beta}\quad+\text{constraint equations}, \tag{17}$$

we immediately see that the metrically inducible mode $U_{\alpha\beta}$ still obeys a *massless* wave equation, while all non-metric modes evolve according to *massive* wave equations. This fact has a great impact on the phenomenology of the non-metric corrections: In e.g. binary systems, massive modes are only produced above a certain threshold[6] and disperse during their propagation. That is, both the *generation* as well as the *detection* of these modes will be suppressed, depending on the exact values of the gravitational constants.

It is remarkable that, in addition to the one metric and two non-metric tt modes, also *solenoidal vector modes* and *scalar modes* propagate. This introduces new ways a gravitational wave can affect test matter, which is seen by calculating the geodesic deviation of dust spheres: a tt mode propagating in $z$-direction deforms geodesic spheres in $x$- and $y$-direction in a volume preserving way. A trace-free scalar mode propagating in the same direction, on the other hand, would lead to volume-preserving oscillations in $z$- and both lateral directions — exhibiting qualitatively completely new behaviour, which is quantitatively suppressed by the unknown mass. Pure trace modes also introduce a new quality as they make the whole volume of a geodesic sphere oscillate.

*How* these modes can be generated has been studied in Ref. 6 for non-gravitationally bound systems, because emission of gravitational waves from a gravitationally bound system is a second order effect which requires knowledge of the gravitational theory to this order. Currently, two different approaches[7] for obtaining the second order of the gravitational theory consistent with GLED are being pursued. Their results will fill this gap and eventually allow to answer the question: What would the non-metric refinements to emission, propagation and detection of gravitational waves be if the vacuum was birefringent?

## References

1. M. Düll, F.P. Schuller, N. Stritzelberger and F. Wolz, Phys. Rev. D 97, 084036 (2018).
2. N. Alex, J. Schneider, F.P. Schuller, N. Stritzelberger and F. Wolz, in preparation.
3. F.W. Hehl and Yu.N. Obukhov, Progress in Mathematical Phyics 33, 430 (2003).
4. H.M. Rieser, in preparation.
5. S. Grosse-Holz, F.P. Schuller and R. Tanzi, `arXiv:1703.07183 [hep-ph]`.
6. M. Möller, Master's thesis (2018).
7. N. Alex and T. Reinhart, in preparation.

# Refinement of Friedmann equations implied by electromagnetic birefringence

Maximilian Düll*

*Zentrum für Astronomie der Universität Heidelberg*
*Philosophenweg 12, 69120 Heidelberg, Germany*
**E-mail: maximilian.duell@uni-heidelberg.de*

The constructive gravity program demands to derive the dynamics for the background geometry of some prescribed matter theory as the solution of a set of partial differential equations – so called gravitational closure equations – whose coefficient functions can be calculated directly from the prescribed matter action. Finding general solutions to the closure equations is hard to achieve in practice. A suitable symmetry reduction of the constructive gravity program will be demonstrated that allows to insert symmetries directly into the closure equations. Using cosmological symmetries – spatial homogeneity and isotropy – enables one to derive the Friedmann equations without Einstein equations for Maxwell electrodynamics as matter input. Studying general linear electrodynamics as a refined matter input, the way towards refined Friedmann equations is laid out.

*Keywords*: constructive gravity, cosmology, pre-metric electrodynamics

## 1. Constructive gravity

Combining the standard model of particle physics with general relativity and comparing them to astrophysical measurements on the energy content of the universe shows a discrepancy; only up to five percent of the universe are actually described by the standard model of particle physics and general relativity in their current form. The remaining 95 percent are attributed to dark matter and dark energy for which no satisfying theoretical model exists until today[1]. This discrepancy strongly indicates that there is something significant that we do not understand about matter, gravity, or their interaction.

In recent years, the constructive gravity program has provided new insights and understanding of the interaction of gravity and matter[2]. Its central tenet is that the gravitational Lagrangian can – and then must – be derived from a prescribed matter action by solving an immutable set of linear partial differential equations, the so called gravitational closure equations, whose coefficient functions depend on the specified matter theory in a simple way. The key physical requirement for the constructive gravity program to work is that the common canonical evolution of data between initial and final data hypersurfaces proceeds in lockstep for both matter and geometry.

A crucial check of the program, namely that using the Lagrangian of the standard model as the matter input yields the Einstein-Hilbert Lagrangian as the unique two-parameter solution of the gravitational closure equations, is passed with flying colours. The scope of the constructive gravity program reaches far beyond the standard model. In fact, for any matter theory that satisfies basic well-posedness and energy conditions, the solution of the associated closure equations yields the

Lagrangian governing the dynamics of whatever (multi-)tensorial background geometry is employed in the formulation of the matter theory. In short, any such matter action, even if a severe modification or extension of the standard model, predicts the suitable gravity theory that must underpin it.

This sheds a new light on the great variety of proposals for modifications and extensions of the standard model. It is useful to spell this out for a physically meaningful modification of the electromagnetic sector of the standard model of particle physics, which replaces Maxwell theory by what is variably called general linear, premetric[3] or area metric electrodynamics. This is the most general theory of electrodynamics with linear field equations that arises from an action. Its action is given by

$$S_{\mathrm{GLED}} = -\frac{1}{8}\int \mathrm{d}^4x\, \omega_G\, F_{ab}F_{cd}\, G^{abcd}\,, \tag{1}$$

where $F$ is the usual electromagnetic field tensor for the gauge potential $A$ and $\omega_G$ a scalar density of weight one built from the background geometry encoded in the fourth rank tensor field $G$. This theory of electrodynamics allows for birefringence in vacuum, an effect that is excluded by construction in Maxwell-Einstein theory. Were one not to provide dynamics for the more general background geometry $G$, the latter would have to be described by hand. This would require to already know whether and how this background geometry would deviate from a metric-induced background of the form

$$G^{abcd} = g^{ac}g^{bd} - g^{ad}g^{bc} - \sqrt{-\det g}\,\epsilon^{abcd}\,. \tag{2}$$

Of course, constructive gravity relieves us from such tinkering, since the gravitational closure equations associated with (1) yield the dynamics for $G$ that allow to predict, rather than stipulate, where the background geometry $G$ takes which values. Closing the matter action by addition of the thus obtained gravitational action then provides the suitable refinement of the Einstein-Maxwell action as it is demanded by dropping the par-force exclusion of vacuum birefringence. This is of course just exactly what constructive gravity is about: learning about the gravitational implications of modifications of the matter theory, which sometimes are nil and more often not (even in cases where intuitive reasoning would suggest no implication for gravity).

## 2. Symmetry-reduced constructive gravity

While it is clear from a theoretical point of view that solving the closure equations yields the desired gravitational Lagrangian, it is at best difficult to do so in practice. Even if one were able to find a general solution to the closure equations, one would not find general solutions to the resulting gravitational field equations, just as in case of the Einstein equations. The way around this, for the purpose of cosmology, is to implement the assumed cosmological symmetries already at the level of the

closure equations rather than later at the level of the gravitational field equations. One can visualize the situation as in figure 1.

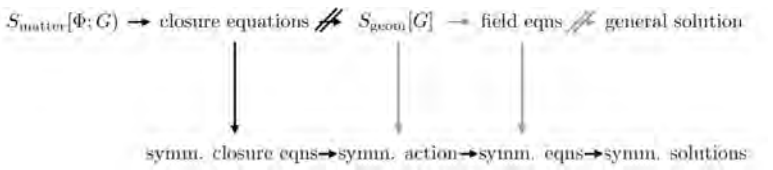



Fig. 1. Key idea of the symmetry reduction in the constructive gravity program. Diagram commutes if symmetric criticality is satisfied [5].

This diagram only commutes if implementing a symmetry at the level of an action or even before – in the closure equations – produces the correct symmetric field equations[a]. Fels and Torre investigated this question of symmetric criticality [5] and derived sufficient conditions for when one may impose a symmetry already at the level of an action (only the simplest of which is if the symmetry group is compact).

With this being a sufficient criterion, also symmetries stemming from non-compact group can be inserted properly into the closure equations. In the case of spatial homogeneity and isotropy, one is left with three different topologies for the symmetric spatial hypersurfaces. Only the case of 3-spheres (closed cosmology) originates from a compact symmetry group. In this work, we will stick with flat cosmology and choose flat planes as the spatially homogeneous and isotropic hypersurfaces. Still, as shown in 3, checking the viability of the symmetry reduction with Maxwell electrodynamics as matter input, produces the correct Friedmann equations with a non-compact underlying symmetry group. Once this is established, attention will turn towards finding the refined Friedmann equations for a universe with general linear electrodynamics in section 4.

## 3. Cosmological closure of Maxwell electrodynamics

Because it is clear that the Friedmann equations are the symmetry-reduced version of the Einstein equations which, in the philosophy of constructive gravity, arise from Maxwell electrodynamics as matter input into the associated closure equations, we employ precisely this standard case as a proof of principle. Apart from providing this proof of concept, the interesting insight is that constructive gravity allows to obtain the two Friedmann equations without using the Einstein equations.

[a]This is not guaranteed a priori. Imagine the situation in classical mechanics. Variation of the action is comparing possible curves with respect to each other and choosing the one making the action stationary. By inserting a symmetry into the action, one compares curves obeying the symmetry – in general fewer than before. By comparing fewer candidate curves, one can produce fake solutions compared to the full setting with symmetry applied in the equations of motion.

In order to identify the form of a Lorentzian metric with spatial homogeneity and isotropy, one solves the Killing condition

$$(\mathcal{L}_{K_i} g)^{ab} = 0 \tag{3}$$

with six Killing vector fields generating this symmetry [4,6,7]. As already mentioned above, we choose flat planes as the spatial hypersurfaces. This allows to write the resulting metric in a simple form

$$g^{ab} = \mathrm{diag}\left(\frac{1}{N^2(t)}, -\frac{1}{a^2(t)}, -\frac{1}{a^2(t)}, -\frac{1}{a^2(t)}\right)^{ab} \tag{4}$$

with the usual scale factor $a(t)$ and a time-dependent lapse function $N$ in an ADM split. Now, an important technical step in the constructive gravity enters the symmetry reduction – namely the introduction of the actual geometric degrees of freedom $\varphi^A$ [b]. The spatial metric $g$ can be expressed in terms of the actual degrees of freedom by virtue of a linear map $\hat{g}^{\alpha\beta}(\varphi^A) = I^{\alpha\beta}{}_A\, \varphi^A$ with a constant matrix $I$ [2]. The resulting gravitational action and its Lagrangian are then built from the actual d. o. f. $\varphi^A$. Another crucial step is the expansion of the gravitational Lagrangian

$$\mathcal{L}^{\mathrm{grav}} = \sum_{N=0}^{\infty} C_{A_1 \ldots A_N}[\varphi^A]\, k^{A_1} \ldots k^{A_N} \tag{5}$$

in terms of generalized velocities $k^A$ and expansion coefficients $C_{A_1 \ldots A_N}$ which are functionals of the actual geometric d. o. f.s and which the closure equations are solved for.

After symmetry reduction, the symmetry-reduced expansion coefficients $C^{\mathrm{symm}}_{A_1 \ldots A_N}$ only depend on the scale factor $a(t)$ which is the only freedom left in flat cosmology. The derivative of a symmetric expansion coefficient can be expressed via a chain rule,

$$C'_{A_1 \ldots A_N}(a) = \frac{\partial C_{A_1 \ldots A_N}}{\partial \varphi^A}\left(\varphi^B(a)\right)\, \frac{\partial \varphi^A}{\partial a}\,. \tag{6}$$

The derivatives $\frac{\partial C_{A_1 \ldots A_N}}{\partial \varphi^A}$ also appear in the gravitational closure equations. The task is now to evaluate the closure equations on symmetric configurations, that is evaluate all equations by setting $\varphi^A = \varphi^A(a)$ and then compute expressions for the derivatives $\frac{\partial C_{A_1 \ldots A_N}}{\partial \varphi^A}$. For each expansion coefficient, this can be calculated in several pages and the procedure will not be presented in this brief article. One ends up at ordinary differential equations for every coefficient. Solving them and

[b] The motivation for introducing actual degrees of freedom stems from the necessity to encode all symmetry conditions and constraints arising from the ADM split into a set of independent variables. For the refined geometry $G$ appearing in (1), performing an ADM split gives rise to non-linear constraints on the spatial geometry [2]. Introducing a suitable parametrization in terms of the actual geometric d. o. f. takes care of these constraints.

plugging all results together, one ends up at an action of the form

$$S_{\text{cosmo}} = \int \mathrm{d}t \left( \frac{\Lambda}{8\pi G} N a^3 + \frac{3}{8\pi G} \frac{a\dot{a}^2}{N} \right) . \tag{7}$$

with two integration constants – the cosmological and the gravitational constant.

In order to complete the symmetry reduction, a suitable matter source has to be determined by applying the Killing condition to the energy momentum tensor

$$T_{ab} := -2 \sqrt{-\det g^{\cdot\cdot}} \, \frac{\delta \mathcal{S}_{\text{matter}}}{\delta g^{ab}} . \tag{8}$$

This tensor then has diagonal form with only two functions of time remaining which can be identified as energy density $\rho$ and pressure $p$ of a perfect fluid in analogy to standard theory. By variation of the total action

$$S_{\text{total}} = S_{\text{cosmo}} + S_{\text{matter}} \tag{9}$$

with respect to lapse and the scale factor, the two Friedmann equations known from standard cosmology are recovered – without using the Einstein equations at any point. This successful proof of principle serves as the starting point for a symmetry reduction of the closure equations associated to general linear electrodynamics.

## 4. Cosmological closure of general linear electrodynamics

The background geometry of general linear electrodynamics is the fourth rank tensor field $G$ for which the Killing condition reads

$$\left(\mathcal{L}_{K_i} G\right)^{abcd} = 0 . \tag{10}$$

Using the six Killing vector fields that describe spatial homogeneity and isotropy, one now finds two scale factors $a(t)$ and $c(t)$ besides the lapse $N(t)$.

Proceeding with the symmetry reduction of the constructive gravity program as in the previous case, one now obtains two partial differential equations

$$\begin{aligned} \frac{\partial C_{A_1 \ldots A_N}}{\partial a} &= \frac{\partial C_{A_1 \ldots A_N}}{\partial \varphi^A} \left(\varphi^B(a,c)\right) \frac{\partial \varphi^A}{\partial a} , \\ \frac{\partial C_{A_1 \ldots A_N}}{\partial c} &= \frac{\partial C_{A_1 \ldots A_N}}{\partial \varphi^A} \left(\varphi^B(a,c)\right) \frac{\partial \varphi^A}{\partial c} \quad \forall N \geq 0 \end{aligned}$$

The procedure is exactly the same as in the previous case. Evaluate the closure equations by setting $\varphi^A = \varphi^A(a,c)$ everywhere to obtain expressions for the derivatives $\frac{\partial C_{A_1 \ldots A_N}}{\partial \varphi^A}$ and solve the two PDEs for every expansion coefficient. The general setting of the coefficient functions of the closure equations is more involved here[2]. For instance, one coefficient function vanishes when setting $\varphi^A = \varphi^A(a,c)$; first taking derivatives of this coefficient and setting $\varphi^A = \varphi^A(a,c)$ after that, produces non-vanishing components. This fact requires to also consider derivatives of the closure equations which are evaluated on $\varphi^A = \varphi^A(a,c)$ after having taken the derivative of the full equation. These calculations are very extensive and currently,

we are working on an automation using computer algebra. The first coefficient $C$ of the expansion (5) of the gravitational Lagrangian is already determined; it is necessary to evaluate the closure equations ($C1$) and the derivative with respect to the geometric degrees of freedom $\varphi^A$ of ($C1$) and ($C5$) in the current numbering[2]. One then finds that $C$ takes the form

$$C = K_0\, a^3\, c^{-9} \tag{11}$$

with $K_0$ a constant of integration[4].

Once all expansion coefficients are calculated, the cosmological action can be constructed from them. It is straightforward to generalize the concept of the energy momentum tensor to refined spacetime geometries[8] and especially the tensor field $G$[4,6]. As there are now two scale factors, we will obtain three refined Friedmann equations by variation with respect to the lapse function and the two scale factors – but only once the cosmological action is calculated.

## 5. Conclusions and Outlook

In this article, it was demonstrated that a suitable symmetry reduction of the constructive gravity program allows to derive Friedmann equations directly from Maxwell electrodynamics. Refined Friedmann equations follow from general linear electrodynamics as matter input once the involved evaluation of the symmetry-reduced gravitational closure equations is completed. Besides, studying the three independent components of the gravitational source tensor have to be studied more closely and a physical interpretation of a perfect fluid in a spacetime with general linear electrodynamics is needed. Once the refined Friedmann equations have been obtained, one can complement them with an equation of state and derive solutions. With these solutions, comparisons and predictions such as the age of the universe can be made. This is work in progress.

## Acknowledgments

Financial support by the Heidelberg University Graduate Academy in order to attend the Marcel Grossmann Meeting is acknowledged.

## References

1. PLANCK collaboration, *Astron. Astrophys.* **594**, A13 (2016).
2. M. Düll, F. Schuller, N. Stritzelberger, F. Wolz, *Phys. Rev.* **D97**, 084036 (2018).
3. F. W. Hehl and Y. N. Obukhov, *Lect. Notes Phys.* **702**, 163-187 (2006).
4. M. Düll, N. L. Fischer, B. M. Schäfer and F. P. Schuller, *in preparation*, (2019).
5. M. E. Fels and C. G. Torre, *Class. Quant. Grav.* **19**, 641 (2002).
6. N. L. Fischer, Master's Thesis, Heidelberg University, 2017.
7. L. Barcaroli et al., *Phys. Rev.* **D95**, 024036 (2017).
8. M. Gotay and J. E. Marsden, *Contemp. Math*, **132**, (2001).

# Non-metric fields from quantum gravity

Kaushik Ghosh*

*Vivekananda College,*
*Kolkata, West Bengal 700091, India*
** E-mail: kaushikhit10@gmail.com*
*www.vckolkata63.org*

In this manuscript, we will discuss the construction of covariant derivative operator in quantum gravity. We will find it is more perceptive to use affine connections more general than metric compatible connections in quantum gravity. We will demonstrate this using the canonical quantization procedure. This is valid irrespective of the presence and nature of sources. General affine connections can introduce new scalar fields in gravity.

*Keywords*: quantization, connections, non-metricity, scalar fields

## 1. Introduction

In this manuscript, we will discuss a few aspects of spacetime geometry relevant to quantum gravity. In Sec.2, we will discuss the construction of covariant derivative operator in quantum gravity. In this article, we will deal with only affine connections and denote general affine connections by affine connections or connection coefficients.[1,2] We will find it is more perceptive to use affine connections more general than metric compatible connections in quantum gravity. We will use the canonical quantization procedure and the Arnowitt-Deser-Misner (ADM) formalism to show this.[3] This is valid irrespective of the presence and nature of sources. This is a general mathematical issue which will be there, in a theory of quantum gravity which is not a quantum field theory in a fixed background, provided some components of metric can be taken as independent variables in a neighborhood of the spacetime manifold. This can be done around any regular point of the spacetime manifold.[4] We will also use the general metric-metric commutators to illustrate this issue.[2]

## 2. Quantum Gravity and Covariant Derivatives

We now consider quantization of gravity by using the canonical quantization procedure. Canonical quantization is important to find the particle spectrum when we quantize a classical theory. In the canonical quantization of gravity, metric becomes operator on a Hilbert space. We represent such operators by carets. Affine connections present in the covariant derivatives act on the tensor operators and we represent them also by the symbols: $\hat{\Theta}^{\alpha}_{\ \mu\nu}$. Affine connections will contain components of metric and their spacetime derivatives and also other fields as evident from the previous discussions.

In a Hamiltonian formulation, induced metric on a set of constant time surfaces is used as dynamical variable. The induced metric on a set of constant time surfaces is given by:

$$h_{\mu\nu} = g_{\mu\nu} + n_\mu n_\nu \tag{1}$$

where $n_\mu$ is the unit normal to the constant time surfaces. An expression for conjugate momenta is given by Eq.(4). We presently use the symbols $\hat{h}, \hat{\pi}$ to denote the corresponding collection of canonical operators. In general, the Levi-Civita connections contain metric and time derivative of metric components and hence, will depend on the canonical conjugate variables $(\hat{h}, \hat{\pi})$. We now express covariant derivative operator in the following form:

$$\hat{\nabla}'_\mu \hat{A}_\nu = [\partial_\mu - \hat{\Gamma}^\alpha{}_{\mu\nu}(\partial\hat{g}, \hat{g})]\hat{A}_\alpha = [\partial_\mu - \hat{\Gamma}^\alpha{}_{\mu\nu}(\hat{\pi}, \hat{h})]\hat{A}_\alpha$$

Here, $\hat{\Gamma}^\alpha{}_{\mu\nu}$ are operator version of the Levi-Civita connections. We adopt the following operator ordering in connection coefficients. Whenever there appears a product between partial derivatives of metric and metric itself, the partial derivative is kept as the first term and metric is kept as the second term. The ordering of the operators $(\hat{h}, \hat{\pi})$ in $\hat{\Gamma}^\alpha{}_{\mu\nu}$ is given to be the same as that written in the above equation, *i.e*, $\hat{h}$ is kept as the successor of $\hat{\pi}$.

We next consider the operator: $q^\mu \hat{\nabla}'_\mu q^\nu$, where $q^\mu$ is a vector field acting as $q^\mu \hat{I}$ on the Hilbert space. This operator contains canonical conjugate pairs of variables when we choose affine connections to be given by the Levi-Civita connections. In this case, we will have the following expression:

$$[q^\mu \hat{\nabla}'_\mu q^\nu]\,|\Psi\rangle \neq 0 \tag{2}$$

remaining valid in a given state $|\Psi\rangle$ with an arbitrary well-behaved vector field $q^\mu$. We will not have a complete set of states for which the expectation value of the operator in the *l.h.s* is zero with negligible fluctuations for all well-behaved vector fields. This will be valid only in the classical limit, and is a subject similar to the familiar Ehrenfest's theorems in non-relativistic quantum mechanics. Similar discussions will remain valid even if we choose affine connections to be given by more general expressions.[5] In general, affine connections will contain canonical pairs of variables from metric sector to have proper classical limit of the Levi-Civita connections, and the concept of geodesics will not remain exact for all vector fields in a quantum state. This will also remain valid for parallel transport and the notion of parallel transport is not exact in a quantum theory of gravity. This is expected and indicates that we can use affine connections more general than the metric compatible connections even in free quantum gravity.

We now consider the metric compatibility conditions. The metric compatibility conditions are to be replaced by the operator identity: $\hat{\nabla}'_\mu[\hat{g}_{\alpha\beta}] \equiv 0$. The action of $\hat{\nabla}'_\mu[\hat{g}_{\alpha\beta}]$ on any state is zero if connection operators are given by the Levi-Civita connection operators and we choose the operator ordering same as that mentioned

in.[5] Here, we always keep metric operators as the successors of the partial derivatives of themselves. The same will also remain valid for $\hat{\nabla}'_{\mu}[\hat{g}_{\alpha\beta}]$. Here, $\hat{g}_{\alpha\beta}$ will be kept at the right of the Levi-Civita connections. We also define the contravariant components of metric as: $\hat{g}^{\alpha\kappa}\hat{g}_{\kappa\beta} = \delta^{\alpha}_{\beta}$. This ordering leads to the operator identity: $\hat{\nabla}'_{\mu}[\hat{g}_{\alpha\beta}] \equiv 0$ irrespective of the ordering of $(\hat{h}, \hat{\pi})$ chosen in the partial derivatives of metric components. However, the operator version of metric compatibility conditions need not be consistent with a canonical quantization condition. We will demonstrate this in the following.

As mentioned above, in a Hamiltonian formulation we use induced metric on a set of constant time surfaces as dynamical variable. Thus, $g_{\mu\nu}$ is replaced by: $h_{\mu\nu} = g_{\mu\nu} + n_{\mu}n_{\nu}$. Here, $n_{\mu}$ is given by $(-N, 0, 0, 0)$, $N$ being the lapse function. The contravariant $n^{\mu}$ is given by: $\frac{1}{N}(1, -N^1, -N^2, -N^3)$; where $N^i$ are the shift functions. We have: $g_{00} = N_k N^k - N^2$, $g_{0i} = N_i$ and $N_i = g_{ij}N^j$.[4] The induced metric on the constant time surfaces coincide with the spatial part of $g_{\mu\nu}$ which are expressed as $g_{ij}$. These fields are taken to dynamical variables in general relativity and we have the following Poisson brackets:

$$\left\{g_{ij}(t, \vec{x}), \pi^{kl}(t, \vec{y})\right\} = \delta^{k}_{(i}\delta^{l}_{j)}[\delta(\vec{x}, \vec{y})] \tag{3}$$

Where, $\vec{x}$ refers to the spatial coordinates and the Poisson bracket is evaluated at equal time. The delta function is defined without recourse to metric. The conjugate momentum is a spatial tensor density and is given by:

$$\pi^{pl} = -\sqrt{|g_{ij}|}(K^{pl} - Kg^{pl}) \tag{4}$$

Where $|g_{ij}|$ is the determinant of the spatial metric, $K_{pl} = -\nabla'_{p}n_l$ is the extrinsic curvature of the spatial sections, $g^{pl}$ is the inverse of $g_{pl}$ and $K$ is the trace of the extrinsic curvature taken *w.r.t* $g_{ij}$.[4] We can also define the Poisson brackets for the lapse and shift functions although their conjugate momenta vanish giving primary constraints.[4] Thus, in a Hamiltonian formulation with the Einstein-Hilbert action, we have constraints and we can not naively replace the Poisson brackets by commutators when we try to quantize the theory.[4] There are two principal approaches to quantize the theory.[4] In the first approach, gauge fixing conditions are introduced to render the complete set of constraints second class.[4] These conditions also determine the lapse and shift functions. We then pick two components of $g_{ij}$ as independent variables and quantize these components using standard commutation relations. We can solve the constraints to evaluate other commutators. The second approach is similar to the Gupta-Bleuler method used to quantize electrodynamics and was initiated by Dirac.[4] In this approach the classical variables are treated as independent variables and the constraints are imposed on the quantum states. In this case, we can replace the classical Poisson brackets by commutators when we quantize the theory.

We now demonstrate that it is appropriate to extend the Levi-Civita connections and metric compatibility conditions as long as we can regard components of spatial

metric on the constant time surfaces as independent physical variables subjected to usual canonical quantization conditions. We will also find that we can not have a Hilbert space on which we can impose the metric compatibility conditions when such quantization conditions remain valid. In the following, we will restrict our attention to a neighborhood around a regular point $'x'$. We can extend the neighborhood to the complete spacetime manifold leaving away singularities and other possible irregular points associated with the constraints.[4] We pick a component of spatial metric, say $g_{pl}$, as an independent physical variable. We then have the following equal time commutator:

$$[\hat{g}_{pl}(t,\vec{x}),\hat{\pi}^{pl}(t,\vec{y})] = i\delta^{p}_{(p}\delta^{l}_{l)}[\delta(\vec{x},\vec{y})] \tag{5}$$

Where, the point $'y'$ belongs to the above mentioned neighborhood of $'x'$. There will be another such commutator for the other independent variable. The *r.h.s* of the commutator is taken to be a distribution that is a spatial tensor density in the spatial coordinates of $\hat{\pi}^{pl}(t,\vec{y})$. The *r.h.s* will be replaced by different expressions when we replace $\hat{g}_{pl}(t,\vec{x})$ by any dependent component of metric including $g_{0\mu}$. These terms are determined by the secondary constraints, gauge fixing conditions, definitions of the lapse and shift functions given before and the fundamental commutators given by the above equation. Covariant derivatives give changes in a tensor when we move from one point to a neighbouring point. If the Levi-Civita connections are consistent with the commutators obtained above, corresponding spatial covariant derivatives of both sides of any of the commutators *w.r.t* the arguments of metric will agree since both sides are equal for all components of metric. We now consider Eq.(5). The action of $\hat{\nabla}'_{xk}$ on the *r.h.s* is same as that on a second rank covariant tensor and will contain spatial partial derivatives of the delta function. It will also contain additional terms dependent on connections and metric that can explicitly depend on time due to explicit time dependence of the gauge fixing conditions.[4] This covariant derivative is not vanishing in general for all values of $\vec{x}$. The left hand side vanishes as can be found from the following expression:

$$\hat{\nabla}'_{xk}\left\{\hat{g}_{pl}(t,\vec{x})\right\}\hat{\pi}^{pl}(t,\vec{y}) - \hat{\pi}^{pl}(t,\vec{y})\hat{\nabla}'_{xk}\left\{\hat{g}_{pl}(t,\vec{x})\right\} = 0 \tag{6}$$

This follows since we are imposing the operator versions of the metric compatibility conditions. This can also be seen by applying the *l.h.s* of the above equation to any state, introducing a sum over a complete set of states between the products of the operators and using the fact that the action of $\hat{\nabla}'_{\mu}[\hat{g}_{\alpha\beta}]$ on any state is zero if connection operators are given by the Levi-Civita connection operators with operator ordering chosen below Eq.(2). Thus, the Levi-Civita connections are not consistent with the canonical commutators given by Eq.(5). We will consider other operator ordering in the Levi-Civita symbols in Appendix:A and we will find that similar inconsistency arise in these cases also. Similar situation will remain valid for any point in the manifold where we can introduce constant time surfaces and assume the existence of a metric component as an independent field in a neighborhood

around that point. The above inconsistency will also arise with a different choice of constant time surfaces. Thus, there will be a multitude of coordinate systems where we can not use the Levi-Civita connections as connection coefficients if we impose the quantization condition given by Eq.(5). This also indicates that we can not use the Levi-Civita connections as connection coefficients in all coordinate systems that are diffeomorphic to these coordinate systems due to the tensorial character of $C^{\alpha}_{\mu\nu}$.

In Dirac's approach to quantize gravity, all the classical variables are independent and we quantize them accordingly. We can find out $[\hat{g}_{0\beta}(t,\vec{x}), \hat{\pi}^{pl}(t,\vec{y})]$ from the definitions of the lapse and shift functions and $[\hat{N}^{\lambda}(t,\vec{x}), \hat{\pi}^{pl}(t,\vec{y})] = 0$, where $N^0 = N$.[4] All spatial components $g_{ij}$ satisfy canonical commutation relations given by Eq.(5). We will again have the inconsistency mentioned above when we use the Levi-Civita connections. In this case, the action of $\hat{\nabla}'_{xk}$ on the *r.h.s* of Eq.(5) is given by expressions like: $i[\partial_{xk}\delta(\vec{x},\vec{y}) - 2\hat{\Gamma}^{0}{}_{kp}(t,\vec{x})\hat{N}^{p}(t,\vec{x})\delta(\vec{x},\vec{y}) - 2\hat{\Gamma}^{p}{}_{kp}(t,\vec{x})\delta(\vec{x},\vec{y})]$, where we have taken $p = l$ and there is no sum over the repeated indices. Also, we can not make the Levi-Civita connections consistent with the commutation relation given by Eq.(5) by introducing additional constraints on the physical Hilbert space. If we demand that the action of $\hat{\nabla}'_{xk}$ to the *r.h.s* of Eq.(5) vanishes on the physical Hilbert space, we will have the constraint:

$$[\partial_{xk}\delta(\vec{x},\vec{y}) - 2\hat{\Gamma}^{0}{}_{kp}(t,\vec{x})\hat{N}^{p}(t,\vec{x})\delta(\vec{x},\vec{y}) - 2\hat{\Gamma}^{p}{}_{kp}(t,\vec{x})\delta(\vec{x},\vec{y})]|\Psi\rangle = 0 \qquad (7)$$

Where, we have again taken $p = l$ and there is no sum over the repeated indices. There will be other similar constraints associated with other commutators. These states also satisfy the secondary constraints. We have used the operator identities $\hat{\nabla}'_{\mu}\hat{g}_{\alpha\beta} = 0$, and the Levi-Civita connections. All the above conditions will lead to singular expressions involving $\delta(\mathbf{0})$ and the partial derivatives $\delta'(\mathbf{0})$ for expectation values of some of the variables: $\hat{g}_{0\mu}, \hat{g}_{ij}, \hat{N}^{p}$ and $\hat{\Gamma}^{\alpha}_{\mu\nu}$. This is valid for all physical states and is physically undesirable. Lastly, the above problems will arise if we use any set of metric compatible connections.

In the first approach to quantize the theory, it is unlikely that there will exist a set of gauges that is time dependent, render the complete set of constraints second class and also remove the inconsistency mentioned above. It is not possible to remove the inconsistency in the second approach to quantize the theory. Also, it is expected that $[\hat{g}_{\alpha\beta}(x^{\mu}), \hat{g}_{\alpha\beta}(y^{\nu})]$ will depend on $(x^{\mu}, y^{\nu})$ non-trivially with non-vanishing covariant derivatives.[2] We can consider semiclassical theories like quantum fields in curved spaces to assume so. Thus, it is more appropriate to use connections more general than metric compatible connections in quantum gravity. The above discussions are valid irrespective of the presence and nature of sources. We can analyze this issue further in the following way. The Levi-Civita connections and metric compatibility conditions are taken as basic assumptions to calculate the scalar curvature when we use the Einstein-Hilbert action to describe classical and quantum gravity. It is better to discuss quantization and non-metricity using the metric-affine action or Palatini action where metric and affine connections are

independent variables. We have found that nen-metricity can give scalar fields in the theory. One of the scalar fields can give negative stress-tensor.[5]

## References


1. D. Lovelock and H. Rund, *Tensors, Differential Forms, and Variational Principals* (Dover Publications, Inc., New York, 1989).
2. R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago and London, 1984).
3. C. W. Misner, K. S. Thorne and J. A. Wheeler, *Gravitation* (W.H. Freeman and company, New York, 1970).
4. K. Sundermeyer, *Constrained Dynamics* (Springer-Verlag, 1995).
5. K. Ghosh, *https://hal.archives-ouvertes.fr/hal-02105422* (accepted in *Physics of the Dark Universe*)

# Extension of GR with a gravitational constitutive tensor

Yakov Itin[1], Yuri N. Obukhov[2], Jens Boos[3], Friedrich W. Hehl[4]

[1] *Institute of Mathematics, The Hebrew University of Jerusalem, and Jerusalem College of Technology, Jerusalem 91160, Israel*

[2] *Theoretical Physics Laboratory, Nuclear Safety Institute Russian Academy of Sciences, B. Tulskaya 52, 115191 Moscow, Russia*

[3] *Theoretical Physics Institute, University of Alberta, Edmonton, AB T6G 2E1, Canada*

[4] *Institute for Theoretical Physics, University of Cologne 50923 Köln, Germany*

In this paper, we present a brief account of a generic premetric gravity model based an the energy-momentum conservation law. Our main goal is to discuss the generic properties of the gravitational constitutive pseudotensor. We present irreducible decomposition of this object under the permutation group, or, equivalently, under the general linear group $GL(4,\mathbb{R})$. Even when the consideration is restricted to the metric ansatz, the nontrivial skewon and axion parts emerge. A special choice of the free dimensionless parameters recovers the standard GR in its premetric version.

*Keywords*: Premetric gravitational model, teleparallel gravity, gravitational constitutive tensor

## 1. Introduction

Gauge field theories such as electromagnetism and Yang-Mills models can be presented in premetric formalism[1–3] that separates the structure into the system of metric-free field equations based on conservation laws and the constitutive law between the basis field variables. In the simplest vacuum case, the constitutive law is expressed in terms of the metric tensor.

At first glance, the standard GR does not allow for such premetric description since it is based essentially on the metric variable. In this paper, we demonstrate how GR can be recast into a form suitable for premetric reformulation. This extension is based on two field variables, the gravitational excitation $H_\alpha$ and the field strength $F^\alpha$, that satisfy the system of field equations that is completely similar to the Maxwell system of electromagnetism. A linear constitutive relation between the vector-valued 2 forms $H_\alpha$ and $F^\alpha$ is established with a sixth order constitutive pseudo-tensor. This new gravitational quantity is a central subject of our consideration. In order to find the physical meaning of the constitutive tensor, we provide its irreducible decomposition under the general linear group $GL(4,\mathbb{R})$. We demonstrate the contributions of different irreducible parts to the coframe Lagrangian, energy-momentum current and dispersion relation. In metric case, we present a most general even and odd metric-dependent ansatz for the constitutive tensor. We find the physical meaning of different irreducible pieces of metric constitutive tensor and recover the standard GR for a special values of parameters.

## 2. Premetric gravitational model

In this section, we present a brief description of the premetric formulation of the gravity model, see Refs. 3, 4, 5, 6. The construction is very similar to the premetric representation of electromagnetism.[1,2]

### 2.1. *Field equation*

We assume the following basic postulate:

*Every viable model of gravity must be based on conservation law of the energy-momentum current of matter and field.*

In the canonical approach, the energy-momentum current is described as a mixed-type tensor or, equivalently, as a covector-valued twisted 3-form $\Sigma_\alpha$. For a system including matter and gravitational field, $\Sigma_\alpha$ consists of two independent pieces – for matter and for field,

$$\Sigma_\alpha = {}^{(m)}\Sigma_\alpha + {}^{(f)}\Sigma_\alpha \,. \tag{1}$$

The inhomogeneous field equation is considered as a consequence of the energy-momentum conservation law. For topologically trivial regions, it is realized as an equation for the excitation variable

$$d\Sigma_\alpha = 0 \qquad \Longrightarrow \qquad \Sigma_\alpha = dH_\alpha \,. \tag{2}$$

The latter equation is a vector-valued analog of the electromagnetic field equation $J = dF$. Thus, we define a twisted covector-valued 2-form of *gravitational excitation* $H_\alpha$ that is expressed in the basis of odd (twisted) 2-forms $\epsilon_{\beta\gamma}$ as

$$H_\alpha = \frac{1}{2} H^{\beta\gamma}{}_\alpha \, \epsilon_{\beta\gamma}. \tag{3}$$

*Gravitational field strength* is defined as an even (untwisted) vector-valued 2-form $F^\alpha$ that is expressed in the basis $\vartheta^{\beta\gamma}$ of even (untwisted) 2-forms as

$$F^\alpha = \frac{1}{2} F_{\beta\gamma}{}^\alpha \vartheta^{\beta\gamma}. \tag{4}$$

In an analogy to the electromagnetic field equation $dF = 0$, the gravitational field strength is assumed to satisfy the closeness field equation

$$dF^\alpha = 0 \qquad \Longrightarrow \qquad F^\alpha = d\vartheta^\alpha \,. \tag{5}$$

For topologically trivial regions, the field equations yields the existence of the *gravitational potential* $\vartheta^\alpha$. It is a vector-valued analog of the potential 1-form $A$ of electromagnetism.

Using the form of the electromagnetic Lorentz force $f_\alpha = (e_\alpha \rfloor F) \wedge J$, we define the *gravitational Lorentz force* as a 4-form relating the field strength to the energy-momentum current of matter

$$f_\alpha = \left(e_\alpha \rfloor F^\beta\right) \wedge {}^{(\mathrm{m})}\Sigma_\beta \,. \tag{6}$$

Also the twisted 3-form of *energy-momentum current of the coframe field*

$$ {}^{(\vartheta)}\Sigma_\alpha = \frac{1}{2}\left[F^\beta \wedge (e_\alpha \rfloor H_\beta) - H_\beta \wedge (e_\alpha \rfloor F^\beta)\right] , \tag{7} $$

and the *gravitational Lagrangian*

$$ {}^{(\vartheta)}\Lambda = -\frac{1}{2} F^\alpha \wedge H_\alpha \tag{8} $$

are defined as full vector-valued analogs of the corresponding quantities of the standard Maxwell electromagnetism.

### 2.2. *Gravitational constitutive pseudo-tensor*

In order to complete the undetermined system of field equations (2) and (5), a constitutive relation between two basic field variables $H_\alpha$ and $F^\alpha$ must be introduced. Such type of a relation depends on physical model of medium. A wide class of electromagnetic models is described by a linear local constitutive relation. In order to describe a most general linear constitutive relation, we use the expansions of the twisted form $H$ and untwisted form $F$ relative to the basis coframe of 1-forms $\vartheta^\alpha$ and the dual basis of vectors $e_\alpha$. Recall a most general linear relation between the components of the electromagnetic fields $H$ and $F$ reads

$$ H^{\alpha\beta} = \frac{1}{2}\chi^{\alpha\beta\gamma\delta} F_{\gamma\delta} , \tag{9} $$

where $\chi^{\alpha\beta\gamma\delta}$ is the *electromagnetic constitutive pseudo-tensor.* It has 36 independent components and irreducibly decomposed into principal part, skewon, and axion parts as $36 = 20 \oplus 15 \oplus 1$.

With this electromagnetic constitutive relation in mind, we postulate a similar local linear constitutive relation between the components of the fields $H_\alpha$ and $F^\alpha$

$$ H^{\beta\gamma}{}_\alpha = \frac{1}{2}\chi^{\beta\gamma}{}_\alpha{}^{\nu\rho}{}_\mu F_{\nu\rho}{}^\mu . \tag{10} $$

The *gravitational constitutive pseudotensor* obeys the symmetries

$$ \chi^{\beta\gamma}{}_\alpha{}^{\nu\rho}{}_\mu = -\chi^{\gamma\beta}{}_\alpha{}^{\nu\rho}{}_\mu = -\chi^{\beta\gamma}{}_\alpha{}^{\rho\nu}{}_\mu \tag{11} $$

so it has 576 independent components.

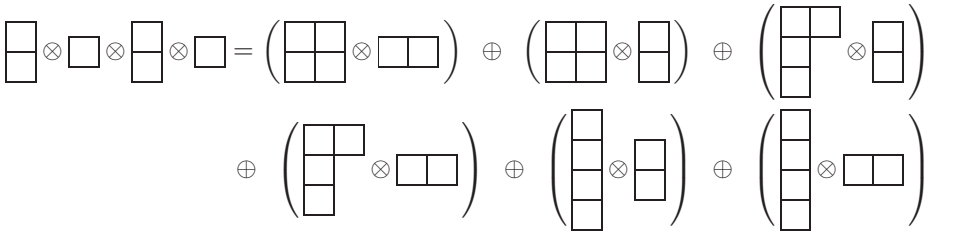
If we additionally assume that the model is completely specified by the Lagrangian (8), we have to restrict our consideration to reversible processes. In this case, we can specialize to a constitutive pseudo-tensor with an additional reciprocity condition

$$ \chi^{\beta\gamma}{}_\alpha{}^{\nu\rho}{}_\mu = \chi^{\nu\rho}{}_\mu{}^{\beta\gamma}{}_\alpha . \tag{12} $$

In this case, the set of the 576 independent components of $\chi^{\beta\gamma}{}_\alpha{}^{\nu\rho}{}_\mu$ is reduced to only 300 ones.

Treating the covariant and contravariant indices as belonging to two separate tensor spaces, the irreducible decomposition of a full (non-restricted) tensor $\chi^{\alpha\beta}{}_\mu{}^{\gamma\delta}{}_\nu$

is defined as a product of two irreducible decompositions. The corresponding Young diagrams are expanded into 6 independent pieces



(13)

Explicitly, the irreducible pieces of the constitutive tensor are expressed as

$$ {}^{[1]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := \chi^{\alpha\beta}{}_{(\mu}{}^{\gamma\delta}{}_{\nu)} - {}^{[3]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} - {}^{[5]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}, \tag{14} $$

$$ {}^{[2]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := \chi^{\alpha\beta}{}_{[\mu}{}^{\gamma\delta}{}_{\nu]} - {}^{[4]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} - {}^{[6]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}, \tag{15} $$

$$ {}^{[3]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := (1/2)\left(\chi^{\alpha\beta}{}_{(\mu}{}^{\gamma\delta}{}_{\nu)} - \chi^{\gamma\delta}{}_{(\mu}{}^{\alpha\beta}{}_{\nu)}\right), \tag{16} $$

$$ {}^{[4]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := (1/2)\left(\chi^{\alpha\beta}{}_{[\mu}{}^{\gamma\delta}{}_{\nu]} - \chi^{\gamma\delta}{}_{[\mu}{}^{\alpha\beta}{}_{\nu]}\right), \tag{17} $$

$$ {}^{[5]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := \chi^{[\alpha\beta}{}_{(\mu}{}^{\gamma\delta]}{}_{\nu)}, \tag{18} $$

$$ {}^{[6]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := \chi^{[\alpha\beta}{}_{[\mu}{}^{\gamma\delta]}{}_{\nu]}. \tag{19} $$

These pieces produce the decomposition of the tensor space into direct sum of invariant subspaces with the dimensions, respectively,

$$ 576 = 200 \ \oplus \ 120 \ \oplus \ 150 \ \oplus \ 90 \ \oplus \ 10 \ \oplus \ 6. \tag{20} $$

Additionally to the described above $S_2 \times S_4$ irreducible decomposition of the constitutive tensor, we study in Ref. 5 an alternative $S_6$ irreducible decomposition based on the pure covariant representation of this tensor.

Let us turn to a possible physical identification of the irreducible parts. We calculate the contributions of the irreducible parts to the coframe Lagrangian, the energy-momentum current and into the dispersion relation. The results are summarized in the following table:

| Irreducible Parts | Nomenclature | Lagrangian | Energy-momentum | Dispersion |
|---|---|---|---|---|
| ${}^{[1]}\chi^{\alpha\beta}{}_{\gamma}{}^{\mu\nu}{}_{\rho}$ | Principal-1 | yes | yes | yes |
| ${}^{[2]}\chi^{\alpha\beta}{}_{\gamma}{}^{\mu\nu}{}_{\rho}$ | Skewon-1 | no | yes | yes |
| ${}^{[3]}\chi^{\alpha\beta}{}_{\gamma}{}^{\mu\nu}{}_{\rho}$ | Skewon-2 | no | yes | yes |
| ${}^{[4]}\chi^{\alpha\beta}{}_{\gamma}{}^{\mu\nu}{}_{\rho}$ | Principal-2 | yes | yes | yes |
| ${}^{[5]}\chi^{\alpha\beta}{}_{\gamma}{}^{\mu\nu}{}_{\rho}$ | Axion-1 | yes | no | no |
| ${}^{[6]}\chi^{\alpha\beta}{}_{\gamma}{}^{\mu\nu}{}_{\rho}$ | Axion-2 | no | yes | no |

## 3. A metric-dependent ansatz

In this section we study the gravitational constitutive tensor constructed from the familiar building blocks – metric tensor $g_{\alpha\beta}$ and the permutation pseudotensor $\epsilon^{\alpha\beta\gamma\delta}$.

### 3.1. *Even and odd parts of the constitutive tensor*

We start with the most general expression that can be constructed from the metric tensor $g^{\alpha\beta}$ alone and apply the symmetry relation constrains. We obtain a *parity conserving ansatz* for the constitutive pseudo-tensor

$$^{\text{even}}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{\sqrt{-g}}{\varkappa}\left(\beta_1\, g^{\gamma[\alpha}g^{\beta]\delta}g_{\mu\nu} + \beta_2\,\delta^{[\alpha}_{\mu}g^{\beta][\gamma}\delta^{\delta]}_{\nu} + \beta_3\,\delta^{[\alpha}_{\nu}g^{\beta][\gamma}\delta^{\delta]}_{\mu}\right). \tag{21}$$

Here, $\varkappa$ is the gravitational constant, $g := \det g_{\mu\nu}$, the parameters $\beta_1, \beta_2, \beta_3$ are arbitrary dimensionless constants.

A most general construction that involves permutation pseudo-tensor $\epsilon^{\alpha\beta\gamma\delta}$ together with metric tensor $g_{\alpha\beta}$ yields a *parity violating ansatz*. Applying the symmetry relation constraints we obtain an expression with three independent terms

$$^{\text{odd}}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{1}{\varkappa}\left[\beta_4\,\varepsilon^{\alpha\beta\gamma\delta}\,g_{\mu\nu} + \beta_5\,\varepsilon^{\alpha\beta[\gamma}{}_{[\mu}\,\delta^{\delta]}_{\nu]} + \beta_6\left(\delta^{[\alpha}_{(\mu}\varepsilon^{\beta]\gamma\delta}{}_{\nu)} - \delta^{[\gamma}_{(\mu}\varepsilon^{\delta]\alpha\beta}{}_{\nu)}\right)\right]. \tag{22}$$

The general metrical constitutive tensor encompasses the parity-even and parity-odd pieces,

$$\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = {}^{\text{even}}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) + {}^{\text{odd}}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g). \tag{23}$$

Computing the irreducible parts, we find two zero terms

$$^{[2]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = {}^{[6]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = 0 \tag{24}$$

and four nontrivial terms:

• Principal-1 part is build up from the even part of the constitutive tensor. It contributes to the Lagrangian, energy-momentum current and dispersion relation

$$^{[1]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{\sqrt{-g}}{\varkappa}\left[\beta_1\, g^{\gamma[\alpha}g^{\beta]\delta}g_{\mu\nu} + (\beta_2+\beta_3)\,\delta^{[\alpha}_{(\mu}g^{\beta][\gamma}\delta^{\delta]}_{\nu)}\right], \tag{25}$$

• Skewon-2 part is constructed from the pure odd part of the constitutive tensor. It is not involved in the Lagrangian but contributes to the energy-momentum current and dispersion relation

$$^{[3]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{\beta_6}{\varkappa}\left(\delta^{[\alpha}_{(\mu}\varepsilon^{\beta]\gamma\delta}{}_{\nu)} - \delta^{[\gamma}_{(\mu}\varepsilon^{\delta]\alpha\beta}{}_{\nu)}\right), \tag{26}$$

• Principal-2 part is similar to the Principal-1 but includes even and odd pieces

$$^{[4]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{1}{\varkappa}\left[(\beta_2-\beta_3)\,\sqrt{-g}\delta^{[\alpha}_{[\mu}g^{\beta][\gamma}\delta^{\delta]}_{\nu]} + \beta_5\,\varepsilon^{\alpha\beta[\gamma}{}_{[\mu}\,\delta^{\delta]}_{\nu]}\right], \tag{27}$$

• Axion-1 part is of the odd origin. It is included into the Lagrangian but does not contribute to the energy-momentum current and to the dispersion relation

$$ {}^{[5]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{\beta_4}{\varkappa}\,\varepsilon^{\alpha\beta\gamma\delta}\,g_{\mu\nu}. \tag{28} $$

Summarizing, a general metric-dependent constitutive tensor encompasses two principal ${}^{[1]}\chi$ and ${}^{[4]}\chi$ parts, one axion ${}^{[5]}\chi$ part, and one skewon ${}^{[3]}\chi$ part.

### 3.2. *GR recovered*

In the framework of the teleparallel equivalent $\mathrm{GR}_{||}$ of general relativity (GR) we have found (see Eq. (88) in Ref. 5) for the parameters $\beta$:

$$ \left.\begin{array}{lll} \beta_1 = 4\,, & \beta_2 = 16\,, & \beta_3 = -8 \\ \beta_4 = 0\,, & \beta_5 = 0\,, & \beta_6 = 0 \end{array}\right\} \quad \mathrm{GR}_{||}\,. \tag{29} $$

The $\mathrm{GR}_{||}$ Lagrangian is distinguished from the other teleparallelism Lagrangians as being *locally Lorentz invariant*,

$$ \Lambda = \chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) F_{\alpha\beta}{}^{\mu} F_{\gamma\delta}{}^{\nu} = R\sqrt{-g} + \text{total derivative}\,. \tag{30} $$

Any other additional term in the gravitational Lagrangian breaks this local invariance. The irreducible parts of this constitutive tensor are Principal-1 and Principal-2 pieces, respectively

$$ {}^{[1]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = 4\frac{\sqrt{-g}}{\varkappa}\Big[\, g^{\gamma[\alpha} g^{\beta]\delta} g_{\mu\nu} + 2\,\delta^{[\alpha}_{(\mu} g^{\beta][\gamma} \delta^{\delta]}_{\nu)} \Big] \tag{31} $$

$$ {}^{[4]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = 24\frac{\sqrt{-g}}{\varkappa}\Big[\, \delta^{[\alpha}_{[\mu} g^{\beta][\gamma} \delta^{\delta]}_{\nu]} \Big], \tag{32} $$

## 4. Conclusion

- We presented a general premetric formalism for gravity based on the energy-conservation law.
- The linear local constitutive equation involves 6-th order constitutive tensor.
- The tensor is decomposed irreducibly and some physical meaning of its independent terms is established. In the linear approximation, it is reminiscent of the treating recently presented in Ref. 7.
- We outlined the embedding of the metric-dependent constitutive tensor in the general scheme.
- The telepallel version of GR is found as a unique Lorentz-invariant model.

## References

1. E. J. Post, *Formal Structure of Electromagnetics – General Covariance and Electromagnetics* (North Holland, Amsterdam, 1962, and Dover, Mineola, NY, 1997).

2. F. W. Hehl and Yu. N. Obukhov, *Foundations of Classical Electrodynamics: Charge, Flux, and Metric* (Birkhäuser, Boston, MA, 2003).
3. F. W. Hehl, Y. Itin and Y. N. Obukhov, "On Kottler's path: origin and evolution of the premetric program in gravity and in electrodynamics," Int. J. Mod. Phys. D **25**, no. 11, 1640016 (2016).
4. Y. Itin, F. W. Hehl and Y. N. Obukhov, "Premetric equivalent of general relativity: Teleparallelism," Phys. Rev. D **95**, 084020 (2017).
5. Y. Itin, Y. N. Obukhov, J. Boos and F. W. Hehl, "Premetric teleparallel theory of gravity and its local and linear constitutive law," Eur. Phys. J. C **78**, no. 11, 907 (2018).
6. Y. Itin, "Energy momentum current for coframe gravity," Class. Quant. Grav. **19**, 173 (2002).
7. V. A. Kostelecký and M. Mewes, "Lorentz and diffeomorphism violations in linearized gravity," Phys. Lett. B **779**, 136 (2018).

# Premetric approach in gravity and electrodynamics

Yuri N. Obukhov

*Russian Academy of Sciences, Nuclear Safety Institute (IBRAE),*
*B. Tulskaya 52, 115191 Moscow, Russia*
*E-mail: obukhov@ibrae.ac.ru*

The basics of the premetric approach are discussed, including the essential details of the formalism and some of its beautiful consequences. We demonstrate how the classical electrodynamics can be developed without a metric in a quite straightforward way: Maxwell's equations, together with the general response law for material media, admit a consistent premetric formulation. Furthermore, we show that in relativistic theories of gravity, the premetric program leads to a better understanding of the interdependence between topological, affine, and metric concepts.

*Keywords*: Premetric theory, electrodynamics, teleparallelism, wave propagation.

## 1. Development of premetric ideas

The study of the metric-free, or premetric, models has a rich and long history. The corresponding timeline of works which developed the premetric approach in physics is as follows. As a starting point, we mention H. Minkowski (1908) who established the special-relativistic (Poincaré-covariant) formalism for Maxwell's theory. F. Kottler (1912) subsequently provided the generally covariant formulation of electromagnetism which was eventually taken by A. Einstein (1916) as a basis for a quasi-premetric formulation of Maxwell's theory. F. Kottler (1922) then actually pioneered the premetric approach by constructing, based on integral conservation laws, metric-free theory for Newton's gravity and Maxwell's electrodynamics. Similar line was followed by É. Cartan (1923) in electrodynamics and by D. van Dantzig (1934) who proposed a general premetric program in physics. Important technical contributions came from I. Tamm (1925) who studied a general linear constitutive law and from A. Sommerfeld (1948) who coined, following Mie, the notions of extensive and intensive quantities. The formal structure of electrodynamics was thoroughly investigated by E.J. Post (1962) and by C. Truesdell and R. Toupin (1960) in the framework of the formal field theory. An important issue of recovering the metric (deriving light-cone) was then clarified in the works of R. Toupin (1965), M. Schönberg (1971), and A. Jadzcyk (1979). W.-T. Ni (1973) was the first to propose an axion-dilaton extension of Maxwell's theory, and subsequently F. Wilczek (1987) looked into experimental issues of axion electrodynamics. See[1–4] for exact references.

Why going metric-free? Quoting Edmund Whittaker (1953): *Since the notion of metric is a complicated one, which requires measurements with clocks and scales, generally with rigid bodies, which themselves are systems of great complexity, it seems undesirable to take metric as fundamental, particularly for phenomena which are simpler and actually independent of it.* Furthermore, as we know (in Einstein's approach), metric is identified with the gravitational field. Thereby, one can say that

a metric formulation of a physical theory is "contaminated" by gravity. Hence, by revealing the metric-free relations one actually finds the most fundamental physical structures. In this sense, premetric approach helps to understand the essential interdependence between topological, affine, and metric concepts. Moreover, from the experimental standpoint, counting is obviously the simplest measurement!

### 1.1. *Principles of premetric approach*

Premetric axiomatics is straightforward. Construction of a metric-free physical model is inherently based on conservation laws. The latter should be phenomenologically verified by means of a certain counting process without use of the metric. Next, all physical objects are divided into two sets: extensive variables (*how much?*) are distinguished from intensive variables (*how strong?*). Based on the conservation laws, the two sets of *fundamental equations* then naturally arise for excitations (extensities) and for strengths (intensities). Final step: to convert the model into a predictive theory, one adds the *linking equations*, or constitutive relations, between excitations and field strengths.

The essence of the premetric art is formulated as follows: Fundamental equations are metric-free; the metric only enters via linking equations (constitutive relations).

## 2. Premetric electrodynamics

Classical electrodynamics admits a consistent premetric formulation[1]. Based on the conservation of electric charge $dJ = 0$, one arrives at $dH = J$, and the magnetic flux conservation results in $dF = 0$. Conservation laws require just *counting procedures* – of electric charges and magnetic flux lines. No distance concept is needed, therefore the premetric framework arises naturally.

Excitations $H = (\mathcal{D}, \mathcal{H})$ are measurable via the *charge*. These are extensive variables (how much?). Field strengths $F = (E, B)$ are measurable via the *force*. They are intensive variables (how strong?).

*Fundamental equations* of electrodynamics are metric-free:

$$dH = J, \qquad dF = 0. \tag{1}$$

*Linking equations* (constitutive relations) yield a predictive theory

$$H = \kappa\,[F]\,, \qquad H_{\alpha\beta} = \kappa_{\alpha\beta}{}^{\mu\nu} F_{\mu\nu}. \tag{2}$$

The metric is hidden/encoded in the constitutive tensor $\kappa_{\alpha\beta}{}^{\mu\nu}$. Constitutive relation can be more general – nonlinear and even nonlocal.

Using the local coordinates $x^i = (t, x^a)$, we have the $(1+3)$-decompositions

$$H = \mathcal{H}_a dx^a \wedge dt + \mathcal{D}^a \epsilon_a, \qquad F = E_a dx^a \wedge dt + B^a \epsilon_a, \tag{3}$$

(with $dx^a$ and $\epsilon_a$ bases of spatial 1- and 2-forms) and the local and linear constitutive law (2) is recast into

$$\mathcal{H}_a = -\,\mathfrak{C}^b{}_a\, E_b + \mathfrak{B}_{ba}\, B^b\,, \qquad \mathcal{D}^a = -\,\mathfrak{A}^{ba}\, E_b + \mathfrak{D}_b{}^a\, B^b\,. \tag{4}$$

With $\varepsilon^{ab} = \varepsilon^{ba}$, $\mu^{-1}_{ab} = \mu^{-1}_{ba}$, $\gamma^c{}_c = 0$, a convenient parametrization reads as[1]:

$$\mathfrak{A}^{ba} = -\,\varepsilon^{ab} + \epsilon^{abc} n_c, \qquad \mathfrak{B}_{ba} = \mu^{-1}_{ab} - \epsilon_{abc} m^c, \tag{5}$$

$$\mathfrak{C}^b{}_a = \gamma^b{}_a - s_a{}^b + \delta^b_a\, s_c{}^c + \alpha\, \delta^b_a, \qquad \mathfrak{D}_b{}^a = \gamma^a{}_b + s_b{}^a - \delta^a_b\, s_c{}^c + \alpha\, \delta^a_b. \tag{6}$$

Thereby the fine structure of the constitutive tensor is eventually revealed:

$$\kappa = \underbrace{\begin{pmatrix} \gamma^b{}_a & \mu^{-1}_{ab} \\ -\varepsilon^{ab} & \gamma^a{}_b \end{pmatrix}}_{\textit{principal}\text{ part 20 comp.}} + \underbrace{\begin{pmatrix} -s_a{}^b + \delta^b_a s_c{}^c & -\epsilon_{abc} m^c \\ \epsilon^{abc} n_c & s_b{}^a - \delta^a_b s_c{}^c \end{pmatrix}}_{\textit{skewon}\text{ part 15 comp.}} + \underbrace{\alpha \begin{pmatrix} \delta^b_a & 0 \\ 0 & \delta^a_b \end{pmatrix}}_{\textit{axion}\text{ part 1 comp.}} \tag{7}$$

## 3. Premetric formulation of gravity

Kottler in 1922 formulated a premetric nonrelativistic (Newton's) gravity. However, the spacetime metric = gravity in general relativity (GR) theory, and one may ask whether a premetric *relativistic* gravity makes any sense? In GR, the answer is negative. However, a *teleparallel* framework offers a viable opportunity[2–4]. Qualitatively, one proceeds by replacing the electric charge with the "gravitational charge" = mass (energy-momentum). The conservation of the gravitational charge $d\Sigma_\alpha = 0$ yields $dH_\alpha = \Sigma_\alpha$, thereby introducing the gravitational excitation 2-form $H_\alpha = \frac{1}{2} H_{ij\alpha} dx^i \wedge dx^j = \frac{1}{2} \check{H}^{\rho\sigma}{}_\alpha \epsilon_{\rho\sigma}$. Furthermore, the gravitational flux conservation results in $dF^\alpha = 0$, introducing the gravitational field strength 2-form $F^\alpha = \frac{1}{2} F_{ij}{}^\alpha dx^i \wedge dx^j = d\vartheta^\alpha$, so that the coframe $\vartheta^\alpha$ plays a role of the gravitational potential. The premetric formulation of gravity can be then consistently constructed along the lines of premetric electromagnetism. The corresponding gravitational-electromagnetic analogy is summarized in Table 1.

Table 1. Premetric electromagnetism-gravity analogy

| Objects and Laws | Electromagnetism | Gravity |
|---|---|---|
| Source current 3-form | $J$ | $\Sigma_\alpha$ |
| Conservation law | $dJ = 0$ | $d\Sigma_\alpha = 0$ |
| Excitation 2-form | $H$ | $H_\alpha$ |
| Inhomogeneous field equation | $dH = J$ | $dH_\alpha = \Sigma_\alpha = {}^{(\vartheta)}\Sigma_\alpha + {}^{(\mathrm{m})}\Sigma_\alpha$ |
| Field strength 2-form | $F$ | $F^\alpha$ |
| Homogeneous field equation | $dF = 0$ | $dF^\alpha = 0$ |
| Potential 1-form | $A$ | $\vartheta^\alpha$ |
| Potential equation | $dA = F$ | $d\vartheta^\alpha = F^\alpha$ |
| Lorentz force | $f_\alpha = (e_\alpha \rfloor F) \wedge J$ | $f_\alpha = \left(e_\alpha \rfloor F^\beta\right) \wedge {}^{(\mathrm{m})}\Sigma_\beta$ |
| Energy-momentum 3-form | $\frac{1}{2}(F \wedge e_\alpha \rfloor H - H \wedge e_\alpha \rfloor F)$ | $\frac{1}{2}(F^\beta \wedge e_\alpha \rfloor H_\beta - H_\beta \wedge e_\alpha \rfloor F^\beta)$ |
| Lagrangian 4-form | $\Lambda = -\frac{1}{2} F \wedge H$ | $\Lambda = -\frac{1}{2} F^\alpha \wedge H_\alpha$ |
| Constitutive tensor | $\chi^{\alpha\beta\gamma\delta}$ | $\chi^{\beta\gamma}{}_\alpha{}^{\nu\rho}{}_\mu$ |

### 3.1. *Gravitational constitutive tensor*

The local and linear constitutive law

$$\check{H}^{\alpha\beta}{}_{\mu} = \frac{1}{2}\chi^{\alpha\beta}{}_{\mu}{}^{\rho\sigma}{}_{\nu}\, F_{\rho\sigma}{}^{\nu} \tag{8}$$

describes a rich class of gravity models. The symmetry properties $\chi^{\beta\gamma}{}_{\alpha}{}^{\nu\rho}{}_{\mu} = -\chi^{\gamma\beta}{}_{\alpha}{}^{\nu\rho}{}_{\mu} = -\chi^{\beta\gamma}{}_{\alpha}{}^{\rho\nu}{}_{\mu}$ yield a simple number count $(6\times 4)\times(6\times 4) = 576$ of components of the constitutive tensor. The constitutive relation is *reversible* when $\chi^{\beta\gamma}{}_{\alpha}{}^{\nu\rho}{}_{\mu} = \chi^{\nu\rho}{}_{\mu}{}^{\beta\gamma}{}_{\alpha}$, with a reduced number of components: $24(24+1)/2 = 300$.

Splitting $\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{+}{\chi}{}^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} + \overset{-}{\chi}{}^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}$ into reversible and irreversible parts,

$$\overset{+}{\chi}{}^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := \frac{1}{2}(\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} + \chi^{\gamma\delta}{}_{\nu}{}^{\alpha\beta}{}_{\mu}), \quad \overset{-}{\chi}{}^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} := \frac{1}{2}(\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} - \chi^{\gamma\delta}{}_{\nu}{}^{\alpha\beta}{}_{\mu}), \tag{9}$$

one derives a decomposition $\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \sum_{I=1}^{8} {}^{[I]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}$ of the constitutive tensor into six irreducible parts:

$${}^{[1]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{+}{\chi}{}^{\alpha\beta}{}_{(\mu}{}^{\gamma\delta}{}_{\nu)} - \overset{+}{\chi}{}^{[\alpha\beta}{}_{(\mu}{}^{\gamma\delta]}{}_{\nu)}, \quad {}^{[2]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{-}{\chi}{}^{\alpha\beta}{}_{[\mu}{}^{\gamma\delta}{}_{\nu]} - \overset{-}{\chi}{}^{[\alpha\beta}{}_{[\mu}{}^{\gamma\delta]}{}_{\nu]}, \tag{10}$$

$${}^{[3]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{-}{\chi}{}^{\alpha\beta}{}_{(\mu}{}^{\gamma\delta}{}_{\nu)}, \qquad {}^{[4]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{+}{\chi}{}^{\alpha\beta}{}_{[\mu}{}^{\gamma\delta}{}_{\nu]}, \tag{11}$$

$${}^{[5]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{+}{\chi}{}^{[\alpha\beta}{}_{(\mu}{}^{\gamma\delta]}{}_{\nu)}, \qquad {}^{[6]}\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu} = \overset{-}{\chi}{}^{[\alpha\beta}{}_{[\mu}{}^{\gamma\delta]}{}_{\nu]}. \tag{12}$$

We call ${}^{[1]}\chi$ a reversible symmetric principal part (principal-1), ${}^{[4]}\chi$ a reversible antisymmetric principal part (principal-2), ${}^{[5]}\chi$ a reversible axion (axion-1), ${}^{[2]}\chi$ a skewon antisymmetric principal part (skewon-1), ${}^{[3]}\chi$ a skewon symmetric principal part (skewon-2), and ${}^{[6]}\chi$ a skewon axion (axion-2).

Comparing to electromagnetism, we have three parts in Maxwell's theory (7): $\chi^{\mu\nu\alpha\beta} = \frac{1}{2}\epsilon^{\mu\nu\rho\sigma}\kappa_{\rho\sigma}{}^{\alpha\beta} = {}^{(1)}\chi^{\mu\nu\alpha\beta} + {}^{(2)}\chi^{\mu\nu\alpha\beta} + {}^{(3)}\chi^{\mu\nu\alpha\beta}$: the principal ${}^{(1)}\chi$ piece and the axion ${}^{(3)}\chi$ (both reversible); and the skewon ${}^{(2)}\chi$ (irreversible).

When the metric exists on spacetime, the most general linear constitutive tensor

$$\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}(g) = \frac{1}{\varkappa}\Big[\beta_1\, g^{\gamma[\alpha}g^{\beta]\delta}g_{\mu\nu} + \beta_2\,\delta^{[\alpha}_{\mu}g^{\beta][\gamma}\delta^{\delta]}_{\nu} + \beta_3\,\delta^{[\alpha}_{\nu}g^{\beta][\gamma}\delta^{\delta]}_{\mu} \tag{13}$$

$$+\beta_4\,\varepsilon^{\alpha\beta\gamma\delta}\,g_{\mu\nu} + \beta_5\,\varepsilon^{\alpha\beta[\gamma}{}_{[\mu}\,\delta^{\delta]}_{\nu]} + \beta_6\left(\delta^{[\alpha}_{(\mu}\varepsilon^{\beta]\gamma\delta}{}_{\nu)} - \delta^{[\gamma}_{(\mu}\varepsilon^{\delta]\alpha\beta}{}_{\nu)}\right)\Big] \tag{14}$$

encompasses four irreducible parts: principal-1 proportional to $\beta_1$ and $(\beta_2+\beta_3)$, principal-2 proportional to $\beta_5$ and $(\beta_2-\beta_3)$, skewon-2 $\sim\beta_6$, and axion-1 $\sim\beta_4$.

### 3.2. *Propagation of gravitational waves*

In a geometric optics approximation, $F^{\alpha} = \mathcal{F}^{\alpha}e^{i\Phi}$ with rapidly varying phase $\Phi$ and slowly changing amplitude $\mathcal{F}^{\alpha}$ (eikonal ansatz). Alternatively, in Hadamard's theory, wave is described as a discontinuity across characteristic hypersurface (wave front). In any case, vacuum field equations reduce to algebraic system for amplitudes

$$dH_{\alpha} = 0,\ dF^{\alpha} = 0 \qquad \Longrightarrow \qquad \chi^{\mu\nu}{}_{\alpha}{}^{\rho\sigma}{}_{\beta}\,\mathcal{F}_{\rho\sigma}{}^{\beta}\,q_{\nu} = 0,\ \epsilon^{\mu\nu\rho\sigma}\,\mathcal{F}_{\rho\sigma}{}^{\alpha}\,q_{\nu} = 0, \tag{15}$$

with the wave covector $q = d\Phi = q_\mu \vartheta^\mu$. As a result, $\mathcal{F}_{\rho\sigma}{}^\alpha = A_\rho{}^\alpha q_\sigma - A_\sigma{}^\alpha q_\rho$, and from (15) for the amplitude $A_\mu{}^\nu$ we find a *characteristic equation*

$$M^\mu{}_\alpha{}^\nu{}_\beta \, A_\nu{}^\beta = 0\,, \qquad M^\mu{}_\alpha{}^\nu{}_\beta := \chi^{\mu\rho}{}_\alpha{}^{\nu\sigma}{}_\beta \, q_\rho \, q_\sigma. \tag{16}$$

The dispersion relation (Fresnel equation) arises as a solvability condition of (16). Decomposing $A_\alpha{}^\beta = \mathring{A}_\alpha{}^\beta + A\delta_\alpha^\beta$ into traceless part ($\mathring{A}_\alpha{}^\alpha = 0$) and trace $A = \frac{1}{4} A_\gamma{}^\gamma$:

$$N^\mu{}_\alpha{}^\nu{}_\beta \, \mathring{A}_\nu{}^\beta = 0, \qquad N^\mu{}_\alpha{}^\nu{}_\beta = M^\mu{}_\alpha{}^\nu{}_\beta M^\rho{}_\rho{}^\sigma{}_\sigma - M^\mu{}_\alpha{}^\rho{}_\rho M^\sigma{}_\sigma{}^\nu{}_\beta. \tag{17}$$

Notice the gauge freedom: $A_\nu{}^\beta \to A_\nu{}^\beta + q_\nu C^\beta$ leaves (16) invariant for any $C^\beta$.

In the metric-dependent case (14), for generic class of models with $2\beta_1 - \beta_2 - \beta_3 \neq 0$ and $2\beta_1 + \beta_3 \neq 0$ the characteristic equation splits into decoupled equations

$$(2\beta_1 - \beta_3) \left\{ 3 \left(q^2 \delta^\mu_\alpha - q^\mu q_\alpha\right) \left(q^2 \delta^\nu_\beta - q^\nu q_\beta\right) - (q^2 g_{\alpha\beta} - q_\alpha q_\beta) q^\mu q^\nu \right\} \mathring{A}_{(\mu\nu)} = 0, \tag{18}$$

$$(2\beta_1 + \beta_3) \left(q^2 \delta^\mu_\alpha - q^\mu q_\alpha\right) \left(q^2 \delta^\nu_\beta - q^\nu q_\beta\right) \mathring{A}_{[\mu\nu]} = 0, \tag{19}$$

and the scalar mode is recovered from $3q^2 A = \mathring{A}_\nu{}^\mu q_\mu q^\nu$. Therefore, only symmetric $\mathring{A}_{(\mu\nu)}$ mode is dynamical if $2\beta_1 = -\beta_3$, whereas only antisymmetric $\mathring{A}_{[\mu\nu]}$ mode remains dynamical when $2\beta_1 = \beta_3$.

The teleparallel equivalent general relativity model $\mathrm{GR}_{||}$ is a special case when both $2\beta_1 - \beta_2 - \beta_3 = 0$ and $2\beta_1 + \beta_3 = 0$. Explicitly: $\beta_4 = \beta_5 = \beta_6 = 0$ and

$$\beta_1 = -1\,, \qquad \beta_2 = -4\,, \qquad \beta_3 = 2. \tag{20}$$

For $\mathrm{GR}_{||}$ coupling constants (20), the characteristic equation (16) reduces to

$$q^2 h_{\alpha\beta} - q_\alpha q^\gamma h_{\gamma\beta} - q_\beta q^\gamma h_{\alpha\gamma} = 0, \tag{21}$$

which describes the usual spin 2 graviton mode $h_{\alpha\beta} = A_{(\alpha\beta)} - \frac{1}{2} g_{\alpha\beta} A_\gamma{}^\gamma$.

### 3.3. *Recovering General Relativity*

The teleparallel equivalent $\mathrm{GR}_{||}$ theory (20) is distinguished among other coframe models by the following remarkable properties:

- Field equations are invariant under local Lorentz group [5,6]
  $\vartheta^\alpha \longrightarrow L^\alpha{}_\beta(x)\vartheta^\beta, \quad L^\alpha{}_\beta \in SO(1,3)$.
- Black hole solutions exist [7].

The gravitational Lagrangian of $\mathrm{GR}_{||}$ actually reduces to Hilbert's Lagrangian $\Lambda = -\frac{1}{8}\chi^{\beta\gamma}{}_\alpha{}^{\rho\sigma}{}_\mu F_{\beta\gamma}{}^\alpha F_{\rho\sigma}{}^\mu = -\frac{1}{2\varkappa} R +$ total derivative. The constitutive tensor has only two nontrivial irreducible parts – principal-1 and principal-2.

*Premetric gravity summarized.* Fundamental equations are metric-free:

$$dH_\alpha = \Sigma_\alpha, \qquad dF^\alpha = 0. \tag{22}$$

Metric enters only in linking (constitutive law) equations (8), where [8] for $\mathrm{GR}_{||}$:

$$\chi^{\alpha\beta}{}_\mu{}^{\gamma\delta}{}_\nu = \frac{1}{\varkappa} \left( -g^{\gamma[\alpha} g^{\beta]\delta} g_{\mu\nu} - 4\delta^{[\alpha}_\mu g^{\beta][\gamma} \delta^{\delta]}_\nu + 2\delta^{[\alpha}_\nu g^{\beta][\gamma} \delta^{\delta]}_\mu \right). \tag{23}$$

## 4. Conclusions and outlook: On Kottler's path

Following the pioneering contributions by Kottler, Cartan and van Dantzig, the premetric program works universally in physics, embracing classical particle mechanics, kinetic theory, and (most importantly) electromagnetism and gravity. A feasible premetric gravity theory can be constructed on the basis of the energy-momentum conservation law as a teleparallel coframe model. The corresponding general local and linear constitutive relation encompasses *six* irreducible parts (2 principal, 2 skewon, and 2 axion).

In premetric approach, one can view constitutive tensor $\chi^{\alpha\beta}{}_{\mu}{}^{\gamma\delta}{}_{\nu}$ as an independent variable. This opens new perspectives in electromagnetism and gravity theory such as:

- Natural extensions with *axion, skewon, and dilaton* fields[9]
- Natural extension to *parity odd* contributions[5,10]
- Convenient framework to discuss *Lorentz violating* models[11]
- Nonlinear gravity models, $f(R)$ and $f(T)$, in particular[12,13]
- Nonlocal constitutive laws[14,15]


### Acknowledgments

This talk is essentially based on the papers[2–4] written in collaboration with Friedrich Hehl, Yakov Itin and Jens Boos. I am grateful to Marcus Werner for the kind invitation and support which made it possible for me to attend MG15.


### References


1. F.W. Hehl and Yu.N. Obukhov, *Foundations of Classical Electrodynamics: Charge, Flux, and Metric* (Birkhäuser: Boston, 2003).
2. F.W. Hehl, Y. Itin, Yu.N. Obukhov, *Int. J. Mod. Phys. D* **25**, 1640016 (2016).
3. Y. Itin, F.W. Hehl, Yu.N. Obukhov, *Phys. Rev. D* **95**, 084020 (2017).
4. Y. Itin, Yu.N. Obukhov, J. Boos, F.W. Hehl, *Eur. Phys. J. C* **78**, 907 (2018).
5. C. Pellegrini, J. Plebanski, *Mat. Fys. Skr. Dan. Vid. Selsk.* **2**, no. 4 (1963).
6. Y.M. Cho, *Phys. Rev. D* **14**, 2521 (1976).
7. Yu.N. Obukhov, J.G. Pereira, *Phys. Rev. D* **67**, 044016 (2003).
8. R. Ferraro, M.J. Guzmán, *Phys. Rev. D* **94**, 104045 (2016).
9. W.-T. Ni, *Int. J. Mod. Phys. D* **25**, 1630002 (2016).
10. F. Müller-Hoissen, J. Nitsch, *Phys. Rev. D* **28**, 718 (1983).
11. V.A. Kostelecký, M. Mewes, *Phys. Lett. B* **779**, 136 (2018).
12. S. Capozziello, M. De Laurentis, *Phys. Reports* **509**, 167 (2011).
13. M. Hohmann, L. Järv, M. Krššák, C. Pfeifer, *Phys. Rev. D* **97**, 104042 (2018).
14. B. Mashhoon, *Nonlocal gravity* (Oxford University Press: Oxford, 2017).
15. D. Puetzfeld, Yu.N. Obukhov, F.W. Hehl, *Constitutive law of nonlocal gravity*, arXiv:1903.04023 (2019).

# Observers' measurements of time and length in premetric electrodynamics

Christian Pfeifer

*Laboratory of Theoretical Physics, Institute of Physics, University of Tartu, W. Ostwaldi 1 50411 Tartu, Estonia*
*E-mail: christian.pfeifer@ut.ee*

The notion of observers' and their measurements is closely tied to the Lorentzian metric geometry of spacetime, which in turn has its roots in the symmetries of Maxwell's theory of electrodynamics. Modifying either the one, the other or both ingredients to our modern understanding of physics, requires also a reformulation of the observer model used. In this presentation we will consider a generalized theory of electrodynamics, so called local and linear premetric, or area metric, electrodynamics and its corresponding spacetime structure. On this basis we will describe an observer's measurement of time and spatial length. A general algorithm how to determine observer measurements will be outlined and explicitly applied to a first order premetric perturbation of Maxwell electrodynamics. The latter contains for example the photon sector of the minimal standard model extension. Having understood an observer's measurement of time and length we will derive the relativistic observables time dilation and length contraction. In the future a modern relativistic description of the classical test of special relativity shall be performed, including a consistent observer model.

*Keywords*: premetric electrodynamics, relativity, observers, radar experiment, time dilation, length contraction

## 1. From electrodynamics to relativity

The origins of special and general relativity and, in particular, of the relation between observers by Lorentz transformations, lies in Maxwell's theory of electrodynamics. It predicts the propagation of light on the integral curves of the directions forming the null cones cones of a Lorentzian metric, and the Lorentz transformations are those transformations which leave this structure invariant, globally for the flat Minkowski metric, locally for any non-flat Lorentzian metric g encoding gravity. Thus historically, first there was a viable relativistic matter field theory which described the behaviour of the electromagnetic field correctly, and then, a theory of gravity consistent with the relativity principles of the matter field theory was constructed[1].

Modifications and extensions of general relativity, and of the field theories forming the standard model of particle physics, are numerous and disperse in various directions. However, it is rarely discussed what happens to the description of observers and their measurements, when one changes the theoretical models of gravity or matter. In approaches where gravity is still encoded into a spacetime metric and matter fields are minimally coupled to this metric, such as f(R)-gravity or other modifications of the Einstein-Hilbert action only involving the metric[2], the observer description from general relativity is still applicable. But, when one changes the field which encodes gravity away from the metric, or introduces couplings of this

field to matter fields, like in the standard model extension (SME) framework[3,4], the Robertson-Mansouri-Sexl (RMS) framework[5,6], Finsler geometries[7] or premetric electrodynamics[8,9], the influence of the added concepts on observers, their measurements and their relation between each other has to be investigated and a consistent observer model has to be constructed.

Here we summarize how one obtains an observer's measurement of time and spatial lengths starting from a theory of electrodynamics in the following algorithm, using the techniques developed in the articles[10,11]:

I. Derive the geometric optics limit of the theory of electrodynamics.
II. Derive the Lagrange functions $L^{\#}$ and $L^{*}$, which define the motion of massless and massive particles on the manifold, from the geometric optic limit.
III. Use $L^{*}$ to realize the clock postulate, i.e. to identify the proper time normalization of observer worldlines $x$ by choosing $L^{*}(\dot{x}) = 1$.
IV. Model the radar experiment by demanding that for a spatial direction $X$ the vectors $N^{+} = \ell_{\dot{x}}(X)\dot{x} + X$ and $N^{-} = \tilde{\ell}_{\dot{x}}(X)\dot{x} + X$ are the tangents of the light rays of the radar signal, i.e. are null-vectors of $L^{\#}$.
V. A solution of $L^{\#}(N^{\pm}) = 0$ defines the radar length $\mathcal{L}_U(X) = \ell_{\dot{x}}(X) + \tilde{\ell}_{\dot{x}}(X)$ an observer on worldline $x$ associates to an object represented by the vector $X$.

To demonstrate how the general algorithm works in practice we apply it to first order modification of Maxwell electrodynamics caledl weak premetric electrodynamics.

Throughout this article we use the convention that $\eta = \mathrm{diag}(1, -1, -1, -1)$.

## 2. Weak premetric electrodynamics

The field equations of weak premetric electrodynamics are[12]

$$(\eta^{ac}\eta^{bd} + \mathcal{K}^{abcd})\partial_b F_{cd} = 0, \quad F_{cd} = \partial_{[c}A_{d]}\,, \tag{1}$$

the tensor $\mathcal{K}^{abcd}$ parametrizes the deviations from metric electrodynamics and has the following properties $\mathcal{K}^{abcd} = \mathcal{K}^{[ab][cd]} = \mathcal{K}^{[cd][ab]}$ and $\mathcal{K}^{a[bcd]} = 0$.

**Step I. Plane waves: The geometric optic limit.** Wave covectors $k$ of the plane waves solving the field equations are the roots of the Fresnel polynomial $\mathcal{G}(k)$[13]

$$\mathcal{G}(k) = \eta^{-1}(k,k)^2 - \eta^{-1}(k,k)\mathcal{K}(k,k) + \frac{1}{2}\big(\mathcal{K}(k,k)^2 - \mathcal{J}(k,k,k,k)\big)\,. \tag{2}$$

The tensor fields appearing are derived from the perturbation tensor defining the field equations

$$\mathcal{K}(k,k) = \mathcal{K}^{ac}k_a k_c = \mathcal{K}^{abc}{}_b k_a k_c, \tag{3}$$

$$\mathcal{J}(k,k,k,k) = \mathcal{J}^{acef}k_a k_c k_e k_f = \mathcal{K}^{abcd}\mathcal{K}^{e}{}_{b}{}^{f}{}_{d}k_a k_c k_e k_f\,. \tag{4}$$

Observe that, in order to study the correction in the propagation of light predicted by the perturbation of the field equations to first order, it is necessary to consider the

Fresnel polynomial to second order. For the following, in particular the condition $\det(\partial_{k_a}\partial_{k_b}\mathcal{G}(k)) \neq 0$ must hold[10].

**Step II.: The determination of the Lagrange functions $L^{\#}$ and $L^{*}$.** The Lagrangian which determines the propagation of light is given by the dual polynomial of $\mathcal{G}(k)$, i.e. by the unique polynomial satisfying $L(\dot{x}(k)) = Q(k)\mathcal{G}(k)$, where $\dot{x}^a(k) = \partial_{k_a}\mathcal{G}(k)$. Employing $\dot{x}^\flat = \eta(x,\cdot)$ we find

$$\begin{aligned} L^{\#}(\dot{x}) &= \eta(\dot{x},\dot{x})^2 + 3\mathcal{K}(\dot{x}^\flat,\dot{x}^\flat)\eta(\dot{x},\dot{x}) \\ &\quad + \frac{1}{2}(3\mathcal{J}(\dot{x}^\flat,\dot{x}^\flat,\dot{x}^\flat,\dot{x}^\flat) + 4\mathcal{K}(\dot{x}^\flat,\dot{x}^\flat)^2 + 5\mathcal{K}_{ac}\mathcal{K}^c{}_b\dot{x}^a\dot{x}^b\eta(\dot{x},\dot{x})) \qquad (5) \\ &= \eta(\dot{x},\dot{x})^2 + h_1(\dot{x},\dot{x})\eta(\dot{x},\dot{x}) + h_2(\dot{x},\dot{x},\dot{x},\dot{x})\,. \qquad (6) \end{aligned}$$

The abbreviations $h_1$ and $h_2$ were introduced to display the radar length formula later in a compact form. The motion of observers and their geometric clock is given by a second Lagrangian $L^*$ which can be obtained from the Helmholtz action

$$S[x,k,\lambda] = \int \mathrm{d}\tau\, \Big(k_a\dot{x}^a - \lambda \ln\big(\mathcal{G}(\tfrac{k}{m})\big)\Big), \qquad (7)$$

by eliminating $k$ and the Lagrange multiplier $\lambda$ with help of the corresponding equations of motion. Doing so we find the length functional

$$S[x] = \int \mathrm{d}\tau\; L^*(\dot{x}) = \int \mathrm{d}\tau\; \sqrt{\eta(\dot{x},\dot{x})}\left(1 + \frac{1}{4}\frac{\mathcal{K}(\dot{x}^\flat,\dot{x}^\flat)}{\eta(\dot{x},\dot{x})}\right). \qquad (8)$$

Physically this parametrization invariant action integral measures the observer's proper time, and mathematically speaking $L^*$ is a Finsler function.

**Step III.: The clock postulate.** An observer worldline is a cuvre $x(\tau)$ whose tangent satisfies $L^*(\dot{x}) = 1$, or explicitly $\eta(\dot{x},\dot{x}) = 1 - \frac{1}{2}\mathcal{K}(\dot{x}^\flat,\dot{x}^\flat)$. The directions $X$ which are spatial w.r.t. $x$ satisfy $X^a\partial_{\dot{x}^a}L^* = \eta(\dot{x},X) + \frac{1}{2}\mathcal{K}(\dot{x}^\flat,X^\flat) = 0$.

## 3. The radar length

**Step IV.: Setup the radar experiment.** An observer on a wordline $x(\tau)$ with tangent $\dot{x}$ emits light along the curve with tangent $N^+$ towards the end of an object, which is modelled by a vector $X$ spatial w.r.t. $\dot{x}$. There the light gets reflected and propagates back to the observer along the curve with tangent $N^-$. The observer measures the time of flight of the light between its emission and return. This time is the radar length the observer associates to the object. The tangents of the light curves $N^+ = \ell_{\dot{x}}(X)\dot{x} + X$ and $N^- = \tilde{\ell}_{\dot{x}}(X)\dot{x} - X$ must satisfy $L^{\#}(N^+) = 0 = L^{\#}(N^+)$, which can be solved for $\ell_{\dot{x}}(X)$ and $\tilde{\ell}_{\dot{x}}(X)$.

**Step V.: Solve for the radar length.** The radar length then is given by $\mathcal{L}_{\dot{x}}(X) = \ell_{\dot{x}}(X) + \tilde{\ell}_{\dot{x}}(X)$.

The influence of $L^{\#}$ on the radar experiment is obvious, however it is important to stress that also $L^*$ enters the calculation by the demands that $\dot{x}$ satisfies $L^*(\dot{x}) = 1$ and the fact that $X$ shall be spatial w.r.t. $\dot{x}$, see the comments below (8).

Using the explicit form of $L^{\#}$ and $L^{*}$ in (5) and (8), the desired radar length is given by, see[12] for the derivation,

$$\mathcal{L}_{\dot{x}}(X)_{\sigma\tilde{\sigma}} = 2\sqrt{-\eta(X,X)} \tag{9}$$
$$+\frac{\epsilon}{4}\frac{\left(\sigma\sqrt{B(X,\dot{x})}+\tilde{\sigma}\sqrt{\tilde{B}(X,\dot{x})}\right)}{\sqrt{-\eta(X,X)}}-\frac{\epsilon}{2}\frac{h_1(X,X)-2h_1(\dot{x},\dot{x})\eta(X,X)}{\sqrt{-\eta(X,X)}}.$$

The abbreviations $B$ and $\tilde{B}$ are lengthy functions of the perturbation tensors $h_1$ and $h_2$ defined in (6), they are displayed in[12]. The labels $\sigma$ and $\tilde{\sigma}$ can each take the values $+1$ or $-1$ and label which polarization of light has been used for $N^{+}$ resp. $N^{-}$. For example $\sigma=\tilde{\sigma}$ yields the radar length deermined by the same polarization used for $N^{+}$ and $N^{+}$, while $\sigma=-\tilde{\sigma}$ represents the case when the reflection at the end of the object changes the polarization of the light.

## 4. Relativistic Observables

Having clarified an observers measurement of time and spatial lengths we can derive the classical relativistic observables time dilation and length contraction. To do so, consider two observers on inertial wordlines $x_1$ and $x_2$, which have met at some point in spacetime. There they synchronized their clocks. After a proper time $t_1$ the first observer may decompose the tangent of the worldline of the second observer in terms of an equal time displacement vector $X=t_1V$ as $t_2\dot{x}_2=t_1(\dot{x}_1+V)$, where $V$ is the relative velocity between the observers.

The relation between the proper time $t_2$ of the second observer and the proper time $t_1$ then is

$$t_2L^{*}(\dot{x}_2)=t_2=t_1L^{*}(\dot{x}_1+V)=t_1\sqrt{1+\eta(V,V)}\left(1+\epsilon\frac{1}{12}\frac{h_1(V,V)}{(1+\eta(V,V))}\right). \tag{10}$$

Most interestingly, for weak premetric theories of electrodynamics for which $h_1^{ab}=\mathcal{K}^{acb}{}_{d}=0$, no change in the time dilation between observers compared to special relativity appears.

The calculation of the length contraction is more involved. Let observer $x_1$ carry a spatial object $Y$ to which it associates the radar length $\mathcal{L}_{\dot{x}_1}(Y)$. With respect to the observer $x_2$ the object $Y$ is not spatial, so first one needs to determine the projection $\bar{Y}$ of $Y$ onto the spatial directions of $x_2$. Then, $x_2$ can associate the radar length $\mathcal{L}_{\dot{x}_2}(\bar{Y})$ to $\bar{Y}$. The length contraction factor is given by the fraction $\frac{\mathcal{L}_{\dot{x}_1}(Y)}{\mathcal{L}_{\dot{x}_2}(\bar{Y})}$. The result is a complicated function of the perturbations $h_1$ and $h_2$ which can be found in[12]. The interesting observation is that, in contracts to the time dilation, even for $h_1^{ab}=\mathcal{K}^{acb}{}_{d}=0$ the effect does acquire a first order modification compared to the special relativistic result.

## 5. Conclusion

A modification of the geometry away from metric spacetime geometry requires a careful revision of the observer model used to compare observations with theoretical predictions. The example of the radar length demonstrates nicely that, in particular for Finslerian spacetime geometries it is not viable to simply exchange the Lorentzian metric employed in general relativity with the Finslerian metric evaluated at the observer curve on the tangent bundle for a description of observables.

The future task for spacetime geometries based on a Finsler function or a general dispersion relation is to find a description of observers from these quantities and to clarify the transformations mapping them onto each other.

## Acknowledgments

This talk is based on the article[12] written in collaboration with Norman Gürlebeck.

## References

1. A. Einstein and H. A. Lorentz and H. Minkowski and H. Weyl, "The Principle of Relativity: A Collection of Original Memoirs on the Special and General Theory of Relativity," Dover Publications, NewYork, 1952.
2. S. Nojiri, S. D. Odintsov and V. K. Oikonomou, Phys. Rept. **692** (2017) 1 doi:10.1016/j.physrep.2017.06.001 [arXiv:1705.11098 [gr-qc]].
3. D. Colladay and V. A. Kostelecky, Phys. Rev. D **58** (1998) 116002 doi:10.1103/PhysRevD.58.116002 [hep-ph/9809521].
4. R. Bluhm, Lect. Notes Phys. **702** (2006) 191 doi:10.1007/3-540-34523-X_8 [hep-ph/0506054].
5. H. P. Robertson, Rev. Mod. Phys. **21** (1949) 378. doi:10.1103/RevModPhys.21.378
6. R. Mansouri and R. U. Sexl, Gen. Rel. Grav. **8** (1977) 497. doi:10.1007/BF00762634
7. C. Pfeifer and M. N. R. Wohlfarth, Phys. Rev. D **85** (2012) 064009 doi:10.1103/PhysRevD.85.064009 [arXiv:1112.5641 [gr-qc]].
8. F. W. Hehl and Y. N. Obukhov, "Foundations of Classical Electrodynamics," Birkhäuser, 2003.
9. F. P. Schuller, C. Witte and M. N. R. Wohlfarth, Annals Phys. **325** (2010) 1853 doi:10.1016/j.aop.2010.04.008 [arXiv:0908.1016 [hep-th]].
10. D. Raetzel, S. Rivera and F. P. Schuller, Phys. Rev. D **83** (2011) 044047 doi:10.1103/PhysRevD.83.044047 [arXiv:1010.1369 [hep-th]].
11. C. Pfeifer, Phys. Rev. D **90** (2014) no.6, 064052 doi:10.1103/PhysRevD.90.064052 [arXiv:1408.5306 [gr-qc]].
12. N. Gürlebeck and C. Pfeifer, Phys. Rev. D **97** (2018) no.8, 084043 doi:10.1103/PhysRevD.97.084043 [arXiv:1801.07724 [gr-qc]].
13. G. F. Rubilar, Annalen Phys. **11** (2002) 717 doi:10.1002/1521-3889(200211)11:10/11<717::AID-ANDP717>3.0.CO;2-6 [arXiv:0706.2193 [gr-qc]].

# Covariant constructive gravity

Tobias Reinhart* and Nils Alex

*Department Physik, Friedrich-Alexander Universität Erlangen-Nürnberg, 91058 Erlangen, Germany*
**E-mail: tobi.reinhart@fau.de*

We present a method of constructing perturbative equations of motion for the geometric background of any given tensorial field theory. Requiring invariance of the gravitational dynamics under spacetime diffeomorphisms leads to a PDE system for the gravitational Lagrangian that can be solved by means of a power series ansatz. Furthermore, in each order we pose conditions on the causality of the gravitational equations, that ensure coevolution of the matter fields and the gravitational background is possible, i.e. gravitational equations and matter equations share the same initial data hypersurfaces.

*Keywords*: Constructive gravity; diffeomorphism invariance; causality

## 1. Introduction

Given a matter field theory in terms of an action functional $S_{mat}[\Phi; G)$ for a tensorial matter field $\Phi$ whose geometric background is provided by an additional tensorfield $G$, the problem of constructive gravity consists of completing the matter field equations with EOM for $G$ that make the coupled system of matter dynamics and supplemented gravitational dynamics a predictive theory. In contrast to the existing canonical framework described in Ref. 1 , we outline in the following how this problem can be tackled perturbatively from the spacetime covariant point of view.

## 2. Perturbative diffeomorphism invariance

Let $M$ be a smooth 4-dimensional manifold that represents spacetime. The gravitational field is described as a tensorfield $G$ on $M$, i.e. as a section of a bundle $\pi_F : F \to M$, with adapted coordinate functions $(x^m, v_A)$, where $F \subset T^m_n M$ for some $m$ and $n$. The equations we want to construct for $G$ shall be given as Euler-Lagrange-Equations of an up to now unknown Lagrangian $L : J^1F \to \Lambda^4 M$, which is a bundle map from the first order jet bundle over $F$ with coordinates $(x^m, v_A, v_{Ap})$ to the bundle of volume forms on $M$. One of the key requirements we pose on the to-be-constructed gravitational EOM is the invariance under spacetime diffeomorphisms. The action of $\mathrm{Diff}(M)$ on M extends naturally to $F$ by pullback and can then be lifted to the first jet bundle in standard ways.[3] Doing so, given $f \in \mathrm{Diff}(M)$ that acts on $M$, we can construct bundle automorphisms $f_*$ and $j^1 f_*$ that act on $F$ and $J^1F$, respectively. Diffeomorphism invariance of the theory is given if the Lagrangian is equivariant w.r.t. the lifted action on $J^1F$ and the action by pullback on $\Lambda^4 M$, i.e. for all $f \in \mathrm{Diff}(M)$ it holds that:

$$L \circ j^1 f_* = f_* \circ L. \tag{1}$$

On the Lie algebra level, (1) can be shown to be equivalent to the following set of first order partial differential equations for $L$:[3]

$$\begin{aligned} 0 &= L_{:m} \\ 0 &= L^{:A} C^{Bm}_{An} v_B + L^{:Ap} C^{Bm}_{An} v_{Bp} - L^{:Am} v_{An} + L\delta^m_n \\ 0 &= L^{:A(p|} C^{B|m)}_{An} v_B, \end{aligned} \tag{2}$$

where the constant tensor $C^{Bm}_{An}$ is obtained from the lie-algebra action of vector fields $\xi$ on $F$, $\xi_F = \xi^m \partial_m + C^{Bm}_{An} V_B \partial_m \xi^n \partial^A$ and hence depends on the specific nature of $G$. Here and in the following we use the notation $L_{:m} := \partial_m L$, $L^{:A} := \partial^A L$, etc. Later on we are going to use the equivariance condition of the Lagrangian to explicitly construct gravitational EOM. It is then often advantageous to work directly on the level of the EOM. Given a Lagrangian $L(x^m, v_A, v_{Ap})$ we obtain the corresponding EOM by applying the variational derivative to $L$ and equating to zero:

$$\begin{aligned} 0 =& E^A(x^m, v_A, v_{Ap}, v_{Apq}) := \frac{\delta L}{\delta v_A} = L^{:A} - D_m(L^{:Am}) \\ & D_m := \partial_m + v_{Am} \cdot \partial^A + v_{Apm} \cdot \partial^{Ap}. \end{aligned} \tag{3}$$

The EOM are now described by a function $E^A$ on $J^2F$. As consequence of (1) the EOM satisfy the following first order partial differential equation system:

$$\begin{aligned} 0 =& E^A_{:m} \\ 0 =& E^{A:B} C^{Cm}_{Bn} v_C + E^{A:Bp} C^{Cm}_{Bn} v_{Cp} - E^{A:Bm} v_{Bn} \\ & + E^{A:Bpq} C^{Cm}_{Bn} v_{Cpq} - 2E^{A:Bpm} v_{Bpn} + E^A \delta^m_n + E^B C^{Am}_{Bn} \\ 0 =& E^{A:B(p|} C^{C|m)}_{Bn} v_C + E^{A:B(p|q} C^{C|m)}_{Bn} v_{Cq} - E^{A:Bpm} v_{Bn} \\ 0 =& E^{A:B(pq|} C^{C|m)}_{Bn} v_C. \end{aligned} \tag{4}$$

The problem of constructing diffeomorphism invariant theories for a given tensor field $G$ is thereby translated into the equivalent problem of finding solutions to the equivariance equations (2) on the Level of the Lagrangian or the equations (4) on the level of the EOM. Details concerning the derivation of these equations can be found in Ref. 2. In most cases of physical interest, obtaining general solutions to either one of the two systems of equations is strikingly difficult. Therefore we are going to construct perturbative solutions, more precisely we are seeking power series expansions of the Lagrangian (or the EOM) that solve the equations (2) (or the equations (4)) up to the desired order.[4] We choose an expansion point $x_0$ that is induced by the flat Minkowski metric $\eta_{ab}$, i.e. $x_0 = (x^m, N_A, 0, ..)$, where the coordinates $N_A = N_A(\eta_{ab})$ are functions of the flat Minkowski metric and the derivative coordinates are zero. This allows the perturbatively constructed theory of gravity to be interpreted as an expansion around flat Minkowski spacetime. Discharging any explicit $x^m$ dependency that is prohibited due to the first equation in (2) from the very beginning, the most general expansion of the Lagrangian is

given as:

$$\begin{aligned} L = {} & L^{:A:B}\big|_{x_0}(v_A - N_A)(v_B - N_B) + L^{:Ap:Bq}\big|_{x_0} v_{Ap} v_{Bq} \\ & + L^{:A:B:C}\big|_{x_0}(v_A - N_A)(v_B - N_B)(v_C - N_C) \\ & + L^{:Ap:Bq:C}\big|_{x_0} v_{Ap} v_{Bq}(v_C - N_C) + \mathcal{O}(4). \end{aligned} \tag{5}$$

Here any linear terms are neglected as they do not contribute to the EOM. Terms with an odd number of indices are neglected as these can otherwise be eliminated using a result that can be derived from the equivariance equations:[2] If the chosen expansion point is eta-induced, the expansion coefficients in (5), i.e. $L^{:A:B}|_{x_0}$, etc. are linear combinations of sums of products of $\eta^{ab}$ and $\epsilon^{abcd}$. Inserting the induced expansion of the EOM into (4) and evaluating at $x_0$ yields a set of linear equations for the expansion coefficients that contribute to the EOM in linear order. Prolonging (4), i.e. deriving with respect to the coordinates of $J^2F$, inserting the expansion of the $E^A$ and again evaluating at $x_0$, we obtain a system of linear equations for the second order expansion coefficients. In total, by doing so we end up with the following set of linear equations for the linear order expansion coefficients:

$$\begin{aligned} 0 &= L^{:A:B}\big|_{x_0} C^{Cm}_{Bn} N_C \\ 0 &= \big|_{(pqm)} L^{:Ap:Bq}\big|_{x_0} C^{Cm}_{Bn} N_C, \end{aligned} \tag{6}$$

where $... = \big|_{(pqm)}$ denotes the total symmetrization in $pqm$ for the whole equation. In quadratic order we obtain:

$$\begin{aligned} 0 = {} & -2L^{:A:B}\big|_{x_0} C^{Cm}_{Bn} + 6L^{:A:B:C}\big|_{x_0} C^{Dm}_{Bn} N_D + 2L^{:B:C}\big|_{x_0} C^{Am}_{Bn} - 2L^{:A:C}\big|_{x_0} \delta^m_n \\ 0 = {} & \big|_{(pq)} 2L^{:Ap:Cm}\big|_{x_0} \delta^q_n + L^{:Ap:Bq}\big|_{x_0} C^{Cm}_{Bn} + L^{:Ap:Cq:B}\big|_{x_0} C^{Dm}_{Bn} N_D \\ & + L^{:Bp:Cq}\big|_{x_0} C^{Am}_{Bn} + L^{:Ap:Cq}\big|_{x_0} \delta^m_n \\ 0 = {} & \big|_{(pm)} L^{:Ap:Cm}\big|_{x_0} \delta^q_n - 2L^{:Ap:Bq}\big|_{x_0} C^{Cm}_{Bn} \\ & + \big[L^{:Bp:Cq:A}\big|_{x_0} - L^{:Ap:Cq:B}\big|_{x_0} - L^{:Bp:Aq:C}\big|_{x_0}\big] C^{Dm}_{Bn} N_D \\ 0 = {} & \big|_{(pqm)} L^{:Ap:Bq}\big|_{x_0} C^{Cm}_{Bn} + L^{:Ap:Bq:C}\big|_{x_0} C^{Dm}_{Bn} N_D. \end{aligned} \tag{7}$$

Note that the complex quest of solving the PDE system (4) which guarantees diffeomorphism invariance is now translated, at least perturbatively, into the much simpler task of solving a system of linear equations. Furthermore the presented equations take this form irrespective of the specific gravitational field $G$. The only quantity that depends on $G$ is the constant tensor $C^{Am}_{Bn}$ that can easily be computed once $G$ is specified. Solving this linear system can be simplified even further by again exploiting the fact that the expansion coefficients are built from $\eta^{ab}$ and $\epsilon^{abcd}$. We construct the most general such expression for each coefficient. These expressions obviously involve undetermined constants. Inserting into (7) yields a system of linear equations for these constants.

## 3. Examples of the construction procedure

In the following we present two examples of the previously described construction procedure.

### 3.1. *Metric gravity*

Given the action functional of a Klein-Gordon matter field on a fixed metric background

$$S_{\mathrm{KG}}[\Phi; g_{ab}) = \int \mathrm{d}^4x\sqrt{-g}g^{ab}\partial_a\Phi\partial_b\Phi, \tag{8}$$

we want to construct a second order expansion of diffeomorphism invariant EOM for $g_{ab}$ around $\eta_{ab}$. Starting with the linear order equations (6), the first step is computing the most general expressions built from $\eta^{ab}$ and $\epsilon^{abcd}$ for the two linear expansion coefficients. Using the induced symmetries of these one readily finds that:

$$L^{:ab:cd}\big|_{x_0} = \mu_1 \cdot \eta^{ab}\eta^{cd} + \mu_2 \cdot \eta^{c(a}\eta^{b)d}. \tag{9}$$

In similar ways we compute the expression for $L^{:abp:cdq}$ with constants $\nu_1, .., \nu_4$. Inserting in (6) and solving for the unknown constants we get:

$$\begin{aligned} &\mu_1 = 0,\ \mu_2 = 0 \\ &\nu_2 = -\nu_1,\ \nu_3 = 2\nu_1\ \nu_4 = -2\nu_1. \end{aligned} \tag{10}$$

Performing the same computation for the quadratic order expansion coefficients and the corresponding equations (7), we are left with one overall constant $\nu_1$ which is irrelevant once the EOM are equated to zero. Furthermore the two terms in the expansion of $E^A$ that have no contributions of derivatives of $g_{ab}$ are identically zero.[a] Comparing the obtained quadratic order expansion of the EOM with the Einstein field equations we conclude that the two systems coincide.

### 3.2. *Area metric gravity*

As a second example we consider the action functional of general linear electrodynamics, in short GLED. The action is given as

$$S_{\mathrm{GLED}}[A_m; G_{abcd}) = \int \mathrm{d}^4x\omega_G G^{abcd}F_{ab}F_{cd}, \tag{11}$$

where $A$ is the electromagnetic potential, $F = dA$ is the field strength, $G$ is the so called area metric, a (0,4) tensor field with the symmetries $G_{abcd} = G_{cdab} = -G_{bacd}$ and $\omega_G$ is a density of weight 1 constructed from $G$. The components of the inverse

[a] In the following such terms are referred to as mass terms and the corresponding constants are called masses, whereas the remaining terms and constants are called kinetic terms and gravitational constants respectively.

area metric $G^{abcd}$ are defined by $G^{abcd}G_{cdef} = 4\delta^{[a}_{e}\delta^{b]}_{f}$. We choose $N_{abcd} = \eta_{ac}\eta_{bd} - \eta_{ad}\eta_{bc} - \epsilon_{abcd}$ as expansion point. With this choice, we recover—to zeroth order in the gravitational field—Maxwell electrodynamics on a Minkowski background. As before we construct ansätze for the expansion coefficients from $\eta^{ab}$ and $\epsilon^{abcd}$. Solving the equations (6) and (7) we obtain 3 masses and 7 gravitational constants in linear order, 5 additional masses and 36 additional gravitational constants in quadratic order. Details regarding this particular example can be found in Ref. 2.

## 4. Causal compatibility of gravitational and matter field equations

Apart from diffeomorphism invariance, the second crucial requirement that constructive gravity poses on the gravitational dynamics is their causal compatibility with the given matter theory.[1] Physically meaningful matter theories have hyperbolic EOM, i.e. given initial data on an initial data hypersurface in $M$, the EOM can be used to evolve this initial data and thereby uniquely predict the values of the matter field away from the hypersurface. Assuming this for the given matter theory, one can construct a certain polynomial function on the symmetric product $S^k(T^*M)$, the so called principal polynomial of the matter equations $P$. From the vanishing set of $P$ a convex cone in each $T^*_xM$ is constructed. It can be shown that precisely those hypersurfaces that have at each point a co-normal vector lying inside this cone are admissible initial data hypersurfaces.[1,4] If the gravitational EOM shall complete the given matter dynamics to a predictive theory, we must require that there exist initial data hypersurfaces that are common to both EOMs. This is certainly the case if the principal polynomial of the gravitational EOM has the same vanishing set as the principal polynomial of the matter theory. In the following we are going to explore how this additional requirement might or might not restrict the gravitational constants and masses that were obtained in the previous two examples. We start with metric gravity. The principal polynomial of the Klein-Gordon theory is, up to an overall factor, simply given by

$$P_{\mathrm{KG}}(k) = g^{ab}k_ak_b = \eta^{ab}k_ak_b - \eta^{ac}\eta^{bd}h_{cd}k_ak_b + \mathcal{O}(2), \tag{12}$$

where we introduced $h_{ab} := g_{ab} - \eta_{ab}$ and expanded the polynomial in powers of $h_{ab}$. Computing the principal polynomial of the perturbative expansion of metric gravity obtained before, we find that

$$P_{\mathrm{metric}}(k) = \left\{(\eta^{ab}k_ak_b - \eta^{ac}\eta^{bd}h_{cd}k_ak_b)\right\}^2 + \mathcal{O}(2). \tag{13}$$

Obviously this polynomial describes the same vanishing set as $P_{\mathrm{KG}}$. Hence we get no additional condition from the causality requirement. This is not too surprising as except from an overall gravitational constant there is no freedom left in the perturbative metric EOM in the first place.

The principal polynomial of GLED was first computed by Rubilar[5] and is given by

$$\begin{aligned} P_{\mathrm{GLED}}(k) &= -\frac{1}{24}\omega_G^2 \epsilon_{mnpq}\epsilon_{rstu} G^{mnra} G^{bpsc} G^{dqtu} k_a k_b k_c k_d \\ &= \big[1 - A\eta(H) - \frac{1}{2}A\epsilon(H) + \frac{1}{12}\epsilon(H)\big]\eta(k)^2 - H(k)\eta(k) + \mathcal{O}(2) \\ &= \big\{\big[1 - \frac{1}{2}A\eta(A) - \frac{1}{4}A\epsilon(H) + \frac{1}{24}\epsilon(H)\big]\eta(k) - \frac{1}{2}H(k)\big\}^2 + \mathcal{O}(2), \end{aligned} \tag{14}$$

where $H_{abcd} := G_{abcd} - N_{abcd}$, $\eta(H) := \eta^{ac}\eta^{bd}H_{abcd}$, $\epsilon(H) := \epsilon^{abcd}H_{abcd}$, $H(k) := \eta^{ac}\eta^{bp}\eta^{cq}H_{abcd}k_p k_q$, $\eta(k) := \eta^{pq}k_p k_q$ and $A$ is a constant that depends on the specific choice for the density $\omega_G$. Note that, as we are only interested in the vanishing set of the principal polynomial, we can multiply a given polynomial with a non vanishing density of arbitrary weight. Such a density admits the general form

$$X = 1 + b_1 \cdot (\eta(H) + \frac{1}{2}\epsilon(H)) + \frac{b_2}{12}\epsilon(H) + \mathcal{O}(2), \tag{15}$$

for arbitrary constants $b_1$ and $b_2$. Multiplying (14) with such a density we see that perturbatively the GLED polynomial describes the same vanishing set as

$$\begin{aligned} \widetilde{P}_{\mathrm{GLED}}(k) = X \cdot P_{\mathrm{GLED}} = \big\{\big[1 - \frac{1}{2}(A - b_1)\eta(H) - \frac{1}{4}(A - b_1)\epsilon(H)) + \\ \frac{1 + b_2}{24}\epsilon(H)\big]\eta(k) - \frac{1}{2}H(k)\big\}^2 + \mathcal{O}(2). \end{aligned} \tag{16}$$

The principal polynomial of the constructed, diffeomorphism invariant expansion of area metric gravity to second order is calculated as

$$P_{\mathrm{area}} = \big\{\big[1 - \frac{1}{2}C\eta(H) - \frac{1}{4}C\epsilon(H) + \frac{7}{12 \cdot 13}\epsilon(H)\big]\eta(k) - \frac{1}{2}H(k)\big\}^{13} + \mathcal{O}(2), \tag{17}$$

where $C$ is an expression that depends on the gravitational constants of the theory, but not on $H_{abcd}$.[2] Comparing (16) and (17) shows that, multiplying the GLED polynomial with a density with constants $b_1 = A - C$ and $b_2 = \frac{1}{13}$, the two polynomials are products of the same factor and therefore describe the same vanishing set in $\mathcal{O}(2)$. Hence already the required diffeomorphism invariance fixes the principal polynomial of area metric gravity to the point where, perturbatively, the EOM have the same initial data hypersurfaces as those of GLED.

## References


1. M. Düll, F.P. Schuller, N. Stritzelberger and F. Wolz, (Phys. Rev. D 97, 084036, 2018), `arXiv:1611.08878 [gr-qc]`.
2. N. Alex and T. Reinhart, in preparation.
3. M.J. Gotay, J. Isenberg, J.E. Marsden and R. Montgomery, `arXiv:physics/9801019 [math-ph]`, 2004.
4. W.M. Seiler and R.W.Tucker, (J.Phys. A28 4431-4452, 1995), `arXiv:hep-th/9506017 [hep-th]`.
5. G.F. Rubilar, Y.N. Obukov and F.W. Hehl, (Int. J. Mod. Phys. D11, 1227, 2002), `arXiv:gr-qc/0109012 [gr-qc]`.

# Area metric gravity: Confrontation with observations of galaxy rotation curves

Hans-Martin Rieser* and Björn M. Schäfer

*Zentrum für Astronomie der Universität Heidelberg,*
*Astronomisches Rechen-Institut*
**E-mail: hans-martin.rieser@uni-heidelberg.de*

Frederic P. Schuller

*Constructive Gravity Group*
*FAU Erlangen, Germany*

In the young field of constructive gravity, the gravity theory constructed from general linear electrdynamics has been frequently used as a physically meaningful text case for the entire approach. Based on a Newtonian approximation of area metric gravity, we use the rotation curves of low surface brightness galaxies to determine ranges for the gravitational constants in this weak field limit. Thus we are able to show that even by using very simple models for the distribution of visible matter in the galaxies, Newtonian area metric gravity yields reasonable rotation curves and consistent estimates for the gravitational constants. Area-metric gravity turns out to not be an alternative to dark matter, but to reduce the dark matter fraction and to remove the need for a distribution subtantially different from visible matter.

*Keywords*: constructive gravity, rotation curves, dark matter, area metric

## 1. Introduction

The rotation properties of spiral galaxies have puzzeled astronomers for about 100 years now.[1] Without the introduction of dark matter, Newtonian physics is not able to explain the flattened behaviour of the velocity curve for large distances from the centre. Until today, neither have dark matter particles been detected, nor has a modified theory been found which all by itself could explain all phenomena connected to dark matter.

In this work, we will assess whether a modified theory using the constructive gravity algorithm yields flat rotation curves in a universe with area metric matter. We will also look for estimates on the gravitational constants under the assumption that gravity is not governed by general relativity but its area metric refinement instead.

## 2. Newtonian Area Metric Gravity

The constructive gravity programme offers a fundamental connection between the matter content of a universe and the dynamics of its spacetime structure. Indeed using the toolset given by the construction equations,[2] one may take a matter theory and construct the corresponding theory of gravity.

Area metric gravity is constructed – in this fashion – as the gravity theory that underlies a refinement of Maxwellian electrodynamics, namely the general linearized

electrodynamics defied by the Lagrangian

$$\mathcal{L}_{\mathrm{GLED}} = \omega_G G^{\alpha\beta\gamma\delta} F_{\alpha\beta} F_{\gamma\delta}, \tag{1}$$

with the tensor $G^{\alpha\beta\gamma\delta} = G^{([\alpha\beta][\gamma\delta])}$ being antisymmetric in the first and the second index pair and symmetric to an exchange of both pairs. To date, no general solution is known for the construction equations that one needs to set up for this matter theory in order to find the dynamics of the fourth-rank tensor geometry encoded in $G$. However, the equations have been solved for a linear perturbation around the standard metric geometry.[2] Ref. 2 also finds the 3D screen manifold action

$$\begin{aligned} S_{\mathrm{grav,GLED}}(\lambda) = -\int \mathrm{d}^4x \frac{m}{2}\gamma_{\alpha\beta}\dot{\lambda}^\alpha\dot{\lambda}^\beta + mA(\lambda(t)) + m\Omega_{\alpha\beta}(\lambda(t))\dot{\lambda}^\alpha\dot{\lambda}^\beta \\ + \frac{m}{2}A(\lambda(t))\gamma_{\alpha\beta}\dot{\lambda}^\alpha\dot{\lambda}^\beta \end{aligned} \tag{2}$$

for the motion of a point particle with trajectory $\lambda(t)$ and mass $m$ on a weak area metric spacetime, where the fields $A(\lambda)$ and $\Omega(\lambda)$ are found from the construction equations. Starting from this solution, many known phenomena can be reassessed to check for a possible realization of area metric gravity. Many astrophysical objects even do not require a full relativistic form as in eq. 2, but can be evaluated using a small velocity and weak field limit[3]

$$\mathcal{L}_{\mathrm{grav,GLED}} \approx -\frac{m}{2}\gamma_{\alpha\beta}\dot{\lambda}^\alpha\dot{\lambda}^\beta + mA(\lambda(t)) \tag{3}$$

with a point particle potential

$$A(r) = -\frac{G}{|r-r'|}\left(1 + \nu e^{-\mu|r-r'|}\right),$$

that will be referred to as the "Newtonian limit" further on. This Yukawa correction term introduces a gravitational length scale with the constant $\mu$.

## 3. Galaxy Rotation Curves

### 3.1. *Symmetric potentials*

To determine galaxy rotation curves, one needs to find the orbital velocities of test bodies within a galactic potential $\Phi$. For simplicity, the potential is assumed to be cylindrically symmetric around the axis of rotation. For a spiral galaxy, two basic types of models can be distinguished: Core-centered models use a spherical mass distribution with a disk of negligible mass, while disk models only contain a flat disk without a central bulge.

The rotation curves for spherical distributions can be calculated analytically, or at least symbolically, for a wide range of density distributions. As the shell theorem does not hold for the area metric correction, contributions from outside the test

mass orbit arise. We get the general expression

$$\Phi[\rho](r) = \tilde{\Phi}(M|_r) - \frac{4\pi\nu G}{\mu r}\left(e^{-\mu r}\int_0^r \mathrm{d}r'\ r'\rho(r')\sinh(\mu r') + \sinh(\mu r)\int_r^\infty \mathrm{d}r'\ r'\rho(r')e^{-\mu r'}\right) \tag{4}$$

for the sphere's gravitational potential[3] with $\bar{\Phi}$ being the standard Newtonian potential. Potentials of disklike distributions can be assessed by using a multipole expansion. Now the rotation curves may be calculated via the relation

$$v^2 = r\frac{\partial\Phi}{\partial r}. \tag{5}$$

### 3.2. *Models and methods*

In this study three common spherical models were used. While the isothermal model is a simple model for a dark matter halo, the Hernquist and Plummer models are more realistic for visible matter distrbutions. The disklike approach was done with a Kuzmin disk. The following table shows the mass density distributions for the four models:

Table 1. Models used for the galactic fits

| | |
|---|---|
| Isothermal | $\rho_I(r) = \frac{\rho_0 a^2}{r^2}$ |
| Hernquist | $\rho_H(r) = \frac{Ma}{2\pi r(a+r)^3}$ |
| Plummer | $\rho_P(r) = \frac{3Ma^2}{4\pi(a^2+r^2)^{\frac{5}{2}}}$ |
| Kuzmin disk | $\rho_K(z,R) = \delta(z)\frac{aM}{2\pi(a^2+R^2)^{\frac{3}{2}}}$ |

Using the emcee Monte Carlo Markov Chain package,[4] galaxy rotation curves were fit to the velocity distributions of 22 low surface brightness galaxies from Ref. 5.

Three such fits were performed. The first fit, with the weak field standard metric theory, was used for comparison and for generating starting values for the gravitational constants. The second fit, with the area metric corrections, yielded individual best fits for the 22 galaxies. The third fit, using the weighted mean of the best-fit values for the gravitational constants from step two, was done to check whether this mean generates consistent rotation curves.

### 3.3. *Individual galactic fits*

While only the isothermal sphere – and for some galaxies the Kuzmin disk – yielded adequate rotation curves, in the area metric regime all four models provide reasonable fits for the galaxies with a sufficient sample size (21 of 22).

For the fits using the weighted mean gravitational constants $\mu$ and $\nu$, we find that most galaxies have a good fit in this parameter range (see Fig. 1). However, Hernquist and Plummer spheres show a bump in their rotation curves for small radii. In seven datasets this small bump is also present in the data, the other galaxies do not show this behaviour. There are two possible explanations. On the one hand, there might be no data available for such small radii where the galaxy shows this bump, on the other hand, this might be a sign for a misfit in the density model. This could mean that the core behaviour cannot be described with these simple models or that some galaxies are more disklike and others are bulge dominated.

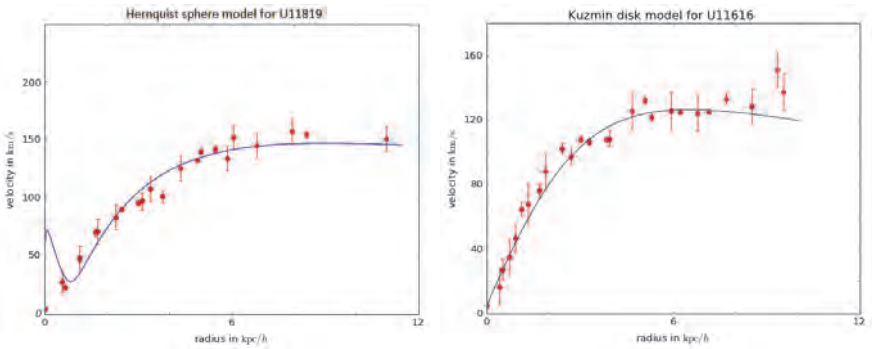



Fig. 1. Sample fits for a Hernquist and a Kuzmin model with area metric correction.

### 3.4. *Implications for area metric gravity*

The four models provide reasonable fits for most of the low surface brightness galaxy data used. The isothermal sphere model reproduces the metric result with vanishing area metric correction. This was expected as the isothermal model already provided a good fit in the first run using Newton's theory.

However, in area metric gravity both density distributions that are close to a typical distribution of visible matter also yield reasonable results for rotation curves over the observable distance range without the need for some particular dark matter distribution. Furthermore, the total galactic mass within area metric gravity is up to 30% lower in comparison to the metric results, reducing the amount of dark matter needed for explaining the observations.

The weighted mean gravitational constants using the spread of the MCMC-samples gives results that differ by about two standard deviations from zero (see Fig. 2), except for the isothermal sphere result. The remaining three models yield comparable constants:

Combining the results we find the guesses $\mu \approx 0.023 \pm 0.010\ kpc^{-1}$ and $\nu \approx 0.011 \pm 0.05$ for area metric gravity. This is not yet a significant difference from zero, but shows a trend that is worth to be investigated further.

Table 2. Area metric gravitational constants for different galaxy model fits

| | $\mu \pm \sigma_\mu$ [h/kpc] | $\nu \pm \sigma_\nu$ |
|---|---|---|
| Isothermal | n/a | $0.000002 \pm 0.00010$ |
| Plummer | $0.025 \pm 0.014$ | $0.017 \pm 0.009$ |
| Hernquist | $0.021 \pm 0.011$ | $0.0021 \pm 0.0028$ |
| Kuzmin | $0.008 \pm 0.002$ | $0.0044 \pm 0.0013$ |

Due to the exponential behaviour of the correction term, effects of area metric gravity are restricted to phenomena near the intrinsic length scale $1/\mu$ of the theory. Compared to the current uncertainty in Newton's constant, the correction does not play a role for phenomena four orders of magnitude off the intrinsic scale. This makes the Newtonian effects of area metric gravity basically unmeasurable in a laboratory. On cosmological scales, however, one would observe a slightly altered gravitational constant by a factor $1 + \nu$, which makes the Newtonian limit of area metric gravity a MOND-like theory.

The theory's scaling behaviour leads to another insight: If the intrinsic length scale is not within the range of galaxy diameters, the galactic dynamics would not be affected and area metric rotation curves would be the same as for standard Newtonian gravity.

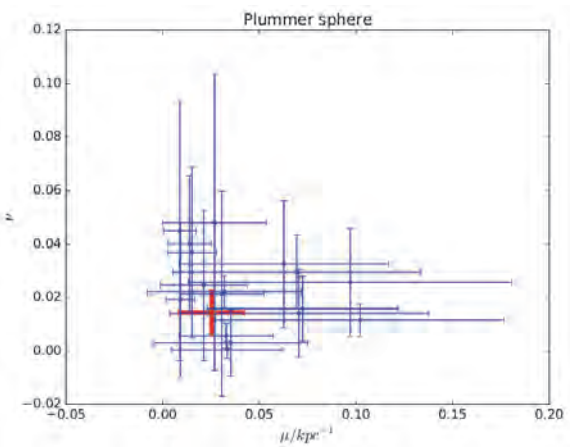



Fig. 2. Weighted mean area metric gravitational constants for a Plummer sphere model.

## 4. Outlook

Several topics connected to the rotation curve issue should be inspected more closely in the future. On the one hand, the density distributions used are mostly based on random motion in a metric universe. These may not be stable anymore for an area

metric geometry which may introduce additional corrections. On the other hand, further phenomena should be used to determine better range estimates for the new gravitational constants.

There are three areas of interest for a determination of the constants. The first is astrophysical phenomena based on linearized area metric gravity like the one presented here, including relativistic phenomena. The second area is effects based on symmetry reduced solutions. In particular, area metric cosmology has gained some attention so far, but has not been resolved yet completely. The third area is the inspection of the matter side. As matter and gravitational dynamics are linked, non-metric quantum effects may be used to restrict the complimentary gravitational constants.[6]

## References

1. F. Zwicky, *Helvetica Physica Acta* **6**, 1933 (2009).
2. `arXiv:gr-qc/1708.03870v1` *Gravitational Closure of Weakly Birefringent Electrodynamics*, 2017.
3. H-M. Rieser and B. Schäfer and F. Schuller, in preparation, 2019.
4. `arXiv:astro-ph/1202.3665v4` *emcee: The MCMC Hammer*, 2013.
5. S. McGaugh and V. Rubin and W. de Blok, *Astron. J.* **122**, 2001.
6. `arXiv:hep-ph/1703.07183v2` Quantum signatures of ray-optically invisible non-metricities.

# Constructive gravity: Foundations and applications

Frederic P. Schuller

*Department of Applied Mathematics, University of Twente,*
*P.O. Box 217, 7500 AE Enschede, The Netherlands*
*E-mail: f.p.schuller@utwente.nl*

Constructive gravity allows to calculate the Lagrangian for gravity, provided one previously prescribes the Lagrangian for all matter fields on a spacetime geometry of choice. We explain the physical and mathematical foundation of this result and point out how to answer questions about gravity that previously could not be meaningfully asked.

*Keywords*: canonical gravity; standard model extension; hyperbolic equations; causal structure; generalized geometry; dispersion relation; modified gravity; quantum field theory, hydrogen hyperfine structure, gravitational waves.

## 1. Introduction

Constructive gravity reveals a connection between matter dynamics and gravitational dynamics, which is deeper than previously appreciated. The key result[1] is this: Three physically mild conditions on a given action

$$S_{\text{matter}}[A,G) \tag{1}$$

for a matter field $A$ (on which the action depends locally) on a geometric background described by one or several tensor fields $G$ (on which the action depends only ultralocally), suffice to derive a universal system of linear homogeneous partial differential equations whose coefficient functions depend on the given matter action and whose solutions provide diffeomorphism-invariant actions

$$S_{\text{geometry}}[G] \tag{2}$$

for the geometry that are causally consistent with the initially prescribed matter dynamics and must be be added to the latter in order to provide a closed system of equations for both the matter and the geometry.

The mechanism behind this causally consistent closure of given matter field dynamics is simple. First, the matter dynamics unequivocally determine[2] — no matter how complicated the background geometry on which they are formulated and as long as elementary well-posedness and energy conditions hold — all ways in which the spacetime may be foliated into initial data hypersurfaces for the matter degrees of freedom. The only further step then consists in a straightforward technical exploitation[3] of the requirement[4] that the geometric background be furnished with canonical dynamics that evolve the pertinent *geometric* degrees of freedom between any two of the initial data hypersurfaces for the stipulated *matter* dynamics, thus making these the common initial data hypersurfaces for both the geometry and the matter fields. Without this requirement, no sustained prediction for the total matter-geometry dynamics could possibly be made. With this requirement, one has a condition so strong as to typically determine the action for the geometry

up to only a few constants of integration. After adding the thus obtained action for the geometry to the matter action, the resulting coupled field equations describe how the geometry is influenced by the presence of dynamical matter fields (which amounts to the generation of a gravitational field) and, conversely, how the motion of matter is influenced by the geometry (which amounts to the influence exerted by gravity on matter).

The thus constituted constructive gravity program hence simply takes a matter action as an input and provides a canonically compatible gravitational action as its output. Starting the program with standard model matter[5], one obtains the Einstein-Hilbert action with two undetermined constants in place of the gravitational and the cosmological constant. Starting the program, instead, with modifications of standard model matter, a correspondingly modified action for the underlying geometry can be obtained and provides the gravity theory for all geometric degrees of freedom that is selected by causal consistency. In any case, gravitational field dynamics are revealed to be a mere consistency condition, imposed by the entirety of matter field dynamics one postulates and tailor-made for the geometric background fields employed by the matter field equations. In programmatic brevity, the philosophy underlying constructive gravity is: Matter first, gravity second.

### 1.1. *A previously not solvable problem*

In order to appreciate the far-reaching implications of what is said above, consider the following phenomenologically interesting observational scenario. Assume that future advances in radioastronomy reveal that electromagnetic waves propagating through vacuum regions of space suffer birefringence effects of various strengths, however small, while there is no indication at all that there would be any violation of the linear superposition principle in the observed energy range. This scenario presents one small step for the matter phenomenologist, but one giant leap for the gravitational theorist. It is easy for the matter phenomenologist, since the most general electrodynamics action that generates linear field equations on a tensorial background geometry takes the form[6]

$$S_{\text{matter}}[A, G) = \int d^4x\, \omega_G\, G^{abcd} F_{ab} F_{cd}\,, \tag{3}$$

where $A$ is the familiar electromagnetic gauge potential with the associated field strength $F = dA$, while $G$ is an, at first arbitrary, fourth-rank tensor field and $\omega_G = (\epsilon_{ijkl} G^{ijkl})^{-1}$ is a weight-one scalar density constructed from it. The quadratic appearance of the field strength $F$ and its two-form character render only those components of $G$ relevant that conform to the algebraic symmetry conditions

$$G^{abcd} = G^{cdab} \quad \text{and} \quad G^{abcd} = -G^{bacd}\,, \tag{4}$$

while finiteness of the density factor imposes the open condition $\epsilon_{ijkl} G^{ijkl} \neq 0$. Any observed birefringence effect can now be modelled in one way or another by suitably

adapting the 21 independent components of the tensor field $G$ at each point within the spacetime region of the electromagnetic wave such as to fit the obervational data. Given the described hypothetical observations, this is certainly the correct classical matter model[7]. But in this form, it is yet of little predictive power, since both the location and strength of vacuum birefringence and further effects beyond Maxwellian electrodynamics entirely depend on the values taken by the fourth-rank tensor field $G$. Without a way to predict the values taken by the tensor field $G$, one is thus not able to predict the electromagnetic field either.

The only way to predict the values of the geometric field $G$ (up to equivalence under diffeomorphisms) with the least possible prejudice is to furnish $G$ with dynamics of its own, by extending the action (3) to the total action

$$S[A,G] = S_{\text{matter}}[A,G) + S_{\text{geometry}}[G] \tag{5}$$

such that the stationarity conditions

$$\frac{\delta S_{\text{matter}}}{\delta A_a}[A,G) = 0 \qquad \text{and} \qquad \frac{\delta S_{\text{geometry}}}{\delta G^{abcd}}[G] = -\frac{S_{\text{matter}}}{\delta G^{abcd}}[A,G) \tag{6}$$

recover both the phenomenologically enforced general linear electromagnetic field equations and the gravitational field equations. The central problem to solve in order to make this work, of course, is the identification of all physically consistent choices of the action $S_{\text{geometry}}[G]$ for the geometry. But finding the gravitational actions that can underpin given matter dynamics is the very problem solved by constructive gravity.

### 1.2. *A familiar problem solved a century ago*

It will be enlightening to see how standard general relativity arises, in the philosophy of constructive gravity, from standard model matter. Other than for the previous example, the reader will not need to have mastered the general machinery described in this article in order to follow the steps of the constructive gravity program in some more detail for the present case, because one simply recovers known concepts from general relativity – even if from a slightly different conceptual perspective: all mathematical objects of the general theory reduce here to their familiar form. Simplifying as much as possible for the purpose of clarity, we consider the matter field action

$$S_{\text{matter}}[\varphi, g) = \int d^4x \sqrt{\det g}\, \left(g^{ab}\partial_a\varphi\partial_b\varphi - m^2\varphi^2\right) \tag{7}$$

for a scalar field $\varphi$ and a second rank tensor field $g$, about which we do not need to assume anything a priori beyond the symmetry and non-degeneracy conditions

$$g^{ab} = g^{ba} \qquad \text{and} \qquad \det g \neq 0\,. \tag{8}$$

In order to find the elementary well-posedness and energy conditions on the matter theory that need to be satisfied for the constructive gravity program to apply, one first calculates the principal polynomial of the postulated matter field equations in

each spacetime cotangent space (see section 2 for an outline of the general theory), which information is equivalent to the one held in an even-rank totally symmetric contravariant tensor field $P$. For the dynamics defined by (7), this principal tensor field happens to be of second rank and turns out to be given by

$$P^{ab} = g^{ab} \,. \tag{9}$$

From the point of view of the general theory, it is a pure coincidence that this principal tensor field $P$ has the same rank as the fundamental geometry $g$ and additionally that it is in fact identical to it. (That this truly is a coincidence is impressively illustrated by the fact that for the previously considered general linear electrodynamics, the principal tensor field $P^{abcd} = \omega_G^2 \epsilon_{mnpq} \epsilon_{rstu} G^{mnr(a} G^{b|ps|c} G^{d)qtu}$ is found to be cubic in the fundamental geometry $G$ underlying that matter theory and as a totally symmetric tensor certainly does not share its index symmetries, although it coincidentally has the same rank.) Indeed, some occurrences of the inverse metric in general relativity are in fact occurrences of the principal tensor field, while others are occurrences of the fundamental geometry. Failure to recognize this degeneracy as coincidental goes with impunity only in general relativity proper, but in fact lies at the heart of causality problems of various generalized theories of gravity and matter[8]. The general well-posedness and energy conditions[1,2] that a matter theory must satisfy in order to be a viable starting point for the constructive gravity program boil down, in the present special example, to the condition that the principal polynomial $P$ be of Lorentzian signature. That also the *fundamental geometry* $g$ must have Lorentzian signature is only due to the very particular coincidence (9).

It is obvious that for the previous subsection's general linear electrodynamics with its pertinent fourth-rank principal tensor $P$, a more general algebraic classification than the signature classification of symmetric bilinear forms needs to kick in and that the implications for the underlying fundamental geometry $G$ are even more intricate to extract, but ultimately obtainable[9]. From a practical point of view, here and in general, one may perfectly ignore the rather extensive theoretical machinery running in the background and write down the the so-called gravitational closure equations (in either functional differential or partial differential form) — which determine the desired gravitational Lagrangian within the constructive gravity program and are informed by the assumed matter theory — by ultimately employing only the pertinent fundamental geometry and the calculated principal polynomial in the coefficient functions of the closure equations. Their solution[1,3] for the present case of Klein-Gordon dynamics on a Lorentzian background yields – without any a priori knowledge of metric geometry whatsoever – the two-parameter family

$$S^{\kappa,\Lambda}_{\text{geometry}}[g] = \kappa \int d^4x \sqrt{\det g}\,(R - 2\Lambda) \tag{10}$$

of gravitational actions, which one recognizes as the Einstein-Hilbert action with both the gravitational and cosmological constant left to be determined by

experiment. The Ricci scalar, or rather its definition in terms of the metric, arises automatically in the solution of the closure equations, which are just informed about the matter dynamics though the principal polynomial of the latter and indeed the underlying fundamental geometry. In the parlance of the constructive gravity program, Einstein-Hilbert gravity arises as the gravitational closure of Klein-Gordon theory. The same result is obtained by starting from Maxwell theory or indeed the entire standard model[5].

## 2. Principal polynomials of matter field equations

We now turn to an exposition of the general theory, for which the previously mentioned examples present just special cases.

Starting point of the gravitational closure mechanism is a matter action of the form

$$S_{\text{matter}}[A, G] = \int_M d^4x\, \mathcal{L}(A(x), \partial A(x), \ldots, \partial^N A(x), G(x))\,, \tag{11}$$

where $\mathcal{L}$ is a scalar density, $A$ is a smooth $GL(4,\mathbb{R})$-irreducible tensor field (or, more generally, a finite collection of various such) representing the matter whose dynamics is determined by $S_{\text{matter}}$ and $G$ is a smooth tensor field (or, again, a finite collection of various such), to which we will refer as the geometry on the smooth four-dimensional manifold $M$. Note that we assume that our matter actions depend locally on the matter fields and ultralocally on the geometry.

Variation with respect to the matter field $A$ yields the tensor-density equations of motion

$$0 = \frac{\delta S_{\text{matter}}}{\delta A^{\mathcal{M}}(x)} \equiv \sum_{n=0}^{N} Q^{a_1\ldots a_n}_{\mathcal{M}\mathcal{N}}[A(x), G(x)]\partial_{a_1}\cdots\partial_{a_n} A^{\mathcal{N}}(x)\,, \tag{12}$$

where $A^{\mathcal{M}}(x)$ indicates components with respect to the $GL(4,\mathbb{R})$ representation space in which $A(x)$ takes its values at any point $x \in M$ and the $Q$ are $N+1$ coefficient functionals with local dependence of $A$ and $G$. For field equations that are linear in their highest order derivative term, the corresponding coefficient functions

$$Q^{a_1\ldots a_N}_{\mathcal{M}\mathcal{N}}(G(x)) \tag{13}$$

do not depend on the field $A$ at all, and only ultralocally on the geometry $G$. In order to avoid inessential technical complications, we restrict attention here to such matter models.

We first discuss the case in which the field equations (12) do not feature any hidden integrability conditions, which could otherwise alter the highest derivative coefficient functions. In this straightforward case, the causal structure of the matter field dynamics is encoded entirely in the functions (13). In particular, well-posedness for the matter field dynamics described by $S_{\text{matter}}$ on a fixed geometry $(M, G)$ — in other words, the question of whether one can find a suitable foliation of spacetime into hypersurfaces such that prescription of initial field data on one

such hypersurface suffices for the dynamics to predict the data on another, neighbouring, such hypersurface such that the resulting spacetime matter field solves the field equations — requires that the $a$-solution space of the infinite frequency limit Wentzel-Kramers-Brillouin condition

$$Q^{a_1\dots a_N}_{\mathcal{M}\mathcal{N}}(G(x))k_{a_1}\dots k_{a_N}a^{\mathcal{N}} = 0 \tag{14}$$

is at least $(S+1)$-dimensional, where $S$ is the dimension of the gauge orbits featured by the matter theory[1]. It can be shown[10,11] that this condition can always be written as a polynomial condition

$$\widetilde{P}^{i_1\dots i_{\deg \tilde{P}}}(x)k_{i_1}\cdots k_{i_{\deg \tilde{P}}} = 0 \qquad \text{for } k \in T^*_xM\,, \tag{15}$$

where $\widetilde{P}^{i_1\dots i_{\deg P}}(x)$ are the coefficient functions for some totally symmetric tensor density $P$ of weight one. Since this is a homogeneous condition we may and will regularly de-densitize this condition by use of some meaningful scalar density that can be constructed ultralocally in terms of the geometry $G$ and then denote the resulting totally symmetric tensor field by $P$. Since, in the language of the theory of partial differential equations, $P(x)$ is the principal polynomial (in the cotangent fibre variable $k$) of the field equations (12), we refer to $P$ as the principal tensor field.

In the presence of hidden integrability conditions – which are only revealed by systematically repeated differentiation and elimination of the equations that were originally obtained by variation – the highest order coefficient functions may be altered once the integrability conditions have been made explicit. The simplest example illustrating this is the system

$$A_x + A_{yy} = 0 \qquad \text{and} \qquad A_y + A_{yx} = 0\,, \tag{16}$$

for which only differentiation and elimination reveals the contained implication $A_{xx} - A_{yy} = 0$, which crucially alters the highest order coefficient functions of the equations of motion (even making them into a non-square matrix) and hence the calculation and final form of the principal tensor.

Extending a given system of partial differential equation, such that all hidden integrability conditions are made explicit, is achieved by the Cartan-Kuranishi algorithm[12]. From the highest order coefficient functions of a so obtained system, which is then termed involutive, one may then calculate the principal tensor, by a slight generalization of the prescription given before for systems without hidden integrability conditions. The algorithm for taking equations of motion $\Phi_B[A] = 0$ for fields $u^{\mathcal{M}}$ into involutive form revolves around repeated calculation of the so-called geometric symbols

$$(M_q)_{B\mathcal{N}}{}^{\nu} := \frac{\partial \Phi_B}{\partial A^{\mathcal{N}}{}_{,\nu}}[A] \qquad \text{for } q = \nu_1 + \cdots + \nu_{\dim M}\,, \tag{17}$$

where the derivative $A^{\mathcal{N}}{}_{,\nu}$ with respect to the multi-index $\nu = (\nu_1, \dots, \nu_{\dim M})$ denotes the $q$-th partial derivative $\partial_1^{\nu_1}\partial_2^{\nu_2}\dots\partial_{\dim M}^{\nu_{\dim M}}A^{\mathcal{N}}$ of the field component $A^{\mathcal{N}}$,

for $q$ being the currently highest derivative order of the intermediate system of equations generated in each step of the now easily performed Cartan-Kuranishi algorithm:

(1) Having calculated the components of the the geometric symbol for the current set of equations (starting with the initially given set of equations if no other set has been generated yet by the algorithm), they are arranged into a matrix $M_q$ whose rows are labeled by the index $B$ and whose columns are labelled by the combination of the indices $\mathcal{N}$ and $\nu$. The only rule for how this labelling is done is that the resulting column indices (non-strictly) decrease in the class $1 \leq c(\nu) \leq \dim M$ of the multiindex $\nu$, which is defined as the smallest $i$ for which $\nu_i$ is non-zero. The actual calculational step then consists in taking the thus constructed matrix to row echelon form by judicious linear combinations of row vectors only. For any $i = 1, \ldots, M$ one then reads off the set of coefficients

$$\beta_q^{(i)} := \text{number of pivot elements in all columns of class } i\,. \tag{18}$$

(2) Prolongate the current system, i.e., combine it with all $\dim M$ possible first order partial derivatives of each of its current equations and calculate the matrix $M_{q+1}$ for this prolongated system. This allows to determine whether the system, as it was before this last prolongation, is pre-involutive. This is the case if the beta coefficients satisfy the pre-involutivity condition

$$\sum_{i=1}^{\dim M} i\beta_q^{(i)} = \operatorname{rank}(M_{q+1})\,. \tag{19}$$

If this is not the case, consider the just calculated prolongated system the new current system and return to the first step with the thus updated system of equations. It is guaranteed that the above equality will be satisfied after a finite number of iterations on steps 1 and 2, in which case one then proceeds to step 3.

(3) Consider the prolongated system that has just been calculated to confirm that the pre-involutivity condition of step 2 has been satisfied, but still consider the unprolongated system as the current system. If no integrability condition (an equation of equal or lower derivative order than the current system that is linearly independent of the latter) can be derived from the prolongated systems, the current system is called involutive and the algorithm terminates. If, however, integrability conditions are found, they are appended to the current system and the such extended system is handed as the new current system to step 1 of the algorithm. It is guaranteed that after a finite number of iterations of steps 1, 2 and 3, the algorithm terminates.

Application of the algorithm to the system (16) yields beta-coefficients $\beta_2^{(1)} = 2$ and $\beta_2^{(2)} = 1$ and rank 3 for the prolongated system, which identifies the initially given system as already pre-involutive in step 2. Since the prolongated system, however,

turns up the integrability conditon $A_{xx} - A_{yy} = 0$, the system is not involutive yet. Adding the integrability condition to the original system and repeating steps 1 and 2 one obtains the new beta coefficients ${\beta'}_2^{(1)} = 2$ and ${\beta'}_2^{(2)} = 1$ and rank 4 for the prolonged system. Thus the original system extended by the found integrability condition is found pre-involutive in step 2 and one indeed finds no further integrability condition in step 3. Thus

$$A_x + A_{yy} = 0 \qquad \text{and} \qquad A_y + A_{yx} = 0 \qquad \textit{and} \qquad A_{xx} - A_{yy} = 0 \tag{20}$$

is the involutive form of the original system (16) as obtained by the Cartan-Kuranishi algorithm.

For equations of motion $\Phi_B[A] = 0$ that follow from a matter action by variation, but are not already involutive, a slight adaptation of the calculation of the principal tensor is required, since the principal symbol

$$T_{\overline{B}\mathcal{N}}(k) := \sum_{\nu_1+\cdots+\nu_{\dim M}=\overline{q}} \frac{\partial \Phi_{\overline{B}}}{\partial A^{\mathcal{N}}_{,\nu}} (k_1)^{\nu_1} \cdots (k_{\dim M})^{\nu_{\dim M}} \tag{21}$$

of their involutive form $\Phi_{\overline{B}}[A] = 0$ — where the index $\overline{B}$ now not only covers the range of the original $B$ but also all the additional equations that had to be added in order to achieve involutive form and $\overline{q}$ is the highest derivative order encountered in the involutive system — generically constitutes a non-square matrix $T(k)$. In any case, the principal tensor density can be shown the be determined in this case by taking the determinant of the Gramian matrix of $T(k)$,

$$\widetilde{P}^{a_1 \ldots a_{\deg \tilde{P}}} k_{a_1} \cdots k_{a_{\deg \tilde{P}}} := \det(T^t(k) T(k)) \,. \tag{22}$$

Note that the Gramian matrix is a square matrix whose rows and columns are labeled by some $GL(4, \mathbb{R})$ representation, so that its determinant transforms as a scalar density of the appropriate weight.

For technical reasons, and since it does not affect the information encoded in the principal scalar density $\widetilde{P}(k)$, whether obtained from (15) or more generally from (22), we will not only routinely de-densitize it by multiplication with a suitable density factor, but also reduce its degree as much as possible by dropping repeated factors, so that

$$\widetilde{P}(k) = \omega_G^m P_1^{n_1}(k) P_2^{n_2}(k) \cdots P_F^{n_F}(k) \,, \tag{23}$$

where $m$ is an integer and $n_1, \ldots, n_F$ are positive integers while $P_1(k), \ldots, P_F(k)$ are irreducible polynomials in $k$ that transform as scalar fields for any substitution of a covector field $k$, gives rise to the reduced principal tensor $P$ uniquely determined by

$$P^{a_1 \ldots a_{\deg P}} k_{a_1} \ldots k_{a_{\deg P}} := P_1(k) P_2(k) \cdots P_F(k) \,. \tag{24}$$

Only this reduced form of the principal tensor $P$ will play a role in constructiuve gravity and will simply be called the principal tensor of the matter field equations in the following.

## 3. Kinematics induced by matter field equations

Three technical conditions on any given matter field equations must be satisfied in order to derive a complete kinematical interpretation for the underlying spacetime geometry. They all concern the principal tensor $P$ of the matter field equations obtained from an action $S_{\text{matter}}[A, G)$, and thus implicitly impose conditions on the geometric tensor $G$ in terms of which $P$ is expressed. Physically, these conditions correspond to classically hardly negotiable necessary requirements for the field equations and their geometric optical limit, namely that (a) there exist initial data surfaces for the field equations, (b) the spacetime structure can be interpreted in a temporal-spatial way and (c) that the spacetime structure allows for time-orientability and corresponding energy-orientability. Technically, these physical conditions amount to

(a) The homogeneous polynomial $P_x : T^*_x M \to \mathbb{R}$ defined in terms of the principal tensor by $k \mapsto P^{a_1 \dots a_{\deg P}}(x) k_{a_1} \cdots k_{a_{\deg P}}$ at each point $x$ of the manifold $M$ must be hyperbolic. This means, by definition, that there exists some $h \in T^*_x M$ with $P_x(h) \neq 0$ such that for all $q \in T^*_x M$ the equation

$$P_x(q + \lambda h) = 0 \tag{25}$$

possesses $\deg P$ real solutions $\lambda_1, \dots, \lambda_{\deg P}$, counting algebraic rather than geometric multiplicity. One can show that if $h$ satisfies the above condition, then there is an entire connected set of such vectors which constitute an open and convex cone $C_x(h)$, the so-called hyperbolicity cone of $P_x$ that contains $h$. Note that the hyperbolicity cones are uniquely determined by the principal tensor and that the latter has been uniquely constructed in terms of the spacetime geometry $G$, in a way that is crucially informed by the particular equations of motion of the matter field $A$, but is functionally independent of the latter.
Clearly, one can choose any other covector $h'$ in this hyperbolicity cone as an alternative representative, since $C_x(h) = C_x(h')$. Moreover, since $-h$ satisfies condition (25) if and only if $h$ does, but since $P(h) \neq 0$, $-h$ does not lie in the same connected component as $h$ and thus $C_x(h) \cap C_x(-h) = \emptyset$; thus there is always an even number of hyperbolicity cones at each point of spacetime. A time-orientation of the spacetime is chosen by prescription of some smooth and everywhere hyperbolic covector field $n$, which singles out one particular cone $C_x := C_x(n_x)$ at spacetime point $x$. We call these cones the local observer cones, since they contain all possible conormals (which set one may geometrically think of as all possible tangent hyperplanes) to initial data hypersurfaces through the respective spacetime point to which they are attached. This, indeed, is the relation to the question of well-posedness of the field equations.
We impose a positive sign convention, which will come in handy later on, namely that $P_x(C_x) > 0$ for all $x \in M$. If this does not hold in the first place, then we necessarily have $P_x(C_x) < 0$ for all $x \in M$, due to the continuity of the time-orienting hyperbolic covector field $n_x$ and the continuity of the principal

tensor field $P$, so that replacing the principal tensor $P$ by $-P$, which has no effect on the hyperbolicity condition, arranges for the desired sign.

(b) The dual polynomial $P^\#_x : T_xM \to \mathbb{R}$, which is uniquely determined by a given hyperbolic polynomial $P_x : T^*_xM \to \mathbb{R}$ up to a non-zero factor (which turns out to cancel for any use the dual polynomial is put to) must be hyperbolic. The dual polynomial $P^\#_x$ of a hyperbolic polynomial $P_x = (P_1)_x \cdots (P_F)_x$, which is decomposed into mutually different irreducible factor polynomials $P_1, \ldots, P_F$, is defined as the product

$$P^\#_x := (P_1)^\#_x \cdots (P_F)^\#_x \,, \tag{26}$$

of the duals of these irreducible factors. Thus it suffices to define the dual polynomial $Q^\#$ associated with an irreducible polynomial $Q$, namely as the likewise irreducible polynomial for which

$$Q^\#_x(DQ_x(k)) = 0 \quad \text{holds for all } k \in T^*_xM \text{ with } Q_x(k) = 0 \text{ and } DQ_x(k) \neq 0 \,, \tag{27}$$

where $DQ_x$ denotes the derivative of $Q_x$ with respect to the cotangent space fibre at $x$. The very existence of the dual polynomial $P^\#_x$ hinges on the hyperbolicity of $P_x$, which is equivalent to the hyperbolicity of each irreducible factor polynomial. The immediate physical relevance of the dual polynomial is revealed by the real projective relation

$$[DP^\#_x([DP_x([k])])] = [k] \tag{28}$$

for all $k \in T^*_xM$ with $P_x(k) = 0$ and $DP_x(k) \neq 0$ and $DDP_x(k) \neq 0$, where $[\cdot]$ denotes projective equivalence, since it reveals that any projective $P_x$-null covector $[k]$ (safe such on an exceptional subset of measure zero) is bijectively mapped to a projective vector $[DP_x([k])]$, with the inverse map given by $[DP^\#_x]$. Due to the generic non-linearity of these maps, this is highly non-trivial. In physics language, however, it establishes a easily understood fact: for each $P^\#$-null ray direction, which emerges in the geometric-optical limit of the underlying hyperbolic matter field theory, there is a unique $P_x$-null wave surface and vice versa. We will refer to $P_x$-null covectors also as massless momenta.

(c) The positive energy cone $E^+_x$ at each point $x$ of the spacetime $M$ is defined for any hyperbolic principal polynomial $P_x$ with hyperbolic dual polynomial $P^\#_x$, as required by conditions (a) and (b), as the closed convex cone

$$E^+_x := \{e \in T^*_xM \,|\, e(X) > 0 \text{ for all } X \in \ell_x^{-1}(C_x)\} \,. \tag{29}$$

The physical rationale for this definition is clear: a massless or massive momentum $p$ can only be said to be of positive energy if all observers (represented here by all possible oberserver worldline tangent vectors $X$) agree on the sign of the respectively seen energy $e(X)$. Any massive momentum $p$, as defined under (b) above, is automatically of positive energy, by construction of $C^\#_x$.

The final condition on the principal polynomial, and thus the underlying spacetime geometry $G$, is that any massless momentum $p$ at some point $x \in M$ must

lie either in the positive energy cone $E_x^+$ or the negative energy cone $E_x^-$. This energy distinguishing condition is physically required, since it is necesary to have all observer agree on whether a decay that involves a massless particle is kinematically possibly or not.
Together with the previously adotped hyperbolicity condition on the dual polynomial $P_x^\#$, the energy distinguishing conditions allows to unquely identify the hyperbolicity cone $C_x^\# \subset T_xM$ of $P_x^\#$ for which the observer cone $C_x \subseteq \ell_x(C_x^\#)$, where $\ell_x$ denotes the invertible Legendre map

$$\ell_x : C_x^\# \to \ell_x(C_x^\#) \subset T_x^*M\,, \qquad X \mapsto -\frac{1}{\deg P_x^\#}\frac{DP_x^\#(X)}{P_x^\#(X)}\,. \tag{30}$$

The corresponding restriction of the inverse $\ell_x^{-1}$ of this map to the observer cone $C_x$ is physically easily understood as the bijective map between the massive momenta $p$ of mass $m$ at the point $x \in M$, which are characterized by $p \in C_x$ and $P_x(p) = m^{\deg P}$ for some positive mass $m$, and the tangent vectors $\ell_x^{-1}(p)$ of their respective worldlines.

Matter dynamics satifying the above three conditions impose the kinematical interpretation of the spacetime geometry $(M, G)$, with the relevant information coming from the matter dynamics being encoded in the principal polynomial. The three kinematical constructions of immediate practical importance are

(1) A local observer is given by a curve $e : (a, b) \longrightarrow LM$ in the spacetime frame bundle $\pi : LM \xrightarrow{\pi} M$ such that *(i)* the first frame vector $e_0(\lambda)$ lies in the Legendre dual $\ell^{-1}_{(\pi\circ e)(\lambda)}(C_{(\pi\circ e)(\lambda))})$ of the cotangent space observer cone for all $\lambda \in (a, b)$ and *(ii)* the other frame vectors $e_1(\lambda), e_2(\lambda), e_3(\lambda)$ are Legendre-orthogonal to $e_0(\lambda)$, which is to say that they lie in the kernel of $\ell_{(\pi\circ e)(\lambda)}(e_0(\lambda))$ for every $\lambda \in (a, b)$. Physically, this means that for every point of every initial data hypersurface, one can find a local observer whose worldline $\pi \circ e$ pierces the hypersurface at this point and whose worldline tangent vector $e_0$ is Legendre orthogonal to the hypersurface's tangent directions. These latter tangent directions are the purely spatial directions seen by this particular observer.
(2) A first order action for the worldline $x : \mathbb{R} \longrightarrow M$ of a massless point particle is immediately implied by the dispersion relation $P(k) = 0$ that must hold for the momentum of such a particle, namely

$$S_{\text{massless}}[x, k, \mu] = \int d\lambda\,[k_a(\lambda)\dot{x}^a(\lambda) - \mu(\lambda)P_{x(\lambda)}(k(\lambda))]\,, \tag{31}$$

where $\mu$ is a Lagrange multiplier. But solving the corresponding equations of motion requires solving for $k$. By virtue of the inverse Gauss map $[DP^\#]$ and the homogeneity of $P^\#$, one can indeed solve for $k = \sigma DP^\#$ in terms of another Lagrange multiplier $\sigma$ (which absorbs the projective scaling ambiguity). Encoding this elimination directly into the action one finds the equivalent second

oder action

$$S_{\text{massless}}[x,\sigma] = \int d\lambda\, \sigma(\lambda) P^{\#}_{x(\lambda)}(\dot{x}(\lambda))\,, \tag{32}$$

which can be straighforwardly varied without knowledge of the generically non-linear kinematical machinery running in the background.

(3) For massive point particles, the same philosophy applies, but entirely different mathematics are at work. Instead of projective algebraic geometry and projective Gauss maps, as for the massless particle, it is now convex analysis and the Legendre map and its inverse that play the crucial role. Also in this case, an obvious first order action

$$S_{\text{massive}}[x,k,\mu] = \int d\lambda\, \left[k_a(\lambda)\dot{x}^a(\lambda) - \mu(\lambda)(\ln P_{x(\lambda)})\left(\frac{k(\lambda)}{m}\right)\right] \tag{33}$$

leads to the problem of inverting a non-linear velocity-momentum relation, which is now achieved by virtue of the Legendre map $k = m\ell_x(\dot{x}/(\lambda \deg P^{\#}))$, which can be used to arrive at the equivalent second order action

$$S_{\text{massive}}[x] = \int d\lambda\, P(\ell_{x(\lambda)}(\dot{x}(\lambda)))^{-1/(\deg P^{\#})}\,. \tag{34}$$

It is instructive to note how the kinematics of standard general relativity follow from the above general theory, starting from even the simplest possible field dynamics one wishes to have available on a Lorentzian metric manifold of signature $(+---)$,

$$S_{\text{KG matter}}[\phi] = \int d^4x \sqrt{-\det g(x)}\left(g^{ab}(x)\partial_a\phi(x)\partial_b\phi(x) - m^2\phi^2(x)\right) \tag{35}$$

whose equations of motion yield a second rank principal tensor field with components $P^{ab} = g^{ab}$, so that hyperbolicity of $P$ is equivalent to the supposed Lorentzian signature of the metric. The hyperbolicity cones of $P_x$ at each point $x$ are two disjoint open convex cones of covectors $k$ for which $g^{ab}_x > 0$ and a time-orientation identifies one of them as the observer cone $C_x$. The positive sign convention $P_x(C_x) > 0$ is immediately satisfied because of the mainly minus signature chosen for the metric. The dual polynomial is $P^{\#}_{ab} = g_{ab}$ and the projective Gauss maps $[D^a P(k)] = [g^{ab}k_b]$ and $[D^P_a \#(X)] = [g_{ab}X^b]$ simply raise and lower the index and are linear in this simple case. The positive energy cone is the closure $\overline{C_x}$ of the observer cone, and thus indeed captures all massive momenta contained in $C_x$ and all massless momenta on the boundary $\partial C_x$ that is not part of the open cone $C_x$ and none of the momenta in $-\overline{C_x}$ are captured. The Legendre map and its inverse evaluate to $(\ell_x(X))_a = -g_{ab}X^b/g_{mn}X^mX^n$ and $(\ell^{-1}_x(k))^a = -g^{ab}k_b/g^{mn}k_mk_n$, yielding precisely the velocity-momentum relation for massive particles and reveal the generically required Legendre orthogonality of purely spatial observer frame vectors to the temporal frame vector as simply their Lorentzian metric orthogonality in this simple case. The general action for massless and massive particles reduces to the known actions in general relativity,

$$S_{\text{massless}}[x,\sigma] = \int d\lambda\, \sigma(\lambda) g_{ab}(x(\lambda))\dot{x}^a(\lambda)\dot{x}^b(\lambda)$$

and

$$S_{\text{massive}}[x] = \int d\lambda \sqrt{g_{ab} x^a(\lambda) \dot{x}^b(\lambda)}\,.$$

The general theory presented before generalizes these cornerstones of general relativity to generalized tensorial spacetimes, in a way that is informed by the specific matter dynamics one stipulates on the given spacetime. From this point of view — which is indeed the view taken by Einstein when he distilled the kinematical lessons conveyed by Maxwell dynamics into the spacetime structure and its interpretation — the kinematical interpretation of a spacetime geometry $(M, G)$ cannot be extracted from, or be assigned to, the geometry per se. Indeed, had we not chosen Klein-Gordon theory (or any other standard model field) as the matter field theory on our Lorentzian manifold, but instead, say, a Proca theory with quartic self-interaction, we would have obtained a vastly different kinematical interpretation of the very same Lorentzian manifold. In essence, the kinematics impressed on a geometry follows only from the triple $(M, G, S_{\text{matter}}[A, G])$. This is an insight that was so far essentially ignored in the overwhelming majority of attempts to construct gravity theories beyond general relativity.

## 4. Gravitational closure

Consider a foliation of the spacetime $M$ into initial data surfaces, described by a one-real-parameter family of smooth embedding maps $X_t : \Sigma \longrightarrow M$, where $\Sigma$ is a smooth three-dimensional manifold such that the image $X_t(\Sigma)$ is an initial data surface for the matter theory. We now define projection frames by choosing coordinate maps $y^\alpha$, for $\alpha = 1, \ldots, 3$ on $\Sigma$, whence we obtain a spacetime tangent basis

$$e_0(t,\sigma) = \ell^{-1}_{X_t(\sigma)}(n_t(\sigma)) \qquad \text{and} \qquad e_\alpha(t,\sigma) = X_{t*}(\left(\frac{\partial}{\partial y^\alpha}\right)_\sigma) \tag{36}$$

along each embedded hypersurface $X_t(\Sigma)$, where each $n(t,\sigma)$ is a spacetime covector field normal to the hypersurface and normalized with respect to $P^\#$,

$$n(t,\sigma)(e_\alpha(t,\sigma)) = 0 \qquad \text{and} \qquad P^\#_{X_t(\sigma)}(\ell^{-1}_{X_t(\sigma)}(n(t,\sigma)) = 1\,. \tag{37}$$

Employing this frame $e_a(t,\sigma)$ and the unique dual frame $\epsilon(t,\sigma)$, we obtain an obvious projection of the spacetime geometry $G$ to several one-real-parameter families of induced tensor fields on $\Sigma$. For a geometry dexcribed by a $(1,3)$-tensor field, for instance, we obtain eight one-real-parameter families of tensor fields on $\Sigma$ with valence $(0,0)$, $(0,1)$, $(0,1)$, $(0,2)$, $(1,0)$, $(1,1)$, $(1,1)$ and $(1,2)$, namely

$$\mathbf{g}_t^{\ {}^p{}_{qr}}(\sigma) = G_{X_t(\sigma)}(\epsilon^p(t,\sigma), e_q(t,\sigma), e_r(t,\sigma)) \tag{38}$$

for ${}^p{}_{qr} \in \{{}^0{}_{00}, {}^0{}_{0\gamma}, {}^0{}_{\beta 0}, {}^0{}_{\beta\gamma}, {}^\alpha{}_{00}, {}^\alpha{}_{0\gamma}, {}^\alpha{}_{\beta 0}, {}^\alpha{}_{\beta\gamma}\}$. Analogously for geometric tensor fields of different valence or even several such tensor fields of various valences. It proves useful to notationally collect the occurring index combinations in one caligraphy index $\mathcal{A}$, such that we can write $\mathbf{g}_t^{\mathcal{A}}(\sigma)$. Analogously, the dual $P^\#$ of the

principal tensor field yields $\deg P^{\#}+1$ one-parameter families of totally symmetric covariant tensor fields $\mathbf{p}_{t\,\alpha_1\ldots\alpha_n}(\sigma)$ for $n=0,1,\ldots,\deg P^{\#}$. An important point is that our projection frames are constructed such that the first two projections always take the values

$$\mathbf{p}_t(\sigma)=1 \qquad \text{and} \qquad \mathbf{p}_t{}^{\alpha}(\sigma)=0\,. \tag{39}$$

But since the dual of the principal polynomial is given ultralocally in terms of the spacetime geometry, one can write these two induced tensor fields, in particular, as functions of the induced geometric fields,

$$\mathbf{p}_t(\sigma)=p(\mathbf{g}_t^{\mathcal{A}}(\sigma)) \qquad \text{and} \qquad \mathbf{p}_t{}^{\alpha}(\sigma)=p^{\alpha}(\mathbf{g}_t^{\mathcal{A}}(\sigma))\,. \tag{40}$$

Thus the properties (39) impose a (generically non-linear) algebraic relation between the projected fields $\mathbf{g}_t^{\mathcal{A}}$. While these are automatically satisfied once the definition of the $\mathbf{g}_t^{\mathcal{A}}$ are employed, they become non-trivial if one turns to the canonical view of dynamics, which no longer considers the spacetime geometry $G$ as the fundamental variables of the theory and the projected fields $\mathbf{g}_t^{\mathcal{A}}$ as derived quantities, but precisely the other way around. This change in perspective is reflected by introducing tensor fields $g_t^{\mathcal{A}}$ on $\Sigma$ whose tensorial structure mimics that of the projections $\mathbf{g}_t^{\mathcal{A}}$, including the algebraic index symmetries the projections once inherited from the spacetime geometry, but are functionally no longer related to the spacetime geometry. But then the generically non-linear conditions (39) must be imposed explicitly as

$$p(g_t^{\mathcal{A}}(\sigma))=1 \qquad \text{and} \qquad p^{\alpha}(g_t^{\mathcal{A}}(\sigma))=0\,, \tag{41}$$

since they no longer follow automatically. But instead of dealing with such non-linear constraints, we introduce generalized configuration fields $\varphi^1,\ldots,\varphi^F$ on $\Sigma$ and parametrization maps $\widehat{g}^{\mathcal{A}}(\phi^1,\ldots,\phi^F)$ such that the tensor fields

$$g_t^{\mathcal{A}}=\widehat{g}^{\mathcal{A}}(\phi_t^1,\ldots,\phi_t^F) \tag{42}$$

generated from from one-real-parameter families of these configuratiomn fields satisfy the linear symmetry conditions and generically non-linear frame conditions (39) while the configuration variables diffeomorphically parametrize the remaining degrees of freedom. The latter is ensured by requiring also the existence of inverse maps $\widehat{\varphi}^A$ with for $A=1,\ldots,F$, such that

$$\widehat{g}^{\mathcal{A}}(\widehat{\varphi}^A(g))=g^{\mathcal{A}} \qquad \text{and} \qquad \widehat{\varphi}^A(\widehat{g}^{\mathcal{A}}(\varphi)=\varphi^A\,, \tag{43}$$

from which the important relation

$$\frac{\partial\widehat{\varphi}^A}{\partial g^{\mathcal{A}}}(\widehat{g}(\varphi))\frac{\partial\widehat{g}^{\mathcal{A}}}{\partial\varphi^B}(\varphi)=\delta^A_B \tag{44}$$

follows.

With the above preparations made, we can now calculate the two coefficient functions $F^A{}_{\mu}{}^{\gamma}(\varphi)$ and $M^{A\gamma}(\varphi)$ of the countable set of linear homogeneous partial

differential equations that must be solved in order to obtain the gravitational actions that are causally consistent with the given matter field dynamics. The first of these coefficients can be read off the right hand side of

$$(\mathcal{L}_{\vec{N}}\widehat{g})^{\mathcal{A}}(\varphi)\frac{\partial\widehat{\varphi}^{A}}{\partial g^{\mathcal{A}}}(\widehat{g}(\varphi)) =: N^{\mu}\partial_{\mu}\varphi^{A} + \partial_{\gamma}N^{\mu}F^{A}{}_{\mu}{}^{\gamma}(\varphi)\,, \tag{45}$$

where $\vec{N}$ is some vector field on $\Sigma$. The second one is calculated directly from

$$M^{A\gamma}(\varphi) = \frac{\partial\mathbf{g}^{\mathcal{A}}}{\partial\partial_{\gamma}X^{a}}e_{0}^{a}\frac{\partial\widehat{\varphi}^{A}}{\partial g^{\mathcal{A}}}(\widehat{g}(\varphi))\,, \tag{46}$$

which is easily calculated by expressing $\mathbf{g}^{\mathcal{A}}$ in terms of the spacetime geometry $G$ (for which $\partial G/\partial\partial_{\gamma}X^{a}$ vanishes) and using the relations

$$\frac{\partial e_{0}^{m}}{\partial\partial_{\gamma}X^{a}} = \frac{1}{1-\deg P^{\#}}e_{\sigma}^{m}p^{\sigma\gamma}(\widehat{g}(\varphi))\,, \tag{47}$$

$$\frac{\partial e_{\mu}^{m}}{\partial\partial_{\gamma}X^{a}} = \delta_{a}^{m}\delta_{\mu}^{\gamma}\,, \tag{48}$$

$$\frac{\partial\epsilon_{m}^{0}}{\partial\partial_{\gamma}X^{a}} = -\epsilon_{a}^{0}\epsilon_{m}^{\gamma}\,, \tag{49}$$

$$\frac{\partial\epsilon_{m}^{\mu}}{\partial\partial_{\gamma}X^{a}} = -\epsilon_{a}^{\mu}\epsilon_{m}^{\gamma} + \frac{1}{\deg P^{\#}-1}\epsilon_{m}^{0}\epsilon_{a}^{0}p^{\mu\gamma}(\widehat{g}(\varphi))\,. \tag{50}$$

The key result[1] of constructive gravity is that the gravitational dynamics for the spacetime geometry are given by the action

$$S_{\text{geometry}}[G(\varphi_{t},N_{t},\vec{N}_{t})] = \int d^{4}x\,N\mathscr{L}\left(\varphi_{t},K[\varphi_{t},N_{t},\vec{N}_{t}]\right) \tag{51}$$

where

$$K^{A}[\varphi,N,\vec{N}] := \frac{1}{N}\Big(\dot{\varphi} - (\partial_{\gamma}N)M^{A\gamma}[\varphi] - N^{\mu}\partial_{\mu}\varphi^{A} + (\partial_{\gamma}N^{\mu})F^{A}{}_{\mu}{}^{\gamma}(\varphi)\Big) \tag{52}$$

and the scalar density $\mathscr{L}$ of weight one is a functional of the $\varphi$ and a function of the $K$ that is determined by the gravitational closure equations, which in functional differential form are just the two following equations

$$\begin{aligned}
0 = {} & -K^{B}(y)\frac{\delta\mathscr{L}(x)}{\delta\varphi^{B}(y)} + (\partial_{\gamma}\delta_{x})\,(y)K^{B}(y)M^{A\gamma}{}_{:B}(x)\frac{\partial\mathscr{L}}{\partial K^{A}}(x) + \partial_{\mu}\left(\frac{\delta\mathscr{L}(x)}{\delta\varphi^{B}(\cdot)}M^{B\mu}\right)(y) \\
& + \partial_{\mu}\frac{\partial\mathscr{L}}{\partial K^{A}}(x)\Big[(\deg P^{\#}-1)^{-1}p^{\rho\mu}F^{A}{}_{\rho}{}^{\nu} - M^{B[\mu|}M^{A|\nu]}{}_{:B}\Big](x)\,(\partial_{\nu}\delta_{x})\,(y) \\
& - \frac{\partial\mathscr{L}}{\partial K^{A}}(x)\Big[(\deg P^{\#}-1)^{-1}p^{\rho\nu}\left(\partial_{\rho}\varphi^{A} + F^{A}{}_{\rho}{}^{\gamma}{}_{,\gamma}\right) \\
& + \partial_{\mu}\left(M^{B[\mu|}M^{A|\nu]}{}_{:B}\right)\Big](x)\,(\partial_{\nu}\delta_{x})\,(y) \\
& - (x \longleftrightarrow y).
\end{aligned} \tag{53}$$

and

$$
\begin{aligned}
0 = {} & \frac{\partial \mathscr{L}}{\partial K^B}(y)\, K^A(y)\, \left(\delta^B_A \delta^\gamma_\mu + F^B{}_\mu{}^\gamma{}_{:A}\right)(y)(\partial_\gamma \delta_y)(x) \\
& - K^A(y)\partial_\gamma \frac{\partial \mathscr{L}}{\partial K^B}(y) F^B{}_\mu{}^\gamma{}_{:A}(y)\delta_y(x) \\
& - \left(K^A \frac{\partial \mathscr{L}}{\partial K^A} - \mathscr{L}\right)(y)(\partial_\mu \delta_y)(x) + \partial_\mu \left(K^A \frac{\partial \mathcal{L}}{\partial K^A} - \mathscr{L}\right)(y)\delta_y(x) \\
& + \left(\partial_\mu \varphi^A + F^A{}_\mu{}^\gamma{}_{,\gamma}\right)(x) \frac{\delta \mathscr{L}(y)}{\delta \varphi^A(x)} + F^A{}_\mu{}^\gamma(x)\partial_\gamma \left(\frac{\delta \mathscr{L}(y)}{\delta \varphi^A(\cdot)}\right)(x)\,, \qquad (54)
\end{aligned}
$$

where the shorthand $Q_{:A}{}^{\alpha_1 \ldots \alpha_N} := \partial Q / \partial \partial_{\alpha_1 \ldots \alpha_N} \varphi^A$ has been used in both. Solving these equations for $\mathscr{L}$ then completely determines the gravitational action (51) that provides dynamics for the geometry employed in the matter action which is causally consistent with the initially stipulated matter field dynamics.

## 5. Solution techniques for gravitational closure equations

***General gravitational closure.*** While it is straightforward to set up the gravitational closure equations for any matter field action on any tensorial background geometry – as long as the principal polynomial of the resulting matter field equations of motion satisfy the three physicality conditions, which may require restriction of the geometry – it appears generically prohibitively hard to solve this countable set of linear homogeneous partial differential equations.

A notable exception is provided by standard model matter on a metric background. The restriction on the geometry, which makes the standard model fields satisfy the physicality conditions, boils down to the metric having Lorentzian signature. In this case, the closure equations are not only set up as swiftly as in any other case, but they can also be solved without further assumptions.

Maybe the difficulty to solve the gravitational closure equations for other matter models of physical interest, such as general linear – and thus birefringent – electrodynamics, is in part due to an unfortunate choice of field parametrization (42) of the field degrees of freedom. A well-considered choice of parametrization might render a general solution less difficult.

The most promising line of attack, however, is to better understand the structure of the closure equations themselves. A theoretically invaluable step would be to take them to involutive form, maybe in general and maybe case by case. At any rate, this would allow, for instance, a direct calculation of the dimension of their linear solution space. This dimension would then of course be equal to the number of physical constants that remain undetermined by constructive gravity and must hence be measured in experiments. While this number is 2 in the case of standard model matter – corresponding to an undetermined gravitational constant and an undetermined cosmological constant — we know that it is at least 11 for the gravitational theory that underlies general linear electrodynamics. Clearly, any gravitational theory with infinitely many undetermined constants is non-predictive

in general, although the number of relevant constants might become finite under special circumstances, such as symmetry assumptions.

From a practical point of view, however, a general solution to the gravitational closure equations for some given matter dynamics beyond the standard model is not required. For even if it was available, it would give rise to field equations that are at least as difficult to solve as Einstein's equations. At that stage, these elusive general gravitational field equations would have to be solved either by symmetry assumptions or perturbation theory. The obvious idea is to implement any desired symmetry assumption or perturbative technique already at the level of the gravitational closure equations. This is not entirely straightforward, and the following two subsections briefly outline the problem an the solution.

***Symmetric gravitational closure.*** Since the gravitational closure equations yield a gravitational action, any implementation of a Killing symmetry

$$\mathcal{L}_K G = 0$$

for the pertinent tensorial spacetime geometry at the level of the closure equations, if properly implemented, will be passed down to the action.

But this means that variation of the action, with the aim to obtain the gravitational field equations, must now be performed with respect to the symmetric field configurations that appear in the already symmetry-reduced action. Thinking of variation in field theory as a competition between a candidate field and slightly deformed competitors, one notes that the competition is severely limited by only considering competitors that also already satisfy the imposed symmetry condition. It is thus clear that variation of a symmetry-reduced action produces weaker symmetrized field equations than variation of the full action and subsequent symmetrization would have. This is the known issue of symmetric criticality[13] and further useful necessary and sufficient conditions for when a symmetry reduction at the level of the action yields the correct field equations have been identified[14,15].

It is clear from the above discussion that the implementation of spacetime symmetries at the level of the gravitational closure equations is possible under precisely the same conditions that apply to their implementation at the level of an action. Symmetric gravitational closure is discussed and illustrated in[16], where it is shown that one obtains the Friedmann equations as the cosmologically symmetry-reduced gravitational closure of standard model matter dynamics, without ever having to know Einstein's equations.

For matter models beyond the standard model, the simplification of the corresponding gravitational closure equations are tremendous[17] and thus put a solution in reach, see the contribution by Düll to these Proceedings.

***Perturbative gravitational closure.*** Perturbative solutions of gravitational closure equations my be performed under precisely the same assumptions that render them meaningful at the level of the equations of motion. Since the action needs to be known, roughly speaking, to second order if one wishes to derive field equations

that are valid to first order perturbations, care needs to be taken in the determination of terms that can be dropped or not. In practice, this is based on a subtle interplay of the order to which the coefficient functions of the gravitational closure equations must be expanded, the number of derivatives that act afterwards and the resulting overall order of a particular closure equation.

The perturbative treatment yields the 11-parameter weak gravitational field equations that underlie general linear electrodynamics[18] as well as the interesting bimetric gravity theory as it underlies the only superficially trivial matter model of two Klein-Gordon fields that couple to two different Lorentzian metrics, see the contribution by Wierzba to these Proceedings.

## 6. Applications

***Birefringence in the weak gravitational field of a point mass.*** The perturbative gravitational closure of general linear electrodynamics, which served as our opening example for a problem that was previously not solvable, yields a spacetime geometry

$$G^{abcd} = 2\eta^{a[c}\eta^{d]b} - \epsilon^{abcd} + H^{abcd}$$

that describes the gravitational field not too close to a point mass $m$ with the perturbative deviation $H^{abcd}$ from Minkowski space $H^{abcd}$ given in[18]. No birefringence is seen, to first order perturbation theory, where and only where[19]

$$\frac{1}{2}H^{ab}H^{cd} = \eta_{mr}\eta_{ns}H^{ambn}H^{crds} \tag{55}$$

holds to second order. Thus allowing for birefringence in principle, the weak gravitational field generated by a point mass indeed generates birefringence, whose strength depends on the mass and four more independent constants to be determined by only four experiments in that gravitational field; see also the contribution by Stritzelberger to these Proceedings.

***Gravitational effects in birefringent quantum electrodynamics.*** The quantization of birefringent electrodynamics is renormalizable in a gauge-invariant way to all perturbative orders on a flat background[20]. This result can be used to reliably calculate quantum field theoretic processes in locally essentially flat regions of a globally non-flat area metric geometry that arises as the solution of the underyling gravity theory that one obtains by gravitational closure of classical birefringent electrodynamics. This allows to search for signatures of birefringence in localized quantum electrodynamic processes, which are now dependent of the spacetime region where they occur. Of particular interest is a modification of the observationally important 21.1 cm line of hydrogen, which in the presence of birefringence is seen to depend qualitatively and quantitatively – in a way precisely predicted by constructive gravity – on the location of the hydrogen; see the contribution by Tanzi to these Proceedings.

***Gravitational radiation.*** An notable result[21] concerning the gravitational closure of birefringent electrodynamics is that the gravitational radiation emitted by two circularly orbiting masses contains only one massless trace-free tensor mode, as in general relativity, while additional scalar, vector and tensor modes are all massive. The production of these massive waves is shown to be significantly suppressed, since it requires the orbital frequency to surpass a certain threshold. Slowly orbiting binaries therefore only radiate waves of the type predicted by general relativity.

***Etherington distance relation.*** On a Lorentzian spacetime, the Etherington distance duality relation[22] connects the luminosity distance, angular diameter distance and redshift of an astrophysical light source independent of the gravitational dynamics. This is not the case for the refined spacetime geometry that underlies birefringent linear electrodynamics. Direct calculation, from the field equations obtained by gravitational closure, yields a modification of Etherington's relation that depends on the gravitational dynamics and indeed the particular spacetime solution[23]. This opens up the possibility of deriving new gravitational lensing effects or indeed a pathway for the explanation of known anomalies that is directly connected to carefully studied extensions of the standard model[24]. The reader is referred to the contribution by Werner to these Proceedings.

***Parametrodynamics.*** Parameters in matter field dynamics are usually considered as constants whose values must be determined by experiment. Gravitational closure, however, can be used to predict the values of non-scalar parameters that appear in any local field theory that is amenable to the closure equations[5]. To this end, one first promotes the constant parameters to fields, analogous to the promotion of the flat spacetime Minkowski metric in Maxwell theory to the metric tensor field in the general theory of relativity. Gravitational closure of the this modified matter action then yields a multi-parameter family of actions for these parameters (and, if one so chooses, also for the underlying geometry) as the unique dynamics that enjoys a consistent co-evolution with the matter fields of the initially stipulated matter theory. See the contribution by Wierzba to these Proceedings.

## 7. Conclusions

Constructive gravity is a method to determine, by calculation rather than stipulation, a family of gravitational actions that are compatible with a large class of matter field theories. Remarkably, it is any concrete representative of these matter field theories themselves that provides the relevant information for the calculation of the gravitational theory, essentially based on the requirement that the latter have a diffeomorphism gauge group and possess a canonical evolution that shares its initial data surfaces with those of the chosen matter theory. The availability of such a procedure allows to ask and answer questions that could not be posed before and reveal gravitational dynamics as a mere consistency requirement once the matter

contents of the universe is specified. While for a universe filled with standard model matter, there is no new physics predicted by the gravitational closure mechanism, this significantly changes once matter beyond the standard model is considered. This is where constructive gravity will likely find its key application.

**Acknowledgments**

The author thanks Marcus Werner for organizing a wonderful session on constructive gravity at the 15th Marcel Grossmann Meeting in Rome and for the invitation to contribute this rapporteur article to the Proceedings. He also thanks all contributors to this session for their presentations and discussions on their take of the subject and particularly Maximilian Düll, Florian Wolz, Nadine Stritzelberger, Alexander Wierzba, Jonas Schneider, Moritz Möller, Nils Fischer, Nils Alex and Hans-Martin Rieser who have contributed to the results mentioned in this article or based further work on them.

**References**

1. M. Düll, F. P. Schuller, N. Stritzelberger and F. Wolz, *Phys. Rev.* **D97** (2018) no.8, 084036 doi:10.1103/PhysRevD.97.084036
2. D. Raetzel, S. Rivera and F. P. Schuller, *Phys. Rev.* **D83** (2011) 044047 doi:10.1103/PhysRevD.83.044047
3. K. Kuchar, *J. Math. Phys.* **15** (1974) 708. doi:10.1063/1.1666715
4. S. A. Hojman, K. Kuchar and C. Teitelboim, *Annals Phys.* **96** (1976) 88. doi:10.1016/0003-4916(76)90112-3
5. A. Wierzba, "Parametrodynamics," *Master thesis*, Ludwig-Maximilians-Universität München, 2018
6. S. Rivera and F. P. Schuller, *Phys. Rev.* **D83** (2011) 064036 doi:10.1103/PhysRevD.83.064036 [arXiv:1101.0491 [hep-th]].
7. F. W. Hehl and Y. N. Obukhov, "Foundations of Classical Electrodynamics – Charge, Flux, and Metric," Birkhäuser (2003).
8. G. Velo and D. Zwanziger, *Phys. Rev.* **188** (1969) 2218. doi:10.1103/PhysRev.188.2218
9. F. P. Schuller, C. Witte and M. N. R. Wohlfarth, *Annals Phys.* **325** (2010) 1853 doi:10.1016/j.aop.2010.04.008
10. F. W. Hehl, Y. N. Obukhov and G. F. Rubilar, *Int. J. Mod. Phys.* **A17** (2002) 2695 doi:10.1142/S0217751X0201162X
11. Y. Itin, *J. Phys.* **A42**, 475402 (2009) doi:10.1088/1751-8113/42/47/475402
12. M. Kuranishi, "On É. Cartan's prolongation theorem of exterior differential systems," *Amer. J. Math.*, vol. 79, 1957, p. 147.
13. R. S. Palais, *Commun. Math. Phys.* **69**, no. 1, 19 (1979). doi:10.1007/BF01941322
14. M. E. Fels and C. G. Torre, *Class. Quant. Grav.* **19**, 641 (2002) doi:10.1088/0264-9381/19/4/303

15. C. G. Torre, *AIP Conf. Proc.* **1360**, no. 1, 63 (2011) doi:10.1063/1.3599128
16. M. Düll, N. L. Fischer, B. M. Schaefer and F. P. Schuller, arXiv:2003.07109 [gr-qc].
17. M. Düll, *PhD thesis*, Ruprecht-Karls-Universität Heidelberg, 2020
18. J. Schneider, F. P. Schuller, N. Stritzelberger and F. Wolz, arXiv:1708.03870 [hep-th].
19. V. A. Kostelecky and M. Mewes, *Phys. Rev. Lett.* **87**, 251304 (2001) doi:10.1103/PhysRevLett.87.251304
20. S. Grosse-Holz, F. P. Schuller and R. Tanzi, arXiv:1703.07183 [hep-ph].
21. M. Möller, "The Generation of Gravitational Waves by Orbiting Charges," *Master thesis*, Friedrich-Alexander-Universität Nürnberg-Erlangen, 2018
22. I. M. H. Etherington, *Phil. Mag.* 15, 761 (1933).
23. F. P. Schuller and M. C. Werner, *Universe* **3**, no. 3, 52 (2017) doi:10.3390/universe3030052
24. D. Colladay and V. A. Kostelecky, *Phys. Rev.* **D58**, 116002 (1998) doi:10.1103/PhysRevD.58.116002

# Is there weak birefringence of light in vacuo?

Nadine Stritzelberger

*Department of Applied Mathematics, University of Waterloo, 200 University Ave W, Waterloo, ON, Canada N2L 3G1*

*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5*

*Institute for Quantum Computing, University of Waterloo, 200 University Ave W, Waterloo, ON, Canada N2L 3G1*
*E-mail: nadine.stritzelberger@cantab.net*

There are numerous reasons to study modifications of general relativity and the Standard Model of particle physics, ranging from modelling inflation to exploring galaxy rotation curves and the nature of dark matter. Here we study the most general linear theory of electromagnetism, which admits vacuum birefringence, and derive weak gravitational field equations for the underlying area metric spacetime geometry. We discuss the weak gravitational field sourced by a point mass in an area metric spacetime and find first order corrections to the linearized Schwarzschild metric of general relativity.

*Keywords*: General linear electrodynamics; vacuum birefringence; area metric; constructive gravity; modified gravity.

## 1. Introduction

Could there be weak birefringence of light in vacuo? How can we predict whether and where such splitting of light rays in vacuo occurs? Standard Maxwell electrodynamics, for which the electromagnetic field couples to a metric spacetime geometry, *a priori* excludes the possibility of vacuum birefringence. The spacetime geometry must be refined, from a metric to a tensor field of the fourth rank, in order to obtain a linear theory of electrodynamics allowing for vacuum birefringence to occur. The such obtained matter theory is known as the theory of *general linear electrodynamics*[1–3] and the fourth rank spacetime tensor field as an *area metric*. Because the matter field equations of general linear electrodynamics depend on the area metric background, finding any concrete solution for the electromagnetic field requires to simultaneously solve some yet-to-be-determined field equations for the area metric, which thus take the role played by Einstein's field equations for a metric background. In particular, also the splitting of light rays depends crucially on such gravitational field equations underlying general linear electrodynamics. We can thus rephrase the question of whether and where birefringence of light in vacuo occurs as follows: What are the gravitational field equations underlying general linear electrodynamics, and what are the solutions to these field equations?
This question can be answered by means of the recently discovered procedure of gravitational closure[4,5], which employs the causal structure of given matter field equations in order to provide causally compatible canonical dynamics[6,7] for the underlying geometry. For instance, starting from Maxwell electrodynamics, one

obtains the familiar Einstein equations in this way. Exactly along the same lines, when starting instead from general linear electrodynamics, one obtains the gravitational field equations for an area metric geometry. Based on Ref. [8], this paper reviews the derivation of the *weak* gravitational field equations underlying general linear electrodynamics, followed by a discussion of a concrete solution to these equations, namely the weak gravitational field sourced by a point mass in an area metric spacetime, revealing first order corrections to the linearized Schwarzschild metric.

## 2. General linear electrodynamics

We start our discussion with a brief review of the theory of general linear electrodynamics[1–3,9–11]. The action of this matter theory is

$$S_{matter}[A, G)] = -\frac{1}{8}\int_M d^4x\, \omega_G F_{ab} F_{cd} G^{abcd}\,, \tag{1}$$

where $M$ is an orientable four-dimensional spacetime manifold, $F = \mathrm{d}A$ denotes the electromagnetic field strength tensor, the fourth-rank contravariant tensor field $G$ denotes the area metric spacetime geometry, and the scalar density $\omega_G = 4!(\epsilon_{abcd}G^{abcd})^{-1}$ is constructed by virtue of the canonical top form density $\epsilon$. In a spacetime equipped with a tensorial spacetime geometry, general linear electrodynamics is the most general theory of electrodynamics for which the classical linear superposition principle still holds. By virtue of the appearance of the fourth-rank tensor field in the above matter action, the algebraic symmetries

$$G^{abcd} = G^{cdab} \quad \text{and} \quad G^{abcd} = -G^{bacd} \tag{2}$$

may be assumed without loss of generality. Consequently, an area metric spacetime geometry features 21 independent degrees of freedom at each spacetime point, as opposed to only 10 independent degrees of freedom for a metric spacetime geometry. It is ultimately due to these additional degrees of freedom that general linear electrodynamics allows not only to describe all of Maxwell electrodynamics, but also various new phenomena such as vacuum birefringence.

The causal structure of the equations of motion of general linear electrodynamics is prescribed by the principal tensor field $P_G$, which was shown in Ref. [2, 3] to be a totally symmetric tensor field density of the fourth rank, determined entirely in terms of the spacetime geometry,

$$P_G^{abcd} = -\frac{1}{4!}\omega_G^2\, \epsilon_{mnpq}\epsilon_{rstu} G^{mnr(a} G^{b|ps|c} G^{d)qtu}\,. \tag{3}$$

The principal tensor field is the reason why generic area metric spacetime geometries feature vacuum birefringence: The null cone of general linear electrodynamics, that is the set of covectors annihilating the principal tensor field, is a quartic surface. However, if the splitting of light rays in vacuo was a large effect, it would most likely have been observed by now. Therefore, instead of studying generic area

metric spacetimes, this paper focuses on spacetime geometries which are small area metric perturbations $H$ around metric Minkowski spacetime $\eta$,

$$G^{abcd} = \eta^{ac}\eta^{bd} - \eta^{ad}\eta^{bc} - \epsilon^{abcd} + H^{abcd} \,. \tag{4}$$

Once we know the linearized gravitational field equations underlying general linear electrodynamics, we can derive the area metric perturbations $H$ for given small matter distributions. In the following two sections, we therefore discuss the derivation of these linearized gravitational field equations by means of gravitational closure.

## 3. Canonical geometry

The procedure of gravitational closure enables the derivation of causally compatible gravitational field equations for the spacetime geometry underlying any predictive and quantizable matter field theory[4–7]. The first step in this gravitational closure procedure is to determine the canonical geometry for which one ultimately wishes to obtain canonical dynamical field equations. The canonical geometry of an area metric spacetime geometry can be determined as follows. Let $\Sigma$ be a three-dimensional manifold and let $X_t : \Sigma \hookrightarrow M$ be a one-real-parameter family of embedding maps, which specifies a foliation of the spacetime manifold into initial data surfaces. For each point $\sigma \in \Sigma$ and each embedding parameter $t \in \mathbb{R}$, one can construct an orthonormal frame $\{\epsilon^0(t,\sigma), \epsilon^\alpha(t,\sigma)\}$, where $\alpha \in \{1,2,3\}$, which allows to project the area metric spacetime geometry onto the initial data surfaces,

$$(g_1)^{\alpha\beta}[X_t] := -G(\epsilon^0, \epsilon^\alpha, \epsilon^0, \epsilon^\beta) \,, \tag{5}$$

$$(g_2)_{\alpha\beta}[X_t] := \frac{1}{4}\frac{\epsilon_{\alpha\mu\nu}}{\sqrt{\det(g_1)}}\frac{\epsilon_{\beta\sigma\tau}}{\sqrt{\det(g_1)}} G(\epsilon^\mu, \epsilon^\nu, \epsilon^\sigma, \epsilon^\tau) \,, \tag{6}$$

$$(g_3)^\alpha{}_\beta[X_t] := \frac{1}{2}\frac{\epsilon_{\beta\mu\nu}}{\det(g_1)} G(\epsilon^0, \epsilon^\alpha, \epsilon^\mu, \epsilon^\nu) - \delta^\alpha_\beta \tag{7}$$

resulting in three one-real-parameter families of induced tenor fields. By construction, the induced tensor fields $g_1$ and $g_2$ are symmetric,

$$(g_1)^{[\alpha\beta]} = 0 \quad \text{and} \quad (g_2)_{[\alpha\beta]} = 0 \,. \tag{8}$$

Moreover, the orthonormality conditions satisfied by the four frame fields imply

$$(g_1)^{\sigma[\alpha}(g_3)^{\beta]}{}_\sigma = 0 \quad \text{and} \quad (g_3)^\alpha{}_\alpha = 0 \,. \tag{9}$$

Therefore, only 17 of the components of the induced tensor fields are independent degrees of freedom. In order to obtain the canonical geometry from this induced geometry, we need to parametrize the three induced tensor fields $\{g_1, g_2, g_3\}$ in terms of unconstrained canonical configuration variables $\{\varphi^1, \ldots, \varphi^{17}\}$. Note that since the four frame conditions (9) are non-linear, any parametrization $\{g_1(\varphi), g_2(\varphi), g_3(\varphi)\}$ respecting the symmetry conditions (8) and frame conditions (9) is non-linear in the canonical configuration variables. One possible parametrization—particularly

suited for the perturbative approach pursued in this paper—is the following. Consider a constant intertwiner $\mathcal{I}^{\alpha\beta}{}_A$, distributing the unconstrained configuration variables $\{\varphi^1, ..., \varphi^{17}\}$ over three second rank tensor fields,

$$(\varphi_1)^{\alpha\beta} := \sum_{A=1}^{6} \mathcal{I}^{\alpha\beta}{}_A \varphi^A, \quad (\varphi_2)^{\alpha\beta} := \sum_{A=7}^{12} \mathcal{I}^{\alpha\beta}{}_A \varphi^A, \quad (\varphi_3)^{\alpha\beta} := \sum_{A=13}^{17} \mathcal{I}^{\alpha\beta}{}_A \varphi^A. \tag{10}$$

Now let $\gamma$ denote a flat three-dimensional Euclidean metric on $\Sigma$ and let the intertwiner $\mathcal{I}^{\alpha\beta}{}_A$ and its inverse $\mathcal{I}^A{}_{\alpha\beta}$ satisfy the completeness relations

$$\mathcal{I}^{\alpha\beta}{}_A \mathcal{I}^B{}_{\alpha\beta} = \delta^B_A \quad , \qquad \sum_{A=1}^{6} \mathcal{I}^{\alpha\beta}{}_A \mathcal{I}^A{}_{\mu\nu} = \delta^{(\alpha}_{\mu} \delta^{\beta)}_{\nu},$$

$$\sum_{A=7}^{12} \mathcal{I}^{\alpha\beta}{}_A \mathcal{I}^A{}_{\mu\nu} = \delta^{(\alpha}_{\mu} \delta^{\beta)}_{\nu} \quad , \qquad \sum_{A=13}^{17} \mathcal{I}^{\alpha\beta}{}_A \mathcal{I}^A{}_{\mu\nu} = \delta^{(\alpha}_{\mu} \delta^{\beta)}_{\nu} - \frac{1}{3} \gamma^{\alpha\beta} \gamma^{\mu\nu} \ . \tag{11}$$

By means of this particular choice of intertwiner, the tensor fields $\varphi_1$ and $\varphi_2$ are symmetric, while $\varphi_3$ is symmetric and trace-free. We find that the parametrization

$$(g_1(\varphi))^{\alpha\beta} := \gamma^{\alpha\beta} + (\varphi_1)^{\alpha\beta}, \tag{12}$$

$$(g_2(\varphi))^{\alpha\beta} := \gamma^{\alpha\beta} + (\varphi_2)^{\alpha\beta}, \tag{13}$$

$$(g_3(\varphi))^{\alpha\beta} := (\varphi_3)^{\alpha\beta} + \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{2^n} \underbrace{\Big\{\varphi_1, \Big\{\varphi_1, \Big\{ \ldots, \Big\{}_{n-1 \text{ anti-commutator brackets}} \underbrace{\varphi_1, \big[\varphi_1, \varphi_3\big]}_{\text{commutator bracket}} \Big\} \ldots \Big\}\Big\}\Big\}^{\alpha\beta} \tag{14}$$

meets all symmetry and frame conditions, leaving us with an explicit expression of the canonical geometry in terms of the 17 independent geometric degrees of freedom.

## 4. Linearized gravitational dynamics

The second step in the gravitational closure procedure is to set up and solve a set of countably many linear homogeneous partial differential equations—the *gravitational closure equations*—whose coefficients are determined by the principal tensor field expressed in terms of the canonical configuration variables. A derivation of these coefficients can be found in section IV.C of [4]. Starting from general linear electrodynamics, the principal tensor field (3), together with the parametrization (12-14), allows to determine these coefficients in terms of the canonical configuration variables $\{\varphi^1, \ldots, \varphi^{17}\}$. The gravitational closure equations then in turn determine the coefficient functionals $C_{A_1 \ldots A_N}[\varphi]$ of a power series ansatz for the gravitational Lagrangian density

$$\mathcal{L}_{grav}[\varphi, k] = \sum_{N=0}^{\infty} C_{A_1 \ldots A_N}[\varphi]\, k^{A_1} \ldots k^{A_N}, \tag{15}$$

where $k$ are the velocities associated with the canonical configuration variables $\varphi$. Deriving gravitational field dynamics hence amounts to solving the gravitational

closure equations—which turns out to be a difficult task for a non-metric geometry. However, we here only wish to study small area metric perturbations, which require just *linearized* gravitational field equations. In order to obtain the gravitational field equations underlying general linear electrodynamics to linear order in $\varphi$ and $k$, we only need to derive the Lagrangian density to second order

$$\mathcal{L}_{grav}[\varphi, k] = C[\varphi] + C_A[\varphi]k^A + C_{AB}[\varphi]k^A k^B + \mathcal{O}(3)\,. \tag{16}$$

That is, we only need to solve the gravitational closure equations for $C[\varphi]$ to second order, $C_A[\varphi]$ to first order and $C_{AB}[\varphi]$ to zeroth order in $\varphi$. We now also see why the parametrization (12-14) is particularly suited for this perturbative approach: The configuration variables $\varphi$ can be directly employed as the perturbative degrees of freedom of an area metric perturbed around flat Minkowski spacetime $\eta = \mathrm{diag}(1, -1, -1, -1)$. We then perturbatively expand both the coefficients of the gravitational closure equations and the coefficient functionals $C_{A_1 \ldots A_N}[\varphi]$ in orders of $\varphi$, and evaluate the equations order by order. The such obtained gravitational Lagrangian, to second order in $\varphi$ and $k$, and the therefrom obtained linearized gravitational field equations can respectively be found in section III.D and IV.B of [8]. These field equations can now be solved for specific matter distributions of interest.

## 5. Solution around a point mass

For instance, let us consider the Hamiltonian $H_{matter}$ of a point mass $M$ at rest,

$$\frac{\delta H_{matter}}{\delta A(x)} = -M\delta^{(3)}(x)\,, \tag{17}$$

where $A$ denotes the perturbation of the lapse vector field. Solving the linearized gravitational equations of motion for this particular small matter distribution yields the following area metric perturbations,

$$H^{0\beta 0\delta} = \gamma^{\beta\delta}(2A - \tfrac{1}{2}U - \tfrac{1}{2}V) + \mathcal{O}(2)\,, \tag{18}$$

$$H^{0\beta\gamma\delta} = \epsilon^{\beta\gamma\delta}(-A + \tfrac{3}{4}U + \tfrac{3}{4}V) + \mathcal{O}(2)\,, \tag{19}$$

$$H^{\alpha\beta\gamma\delta} = 2\gamma^{\alpha[\gamma}\gamma^{\delta]\beta}(U + 2V) + \mathcal{O}(2)\,, \tag{20}$$

where the scalar perturbations $A$, $U$ and $V$ can be expressed in terms of the Euclidean radial distance $r$ and undetermined integration constants $\alpha$, $\beta$, $\gamma$ and $\mu$:

$$U = \frac{M}{4\pi r}\left(\alpha + \beta e^{-\mu r}\right), \quad V = \frac{M}{4\pi r}\gamma e^{-\mu r}\,, \quad A = \frac{M}{16\pi r}\left(\alpha - (\beta + 3\gamma)\, e^{-\mu r}\right). \tag{21}$$

Just like the gravitational and cosmological constants in the Einstein field equations of general relativity, these constants need to be determined by experiment. For the principal tensor field of an area metric spacetime sourced by a point mass we obtain,

$$P_G^{0000} = 1 + 2\Phi + \mathcal{O}(2)\,, \quad P_G^{\alpha 000} = \mathcal{O}(2)\,, \quad P_G^{\alpha\beta 00} = -\tfrac{1}{6}\gamma^{\alpha\beta} + \mathcal{O}(2)\,,$$
$$P_G^{\alpha\beta\gamma 0} = \mathcal{O}(2)\,, \quad P_G^{\alpha\beta\gamma\delta} = (1 - 2\Phi)\gamma^{(\alpha\beta}\gamma^{\gamma\delta)} + \mathcal{O}(2)\,, \tag{22}$$

where we defined $\Phi := \frac{M}{4\pi r}\left[-\frac{\alpha}{2} + \left(\frac{\beta}{2} + \frac{3\gamma}{4}\right) e^{-\mu r}\right]$. In order to decide whether light rays split, one would need to determine the solution of the gravitational field equations for the area metric to second order [12]. However, already to linear order, we find corrections to the usual linearized Schwarzschild metric in the form of a Yukawa potential. Allowing for vacuum birefringence, by assuming the electromagnetic field to be described by general linear electrodynamics, hence leaves its imprints on the spacetime geometry around a point mass. With the solution (22) one could now for instance study weak gravitational lensing in area metric spacetimes.

## 6. Conclusions

The weak gravitational field equations for the area metric spacetime geometry underlying general linear electrodynamics can be obtained by means of the gravitational closure mechanism. We find that modifying the spacetime geometry from a metric to an area metric, thereby allowing for vacuum birefringence, affects the canonical dynamical field equations and their solutions: For instance, for the weak field around a point mass in an area metric spacetime, we find first order Yukawa corrections to the Schwarzschild metric. The linearized gravitational field equations now also enable further studies, e.g. the propagation of gravitational waves [8] or modifications of the standard Etherington relation [13] in area metric spacetimes.

## Acknowledgments

This paper is based on the research article [8] and a talk presented in the Constructive Gravity session AT5 at the 15th Marcel Grossmann meeting. The author would like to thank F. P. Schuller for most valuable discussions and advice. The author gratefully acknowledges support from the Studienstiftung des deutschen Volkes and the Ontario Trillium Scholarship (OTS) program.

## References

1. F. W. Hehl and Y. N. Obukhov, *Foundations of classical electrodynamics* (Birkhäuser, 2003)
2. G.F. Rubilar, *Annals Phys.* **11**, 717 (2002)
3. G.F. Rubilar, Y.N. Obukhov, F. W. Hehl, *Int. J. Mod. Phys.* D **11**, 1227 (2002)
4. M. Düll, F. P. Schuller, N. Stritzelberger and F. Wolz, *Phys. Rev.* D **97**, 084036 (2018)
5. K. Giesel, F. P. Schuller, C. Witte and M. N. R. Wohlfarth, *Phys. Rev.* D **85**, 104042 (2012)
6. K. Kuchar, *J. Math. Phys.* **15**, 708 (1974)
7. S. A. Hojman, K. Kuchar and C. Teitelboim, *Annals Phys.* **96**, 88 (1976)
8. J. Schneider, F. P. Schuller, N. Stritzelberger and F. Wolz, `arXiv:hep-th/1708.03870` (2017)

9. V. A. Kostelecky and M. Mewes, *Phys. Rev.* D **80**, 015020 (2009), 0905.0031
10. F. P. Schuller, C. Witte and M. N. R. Wohlfarth, *Annals Phys.* **325**, 9 (2010)
11. R. Rätzel, S. Rivera and F. P. Schuller, *Phys. Rev.* D **83**, 044047 (2011)
12. S. Grosse-Holz, F. P. Schuller and R. Tanzi, `arXiv:hep-ph/1703.07183v2` (2017)
13. F. P. Schuller and M. C. Werner, *Universe* **3**, 52 (2017)

# Quantum signatures of area-metric deviations from a metric

Roberto Tanzi*

*University of Bremen, Center of Applied Space Technology and Microgravity (ZARM)*
*28359 Bremen, Germany*
**E-mail: roberto.tanzi@zarm.uni-bremen.de*

The most general theory of electrodynamics with linear field equations introduces a new geometry, the area metric, that regulates the propagation of light rays and massive particles instead of the usual Lorentzian metric. In the majority of the experimental situations, the area metric is expected to be a small perturbation around a metric background. In this perturbative case, two interesting results can be achieved. First, the dynamics of the area metric can be found explicitly. Second, the relative quantum theory of electrodynamics can be shown to be renormalizable and can be used to compute various fundamental processes.

I will show that, when one combines the results of quantum electrodynamics with the dynamics of an area-metric perturbation, the anomalous magnetic moment of the electron, the cross sections of Bhabha scattering, and the hyperfine splitting of the hydrogen pick up a dependence on the position. This way, measurements of the position dependence of these quantities provide a new channel to investigate area-metric deviations from a metric spacetime.

*Keywords*: general linear electrodynamics; constructive gravity; area metric; SME.

## 1. General Linear Electrodynamics and the Area Metric

The main subject of our discussion is the most general theory of electrodynamics with linear field equations, i.e., the most general theory of electrodynamics that still features the superposition principle. In particular, we wish to study the quantum effects that are produced when one considers this general linear electrodynamics as the fundamental description of the electromagnetic interactions instead of Maxwell's theory.

General linear electrodynamics has been extensively studied in the past years and many effects are currently known.[1] The starting point of our discussion is the action

$$S_{\text{GLED}}[A;\omega,G] = -\frac{1}{8}\int d^4x\,\omega F_{ab}F_{cd}G^{abcd}\,, \tag{1}$$

where $A$ is the one-form electromagnetic potential, $F = dA$ is the field strength, $G^{abcd}$ is a fourth-rank contravariant tensor field called the *area metric*, and $\omega$ is a scalar density employed to build the volume form. The area metric $G^{abcd}$ is required to satisfy, without loss of generality, the algebraic symmetries $G^{abcd} = G^{[ab][cd]} = G^{[cd][ab]}$ and to be cyclic, i.e., $G^{abcd}\epsilon_{abcd} = 0$.

Further conditions on the area metric come from the requirements that the field equations ensuing from (1) have a well-posed initial-value problem, that the spacetime is time orientable, and that general linear electrodynamic is canonically quantizable. We leave out the details of this important points since they do not play a relevant role in the present discussion. Indeed, starting from the next section, we

will consider only small area-metric perturbations around a flat metric background, in which case all the three conditions are satisfied.

At this point, it is important to note that not only does the action (1) cause a departure from Maxwell electrodynamics, but it also implies that the geometry of the spacetime is no longer described by a metric. One way to see this is by considering the geometric-optical limit of the field equations, which restricts the wave covector $k$ of light rays to satisfy the quartic dispersion relation

$$P(G,\omega)^{abcd}k_a k_b k_c k_d = 0 \tag{2}$$

rather than the familiar quadratic one $g^{ab}k_a k_b$ enforced by a Lorentzian metric $g$. In the expression above, the completely-symmetric tensor field that defines the dispersion relation is explicitly given in terms of $G^{abcd}$ and $\omega$ by the expression

$$P^{abcd} = -\frac{1}{24}\omega^2 \epsilon_{mnpq}\epsilon_{rstu} G^{mnr(a} G^{b|ps|c} G^{d)qtu}\,.$$

Furthermore, the motion of massive particles is affected in a similar way, as it was shown in Ref. 2. In particular, the momentum $p$ of a particle of mass $m$ satisfies the dispersion relation

$$P(G,\omega)^{abcd}k_a k_b k_c k_d = m^4\,. \tag{3}$$

Therefore, we see that the area metric $G^{abcd}$, together with the scalar density $\omega$, regulates the motion of point particles instead of the metric and, therefore, replaces this latter as the geometry of the spacetime.

Finally, in order to study an interacting theory of electrodynamics, it is necessary to introduce Dirac particles on an area-metric background. The general procedure to do so can be found in Ref. 3. Here, we limit ourself in quoting the fact that the $\gamma$-matrices need to satisfy the quaternary algebra

$$\gamma^{(a}\gamma^b\gamma^c\gamma^{d)} = P(G,\omega)^{abcd}\, I \tag{4}$$

rather than the usual binary one $\gamma^{(a}\gamma^{b)} = \eta^{ab}\, I$, being $I$ the identity matrix. This ensures that the Dirac field equations share the same initial-value surfaces of the field equations of general linear electrodynamics and that the causal structure of the spacetime is well defined.

### 1.1. *Linear deviations from a flat metric background*

When considered in its generality, area-metric electrodynamics (1) can produce effects that greatly differ, qualitatively and quantitatively, from Maxwell electrodynamics. However, the majority of these effects are experimentally excluded in vacuum with a fairly good accuracy. For this reason, we study the case in which the area metric is a small perturbation around Maxwell electrodynamics in a flat metric background. In particular, we consider the case in which

$$G^{abcd} = \eta^{ac}\eta^{bd} - \eta^{ad}\eta^{bc} + E^{abcd} \qquad \text{and} \qquad \omega = 1 + e\,, \tag{5}$$

where second-order corrections in $E^{abcd}$ and $e$ will be discarded.

In this perturbative regime, a great simplification is possible. It was indeed shown in Ref. 4 that it is always possible to find a (local) frame where the area-metric perturbations satisfy the conditions

$$E^{amb}{}_{m} = \frac{1}{4} E^{ij}{}_{ij} \eta^{ab} \qquad \text{and} \qquad e = -\frac{1}{8} E^{ij}{}_{ij} \,. \tag{6}$$

This frame has two major advantages. First, it is an observer frame as defined in Ref. 5, which means that results expressed in it are those measured by an observer. Second, the massive dispersion relation and the Dirac algebra reduce to the usual

$$\eta^{ab} k_a k_b = m^2 \qquad \text{and} \qquad \gamma^{(a} \gamma^{b)} = \eta^{ab} \, I \tag{7}$$

up to second order corrections. In particular, the description of massive fermions, such as the electron, is the same as in the standard theory. Finally, note that there is more than one frame in which (6) and (7) are satisfied; these frames are all related to one another by means of Lorentz transformations.

### 1.2. *The dynamics of the area metric*

So far, the area metric has been an *a priori* given and fixed geometry and we have mentioned some of the effects it produces on matter, such as the propagation of massless and massive particles. This still leaves open the questions of how matter affects the area metric and what the precise dynamics of this new geometry is.

The good news is that it is possible to *calculate*, without any need to postulate, the precise dynamics of the area metric starting from the matter action (1). This is precisely the purpose of the gravitational closure mechanism. [5] The (purely computational) bad news is that the gravitational action emerges as the solution of countably-many partial differential equations that have not been solved, so far, for a general area metric. Nevertheless, the solution to these differential equations was found explicitly in the case of linear area-metric perturbations around a flat metric background. [6] With these perturbative solutions, it is possible to find, in particular, the area metric linear perturbations sourced by a point mass. [7]

We now make the two assumptions that our experimental apparatus is at a fixed distance $R$ from the point mass and that, inside the laboratory, the area metric can be considered to be constant. Then, in one of the special frames described in the previous section, the area metric perturbations around a point mass take the form

$$\begin{gathered} E^{0\alpha 0\beta} = U(R)\, \delta^{\alpha\beta} \,, \qquad E^{0\alpha\beta\gamma} = 0 \,, \qquad E^{\alpha\beta\gamma\delta} = -U(R) \big( \delta^{\alpha\gamma} \delta^{\beta\delta} - \delta^{\alpha\delta} \delta^{\beta\gamma} \big) \,, \\ e = 3/2\, U(R) \,, \quad \text{where} \quad U(R) := -\frac{M\varepsilon}{4\pi R} e^{-\mu R} \,. \end{gathered} \tag{8}$$

In the expression above, $\varepsilon$ and $\mu$ are two integration constants of the countably-many partial differential equations and they need to be fixed by experiments.

All the results we derive will be expressed both in terms of a generic area-metric linear perturbation $E^{abcd}$ satisfying the restrictions (6) and in terms of the particular area-metric perturbation around a point mass displayed in (8).

### 1.3. *Quantization and renormalization*

The theory we consider is the general linear electrodynamics (1) in the linear-perturbation regime (5) with the restrictions (6) minimally coupled to the usual fermions. The Feynman rules can be derived with no difficulties using the path-integral formalism, as it is done in Ref. 4. One sees that the fermion propagator and the vertex are the same as in standard QED, whereas the photon propagator contains a new term linear in the area-metric perturbation. Moreover, fermion external legs can be treated in the same way as in standard QED. Photons external legs would require some extra care, but they are not needed in the processes we consider.

Finally, it is possible to prove that quantum electrodynamic with linear area-metric deviations from a flat metric background is renormalizable at every loop order in a gauge-invariant way. The complete proof, which uses the Batalin-Vilkovisky formalism, can be found in Ref. 4.

## 2. Results

We collect in this section the various processes that we have studied and the results that we have found. The derivation of these with detailed discussions can be found in Ref. 4.

### 2.1. *Bhabha scattering*

The first process is Bhabha scattering, that is the elastic scattering of an electron and a positron: $e^+e^- \to e^+e^-$. The tree-level cross section $d\sigma/d\Omega$ picks up a dependence on the area metric perturbations $E^{abcd}$ and on the angle $\phi$. Integrating out this latter, one finds

$$\frac{d\sigma}{d\cos\theta} = \frac{\pi\alpha^2}{s}\left[\frac{1}{2}(1+\cos^2\theta) + \frac{1+\cos^4(\theta/2)}{\sin^4(\theta/2)} - 2\frac{\cos^4(\theta/2)}{\sin^2(\theta/2)} + \right.$$
$$\left. - \frac{\cos^4\theta + 12\cos^2\theta + 3}{8\sin^4(\theta/2)} E^{0303} - \frac{(1+\cos^2\theta)(7+\cos^2\theta)}{32\sin^4(\theta/2)} E^{ij}{}_{ij}\right] \quad (9)$$

for a generic area metric perturbation.

When one considers the area metric perturbations (8) caused by a point mass, one finds that the tree-level cross section loses its dependence on the angle $\phi$ and, actually, becomes proportional to the usual cross section of standard QED. In particular,

$$\frac{d\sigma}{d\Omega} = \Big(1 + 2U(R)\Big)\left[\frac{d\sigma}{d\Omega}\right]_{\text{QED}}, \quad (10)$$

where the proportionality factor $1 + 2U(R)$ depends on the distance $R$ from the point mass.

### 2.2. *The anomalous magnetic moment*

The second process is the scattering of an electron in an external magnetic field. Using the Born approximation on the one-loop non-relativistic scattering amplitude, one finds the interaction potential

$$V(\vec{x}) = 2\left[1 + \frac{\alpha}{2\pi}\left(1 - \frac{E^{ij}{}_{ij}}{12}\right)\right]\left\langle\frac{e\vec{s}}{2m}\right\rangle \cdot \vec{B} \; - \frac{\alpha}{2\pi}\widehat{E}^{\alpha\beta}\left\langle\frac{es_\alpha}{2m}\right\rangle B_\beta \qquad (11)$$

between the magnetic field $\vec{B}$ and the spin of the electron $\vec{s}$.

The first summand in (11) is the usual spin-magnetic field coupling which gives the anomalous magnetic moment in terms of the Schwinger correction $\alpha/2\pi$ times the area-metric-dependent factor $1 - E^{ij}{}_{ij}/12$. Around a point mass, the one-loop anomalous magnetic moment becomes

$$a_{\text{p.m.}}^{(1-\text{loop})} = \frac{\alpha}{2\pi}\Big(1 + U(R)\Big)\,, \qquad (12)$$

whose behavior is depicted in Fig. 1.

The second summand in (11) shows the emergence of a new interaction between the spin and the magnetic field by means of the tracefree tensor $\widehat{E}^{\alpha\beta} := E^{0\alpha0\beta} + E^{ij}{}_{ij}/12$. This qualitatively-new interaction is, however, not present in the case of the area metric perturbations (8) around a point mass.

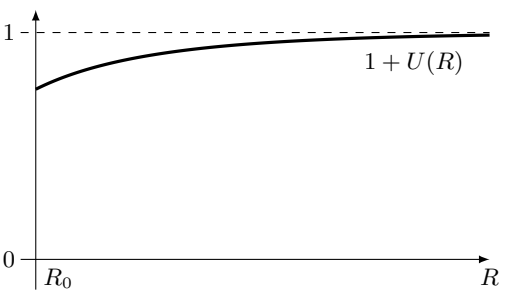



Fig. 1. The position dependence of the anomalous magnetic moment at one loop.

### 2.3. *Hyperfine splitting in hydrogen*

Finally, we consider the hyperfine splitting in the hydrogen atom. The situation is the one depicted in Fig. 2. The trace of the perturbation $E^{ij}{}_{ij}$ is responsible for a stretching of the triplet-singlet separation. The tracefree part $\widehat{E}^{\alpha\beta}$ removes the degeneracy of the triplet states. A quantitative description of these effects can be found in Ref. 4.

For the area metric (8) around a point mass, we see an explicit dependence of the singlet-triplet separation on the distance $R$ from the point mass, whereas the degeneracy of the triplet states is *not* removed.

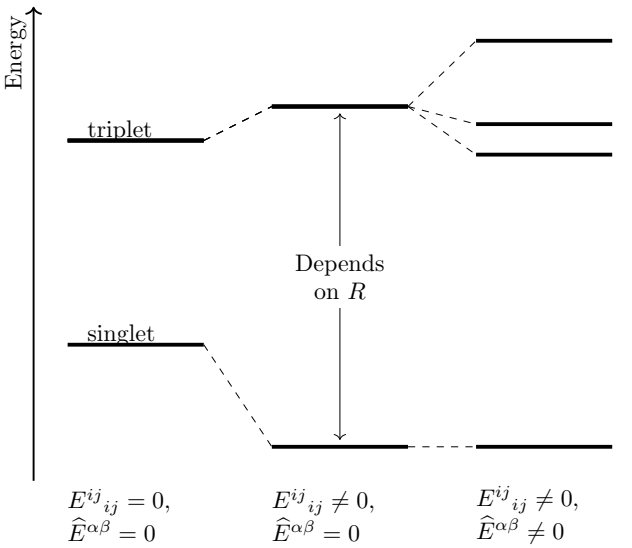



Fig. 2. The hyperfine splitting in the hydrogen atom due to the area metric perturbations.

## Acknowledgments

This talk is based on the paper[4] written in collaboration with Simon Grosse-Holz and Frederic P. Schuller. The results presented in section 2.3 were derived by Simon Grosse-Holz in his Master's thesis.

## References

1. F.W. Hehl and Y.N. Obukhov, "Foundations of Classical Electrodynamics: Charge, Flux, and Metric," Progress in Mathematical Physics , Birkhäuser Boston, 2003.
2. D. Raetzel, S. Rivera, and F. P. Schuller, "Geometry of physical dispersion relations," Phys. Rev. D **83** (2011) 044047
3. S. Rivera, "Tensorial spacetime geometries carrying predictive, interpretable and quantizable matter dynamics," PhD thesis (2012), https://publishup.uni-potsdam.de/opus4-ubp/frontdoor/index/index/docId/4495
4. S. Grosse-Holz, F. P. Schuller, and RT, "Quantum signatures of ray-optically invisible non-metricities," arXiv:1703.07183 [hep-ph], 2017.
5. M. Düll, F. P. Schuller, N. Stritzelberger, and F. Wolz, "Gravitational closure of matter field equations," arXiv:1611.08878 [gr-qc], 2016.
6. J. Schneider, F. P. Schuller, N. Stritzelberger, and F. Wolz, "Gravitational closure of weakly birefringent electrodynamic," 2017.
7. F. P. Schuller and M. C. Werner, "Etherington's Distance Duality with Birefringence," Universe **3** (2017) no. 3, 52.

# Two suprisingly not-so-different employments of bimetric geometry

Alexander A. Wierzba

*Fakultät für Physik, Ludwig–Maximilians–Universität München, München, Bayern 80799, Germany*
**E-mail: alexander.wierzba@gmx.de*

The gravitational closure of given matter field dynamics provides diffeomorphism-invariant dynamics for the very background geometry that is ultralocally employed in the matter action. Conceptual and technical key to this construction is the principal polynomial of the initially given matter field equations, which crucially depends on how exactly the background geometry makes its appearance in the matter action. In this talk, we consider two very different matter theories that employ a background geometry consisting of two Lorentzian metrics in vastly different ways. Applying the gravitational closure mechanism we derive the remarkable result that they share the same underlying gravitational dynamics.

*Keywords*: Bimetric theories; Gravity; Gravitational closure.

## 1. Gravitational Closure

In theoretical physics, one usually postulates separately, on the one hand, dynamics for a collection of matter fields on a particular four-dimensional spacetime manifold and, on the other hand, dynamics for the geometry of that spacetime which is utilized in the matter field dynamics. The prime example for this is the standard model of particle physics with its matter fields, and the — independently postulated — theory of general relativity making dynamical the spacetime metric employed in the standard model Lagrangian. The same reasoning is applied in situations where either different matter theories or different gravity theories are used.[1]

This view, however, is ill-fated, as the two theories — matter and geometry — are not independent of each other. The need for the existence of a well-defined initial value problem of the combined full theory requires the existence and co-evolution of common initial data hypersurfaces. One can show this severely restricts the admissible dynamics for the spacetime geometry.

The *gravitational closure mechanism* provides a way of implementation of these restrictions and allows to *derive* the unique gravitational dynamics for the geometrical background that is compatible with the stipulated matter field theory.[3] Thus, if having been given a matter field theory, one must not postulate the dynamical theory for the geometric background used, but instead calculate that theory.

Indeed, if one starts with Maxwell electrodynamics as the matter field theory and the spacetime metric appearing in its action as the geometric background, application of the gravitational closure mechanism will give rise to the Einstein-Hilbert action of general relativity, with two undetermined constants, the gravitational and the cosmological, appearing as constants of integration in the process.

Similarly, if one now wants to use a different matter theory, one can immediately calculate the gravity theory determining its geometric background.[9]

This contribution will study two specific, considerably different toy model matter theories which both use a geometric background consisting of two spacetime metrics.[6,7] The focus will lie on the calculation of the principal polynomial of these matter field theories, which constitutes the first crucial step of the gravitational closure mechanism. In particular, it will also look at the case of a constrained system of matter field equations and the implications this has for the application of the mechanism.

## 2. Principal Polynomial

The central object underlying the gravitational closure program is the principal polynomial of the matter field theory one takes as a start. Its origin lies in the study of approximate local wave-like solutions to the field equations and the resulting dispersion-relations for massless modes.[4] Having linear field equations of the form

$$\sum_{i=0}^{n} Q^{a_1 \dots a_i}_{\mathcal{AB}}(x) \partial_{a_1 \dots a_i} A^{\mathcal{B}}(x) = 0, \tag{1}$$

for matter fields $A^{\mathcal{B}}$, one can read off, using the highest order derivative term, the *principal symbol* of these equations, as the matrix

$$T(k)_{\mathcal{AB}} := Q^{a_1 \dots a_n}_{\mathcal{AB}}(x) k_{a_1} \dots k_{a_i}, \tag{2}$$

with $k$ being a covector. The *principal polynomial* of the matter field equations is then the cotangent space polynomial given by

$$P(k) := \det\left(T(k)\right). \tag{3}$$

The importance of this polynomial in the study of the matter field theory, stems from the fact that it has to be a *hyperbolic* polynomial everywhere in spacetime in order to ensure the existence of a well-posed initial value problem.[4] Furthermore, it enters the hypersurface deformation algebra, and thus the Poisson algebra, used in the derivations of the gravitational closure mechanism.[3]

## 3. Constrained systems

An important observation is, that the method for calculation of the principal polynomial just described, has to be refined slightly for cases in which one deals with a constrained system of field equations, such as the one presented in section 5.2. In such a case there is information present in the lower derivative order terms of the equations that is of importance to e.g. the question of existence of a well-posed initial value problem. This information, however, is clearly not accessible by considering only the highest order derivative term as in equation (3).

The reason for this is that the field equations in such a case are not in *involutive* form, in which all the information implicitly contained in them is explicitly expressed in the form of additional partial differential equations.

Fortunately, every system of partial differential equations can be brought to involutive form by application of the *Cartan-Kuranishi algorithm.*[5] This method adds to the original system all the implicitly contained information in the form of additional partial differential equations. These additional equations are always satisfied for solutions of the original system of equations and thus do not change the solution space. Furthermore, applying this algorithm also works as a test, of whether or not a system of differential equations is already in involutive form in the first place.

Only after this step has been taken, one can calculate the principal polynomial of the matter field theory, starting from the derived enlarged system of field equations that now contains all necessary information explicitly.

## 4. Principal Polynomial of Non-Quadratic Systems

The enlarged system of partial differential equations derived using the Cartan-Kuranishi algorithm will, if the theory contained constraints, be larger than the original system. In particular, since the original system of equations derived by variation of an action contains as many equations as unknowns, this enlarged system will contain more equations than there are unknowns. Thus, the principal symbol $T(k)$ one reads off from the equations is a non-square matrix. Hence, for such a system, one cannot apply equation (3) to calculate the principal polynomial.

The underlying origin of the determinant in equation (3) is to encode a condition for the *non-maximality* of the rank of $T(k)$. This condition can be written in polynomial form also for a non-square $T(k)$, by either — but not exclusively — one of the three methods presented in the next sections. While all of these will in the end lead to the same principal polynomial, they may differ in the ease of applicability and/or the difficulty of calculations necessary, from case to case.

### 4.1. *Method 1: Elimination of a pure constraint*

The first possibility is to reduce the number of equations by dropping pure constraint equations. That is equations that will only restrict the initial data that may be provided on an initial value surface, but are immediately satisfied everywhere else as a result of the remaining equations of motions.

The process of finding such pure constraints differs from case to case and may at times be tedious. Examples for this approach can be found in Ref. 8 where it is illustrated for example for Proca electrodynamics.

### 4.2. *Method 2: Minors or Gramian matrix*

It is clear that a matrix of arbitrary form has non-maximal rank if and only if all its quadratic minors of maximal size have non-maximal rank, i.e., have determinant zero. Thus, for each such minor $T_i(k)$ one finds a polynomial condition $P_i(k) := \det(T_i(k)) \stackrel{!}{=} 0$, which all have to be satisfied simultaneously. This is the case if and

only if $P(k) := \sum_i P_i(k)^2 \stackrel{!}{=} 0$, where the second expression is, by the theorem of Cauchy-Binet, just the determinant of the Grammian matrix $T(k)^t T(k)$.

### 4.3. *Method 3: Lagrange multipliers*

For each equation $C[A] = 0$ in the enlarged system of field equations that does not follow directly from variation, but is added during the course of the involutive completion, one is free to add a term $\lambda C[A]$ to the Lagrangian, introducing a Lagrange multiplier $\lambda$. This does not change the solution space of the matter field theory, as we are merely expressing explicitly, at the level of the Lagrangian, equations that have to be satisfied as a result of the original field equations anyway.

For each equation that was added to the original system, one thus introduces one new unknown quantity, the corresponding Lagrange multiplier. Varying this new Lagrangian and reading off the principal symbol of the resulting equations will thus give rise to a quadratic matrix and one can use the expression from equation (3) for calculating the principal polynomial.

## 5. Bimetric Matter Theories

As a possible model application for the gravitational closure mechanism and in particular to showcase the importance and possible implications of the necesssary application of the Cartan-Kuranishi algorithm, we will consider two — at the first glance — completely different matter theories, which both employ a geometric background consisting of two metric tensor fields.

### 5.1. *Bi-Klein-Gordan theory*

The first bimetric matter theory is one for two scalar fields $\varphi$ and $\psi$, with dynamics being governed by the action

$$\mathcal{S}_{\mathrm{BKG}}\left[\varphi,\psi;g,h\right] = \int \mathrm{d}^4x \, \left(\sqrt{-g}\, g^{ab}\left(\partial_a\varphi\right)\left(\partial_b\varphi\right) + \sqrt{-h}\, h^{ab}\left(\partial_a\psi\right)\left(\partial_b\psi\right)\right), \qquad (4)$$

combining two individual Klein-Gordan actions. However each of the scalar fields couples to a different spacetime metric.

Variation of the action gives rives to the matter field equations

$$0 = \partial_a\left(\sqrt{-g}g^{ab}\partial_b\varphi\right), \qquad (5a)$$

$$0 = \partial_a\left(\sqrt{-h}h^{ab}\partial_b\psi\right), \qquad (5b)$$

which, upon application of the Cartan-Kuranishi algorithm, are found to be in involutive form, allowing to directly read off the principal symbol

$$T_{\mathrm{BKG}}(k) = \begin{pmatrix} \sqrt{-g}g^{ab}k_ak_b & 0 \\ 0 & \sqrt{-h}h^{ab}k_ak_b \end{pmatrix}. \qquad (6)$$

Upon taking the determinant and dropping the density factor (which is not of importance for the calculations of the gravitational closure mechanism), one thus finds the principal polynomial

$$P_{\mathrm{BKG}}(k) := g^{ab}h^{ab}k_a k_b k_c k_d. \tag{7}$$

### 5.2. *Refined Proca Theory*

The second matter theory making use of a bimetric geometry is a refinement of Proca theory for a covector field $A$. Instead of adding an "isotropic" mass term $\frac{m}{2}g^{ab}A_aA_b$ to the Maxwell Lagrangian, one introduces an "anisotropic" mass term utilising a second spacetime metric, giving rise to the action

$$\mathcal{S}_{\mathrm{RPT}}\left[A; g, h\right) = \int \mathrm{d}^4x\,\sqrt{-g}\left(-\frac{1}{4}g^{ab}g^{cd}F_{ac}F_{bd} + \frac{m}{2}h^{ab}A_aA_b\right), \tag{8}$$

and upon variation the matter field equations

$$0 = \partial_a\left(\sqrt{-g}g^{ab}g^{cd}F_{bd}\right) + m\sqrt{-g}h^{cd}A_d. \tag{9}$$

Two crucial differences to the first example theory are, first, the presence of a lower order derivative term in the field equations and, second, the highest order derivative term being completely determined by only one of the two metrics. This may lead one to the, as it will turn out wrong, assumption that the other metric $h$ cannot enter into the principal polynomial.

However, most importantly, these field equations constitute a constrained system and application of the Cartan-Kuranishi algorithm will lead to the enlarged system of field equations

$$0 = \partial_a\left(\sqrt{-g}g^{ab}g^{cd}F_{bd}\right) + m\sqrt{-g}h^{cd}A_d, \tag{10a}$$

$$0 = \partial_c\left(\sqrt{-g}h^{cd}A_d\right), \tag{10b}$$

$$0 = \partial^2_{mc}\left(\sqrt{-g}h^{cd}A_d\right). \tag{10c}$$

Using one of the proposed methods for calculating the principal polynomial, one finds that it takes the form

$$P_{\mathrm{RPT}}(k) := g^{ab}h^{ab}k_a k_b k_c k_d, \tag{11}$$

which is identical to the one found in the first model theory. In particular, one realises that the second metric $h$ does indeed enter into the principal polynomial as a result of the necessary application of the Cartan-Kuranishi algorithm to get an involutive set of field equations.

## 6. Conclusions

The previous sections showed that remarkably the two, very differently looking employments of a bimetric geometric background, have the same principal polynomial. Thus they provide the same starting point for the application of the gravitational

closure mechanism. As a result one will find that the gravitational dynamics governing the two theories are indeed the same. Using the mechanism to derive the linearised version of these dynamics, one will end up with equations containing 13 scalar constants, taking the role of the two constants found in general relativity.[6,7]

The study of the refined Proca theory furthermore showcased the importance of applying the Cartan-Kuranishi algorithm before trying to calculate the principal polynomial, and thus also before applying of the gravitational closure mechanism. In particular, this study demonstrates the impact that the addition of lower order derivative terms can have on the involutivity and thus on the calculation of the principal polynomial.

Such lower order terms also appear in that matter theory which is currently considered to be the most likely one to describe the physical reality, the standard model of particle physics. Therefore, if wanting to give dynamics to the spacetime metric (or possibly also to additional parameters like mass matrices or coupling constants) appearing in the standard model lagrangian, it is vital to apply the Cartan-Kuranishi algorithm here as well.[7]

## Acknowledgments

The author thanks the TMP study program at LMU Munich for the financial support given towards attending the Marcel Grossman Meeting.

## References

1. D. Colladay and V.A. Kostelecky, *Lorentz violating extension of the standard model*, *Phys. Rev. D* **58**, 116002 (1998).
2. F.W. Hehl and Y.N. Obukhov, *Foundations of Classical Electrodynamics* (Birkhäuser, Basel, 2003).
3. F.P. Schuller, N. Stritzelberger, F. Wolz and M. Düll, *Gravitational closure of matter field equations*, *Phys. Rev. D* **97**, 084036 (2016).
4. D. Raetzel, S. Rivera and F.P. Schuller, *Geometry of physical dispersion relations*, *Phys. Rev. D* **83**, 044047 (2011).
5. W.M. Seiler, *Involution — The Formal Theory of Differential Equations and its Applications in Computer Algebra* (Springer, Berlin Heidelberg, 2010)
6. U. Beier, *Perturbative gravitational closure of a bi-metric matter theory*, MSc thesis (FAU Erlangen-Nürnberg, 2018).
7. A. Wierzba, *Parametrodynamics*, MSc thesis (LMU München, 2018).
8. G. Velo and D. Zwanziger, *Noncausality and Other Defects of Interaction Lagrangians for Particles with Spin One and Higher*, *Phys. Rev.* **188**, 2218-2222 (1969).
9. J. Schneider, F.P. Schuller, N. Stritzelberger and F. Wolz, *Gravitational closure of weakly birefringent electrodynamics*, `arXiv:hep-th/1708.03870` (2017).

# Causal structure of matter field equations

Florian Wolz

*Institute for Theoretical Physics, University of Hannover,*
*Appelstrasse 2, D-30167 Hannover, Germany*
**E-mail: florian.wolz@fau.de*

The causal structure of given matter field equations provides the crucial input from which the constructive gravity program starts. It is therefore of paramount importance to successfully address a number of subtle issues, which do not arise in the simplest examples in the mathematical literature, but urgently must be taken into account for physically realistic models. This talk presents a recipe to handle these complications. In particular, we will focus on how to deal with non-scalar systems of equations, gauge symmetries and implicit information that needs to be made explicit before the causal structure, encoded in the so-called principal polynomial, can be calculated correctly.

*Keywords*: Constructive Gravity

## 1. Introduction

The causal structure of field equations tells every physicist what events are connected to which events, and is intimately linked to the notion of initial value formulations. Not any mathematical theory, in the form of a partial differential equation, can be used to formulate a predictive theory, as is demonstrated already in the case of Maxwell electrodynamics on a metric background: only for a metric of Lorentzian signature we find an initial value formulation. This was used in the whole standard model of particle physics and general relativity by making sure that the causal structure does not differ from the causality that light propagates on. Unfortunately, this process led to the idea that as long as a Lagrangian is constructed in a Lorentz-invariant way, the resulting equations of motion will possess the well-known causal structure.

However, a surprising example that a Lagrangian that is constructed in a fully Lorentz-invariant fashion but still has a causality that differs from the standard Lorentzian causal structure was given by Velo-Zwanziger[1]. The analyzed Lagrangian is a modification of covariant Maxwell electrodynamics by a Proca mass term and a quartic term in the covector field $A$, i.e.

$$\mathcal{L}_{\mathrm{VZ}}[A] = -\frac{1}{4}F_{ab}[A]F^{ab}[A] + \frac{1}{2}m^2A^2 + \frac{1}{2}\lambda\left(A^2\right)^2 \tag{1}$$

The causal structure, encoded in the principal polynomial that describes the characteristic surfaces of the equations of motions, is given by

$$P(n) = (n^2)^3 \cdot \left[n^2 + \lambda m^{-2}(n^2A^2 + 2(n\cdot A)^2)\right] . \tag{2}$$

This is a product of the standard term $n^2$ and a modification term, resulting in massless modes propagating with another speed of light that – worst-of-all – depends on the particular solution $A_a$ to the equations of motion. Although the highest derivative order coefficient is the input for the calculation of the principal polynomial

and is seemingly not influenced by the modification terms, Velo-Zwanziger showed that one can derive a constraint that *does* change the characteristic surfaces.

Nowadays this is not surprising at all and is related to the fact that the equations of motion are no *involutive differential equations*. Intuitively, this type of differential equations has all information that is required to construct a formal power series solution explicitely present. In other words, they do not contain any *hidden* information that needs to be extracted by tedious steps, as is the case in the Velo-Zwanziger example. Due to the *Cartan-Kuranishi algorithm* it turns out, that any differential equation can be brought into involutive form in finitely many steps.

This article will first present the notion of involutive differential equations and explain the Cartan-Kuranishi algorithm [4]. Afterwards we tackle a practical example that represents a generalization of the case analyzed by Velo-Zwanziger.

## 2. Involutive differential equations

Before we can properly introduce the notion of an involutive differential equation, let us first give some definitions of the underlying mathematical theory. A system of differential equation – in the following simply referred to as differential equation – is formulated over *independent variables* $x^1, \ldots, x^n$ and *dependent variables* $u^1, \ldots u^m$. In the case of point particle equations of motions the only independent variable is the time coordinate $t$ and the generalized coordinates $q^i$ being the dependent variables. For field theories, the independent variables are given by spacetime points and the field values are the dependent variables of the theory. In the case of the system of differential equations called the gravitational closure equations of the constructive gravity program [3] the independent variables correspond to the local degrees of freedom of the gravitational field and their derivatives and the dependent variables are the coefficients of the gravitational Lagrangian.

Locally we treat a differential equations as a collection of implicit functions that depend on the independent variables $x^i$, the dependent variables $u^\alpha$ and their derivatives, i.e.

$$\Phi^\tau(x, u^\alpha, p^\alpha_\mu) = 0\,, \tag{3}$$

where $p^\alpha_\mu = \frac{\partial^{|\mu|} u^\alpha}{\partial x^\mu}$ are the corresponding derivatives in the direction of the independent variables, and $\mu$ being a multi-index. The order of the differential equations will be denoted by $q$ in this article.

An object of incredible importance is the *geometric symbol* of an differential equation. It is defined by the linear equation [2]

$$\mathcal{M}_q: \quad \sum_{|\mu|=q} \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}\, v^\alpha_\mu = 0 \tag{4}$$

which is the kernel of the coefficient $Q^{\tau\,\mu}_{\ \ \alpha} := \frac{\partial \Phi^\tau}{\partial p^\alpha_\mu}$. This matrix can be brought into reduced row echelon form in order to read off which coefficients $p^\alpha_\mu$ can be solved

for in the differential equation. For everything that follows it suffices to summarise this information in a collection of coefficients that are defined in the following way:

For each column of the matrix $Q$, i.e. the different derivatives of the dependent variables, we associate a *class*. For each multi-index $\mu = [i_1, \ldots, i_n]$ we define its class as the index $k$ of the first non-vanishing $i_k$. Then the columns of $Q$ can be sorted in a class respecting order (from highest to lowest), and after bringing the matrix into row echelon form we can look for the first non-zero entry of each row. The corresponding $p^\alpha_\mu$ of this entry is called the leader of the row of class $k$. Then let $\beta^{(k)}_q$ denote the leaders of a row of class $k$, i.e. the number of highest order derivative terms of class $k$ that could be solved for[2].

The symbol of the differential equation is then called *involutive* if the rank of the prolongation of the differential equation, i.e. the derivatives of the differential equation $\Phi^\tau$ by all independent variables, can be determined as the following sum[2]

$$\operatorname{rank} \mathcal{M}_{q+1} = \sum_{k=1}^{n} k \cdot \beta^{(k)}_q \tag{5}$$

Is this the case, one can derive a recursion relation to calculate the coefficients $\beta^{(k)}_{q+1}$ directly from the coefficients $\beta^{(k)}_q$, meaning that we can completely predict what coefficients we can solve for in the prolongations of the differential equation. If the symbol $\mathcal{M}_q$ is involutive, it turns out that the same will be true for all further prolongations $\mathcal{M}_{q+r}$, for any $r \geq 0$. Furthermore, it can even be shown that there always exist an integer $r$ such that the symbol $\mathcal{M}_{q+r}$ will be involutive, and thus all further prolongations. Note that this, however, does increase the differential order of the equation. In practise, this can lead to the scenario where a seemingly second order field equation is in fact of higher order and is then possible equipped with Ostrogradski ghosts[6].

"Hidden information" can still be contained in the differential equations. If we can algebraically manipulate the prolongation of the differential equations – equations of order $q+1$, such that we obtain an equation of order $q$ then this is considered an *integrability condition* and needs to be added to the system[2]. Of course, no new information is added, it is simply made explicit and can then properly be considered in all further calculations – as in the calculation of the causal structure.

If we revealed hidden integrability conditions the symbol of the equation at order $q$ possibly changed, so we need to make sure again that it is involutive in the way described above. Once the symbol is involutive again, we can search for new integrability conditions. If no condition exists, then it can be shown that for no further prolongation an integrability condition can be found: all hidden information is made explicit and the differential equations (with all the integrability conditions identified) is called *involutive*[2].

The algorithm to bring the differential equation into involutive form is the so-called *Cartan-Kuranishi* algorithm and can be performed for any differential equations[4]. Once this step is performed all information that can contribute to

the characteristic surfaces and the causal structure of the differential equation is present.

As a result, we can then properly calculate the principal polynomial of the differential equation and obtain the causal structure. If one applies all the steps of the Cartan-Kuranishi algorithm for the Euler-Lagrange equations of the Lagrangian considered by Velo-Zwanziger, it turns out that there are indeed hidden integrability conditions. Calculation of the principal polynomial then recovers their result. Note, however, that the result was obtained in a completely algorithmic fashion.

## 3. Example: Generalizations of Maxwell electrodynamics

We now want to put the algorithm presented above to good use by analyzing a specific example. The setup is the following:

$$\mathcal{L}[A;\eta) = -\frac{1}{4}F^{ab}F_{ab} + V(A)\,, \tag{6}$$

i.e. we take the Lagrangian of Maxwell electrodynamics on a flat Minkowskian background $\eta$ and add an arbitrary function of the covector field $A_m$.

The equations of motions are calculated by variation of the action with respect to $A$. This yields a differential equation of $2^{\text{nd}}$ order

$$0 = 2\,\eta^{m[n}\eta^{p]a}\partial_{mn}A_p + \frac{\partial V}{\partial A_a}(A)\,. \tag{7}$$

In this case, by the choice of our function $V$, we see that the geometric symbol coincides with the term obtained for standard electrodynamics. To obtain the four $\beta_2^{(i)}$ coefficients, we first pick an order for our independent variables. Since we want to solve for as many $\ddot{A}_m$ terms as possible in order to be able to distinguish between evolution equations and constraints, we make the choice $(x^\alpha, t)$, i.e. time comes *after* the spatial components and has the highest class. As the next step, we need to write down the geometric symbol and sort the columns in our chosen order. From this one finds that the coefficients read

$$\beta_2^{(1)} = 0 \qquad \beta_2^{(2)} = 0 \qquad \beta_2^{(3)} = 1 \qquad \beta_2^{(4)} = 3 \tag{8}$$

This just reflects the well-known fact that in Maxwell's equations we have three evolution equations and one constraint. In order to check if the geometric symbol is involutive, we need to calculate the rank at the next order and see if it is given by

$$\text{rank}\mathcal{M}_3 = \sum_{k=1}^{4} k\cdot\beta_2^{(k)} = 15 \tag{9}$$

Although tedious, the calculation can easily be performed by computer algebra software, with the result that the symbol is indeed involutive.

The next step consists in checking for hidden integrability conditions. By considering the divergence of the equations of motion (7) we find the integrability

condition

$$0 = \frac{\partial^2 V}{\partial A_m \partial A_n}(\partial_n A_m) =: H^{mn}(\partial_n A_m)\,, \tag{10}$$

which is of 1st order. This means we have to add it to the system to make the information explicit. For our new system, we repeat the steps from above, i.e. check if the symbol is involutive. Since the new equation (10) is of 1st order, the symbol remains unchanged and so it is still involutive. But clearly, if we prolong the integrability condition (10) we find another 2nd order equation that also has to be added to the system. The geometric symbol of the system has then changed, so we need to read off the $\beta_2^{(k)}$ coefficients again. We find that

$$\beta_2^{(1)} = 1 \qquad \beta_2^{(2)} = 1 \qquad \beta_2^{(3)} = 2 \qquad \beta_2^{(4)} = 4 \tag{11}$$

and that the geometric symbol is still involutive. If we now check for integrability conditions we find that there are no further conditions. As a result our final system

$$\begin{aligned} 0 &= 2\eta^{m[n}\eta^{p]a}(\partial_{mn}A_p) + \frac{\partial V}{\partial A_a}(A) \\ 0 &= H^{mn}(A)(\partial_n A_m) \\ 0 &= H^{mn}(A)(\partial_{np}A_m) + (\partial_p H^{mn}(A))(\partial_n A_m) \end{aligned} \tag{12}$$

is involutive and thus, all hidden information is made explicit.

We can now analyze the causal structure of this system. By making a Wentzel-Kramers-Brillouin approximation to lowest order via

$$A_m(x) = \mathrm{Re}\left\{\exp\left(iS(x)/\lambda\right)\left[a_m(x) + \mathcal{O}(\lambda)\right]\right\}\,, \tag{13}$$

only the highest derivative order terms survive and the derivative operators are replaced by the covector $n_a$ that is the derivative of the eiknonal function $S(x)$. This equation has non-trivial solutions in case that the principal polynomial, defined by

$$P(n) = \det\left(T^t(n)\cdot T(n)\right)\,, \tag{14}$$

vanishes, where $T$ are the coefficients of the differential equation in front of the amplitudes $a_m$ and the operation $t$ denotes the matrix transpose. One finds that in general we have the principal polynomial

$$P(k) = (\eta^{mn}k_m k_n)^3 \cdot (H^{pq}(A)k_p k_q)\,. \tag{15}$$

This gives, if we eliminate repeated factors, a principal polynomial of degree 4. We see that in general this is not the standard Lorentzian causal structure, but resembles the causality of a bi-metric theory, with the second metric being constructed by the four-potential $A$ itself. This particularly indicates that in this case the dichotomy of matter fields and a geometric field $\eta$ breaks down, since $A$ acts as both.

Note that in order to have a well-defined initial value problem and a canonically quantizable theory it is necessary that all factors of the principal polynomial are hyperbolic. This tells us that $H^{pq}$ has to be of Lorentzian signature[5].

Only for two special cases we recover the standard principal polynomial of the standard model: Since only the $2^{\text{nd}}$ derivative of $V$ appears in the principal polynomial, any linear term in the gauge field clearly obviously leaves the causal structure unchanged. This is not surprising, since this just contributes a coupling of the covector field $A_m$ to a current, i.e.

$$V(A) = j^m A_m \,. \tag{16}$$

The other case occurs in case the second derivative is proportional to the flat metric $\eta$, i.e.

$$\frac{\partial^2 V}{\partial A_p \partial A_q} \propto \eta^{pq} \,. \tag{17}$$

In this case $V$ only contributes as a repeated factor to the principal polynomial. This corresponds to a Proca mass term

$$V(A) = m^2 \, \eta^{ab} A_a A_b \,. \tag{18}$$

While this does break the $U(1)$ gauge invariance – which can be seen in the involutive system from the $\beta_2^{(4)}$ coefficient, since we now have four evolution equations and no constraint – the causal structure is not altered.

## 4. Conclusion

In this article, we presented how to perform the Cartan-Kuranishi to bring differential equations, especially equations of motions into involutive form in order to make all hidden information in the differential equations explicit. This is necessary in order to correctly calculate the principal polynomial that encodes the causal structure of the theory one considers.

We applied this to the example of a modification of Maxwell electrodynamics by adding an arbitrary function $V(A)$. If $V$ is not a linear term in $A_a$, i.e. a coupling to a current or a Proca mass term, in general the causal structure is modified. If one had not revealed all the hidden integrability conditions of the equations of motion, one would not have ended up with the result.

## References


1. G. Velo and D. Zwanziger, *Phys. Rev.* **188**, (1969) 2218.
2. W. Seiler, *Involution: The Formal Theory of Differential Equations and Its Applications in Computer Algebra*, 1st edition (Springer Publishing Company, Incorporated, 2009).
3. M. Düll, F.P. Schuller, N. Stritzelberger and F. Wolz, *Phys.Rev.* **D97** (2018) no.8, 084036.
4. M. Kuranishi, *American Journal of Mathematics*, **vol. 79**, no. 1, (1957).
5. D. Raetzel, S. Rivera, F. P. Schuller, *Phys.Rev.* **D83** (2011) 044047.
6. M. Ostrogradsky, *Mem. Ac. St. Petersbourg* **VI** 4 (1850) 385.

PART B

PROCEEDINGS OF THE FIFTEENTH

# MARCEL GROSSMANN MEETING ON GENERAL RELATIVITY

Editors

**Elia S. Battistelli**

**Robert T. Jantzen**

**Remo Ruffini**

World Scientific

# THE FIFTEENTH
# MARCEL GROSSMANN MEETING

**On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories**

This page intentionally left blank

PART B

# THE FIFTEENTH
# MARCEL GROSSMANN MEETING

## On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories

Proceedings of the MG15 Meeting
on General Relativity
University of Rome "La Sapienza", Italy 1–7 July 2018


Editors

**Elia S. Battistelli**
University of Rome "La Sapienza"
Rome, Italy

**Robert T. Jantzen**
Villanova University
Villanova, PA, USA

Series Editor

**Remo Ruffini**
International Center for Relativistic Astrophysics (ICRA)
University of Rome "La Sapienza"
Rome, Italy
International Center for Relativistic Astrophysics
Network (ICRANet)
Pescara, Italy



World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

*Published by*

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

*USA office:* 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

*UK office:* 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Control Number: 2022906140

**British Library Cataloguing-in-Publication Data**
A catalogue record for this book is available from the British Library.

**THE FIFTEENTH MARCEL GROSSMANN MEETING**
**On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories**
**(In 3 Volumes)**



ISBN 978-981-125-824-4 (set_hardcover)
ISBN 978-981-125-825-1 (set_ebook for institutions)
ISBN 978-981-125-826-8 (set_ebook for individuals)

ISBN 978-981-125-661-5 (vol. 1_hardcover)

ISBN 978-981-125-662-2 (vol. 2_hardcover)

ISBN 978-981-125-663-9 (vol. 3_hardcover)

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/12843#t=suppl

Desk Editor: Ng Kah Fee

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

**THE MARCEL GROSSMANN MEETINGS**

**Series Editor:** REMO RUFFINI

**Publications in the Series of Proceedings**

Proceedings of the Fifteenth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2018)
Edited by E.S. Battistelli, R.T. Jantzen, R. Ruffini
World Scientific, 2021

Proceedings of the Fourteenth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2015)
Edited by M. Bianchi, R.T. Jantzen, R. Ruffini
World Scientific, 2017

Proceedings of the Thirteenth Marcel Grossmann Meeting on General Relativity
(Stockholm, Sweden, 2012)
Edited by K. Rosquist, R.T. Jantzen, R. Ruffini
World Scientific, 2015

Proceedings of the Twelfth Marcel Grossmann Meeting on General Relativity
(Paris, France, 2009)
Edited by T. Damour, R.T. Jantzen, R. Ruffini
World Scientific, 2012

Proceedings of the Eleventh Marcel Grossmann Meeting on General Relativity
(Berlin, Germany, 2006)
Edited by H. Kleinert, R.T. Jantzen, R. Ruffini
World Scientific, 2007

Proceedings of the Tenth Marcel Grossmann Meeting on General Relativity
(Rio de Janiero, Brazil, 2003)
Edited by M. Novello, S. Perez-Bergliaffa, R. Ruffini
World Scientific, 2005

Proceedings of the Ninth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2000)
Edited by V.G. Gurzadyan, R.T. Jantzen, R. Ruffini
World Scientific, 2002

Proceedings of the Eighth Marcel Grossmann Meeting on General Relativity
(Jerusalem, Israel, 1997)
Edited by T. Piran
World Scientific, 1998

Proceedings of the Seventh Marcel Grossmann Meeting on General Relativity
(Stanford, USA, 1994)
Edited by R.T. Jantzen and G.M. Keiser
World Scientific, 1996

Proceedings of the Sixth Marcel Grossmann Meeting on General Relativity
(Kyoto, Japan, 1991)
Edited by H. Sato and T. Nakamura
World Scientific, 1992

Proceedings of the Fifth Marcel Grossmann Meeting on General Relativity
(Perth, Australia, 1988)
Edited by D.G. Blair and M.J. Buckingham
World Scientific, 1989

Proceedings of the Fourth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 1985)
Edited by R. Ruffini
World Scientific, 1986

Proceedings of the Third Marcel Grossmann Meeting on General Relativity
(Shanghai, People's Republic of China, 1982)
Edited by Hu Ning
Science Press – Beijing and North-Holland Publishing Company, 1983

Proceedings of the Second Marcel Grossmann Meeting on General Relativity
(Trieste, Italy, 1979)
Edited by R. Ruffini
North-Holland Publishing Company, 1982

Proceedings of the First Marcel Grossmann Meeting on General Relativity
(Trieste, Italy, 1975)
Edited by R. Ruffini
North-Holland Publishing Company, 1977

## SPONSORS

International Center for Relativistic Astrophysics Network (ICRANet)
International Center for Relativistic Astrophysics (ICRA)
International Centre for Theoretical Physics (ICTP)
Istituto Nazionale di Astrofisica (INAF)
International Union of Pure and Applied Physics (IUPAP)

## FREEDOM OF MOVEMENT FOR SCIENTISTS

The Marcel Grossman Meetings were founded with the premise that scientists of all nations have a right to meet to exchange knowledge independent of national borders. As such we affirm the IUPAP declaration: "To secure IUPAP sponsorship, the organisers have provided assurance that MG15 will be conducted in accordance with IUPAP principles as stated in the IUPAP resolution passed by the General Assembly in 2008. In particular, no bona fide scientist will be excluded from participation on the grounds of national origin, nationality, or political considerations unrelated to science."

## ACKNOWLEDGEMENTS

We acknowledge the outstanding job done before, during and after the meeting by the ICRANet/ICRA administrative and secretarial staff: Cristina Adamo, Silvia Latorre, Elisabetta Natale, and Cinzia di Niccolo and the conference agency support of Sistema Congressi (Marisa Santori, Ilaria Coccato, Lorena Moretti). Finally this meeting and its proceedings could not have functioned without the dedicated IT support of the ICRA system manager Vittorio Vannini for every MG meeting this century and the ICRANet system manager Gabriele Brandolini, with some temporary assistance from Valerio Antonucci and Damiano Verzulli.

## ORGANIZING BODIES OF THE FIFTEENTH MARCEL GROSSMANN MEETING:

### INTERNATIONAL ORGANIZING COMMITTEE

David Blair, Yvonne Choquet-Bruhat, Thibault Damour, Paolo De Bernardis, Francis Everitt, Riccardo Giacconi, Theodor Haensch, Stephen Hawking, Marc Henneaux, Christine Jones Forman, Roy Kerr, Hagen Kleinert, Jutta Kunz, Claus Laemmerzahl, Malcolm Longair, Tsvi Piran, Jorge Rueda, Remo Ruffini (Chair), Misao Sasaki, Humitaka Sato, Rashid Sunayev, Gerard 't Hooft, Steven Weinberg, Shing-Tung Yau

### LOCAL ORGANIZING COMMITTEE

Amati, L., Angelantonj, C., Barbiellini, G., Bassan, M., Battistelli, E. S. (chair), Belinski, V.A., Belli, L., Benedetti, R., Bernabei, R., Bianchi, M., Bianco, C.L. (co-chair), Bini, D., Buchert, T., Burgio, F., Capozziello, S., Chakrabarti, S.K., Chardonnet, P., Dall'Agata, G., De Angelis, A., De Bernardis, P., Della Valle, M., Di Virgilio, A., Frasca S., Fraschetti, F., Fré, P., Frontera, F., Giavalisco, M., Giommi, P., Gionti, G., Ingrosso, G., Jantzen, R.T., Jetzer, P., Lee, H.-W., Lerda, A., Liberati, S., Longo, R., Luzzi, G., Mandolesi, N., Marmo, G., Menotti, P., Merafina, M., Morselli, A., Pani, P., Pelster, A., Piacentini, F., Pian, E., Procesi, C., Quevedo, H., Ricci, F., Riccioni, F., Rosati, P., Ruffo, S., Scarpetta, E.V., Tavani, M., Titarchuk, L.G., Vereshchagin, G.V., Xue, S.-S., Zen Vasconcellos, C.

### INTERNATIONAL COORDINATING COMMITTEE

ALBANIA: Hafizi, M. ARGENTINA: Arguelles, C., Ghezzi, C.R., Mirabel, F., Romero, G.E. ARMENIA: Sahakyan, N. AUSTRALIA: Blair, D., Ju, L., Lun, A., Manchester, D., Melatos, A., Quinn, P., Scott, S.M., Steele, J.D., AUSTRIA: Aichelburg, P.C., Schindler, S. BELARUS: Kilin, S., Minkevich, A.V., Siutsou, I.A. BELGIUM: Henneaux, M. BOLIVIA: Aguirre, C.B. BOSNIA AND HERZEGOVINA: Pasic, V. BRAZIL: Aguiar, O., Barres de Almeida, U., Berkovits, N.J., Carneiro da Cunha, B., de Gouveia Dal Pino, E.M., de Souza Oliveira Filho, K., Guzzo, M., Hadjimichef, D., Klippert, R., Malheiro, M., Negreiros, R., Peres Menezes, D., Romero Filho, C.A, Shellard, R., Villela, T., Wuensche, C.A., Zen Vasconcellos, C. BULGARIA: Yazadjiev, S. CANADA: Singh, D., Smolin, L., Turok, N. CHILE: Bunster Weitzman, C., Giacomini, A., Reisenegger, A. CHINA: Cai, R.-G., Cao, Z., Chang, J., Chen, J.-S., Chen, X.-L, Dai, Z.-G, Feng, L.-L., Han, W.-B., Jing, Y.-P., Li, T.-P., Lin, W.-B., Lou, Y.-Q., Luo, J., Mei, J.-W., Wang, Y., Wu, X.-P., Wu, Y.-L., Yuan, Y.-F., Zhang, B., Zhang, S.-N., Zhao, G. CHINA (TAIWAN): Chen, C.-M., Chen, P., Lee, D.-S., Lee, W.-L., Ni, W.-T. COLOMBIA: Bargueño de Retes, P., Gonzalez, G., Granda Velasquez, L.N., Núñez, L., Pachón Contreras, L.A., Portilla, B., Valenzuela Toledo, C.A. CROATIA: Dominis Prester,

D., Milekovic, M., Smolčič, V., Smolić, I., Surić, T. CUBA: Pérez Martínez, A., Pérez Rojas, H., Quiros, I. CZECH REPUBLIC:Bicak, J. DENMARK: Naselsky, P. EGYPT: Tawfik, A.N., Wanas, M.I. ESTONIA: Einasto , J., Saar, E. FINLAND: Poutanen, J., Volovik, G. FRANCE: Brillet, A., Buchert, T., Chardonnet, P., Coullet, P., de Freitas Pacheco, J.A., Deruelle, N., Iliopoulos, J., Mignard, F., GEORGIA: Lavrelashvili, G., Machabeli, G. GERMANY: Biermann, P., Blumlein, J., Di Piazza, A., Fritzsch, H., Genzel, R., Gilmozzi, R., Hehl, F., Keitel, C., Kiefer, C., Mirzoyan, R., Neugebauer, G., Nicolai, H., Renn, J., Ringwald, A., Ruediger, A. GREECE: Batakis, N.A., Cotsakis, S., Vagenas, E.C. HUNGARY: Fodor, G., Levai, P. ICELAND: Bjornsson, G., Jakobsson, P. INDIA: Chakrabarti, S.K.,Iyer, B., Padmanabhan, T., Souradeep, T. IRAN: Mansouri, R., Mashhoon, B., Sobouti, Y. IRELAND: O'Murchada, N., Pe'er, A. ISRAEL: Milgrom, M., Nakar, E., Piran, T. ITALY: Belinski, V.A., Bianchi, M., Bianco, C.L., Cherubini, C., Della Valle, M., Falciano, S., Filippi, S., Menotti, P., Merafina, M., Pani, P., Ricci, F., Treves, A., Vereshchagin, G.V., Vitale, S., Xue, S.-S. JAPAN: Fujimoto, M.-K., Makishima, K., Nakamura, T., Sato, K., Shibata, M. KAZAKHSTAN: Abishev, M., Aimuratov, Y., Boshkayev, K., Mychelkin, E.G., Spitaleri, C. KOREA (PYEONGYANG): Kim, J.S. KOREA (SEOUL): Kim, S.P., Kim, S.-W., Lee, H.K., Lee, H.-W., van Putten, M.KYRGYZSTAN: Gurovich, V.Ts. LIBYA: Gadri, M. MEXICO: Breton, N., García-Diaz, A.A., Macías Alvarez, A., Mielke, E.W., Quevedo, H., Rodriguez, L.F. NEW ZEALAND: Visser, M., Wiltshire, D. NORWAY: Elgarøy, Ø., Fonseca Mota, D., Knutsen, H. POLAND: Belczynski, K., Demianski, M., Lewandowski, J., Nurowski, P., Sokolowski, L. PORTUGAL: Costa, M., da Silva, A., Lemos, J.P.S., Lobo, F., Moniz, P., Silva, L.O. ROMANIA: Visinescu, M. RUSSIA: Aksenov, A.G., Arkhangelskaja, I., Bisnovatyi-Kogan, G.S., Blinnikov, S.I., Chechetikin, V.M., Cherepaschuk, A.M., Khriplovich, I.B., Lipunov, V.M., Lukash, V.N., Novikov, I.D., Rudenko, V.N., Starobinsky, A.A. SERBIA: Djordjevic, G., Jovanovic, P., Knežević, Z., Pankov-Hzvojevic, M., Popovic, L., Prodanovic, T., M., Sijacki, D., Simic, S. SLOVAKIA: Balek, V. SLOVENIA: Cadez, A., Gomboc, A., Zavrtanik, D. SOUTH AFRICA: Colafrancesco, S., Larena, J., Maharaj, S., SPAIN: Elizalde, E., Ibanez, J., Perez Mercader, J., Verdaguer, E. SWEDEN: Abramowicz, M.A., Marklund, M., Ryde, F. SWITZERLAND: Durrer, R., Jetzer, P. TURKEY: Aliev, A., Gurses, M. UKRAINE: Novosyadlyj, B., Zaslavski, O.B., Zhuk, A. UK: Cruise, A.M., Frenk, C.S., Green, M., Mavromatos, N., Perry, M., Willingale, R. USA: Abel, T., Ashtekar, A., Bardeen, J.,Carlstrom, J., Cornish, N., Dermer, C., Fan, X., Flanagan, E., Fraschetti, F., Fryer, C., Incera, V., Jantzen, R.T. (Chairperson), Kolb, R., Laguna, P., Longair, M., Lousto, C., Madau, P., Mathews, G., Matzner, R., Melia, F., Mester, J., Michelson, P., Nordtvedt, K., Parker, L., Pretorius, F., Pullin, J., Shapiro, I., Shapiro, S., Shoemaker, D., Smoot, G., Stiavelli, M., Teukolsky, S., van Nieuwenhuizen, P., Zhang, B. UZBEKISTAN: Ahmedov, B., Zalaletdinov, R.M. VATICAN CITY: Gionti, G. VENEZUELA: Herrera, L. VIETNAM: Long, H.N.

## MARCEL GROSSMANN AWARDS

## FIFTEENTH MARCEL GROSSMANN MEETING

### Institutional Award

Planck Scientific Collaboration (ESA) (ESA)

*"for obtaining important constraints on the models of inflationary stage of the Universe and level of primordial non-Gaussianity; measuring with unprecedented sensitivity gravitational lensing of Cosmic Microwave Background fluctuations by large-scale structure of the Universe and corresponding Bpolarization of CMB, the imprint on the CMB of hot gas in galaxy clusters; getting unique information about the time of reionization of our Universe and distribution and properties of the dust and magnetic fields in our Galaxy."*
— presented to its Director General Johann-Dietrich Woerner

### Institutional Award

Hansen Experimental Physics Laboratory (HEPL) at Stanford University

*"for having developed interdepartmental activities at Stanford University at the frontier of fundamental physics, astrophysics and technology."*

### Individual Awards

Lyman Page
*"for his collaboration with David Wilkinson in realizing the NASA Explorer WMAP mission and who now leads the Atacama Cosmology Telescope as its project scientist."*

Rashid Alievich Sunyaev
*"for the development of theoretical tools in the scrutinising, through the CMB, of the first observable electromagnetic appearance of our Universe."*

Shing-Tung Yau
*"for the proof of the positivity of total mass in the theory of general relativity and perfecting as well the concept of quasi-local mass, for his proof of the Calabi conjecture, for his continuous inspiring role in the study of black holes physics ."*

Each recipient is presented with a silver casting of the TEST sculpture by the artist A. Pierelli. The original casting was presented to His Holiness Pope John Paul II on the first occasion of the Marcel Grossmann Awards.

**FOURTEENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

European Space Agency (ESA)

**Individual Awards**

Frank Chen Ny Yang, Tsung Dao Lee, Ken'ichi Nomoto, Martin Rees, Yakov G. Sinai, Sachiko Tsuruta

**THIRTEENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

ALBANOVA UNIVERSITY CENTER, STOCKHOLM

**Individual Awards**

David Arnett, Vladimir Belinski and Isaak M. Khalatnikov, Filippo Frontera

**TWELFTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Institut des Hautes Études Scientifique (IHES)

**Individual Awards**

Jaan Einasto, Christine Jones, Michael Kramer

**ELEVENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Freie Universität Berlin

**Individual Awards**

Roy Kerr, George Coyne, Joachim Trümper

**TENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

CBPF (Brazilian Center for Research in Physics)

**Individual Awards**

Yvonne Choquet-Bruhat, James W. York, Jr., Yval Ne'eman

**NINTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Solvay Institutes

**Individual Awards**

Riccardo Giacconi, Roger Penrose, Cecile and Bryce DeWitt

**EIGHTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Hebrew University of Jerusalem

**Individual Awards**

Tullio Regge, Francis Everitt

**SEVENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Hubble Space Telescope Institute

**Individual Awards**

SUBRAHMANYAN CHANDRASEKHAR, JIM WILSON

**SIXTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Research Institute for Theoretical Physics (Hiroshima)

**Individual Awards**

Minora Oda, Stephen Hawking

**FIFTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The University of Western Australia

**Individual Awards**

Satio Hayakawa, John Archibald Wheeler

**FOURTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Vatican Observatory

**Individual Awards**

William Fairbank, Abdus Salam

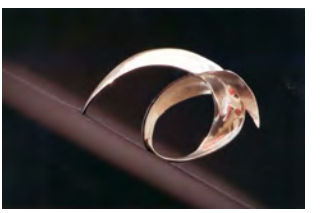


Fig. 1. TEST: sculpture by Attilio Pierelli

**PLANCK SCIENTIFIC COLLABORATION (ESA)**
presented to **Jean-Loup Puget**, the Principal Investigator of the High Frequency Instrument (HFI).

*"for obtaining important constraints on the models of inflationary stage of the Universe and level of primordial non-Gaussianity; measuring with unprecedented sensitivity gravitational lensing of Cosmic Microwave Background fluctuations by large-scale structure of the Universe and corresponding B-polarization of CMB, the imprint on the CMB of hot gas in galaxy clusters; getting unique information about the time of reionization of our Universe and distribution and properties of the dust and magnetic fields in our Galaxy"*

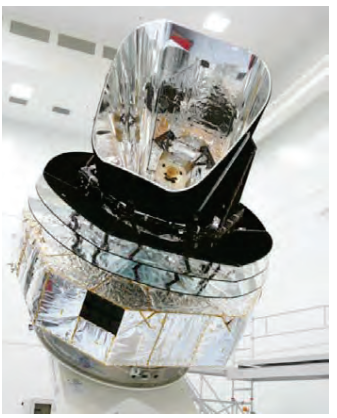
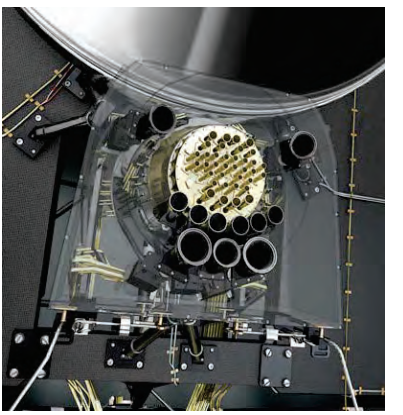


*Photo of the Planck satellite (Courtesy of ESA).*



*Planck focal plane (Courtesy of ESA)*

Planck ESA's mission, was designed to image the temperature and polarization anisotropies of the CMB over the whole sky, with unrivalled angular resolution and sensitivity, pushing the technology to unprecedent limits. In the framework of the highly precision experimental cosmology the legacy Planck results on testing theories of the early universe and the origin of cosmic structure, has provided a major source of information crucial to many cosmological and astrophysical issues. Planck carried out two instruments:

- the High Frequency Instrument (HFI), Principal Investigator: Jean Loup Puget;
- the Low Frequency Instrument (LFI), Principal Investigator: Nazzareno Mandolesi.

The instruments were complementary and using different technology to cross check independently final results and systematic errors. They worked together to produce the overall mission results. The Planck space mission (ESA) has been a wonderful example of Team effort in a large international collaboration, involving scientific, technical and managerial aspects. The unprecedented accuracy of the Planck measurements have established new standards in the determination of fundamental cosmological parameters, as well as new insight in Galactic and extragalactic astrophysics. The Planck full-sky maps in temperature and polarization will remain a lasting legacy for at least dozen years to come. More than 100 papers signed by Planck collaboration have already 30 000 citations in scientific literature. The success of Planck HFI and LFI would not have been possible without the contribution of a large number of talented and dedicated scientists and engineers from many countries of Europe, USA and Canada. HFI was designed to produce high-sensitivity, multi-frequency measurements of the diffuse radiation permeating the sky in all directions in the frequency range of 84 GHz to 1 THz cooled at 100 mK. The instrument consisted of an array of 52 bolometric detectors placed in the focal plane of the telescope. LFI, a microwave instrument, was designed to produce high-sensitivity, multi-frequency measurements of the microwave sky in the frequency range of 27 to 77 GHz. The instrument consisted of an array of 22 tuned radio receivers located in the focal plane of the telescope, cooled at 20 K.

*Jean Loup Puget - PI of the HFI.*

*Nazzareno Mandolesi - PI of the LFI.*

## HANSEN EXPERIMENTAL PHYSICS LABORATORY AT STANFORD UNIVERSITY

presented to **Leo Hollberg**, HEPL Assistant Director

*"to HEPL for having developed interdepartmental activities at Stanford University at the frontier of fundamental physics, astrophysics and technology"*

### Brief History of Stanford's HEPL and Ginzton Laboratories

1947: WW Hansen (right) and team with the Mark I linear accelerator

In 1947, working in the Stanford Physics Department's Microwave Lab, Physics Professor, William W. Hansen and his research team, along with Assistant Professor of Physics and microwave expert, Edward L. Ginzton, completed development on the world's first traveling wave electron linear accelerator. Dubbed the Mark I (see photo) it generated a 1.5 million electron volt (MeV) beam. Hansen's entire report to the U.S. Office of Naval Research (ONR) that funded the project was just one sentence: "We have accelerated electrons."

This successful first step in linear electron acceleration spawned the birth of Stanford's High Energy Physics Lab (HEPL) and Ginzton labs. In 1990, HEPL was renamed the WW Hansen Experimental Physics Lab (also HEPL). HEPL and Ginzton were setup as Stanford's first independent labs. They were organized to facilitate cross-disciplinary research, enabling scientists, engineers, staff and students to work towards common research goals using cutting edge lab equipment and technologies on medium-scale projects. For the past 70 years, the HEPL and Ginzton Labs have spearheaded Stanford's leadership in cross-disciplinary physics and become nurturing homes to a variety of physics-based, research projects: including the following examples:

### Robert Hofstadter's Nobel Prize & Later Work

In 1961, Stanford Professor Robert Hofstadter was awarded the Nobel Prize for his HEPL Mark III Linear Accelerator work on nuclear form factors (nucleons). In the 1980s, Hofstadter became interested in astrophysics and helped design the EGRET telescope in the NASA Compton Gamma Ray Observatory (CGRO).

### Gravity Probe B (GP-B)

2003: Francis Everitt (left) and Brad Parkinson holding a GP-B gyro

In 1959, Physics Department Chair, Leonard Schiff, became interested in using gyroscopes in a satellite to measure the Earth's geodetic effect and the miniscule frame-dragging effect predicted by Albert Einstein's general theory of relativity. Schiff discussed this project with Stanford cryogenic physicist, William Fairbank, and gyroscope expert, Robert Cannon (Aero-Astro department).

In 1962, Fairbank invited post-doc, Francis Everitt, to join the research effort. The team sent a proposal to NASA's Office of Space Sciences requesting funding to develop gyroscopes and a satellite to carry out this unprecedented test. It took 40 years of R&D at Stanford and other places to create and ready the cryogenic satellite and all of its cutting-edge technologies for launch. In 1975, Leonard Schiff moved the GP-B program to HEPL, breathing new life into the lab. In 1981, Francis Everitt became Principal Investigator, a position he still holds. In 1984, Brad Parkinson became Project Manager and a Co-PI, along with Co-PI's John Turneaure and Daniel DeBra.

On 20 April 2004, GP-B launched from Vandenberg AFB into a polar orbit. Data collection began on 28 August 2004 and lasted 50 weeks. Data analysis took five years in order to remove confounding factors in the data. The final results, published in PRL on 31 May 2011, yielded highly accurate geodetic and frame dragging measurements, with 0.28% and 19% margins of error, respectively.

### GPS Spinoffs from GP-B

In the 1990s, Brad Parkinson's research on GPS solutions for positioning the GP-B satellite led to two revolutionary spin-off projects: 1) Wide Area Augmentation System (WAAS) provides highly precise positioning accuracy and itegrity for navigation and the automatic landing of airplanes and 2) Precision Farming adding GPS technology to tractors has enabled the automation of many aspects of farming and has spawned a $1 billion/year industry.

### Fermi Gamma Ray Space Telescope (GLAST)

Peter Michelson, Fermi LAT Principal Investigator

Stanford Physics Professor, Peter Michelson, is a former HEPL Director and the Principal Investigator for the Large Area Telescope (LAT) on board NASA's Fermi Gamma Ray Space Telescope, the successor to CGRO/EGRET. Launched on 11 June 2008, Fermi has been highly successful mapping the gamma-ray sky. Under Michelson's guidance, HEPL's collaborations with Italy are noteworthy. The development of cryogenic bar detectors of gravitational waves, in collaboration with Edoardo Amaldi and his colleagues, established new stringent upper limits to the gravitational waves incident on the Earth. Likewise, the Fermi LAT was developed by a collaboration between Italian INFN and ASI, NASA, and international partners in France, Japan, and Sweden, and used tracking detectors developed, integrated, tested, and qualified for the mission by Italy. GP-B provided the first evidence of frame-dragging on a spinning, superconducting gyroscope. The Fermi detector offers the potential of seeing, through the GeV emission in the Binary Driven Hypernova subclass of long GRBs, the emission from a newly born Black Hole, originating in the induced gravitational collapse of a supernova hypercritically accreting on a binary neutron star companion.

### Robert Byer's LIGO and ACHIP Projects

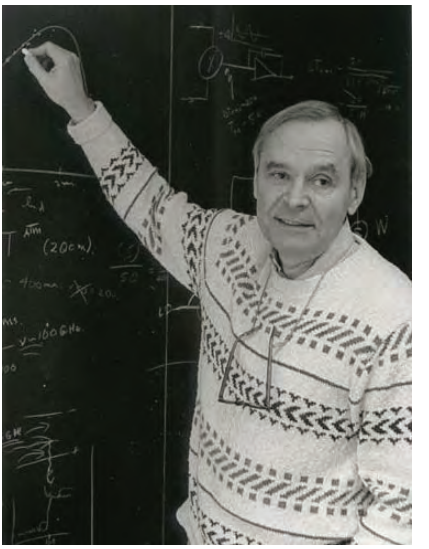
Professor Robert Byer

Robert Byer, former Stanford Dean of Research and former HEPL Director, nurtured the GP-B, GPS and Fermi programs to success during his tenure. He is currently an Applied Physics Professor specializing in lasers and optics. His LIGO Group provided seismic isolation, coatings and materials for the LIGO observatories. His ACHIP project is developing a particle accelerator on a microchip—bringing the HEPL/Ginzton Labs full circle to Hansen's 1947 research.

Professor **LYMAN PAGE**

*"for his collaboration with David Wilkinson in realizing the NASA Explorer WMAP mission and as founding director of the Atacama Cosmology Telescope"*

*Lyman Page*



*David Wilkinson*

This award is given in recognition of Lyman Page's pivotal role in transforming cosmic microwave background observations into a high-precision experimental science over the past two decades. In particular Page provided major contributions to the success of the Wilkinson Microwave Anisotropy Probe (WMAP) space mission, which delivered outstanding measurements of the CMB anisotropy and polarization pattern. He is now continuing his effort by promoting a new generation of experiments like the Atacama Cosmology Telescope to study CMB polarization to greater precision.

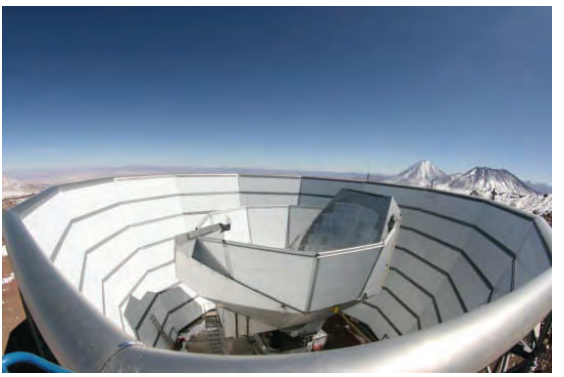


*Photo of the Atacama Cosmology Telescope*

The CMB, the faint afterglow of the Big Bang, is the most powerful probe of the early universe. From its study, we have learned the age of the universe, its major constituents, and have characterized the fundamental fluctuations in gravity that gave rise to cosmic structure. The desire to measure the CMB ever more precisely has driven the development of extraordinary detectors and techniques which will be reviewed in the Lectio Magistralis by Lyman Page. He will describe what we might hope to learn from the CMB in the next decade, including detecting gravitational waves from the birth of the universe if they exist at sufficient amplitude.

Professor **RASHID ALIEVICH SUNYAEV**
*"for the development of theoretical tools in the scrutinising, through the CMB, of the first observable electromagnetic appearance of our Universe"*

*Rashid Sunyaev*

*Rashid Sunyaev and Yakov Borisovich Zeldovich*

Rashid Sunyaev gave extraordinary contributions to the understanding of physical processes in the universe which identified new and uniquely informative targets for observational cosmology. In particular, the Sunyaev-Zeldovich effect, now observed in thousands of clusters of galaxies over the entire sky, has become a cornerstone of cosmology and extragalactic astrophysics, so much so that it is now considered a research field in its own right. Furthermore, Sunyaev's studies of processes in the early universe responsible for angular anisotropy and frequency distortions of the cosmic microwave background have left a profound and lasting legacy for cosmology. In particular, Sunyaev and Zeldovich predicted the presence of acoustic peaks in the CMB angular fluctuation power spectrum and the existence of baryonic acoustic oscillations.
He is currently the project scientist leading the scientific team of the international high-energy astrophysics observatory Spektr-RG being built under the direction of the Russian Space Research Institute.

*Yakov Zeldovich and Remo Ruffini at the audience with Pope John Paul II*

Professor **SHING-TUNG YAU**
*"for the proof of the positivity of total mass in the theory of general relativity and perfecting as well the concept of quasi-local mass, for his proof of the Calabi conjecture, for his continuous inspiring role in the study of black holes physics"*

*Shing-Tung Yau*

Shing-Tung Yau has made fundamental contributions to differential geometry which have influenced a wide range of scientific disciplines, including astronomy and theoretical physics. With Richard Schoen, Yau solved a longstanding question in general relativity by proving that the combined total energy of matter and gravitational field in an asymptotically flat universe is positive. In 1982 Yau was awarded the Fields Medal, the highest award in mathematics, and in 1994 he shared with Simon Donaldson the Crafoord Prize of the Royal Swedish Society in recognition of his development of nonlinear techniques in differential geometry leading to the solution of several outstanding problems.

Another outstanding achievement of Yau is his proof of the Calabi conjecture which allowed physicists to show that string theory is a viable candidate for a unified theory of nature. Furthermore in 2008 Yau (with M.T. Wang) introduced the concept of "quasi-local mass" in general relativity which can be of help to get around the old conundrum — the non-locality of the energy density in relativistic gravity.

During his scientific carrier Yau had more than 50 successful PhD students. At present he is a professor of mathematics at Harvard University where along with research he continues many pedagogical activities. For example, he has created the "Black Hole Initiative", an interdisciplinary center at Harvard University involving a collaboration between principal investigators from the fields of astronomy (Sheperd Doeleman, Avi Loeb and Ramesh Narayan), physics (Andrew Strominger), mathematics (Shing-Tung Yau) and philosophy (Peter Galison). This "Black Hole Initiative" is the first center worldwide to focus on the study of the many facets of black holes.

MG16 in 2021 will mark the $50^{th}$ anniversary of the mass-energy formula for black holes based on the Kerr metric. This timing is an omen that Yau and his school will soon further enlarge our knowledge of this formula with their powerful mathematical analysis.

# PREFACE

The Marcel Grossmann Meetings on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories have always had the goal of bringing together scientists from all over the world enabling them to share recent developments in general relativity and cosmology, paying attention to the interplay between physical predictions and mathematical foundations.

More than 800 scientists met at the University of Rome "la Sapienza" for the Fifteenth Marcel Grossmann Meeting (MG15) during the week of July 1–7, 2018 to discuss theoretical topics and the status of the experimental testing and observations of Einstein's theory of gravitation together with the broad spectrum of gravitational physics related phenomena. The topics discussed ranged from classical and quantum aspects of gravity, relativistic astrophysics, branes and strings, inflation theory, the thermal history of the Universe, to more concrete astrophysics experiments, observations, and modeling, reviewing the current state of the art in theory, observations, and experiments related to relativistic gravitation.

The meeting featured 39 plenary talks during the six mornings of the meeting, plus 6 public lectures given by experts in gravitation and cosmology. 73 parallel sessions, some of which were split over multiple days, kept participants busy with the crucial support of 33 students from Sapienza who managed the coordination of each parallel session and offered technical help.

Three scientists and two collaborations were presented with Marcel Grossmann awards. These were: Lyman Page, Rashid Alievich Sunyaev, Shing-Tung Yau, the Planck Scientific Collaboration (ESA) represented by Jean-Loup Puget, and the Hansen Experimental Physics Laboratory (HEPL) at Stanford University, represented by Leo Hollberg,

The detailed program of plenary and public talks is as follows:

**Shing-Tung Yau** (Harvard University): Quasi-local mass at null infinity
**Malcolm J. Perry** (University of Cambridge): Black hole entropy and soft hair
**Thomas Hertog** (KU Leuven): A smooth exit form eternal inflation
**Jean-Luc Lehners** (Max Planck Institute for Gravitational Physics): No smooth beginning for spacetime
**Ivan Agullo** (Louisiana State University): Loop quantum cosmology and the cosmic microwave background
**Elena Pian** (IASF Bologna):Kilonovae: the cosmic foundries of heavy elements
**Nial Tanvir** (University of Leicester): A new era of gravitational-wave / electromagnetic multi-messenger astronomy
**Tsvi Piran** (Hebrew University of Jerusalem): Mergers and GRBs: past present and future
**Stephan Rosswog** (Stockholm University): Neutron star mergers as heavy element production site

**David Shoemaker** (MIT LIGO Laboratory): LIGO's past and future observations of black hole and neutron star binaries
**Yu Wang** (ICRANet): On the role of binary systems in GW170817/ GRB170817A/AT2017gfo
**Hao Liu** (University of Copenhagen): An independent investigation of gravitational wave data
**Stefano Vitale** (University of Trento): Gravitation wave astronomy in ESA science programme
**Takaaki Kajita** (University of Tokyo): Status of KAGRA and its scientific goals
**Masaki Ando** (University of Tokyo): DECIGO : Gravitational-Wave observations from space
**Jun Luo** (Sun Yat-Sen University): TianQin: a space-borne gravitational wave detector
**Jo Van Den Brand** (Dutch National Institute for Subatomic Physics Nikhef, and VU University Amsterdam): Gravitational wave science and Virgo
**Ernst Maria Rasel** (Leibniz Universität Hannover): Science fiction turns into reality: Interferometry with Bose-Einstein condensates on ground and in space
**Manuel Rodrigues** (Université Paris Saclay): The first results of the MICROSCOPE test of the equivalence principle in space
**Victoria Kaspi** (McGill University): Fast radio bursts
**Bing Zhang** (University of Nevada): From gamma-ray bursts to fast radio bursts: unveiling the mystery of cosmic bursting sources
**Jean-Loup Puget** (CNRS): The Planck mission
**Jorge Armando Rueda Hernandez** (ICRANet): Binary-driven hypernovae and the understanding of gamma-ray bursts
**Remo Ruffini** (ICRANet): The essential role of the nature of the binary progenitors for understanding gamma ray bursts
**Heino Falcke** (Radboud University Nijmegen): Imaging black holes now and in the future
**Luc Blanchet** (Institut d'Astrophysique de Paris): Post-Newtonian theory and gravitational waves
**Jean-Loup Puget** (Université Paris): Frome Planck to Atacama Cosmology Telescope
**Razmik Mirzoyan** (Max Planck Institute for Physics): Gamma-ray and multimessenger highlights with MAGIC
**Elisa Resconi** (Technical University Munich): Neutrino astronomy in the multimessenger era
**Francis Halzen** (University of Wisconsin-Madison): IceCube: opening a new window on the universe from the South Pole
**James Lattimer** (Stony Brook University): The history of R-process
**Ralph Engel** (Karlsruhe Institute of Technology): What have we learned about ultra-high-energy cosmic rays from the Pierre Auger Observatory?

**Paolo De Bernardis** (Sapienza - University of Rome):
**Fabio Gargano** (INFN Bari): DAMPE and its latest results
**Markus Arndt** (University of Vienna): Experiments to probe quantum linearity at the interface to gravity & complexity
**Tobias Westphal** (University of Vienna): Micro-mechanical measurements of weak gravitational forces
**Shu Zhang** (Institute of High Energy Physics, Chinese Academy of Sciences):
**Lorenzo Amati** (INAF - OAS Bologna): Cosmology and multi-messenger astrophysics with Gamma-Ray Bursts
**Elisabetta Cavazzuti** (ASI): Gev LAT observations from GRBs and active galactic nuclei
**Remo Ruffini** (ICRANet): Concluding Remarks
**Roy Kerr** (University of Canterbury): Towards MG16

**Public Lectures**

**Jeremiah Ostriker** (Columbia University): Ultra-light scalars as cosmological dark matter
**Malcolm Longair** (University of Cambridge): Ryle and Hewish: 50 and 100 year anniversaries [Radio astrophysics and the rise of high energy astrophysics]
**Lyman Page** (Princeton University): Measuring the Cosmic Microwave Background
**Marc Henneaux** (Université Libre de Bruxelles): The cosmological singularity
**Anne Archibald** (Newcastle University): Does extreme gravity affect how objects fall?

# Concluding remarks by Prof. Remo Ruffini

For over more than four decades the Marcel Grossmann Meetings have been fostering the interaction between mathematics, relativistic field theories and observations in physics and astrophysics and has witnessed the birth and exponential growth of various new subfields within astrophysics. This has also occurred in 2018 during the MG15 (see Fig. 1), which has seen the participation of approximately one thousand participants from 71 different countries (see Fig. 2). In MG15, as is clear from the above program, we have seen considerable progress in theory on topics ranging from models of quantum gravity to the mathematical structure of Einstein's equations.


Fig. 1. MG15 official Awards Ceremony on Monday, July 2, 2018.

Fig. 2. Group photo of MG15 participants in the Aula Magna, Sapienza University of Rome.

In particular, an MG15 individual award was given to Prof. Rashid Sunyaev, as a member of the relativistic astrophysics school of Zel'dovich-Sunyaev (see Fig. 3) and to Prof. Shing-Tung Yau, as a representative of the Chinese-American school of differential geometry and general relativity (see Fig. 4).

Fig. 3. Rashid Sunyaev receiving the MG15 individual award from Roy Kerr.

Fig. 4. Shing-Tung Yau receiving the MG15 individual award from Roy Kerr.

This has been accompanied by extensive developments in the exploration of the cosmic microwave background with announcements of the final results of the Planck satellite that followed earlier predecessors, like WMAP. In this sense, particularly meaningful have been the MG15

Awards presented to Prof. Jean-Loup Puget, who has successfully accomplished the ESA Plank mission (see Fig. 5) and to Prof. Lyman Page, one of the main participants of the glorious NASA WMAP mission, successfully lead by David Wilkinson (see Fig. 6). I was very fortunate to be in Princeton as an assistant professor at the time WMAP was initially conceived.

Fig. 5. Jean-Loup Puget accepting the MG15 institutional award on behalf of the Planck Scientific Collaboration (ESA) from Roy Kerr.

Fig. 6. Lyman Page accepting the MG15 individual award from Roy Kerr.

X-ray astrophysics gradually successfully evolved from the first rockets of Riccardo Giacconi discovering Scorpius X-1 using well tested Geiger counters, moving on to larger missions using X-ray mirrors in the Einstein Observatory, in XMM and on to the NASA Chandra mission, each new mission introducing new technologies based on previous successes. The

Chandra data on GRB 170817A were amply discussed in the MG15 meeting. Analogously, starting from the earlier gamma ray detectors on the Vela satellite, gamma ray astrophysics reached maturity with the Compton Observatory with the BATSE instruments on board. These two fields joined their separate expertise together in the hybrid Beppo SAX satellite. Equally impressive was the transition from Beppo SAX to the Neil Gehrels Swift Observatory and the Fermi Gamma-Ray Space Telescope that followed EGRET and opened up the field of high energy gamma-ray astrophysics, adding further successes to this ongoing story. One of the most significant contributions to the success of the Fermi mission has been the LAT detector, jointly led by the Hansen Experimental Physics Laboratory (HEPL) at Stanford University, which received the MG15 institutional award, presented to Prof. Leo Hollberg (see Fig. 7).

Fig. 7. Leo Hollberg accepting the MG15 institutional award from Roy Kerr on behalf of the Hansen Experimental Physics Laboratory at Stanford University.

More recently we are witnessing the birth of TeV astrophysics springing from the ground-based MAGIC, HESS and Whipple telescopes, as well as neutrino astrophysics associated with underground (ICE Cube) and underwater laboratories. New tantalizing results concerning the spectrum and composition of cosmic rays have been reported by the Alpha Magnetic Spectrometer experiment on the international space station. All of these

together mark the gradual development of the largest observational effort in history, well recorded in these MG meetings (see Fig. 8). Today relativistic astrophysics is offering us the possibility of utilizing all of these multi-wavelength observations simultaneously in the study of GRBs and active galactic nuclei. Indeed the description of the distinct temporally discrete episodes for the GRB 130427A, one of the most complex astrophysical objects ever observed, requires the simultaneous knowledge of observations from all of these various wavelengths. This exponential growth in all these areas of astrophysics has been accompanied by attempts at data acquisition of gravitational wave signals arriving on the Earth. After the failure of the initial attempt by Misner and Weber in 1972, the new attempt by LIGO-VIRGO was discussed extensively in this MG15 meeting.

Fig. 8. From the left to right: Leo Hollberg, Rashid Sunyaev, Shing-Tung Yau and Remo Ruffini during the MG15 official awards ceremony.

Indeed in addition to the many topics discussed in the plenary lectures, there was a special session dedicated to kilonova and gravitational waves chaired by Enrico Costa on July 3 and a session on gravitational waves chaired by Claus Lämmerzahl on July 4. Many intense discussions took place at the meeting, both in the plenary and parallel sessions, bringing together different points of view, at times controversial, which stimulated a large number of articles published in leading scientific journals in the months following the meetings. The impossibility of finding these discussions in these proceedings is simply explained: it is due to the incomprehensible decision by Sapienza University not to renew the collaboration agreement with ICRANet. What is incomprehensible is that both the research and teaching activities of ICRA and ICRANet are

recognized internationally at a worldwide level, in more than 70 nations. Locks were put on the ICRA Offices (see Fig. 9).

Fig. 9. Remo Ruffini shows to Barry Barish the locked offices of ICRA at Sapienza University.

The preparation of the proceedings was halted. It was not possible to follow up with post-conference interactions and the impossibility of access to our offices and interact with our students. The invited lecturers, clearly indicated in the conference program and some of the world leaders in their fields who had just presented their extremely interesting lectures stimulating widespread interest, mostly chose to publish their results in international journals giving rise to some of the most highly cited articles in recent times. Only a minority were published here. This was also the case for some of the ICRA-ICRANet lectures, which appeared in:

- J.F. Rodriguez, J.A. Rueda, R. Ruffini, *On the final gravitational wave burst from binary black hole mergers*, Astron. Rep. 62, 940 (2018).
- J.A. Rueda, R. Ruffini, Y. Wang, Y. Aimuratov, U. Barres de Almeida, C.L. Bianco, Y.C. Chen, R.V. Lobato, C. Maia, D. Primorac, R. Moradia, and J.F. Rodriguez, *GRB 170817A-GW170817-AT 2017gfo and the observations of NS-NS, NS-WD and WD-WD mergers*, JCAP 10, 006 (2018); and J.A. Rueda, R. Ruffini, Y. Wang, C.L. Bianco, J.M. Blanco-Iglesias, M. Karlica, P. Lorén-Aguilar, R. Moradi and N. Sahakyan, *Electromagnetic emission of white dwarf binary mergers*, JCAP 03, 044 (2019).

- R. Ruffini, R. Moradi, J. A. Rueda, L. Becerra, C. L. Bianco, C. Cherubini, S. Filippi, Y. C. Chen, M. Karlica, N. Sahakyan, Y. Wang, and S. S. Xue, *On the GeV Emission of the Type I BdHN GRB 130427A*, The Astrophysical Journal, 886, 82 (2019).

The dialogue we had started at MG15 undoubtedly had a very stimulating effect on the community which have been followed up and expanded upon in the subsequent MG16 meeting in 2021. Paradoxically, the moments of greatest difficulty for ICRA members have coincided with the flow of new ideas which promoted the steps forward towards the final identification of some fundamental laws of black hole electrodynamics. The very night ICRA offices were locked up and its electronic communications cut, we had submitted a paper to The Astrophysical Journal for publication containing the first fundamental steps for identifying the "inner engine" of GRBs as discussed in talks in this conference. After one year and four months a judge ordered Sapienza to return the ICRA offices.

I am looking forward to a renewed successful collaboration with Sapienza University, enjoyed for more than three decades, to pursue a new common "path" for the rapidly expanding knowledge of the largest distant objects in our Universe. There is still room for improvement: now that the differences have been identified in published articles, a joint effort on an attentive scientific analysis on both sides can lead to the understanding of the nature of these differences and to jointly converge to common solutions. For this to occur, dialogue and not obstruction of office space is needed. As soon as these clarifications will be achieved, we will be able to proceed in a broad scientific effort of common topics of interest in Einstein's theory, in relativistic astrophysics, ranging from the astrophysics of black holes, neutron stars, dark matter and primordial cosmology, observing the rules of basic academic relations.

I am grateful to all members of the IOC and to Carlo Luciano Bianco, Nathalie Deruelle, Rahim Moradi, Tsvi Piran, Jorge Rueda and Narek Sahakyan for advice and discussion.

Remo Ruffini
Chair of the MG International Organizing Committee
Director of ICRANet
President of ICRA

January 10, 2022

This page intentionally left blank

# CONTENTS



## PART A
## PLENARY AND REVIEW TALKS

## PARALLEL SESSIONS

## PART B

## PART C

# PART B

This page intentionally left blank

# Some physics of the kinetic-conformal Hořava theory

Jorge Bellorín[1*], Byron Droguett[1**] and Alvaro Restuccia[1,2***]

[1] *Department of Physics, Universidad de Antofagasta, Antofagasta, 1240000, Chile*

[2] *Department of Physics, Universidad Simón Bolívar, Caracas, 1080-A, Venezuela*

[*] *E-mail: jbellori@gmail.com,* [**] *Email: bjdrogue@gmail.com,*

[***] *E-mail: alvaro.restuccia@uantof.cl*

*Keywords*: Horava theory; Modifications of General Relativity, Newton-Cartan Geometry; Quantum Gravity.

## 1. Definition of the theory and its canonical formulation

With the aim of obtaining simultaneously perturbative renormalizability and unitarity, Hořava[1] has introduced a gravitational theory with higher orders only in spatial derivatives, breaking the symmetry of general diffeomorphisms over the spacetime. Indeed, the very concept of spacetime is substituted by the one of foliation of spacelike hypersurfaces along an absolute line of time. The theory is defined in terms of the ADM (Arnowitt-Deser-Misner) variables $N$, $N_i$ and $g_{ij}$. They are understood as tensors over the spacelike hypersurfaces that evolve in time. The gauge symmetry is given by the FDiff (diffeomorphisms that preserve the foliation), defined on the ADM varaibles. The Lagrangian has a kinetic terms that is of second order in time derivatives,

$$\mathcal{L}_K = \sqrt{g} N G^{ijkl} K_{ij} K_{kl}\,, \quad G^{ijkl} \equiv \frac{1}{2}(g^{ik}g^{jl} + g^{il}g^{jk}) - \lambda g^{ij}g^{kl}\,, \tag{1}$$

and $K_{ij}$ is the extrinsic curvature of the hypersurfaces. This kinetic term is FDiff covariant for any value of the dimensionless coupling constant $\lambda$. Our concern in this paper is a particular formulation of the Hořava theory, which we call the kinetic-conformal theory, given by a critical value of $\lambda$. For spatial hypersurfaces of dimension 3, the critical value for $\lambda$ we refer to is $\lambda = 1/3$. This value defines a dynamically different formulation of the Horava theory, in the sense that the structure of constraints is discontinuous to the generic formulation with $\lambda \neq 1/3$. That is, the constraints of the kinetic-conformal theory cannot be obtained by continuously varying $\lambda$. At $\lambda = 1/3$ the hypermatrix $G^{ijkl}$ becomes degenerated and this leads to the raising of the primary constraint $\pi \equiv g_{ij}\pi^{ij} = 0$, where $\pi^{ij}$ is the canonically conjugate of $g_{ij}$. There is an additional secondary constraint that emerges when the time preservation of $\pi$ is imposed. As a consequence, the kinetic-conformal theory propagates less physical degrees of freedom than the generic formulation of the nonprojectable Hořava theory with $\lambda \neq 1/3$. It propagates two physical modes, the same number of General Relativity. The so-called extra mode of the generic formulation theory is absent in the kinetic-conformal formulation. We consider this an interesting feature that deserves to be explored. Furthermore, at $\lambda = 1/3$ the

kinetic term (1) gets an *anisotropic* conformal symmetry defined by the anisotropic Weyl transformations [1]

$$\tilde{g}_{ij} = \Omega^2 g_{ij}\,, \qquad \tilde{N} = \Omega^3 N\,, \qquad \tilde{N}_i = \Omega^2 N_i\,, \tag{2}$$

where $\Omega = \Omega(t, \vec{x})$. This motivates the name kinetic-conformal. The full theory is not conformally invariant since the potential is not, unles it is defined with specific conformal terms.

In Ref. 2 several versions of the Hořava theory were associated to the algebra of the Newton-Cartan geometry. In summary, depending of the connection used in the gauging and the central extension on the side of the Newton-Cartan algebra, the correspondence is

- Torsionless connection: projectable Hořava theory.
- Twistless-torsional connection: nonprojectable Hořava theory.
- Central Bargmann extension: $U(1)$ extension of the Hořava theory.

Thus, the kinetic-conformal formulation can be found in the twistless-torsional-gauging case, by fixing the values of some coupling constants on the side of the Newton-Cartan dynamics. The extrinsic curvature arises via the covariant derivatives of the inverse vielbein denoted by $v^\mu$ in Ref. 2. Then the kinetic term (1) emerges from the terms that are quadratic in derivatives of $v^\mu$,

$$c_3 \nabla_\mu v^\mu \nabla_\nu v^\nu + c_4 \nabla_\nu v^\mu \nabla_\mu v^\nu = c_4 \left( K_{ij} K^{ij} - \frac{c_3}{c_4} K^2 \right) . \tag{3}$$

Thus, the kinetic-conformal theory is achieved by adjusting $c_3$ and $c_4$ according to

$$\frac{c_3}{c_4} = \frac{1}{3}\,. \tag{4}$$

The action of the nonprojectable Hořava theory is [1,3]

$$S = \int dt d^3x \sqrt{g} N \left( G^{ijkl} K_{ij} K_{kl} - \mathcal{V} \right), \tag{5}$$

where to define the kinetic-conformal theory we consider that the value $\lambda = 1/3$ has been fixed. The potential $\mathcal{V}$ should include all the inequivalent terms that are FDiff covariant and up to sixth order in spatial derivatives ($z = 3$ terms), as required for the power-counting renormalizability [1]. Among them, the terms that contribute to the action of second order in perturbations, hence contributing to the propagator, are [4]

$$-\mathcal{V}^{(z=1)} = \beta R + \alpha a_i a^i\,, \tag{6}$$

$$-\mathcal{V}^{(z=2)} = \alpha_1 R \nabla_i a^i + \alpha_2 \nabla_i a_j \nabla^i a^j + \beta_1 R_{ij} R^{ij} + \beta_2 R^2\,, \tag{7}$$

$$-\mathcal{V}^{(z=3)} = \alpha_3 \nabla^2 R \nabla_i a^i + \alpha_4 \nabla^2 a_i \nabla^2 a^i + \beta_3 \nabla_i R_{jk} \nabla^i R^{jk} + \beta_4 \nabla_i R \nabla^i R\,, \tag{8}$$

where $a_i = \partial_i \ln N$.

In order to determine the dynamical consistency of the theory we have performed its Hamiltonian formulation [5]. The phase space is spanned by the conjugate pairs

$(g_{ij}, \pi^{ij})$ and $(N, P_N)$. The Hamiltonian, with the primary constraints added, is given by

$$H = \int d^3x \left( \frac{N}{\sqrt{g}} \pi^{ij} \pi_{ij} + \sqrt{g} N \mathcal{V} + N_i \mathcal{H}^i + \mu \pi + \sigma P_N \right) , \tag{9}$$

and the full set of constraints is

$$\mathcal{H}^j \equiv -2 \nabla_i \pi^{ij} + P_N \partial^j N = 0 \, . \tag{10}$$

$$P_N = 0 \, , \quad \pi = 0 \, , \tag{11}$$

$$\frac{1}{\sqrt{g}} \mathcal{H} \equiv \frac{1}{g} \pi^{ij} \pi_{ij} - \beta R + 2\alpha \frac{\nabla^2 N}{N} - \alpha a_i a^i = 0 \, , \tag{12}$$

$$\frac{1}{\sqrt{g}} \mathcal{C} \equiv \frac{3}{2g} \pi^{ij} \pi_{ij} + \frac{\beta}{2} R - 2\beta \frac{\nabla^2 N}{N} + \frac{\alpha}{2} a_i a^i = 0 \, . \tag{13}$$

$\mathcal{H}^i$ is a first-class constraint whereas the four constraints $P_N$, $\pi$, $\mathcal{H}$ and $\mathcal{C}$ are of second class. In the Hamiltonian formulation the shift vector $N_i$ plays the role of Lagrange multiplier (as in GR), as well as $\mu$ and $\sigma$. When applying Dirac's programme for extracting the constraints, the time preservation of the second-class constraints leads to equations for $\mu$ and $\sigma$. Considering the complete potential $\mathcal{V}$, it can be shown[5] that these are elliptic (sixth-order) equations, hence they can be consistently solved with appropiated boundary conditions. This ends Dirac's programme, the set of constraints is closed. The constraints $\mathcal{H}$ and $\mathcal{C}$ can also be casted as elliptic equations if the appropiated field variables are chosen to solve them, see[6]. This analysis shows that the Hamiltonian formulation of the theory is consistent. Considering the second-class nature of $P_N$, $\pi$, $\mathcal{H}$ and $\mathcal{C}$, it results that the theory propagates two physical modes, coinciding with the number of modes of GR.

The presence of the $\pi = 0$ constraint is intriguing, since it generates the Weyl scalings on $g_{ij}$ and $\pi^{ij}$, but this theory is not conformally invariant. This is in agreement with the fact that $\pi$ is of second-class, hence it is not the generator of gauge symmetries. Contrasting with an exactly anisotropic conformal Horava theory, which is given by a conformal potential, we have that $\pi = 0$ can be combined with the $P_N = 0$ constraint to form the full generator of the anisotropic conformal transformations (2), which is

$$\varpi = \pi + \frac{3}{2} N P_N \, . \tag{14}$$

In the anisotropic conformal theory this is a first-class constraint. It is preserved without further conditions, hence no furhter constraints are generated. Therefore, in the exact anisotropic conformal case there is a symmetry more than in the kinetic-conformal formulation, the anisotropic Weyl scalings, but a second-class constraint less (the $\mathcal{C}$ constraint). Hence the number of physical degrees of freedom is the same in both cases (at least in the sense of classical field theories, since there is evidence[7] for a conformal anomaly in the anistropic conformal theory), and it is the same of

GR. Further developments in the relationship bewteen the kinetic-conformal theory and the exact anisotropic conformal formulation can be found in[8].

## 2. Quantization: propagators and the superficial degree of divergence

A perturbative analysis considering all the terms (6) - (8) allows to check explicitly the consistency of the Hamiltonian formulation of the theory and to obtain the propagators of the physical modes. We work on the transverse gauge $\partial_i h_{ij} = 0$, where $h_{ij}$ represents the perturbative metric around the flat background. The momentum constraint eliminates the longitudinal sector of the canonical momentum at first order in perturbations. Constraints $\mathcal{H}$ and $\mathcal{C}$ are consistently solved for $h_{kk}$ and the pertubative version of the lapse function, $n$, fixing these variables to zero at first order in perturbations (with asymptotically flat conditions). Constraint $\pi = 0$ elminates the trace of the canonical momentum. There remains the transverse-traceless sector $(h_{ij}^{TT}, p_{ij}^{TT})$ as the independent propagating physical modes. The corresponding propagators are[6]

$$\left\langle h_{ij}^{TT} h_{kl}^{TT} \right\rangle = \frac{P_{ijkl}^{TT}}{\omega^2 - \beta \vec{k}^{\,2} + \beta_1 \vec{k}^{\,4} + \beta_3 \vec{k}^{\,6}}\,, \tag{15}$$

where $P_{ijkl}^{TT}$ is the transverse-traceless projector.

Another important issue of the perturbative quantization is the distribution of Fourier momentum in Feynmann diagrams. We recall that this is a theory with second-class constraints, hence standard techniques of gauge field theories which only have first-class constraints do not apply. One plausible scheme of quantizaton is to solve the second-class constraints perturbatively. It can be shown that the perturbative field variables used to solve the constraints end with a balance of zero Fourier momentum[6]. Therefore when the solutions are substituted in the Lagrangian they do not alter the order in momentum of the interacting terms, the weight in Fourier momentum of the vertices is the same of the off-shell theory. These results allow to evaluate the superficial degree of divergence of Feynman diagrams, since the badly divergent diagrams are those with vertices corresponding to the terms of the Lagrangian of highest order, which is the sixth order. The internal lines scale also with sixth order according to the propagator (15). With these considerations we may show[6] that the superficial degree of divergence of the badly divergent diagrams is given by the order 6. This implies that counter-terms of 6th order in spatial derivatives must be added to the bare Lagrangian, but this is precisely the order of the bare Lagrangian designed for power-counting renormalizability. Thus, the theory passes the criterium given by the superficial degree of divergence needed for the renormalization of the theory.

## 3. Gravitational waves and observational bounds

The coincidence in the number physical degrees of freedom between the kinetic-conformal Hořava theory and GR raises interest on the behavior of the gravitational waves in this theory. This was analyzed in[9] using the equivalence of the large-distance effective action with the Einstein-aether theory[10]. The Einstein-aether theory implements the breaking of the Lorentz invariance keeping the gauge symmetry of general diffeomorphisms over spacetime by using a dynamic unit time-like vector (the aether). The analysis of gravitational waves in the kinetic-conformal Hořava theory done in[9] was achieved in a gauge-invariant way, involving the aether field in the construction of the gauge invariants. The first main result is that the transverse-traceless sector is propagated with a wave equation, $\sqrt{\beta}$ being the speed of the gravitational waves.

For the case of an isolated source, the dominant mode of its gravitational radiation in the far zone can be deduced by applying the same techniques of GR. The considerations on the source are the standard ones for a weak source: small mass, slow velocity and negligible self-gravity. If $I_{ij}$ is the quadrupole moment given by the 00 component of the energy-momentum tensor, then it is found[9] that the leading contribution for the generation of gravitational waves has the same structure of Einstein's quadrupole formula of GR,

$$h_{ij}^{TT} = \frac{\kappa_{\rm H}}{4\pi\beta r} P_{ijkl}^{TT} \frac{d^2 I_{kl}(t - r/\sqrt{\beta})}{dt^2}\,, \tag{16}$$

where $\kappa_{\rm H}$ is the coupling constant arising in front of the Hořava action (which we set equal to one in Eq. (5), since that is a vacuum action). To get an exact matching with the quadrupole formula of GR, we must set the coupling constants $\kappa_{\rm H}$ and $\beta$ equal to their GR values, $\kappa_{\rm H} = 8\pi G_{\rm N}$ and $\beta = 1$.

We may apply the analysis of the parameterized-post-Newtonian (PPN) expansion for solar-system tests to the kinetic-conformal theory. It turns out that the theory reproduces the same values of the PPN parameters of GR, except for the parameters $\alpha_1^{\rm PPN}$ and $\alpha_2^{\rm PPN}$, whose deviations from the zero value indicate Lorentz-symmetry violation. For the kinetic-conformal Hořava theory these two constants are given by[9]

$$\alpha_2^{\rm PPN} = \frac{1}{8}\alpha_1^{\rm PPN} = \beta - 1 - \frac{\alpha}{2}\,. \tag{17}$$

The current observational bounds[12] on these parameters are $|\alpha_1^{\rm PPN}| < 10^{-4}$ and $|\alpha_2^{\rm PPN}| < 10^{-9}$. Relation (17) demands that the strong bound, which is the one on $\alpha_2^{\rm PPN}$, must be satisfied by both parameters. This condition is met if

$$\alpha = 2(\beta - 1) + \delta\,, \tag{18}$$

where $\delta$ represents the narrow observational window for the $\alpha_2^{\rm PPN}$ parameter, i. e. $|\delta| < 10^{-9}$.

**Acknowledgments**

This work was supported by MINEDUC-UA project, Cod. ANT 1855. A. R. is partially supported by grant Fondecyt No. 1161192, Chile.

**References**

1. P. Hořava, *Phys. Rev. D* **79**, 084008 (2009).
2. J. Hartong and N. A. Obers, *JHEP* ***1507***, 155 (2015).
3. D. Blas, O. Pujolàs and S. Sibiryakov, *Phys. Rev. Lett.* **104**, 181302 (2010).
4. M. Colombo, A. E. Gümrükçuoğlu and T. P. Sotiriou, *Phys. Rev. D* **91**, 044021 (2015).
5. J. Bellorín, A. Restuccia and A. Sotomayor, *Phys. Rev. D* **87**, 084020 (2013).
6. J. Bellorín and A. Restuccia, *Phys. Rev. D* **94**, 064041 (2016).
7. I. Adam, I. V. Melnikov and S. Theisen, *JHEP* ***0909***, 130 (2009).
8. J. Bellorín and B. Droguett, *Phys. Rev. D* **98**, 086008 (2018).
9. J. Bellorín and A. Restuccia, *Int. J. Mod. Phys. D* **27**, 1750174 (2017).
10. T. Jacobson and D. Mattingly, *Phys. Rev. D* **64**, 024028 (2001).
11. D. Blas and H. Sanctuary, *Phys. Rev. D* **84**, 064004 (2011).
12. C. M. Will, *Living Rev. Rel.* **17**, 4 (2014).

# String theory and non-relativistic gravity

Eric A. Bergshoeff* and Ceyda Şimşek†

*Van Swinderen Institute, University of Groningen,*
*Nijenborgh 4, 9747 AG Groningen, The Netherlands*
**E-mail: E.A.Bergshoeff@rug.nl, website: www.ericbergshoeff.nl*
†*E-mail: c.simsek@rug.nl*

We identify the proper geometry underlying non-relativistic string theory as 'String Newton-Cartan Gravity'. It has the distinguishing feature that one-dimensional foliation in Newton-Cartan geometry is replaced by a two-dimensional foliation on the target space. We discuss some basic properties of the string Newton-Cartan space-time, in particular it's behaviour under T-duality transformations. This leads to interesting non-relativistic conformal field theories at the boundary that have applications in condensed matter physics.

*Keywords*: Newton-Cartan gravity; Non-relativistic string theory.

## 1. Introduction

Assuming small spacetime curvature, the low-energy limit of superstring theory is described by a supergravity theory. Using this connection our knowledge about supergravity has teached us a lot about some basic properties of superstring theory such as T-duality, D-branes, exotic branes etc. Via it's embedding into string theory, supergravity in an AdS background has an intriguing holographic relation with a conformal field theory in one dimension lower. This latter relationship has also been investigated at the non-relativistic level by considering background geometries in the bulk that have non-relativistic isometries such as Lifshitz or Schrödinger symmetries. This leads to non-relativistic conformal field theories at the boundary that have applications in condensed matter physics.[a]

A different and less well studied corner of non-relativistic holography occurs when the gravity in the bulk itself becomes non-relativistic. It has been argued that the relevant non-relativistic gravity is an extension of Newton-Cartan (NC) gravity, i.e. the frame-independent formulation of Newtonian gravity. This new kind of holography leads to non-relativistic CFTs at the boundary with a different set of symmetries such as the (infinite-dimensional) Galilean conformal symmetries, see Ref. 1. We are especially interested in the possibility that the non-relativist gravity theory in the bulk is connected to a non-relativistic string theory in the same way that general relativity is related to relativistic string theory. This non-relativistic string theory could then perhaps be used as at starting point by itself, independent of the relativistic string theory, to study non-relativistic holography.

Independent of holography, it was realized in the condensed matter community that reparametrization invariance is not only relevant to gravity but also plays

[a]It has been found that the non-relativistic conformal field theories at the boundary couple to a Newton-Cartan geometry with twistless torsion in Ref. 13.

a crucial role in Effective Field Theories (EFTs), where one writes effective Lagrangians in an expansion organized according to the number of derivatives. Coupling an EFT to an arbitrary gravitational background enables one to study generic features of this EFT in an arbitrary frame. The application of non-relativistic gravity in this new context was pioneered by Son and coworkers in Ref. 2. The detailed understanding of how to couple a non-relativistic EFT to a general NC gravitational background can be obtained systematically by starting from a relativistic EFT in a general gravitational background that includes an auxiliary zero-flux vector field as in Ref. 3. By taking a special limit of this EFT in a frame-independent way, one automatically obtains a non-relativistic EFT in a Newton-Cartan background.

In the first part we recall how the kinematics of NC gravity follows from gauging a specific non-relativistic algebra. We extend this analysis to String NC gravity. In the second part we discuss the non-relativistic string theory that naturally couples to the String NC geometry. In particular, we discuss the T-duality properties of this new kind of geometry. At the end we give our conclusions.

## 2. Newton-Cartan Gravity

It is well-known that the symmetries connecting non-relativistic inertial frames are given by the following set of Galilei symmetries:

- time translations: $\delta t = \xi^0$,
- space translations: $\delta x^i = \xi^i$, $\quad i = 1, 2, 3$,
- spatial rotations: $\delta x^i = \lambda^i{}_j \, x^j$,
- Galilean boosts: $\delta x^i = \lambda^i t$.

Here $\xi^0, \xi^i, \lambda^i{}_j$ and $\lambda^i$ are the constant parameters associated with the time translation, space translations, spatial rotations and Galilean boosts, respectively. These symmetries are identical to the Poincaré symmetries except for the Galilean boosts which differ from the Lorentzian boosts in the sense that under Galilean boosts the spatial coordinates $x^i$ transform to the absolute time coordinate $t$, but $t$ does not transform back into the spatial coordinates $x^i$.

It turns out that gauging the Galilei algebra associated with the above symmetries is not sufficient to obtain NC gravity. One way to see this is to realize that the spin-connection fields associated with the spatial rotations and boost symmetries are dependent fields in General Relativity and one expects this to be the same in NC gravity. In General Relativity this is obtained by setting the torsion equal to zero which yields the 24 so-called conventional constraints needed to solve the 24 spin-connection fields associated with the Lorentz transformations. The problem with the Galilei algebra is that the curvature of time translations does not contain a spin-connection field and hence cannot be used for obtaining a conventional constraint. Therefore, there are not enough conventional constraints to solve for all spin-connections. One way to circumvent this issue is to work with a centrally extended Galilei algebra, called the Bargmann algebra. It turns out that

the curvature corresponding to the central charge transformations does contain the boost spin-connection fields. By setting this curvature to zero one acquires the 6 missing conventional constraints. Together with the 18 conventional constraints that follow from setting the curvature corresponding to the spatial translations equal to zero, one is now able to solve for the 12 spatial rotation connection fields $\Omega_\mu{}^{ab}$ ($\mu = 0, 1, 2, 3; a = 1, 2, 3$) and the 12 Galilean boost connection fields $\Omega_\mu{}^a$. The central charge introduced by the Bargmann algebra also has a physical interpretation. When coupled to matter, the U(1) central charge symmetry leads to a conserved Noether charge expressing the conservation of particle number. It also leads to the fact the the Schrödinger equation is necessarily complex providing a projective representation of the Galilei algebra.

We now consider the gauging of the Bargmann algebra, see Ref. 4. Our starting point is the set of commutation relations defining the Bargmann algebra

$$
\begin{aligned}
&[J_{ab}, P_c] = -2\delta_{c[a}P_{b]}\,, && [J_{ab}, G_c] = -2\delta_{c[a}G_{b]}\,,\\
&[G_a, H] = -P_a\,, && [G_a, P_b] = -\delta_{ab}Z\,,
\end{aligned}
\tag{1}
$$

where $\{H, P_a, J_{ab}, G_a, Z\}$ with $a = 1, 2, 3$ are the generators of time translation, space translations, spatial rotations, Galilean boosts and central charge transformation, respectively. In this gauging procedure we associate to every generator/symmetry a gauge field and gauge parameters that are arbitrary functions of spacetime together with the covariant curvatures, see Table 1 below. Note that we have left out the parameters corresponding to the time and space translations. Instead, we assume that all gauge fields transform as covariant vectors under general coordinate transformations with parameters $\xi^\mu(x)$. The two symmetries ($H$ and $P$ transformations versus general coordinate transformations) are related to each other via a so-called trivial or Zilch symmetry in Ref. 5.

Table 1. This table indicates for every symmetry of the Bargmann algebra the corresponding generators, gauge fields, local gauge parameters and covariant curvatures.

| symmetry | generators | gauge field | parameters | curvatures |
|---|---|---|---|---|
| time translation | $H$ | $\tau_\mu$ | – | $R_{\mu\nu}(H)$ |
| space translations | $P_a$ | $E_\mu{}^a$ | – | $R_{\mu\nu}{}^a(P)$ |
| Galilean boosts | $G_a$ | $\Omega_\mu{}^a$ | $\lambda^a(x^\nu)$ | $R_{\mu\nu}{}^a(G)$ |
| spatial rotations | $J_{ab}$ | $\Omega_\mu{}^{ab}$ | $\lambda^{ab}(x^\nu)$ | $R_{\mu\nu}{}^{ab}(J)$ |
| central charge transf. | $Z$ | $M_\mu$ | $\sigma(x^\nu)$ | $R_{\mu\nu}(Z)$ |

One can now show that by imposing the 24 conventional constraints

$$
R_{\mu\nu}{}^a(P) = R_{\mu\nu}(Z) = 0 \tag{2}
$$

the 24 connection fields $\Omega_\mu{}^{ab}$ and $\Omega_\mu{}^a$ can be solved in terms of independent gauge fiels . Furthermore, we impose the 6 geometric constraints

$$R_{\mu\nu}(H) = 2\partial_{[\mu}\tau_{\nu]} = 0 \qquad \rightarrow \qquad \tau_\mu = \partial_\mu \tau \,, \tag{3}$$

defining absolute time. Due to this constraint the time difference $\Delta T$ between two events becomes independent of the path $\mathcal{C}$ that the observer has traveled between these two events:

$$\Delta T = \int_{\mathcal{C}} \mathrm{d}x^\mu \tau_\mu = \int_{\mathcal{C}} \mathrm{d}\tau \,. \tag{4}$$

One thus ends up with three independent gauge-fields $\{\tau_\mu, E_\mu{}^a, M_\mu\}$ that transform under general coordinate transformations, with parameters $\xi^\mu$, as covariant vectors and under spatial rotations, Galilean boosts and central charge transformation as follows:

$$\begin{aligned} \delta\tau_\mu &= \xi^\lambda \partial_\lambda \tau_\mu + \partial_\mu \xi^\lambda \tau_\lambda \,, \\ \delta E_\mu{}^a &= \xi^\lambda \partial_\lambda E_\mu{}^a + \partial_\mu \xi^\lambda E_\lambda{}^a + \lambda^a{}_b \, E_\mu{}^b + \lambda^a \tau_\mu \,, \\ \delta M_\mu &= \xi^\lambda \partial_\lambda M_\mu + \partial_\mu \xi^\lambda M_\lambda + \partial_\mu \sigma + \lambda_a \, E_\mu{}^a \,. \end{aligned} \tag{5}$$

One may define the following Bargmann-invariant metrics

$$\tau_{\mu\nu} = \tau_\mu \tau_\nu \,, \qquad\qquad h^{\mu\nu} = E^\mu{}_a E^\nu{}_b \, \delta^{ab} \,, \tag{6}$$

one in the time direction and a separate one in the spatial directions. One cannot define a boost-invariant metric with upper indices in the time direction or lower indices in the spatial directions. A Galilean boost-invariant metric can be obtained by adding terms proportional to the central charge gauge field $M_\mu$ as follows:

$$H_{\mu\nu} = E_\mu{}^a E_\nu{}^b \delta_{ab} - 2\tau_{(\mu} M_{\nu)} \,. \tag{7}$$

However, in that case one ends up with a metric that is not invariant under the central charge gauge transformation.

Newton-Cartan gravity can be obtained as a non-relativistic limit of general relativity, much in the same way as the Bargmann algebra can be obtained as the Inönü-Wigner contraction of the direct product of the Poincaré algebra and a U(1) generator. The latter has the physical interpretation that, when coupled to matter, its Noether charge corresponds to the conservation of particles minus anti-particles. To describe this particular Inönü-Wigner contraction we consider the Poincaré algebra plus an additional U(1) generator $\hat{\mathcal{Z}}$ that commutes with all the Poincaré generators:

$$\big[\hat{P}_{\hat{A}}, \hat{M}_{\hat{B}\hat{C}}\big] = 2\,\eta_{\hat{A}[\hat{B}}\,\hat{P}_{\hat{C}]} \,, \quad \big[\hat{M}_{\hat{A}\hat{B}}, \hat{M}_{\hat{C}\hat{D}}\big] = 4\,\eta_{[\hat{A}[\hat{C}}\,\hat{M}_{\hat{D}]\hat{B}]} \quad \text{plus} \quad \hat{\mathcal{Z}} \,. \tag{8}$$

Here $\{\hat{P}_{\hat{A}}, \hat{M}_{\hat{A}\hat{B}}\}$ are the generators of spacetime translations and Lorentz transformations, respectively. We have indicated all relativistic generators with a hat to distinguish them from the non-relativistic case. The extra U(1) generator $\hat{\mathcal{Z}}$ is needed because the Bargmann algebra contains one generator more than the Poincaré algebra. Next, we decompose $\hat{A} = (0, a)$ and relate the Poincaré $\otimes$ U(1) generators $\{\hat{P}_0, \hat{P}_a\, \hat{M}_{a0}, \hat{M}_{ab}\}$ and $\hat{\mathcal{Z}}$ to the non-relativistic Bargmann generators $\{H, P_a, G_a, J_{ab}, Z\}$ as follows:

$$\hat{P}_0 = \frac{1}{2\omega}\, H + \omega\, Z\,, \qquad\qquad \hat{\mathcal{Z}} = \frac{1}{2\omega}\, H - \omega\, Z\,, \tag{9}$$

$$\hat{P}_a = P_a\,, \qquad \hat{M}_{ab} = J_{ab}\,, \qquad \hat{M}_{a0} = \omega\, G_a\,, \tag{10}$$

where we have introduced a contraction parameter $\omega$. In a second step, taking the limit $\omega \to \infty$, we obtain the Bargmann algebra including the following commutator containing the central charge generator $Z$:

$$\big[P_A, G_B\big] = \delta_{AB}\, Z\,. \tag{11}$$

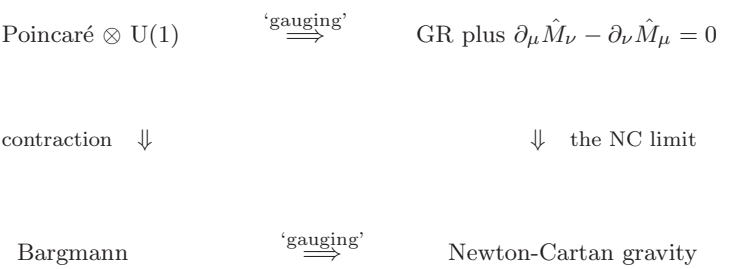



Fig. 1. This figure compares the Inönü-Wigner contraction of the Poincaré algebra times a U(1)with the non-relativistic limit of general relativity plus a zero flux gaugefield $\hat{M}_\mu$ discussed in the text.

The non-relativistic limit of general relativity, which is inspired by the above Inönü-Wigner contraction, is introduced as follows, see Figure 1. Using a second-order formulation of general relativity, we first introduce, on top of the Vierbein field, a vector field $\hat{M}_\mu$ with $\partial_{[\mu}\hat{M}_{\nu]} = 0$, see Ref. 3. Next, we relate the relativistic gauge fields $\{\hat{E}_\mu{}^{\hat{A}}, \hat{M}_\mu\}$ to the non-relativistic gauge fields $\{\tau_\mu, E_\mu{}^a, M_\mu\}$ of NC gravity as follows:

$$\hat{E}_\mu{}^0 = \omega\, \tau_\mu + \frac{1}{2\omega}\, M_\mu\,, \qquad \hat{M}_\mu = \omega\, \tau_\mu - \frac{1}{2\omega}\, M_\mu\,, \qquad \hat{E}_\mu{}^a = E_\mu{}^a\,. \tag{12}$$

In a second step we take the limit $\omega \to \infty$ in the relativistic transformation rules. In this way we obtain the correct non-relativistic transformation rules (5). The

same limit can be used to derive the NC gravity equations of motion from the Einstein equations. Note that the standard textbooks on general relativity usually go straight from General Relativity to Newtonian gravity skipping the general frame formulation of Newtonian gravity.

A striking feature of the non-relativistic limit we just defined is that it requires the introduction of the auxiliary gauge field $\hat{M}_\mu$. Another way to see the necessity of this additional gauge field for taking the limit is considering the action of a particle moving in a general relativistic gravitational background. In order to define a limit without fatal divergencies one needs to couple the particle not only to general relativity but also to the gauge field $\hat{M}_\mu$ via a so-called Wess-Zumino term. This brings us to the remarkable conclusion that the limit of general relativity we just defined is particular to particles. It will not work for other extended objects such as strings. This leads to the following question:

*If NC gravity couples to particles what is the non-relativistic gravity that couples to strings?*

We will answer this question in the next section.

## 3. String Newton-Cartan Gravity

A basic difference between particles and strings is that a particle naturally couples to a 1-form gauge field $\hat{M}_\mu$ whereas a string couples to a 2-form gauge field $\hat{M}_{\mu\nu}$. In order to cancel the leading divergence that occurs in the particle limit one needs to add a Wess-Zumino term to the particle kinetic term with

$$\hat{M}_\mu = \omega\tau_\mu + \cdots \tag{13}$$

In the case of strings the cancellation of the leading divergence requires that the 2-form gauge field $\hat{M}_{\mu\nu}$ starts as $\omega^2$ times something quadratic in $\tau_\mu$ but, due to the anti-symmetry of the 2-form, one cannot write down such an expression. One way to avoid this problem is to introduce *two* Vierbeine ${\tau_\mu}^A$ where $A = 0, 1$ refer to the two (timelike and spacelike) directions longitudinal to the string. One writes

$$\hat{M}_{\mu\nu} = \omega^2 {\tau_\mu}^A {\tau_\nu}^B \epsilon_{AB} + \mathcal{O}(\omega^0)\,. \tag{14}$$

The ${\tau_\mu}^A$ are two generalized clock functions that define a so-called string NC geometry. The basic difference with the particle case is that the relativistic index $\hat{A}$ is divided up differently in longitudinal directions $A$ and transverse directions $a$ as follows:

$$\text{particle:} \quad \hat{A} = (0, a) \;\text{ with }\; a = 1, \cdots, d \tag{15}$$

$$\text{string:} \quad \hat{A} = (A, a) \;\text{ with }\; A = 0, 1 \;\text{ and }\; a = 2, \cdots, d-1\,. \tag{16}$$

The string Galilei symmetries underlying this geometry are given by

| | |
|---|---|
| longitudinal translations | $H_A$ |
| transverse translations | $P_a$ |
| string Galilei boosts | $G_{Ab}$ |
| longitudinal Lorentz rotations | $M_{AB}$ |
| transverse spatial rotations | $J_{ab}$ |

In gauging the algebra corresponding to these symmetries one again encounters extensions that in this case are non-central:

$$[G_{Aa}, P_b] = 0 \qquad \rightarrow \qquad [G_{Aa}, P_b] = \delta_{ab} Z_A \tag{17}$$

and leads to non-central charge gauge fields $M_\mu{}^A$. To obtain irreducibility one needs to impose a set of conventional and geometric constraints. A difference with the particle case is that the zero torsion constraint now becomes a mixture of conventional and geometric constraint:

$$R_{\mu\nu}{}^A(H) = D_{[\mu}(\Omega)\tau_{\nu]}{}^A = 0. \tag{18}$$

For instance, in 4 spacetime dimensions, i.e. $d = 3$, the 12 curvature constraints (18) is a mixture of 4 conventional constraint that can be used to solve for the 4 components of longitudinal spin-connection fields

$$\Omega_\mu{}^{AB} = \epsilon^{AB}\Sigma_\mu \tag{19}$$

and 8 geometric constraints.

After gauging, one ends up with the independent string NC gauge fields $\{\tau_\mu{}^A, E_\mu{}^a, m_\mu{}^A\}$ that transform under the different symmetries as follows:

$$\begin{aligned} \delta\tau_\mu{}^A &= \Lambda^A{}_B \tau_\mu{}^B\,, \\ \delta E_\mu{}^a &= \Lambda^a{}_b\, E_\mu{}^b - \Sigma_A{}^a \tau_\mu{}^A\,, \\ \delta m_\mu{}^A &= D_\mu \sigma^A + \Sigma^A{}_a\, E_\mu{}^a\,. \end{aligned} \tag{20}$$

Furthermore, all gauge fields transform as covariant vectors under general coordinate transformations. These transformations rules are the generalization of the particle transformation rules (5) to strings and will play a prominent role as the symmetries on the non-relativistic Polyakov string we discuss in the next section.

We note that the string analogue of the longitudinal metric (6) and boost invariant metric (7) are given by

$$\text{longitudinal metric}: \qquad \tau_{\mu\nu} \equiv \tau_\mu{}^A \tau_\nu{}^B \eta_{AB}\,, \tag{21}$$

$$\text{boost invariant metric}: \qquad H_{\mu\nu} \equiv E_\mu{}^a E_\nu{}^b \delta_{ab} + \left(\tau_\mu{}^A m_\nu{}^B + \tau_\nu{}^A m_\mu{}^B\right)\eta_{AB}\,. \tag{22}$$

Like in the particle case, $H_{\mu\nu}$ is invariant under string Galilei boost transformations but transforms under the non-central charge gauge transformations with parameter $\sigma^A$.

## 4. The Non-relativistic Polyakov String

To understand the non-relativistic Polyakov string as the limit of the relativistic Polyakov string, it is instructive to first consider the particle case. Our starting point is a particle in a Polyakov formulation moving in the fields of general relativity and coupled to the extra gauge field $\hat{M}_\mu$ via a Wess-Zumino (WZ) term:

$$S_{\text{Pol.}} = -\tfrac{1}{2}\int d\tau\Big\{ -\frac{1}{e}\hat{E}_\mu{}^{\hat{A}}\dot{x}^\mu\hat{E}_\nu{}^{\hat{B}}\dot{x}^\nu\eta_{\hat{A}\hat{B}} + M^2 e - 2M\hat{M}_\mu\dot{x}^\mu\Big\}. \tag{23}$$

Here $e$ is the Einbein and $M$ a mass parameter. Substituing the parametrization (12) we obtain a potential leading divergence of order $c^2$:

$$S_{\text{Pol.}}(c^2) = -\tfrac{1}{2}c^2\int d\tau\frac{1}{e}\big[\tau_\mu\dot{x}^\mu - me\big]^2. \tag{24}$$

To circumvent this potential divergence, we introduce by hand a Lagrange multiplier $\lambda$ and rewrite the above term in the following equivalent form:

$$S_{\text{Pol.}}(c^2) = -\tfrac{1}{2}\int d\tau\frac{1}{e}\Big\{\lambda(\tau_\mu\dot{x}^\mu - me) - \frac{1}{4c^2}\lambda^2\Big\}. \tag{25}$$

Solving for $\lambda$ by using it's equation of motion one recovers the Lagrangian (24). The advantage of the Lagrangian (32) is that the quadratic divergence has disappeared. Note that this rewriting can only be done after making the expansion (12). Strictly speaking we can not consider the result we obtain as the limit of the original Polyakov Lagrangian (23).

Continuing in this way, we obtain the following Lagrangian for the non-relativistic Poyakov particle:

$$S_{\text{Pol.}}(\text{N.R.}) = -\tfrac{1}{2}\int d\tau\frac{1}{e}\Big\{\dot{x}^\mu\dot{x}^\nu H_{\mu\nu} + \lambda\big(\tau_\mu\dot{x}^\mu - me\big)\Big\}. \tag{26}$$

Note that the boost invariant metric $H_{\mu\nu}$ defined in eq. (7) occurs in the Lagrangian. The fact that this metric is not invariant under the central charge transformation leads to a total derivative in the Lagrangian. Solving for $e$ by using the equation of motion for $\lambda$ leads to the Lagrangian for the non-relativistic Nambu-Goto particle. Note that the resulting Nambu-Goto Lagrangian is non-linear in the longitudinal time coordinate $\tau_\mu\dot{x}^\mu$ but quadratic in the transverse coordinates that are contracted with the boost invariant metric $H_{\mu\nu}$. This is in contrast to the Lagrangian for the relativistic Nambu-Goto particle which is non-linear in *all* directions.

The Lagrangian (26) for the non-relativistic Polyakov particle has all the features of the non-relativistic Polyakov string. To derive the action for the non-relativistic

Polyakov string, the relativistic gauge fields are redefined as follows;

$$\hat{E}_\mu{}^A = \omega\tau_\mu{}^A + \frac{1}{\omega}\, m_\mu{}^A\,, \qquad\qquad \hat{E}_\mu{}^a = E_\mu{}^a\,, \tag{27}$$

$$\hat{M}_{\mu\nu} = \omega^2 \tau_\mu{}^A \tau_\nu{}^B\, \epsilon_{AB} + B_{\mu\nu} \tag{28}$$

and two Lagrange multipliers $\lambda$ and $\bar{\lambda}$ are introduced to avoid divergences when the limit is taken. We have used here the following definitions. First of all, we have defined a worldsheet metric $h_{\alpha\beta}$ that can be expressed in terms of worldsheet Zweibeine as follows:

$$h_{\alpha\beta} = e_\alpha{}^a e_\beta{}^b \eta_{ab}\,, \tag{29}$$

with $\alpha = 0, 1 (a = 0, 1)$ curved (flat) worldsheet indices.[b] Furthermore, for $e_\mu{}^a$ and $\tau_\mu{}^A$ we use the following lightcone notation:

$$e_\alpha \equiv e_\alpha{}^0 + e_\alpha{}^1\,, \qquad \bar{e}_\alpha \equiv e_\alpha{}^0 - e_\alpha{}^1\,, \tag{30}$$

$$\tau_\mu \equiv \tau_\mu{}^0 + \tau_\mu{}^1\,, \qquad \bar{\tau}_\mu \equiv \tau_\mu{}^0 - \tau_\mu{}^1\,. \tag{31}$$

On top of definitions above, by taking $\omega \to \infty$ limit of relativistic Polyakov action of a string that is coupled to both the target space metric $\hat{G}_{\mu\nu}$ and 2-form gauge field $\hat{M}_{\mu\nu}$, the non-relativistic Polyakov string action

$$S_{\text{Pol.}} = -\frac{T}{2}\int d^2\sigma \left[\sqrt{-h}\, h^{\alpha\beta}\, \partial_\alpha x^\mu \partial_\beta x^\nu H_{\mu\nu} + \epsilon^{\alpha\beta}\left(\lambda\, e_\alpha \tau_\mu + \bar{\lambda}\, \bar{e}_\alpha \bar{\tau}_\mu\right) \partial_\beta x^\mu\right]$$

$$- \frac{T}{2}\int d^2\sigma\, \epsilon^{\alpha\beta} \partial_\alpha x^\mu \partial_\beta x^\nu B_{\mu\nu} \tag{32}$$

is obtained. Using the equations of motion of the Lagrange multipliers $\lambda$ and $\bar{\lambda}$ one is able to solve for the worldsheet metric $h_{\mu\nu}$ in terms of the longitudinal metric $\tau_{\mu\nu}$ up to a scale factor:

$$h_{\alpha\beta} \sim \partial_\alpha x^\mu \partial_\beta x^\nu \tau_{\mu\nu}\,. \tag{33}$$

Substituting this solution back into the non-relativistic Polyakov Lagrangian (32) one obtains the Nambu-Goto formulation of the non-relativistic Polyakov string.

The non-relativistic Polyakov string defined by the Lagrangian (32) is a central result. The nice thing is that, although it was obtained as a kind of limit of the relativistic Polyakov string, one could use the Lagrangian (32) as an independent starting point to investigate the properties of non-relativitic string theory. Imposing the worldsheet conformal gauge

$$\sqrt{-h}\, h^{\alpha\beta} = \eta^{\alpha\beta} \tag{34}$$

and taking the flat spacetime background

$$\tau_\mu{}^A = \delta_\mu^A\,, \qquad E_\mu{}^a = \delta_\mu^a\,, \qquad M_\mu{}^a = 0 \tag{35}$$

[b]Note that we use the index $a$ both for flat worldsheet indices as well as for flat transverse target space indices. Whenever confusion could occur we will specify which index we mean.

one recovers the Lagrangian of the non-relativistic string theory of Ref. 6

$$S = -\frac{T}{2}\int d^2\sigma \left(\partial x^a\, \bar{\partial} x^b \delta_{ab} + \lambda\, \bar{\partial} X + \bar{\lambda}\, \partial \bar{X}\right) . \tag{36}$$

where

$$X = x_0 + x_1 \qquad \bar{X} = x_0 - x_1, \tag{37}$$

$$\partial = \partial_0 + \partial_1 \qquad \bar{\partial} = -\partial_0 + \partial_1. \tag{38}$$

We have generalized this Lagrangian to a string moving in an arbitrary non-relativistic gravitational background and, moreover, we have identified the geometry underlying the non-relativistic string as the string NC geometry discussed in section 2. We could now use this relation to study some of the properties of string NC geometry in the same way as we are doing in the relativistic case. One of these peculiar properties is the behaviour of the geometry under the so-called T-duality tranformations. In the next section we will derive the non-relativistic T-duality properties using the non-relativistic Polyakov Lagrangian (32) as our starting point. The non-relativistic limit introduced above is also applied to the Eintein-Hilbert action plus a term contains an auxiliary two-form and one-form gauge field and it produces the extended string Newton-Cartan gravity in Ref. 11.

## 5. T-duality

Before discussing the non-relativistic case, it is instructive to first go shortly over the relativistic case. Following Refs. 8, 9, 10, we assume that the geometry has a spacelike Killing vector $k^\mu$. Using adapted coordinates with isometry direction $y$

$$x^\mu = (y, x^i) \qquad\qquad k^\mu \partial_\mu = \partial_y \tag{39}$$

we rewrite the non-relativistic Polyakov Lagrangian (32) as the following parent action:

$$S_{\text{parent}} = \underbrace{S_{\text{Pol.}}(\partial_\alpha y \to v_\alpha)}_{\text{quadratic in } v_\alpha!} - T\int d^2\sigma \epsilon^{\alpha\beta}\tilde{y}\partial_\alpha v_\beta\,, \tag{40}$$

where we have replaced everywhere $\partial_\alpha y$ by a new field $v_\alpha$ and, moreover, have introduced a new Lagrange multiplier field $\tilde{y}$ who in a minute is going to play the role of the dual coordinate. The important point is that the Polyakov Lagrangian (32) is quadratic in $\partial y$ and, hence, the parent action as given in (40) is quadratic in $v_\alpha$.

Starting from the parent action (40) we can follow two approaches. One, we impose the equation of motion corresponding to the Lagrange multiplier $\tilde{y}$ leading to the Bianchi identiy for $v_\alpha$ whose solution is given by

$$v_\alpha = \partial_\alpha y\,. \tag{41}$$

Substituting this solution back into the action corresponding to the Polyakov Lagrangian (32) one re-obtains the Polyakov action we started from. Instead of doing

this, we can also solve for the equation of motion of $v_\alpha$. Since the first term in the parent action (40) is quadratic in $v_\alpha$ this leads to the following solution for $v_\alpha$:

$$v_\alpha = \epsilon_\alpha{}^\beta \partial_\beta \tilde{y}\,. \tag{42}$$

Substituting this solution back into the parent action (40) leads in the relativistic case to a Polyakov string moving in a dual geometric background. For instance, taking the ten-dimensional background to be a nine-dimensional Minkowski spacetime times a circle with radius $R$, after the duality transformation one ends up with the same geometry but now with R replaced by $1/R$:

$$R \quad \stackrel{T}{\to} \quad 1/R\,. \tag{43}$$

This is a special case of the so-called Buscher rules obtaned in Refs. 8, 9, 10.

We now wish to perform the same manipulations for the non-relativistic Polyakov string and derive the non-relativistic version of the Buscher rules. First of all, it matters whether we dualise in the longitudinal or transverse direction. This has to do with the fact that the term with the extra Lagrange multipliers $\lambda$ and $\bar{\lambda}$ only depends on the longitudinal embedding coordinates. These extra terms, which are absent in the relativistic case, can therefore only effect the longitudinal T-duality transformations. Following Ref. 7 we find the following results:

(1) The *longitudinal* spatial T-dual of the NR string is the Polyakov string moving in a general relativity background with a lightlike direction.

(2) The *transverse* spatial T-dual of the NR string is again a NR string but with a transverse spatial isometry direction given by the Buscher rules, i.e. a transverse circle with radius $R$ becomes a transverse circle with radius $1/R$.

## 6. Conclusions

We showed that the geometry underlying non-relativistic string theory is given by the string NC geometry as discussed in section 2. In section 4 we derived the action describing a non-relativistic Polyakov string moving in this string NC background. This led to the action (32) which should be the starting point for every effort to study non-relativistic string theory. As an example, we showed in this talk how the action (32) could be used to obtain the non-relativistic T-duality rules which gives information on how a non-relativistic string feels it's own geometry.

There are many other things one can study now, like, e.g., the calculation of the non-relativistic beta functions which would determine the consistent backgrounds in which the non-relativistic string can move. Our hope is that at the end we may use non-relativistic string theory as defined by the Lagrangian (32) to define non-relativistic holography and in this way learn something about the (non-perturbative) properties of non-relativistic conformal quantum field theories with (infinite-dimensional) Galilean conformal symmetries that could not be studied using general relativity in the bulk.

**Acknowledgements**

We thank Jaume Gomis, K. T. Grosvenor, Jan Rosseel, Thomas Zojer for the many stimulating discussions we had with them. We also thank the organizers for a very stimulating atmosphere.

**References**

1. A. Bagchi and R. Gopakumar, *Galilean Conformal Algebras and AdS/CFT, JHEP* **0907** (2009) 037 doi:10.1088/1126-6708/2009/07/037 [arXiv:0902.1385 [hep-th]].
2. D. T. Son and M. Wingate, *General coordinate invariance and conformal invariance in nonrelativistic physics: Unitary Fermi gas, Annals Phys.* **321** (2006) 197 doi:10.1016/j.aop.2005.11.001; C. Hoyos and D. T. Son, *Hall Viscosity and Electromagnetic Response, Phys. Rev. Lett.* **108**, 066805 (2012) doi:10.1103/PhysRevLett.108.066805; D. T. Son, *Newton-Cartan Geometry and the Quantum Hall Effect*, [arXiv:1306.0638 [cond-mat.mes-hall]].
3. E. Bergshoeff, J. Rosseel and T. Zojer, *Non-relativistic fields from arbitrary contracting backgrounds, Class. Quant. Grav.* **33** (2016) no.17, 175010 doi:10.1088/0264-9381/33/17/175010 [arXiv:1512.06064 [hep-th]].
4. R. Andringa, E. Bergshoeff, S. Panda and M. de Roo, *Newtonian Gravity and the Bargmann Algebra, Class. Quant. Grav.* **28** (2011) 105011 doi:10.1088/0264-9381/28/10/105011 [arXiv:1011.1145 [hep-th]].
5. D. Z. Freedman and A. Van Proeyen, *Supergravity,* Cambridge University Press (2012).
6. J. Gomis and H. Ooguri, *Nonrelativistic closed string theory, J. Math. Phys.* **42** (2001) 3127 doi:10.1063/1.1372697 [hep-th/0009181].
7. E. Bergshoeff, J. Gomis and Z. Yan, *Nonrelativistic String Theory and T-Duality, JHEP* **1811** (2018) 133 doi:10.1007/JHEP11(2018)133 [arXiv:1806.06071 [hep-th]].
8. T. H. Buscher, *A Symmetry of the String Background Field Equations, Phys. Lett. B* **194** (1987) 59. doi:10.1016/0370-2693(87)90769-6
9. T. H. Buscher, *Path Integral Derivation of Quantum Duality in Nonlinear Sigma Models, Phys. Lett. B* **201** (1988) 466. doi:10.1016/0370-2693(88)90602-8
10. M. Rocek and E. P. Verlinde, *Duality, quotients, and currents, Nucl. Phys. B* **373** (1992) 630 doi:10.1016/0550-3213(92)90269-H [hep-th/9110053].
11. E. A. Bergshoeff, K. T. Grosvenor, C. Simsek and Z. Yan, *An Action for Extended String Newton-Cartan Gravity, JHEP* **1901** (2019) 178 doi:10.1007/JHEP01(2019)178 [arXiv:1810.09387 [hep-th]].
12. E. Bergshoeff, J. Rosseel and T. Zojer, *Newton-Cartan (super)gravity as a non-relativistic limit, Class. Quant. Grav.* **32** (2015) no.20, 205003 doi:10.1088/0264-9381/32/20/205003 [arXiv:1505.02095 [hep-th]].

13. M. H. Christensen, J. Hartong, N. A. Obers and B. Rollier, *Torsional Newton-Cartan Geometry and Lifshitz Holography*, *Phys. Rev. D* **89** (2014) 061901 doi:10.1103/PhysRevD.89.061901 [arXiv:1311.4794 [hep-th]].
14. R. Andringa, E. Bergshoeff, J. Gomis and M. de Roo, *'Stringy' Newton-Cartan Gravity, Class. Quant. Grav.* **29** (2012) 235020 doi:10.1088/0264-9381/29/23/235020 [arXiv:1206.5176 [hep-th]].

# Non-relativistic expansion of the Einstein–Hilbert Lagrangian

Dennis Hansen[1], Jelle Hartong[2*] and Niels A. Obers[3,4]

[1] *Institut für Theoretische Physik, Eidgenössische Technische Hochschule Zürich Wolfgang-Pauli-Strasse 27, 8093 Zürich, Switzerland*
[2] *School of Mathematics and Maxwell Institute for Mathematical Sciences, University of Edinburgh, Peter Guthrie Tait road, Edinburgh EH9 3FD, UK*
[3] *Nordita, KTH Royal Institute of Technology and Stockholm University, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden*
[4] *The Niels Bohr Institute, Copenhagen University, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark*
[*] *E-mail: j.hartong@ed.ac.uk*

We present a systematic technique to expand the Einstein–Hilbert Lagrangian in inverse powers of the speed of light squared. The corresponding result for the non-relativistic gravity Lagrangian is given up to next-to-next-to-leading order. The techniques are universal and can be used to expand any Lagrangian theory whose fields are a function of a given parameter.

*Keywords*: Newton–Cartan geometry, non-relativistic gravity, post-Newtonian expansion.

## 1. Introduction

Recent developments have significantly improved our understanding of non-relativistic geometries such as Newton–Cartan geometry and their role in approximations and limits of relativistic theories such as general relativity, Chern–Simon theories and string theories. We will show how Newton–Cartan geometry and its torsionful generalization can be used to construct an order by order non-relativistic approximation of the Einstein–Hilbert Lagrangian. This will provide us with a non-relativistic gravity Lagrangian whose equations of motion describe Newtonian gravity and generalizations thereof that involve e.g. gravitational time dilation.

## 2. Expanding fields and Lagrangians

Consider a Lagrangian for a set of fields $\phi^I$ (with spacetime indices suppressed) that is a function of $c, \phi^I, \partial_\mu \phi^I$ where $\phi^I = \phi^I_{(0)} + c^{-2}\phi^I_{(1)} + \cdots$ with $c$ the speed of light. The explicit $c$ dependence is determined by the Lagrangian and is related to our choice of working with a set of fields whose $1/c^2$ expansion starts at order $c^0$. We will only consider even powers series in $c^{-1}$. Assuming the overall power of the Lagrangian is $c^N$ we define $\tilde{\mathcal{L}}(\sigma) = c^{-N}\mathcal{L}(c, \phi^I, \partial_\mu \phi^I)$ where $\sigma = c^{-2}$. We then Taylor expand $\tilde{\mathcal{L}}(\sigma)$ around $\sigma = 0$, i.e.

$$\tilde{\mathcal{L}}(\sigma) = \tilde{\mathcal{L}}(0) + \sigma \tilde{\mathcal{L}}'(0) + \frac{1}{2}\sigma^2 \tilde{\mathcal{L}}''(0) + O(\sigma^3)\,, \tag{1}$$

where the prime denotes differentiation with respect to $\sigma$, which satisfies

$$\frac{d}{d\sigma} = \frac{\partial}{\partial \sigma} + \frac{\partial \phi^I}{\partial \sigma}\frac{\partial}{\partial \phi^I} + \frac{\partial \partial_\mu \phi^I}{\partial \sigma}\frac{\partial}{\partial \partial_\mu \phi^I}\,. \tag{2}$$

Hence if we write

$$\mathcal{L} = c^N \tilde{\mathcal{L}}(\sigma) = c^N \left( \mathcal{L}_{\rm LO} + c^{-2} \mathcal{L}_{\rm NLO} + c^{-4} \mathcal{L}_{\rm NNLO} + O(c^{-6}) \right) , \tag{3}$$

it follows that the LO and NLO Lagrangians take the form

$$\mathcal{L}_{\rm LO} = \tilde{\mathcal{L}}(0) = \mathcal{L}_{\rm LO}(\phi^I_{(0)}, \partial_\mu \phi^I_{(0)}) , \tag{4}$$

$$\mathcal{L}_{\rm NLO} = \tilde{\mathcal{L}}'(0) = \frac{\partial \tilde{\mathcal{L}}}{\partial \sigma}\Big|_{\sigma=0} + \phi^I_{(1)} \frac{\delta \mathcal{L}_{\rm LO}}{\delta \phi^I_{(0)}} . \tag{5}$$

This shows that the equations of motion of the next-to-leading order (NLO) fields of the next-to-leading order Lagrangian are the equations of motion of the leading order (LO) fields of the leading order Lagrangian. A very similar calculation gives for the NNLO Lagrangian the expression

$$\begin{aligned} \mathcal{L}_{\rm NNLO} = {} & \frac{1}{2} \frac{\partial^2 \tilde{\mathcal{L}}}{\partial \sigma^2}\Big|_{\sigma=0} + \phi^I_{(1)} \frac{\delta}{\delta \phi^I_{(0)}} \frac{\partial \tilde{\mathcal{L}}}{\partial \sigma}\Big|_{\sigma=0} + \frac{1}{2} \phi^I_{(2)} \frac{\delta \mathcal{L}_{\rm LO}}{\delta \phi^I_{(0)}} + \frac{1}{2} \Bigg[ \phi^I_{(1)} \phi^J_{(1)} \frac{\partial^2 \mathcal{L}_{\rm LO}}{\partial \phi^I_{(0)} \partial \phi^J_{(0)}} \\ & + \left( \phi^I_{(1)} \partial_\mu \phi^J_{(1)} + \phi^J_{(1)} \partial_\mu \phi^I_{(1)} \right) \frac{\partial^2 \mathcal{L}_{\rm LO}}{\partial \phi^I_{(0)} \partial (\partial_\mu \phi^J_{(0)})} + \partial_\mu \phi^I_{(1)} \partial_\nu \phi^J_{(1)} \frac{\partial^2 \mathcal{L}_{\rm LO}}{\partial (\partial_\mu \phi^I_{(0)}) \partial (\partial_\nu \phi^J_{(0)})} \Bigg] . \end{aligned} \tag{6}$$

The term in square brackets is the second variation of the LO Lagrangian and is a quadratic form involving the Hessian of the LO Lagrangian. It can be shown that

$$\frac{\delta \mathcal{L}_{\rm NNLO}}{\delta \phi^I_{(1)}} = \frac{\delta \mathcal{L}_{\rm NLO}}{\delta \phi^I_{(0)}} . \tag{7}$$

Combining this with the fact that the equations of motion of $\phi^I_{(2)}$ of the NNLO Lagrangian give the equations of motion of the LO Lagrangian we see that the NNLO Lagrangian reproduces all of the equations of motion of the NLO Lagrangian.

## 3. Explicit dependence on speed of light of Einstein–Hilbert Lagrangian

In order to apply these ideas to the expansion of the Einstein–Hilbert (EH) Lagrangian we need to construct fields that start at order $c^0$. To this end we define

$$g_{\mu\nu} = -c^2 T_\mu T_\nu + \Pi_{\mu\nu} \qquad \text{and} \qquad g^{\mu\nu} = -\frac{1}{c^2} T^\mu T^\nu + \Pi^{\mu\nu} , \tag{8}$$

where $c^2 = \hat{c}^2/\sigma$ with $\hat{c}$ equal to the speed of light. The expansion is in $\sigma$ around zero which is the square of the dimensionless slope of the light cone in tangent space. We will set $\hat{c} = 1$ and expand around $1/c^2$. We can always impose the conditions

$$T_\mu \Pi^{\mu\nu} = 0 , \qquad T^\mu \Pi_{\mu\nu} = 0 , \qquad T_\mu T^\mu = -1 , \qquad \Pi_{\mu\rho} \Pi^{\rho\nu} = \delta^\nu_\mu + T^\nu T_\mu . \tag{9}$$

It will be convenient to use the following torsionful connection (instead of the usual Levi-Civita connection)

$$C^\rho_{\mu\nu} = -T^\rho \partial_\mu T_\nu + \frac{1}{2} \Pi^{\rho\sigma} \left( \partial_\mu \Pi_{\nu\sigma} + \partial_\nu \Pi_{\mu\sigma} - \partial_\sigma \Pi_{\mu\nu} \right) . \tag{10}$$

The covariant derivative $\overset{(C)}{\nabla}_\mu$ with respect to this connection obeys

$$\overset{(C)}{\nabla}_\mu T_\nu = 0\,, \qquad \overset{(C)}{\nabla}_\mu \Pi^{\nu\rho} = 0\,, \qquad \overset{(C)}{\nabla}_\mu T^\nu = \frac{1}{2}\Pi^{\nu\rho}\mathcal{L}_T \Pi_{\rho\mu}\,, \tag{11}$$

where $\mathcal{L}_T$ is the Lie derivative along $T^\mu$. The EH Lagrangian can be written as

$$\mathcal{L}_{\rm EH} = \frac{c^3}{16\pi G}\sqrt{-g}R = \frac{c^6}{16\pi G}\tilde{\mathcal{L}}(\sigma, T, \Pi, \partial)\,, \tag{12}$$

where $G$ is Newton's constant. The prefactor $c^3$, as opposed to the usual $c^4$, is due to the presence of $c$ in (8). The Lagrangian $\tilde{\mathcal{L}}$ which is at most order $c^0$ is

$$\tilde{\mathcal{L}} = \sqrt{-\det\left(-T_\alpha T_\beta + \Pi_{\alpha\beta}\right)}\left[\frac{1}{4}\Pi^{\mu\nu}\Pi^{\rho\sigma}T_{\mu\rho}T_{\nu\sigma} + \sigma\Pi^{\mu\nu}\overset{(C)}{R}_{\mu\nu} - \sigma^2 T^\mu T^\nu \overset{(C)}{R}_{\mu\nu}\right]. \tag{13}$$

In here $\overset{(C)}{R}_{\mu\nu}$ is defined with respect to the connection $C^\rho_{\mu\nu}$ and $T_{\mu\nu} \equiv \partial_\mu T_\nu - \partial_\nu T_\mu$.

## 4. Non-relativistic expansion of the metric

The fields $T_\mu$ and $\Pi_{\mu\nu}$, entering the metric, admit the following $1/c^2$ expansion

$$T_\mu = \tau_\mu + c^{-2}m_\mu + c^{-4}B_\mu + O(c^{-6})\,, \qquad \Pi_{\mu\nu} = h_{\mu\nu} + c^{-2}\Phi_{\mu\nu} + O(c^{-4})\,, \tag{14}$$

$$T^\mu = v^\mu + O(c^{-2})\,, \qquad \Pi^{\mu\nu} = h^{\mu\nu} + O(c^{-2})\,, \tag{15}$$

where the LO fields obey the following orthogonality and completeness relations

$$\tau_\mu v^\mu = -1\,, \qquad \tau_\mu h^{\mu\nu} = 0\,, \qquad v^\mu h_{\mu\nu} = 0\,, \qquad h^{\mu\rho}h_{\rho\nu} = \delta^\mu_\nu + v^\mu\tau_\nu\,. \tag{16}$$

This implies that the metric is expanded as

$$g_{\mu\nu} = -c^2\tau_\mu\tau_\nu + \bar{h}_{\mu\nu} + \frac{1}{c^2}\bar{\Phi}_{\mu\nu} + O(c^{-4})\,, \tag{17}$$

where we defined

$$\bar{h}_{\mu\nu} = h_{\mu\nu} - 2\tau_{(\mu}m_{\nu)}\,, \qquad \bar{\Phi}_{\mu\nu} = \Phi_{\mu\nu} - m_\mu m_\nu - B_\mu\tau_\nu - B_\nu\tau_\mu\,. \tag{18}$$

This is equivalent to metric expansions presented in the literature[1,2]. In the decomposition (8) we can perform local Lorentz transformations that act on $T_\mu$ and the vielbeine that form $\Pi_{\mu\nu}$. In the $1/c^2$ expansion these become local Galilean boosts (also known as Milne boosts) at leading order and NLO local boosts at NLO. These leave each order of the metric invariant. The quantities $\tau_\mu$, $\bar{h}_{\mu\nu}$ and $h^{\mu\nu}$ (the LO term of the inverse metric) are Milne boost invariant. The combination (18) in the metric at order $c^{-2}$ is invariant under both local Galilean and NLO local boosts.

The form of the expansions for $T_\mu$ and $\Pi_{\mu\nu}$ should be preserved in any coordinate system. This means that we must also expand the diffeomorphisms in powers of $1/c^2$. If under a diffeomorphism a tensor $X_{\mu\nu}$ changes as $\delta X_{\mu\nu} = \mathcal{L}_\Xi X_{\mu\nu}$ then we should write $\Xi^\mu = \xi^\mu + c^{-2}\zeta^\mu + O(c^{-4})$. All fields in the expansions are then tensors with respect to diffeomorphisms generated by $\xi^\mu$. The LO fields $\tau_\mu$ and $h_{\mu\nu}$

are invariant under the NLO diffeomorphisms $\zeta^\mu$, while the NLO fields transform according to

$$\delta m_\mu = \mathcal{L}_\zeta \tau_\mu = \partial_\mu \Lambda - \Lambda a_\mu + h^{\rho\sigma}\zeta_\sigma(\partial_\rho \tau_\mu - \partial_\mu \tau_\rho)\,, \tag{19}$$

$$\delta \Phi_{\mu\nu} = \mathcal{L}_\zeta h_{\mu\nu} = 2\Lambda K_{\mu\nu} + \check{\nabla}_\mu \zeta_\nu + \check{\nabla}_\nu \zeta_\mu\,, \tag{20}$$

where $\Lambda = \tau_\mu \zeta^\mu$ and $\zeta_\nu = h_{\nu\mu}\zeta^\mu$ and where $K_{\mu\nu} = -\frac{1}{2}\mathcal{L}_v h_{\mu\nu}$ is the extrinsic curvature and $a_\mu = \mathcal{L}_v \tau_\mu = v^\rho(\partial_\rho \tau_\mu - \partial_\mu \tau_\rho)$ is called the torsion vector. The derivative $\check{\nabla}_\mu$ is covariant with respect to the torsionful connection

$$\check{\Gamma}^\rho_{\mu\nu} = C^\rho_{\mu\nu}\big|_{\sigma=0} = -v^\rho \partial_\mu \tau_\nu + \frac{1}{2}h^{\rho\sigma}\left(\partial_\mu h_{\nu\rho} + \partial_\nu h_{\mu\rho} - \partial_\rho h_{\mu\nu}\right)\,. \tag{21}$$

The geometry is thus described by the LO fields $\tau_\mu$ and $h_{\mu\nu}$ and the subleading fields $m_\mu$ and $\Phi_{\mu\nu}$ are treated as gauge fields on this non-relativistic geometry. This setup is called type II Newton–Cartan geometry [3].

## 5. Non-relativistic expansion of the Einstein–Hilbert Lagrangian

We are now in a position to apply the results of section 2 to the case of the EH Lagrangian. Using (13) equation (4) becomes

$$\mathcal{L}_{\rm LO} = \frac{e}{4}h^{\mu\nu}h^{\rho\sigma}\tau_{\mu\rho}\tau_{\nu\sigma}\,, \tag{22}$$

where we defined $\tau_{\mu\nu} = \partial_\mu \tau_\nu - \partial_\nu \tau_\mu$ and $e = \left(-\det\left(-\tau_\mu\tau_\nu + h_{\mu\nu}\right)\right)^{1/2}$.

We can consider $\tau_\mu$, $h_{\mu\nu}$, $m_\mu$ and $\Phi_{\mu\nu}$ to be an independent set of quantities that can be varied independently. The total variation of $\mathcal{L}_{\rm LO}$ is then given by

$$\begin{aligned}\delta\mathcal{L}_{\rm LO} = e\left[\frac{1}{8}h^{\mu\nu}h^{\rho\sigma}\tau_{\mu\rho}\tau_{\nu\sigma}h^{\alpha\beta} - \frac{1}{2}h^{\mu\alpha}h^{\nu\beta}h^{\rho\sigma}\tau_{\mu\rho}\tau_{\nu\sigma}\right]\delta h_{\alpha\beta} \\ + e\left[-\frac{1}{4}h^{\mu\nu}h^{\rho\sigma}\tau_{\mu\rho}\tau_{\nu\sigma}v^\alpha + h^{\rho\sigma}a_\rho h^{\nu\alpha}\tau_{\nu\sigma} - e^{-1}\partial_\nu\left(e h^{\mu\nu}h^{\rho\alpha}\tau_{\mu\rho}\right)\right]\delta\tau_\alpha\,.\end{aligned} \tag{23}$$

We used here that $\delta h^{\mu\nu} = \left(v^\mu h^{\nu\rho} + v^\nu h^{\mu\rho}\right)\delta\tau_\rho - h^{\mu\rho}h^{\nu\sigma}\delta h_{\rho\sigma}$. Contracting the $\tau_\mu$ equation of motion with $\tau_\mu$ tells us that $h^{\mu\nu}h^{\rho\sigma}\tau_{\mu\rho}\tau_{\nu\sigma} = 0$. This is a sum of squares and so it implies the vanishing of the twist tensor $h^{\mu\nu}h^{\rho\sigma}\tau_{\mu\rho} = 0$ and thus that $\tau \wedge d\tau = 0$. This is a causality condition for a non-relativistic spacetime. The spacetime admits a global foliation given by a nowhere vanishing hypersurface orthogonal (clock) 1-form $\tau$. This is known as twistless torsional Newton–Cartan (TTNC) geometry [4].

From equation (13) we find that

$$\frac{\partial\tilde{\mathcal{L}}}{\partial\sigma}\Big|_{\sigma=0} = \sqrt{-\det\left(-T_\alpha T_\beta + \Pi_{\alpha\beta}\right)}\Pi^{\mu\nu}\overset{(C)}{R}{}_{\mu\nu}\big|_{\sigma=0} = e h^{\mu\nu}\check{R}_{\mu\nu}\,. \tag{24}$$

The NLO Lagrangian is of the form (5) and can thus be written as

$$\mathcal{L}_{\rm NLO} = e h^{\mu\nu}\check{R}_{\mu\nu} + \frac{\delta\mathcal{L}_{\rm LO}}{\delta\tau_\mu}m_\mu + \frac{\delta\mathcal{L}_{\rm LO}}{\delta h_{\mu\nu}}\Phi_{\mu\nu}\,, \tag{25}$$

where the Ricci tensor is defined with respect to the connection (21) as

$$\check{R}_{\mu\nu} = \check{R}_{\mu\rho\nu}{}^{\rho} = -\partial_\mu \check{\Gamma}^{\rho}_{\rho\nu} + \partial_\rho \check{\Gamma}^{\rho}_{\mu\nu} - \check{\Gamma}^{\rho}_{\mu\lambda}\check{\Gamma}^{\lambda}_{\rho\nu} + \check{\Gamma}^{\rho}_{\rho\lambda}\check{\Gamma}^{\lambda}_{\mu\nu}\,. \tag{26}$$

In section 2 we have shown that the equations of motion of the NNLO Lagrangian include all the equations of motion of the NLO Lagrangian. Furthermore, as can be seen from (6), the equations of motion of the NNLO fields will reproduce the equations of motion of the LO Lagrangian, i.e. they will impose the TTNC condition. If we are only interested in those equations of motion that involve at most NLO but not NNLO fields then we can ignore the variations of $\tau_\mu$ and $h_{\mu\nu}$ that lead to equations of motion for the NNLO fields. Using the TTNC condition on shell it can be seen that the only $\tau_\mu$ and $h_{\mu\nu}$ variations that do not involve NNLO fields are $\tau_\mu \frac{\delta}{\delta\tau_\mu}$ and $h_{\mu\rho}h_{\nu\sigma}\frac{\delta}{\delta h_{\mu\nu}}$. As far as these variations and those of the other fields are concerned we might as well implement the TTNC condition off shell. We thus conclude that an action for the LO and NLO fields can be obtained by computing the NNLO Lagrangian with TTNC off shell. The TTNC condition itself can always be obtained by adding a Lagrange multiplier of the form $\frac{1}{2}B_\mu \frac{\delta \mathcal{L}_{\text{LO}}}{\delta\tau_\mu}$ where $B_\mu$ is a NNLO field.

We will call the NNLO Lagrangian with TTNC imposed off shell the non-relativistic gravity (NRG) Lagrangian. Using (6) it can be shown to take the form

$$\begin{aligned}\mathcal{L} = e\Bigg[&-v^\mu v^\nu \check{R}_{\mu\nu} - 2m_\nu \check{\nabla}_\mu \left(h^{\mu\rho}h^{\nu\sigma} - h^{\mu\nu}h^{\rho\sigma}\right) K_{\rho\sigma} + \Phi h^{\mu\nu}\check{R}_{\mu\nu} + \frac{1}{4}h^{\mu\rho}h^{\nu\sigma}F_{\mu\nu}F_{\rho\sigma} \\ &-\Phi_{\rho\sigma}h^{\mu\rho}h^{\nu\sigma}\left(\check{R}_{\mu\nu} - \check{\nabla}_\mu a_\nu - a_\mu a_\nu - \frac{1}{2}h_{\mu\nu}h^{\kappa\lambda}\check{R}_{\kappa\lambda} + h_{\mu\nu}e^{-1}\partial_\kappa\left(eh^{\kappa\lambda}a_\lambda\right)\right)\Bigg]\,,\end{aligned} \tag{27}$$

where $\Phi = -v^\mu m_\mu$ is the Newtonian potential and where $F_{\mu\nu} = \partial_\mu m_\nu - \partial_\nu m_\mu - a_\mu m_\nu + a_\nu m_\mu$. When comparing with (6) we see that the first term in (27) agrees with $\frac{1}{2}\frac{\partial^2\tilde{\mathcal{L}}}{\partial\sigma^2}\big|_{\sigma=0}$. The terms linear in $\Phi_{\mu\nu}$ and $m_\mu$ agree with $\phi^I_{(1)}\frac{\delta}{\delta\phi^I_{(0)}}\frac{\partial\tilde{\mathcal{L}}}{\partial\sigma}\big|_{\sigma=0}$. The terms containing NNLO fields in (6) are not present here because we used TTNC off shell. Finally the second order variation of the LO Lagrangian with respect to $\tau_\mu$ is still nontrivial even when we use TTNC off shell and this leads to the field strength squared terms in (27) on the last line. Note that there are no terms quadratic in $\Phi_{\mu\nu}$ because all derivatives with respect to $h_{\mu\nu}$ of the LO Lagrangian give zero upon using the TTNC condition. The equation of motion of $\Phi_{\mu\nu}$ agrees with the equation of motion of $h_{\mu\nu}$ when varying (25) and the $m_\mu$ equation of motion of (27) agrees with the $\tau_\mu$ equation of motion of (25).

There is only one term containing $\Phi = -v^\mu m_\mu$, the Newtonian potential. To find Newtonian gravity and its generalization to TTNC geometries requires computing the variations $\tau_\mu\frac{\delta}{\delta\tau_\mu}$ and $h_{\mu\nu}\frac{\delta}{\delta h_{\mu\nu}}$[3]. If one would like to continue expanding to higher orders then we would need to compute the full NNLO Lagrangian without using TTNC off shell.

In non-relativistic gravity $\tau$ is of the form $\tau = NdT$ where $N$ is a non-relativistic lapse function and $T$ is some time function. This means that gravitational time dilation is a non-relativistic phenomenon[2,5]. For example the Tolman–Oppenheimer–Volkov solution for a fluid star is a solution of NRG coupled to a fluid[6,7]. As a result

the classical tests of general relativity: gravitational redshift, perihelion precession and bending of light are also passed by non-relativistic gravity[5].

We can rewrite (27) in terms of manifest Milne boost invariant tensors as

$$\mathcal{L} = e\Big[-\hat{v}^\mu\hat{v}^\nu\bar{R}_{\mu\nu} + \hat{\Phi}h^{\mu\nu}\bar{R}_{\mu\nu} - \bar{\Phi}_{\rho\sigma}h^{\mu\rho}h^{\nu\sigma}\Big(\bar{R}_{\mu\nu} - \bar{\nabla}_\mu a_\nu - a_\mu a_\nu - \frac{1}{2}h_{\mu\nu}h^{\kappa\lambda}\bar{R}_{\kappa\lambda} + h_{\mu\nu}e^{-1}\partial_\kappa\left(eh^{\kappa\lambda}a_\lambda\right)\Big)\Big], \tag{28}$$

where $\hat{v}^\mu = v^\mu - h^{\mu\nu}m_\nu$ and $\hat{\Phi} = -v^\mu m_\mu + \frac{1}{2}h^{\mu\nu}m_\mu m_\nu$. In here the Ricci tensors are defined with respect to the Milne boost invariant connection

$$\bar{\Gamma}^\rho_{\mu\nu} = -\hat{v}^\rho\partial_\mu\tau_\nu + \frac{1}{2}h^{\rho\sigma}\left(\partial_\mu\bar{h}_{\nu\rho} + \partial_\nu\bar{h}_{\mu\rho} - \partial_\rho\bar{h}_{\mu\nu}\right). \tag{29}$$

The NRG Lagrangian is uniquely fixed by all its gauge symmetries[3].

It would be interesting to make contact with the post-Newtonian approximation by finding a convenient way of expressing the higher orders of the $1/c^2$ expansion of the EH Lagrangian and more generally to see if this off shell and covariant approach can be used for interesting approximations to general relativity. To this end it might be useful to employ a first order formulation[8,9].

## Acknowledgments

We thank Dieter Van den Bleeken for useful discussions. The work of DH is supported by the Swiss National Science Foundation through the NCCR SwissMAP. The work of JH is supported by the Royal Society University Research Fellowship "Non-Lorentzian Geometry in Holography" (UF160197). The work of NO is supported in part by the project "Towards a deeper understanding of black holes with non-relativistic holography" of the Independent Research Fund Denmark (DFF-6108-00340) and Villum Foundation Experiment project 00023086.

## References

1. G. Dautcourt, *Class. Quant. Grav.* **14** (1997) A109 [gr-qc/9610036].
2. D. Van den Bleeken, *Class. Quant. Grav.* **34** (2017) no.18, 185004 [arXiv: 1703.03459].
3. D. Hansen, J. Hartong and N. A. Obers, *Phys. Rev. Lett.* **122** (2019) no.6, 061106 [arXiv:1807.04765].
4. M. H. Christensen, J. Hartong, N. A. Obers and B. Rollier, *Phys. Rev.* D**89** (2014) 061901 [arXiv:1311.4794].
5. D. Hansen, J. Hartong and N. A. Obers, arXiv:1904.05706.
6. D. Van den Bleeken, arXiv:1903.10682.
7. D. Hansen, J. Hartong and N. A. Obers, in preparation.
8. M. Cariglia, *Phys. Rev.* D**98** (2018) no.8, 084057 [arXiv:1811.03446].
9. E. Bergshoeff, J. M. Izquierdo, T. Ortn and L. Romano, arXiv:1904.08304.

# Non-relativistic supergravity

Johannes Lahnsteiner

*Van Swinderen Institute, University of Groningen,*
*Nijenborgh 4, 9747 AG Groningen, The Netherlands*
*E-mail: j.m.lahnsteiner@rug.nl*

Jan Rosseel

*Faculty of Physics, University of Vienna,*
*Boltzmanngasse 5, A-1090, Vienna, Austria*
*E-mail: jan.rosseel@univie.ac.at*

In this talk, we review the construction of two three-dimensional non-relativistic pure supergravity theories. One of these yields a supersymmetrization of Newton-Cartan gravity. The other one is a Chern-Simons theory with different bosonic equations of motion and matter couplings than Newton-Cartan gravity. We comment on the prospect of using these theories to construct non-relativistic supersymmetric field theories in curved backgrounds, that can be amenable to exact non-perturbative analysis via localization techniques.

*Keywords*: Supergravity; Newton-Cartan gravity.

## 1. Introduction

Ever since its inception in Ref. 1, supergravity (sugra) has played an important role in string theory, cosmology and high energy particle physics. Apart from its phenomenological use, sugra has also been instrumental as a tool to study strongly coupled and non-perturbative aspects of Quantum Field Theory (QFT). For example, the anti-de Sitter/Conformal Field Theory (AdS/CFT) correspondence allows one to rephrase the calculation of quantities in $d$-dimensional strongly coupled CFTs as computations in weakly coupled dual $(d+1)$-dimensional sugra theories. Another development where sugra is used in the study of non-perturbative QFT, concerns the construction of supersymmetric QFT in curved space-times. Such QFTs have the advantage that localization techniques often allow quantities such as the partition function to be calculated exactly (see Ref. 2 for a recent collection of reviews). Here, sugra techniques give a particularly convenient and insightful way to obtain the Lagrangians and supersymmetry transformation rules of these theories as a rigid limit of matter-coupled off-shell sugra, as has been shown in Ref. 3.

Sugra is most often discussed in the relativistic regime, that is suitable for high energy physics. In light of the above mentioned applications, it is clear that also non-relativistic (NR) sugra can be useful to study strongly coupled NR effective field theories that appear in Condensed Matter physics in the context of e.g. the Fractional Quantum Hall Effect or quantum critical points. NR versions of the AdS/CFT correspondence have been proposed, in which strongly coupled NR CFTs are given a dual gravity description.[4–6] Most of the ensuing work has focused on non-supersymmetric versions of these dualities. Including supersymmetry can shed

more light on the validity of such conjectured correspondences and this is where NR sugra can show its applicability. Similarly, if one can extend the results of Ref. 3 to the NR realm, one can use NR sugra as a tool to construct supersymmetric NR QFTs in curved space-times, that are amenable to exact analysis using localization techniques.

Despite the fact that relativistic sugra is by now more than 40 years old, NR versions thereof have only recently appeared in the literature. The first non-trivial pure sugra theory that incorporates only NR gravitational and spin-3/2 fields was constructed in Ref. 7. This is a three-dimensional on-shell theory and its equations of motion correspond to a supersymmetric version of Newton-Cartan (NC) gravity, Cartan's diffeomorphism covariant reformulation of the laws of Newtonian gravity.[8] A peculiar feature of NR gravity is that one can formulate different theories, with different local space-time symmetries and equations of motion, that nevertheless are of the same order of derivatives. This is unlike relativistic gravity, where the only possibility for local space-time symmetries is given by Lorentz symmetry, leading to General Relativity as the unique theory of gravity at the two-derivative level. Indeed, a three-dimensional NR gravity theory, that is different from NC gravity, was considered as a Chern-Simons (CS) theory in Ref. 10. The sugra version of this theory was then constructed as a super-CS theory in Ref. 9.

In this talk, we will give a brief overview of the construction and structure of the sugra theories of Refs. 7 and 9. We will then also comment on ongoing and future research regarding the steps that need to be taken to apply these theories to studies of non-perturbative aspects of NR field theories.

## 2. Three-dimensional Newton-Cartan and Extended Bargmann supergravity

### 2.1. *Three-dimensional Newton-Cartan supergravity*

A three-dimensional sugra version of NC gravity can be constructed by gauging a suitable NR supersymmetry algebra. The algebra chosen in Ref. 7 is the three-dimensional $\mathcal{N} = 2$ super-Bargmann algebra.[11] The bosonic part of this algebra is given by the Bargmann algebra, that consists of time and space translations, spatial rotations, NR boosts and a central charge that is associated to mass conservation. The first step of the gauging procedure then consists of associating gauge fields to every generator of the superalgebra and reading off the gauge transformation rules and covariant field strengths from the structure constants of the superalgebra in Ref. 11. See Table 1 for the notation used here.

The bosonic fields $\tau_\mu$, $e_\mu{}^a$ and $m_\mu$ transform under local rotations, boosts and the central charge such that they can be interpreted as the time-like and space-like Vielbeine and the central charge gauge field of the Vielbein formulation of NC geometry.[12,13] The two fields $\omega_\mu{}^{ab}$, $\omega_\mu{}^a$ can be interpreted as supersymmetric versions of the rotation and boost spin connections of NC geometry, once the following

Table 1. Gauging the $\mathcal{N}=2$ super-Bargmann algebra

| symmetry | generators | gauge field | parameters | curvatures |
|---|---|---|---|---|
| time translations | $H$ | $\tau_\mu$ | $\zeta(x^\nu)$ | $\hat{R}_{\mu\nu}(H)$ |
| space translations | $P_a$ | $e_\mu{}^a$ | $\zeta^a(x^\nu)$ | $\hat{R}_{\mu\nu}(P_a)$ |
| boosts | $G_a$ | $\omega_\mu{}^a$ | $\lambda^a(x^\nu)$ | $\hat{R}_{\mu\nu}(G_a)$ |
| spatial rotations | $J_{ab}$ | $\omega_\mu{}^{ab}$ | $\lambda^{ab}(x^\nu)$ | $\hat{R}_{\mu\nu}(J_{ab})$ |
| central charge transf. | $M$ | $m_\mu$ | $\sigma(x^\nu)$ | $\hat{R}_{\mu\nu}(M)$ |
| two supersymmetries | $Q_\alpha^\pm$ | $\psi_{\mu\pm}$ | $\epsilon_\pm(x^\nu)$ | $\hat{\psi}_{\mu\nu\pm}$ |

so-called 'conventional' constraints are imposed

$$\hat{R}_{\mu\nu}(P^a) = 2\partial_{[\mu}e_{\nu]}{}^a - 2\omega_{[\mu}{}^{ab}e_{\nu]b} - 2\omega_{[\mu}{}^a\tau_{\nu]} - \bar{\psi}_{[\mu+}\gamma^a\psi_{\nu]-} = 0\,, \tag{1}$$

$$\hat{R}_{\mu\nu}(M) = 2\partial_{[\mu}m_{\nu]} - 2\omega_{[\mu}{}^a e_{\nu]a} - \bar{\psi}_{[\mu-}\gamma^0\psi_{\nu]-} = 0\,. \tag{2}$$

Solving these constraints to express $\omega_\mu{}^{ab}$, $\omega_\mu{}^a$ in terms of the other fields, one finds that the resulting expressions transform as expected from NC spin connections.

At this point, one is left with the independent fields $\tau_\mu$, $e_\mu{}^a$, $m_\mu$ and $\psi_{\mu\pm}$. The commutator of two local supersymmetries should then close on these fields in order for them to define a NR sugra multiplet. The conventional constraints ensure that the supersymmetry algebra properly closes on $e_\mu{}^a$ and $m_\mu$. In order for this to work on $\tau_\mu$ however, one needs to impose the extra constraint

$$\hat{R}_{\mu\nu}(H) = 2\partial_{[\mu}\tau_{\nu]} - \frac{1}{2}\bar{\psi}_{[\mu+}\gamma^0\psi_{\nu]-} = 0\,. \tag{3}$$

This is the supercovariantization of the so-called 'zero torsion' or 'foliation' constraint that is often imposed in NR gravity, as it expresses the notion of absolute time in a coordinate independent way.

For consistency, one then needs to require that this constraint is invariant under supersymmetry, leading to two additional constraints

$$\hat{R}_{\mu\nu}(H) = 0 \quad \Rightarrow \quad \hat{\psi}_{\mu\nu+} = 0 \quad \Rightarrow \quad \hat{R}_{\mu\nu}(J_{ab}) = 0\,. \tag{4}$$

The constraint $\hat{\psi}_{\mu\nu+} = 0$ ensures algebra closure on $\psi_{\mu+}$. For the superalgebra to close on $\psi_{\mu-}$, one needs to impose two more constraints

$$e^\mu{}_a e^\nu{}_b\hat{\psi}_{\mu\nu-} = 0\,, \qquad \gamma^\mu\tau^\nu\hat{\psi}_{\mu\nu-} = 0\,, \tag{5}$$

where $e^\mu{}_a$, $\tau^\mu$ are projective inverses of $e_\mu{}^a$ and $\tau_\mu$. The last of these constraints can be interpreted as an equation of motion for a NR spin-3/2 field. This interpretation is confirmed by its supersymmetry variation that is given by a supersymmetric version of the covariant Poisson equation, that is the analog of the Einstein equations in NC gravity. At this point, one is then left with a set of independent fields $\tau_\mu$, $e_\mu{}^a$, $m_\mu$, $\psi_{\mu\pm}$ that form a consistent NR on-shell sugra multiplet. This multiplet can be called 'on-shell Newton-Cartan' since it only realizes the superalgebra once the supersymmetric NC equations of motion hold.

### 2.2. *Extended Bargmann supergravity*

One of the drawbacks of ordinary NC supergravity described above, is that there is no simple action principle (involving just the fields given above) known that leads to the NC equations of motion. This complicates discussions of matter couplings and attempts to e.g. construct supersymmetric QFTs in curved space-times by generalizing Ref. 3. Moreover, as noted in eq. 4, on-shell NC sugra implements a foliation constraint $\hat{R}_{\mu\nu}(H) = 0$ from which it follows that also $\hat{R}_{\mu\nu}(J_{ab}) = 0$. These two constraints imply that the theory's space-time manifold is foliated in an absolute time and flat spatial slices, severely limiting the possible backgrounds that could be obtained via a generalization of Ref. 3. Adding matter is unlikely to change this, as matter coupling in NC gravity does not change these constraints.

As mentioned in the introduction, there exists a different three-dimensional NR gravity theory, that is a CS theory with gauge algebra given by a central extension of the Bargmann algebra encountered in the previous subsection.[14] In Ref. 9, it was shown that adding matter density in this theory sources the constraint that implies that the spatial slices are flat, so that matter coupling can now lead to interesting NR curved backgrounds. A sugra generalization might thus serve as a good starting point to generalize the construction of Ref. 3 to NR QFTs.

Such a sugra generalization was constructed in Ref. 9, as a super-CS theory based upon a novel superalgebra that admits a non-degenerate supertrace. With respect to Table 1, this algebra contains 2 extra generators: a bosonic generator $S$ with associated gauge field $s_\mu$ and field strength $\hat{R}_{\mu\nu}(S)$ and a fermionic generator $R$ with associated gauge field $\rho_\mu$ and field strength $\hat{\rho}_{\mu\nu}$. The action is easily constructed following the CS recipe and is given by

$$S = \frac{k}{4\pi}\int \mathrm{d}^3x\epsilon^{\mu\nu\rho}\Big(e_\mu{}^a\hat{R}_{\nu\rho}(G_a) - m_\mu\hat{R}_{\nu\rho}(J) - \tau_\mu\hat{R}_{\mu\nu}(S) + \bar{\psi}_{\mu+}\hat{\rho}_{\nu\rho} + \bar{\rho}_\mu\hat{\psi}_{\nu\rho+} + \bar{\psi}_{\mu-}\hat{\psi}_{\nu\rho-}\Big)\,, \tag{6}$$

where $k$ is the CS coupling constant and $\hat{R}_{\mu\nu}(J) = \frac{1}{2}\epsilon^{ab}\hat{R}_{\mu\nu}(J_{ab})$. Varying this action with respect to $\omega_\mu{}^a$, $\omega_\mu{}^{ab}$ reproduces the conventional constraints 1, that allow one to express these spin connections in terms of the other fields. Variation with respect to $s_\mu$ yields the foliation constraint 3. Since one has an action at hand, matter coupling can be done by adding a suitable matter action, minimally coupled to this sugra background. In such a minimally coupled matter action, the field $m_\mu$ couples to the matter density. Since $m_\mu$ couples to $\hat{R}_{\mu\nu}(J_{ab})$ in the gravitational action 6, one sees that the inclusion of matter can indeed source the curvature $\hat{R}_{\mu\nu}(J_{ab})$ of spatial sections, as advertised.

## 3. Towards non-relativistic supersymmetric field theories in curved space-times

The method of Ref. 3 to obtain supersymmetric QFTs in curved backgrounds starts from matter-coupled off-shell sugra. This has several technical advantages. Firstly,

the off-shell character of the sugra multiplet cleanly separates between dynamical matter fields and sugra background fields. It further allows one to systematically classify all possible backgrounds on which supersymmetry can be defined. This is done by looking for background values for the metric and auxiliary fields of the off-shell sugra multiplet for which the Killing spinor equations, obtained by putting the gravitino supersymmetry transformations equal to zero, have non-trivial solutions. The auxiliary fields can then be physically interpreted as couplings of the supercurrent multiplet to the background geometry.

Both NR sugra multiplets discussed in the previous section are on-shell. In order to maintain the above mentioned technical advantages, it would be desirable to extend these multiplets to off-shell ones. This question has been addressed for the on-shell NC multiplet in Refs. 15, 16. The multiplets constructed in these references are off-shell, in the sense that one no longer needs the second of eqs. 5 to obtain algebra closure. At present, there is however no completely off-shell formulation of the multiplets presented above known, for which algebra closure is guaranteed without the imposition of extra constraints. One always needs to impose the foliation constraint or a so-called twistless torsional (see Ref. 18) generalization thereof. The last constraint of eqs. 4 is modified to include auxiliary field contributions. It would be interesting to investigate how to couple matter to these multiplets. Likewise, it would be interesting to carry out an analysis of the Killing spinor equations, following from these multiplets, to see whether they can support non-trivial curved supersymmetric backgrounds.[17]

The lack of completely off-shell multiplets indicates that an alternative approach to obtaining NR supersymmetric field theories in curved space-times can be useful. In this regard, it is noteworthy that four-dimensional Lorentzian backgrounds on which supersymmetry can be defined exhibit a null Killing vector.[19,20] It is known that dimensional reduction along null isometries leads to NR theories.[12,21] This suggests that three-dimensional NR supersymmetric field theories on curved backgrounds can be obtained by dimensional reduction of relativistic ones over null directions. This approach is currently pursued in work in progress, on which we hope to report soon.[22]

## Acknowledgments

The authors are grateful to the organizers of the MG15 meeting and the organizers of the 'Newton-Cartan gravity' parallel session at the MG15 meeting for the opportunity to present the work described here.

## References

1. D. Z. Freedman, P. van Nieuwenhuizen and S. Ferrara, "Progress Toward a Theory of Supergravity," Phys. Rev. D **13** (1976) 3214.
2. V. Pestun *et al.*, "Localization techniques in quantum field theories," J. Phys. A **50** (2017) no.44, 440301.

3. G. Festuccia and N. Seiberg, "Rigid Supersymmetric Theories in Curved Superspace," JHEP **1106** (2011) 114.
4. D. T. Son, "Toward an AdS/cold atoms correspondence: A Geometric realization of the Schrodinger symmetry," Phys. Rev. D **78** (2008) 046003.
5. K. Balasubramanian and J. McGreevy, "Gravity duals for non-relativistic CFTs," Phys. Rev. Lett. **101** (2008) 061601.
6. S. Kachru, X. Liu and M. Mulligan, "Gravity duals of Lifshitz-like fixed points," Phys. Rev. D **78** (2008) 106005.
7. R. Andringa, E. A. Bergshoeff, J. Rosseel and E. Sezgin, "3D Newton–Cartan supergravity," Class. Quant. Grav. **30** (2013) 205005.
8. É. Cartan, "Sur les Variétés à Connexion Affine et la Théorie de la Relativité Généralisée (suite)," Ann. École Norm. Sup. 40 (1923), 325-412.
9. E. A. Bergshoeff and J. Rosseel, "Three-Dimensional Extended Bargmann Supergravity," Phys. Rev. Lett. **116** (2016) no. 25, 251601.
10. G. Papageorgiou and B. J. Schroers, "A Chern-Simons approach to Galilean quantum gravity in 2+1 dimensions," JHEP **0911** (2009) 009.
11. J. Lukierski, I. Prochnicka, P. C. Stichel and W. J. Zakrzewski, "Galilean exotic planar supersymmetries and nonrelativistic supersymmetric wave equations," Phys. Lett. B **639** (2006) 389.
12. C. Duval, G. Burdet, H. P. Kunzle and M. Perrin, "Bargmann Structures and Newton-Cartan Theory," Phys. Rev. D **31** (1985) 1841.
13. R. Andringa, E. Bergshoeff, S. Panda and M. de Roo, "Newtonian Gravity and the Bargmann Algebra," Class. Quant. Grav. **28** (2011) 105011.
14. J. M. Lévy-Leblond, "Galilei group and Galilean invariance," Group Theory and Applications (Loebl Ed.), **II**, Acad. Press, New York, p. 222 (1972).
15. E. Bergshoeff, J. Rosseel and T. Zojer, "Newton-Cartan (super)gravity as a non-relativistic limit," Class. Quant. Grav. **32** (2015) no. 20, 205003.
16. E. Bergshoeff, J. Rosseel and T. Zojer, "Newton-Cartan supergravity with torsion and Schrdinger supergravity," JHEP **1511** (2015) 180.
17. G. Knodel, P. Lisbao and J. T. Liu, "Rigid Supersymmetric Backgrounds of 3-dimensional Newton-Cartan Supergravity," JHEP **1606** (2016) 028.
18. M. H. Christensen, J. Hartong, N. A. Obers and B. Rollier, "Torsional Newton-Cartan Geometry and Lifshitz Holography," Phys. Rev. D **89** (2014) 061901.
19. J. T. Liu, L. A. Pando Zayas and D. Reichmann, "Rigid Supersymmetric Backgrounds of Minimal Off-Shell Supergravity," JHEP **1210** (2012) 034.
20. D. Cassani, C. Klare, D. Martelli, A. Tomasiello and A. Zaffaroni, "Supersymmetry in Lorentzian Curved Spaces and Holography," Commun. Math. Phys. **327** (2014) 577.
21. B. Julia and H. Nicolai, "Null Killing vector dimensional reduction and Galilean geometrodynamics," Nucl. Phys. B **439** (1995) 291.
22. E. A. Bergshoeff, J. Lahnsteiner and J. Rosseel, "Non-Relativistic Supersymmetry on Three-Manifolds," work in progress.

# Torsional Newton-Cartan gravity and strong gravitational fields

Dieter Van den Bleeken*

*Physics Department, Boğaziçi University,*
*Istanbul, Turkey*
*Institute for Theoretical Physics, KU Leuven*
*Leuven, Belgium*
**E-mail: dieter.van@boun.edu.tr*

We review how the large $c$ expansion of General Relativity leads to an effective theory in the form of Twistless Torsional Newton-Cartan gravity. We show how this is a strong field expansion around the static sector of General Relativity and illustrate this through two examples.

*Keywords*: Nonrelativistic gravity; Strong gravity; Newton-Cartan geometry.

## 1. Introduction

Weak field and/or nonrelativistic approximations of General Relativity (GR) have various applications, see e.g. Ref. 1 for an introduction. These approximations are typically worked out in an adapted coordinate system. To the author's knowledge a covariant expansion of GR in the (inverse) speed of light $c$ was not systematically studied before the work of Dautcourt[2] in the '90s. In that work it is assumed that the coefficient of $c^2$ in the relativistic connection vanishes and with this assumption the leading and subleading orders of the expansion reproduce Newtonian and post-Newtonian gravity respectively, see also Ref. 3. Because the formalism is manifestly covariant one reproduces (post-)Newtonian gravity in a covariant formalism, so called Newton-Cartan (NC) gravity. The analysis of Ref. 2 leaves however two questions: 1) why this assumption on the connection? 2) how come a large $c$ expansion automatically leads to a small $G_{\mathrm{N}}$ result? Indeed, as was recently shown[4], these two questions are closely related: relaxing the assumption on the connection introduces (nonrelativistic) strong gravitational effects, in particular time dilation. Working out the expansion in full detail one obtains at leading order a generalization of Newtonian gravity, where this effect of time dilation is encoded as a torsion in the NC connection.[a]

## 2. Review of the large $c$ expansion

In the large $c$ expansion[2–4] one starts from an expansion of the metric and its inverse in (even[b]) powers of $c^{-1}$:

$$g_{\mu\nu} = \sum_{i=-1}^{\infty} \overset{(2i)}{g}_{\mu\nu} c^{-2i} \qquad g^{\mu\nu} = \sum_{i=0}^{\infty} \overset{(2i)}{g}{}^{\mu\nu} c^{-2i}. \tag{1}$$

[a] Since this talk was presented some further related work appeared[5,6].

[b] It is consistent to consider only even powers of $c$, odd powers are allowed to appear however and including them is interesting[7].

The key physical input – that time is related to space through a factor $c$ – translates to the assumption that $\overset{(-2)}{g}_{\mu\nu}$ is of rank 1 and negative, so we can write

$$\overset{(-2)}{g}_{\mu\nu} = -\tau_\mu \tau_\nu . \tag{2}$$

One needs to impose the condition that $g_{\mu\nu}$ and $g^{\mu\nu}$ are each others inverse. For the leading term of the inverse metric this implies it is purely spatial, i.e. a rank three symmetric tensor orthogonal to the time direction:

$$\overset{(0)}{g}{}^{\mu\nu} = h^{\mu\nu} \qquad \text{with} \qquad h^{\mu\nu}\tau_\nu = 0. \tag{3}$$

At subleading order one finds that[4]

$$\overset{(0)}{g}_{\mu\nu} = -2\hat{\Phi}\tau_\mu\tau_\nu + \hat{h}_{\mu\nu} \qquad \overset{(2)}{g}{}^{\mu\nu} = -\hat{\tau}^\mu\hat{\tau}^\nu + \hat{\beta}^{\mu\nu} \tag{4}$$

where $\tau_\nu\hat{\tau}^\nu + \hat{h}_{\mu\rho}h^{\rho\nu} = \delta^\nu_\mu \qquad \hat{\tau}^\rho\hat{\tau}^\sigma\hat{h}_{\rho\sigma} = 0 \qquad \tau_\mu\hat{\beta}^{\mu\nu} = 0.$

Expanding the metric compatibility equation for the relativistic covariant derivative naturally leads to the nonrelativistic connection[4]

$$\begin{aligned}\overset{(nc)}{\Gamma}{}^\lambda_{\mu\nu} = \frac{1}{2}h^{\lambda\rho}\Big(\partial_\mu\hat{h}_{\rho\nu} + \partial_\nu\hat{h}_{\mu\rho} - \partial_\rho\hat{h}_{\mu\nu} + 2\partial_\rho\hat{\Phi}\tau_\mu\tau_\nu - 4\hat{\Phi}(\tau_\mu\partial_{[\nu}\tau_{\rho]} + \tau_\nu\partial_{[\mu}\tau_{\rho]})\Big) \\ + \hat{\tau}^\lambda\partial_\mu\tau_\nu .\end{aligned} \tag{5}$$

This connection is compatible with the Newton-Cartan structure, i.e. $\overset{(nc)}{\nabla}_\mu h^{\nu\lambda} = 0$ and $\overset{(nc)}{\nabla}_\mu\tau_\nu = 0$, but has torsion $\overset{(nc)}{\mathrm{T}}{}^\lambda_{\mu\nu} = 2\hat{\tau}^\lambda\partial_{[\mu}\tau_{\nu]}$.

One can now insert the expansion (1) into the relativistic Ricci tensor and expand it in turn in powers of $c^{-2}$. Equating each term to zero will provide an expansion of the vacuum Einstein equations[c]. At leading order the Einstein equation is equivalent to

$$\tau_{[\mu}\partial_\nu\tau_{\lambda]} = 0\,. \tag{6}$$

The geometric interpretation of the equation (6) is two-fold: first it implies that $\tau_\mu$ defines a foliation and hence a global time direction, second it implies that the torsion of the connection (5) is twistless[8], and the related geometry is referred to as Twistless Torsional Newton Cartan (TTNC) geometry.

The next to leading order Einstein equation provides an equation for this torsion:

$$h^{\lambda\rho}(\overset{(nc)}{\nabla}_\lambda\hat{a}_\rho - \hat{a}_\lambda\hat{a}_\rho) = 0\,, \qquad \text{where } \hat{a}_\mu = \mathcal{L}_{\hat{\tau}}\tau_\mu\,. \tag{7}$$

At next to next to leading order the expanded relativistic Ricci tensor provides an equation for the Ricci tensor of the Newton-Cartan connection (5), see Ref. 4 for the actual expression, we'll here simply refer to it as the 'Ricci equation'. The component of this Ricci equation proportional to $\tau_\mu\tau_\nu$ provides the Poisson equation for the Newtonian potential $\Phi$. The Ricci equation together with (7) provide a system of equations for the fields determining the TTNC geometry and for this reason we'll refer to the theory formed by these equations as TTNC gravity.

[c]See Ref. 4 for the contributions of the energy momentum tensor.

In summary, up to next to next to leading order in a $c^{-2}$ expansion GR is effectively described by TTNC gravity.

## 3. TTNC gravity as an expansion around static GR

One of the defining features of NC geometry is that it is covariant under 4 dimensional diffeomorphisms. But since in TTNC there is a well defined time direction – specified by $\tau_\mu$ – on which all observers agree, this 4 dimensional diffeomorphism covariance is somewhat artificial. Indeed, from a physical point of view it might be more natural to introduce an explicit time coordinate $t$, such that $\tau_\mu = e^{\lambda(t,x)}\delta^t_\mu$, the most general form allowed by (6). This choice of time breaks 4 dimensional diffeomorphisms to 3 dimensional time dependent diffeomorphisms.[d] It has the advantage that it removes a number of unphysical components of the fields and makes the physical field content more explicit. After this partial gauge fixing the expansion (2, 3, 4) can be rewritten as

$$\begin{aligned} ds^2 = &-c^2\, e^{\lambda} dt^2 + e^{-\lambda} h_{ij} dx^i dx^j - 2e^{\lambda} dt(\Phi dt - C_i dx^i) \\ &+ c^{-2}\left((e^{-\lambda}\beta_{ij} - e^{\lambda} C_i C_j) dx^i dx^j + B_i dt dx^i + \Sigma dt^2\right) + \mathcal{O}(c^{-4}). \end{aligned} \tag{8}$$

The physical fields are then a scalar $\lambda$, a spatial 3-metric $h_{ij}$, the Newtonian potential $\Phi$, a vectorfield $C_i$ and a symmetric 3-tensor $\beta_{ij}$[e]. This formulation makes manifest the connection between the NC torsion and the time-like warpfactor of GR, both determined by $\lambda$. Contrary to the Newtonian potential $\Phi$, which is the contribution to the warp factor at zeroth order the potential $\lambda$ appears at the higher order $c^2$, it encodes a strong gravitational effect that is absent in the weak field post-Minkowskian approximation.

The TTNC gravitational equations, (7) and the Ricci equation, take the following form in this formulation ($K_{ij} = 2\partial_{[i}C_{j]}$ , $G_i = -\partial_i\Phi - e^{-\frac{1}{2}\lambda}\partial_t(e^{\frac{1}{2}\lambda}C_i)$):

$$\nabla_i \partial^i \lambda = 0 \qquad\qquad R_{ij} = \frac{1}{2}\partial_i\lambda\partial_j\lambda \tag{9}$$

$$\frac{1}{2}\nabla^j\left(e^{2\lambda}K_{ij}\right) = -h^{jk}\nabla_{[j}\dot{h}_{i]k} - \frac{1}{2}(h^{jk}\dot{h}_{jk} - \dot{\lambda})\partial_i\lambda - \partial_i\dot{\lambda} + \dot{h}_{ij}\partial^j\lambda \tag{10}$$

$$-e^{\frac{3}{2}\lambda}\nabla^i(e^{\frac{1}{2}\lambda}G_i) = T[\ddot{\lambda}, \ddot{h}, \partial\phi, \nabla\beta]. \tag{11}$$

The first important point to note is the hierarchical nature of the above equations. The first two equations (9) form an independent closed system of equations for $h_{ij}(t,x)$ and $\lambda(t,x)$ wich determines their spatial dependence but leaves the time dependence free. The third equation (10) is an equation that determines the spatial dependence of $C_i(t,x)$ and takes the time derivatives of $h_{ij}$ and $\lambda$ as input sources. The last equation (11) is an equation containing $\Delta\Phi$ and can thus be identified with

[d] See e.g. Ref. 9 for some details on the precise relation between the 4d diffeomorphisms, Milne boosts and 3d time dependent diffeomorphisms.

[e] Note that the fields $B_i$ and $\Sigma$ will not appear in the equations to the order of the expansion we consider in this work.

the Newtonian Poisson equation, interestingly it contains second time derivatives of $h_{ij}$ and $\lambda$, and spatial derivatives of $C_i$ as sources[f].

Secondly one observes[g] that the equations (9) are essentially the static Einstein equations. Indeed any static metric can be written in the form $ds^2 = -c^2\, e^{\tilde{\lambda}(x)} dt^2 + e^{-\tilde{\lambda}(x)} \tilde{h}_{ij}(x) dx^i dx^j$ and for such an ansatz the vacuum Einstein equations reduce to equations of the form (9). The key difference between (9) and the static Einstein equations is that contrary to $\tilde{\lambda}$ and $\tilde{h}_{ij}$ the fields $\lambda$ and $h_{ij}$ are time-dependent. We come to the conclusion that $\lambda$ and $h_{ij}$ correspond to solutions of the static Einstein equations with any integration constants replaced by arbitrary functions of time. Such time-dependent deviations from static solutions then source the equations for the higher order fields, starting a series of corrections. One thus sees that the large $c$-expansion is an expansion around the static sector of GR. In particular any static solution[h] of GR will also be an exact (i.e. contrary to infinite series) solution of the effective TTNC gravity theory (by putting all fields other than $\lambda$, $h_{ij}$ to zero). This makes very concrete how the large $c$-expansion is an extension of Newtonian gravity that also contains strong gravitational effects: indeed all the non-linearities of the static sector of GR are included.

## 4. Non-static examples

This interpretation of the large $c$-expansion as an expansion around the static sector of GR can best be explored in some non-static examples.

### 4.1. *Kerr black hole*

The Kerr black hole is a good starting point as it is no longer static but still stationary. The metric[i] contains two physical parameters $M = m\mathrm{G_N}c^{-2}$ $a = J(mc)^{-1} = Ac^{-1}$. We consider the 'strong field'[4] expansion $c \to \infty$ with $M$ and $A$ fixed. Comparing with (8) one finds

$$\lambda = \log\left(1 - \frac{2M}{r}\right) \qquad h_{ij}dx^i dx^j = dr^2 + \left(1 - \frac{2M}{r}\right) r^2 d\Omega^2 \tag{12}$$

$$C_i dx^i = 2\frac{AM}{r}\sin^2\theta \left(1 - \frac{2M}{r}\right)^{-1} d\phi \qquad \Phi = \frac{A^2 M}{r^3}\left(1 - \frac{2M}{r}\right)^{-1} \cos^2\theta. \tag{13}$$

[f]Note that we refrained from spelling out these source terms explicitely here. They can be obtained from Ref. 4 and as was pointed out there one can remove $\beta_{ij}$ from (11) by a gauge transformation. A better, gauge-independent, understanding has emerged from work[5] that appeared after this talk was presented: at one lower order there will appear an additional equation for $\beta_{ij}$ not including other new fields. Together with the equation (11) this will form a closed system determining the spatial dependence of $\Phi$ and $\beta_{ij}$.

[g]This was pointed out by J. Raeymaekers.

[h]The example of the Schwarzschild black hole was worked out in Ref. 4 and extends to other examples such as Tolmann-Oppenheimer-Volkov fluid stars[10].

[i]We follow the notation of Ref. 11.

This example – of which one can check that it solves (9-11) – shows the Kerr black hole as a nonrelativistic expansion around the Schwarzschild black hole, which is a solution to (12)[4]. The non-zero rotation of Kerr – i.e. $A$ – translates to a non-zero $C_\phi$ and $\Phi$. Note that the fields above only express the first terms in an infinite expansion of the Kerr metric. For these first terms to be a good approximation of the relativistic metric one needs the condition

$$r \gg r_+ + \frac{A^2}{Mc^2} + \mathcal{O}(c^{-4})\,, \tag{14}$$

which means that at a distance $\frac{A^2}{Mc^2}$ from the outer horizon the nonrelativistic approximation breaks down. Interestingly the strong field expansion remains valid for large radii where gravity becomes weak and where it overlaps with the post-Newtonian expansion obtained by keeping $m$ and $J$ fixed as $c \to \infty$. This suggests that the TTNC effective theory is a resummation of the post-Newtonian series.

### 4.2. *Oppenheimer-Snyder collapse*

Realistic dynamic situations can be far from stationary, and so it is interesting to explore what additional effects can appear in their large $c$-expansion. To this aim we shortly present some aspects of the nonrelativistic expansion of the Oppenheimer-Snyder solution[12] which describes the collapse of a ball of dust. This solution depends on two parameters: $r_0$, the initial radius of the ball and $M = m\mathrm{G_N}c^{-2}$, its total mass. It is composed of two metrics – an FRW universe inside the ball and the Schwarzshild metric outside – glued together along the worldtube of the ball's surface. This surface follows a radial geodesic in the Schwarschild geometry, see figure 1, and can be described through the curve $r_{\mathrm{match}}(t)$, with $r$ and $t$ the standard Schwarzschild coordinates.

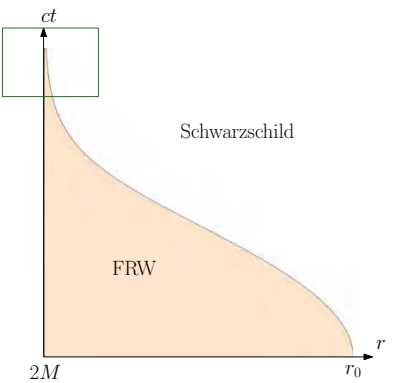



Fig. 1. Space-time of a collapsing ball of dust.

Again we'll be interested in a 'strong field'[4] expansion, where we keep $r_0$ and $M$ fixed as $c \to \infty$, and one finds

$$r_{\text{match}}(t) \approx 2M \left(1 + 2\sqrt{\frac{r_0}{2M} - 1}\, e^{-\frac{ct}{2M}} + \mathcal{O}(e^{-\frac{ct}{M}})\right). \tag{15}$$

Note that the expansion here is of a very different type than the standard Newtonian expansion which holds far away from the ball of dust, just after the collapse has started. In (15) we are close to the event where the ball forms a horizon, both in space and time, the green box in figure 1. There gravity is strong and non-Newtonian, still it is nonrelativistic. That this is the case follows from the well-known – but ever intriguing – fact that an outside observer never sees a horizon form. For such an observer the radial velocity of the shell goes asymptotically to zero again at late times, making it describable by non-relativistic physics. Interestingly though we see that the form of the first correction in (15) is non-perturbative in large $c$. It appears that to describe the nonrelativistic gravitational fields of such a collapsing ball one would need a transseries extension of the ansatz (1).

## Acknowledgments

The author thanks J. Hartong, J. Raeymaekers, and B. Vercnocke for useful discussions and comments. This work was partially supported by the Boğaziçi University Research Fund under grant number 17B03P1 and TUBITAK grant 117F376.

## References

1. E. Poisson and C. Will, *Gravity: Newtonian, Post-Newtonian, Relativistic.* (Cambridge University Press 2014).
2. G. Dautcourt, Acta Physica Polonica B **21** (1991) 755-765, Class. Quant. Grav. **14** (1997) A109 [gr-qc/9610036].
3. W. Tichy and E. E. Flanagan, Phys. Rev. D **84** (2011) 044038 [arXiv:1101.0588].
4. D. Van den Bleeken, Class. Quant. Grav. **34** (2017) no.18, 185004 [arXiv:1703.03459].
5. D. Hansen, J. Hartong and N. A. Obers, Phys. Rev. Lett. **122** (2019) no.6, 061106 [arXiv:1807.04765].
6. M. Cariglia, Phys. Rev. D **98** (2018) no.8, 084057 [arXiv:1811.03446].
7. M. Ergen, E. Hamamcı, D. Van den Bleeken, *to appear.*
8. M. H. Christensen, J. Hartong, N. A. Obers and B. Rollier, Phys. Rev. D **89** (2014) 061901 [arXiv:1311.4794].
9. D. Van den Bleeken and C. Yunus, Class. Quant. Grav. **33** (2016) no.13, 137002 [arXiv:1512.03799].
10. D. Van den Bleeken, *private notes*, D. Hansen, J. Hartong and N. A. Obers, *to appear.*
11. R. M. Wald, *General Relativity.* (Chicago Univeristy Press 1984).
12. J. R. Oppenheimer and H. Snyder, Phys. Rev. **56** (1939) 455.

# Constraints on chameleon dark energy model from SN Ia and BAO observations

Stefano Bellucci*

*INFN - Laboratori Nazionali di Frascati, 1-00044, Frascati Rome, Italy*
** E-mail: Stefano.Bellucci@lnf.infn.it*

Behnaz Fazlpour

*Department of Physics, Babol Branch, Islamic Azad University Babol, Iran*
*E-mail: b.fazlpour@umz.ac.ir*

We have used data sets from Type Ia supernova (SN Ia) and Baryon Acoustic Oscillations (BAO) to place constraints on a chameleon model of dark energy in which tachyon scalar field plays the role of chameleon field. In our analysis the exponential and non-exponential forms for the non-minimal coupling function and tachyonic potential have been assumed.

*Keywords*: Dark energy; Chameleon model; Observational constraints.

## 1. Tachyonic chameleon model

A scalar field with non-minimal coupling to the matter is known as a chameleon field[1,2]. Although, this type of scalar field is heavy in the laboratory environment, on cosmological scale where the matter density is small, the chameleon is light enough to play the role of dark energy.

A chameleon scalar field has many remarkable cosmological features as they pointed out by many authors. For instance, when it couples to an electromagnetic field[3] in addition to the fluid then the fine tuning of the initial condition on the chameleon may be resolved[4]. Chameleon field can also successfully explain a smooth transition from a deceleration to an acceleration epoch for our universe[5]. The scalar field that plays the role of chameleon could be a Brans-Dicke scalar field with interesting cosmological consequences such as explaining the current accelerated expansion of the universe[5–7]. Koury in[8] has summarized some important features of the chameleon field theories. Moreover, phase-space analysis of chameleon scalar field where a quintessence field acts as a chameleon has been investigate in[9]. We have also studied the dynamics of chameleon model where tachyon field plays the role of chameleon[10]. We have utilized the dynamical system tools to obtain the critical points of such theory and showed its interesting cosmological behaviour.

In the present work we apply combined datasets of Type Ia Supernova (SN Ia) and Baryon Acoustic Oscillations to test the tachyonic chameleon model and constrain its parameters. We use $\chi^2$ minimization technique to find the best fit values of the model parameters and to plot the likelihood contours for them.

Tachyonic chameleon model of dark energy is described by the following action

$$S = \int d^4x\sqrt{-g}\Big[\frac{R}{16\pi G} - V(\phi)\sqrt{1-\partial_\mu\phi\partial^\nu\phi} + f(\phi)\mathcal{L}_m\Big], \tag{1}$$

where $g$ is the determinant of the metric tensor $g_{\mu\nu}$, $R$ is the Ricci scalar and $G$ is the bare gravitational constant. $f(\phi)$ is the coupling function between scalar field and the matter. $\mathcal{L}_m$ is the matter Lagrangian and $V(\phi)$ is also the tachyon potential.

Variation of action (1) with respect $g_{\mu\nu}$ in FRW background

$$ds^2 = dt^2 - a^2(dr^2 + r^2d\theta^2 + r^2sin^2\theta d\phi^2), \tag{2}$$

gives the following Friedman equations:

$$3H^2 = \rho_m f + \frac{V(\phi)}{\sqrt{1-\dot{\phi}^2}}, \tag{3}$$

$$2\dot{H} + 3H^2 = -\gamma\rho_m f + V(\phi)\sqrt{1-\dot{\phi}^2}, \tag{4}$$

where $\rho_m$ and $p_m$ are the matter energy density and pressure respectively where $p_m = \gamma\rho_m$.

On the other hand, variation of (1) with respect to the scalar field yields to

$$\ddot{\phi} + (1-\dot{\phi}^2)(3H\dot{\phi} + \frac{V'}{V}) = (1-3\gamma)(1-\dot{\phi}^2)^{\frac{3}{2}}\frac{\rho_m f'}{V}, \tag{5}$$

where a prime stands for derivative with respect to the $\Phi$.

In addition, the continuity equation reads,

$$(\rho_m f)^{\cdot} + 3H(1+\gamma)\rho_m f = -(1-3\gamma)\rho_m\dot{f}. \tag{6}$$

Since, tachyonic chameleon dark energy model exhibits some interesting cosmological implication from dynamical system point of view [10], investigation of such scenario using observational cosmology and constraining the model parameters according to the latest data is not only interesting but also necessary.

## 2. Observational constraints

In this section, we will fit the model parameters with recent observational data from Type Ia Supernova (SN Ia) and Baryon Acoustic Oscillations (BAO) observations.

The total $\chi^2$ for combined data analysis is given by:

$$\chi^2 = \chi^2_{\rm SN} + \chi^2_{\rm BAO}, \tag{7}$$

where

$$\chi^2_{\rm SN} = A - \frac{B^2}{C} \tag{8}$$

$$A = \sum_i \frac{\left[\mu_{\rm obs}(z_i) - \mu_{\rm th}(z_i;\mu_0 = 0)\right]^2}{\sigma_i^2}\,, \tag{9}$$

$$B = \sum_i \frac{\mu_{\rm obs}(z_i) - \mu_{\rm th}(z_i;\mu_0 = 0)}{\sigma_i^2}\,, \tag{10}$$

$$C = \sum_i \frac{1}{\sigma_i^2}\,. \tag{11}$$

$\mu_{obs}$ and $\mu_{th}$ represent the observed and theoretical distance modulus respectively. $\sigma_i$ is also the uncertainly in the distance modulus.

We use (8) for minimization of $\chi^2$ for 580 recent data points of SN Ia [11].

$\chi^2_{\rm BAO}$ is also given by

$$\chi^2_{\rm BAO} = \left(x^{\rm th}_{i,\rm BAO} - x^{\rm obs}_{i,\rm BAO}\right)\left(C^{-1}_{\rm BAO}\right)_{ij}\left(x^{\rm th}_{j,\rm BAO} - x^{\rm obs}_{j,\rm BAO}\right), \tag{12}$$

where the indices $i, j$ are in growing order in $z$, as in Table 1. The WiggleZ collaboration [12] has measured the baryon acoustic scale at three different redshifts, while SDSS and 6DFGS surveys provide data at lower redshift [13].

Table 1. The BAO data used in our analysis.

| | 6dF | SDSS | | WiggleZ | | |
|---|---|---|---|---|---|---|
| $z$ | 0.106 | 0.2 | 0.35 | 0.44 | 0.6 | 0.73 |
| $d_z$ | 0.336 | 0.1905 | 0.1097 | 0.0916 | 0.0726 | 0.0592 |
| $\Delta d_z$ | 0.015 | 0.0061 | 0.0036 | 0.0071 | 0.0034 | 0.0032 |

In addition, $C^{-1}_{\rm BAO}$ was obtained from the covariance data [12] as follows:

$$C^{-1}_{\rm BAO} = \begin{pmatrix} 4444 & 0 & 0 & 0 & 0 & 0 \\ 0 & 30318 & -17312 & 0 & 0 & 0 \\ 0 & -17312 & 87046 & 0 & 0 & 0 \\ 0 & 0 & 0 & 23857 & -22747 & 10586 \\ 0 & 0 & 0 & -22747 & 128729 & -59907 \\ 0 & 0 & 0 & 10586 & -59907 & 125536 \end{pmatrix}. \tag{13}$$

### 2.1. *Four different classes*

Two important functions in the model (1) are the chameleon field potential $V(\phi)$ and coupling function $f(\phi)$. We perform our analysis based on this fact that whether these functions are exponential (power-law) or not. Thus we have four categories given by the following cases,

*case I*: exponential coupling function and power-law potential

$$f(\phi) = f_0 e^{\alpha\phi},$$
$$V(\phi) = V_0\phi^{\beta}. \tag{14}$$

*case II*: exponential potential and a power-law $f(\phi)$

$$f(\phi) = f_0\phi^{\alpha},$$
$$V(\phi) = V_0 e^{\beta\phi}. \tag{15}$$

*case III*: power-law $f(\phi)$ and $V(\phi)$

$$f(\phi) = f_0\phi^{\alpha},$$
$$V(\phi) = V_0\phi^{\beta}. \tag{16}$$

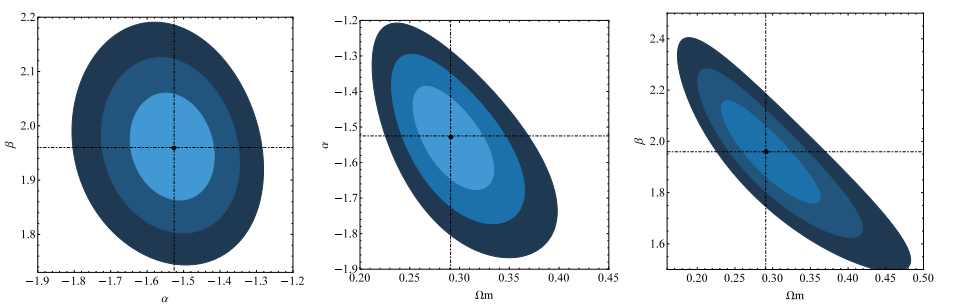
*case IV*: Exponential $f(\phi)$ and $V(\phi)$

$$f(\phi) = f_0 e^{\alpha\phi},$$
$$V(\phi) = V_0 e^{\beta\phi}. \tag{17}$$

We are interested in constraining model parameters $\alpha$ and $\beta$ together with the present values of the density parameters. The best-fit values of $\alpha$, $\beta$ and $\Omega_{m_0}$ in addition to $\chi^2_{min}$ (the minimum value of chi-square) for each cases are presented in Table 2, separately. We have also plotted the likelihood contours for different combinations of the parameters for $1\sigma$ , $2\sigma$ and $3\sigma$ confidence levels in Figures 1-4 in cases $I$ to $IV$ respectively. For simplicity reasons we set $f_0 = V_0 = 1$. As one can see in these figures this model is in agreement with observations.

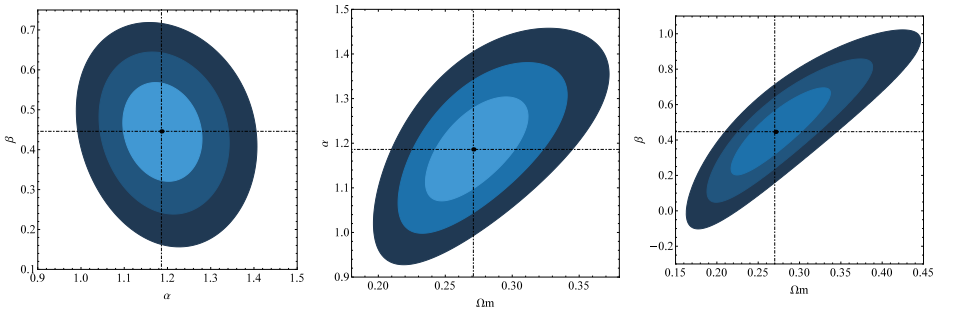
Table 2. The value of $\chi^2_{min}$ and the best fit values of model parameters $\Omega_{m_0}$, $\alpha$ and $\beta$.

| | $\chi^2_{min}$ | $\Omega_{m_0}$ | $\alpha$ | $\beta$ |
|---|---|---|---|---|
| case$I$ | 584.08 | 0.29 | -1.53 | 1.96 |
| case$II$ | 584.08 | 0.270 | 1.18 | 0.44 |
| case$III$ | 584.08 | 0.30 | 1.23 | -0.44 |
| case$IV$ | 584.09 | 0.31 | -1.58 | 2.25 |



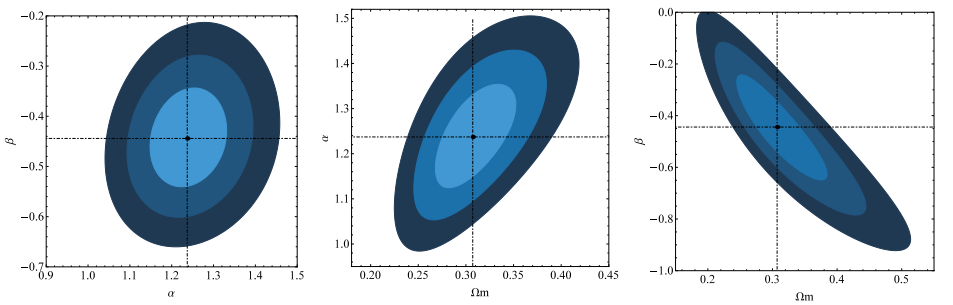

(a) (b) (c)

Fig. 1. Plots of $1\sigma$ (light blue), $2\sigma$ (blue) and $3\sigma$ (dark blue) confidence contours on $\alpha-\beta$ (a), $\alpha-\Omega_{m_0}$ (b) and $\beta-\Omega_{m_0}$ (c) parameter spaces for SN Ia+BAO datasets in the case $I$. The black dot in each plot shows the best fit point.



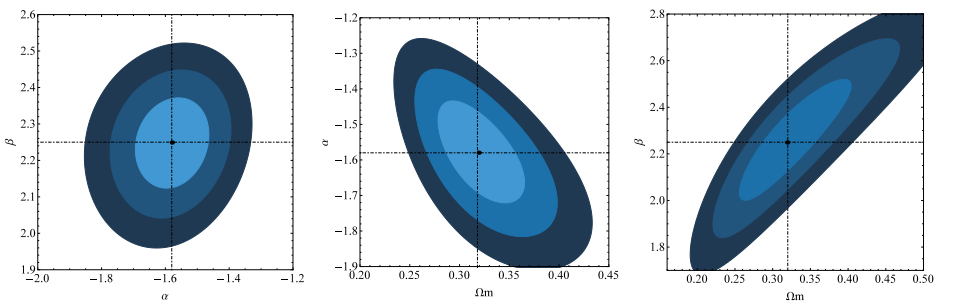

(a) (b) (c)

Fig. 2. Plots of $1\sigma$ (light blue), $2\sigma$ (blue) and $3\sigma$ (dark blue) confidence contours on $\alpha - \beta$ (a), $\alpha - \Omega_{m_0}$ (b) and $\beta - \Omega_{m_0}$ (c) parameter spaces for SN Ia+BAO dataset in the case $II$. The black dot in each plot shows the best fit point.




(a) (b) (c)

Fig. 3. Plots of $1\sigma$ (light blue), $2\sigma$ (blue) and $3\sigma$ (dark blue) confidence contours on $\alpha - \beta$ (a), $\alpha - \Omega_{m_0}$ (b) and $\beta - \Omega_{m_0}$ (c) parameter spaces for SN Ia+BAO dataset in the case $III$. The black dot in each plot shows the best fit point.




(a) (b) (c)

Fig. 4. Plots of $1\sigma$ (light blue), $2\sigma$ (blue) and $3\sigma$ (dark blue) confidence contours on $\alpha - \beta$ (a), $\alpha - \Omega_m$ (b) and $\beta - \Omega_{m_0}$ (c) parameter spaces for SN Ia+BAO dataset in the case $IV$. The black dot in each plot shows the best fit point.

## 3. Conclusion

Tachyonic chameleon dark energy model has the ability to alleviate the coincidence problem via the mechanism of scaling attractors[10]. In this paper we have used the latest observational data to constrain the parameters of this model. Two important functions in our analysis are tachyonic potential $V(\phi)$ and non-minimal coupling function $f(\phi)$ in action (1). In general we have considered two types of these functions i.e power-law and exponential forms.

We have fitted data from Type Ia supernova (SN Ia) and Baryon Acoustic Oscillation (BAO) to constrain the present matter density parameter $\Omega_{m_0}$, the parameter $\alpha$ in functional form of $f(\phi)$ (the non-minimal coupling function $f$ is of the form $f(\phi) \propto e^{\alpha\phi}$ or $f(\phi) \propto \phi^{\alpha}$) and the parameter $\beta$ in $V(\phi)$ (tachyonic potential is assumed to be $V(\phi) \propto e^{\beta\phi}$ or $V(\phi) \propto \phi^{\beta}$). For exponential $f(\phi)$ and power-law $V(\phi)$ we have seen that positive $\alpha$ and negative $\beta$ are favoured by the data while for power-law $f(\phi)$ and exponential $V(\phi)$ both $\alpha$ and $\beta$ should be positive in order to our model be compatible with observations. On the other hand, when $f(\phi)$ and $V(\phi)$ are both power-law functions of scalar field a positive value of $\alpha$ and a negative value of $\beta$ are favoured while when these functions are both in exponential forms then $\alpha$ is negative and $\beta$ is a positive constant. We remark that in all four cases the value of present matter density parameter $\Omega_{m_0}$ is very close to the desired value.

In summay, the scenario of the tachyonic chameleon dark energy is compatible with observations, for all examined scalar field potential and non-minimal coupling functions.

## References


1. S. Das and N. Banerjee, *Phys. Rev. D.* **78**, 043512 (2008).
2. Y. Bisabr, *Phys. Rev. D.* **86**, 127503 (2012).
3. J. Khoury, A. Weltman, et al., *Phys. Lett. B.* **699**, 5 (2011).
4. D. F. Mota and C. A. O. Schelpe, *Phys. Rev. D.* **86**, 123002 (2012).
5. N. Banerjee, S. Das and K. Ganguly, *Pramana.* **74**, L481 (2010).
6. J. Khoury, A. Weltman, et al., *Phys. Rev. D.* **70**, 123518 (2004).
7. S. Das, P. S. Corasaniti and J. Khoury, *Phys. Rev. D.* **73**, 083509 (2006).
8. J. Khoury, *Class. Quant. Grav.* **30**, 214004 (2013).
9. N. Roy, N. Banerjee, *Ann. Phys.* **452**, 356 (2015).
10. A. Banijamali and M. Solbi, *Gen Relativ Gravit.* **49**, 103 (2017).
11. N. Suzuki, D. Rubin, Lidman, et al., *Astrophys. J.* **746**, 85 (2012).
12. C. Blake, E. Kazin, et al., *Mon. Not. Roy. Astron. Soc.* **418**, 1707 (2011).
13. W. J. Percival, et al., *Mon. Not. Roy. Astron. Soc.* **401**, 2148 (2010).

# Density perturbations in $f(R, \phi)$-gravity in general with an application to the (varying power)-law model*

Fayçal Hammad

*Department of Physics and Astronomy & STAR Research Cluster, Bishop's University*
*2600 College Street, Sherbrooke, QC, J1M 1Z7 Canada*
*Physics Department, Champlain College-Lennoxville*
*2580 College Street, Sherbrooke, QC, J1M 0C8 Canada*
* *E-mail: fhammad@ubishops.ca*

Density perturbations in the cosmic microwave background within general $f(R, \phi)$ models of gravity are presented. The general dynamical equations for the tensor and scalar modes in any $f(R, \phi)$ gravity model are provided. An application of the equations to the (varying power)-law modified gravity toy-model is then given. Based on the latest observations of the density perturbations in the sky, the model requires inflation to occur at an energy scale less than the GUT-scale ($10^{14}$ GeV). The density perturbations obtained from observations are recovered naturally, with very high precision, and without fine tuning the model's parameters.

*Keywords*: Modified gravity; Inflation; Density perturbations; (Varying power)-law model.

## 1. Introduction

Some of the great problems of modern cosmology are the problem of dark energy, the problem of dark matter and the problem of the early Universe. A widely pursued axis of research to solve these problems consists in modifying general relativity (GR). The easiest way to modify GR is to replace the Ricci scalar $R$ in the gravitational action by an arbitrary function of $R$ and/or to add a possible extra degree of freedom for spacetime, besides the metric, in the form of a scalar field $\phi$. Such modified GR models are known as $f(R, \phi)$ gravity models. In order to understand the motivation behind the peculiar (varying power)-law model, let us briefly draw a timeline of some of the well-known models of $f(R, \phi)$ modified gravity.

## 2. A brief history of $f(R, \phi)$ gravity

Here is a very simplified sketch of the various $f(R, \phi)$ models and their evolution in time.

*This talk is based on the published paper 1.

$$S \sim \int \mathrm{d}^4x\sqrt{-g}R \quad \text{Hilbert (1915).}$$

$$S \sim \int \mathrm{d}^4x\sqrt{-g}\left(\phi R - \frac{\omega(\phi)}{\phi}\partial_\mu\phi\partial^\mu\phi - V(\phi)\right) \quad \begin{cases}\text{Thiry (1948), Jordan (1955),}\\ \text{Brans-Dicke (1961),}\\ \text{Bergman (1968),}\\ \text{Nordtveldt (1970),}\\ \text{Wagoner (1970),}\\ \text{O'Hanlon (1972),}\\ \text{Horndeski (1974, complicated),}\\ \text{Bekenstein (1977),}\\ \text{Barker (1978).}\end{cases}$$

$$S \sim \int \mathrm{d}^4x\sqrt{-g}f(R) \quad \text{Buchdahl (1970).}$$

$$S \sim \int \mathrm{d}^4x\sqrt{-g}\left(\alpha_0 + \alpha_1 R + \alpha_2(R^2 - 4R_{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}\right) \quad \text{Lovelock (1971).}$$

$$S \sim \int \mathrm{d}^4x\sqrt{-g}\left(R + \frac{R^2}{6M_P^2}\right) \quad \text{Starobinski (1980).}$$

$$S \sim \int \mathrm{d}^4x\sqrt{-g}\left(R + f_1R^2 + f_2R_{\mu\nu}R^{\mu\nu} + f_3R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}\right) \quad \text{Stelle's gravity (1980).}$$

$$S \sim \int \mathrm{d}^4x\sqrt{-g}\left(R + RF_1(\Box)R + R_{\mu\nu}F_2(\Box)R^{\mu\nu} + C_{\mu\nu\rho\rho}F_3(\Box)C^{\mu\nu\rho\sigma}\right) \quad \begin{cases}\text{Biswas-}\\ \text{Gerwick-}\\ \text{Koivisto}\\ \text{(1980).}\end{cases}.$$

The tensor $C_{\mu\nu\rho\sigma}$ consists of the more complicated invariant $\frac{1}{3}R^2 - 2R_{\mu\nu}R^{\mu\nu} + R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$. These expressions arise from the study of the quantum trace anomaly. A pragmatic trend has appeared more recently, however, which consists in trying out an arbitrary function of $f(R)$ inside the Lagrangian such that it could a priori be Taylor-expanded in $R$: $f(R) = \alpha_0 + \alpha_1 R + \alpha_2 R^2 \ldots$ Such models are called $f(R)$ gravity. Below is a list of the most studied ones,

$$f(R) = R - \frac{c}{(R-\Lambda_1)^n} + b(R-\Lambda_2)^m \qquad \text{Nojiri \& Odintsov (2003),}$$

$$f(R) = R + \alpha\ln\left(\frac{R}{\mu^2}\right) + \beta R^m \quad \text{Nojiri \& Odintsov (2004).}$$

$$f(R) = R + \gamma R^{-n}\left(\frac{R}{\mu^2}\right)^m \quad \text{Meng \& Wang (2004).}$$

$$f(R) = \frac{\alpha R^{2n} - \beta R^n}{1 + \gamma R^n} \quad \text{Nojiri \& Odintsov (2007).}$$

The various powers in these models are constrained by solar-system observations one one hand, and the need to reproduce both the early expansion of the universe and the late-time cosmic acceleration on the other. Some models, however, require different (real and fractional) values for the powers whether one applies them to reproduce inflation or the late-time expansion. This suggests that one might as well let the power of the Ricci scalar in the gravitational action to vary in space and time, so that it would adjust itself depending on the epoch in the evolution of the Universe and on the cosmological environment. This leads to a (*varying power*)law model, in which the power of the Ricci scalar is a spacetime-dependent scalar field:

$$S \sim \int \mathrm{d}^4x\sqrt{-g}\left[R - \frac{\mu^2}{2}\left(\frac{R}{\mu^2}\right)^{\phi}\right]. \tag{1}$$

An important remark to make here is that, unlike $f(R)$ gravity models, the redefinitions $\partial_R f = \phi$ and $(2f - R\partial_R f)/3 = dV/d\phi$ do not transform the (varying power)-law model into GR with a scalar field. this is due to the high non-linearity of the model. On the other hand, a Weyl transformation $g_{\mu\nu} \to \phi^2 g_{\mu\nu}$, followed by the field redefintion $\phi^2 \sim \sigma^{-1}$, does transform the model into a Brans-Dicke like theory, in which matter also acquires a coupling with the scalar field $\sigma$, though.

## 3. Scalar modes in any $f(R, \phi)$ model

Start from the field equations of an $f(R, \phi)$ model.[2] Then, perturb the equations using $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$, with $\bar{g}_{\mu\nu} = diag(-1, a^2, a^2, a^2)$ is the Friedmann-Lemaître-Robertson-Walker (FLRW) background metric and $|h_{\mu\nu}| \ll 1$, as well as $\phi = \bar{\phi} + \delta\phi$. Then, choose the Newtonian gauge $\mathrm{d}s^2 = -(1+2\Phi)\mathrm{d}t^2 + a^2(t)(1-2\Psi)\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j$. The dynamical equation for the gravitational potentials $\Psi$ and $\Phi$ is,[1]

$$\begin{aligned}
&\ddot{\Phi} + \ddot{\Psi} + \left(H - \frac{2\ddot{\phi}}{\dot{\phi}} + \frac{3\dot{f}_{,R}}{f_{,R}}\right)(\dot{\Phi} + \dot{\Psi}) - \left(4\dot{H} - \frac{4\ddot{f}_{,R}}{f_{,R}} + \frac{6\dot{f}_{,R}}{f_{,R}}\frac{\ddot{\phi}}{\dot{\phi}}\right)\Phi \\
&+ \left[6\dot{H} + 6H^2 + H\left(\frac{4\dot{f}_{,R}}{f_{,R}} - \frac{2\ddot{\phi}}{\dot{\phi}}\right) - \frac{f - 2\Box f_{,R}}{f_R} - \frac{\ddot{f}_{,R}}{f_{,R}} + \frac{2\dot{f}_{,R}}{f_{,R}}\frac{\ddot{\phi}}{\dot{\phi}} - \frac{\dot{\phi}^2}{f_{,R}} - \frac{\nabla^2}{a^2}\right](\Phi + \Psi) = 0.
\end{aligned} \tag{2}$$

## 4. Tensor modes in any $f(R, \phi)$ model

In order to extract the tensor modes equation, perturb the FLRW background, $\mathrm{d}s^2 = -(1+2\Phi)\mathrm{d}t^2 + \left[a^2(t)(1-2\Psi)\delta_{ij} + h_{ij}\right]\mathrm{d}x^i\mathrm{d}x^j$ and impose the transverse gauge, $\partial_i h^i_j = 0$, with also $h^i_i = 0$. The transverse modes' equation is,[1]

$$\left[\partial_0^2 - \left(H - \frac{\dot{f}_{,R}}{f_{,R}}\right)\partial_0 + 4H^2 - \frac{f - 2\Box f_{,R}}{f_{,R}} - \frac{\nabla^2}{a^2}\right]h_{ij} = 0. \tag{3}$$

## 5. The (varying power)-law model

The action of the model is,[3]

$$S = \frac{M_P^2}{16\pi}\int \mathrm{d}^4x\sqrt{-g}\left[R - (\partial\phi)^2 - m^2\phi^2 - \frac{\mu^2}{2}\left(\frac{R}{\mu^2}\right)^{\phi}\right], \tag{4}$$

where $M_P$ is the Planck mass and $m$, whose order of magnitude is not constrained by the model, is the scalar field's mass. The energy scale $\mu$ is determined by observation and it is found that it should be $\sim 10^{14}$ GeV. At the beginning of inflation $R \sim \mu^2$, so that the action approximates to $S \sim \int \mathrm{d}^4x\sqrt{-g}(R-(\partial\phi)^2)$. At late times, $R \ll \mu^2$, so that the action approximates to, $S \sim \int \mathrm{d}^4x\sqrt{-g}(R-(\partial\phi)^2-m^2\phi^2-\frac{1}{2}\mu^2(R/\mu^2)^{\phi})$, in which the last term could be identified with a cosmological constant, while the scalar field could be identified with dark matter. Thanks to the effective potential,

$$V(R,\phi) = \frac{m^2}{2}\phi^2 + \frac{\mu^2}{4}\left(\frac{R}{\mu^2}\right)^{\phi}, \tag{5}$$

of the scalar field, the mass of the latter is found to depend on the curvature of the environment.[3] In addition, the radiative corrections show that $m_{\mathrm{eff}}$ could be huge for very low curvatures.[3,4]

It turns out that this model easily finds a very nice application for investigating the early Universe.[1] The slow-roll parameters of the model are,

$$\epsilon \equiv \frac{|\dot{H}|}{H^2} \ll 1, \qquad \eta \equiv \frac{|\dot{\phi}|}{H} \ll 1. \tag{6}$$

In usual inflationary models, where the Hubble parameter is given by $H^2 \sim V(\phi)$, the analogue of the $\eta$-condition is $\dot{\phi}^2/H^2 \ll 1$, which, thanks to the scalar field's equation of motion, is just equivalent to the $\epsilon$-condition in (6). However, in the model (4) the Hubble parameter is not simply given by $H^2 \sim V(\phi)$ and, therefore, no condition should *a priori* be imposed on $\dot{\phi}$. However, it turns out that the $\eta$-condition is automatically satisfied throughout all the duration of an inflation. The latter starts in this model at an energy scale $\mu$ and ends at another energy scale not very far off. The end of inflation occurs at that energy scale for which either both or one of the slow-roll parameters in (6) ceases to be small and negligible. It is found that both slow-roll conditions (6) are satisfied up to the scale $R/\mu^2 \sim 10^{-0.05}$. The number of $e$-folds $N$ in this model can be deduced via the formula,[1]

$$N \simeq \frac{13}{4}\ln\left(\frac{\ln\rho_f}{\ln\rho_i}\right). \tag{7}$$

For example, if one assumes inflation to have ended at the energy scale of $\rho \simeq 10^{-0.05}$, then one should allow inflation to start at the energy scale of $\rho \simeq 10^{-10^{-8}}$, in which case one finds, $N \simeq 50$.[1] To have a bigger number of $e$-folds one needs to push back the starting of inflation to scales which are even closer to the $\mu$-scale, if not exactly equal to $\mu$.

### 5.1. *Scalar modes in the (varying power)-law model*

Using the scalar modes equations (2), one easily extracts for the model the following $\alpha_s$, $\nu_s$ and $\beta$ are given by,[1]

$$\alpha_s = \frac{4}{11} - \frac{3\epsilon}{22} + \frac{1501\eta}{968}, \quad \nu_s^2 = \frac{225}{121} + \frac{285\epsilon}{121} + \frac{145555\eta}{21296},$$
$$\beta = 1 + \frac{13\eta}{22}. \tag{8}$$

### 5.2. *Tensor modes in the (varying power)-law model*

Using the tensor modes equations (3), one easily extracts for the model the following $\alpha_t$, $\nu_t$,[1]

$$\alpha_t = \frac{1}{2} - \frac{\eta}{4} \qquad \text{and} \qquad \nu_t^2 = \frac{9}{4} + 3\epsilon - \frac{11\eta}{4}. \tag{9}$$

### 5.3. *The power spectra in the (varying power)-law model*

At Hubble-crossing one finds,[1]

$$r \simeq \eta_*^2, \tag{10}$$

where $\eta_*$ is the slow-roll parameter in (6) evaluated at the Hubble crossing. The scalar and tensor tilts $n_s$ and $n_t$ can be computed from the scalar and tensor power spectra $\mathcal{P}_\mathcal{R}$ and $\mathcal{P}_h$, respectively, as,[1]

$$n_s - 1 = \left.\frac{d\ln\mathcal{P}_\mathcal{R}}{d\ln k}\right|_{k=k_*} = 2 + 2\alpha_s - 2\nu_s, \tag{11}$$

$$n_t = \left.\frac{d\ln\mathcal{P}_h}{d\ln k}\right|_{k=k_*} = 2 + 2\alpha_t - 2\nu_t. \tag{12}$$

One can compute the values of $r$, $n_s$ and $n_t$ in (10), (11) and (12), respectively, for various energy scales comprised between the energy scales $\rho \sim 10^{-10^{-8}}$ and $\rho \sim 10^{-0.05}$. One easily finds that the scale for the Hubble crossing that matches the best with observations is comprised between $\rho \sim 10^{-0.008}$ and $\rho \sim 10^{-0.01}$ for which the values of $r$, $n_s$, and $n_t$ found using expressions (8) and (9) for $\alpha_s$, $\nu_s$, $\alpha_t$ and $\nu_t$ are: $r \in [0.00035, 0.00055]$, $n_s \in [0.965, 0.972]$, and $n_t \in [0.031, 0.040]$.[1]

The specific scale that agrees best with the observed values is found to be $\rho \sim 10^{-0.009}$. Indeed for this scale a computation using again expressions (8) and (9) for $\alpha_s$, $\nu_s$, $\alpha_t$ and $\nu_t$ gives the following results:[1]

$$r \simeq 0.00044, \quad n_s \simeq 0.969, \quad n_t \simeq 0.036. \tag{13}$$

Here, one uses the value of the scalar tilt to find the best match. In fact, the value of the ratio $r$ is only known by its upper boundary value which is $r < 0.1$[5]. Notice that the tensor tilt here is positive in contrast to usual single-field inflationary models

where it comes out negative. The other major difference is that the model predicts a very small value for the tensor-to-scalar ratio $r$.

One can also deduce the energy scale for inflation required by the model after making use of the relation between the scalar power spectrum $\mathcal{P}_{\mathcal{R}}$ and its amplitude $A_S$. The latest observations show that the amplitude of the scalar power spectrum at the Planck pivot scale $k_*$ is $A_S \sim 2 \times 10^{-9}$.[5] This implies that the inflation scale according to the model should be around $\sim 10^{14}\,\mathrm{GeV}$.

## 6. Conclusion

The (varying power)-law model cannot be motivated from particle physics. It has only two free parameters, though, and does not require any fine-tuning. It might unify dark energy, dark matter and the early expansion of the Universe.

## Acknowledgements

This work is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant (RGPIN-2017-05388), as well as by the STAR Research Cluster of Bishop's University.

## References

1. F. Hammad, "Density perturbations in $f(R, \phi)$-gravity with an application to the (varying power)-law model", *Phys. Rev. D* **96**, 064006 (2017) [arXiv:1709.02276].
2. A. De Felice and S. Tsujikawa, "$f(R)$ theories", *Living Rev. Rel.* **13**, 3 (2010) [arXiv:1002.4928].
3. F. Hammad,"A (varying power)-law modified gravity" *Phys. Rev. D* **89**, 044042 (2014) [arXiv:1403.0261].
4. F. Hammad, "Radiative corrections in the (varying power)-law modified gravity", *Phys. Rev. D* **91**, 124027 (2015) [arXiv:1505.06023].
5. P. A. R. Ade et al. [Planck Collaboration], "*Planck 2015 results. XX. Constraints on inflation*", *A&A* **594**, A20 (2016) [arXiv:1502.02114].

# Dark matter with pressure as an alternative to dark energy

Orlando Luongo* and Marco Muccino

*Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, 00044 Frascati, Italy*
**E-mail: orlando.luongo@lnf.infn.it*

We revise the cosmological concordance paradigm assuming dark matter shows non-vanishing pressure. We propose a Lagrangian for matter field with a Lagrange multiplier as constraint. In addition, we add a symmetry breaking effective potential accounting for the classical cosmological constant problem. To face out the fine-tuning issue, we investigate two phases: before and after transition due to the symmetry breaking. Investigating the conserved Noether current due to the *shift symmetry*, we show that our Lagrangian does not depend upon the scalar field $\varphi$ before and after transition. We propose that during the transition the dark matter sector cancels out the effects due to quantum field vacuum energy. This process induces a positive increase of Helmotz's free-energy which naturally leads to a negative pressure, which turns out to be constant since the sound speed naturally vanishes. The order of magnitude of the pressure is determined by visible matter only or by the bare cosmological constant. The numerical bounds over the pressure and matter densities are in agreement with current observations, alleviating the coincidence problem. Finally assuming a thermal equilibrium between the bath and our effective matter fluid, we estimate the mass associated to the dark matter candidate. Our bounds seem to be compatible with the most recent predictions over WIMP masses for fixed spin and temperature.

*Keywords*: The cosmological constant problem; dark energy; dark matter.

## 1. Introduction

It is now a controversial and debatable fact that the cosmic speed up is induced by a vacuum energy cosmological constant $\Lambda$. Quantum fluctuations lead to a huge amount of energy which severely differs from observations. Although $\Lambda$ is the most accepted possibility, it does not represent the unique alternative to describe present-time acceleration[1]. What cancels $\Lambda$ out leads to the well-known *cosmological constant problem* and may suggest the $\Lambda$CDM model could be theoretically incomplete[2], requiring stunning revisions[a].

In the standard paradigm, cold dark matter and baryons are pressureless. We here show that assuming an effective a scalar field Lagrangian for matter with *non-vanishing pressure* it is possible to address the cosmological constant problem if a cancellation mechanism is included. The underlying idea is to assume that during

[a]It is commonly believe that the introduction of *dark energy* represents a possible strategy. As the Universe expands the net fluid gets less dense, whereas $\Lambda$ remains constant over time. Dropping this out, dark energy leads to an evolving equation of state. However, although appealing, this scenario does not explain why quantum fluctuations are not included into Einstein's energy-momentum tensor. Nonetheless, our $\Lambda$-dominated Universe fixes the cosmological constant magnitude to be overwhelmingly close to matter, providing a further issue: *the coincidence problem*.

an early-time phase transition, dark matter pressure counterbalances the effects of $\Lambda$, leaving as unique contribute the pressure of baryons[b], as the transition ends[3].

In this paper, we discuss: 1) under which conditions pressure can be negative without putting it by hand, 2) which mechanism cancels $\Lambda$ out, 3) which predictions on dark matter candidates can be obtained and 4) how thermodynamics and small perturbations fix bounds on our model. Since we show our process does not set $\Lambda$ to zero, but partially removes it in a natural way, we end up there is no need of dark energy and/or modified gravity contributions to accelerate the Universe today. We only invoke that baryons may have non-zero pressure or that the bare cosmological constant can be the only byproduct as the transition ends.

The paper is structured as follows: In section II we discuss the effective field representation of our model. In section III, we discuss thermodynamics and small perturbations. Afterwards we show the predictions of our model in section IV and finally we conclude with our final remarks and perspectives, in section V.

## 2. Effective field representation of matter with pressure

Let us invoke a few number of basic hypotheses and possible clues on the problems plaguing the concordance paradigm: 1) dark energy does not evolve and its existence is not due to quantum fluctuations, but only to *one matter fluid* with non-vanishing pressure, 2) the coupling between matter and $\Lambda$ can produce a mechanism of cancellation for quantum fluctuations, 3) a phase transition enables the aforementioned cancellation, 4) the whole energy is constrained through a Lagrange multiplier, i.e. the fluid flows along geodesics, but with non-zero pressure, 5) thermodynamics *naturally* fixes an *emergent negative pressure* capable of accelerating the Universe. In this respect, the model is thought to degenerate with the $\Lambda$CDM paradigm, although providing different predictions[4].
This suggests

$$\mathcal{L} = K\left(X,\varphi\right) + \lambda Y\left[X,\nu\left(\varphi\right)\right] - V^{\text{eff}}\left(X,\varphi\right) \tag{1}$$

where the scalar field $\varphi$ is associated to *matter only* and $X = \frac{1}{2}g^{\alpha\beta}\nabla_\alpha\varphi\nabla_\beta\varphi$. Here, $\nu(\varphi)$ plays the role of the specific inertial mass and $V^{\text{eff}}$ takes arbitrary forms which produce a symmetry breaking transition. Assuming thermal equilibrium, the simplest possibility is: $V^{\text{eff}}(X,\varphi) = V_0 + \frac{\chi}{4}\left(\varphi^2-\varphi_0^2\right)^2 + \frac{\chi}{2}\varphi_0^2\varphi^2\frac{T^2(X)}{T_c^2}$, with $T_{\text{c}} = \varphi_0\sqrt{\chi/\bar{g}}$ the critical temperature.
The corresponding equations of motion are

$$Y = 0\,, \tag{2}$$

$$\mathcal{L}_\varphi - \nabla_\alpha\left(\mathcal{L}_X\nabla^\alpha\varphi\right) = 0\,, \tag{3}$$

$$T_{\alpha\beta} = \mathcal{L}_X\nabla_\alpha\varphi\nabla_\beta\varphi - \left(K - V^{\text{eff}}\right)g_{\alpha\beta}\,, \tag{4}$$

[b]For the sake of completeness, our model leaves open the possibility that the effective contribution to the net pressure is due to the bare cosmological constant, leaving baryons to be pressureless and dark matter to have a primordial non-vanishing pressure term.

and immediately one gets:

$$\rho\left(\lambda, X, \varphi\right) = 2X\mathcal{L}_X - \left(K - V^{\text{eff}}\right) , \tag{5}$$

$$P\left(X, \varphi\right) = K - V^{\text{eff}} . \tag{6}$$

Non-dissipative fluids are puzzled through the Carter's covariant formulation, in which coordinates of each matter space serve as labels, distinguishing fluid element worldlines[5]. Each scalars can be viewed as Stückelberg fields and are framed with four scalar fields. Hence, all thermodynamic properties are encoded via precise action's symmetries. Below we discuss which symmetries can be used for our purposes[6].

## 3. Thermodynamics and small perturbations

Providing the cosmological principle, the global shift symmetry[7] applies as

$$\varphi \to \varphi + c^0 , \tag{7}$$

with $c^0$ an arbitrary constant, on the zero-component of our fluid only.

To account for perfect fluid thermodynamics we use the first principle, the Gibbs-Duhem relation and the Helmotz free-energy density and we infer that if the Helmotz free energy variation is positive, i.e. $df > 0$, one gets $-PdV > 0$, or alternatively that $P < 0$. Immediately one finds: the internal energy $u = \rho = 2X\mathcal{L}_X - \left(K - V^{\text{eff}}\right)$, the enthalpy $h = u + P = 2X\mathcal{L}_X$, and the Gibbs free-energy $g = f + P = 0$.

It is remarkable to stress that before transition the effective potential remains constant, although after transition: the effective potential becomes a function of $X$, i.e. invariant under global shift symmetry. The study of our model during transition is beyond the purposes of this work, thus we limit our analysis when we are in proximity of minima. Therefore, invoking the Noether's theorem, from the matter Lagrangian part we find a first conserved current:

$$\mathcal{J}_1^\alpha = \sqrt{2X}\left(K_X + \lambda Y_X\right) v^\alpha . \tag{8}$$

while from the potential part only, we define another conserved current:

$$\mathcal{J}_2^\alpha = -\sqrt{2X} V_X^{\text{eff}} v^\alpha . \tag{9}$$

Hence one can see that

$$\mathcal{J}^\alpha \equiv \mathcal{J}_1^\alpha + \mathcal{J}_2^\alpha = \sqrt{2X}\mathcal{L}_X v^\alpha = s^\alpha . \tag{10}$$

The former coincides with the entropy density current $s_\alpha = s v_\alpha$.

It is easy to see that the fluid is irrotational and insentropic and, in the conformal Newtonian gauge, the density perturbations are functions of the $X$ and of the Lagrange multiplier perturbations, whereas pressure perturbations depend only on $\delta X$. We easily obtain:

$$\delta\rho = A(X)\delta X + 2X\lambda Y_{X\nu}\delta\nu + 2XY_X\delta\lambda , \quad \delta P = B(X)\delta X . \tag{11}$$

where $A(X) = (2X\mathcal{L}_X)_X - K_X$ and $B(X) = K_X$. After some manipulations, the entropy perturbation shift, $\Delta = \left(\frac{\delta P}{\delta \rho} - c_{\rm s}^2\right)\frac{\delta \rho}{P} = -\frac{D(X)\delta\nu + E(X)\delta\lambda}{P}$ leads to $c_{\rm s}^2 \equiv 0$ and $P = \text{const}$, i.e. showing a vanishing adiabatic sound speed[8–10]. The functions $D(X)$ and $E(X)$, in particular, are:

$$D(X) = 2X\lambda Y_{X\nu} c_X^2 \,, \tag{12}$$

$$E(X) = 2X\lambda Y_X c_X^2 \,. \tag{13}$$

## 4. Solving the cosmological constant problem

Shortly summarizing the cosmological constant problem is due to the ambiguity over the two possible choices of the off-set $V_0$, which leads to profoundly different physical puzzles[11]. The two possibilities are:

1) $V_0 = -\chi\varphi_0^4/4$. This implies that as the transition finishes $V^{\rm eff} = 0$. We have:

$$P_1 = \begin{cases} K & \text{(BT)} \\ K + \chi\varphi_0^4/4 & \text{(AT)} \end{cases} , \tag{14}$$

$$\rho_1 = \begin{cases} 2X\lambda Y_X - K & \text{(BT)} \\ 2X\lambda Y_X - K - \chi\varphi_0^4/4 & \text{(AT)} \end{cases} , \tag{15}$$

with $K < -\chi\varphi_0^4/4$.

2) $V_0 = 0$. Now we have that before the transition: $V^{\rm eff} = 0$. One has:

$$P_2 = \begin{cases} K - \chi\varphi_0^4/4 & \text{(BT)} \\ K & \text{(AT)} \end{cases} , \tag{16}$$

$$\rho_2 = \begin{cases} 2X\lambda Y_X - K + \chi\varphi_0^4/4 & \text{(BT)} \\ 2X\lambda Y_X - K & \text{(AT)} \end{cases} , \tag{17}$$

so $K < 0$.

In both the aforementioned cases, one has that $K$ is negative, but with different magnitudes. On the one hand, it is possible to demonstrate that $V_0 = 0$ corresponds to the ΛCDM model, plagued by the issues of coincidence and fine tuning. On the other hand, we propose that:

$$K_{\rm DM} \approx -\chi\varphi_0^4/4 \,, \tag{18}$$

$$K_{\rm BM} \ll K_{\rm DM} \,, \tag{19}$$

and easily one infers

$$P_1 \approx \begin{cases} K_{\rm DM} & \text{(BT)} \\ K_{\rm BM} & \text{(AT)} \end{cases} , \tag{20}$$

$$\rho_1 \approx \begin{cases} (\rho_{\rm DM} + \rho_{\rm BM})(1+z)^3 - K_{\rm DM} & \text{(BT)} \\ (\rho_{\rm DM} + \rho_{\rm BM})(1+z)^3 - K_{\rm BM} & \text{(AT)} \end{cases} , \tag{21}$$

where $\rho_{\rm BM} = 2X\lambda_0 Y_{{\rm BM},X}$, $\rho_{\rm DM} = 2X\lambda_0 Y_{{\rm DM},X}$ and $a = \frac{1}{1+z}$.

### 4.1. *The mechanism for cancelling Λ out*

The mechanism of cancellation is defined in Eq. (21). It elides vacuum energy contribution through dark matter. In fact, $\chi > 0$ implies that $K_{\rm DM}$ has opposite sign with respect to vacuum energy. It remains the question: Which fluid contributes to accelerate the Universe today? The answer is at the end of transition: the Universe accelerates *because of the presence of a negative baryonic pressure or in terms of a bare cosmological constant which remains non-zero.* Since the pressure is constant, the corresponding *emergent cosmological constant* is represented by a fluid whose equation of state evolves by $\omega^{-1} \propto \rho$, mimicking a dark fluid, but with no coincidence and fine-tuning[3].

The mechanism works only if dark matter constituents are appropriately chosen. In the simplest case of thermal equilibrium, the energy density of all bosons (b) and fermions (f) becomes

$$\epsilon_{\rm BT} = \left[\Omega_{\rm r}\left(\frac{T_{\rm p}}{T_0}\right)^4 + \Omega_{\rm m}\left(\frac{T_{\rm p}}{T_0}\right)^3 + \Omega_\Lambda\right]\epsilon_{\rm c} + \epsilon_{\rm v}\,, \tag{22}$$

where $\Omega_{\rm r} = (9.16 \pm 0.19) \times 10^{-5}$ is the radiation density parameter, and $\epsilon_{\rm c} = \rho_{\rm c} c^2$, where $\rho_{\rm c} = 3H^2/(8\pi G)$ is the Universe critical density, in which $H$ is the Hubble parameter and $G$ the gravitational constant. The plasma temperature is $T_{\rm p} = (6.6559 \pm 0.0019) \times 10^{14}\ K$.

In the standard picture, primordial dark matter interactions are due to annihilation of a heavier particle $Q$ with its antiparticle $\bar{Q}$. Assuming a mass $M$ and a production of two lighter particles, say $q$ and $\bar{q}$, it is licit to invoke symmetry between $Q$ and $\bar{Q}$ and so: $n_Q \equiv n_{\bar{Q}} \equiv n$, leading to the Boltzmann equation for the evolution of $n$ under the form $\frac{1}{a^3}\frac{d\left(a^3 n\right)}{dt} = -\langle \varkappa v\rangle \left(n^2 - n_{\rm eq}^2\right)$, where $n_{\rm eq}$ is the equilibrium number density and $\langle \varkappa v\rangle$ the thermally averaged cross-section.

Standard treatments[3] on the form of such particles show that the mass interval is $\leq 2$ TeV. These constituents enter the standard puzzle and represent the ones enabling the mechanism to work. Summarizing the model produces intriguing outcomes which have been reported in the list below

$$P = \text{const (always)} \Rightarrow c_{\rm s}^{\rm DM} \equiv c_{\rm s}^{\rm BM} \equiv 0\,, \tag{23}$$

$$P < 0 \text{ (from thermodynamics)}\,, \tag{24}$$

$$T = \text{const (during the transition)}\,, \tag{25}$$

$$P_{\rm DM} \gg P_{\rm BM}\ ,\ P_{\rm DM} \approx \epsilon_{\rm v}\,, \tag{26}$$

$$\rho_\Lambda \equiv P_{\rm DM}\ (\rm BT)\,, \tag{27}$$

$$\rho_\Lambda \equiv P_{\rm BM}\ (\rm AT)\,, \tag{28}$$

$$0.5 \lesssim Mc^2/\text{TeV} \lesssim 1.7\ \text{(Cold Dark Matter)}\,, \tag{29}$$

$$0.81 \leq \langle \varkappa v\rangle/(10^{-26}\text{cm}^3\text{s}^{-1}) \leq 2.42\,. \tag{30}$$

## 5. Conclusions

We demonstrated that there exists only one *perfect, irrotational, and isentropic fluid*, composed by baryons and cold dark matter. If matter is not purely pressureless, a coupling to $\Lambda$ induces a phase transition in which the quantum fluctuations may be removed by means of matter pressure. In particular, the corresponding thermodynamics of our matter fluid *naturally* suggests an emergent negative pressure, which turns out to be constant. This guarantees, from small perturbations, that the sound speed vanishes, degenerating with the $\Lambda$CDM model. Soon after the transition, we can have two possibilities: the emergent cosmological constant is given by the (negative) pressure of baryons only or by the bare cosmological constant which may dominate over baryon pressure. In both cases, the fine-tuning and coincidence issues are overcome and the $\Lambda$CDM effects are mimed in terms of a dark fluid with constant pressure but non-constant density. To address this mechanism to work at thermal equilibrium, we predict the dark matter constituents to lie inside the interval $0.5 \lesssim M \lesssim 1.7$ TeV. This leaves open the possibility to detect in laboratory additional heavier bosons. However, the mass value is strongly influenced by several factors. If non-equilibrium is invoked, for example, its magnitude is expected to change. Future developments will take into account the consequences of our model on structure formation and during early phases of Universe's evolution.

## Acknowledgements

The work was supported in part by the Ministry of Education and Science of the Republic of Kazakhstan, Program 'Fundamental and applied studies in related fields of physics of terrestrial, near-earth and atmospheric processes and their practical application' IRN: BR05236494.

## References

1. S. Weinberg, Rev. Mod. Phys., **61** (1989) 1.
2. E. J. Copeland, M. Sami, S. Tsujikawa, Int. J. Mod. Phys. D, **15** (2006) 1753.
3. O. Luongo, M. Muccino, Phys. Rev. D, **98** (2018) 103520.
4. E. A. Lim, I. Sawicki, A. Vikman, Jour. Cosm. Astrop. Phys., **5** (2010) 012.
5. B. Carter, Lect. Not. Mathem., Berlin Springer Verlag, **1385**, (1989) 1.
6. G. Ballesteros, D. Comelli, L. Pilo, Phys. Rev. D, **94**, (2016) 124023.
7. G. Ballesteros, D. Comelli, L. Pilo, Phys. Rev. D, **94**, (2016) 025034.
8. O. Luongo, H. Quevedo, Int. J. Mod. Phys. D, **23** (2014) 1450012.
9. O. Luongo, H. Quevedo, Astroph. Sp. Scien., **338** (2012) 345.
10. P. K. S. Dunsby, O. Luongo, L. Reverberi, Phys. Rev. D, **94** (2016) 083525.
11. J. Martin, Comptes Rend. Phys., **13**, (2012) 566.

# Nonminimally coupled curvature-matter gravity models and Solar System constraints

Riccardo March

*Istituto per le Applicazioni del Calcolo, CNR, Via dei Taurini 19, Roma, 00185, Italy*
*E-mail: r.march@iac.cnr.it*

Orfeu Bertolami* and Jorge Páramos

*Departamento de Física e Astronomia, Universidade do Porto, Rua do Campo Alegre 687, Porto, 4169-007, Portugal*
** E-mail: orfeu.bertolami@fc.up.pt*

Simone Dell'Agnello

*INFN, Laboratori Nazionali di Frascati (LNF), Via E. Fermi 40, Frascati, 00044 Roma, Italy*
*E-mail: simone.dellagnello@lnf.infn.it*

We discuss constraints to some nonminimally (NMC) coupled curvature-matter models of gravity by means of Solar System experiments.

First we discuss a NMC gravity model which constitutes a natural extension of $1/R^n$ gravity to the nonminimally coupled case. Such a NMC gravity model is able to predict the observed accelerated expansion of the Universe. Differently from the $f(R) = 1/R^n$ gravity case, which is not compatible with Solar System observations, it turns out that this NMC model is a viable theory of gravity.

Then we consider a further NMC gravity model which admits Minkowski spacetime as a background, and we derive the $1/c$ expansion of the metric. The nonrelativistic limit of the model is not Newtonian, but contains a Yukawa correction. We look for trajectories around a static, spherically symmetric body. Since in NMC gravity the energy-momentum tensor of matter is not conserved, then the trajectories deviate from geodesics. We use the NMC gravity model to compute the perihelion precession of planets and we constrain the parameters of the model from radar observations of Mercury.

*Keywords*: Extended theories of gravity; nonminimal coupling; PPN formalism; Mercury precession.

## 1. Introduction

We consider the possibility of constraining some nonminimally coupled (NMC) curvature-matter models of gravity[1] by means of Solar System experiments. The action functional involves two functions $f^1(R)$ and $f^2(R)$ of the Ricci curvature $R$. The function $f^1(R)$ is a nonlinear term which is analogous to $f(R)$ gravity, and the function $f^2(R)$ yields a NMC between the matter Lagrangian density and curvature. For other NMC gravity theories and their applications, see for instance[2–4].

NMC gravity has been applied to several astrophysical and cosmological problems such as dark matter[5,6], cosmological perturbations[7], post-inflationary reheating[8] or the current accelerated expansion of the Universe[9].

First we discuss the application of a perturbative method due to Chiba, Smith and Erickcek[10] to the NMC gravity model by Bertolami, Frazao and Paramos[9],

which constitutes a natural extension of $1/R^n$ gravity to the non-minimally coupled case. Such a NMC gravity model is able to predict the observed accelerated expansion of the Universe. Differently from the $f(R) = R + 1/R^n$ gravity case, which predicts the value $\gamma = 1/2$ for the PPN parameter $\gamma$, so that the $f(R)$ model is not compatible with Solar System observations, it turns out[11] that the NMC gravity model cannot be constrained, for specific choices of the functions $f^1(R)$ and $f^2(R)$, by the perturbative method considered by Chiba *et al.*[10], so that it remains, in this respect, a viable theory of gravity.

Then we consider a further NMC gravity model[12,13], which admits Minkowski spacetime as a background, and we derive the $1/c$ expansion of the metric assuming the functions $f^1(R)$ and $f^2(R)$ analytic at $R = 0$. The nonrelativistic limit of the model is not Newtonian, but contains a Yukawa correction. A parameterized post-Newton plus Yukawa (PPNY) approximation of the NMC model of gravity can be computed. We consider the metric around a static, spherically symmetric body and we look for trajectories of a test body around the spherical body. Since in NMC gravity the energy-momentum tensor of matter is not conserved, then the trajectories deviate from geodesics. We use the NMC gravity model to compute the perihelion precession of planets. Eventually we constrain the parameters of the model from radar observations of Mercury, including data from the NASA orbiter MESSENGER (MErcury Surface, Space ENvironment, GEochemistry and Ranging) spacecraft.

## 2. The NMC gravity action functional

The action functional of NMC gravity is given by[1]

$$S = \int \left[\frac{1}{2} f^1(R) + [1 + f^2(R)]\mathcal{L}_m\right] \sqrt{-g} d^4x,$$

where $f^1(R), f^2(R)$ are functions of the spacetime curvature $R$, $g$ is the metric determinant, $\mathcal{L}_m = -\rho c^2$ is the Lagrangian density of matter, and $\rho$ is mass density.

The function $f^2(R)$ yields a NMC between geometry and matter, and the class of $f(R)$ gravity theories is recovered in the case $f^2(R) = 0$. General Relativity (GR) is recovered by taking:

$$f^1(R) = 2\kappa(R - 2\Lambda), \quad f^2(R) = 0, \quad \kappa = c^4/16\pi G,$$

where $G$ is Newton's gravitational constant and $\Lambda$ is the Cosmological Constant.

The first variation of the action functional with respect to the metric yields the field equations

$$\left(f^1_R + 2f^2_R\mathcal{L}_m\right) R_{\mu\nu} - \frac{1}{2} f^1 g_{\mu\nu} = \nabla_{\mu\nu}\left(f^1_R + 2f^2_R\mathcal{L}_m\right) + \left(1 + f^2\right) T_{\mu\nu}, \qquad (1)$$

where $f^i_R = df^i/dR$ and $\nabla_{\mu\nu} = \nabla_\mu\nabla_\nu - g_{\mu\nu}g^{\sigma\eta}\nabla_\sigma\nabla_\eta$. Such equations will be solved by perturbative methods.

## 3. A model for the accelerated expansion of the Universe

We consider the NMC gravity model proposed by Bertolami, Frazao and Paramos[9] to account for the observed accelerated expansion of the Universe:

$$f^1(R) = 2\kappa R, \qquad f^2(R) = \left(\frac{R}{R_n}\right)^{-n}, \quad n > 0, \tag{2}$$

where $n$ is an integer and $R_n$ is a constant. This NMC gravity model constitutes a natural extension to the non-minimally coupled case of the $1/R^n$ model proposed by Carroll *et al.*[14] as an instance of $f(R)$ model.

Matter is described as a perfect fluid with negligible pressure[15] with Lagrangian density $\mathcal{L}_m = -\rho c^2$. We assume that the metric, which describes the spacetime around the Sun, is a perturbation of a flat Friedmann-Robertson-Walker (FRW) metric with scale factor $a(t)$:

$$ds^2 = -\left[1 + 2\Psi(r,t)\right] dt^2 + a^2(t) \left(\left[1 + 2\Phi(r,t)\right] dr^2 + r^2 d\Omega^2\right), \tag{3}$$

where $|\Psi(r,t)| \ll 1$ and $|\Phi(r,t)| \ll 1$. The NMC gravity model Eq. (2) yields a cosmological solution with a negative deceleration parameter $q < 0$, and the scale factor $a(t)$ of the background metric follows the temporal evolution $a(t) = a_0 \left(t/t_0\right)^{2(1+n)/3}$, where $t_0$ is the current age of the Universe[9].

In the perturbative approach developed by Chiba *et al.*[10] for $f(R)$ gravity, the Ricci curvature of the perturbed spacetime is expressed as the sum

$$R(r,t) = R_0(t) + R_1(r,t),$$

where $R_0$ denotes the scalar curvature of the background FRW spacetime and $R_1$ is the perturbation due to the Sun. The extension of the perturbative method of Chiba *et al.*[10] to NMC gravity consists in the following steps[11]. We assume that functions $f^1(R)$ and $f^2(R)$ admit a Taylor expansion around $R = R_0$, and we linearize the field equations (1) under two conditions:

(i) terms nonlinear in $R_1$ can be neglected in the Taylor expansion of $f^1, f^2$;
(ii) the following inequality

$$|R_1(r,t)| \ll R_0(t), \tag{4}$$

is satisfied both around and inside the Sun.

We compute the functions $\Psi$ and $\Phi$ of the metric Eq. (3), then we find an expression of the parameter $\gamma$ of the PPN (Parameterized Post-Newtonian) formalism[16]. Eventually the validity of the condition (4) is checked a posteriori.

The condition (4) means that the curvature $R$ of the perturbed spacetime remains close to the cosmological value $R_0$ inside the Sun. In GR such a property of the curvature is not satisfied inside the Sun. However, for some $f(R)$ theories condition (4) can be satisfied and that leads to a violation of a constraint on PPN parameter $\gamma$ from Solar System tests of gravity. For instance, the $1/R^n$ $(n > 0)$ gravity model[14] satisfies condition (4)[10,17].

The perturbative solution of the field equations (1) yields the following expression for the PPN parameter $\gamma = -\Phi(r)/\Psi(r)$ [11]:

$$\gamma = \frac{1}{2}\left[\frac{1 + f_0^2 + 4f_{R0}^2 R_0 + 12\Box f_{R0}^2}{1 + f_0^2 + f_{R0}^2 R_0 + 3\Box f_{R0}^2}\right],$$

where $f_0^2 = f^2(R_0)$ and $f_{R0}^2 = df^2/dR(R_0)$. When $f^2(R) = 0$ we find the known result $\gamma = 1/2$ which holds for $f(R)$ gravity theories which satisfy the condition $|R_1| \ll R_0$ [10]. The $1/R^n$ $(n > 0)$ gravity theory [14], where $f(R)$ is proportional to $(R + \text{constant}/R^n)$, is one of such theories that, consequently, have to be ruled out by Cassini measurement [18].

For the NMC gravity model (2), though $|R_1| \ll R_0$ for $n \gg 1$, the solution for $R_1$ inside the Sun shows that non-linear terms in the Taylor expansion of $f^2(R)$ cannot be neglected [11]:

$$f^2(R) = f_0^2\left[1 - n\frac{R_1}{R_0} + \frac{n(n+1)}{2}\left(\frac{R_1}{R_0}\right)^2 - \frac{1}{6}n(n+1)(n+2)\left(\frac{R_1}{R_0}\right)^3\right] + O\left(\left(\frac{R_1}{R_0}\right)^4\right).$$

Hence assumption (i) is contradicted, implying the lack of validity of the perturbative regime. Eventually, by such a contradiction argument the model (2) cannot be constrained by the extension to a NMC of the perturbative method by Chiba *et al.* [10], so that the model (2) remains, in this respect, a viable theory of gravity [11].

## 4. Planetary precession

We now consider a NMC gravity model where the functions $f^1(R)$ and $f^2(R)$ are assumed analytic at $R = 0$ [13], so that they admit the Taylor expansions:

$$f^1(R) = 2\kappa \sum_{i=1}^{\infty} a_i R^i, \qquad a_1 = 1, \qquad f^2(R) = \sum_{j=1}^{\infty} q_j R^j.$$

If $a_i = 0$ for any $i > 1$ and $q_j = 0$ for any $j$, then the action of GR is recovered.

The model admits Minkowski spacetime as a background, and the $1/c$ expansion of the metric can be computed [13], assuming a general distribution of matter with mass density, pressure and velocity. The nonrelativistic limit of the model turns out to be non-Newtonian, but contains also a Yukawa correction. The coefficients $a_2, a_3, q_1, q_2$ are used to compute the metric at the order $O(1/c^4)$ for the $0-0$ component, and are considered as parameters of the NMC gravity model. A parameterized post-Newton plus Yukawa (PPNY) approximation of the NMC model of gravity can be computed [13].

Here we report the result [13] for the metric in vacuum around a static, spherically symmetric body (Sun) with uniform mass density ($g_{0i} = 0$):

$$\begin{aligned} g_{00} &= -1 + 2\frac{GM_S}{rc^2}\left(1 + \alpha e^{-r/\lambda}\right) + \frac{2}{c^4}F(r), \\ g_{ij} &= \left[1 + 2\frac{GM_S}{rc^2}\left(1 - \alpha e^{-r/\lambda}\right)\right]\delta_{ij}, \end{aligned} \tag{5}$$

where $M_S$ is the mass of the spherical body, $F(r)$ is a radial potential, and $\lambda, \alpha$ are the range and strength of the Yukawa potential which depend on the parameters of the NMC gravity model[13]:

$$\lambda = \sqrt{6a_2}, \qquad \alpha = \frac{1}{3}(1-\theta) + \frac{GM_S}{c^2 R_S}\theta\left[\theta\left(\frac{\mu}{2}-1\right) - \frac{2}{3}\nu\right]\left(\frac{\lambda}{R_S}\right)^2 + \ldots, \tag{6}$$

where $R_S$ is the radius of the spherical body, $\theta, \mu, \nu$ are the following dimensionless ratios: $\theta = q_1/a_2$, $\mu = a_3/a_2^2, \nu = q_2/a_2^2$, and dots ... denote smaller contributions[13]. Formula (6) has been obtained for $\lambda \gg R_S$.

Using the metric (5) the effect of NMC gravity on the orbit of a planet is computed. In NMC gravity the energy-momentum tensor is not covariantly conserved[1]:

$$\nabla_\mu T^{\mu\nu} = \frac{f_R^2}{1+f_2}(g^{\mu\nu}\mathcal{L}_m - T^{\mu\nu})\nabla_\mu R \neq 0 \qquad \text{if } f^2(R) \neq 0,$$

consequently, the trajectories deviate from geodesics:

$$\frac{d^2x^\alpha}{ds^2} + \Gamma^\alpha_{\mu\nu}\frac{dx^\mu}{ds}\frac{dx^\nu}{ds} = \frac{f_R^2(R)}{1+f^2(R)}g^{\alpha\beta}R_{,\beta}. \tag{7}$$

Moreover, geodesics are different from GR. The formula for perihelion precession of a planet has been computed[13] for $\lambda \gg L$, where $L$ is the *semilatus rectum* of the unperturbed orbit. Here we report the leading term in the formula[13]:

$$\begin{aligned}\delta\phi_P = {} & \frac{6\pi GM_S}{Lc^2} + (1-\theta)^2\frac{\pi}{3}\left(\frac{L}{\lambda}\right)^2 e^{-L/\lambda} \\ & + (1-\theta)\frac{\pi GM_S}{3Lc^2}\theta\left[3\theta\left(\frac{\mu}{2}-1\right) - 2\nu\right]\left(1-\frac{L}{\lambda}\right)\left(\frac{L}{R_S}\right)^3 + \ldots,\end{aligned} \tag{8}$$

where the terms in the first row are the GR precession and the nonrelativistic Yukawa precession, respectively, and the term in the second row is the leading contribution from the NMC relativistic correction. Dots ... denote smaller contributions[13]. Eq. (8) reduces to the GR expression if $\theta = 1$.

Using Eq. (8), bounds on PPN parameters from the Cassini experiment[18] and fits to planetary data, including data from Messenger spacecraft[19] orbiting around Mercury, it follows that the additional perihelion precession due to NMC deviations from GR, in the case of Mercury orbit, is bounded by[13]

$$-5.87537 \times 10^{-4} < \delta\phi_P - 42.98'' < 2.96635 \times 10^{-3}.$$

These inequalities define an admissible region in the four-dimensional parameter space with dimensionless coordinates $\theta, \mu, \nu, R_S/\lambda$. Exclusion plots obtained by slicing the admissible region with two-dimensional planes can be drawn[13].

The admissible region in three-dimensional parameter subspace with coordinates $(\theta, \mu, \nu)$, for $0 < |1-\theta| \ll 1$ and a given $\lambda \gg L$, can be approximated by the region enclosed within the degenerate quadric surfaces

$$\nu = \frac{3}{4}\mu - \frac{3}{2} - 9\left(\frac{R_S}{L}\right)^3 \frac{\varepsilon_i}{(1-L/\lambda)(1-\theta)}, \qquad i = 1, 2,$$

where

$$\varepsilon_1 \frac{6\pi G M_S}{Lc^2} = -5.87537 \times 10^{-4}, \qquad \varepsilon_2 \frac{6\pi G M_S}{Lc^2} = 2.96635 \times 10^{-3}.$$

The intersection of the three-dimensional admissible subregion with a plane $\theta = constant$, with $0 < |1 - \theta| \ll 1$, is a strip enclosed between two lines in the $(\mu, \nu)$ plane. The intersections with the planes $\mu = constant$ and $\nu = constant$ are regions enclosed by pairs of hyperbolae[13].

Eventually, the BepiColombo mission to Mercury should allow for a reduction on the above bounds by approximately one order of magnitude[20].

## Acknowledgments

The work of R.M. and S.DA is, respectively, partially and fully supported by INFN (Istituto Nazionale di Fisica Nucleare, Italy), as part of the MoonLIGHT-2 experiment in the framework of the research activities of the Commissione Scientifica Nazionale n. 2 (CSN2).

## References

1. O. Bertolami, C.G. Böhmer, T. Harko and F.S.N. Lobo, *Phys. Rev. D* **75**, 104016 (2007).
2. D. Puetzfeld and Y.N. Obukhov, *Phys. Rev. D* **87**, 044045 (2013).
3. D. Puetzfeld and Y.N. Obukhov, *Phys Lett A* **377**, 2447 (2013).
4. D. Puetzfeld and Y.N. Obukhov, *Phys. Rev. D* **88**, 064025 (2013).
5. O. Bertolami and J. Páramos, *JCAP* **03**, 009 (2010).
6. O. Bertolami, P. Frazão and J. Páramos, *Phys. Rev. D* **86**, 044034 (2012).
7. O. Bertolami, P. Frazão and J. Páramos, *JCAP* **05**, 029 (2013).
8. O. Bertolami, P. Frazão and J. Páramos, *Phys. Rev. D* **83**, 044010 (2011).
9. O. Bertolami, P. Frazão and J. Páramos, *Phys. Rev. D* **81**, 104046 (2010).
10. T. Chiba, T.L. Smith and A.L. Erickcek, *Phys. Rev. D* **75**, 124014 (2007).
11. O. Bertolami, R. March and J. Páramos, *Phys. Rev. D* **88**, 064019 (2013).
12. N. Castel-Branco, J. Páramos and R. March, *Phys. Lett. B* **735**, 25 (2014).
13. R. March, J. Páramos, O. Bertolami and S. Dell'Agnello, *Phys. Rev. D* **95**, 024017 (2017).
14. S.M. Carroll, V. Duvvuri, M. Trodden and M.S. Turner, *Phys. Rev. D* **70**, 043528 (2004).
15. O. Bertolami, F.S.N. Lobo and J. Páramos, *Phys. Rev. D* **78**, 064036 (2008).
16. C.M. Will, *Theory and Experiment in Gravitational Physics, Revised Ed.*, (Cambridge University Press, 1993).
17. K. Henttunen, T. Multamäki and I. Vilja *Phys. Rev. D* **77**, 024040 (2008).
18. B. Bertotti, L. Iess and P. Tortora, *Nature* **425**, 374 (2003).
19. A. Fienga *et al.*, *Celest. Mech. Dyn. Astr.* **111**, 363 (2011).
20. F. De Marchi, G. Tommei, A. Milani and G. Schettino, *Phys. Rev. D* **93**, 123014 (2016).

# CPV in the space

G. M. Piacentino*[1,2,3] G. Di Sciascio*[2] A. Gioiosa*[1,2] D. Hajdukovic*[4] A. Palladino*[5]
G. Venanzoni*[6]

[1] *Università degli Studi del Molise, Campobasso, Italy*
[2] *INFN, Sezione di Tor Vergata, Roma, Italy*
[3] *INAF, Sezione di Milano, Milano, Italy*
[4] *CERN, Geneve, CH*
[5] *Boston University, Boston, USA*
[6] *INFN, Sezione di Pisa, Pisa, Italy*
* *E-mail: giovanni.piacentino@unimol.it*

Our proposal is suggested by the presence of still open problems in Astrophysics and Cosmology. In fact the Standard Cosmological Model, even if in agreement with the experiments on several Phenomena, features strange structure, with the presence of the "Magic" Contribute of Dark Matter and Dark Energy that reaches 96% of the total. At a closer glance the principal open problems in Astrophysics and Cosmology are:

*i.* In the observed Universe the matter prevails on antimatter even if both are always created together;
*ii.* CMB is not anisotropic nor inhomogeneous enough to be compatible with the Big Bang model without the introduction of a still unknown interaction driving the inflation;
*iii.* Given the gravity we expect a negative acceleration of the expansion. On the contrary that seems to accelerate;
*iv.* The gravitational field of Galaxies, clusters and even of the Solar system seems much stronger that the one due to the visible matter.

At present the possible suggested solutions are:

*i.* The mechanism suggested by Sakarov for matter/antimatter asymmetry is connected to CPV but experimentally this phenomenon is far too weak;
*ii.* Models have been proposed to justify inflation by supersymmetric vacuum energy and SSB but at present no evidence for supersymmetry has been found yet;
*iii.* Dark Energy has been introduced by hand in order to give a motivation to the accelerated expansion of the Universe;
*iv.* Dark Matter has been introduced in order to give a motivation to the observed discrepancies between theory and measurements of the orbital speed of the stars of the external part of the galaxies.

This situation is far from being satisfactory from the point of view of the elegance:

*i.* As many hypothesis as problems;

*ii.* Most of them just put by hand into the theory;
*iii.* Dark Matter and Dark Energy hypotheses are artificial.

The hypothesis that antimatter could have a different coupling to the gravitational field has fascinated physicists since the discovery of the first antiparticles. Several authors and one of the authors of this paper, Hajdukovic, have shown that the possible gravitational repulsion between matter and antimatter could offer at least a partial explanation for this number of cosmological problems, including those connected to dark matter and dark energy [1–15] with this single hypothesis. At present, the state-of-the-art is not very different from the framework summarized in the article by Nieto et al. in 1991 [16]. Limits on repulsive gravity have been calculated based on measurements [17] and [18]. A relatively large number of experiments on the gravitational interaction of antimatter have been proposed and even started, e.g. AEGIS [19], ALPHA [20], ATRAP [21], and GBAR [22]. Our proposal is very simple and is completely model independent. It is inspired by Good's argument [28]. We in fact propose to study a possible dependence of CP violation on the gravitational interaction in the Ks-KL system in space. The magnitude of any difference between the CP violation parameter, measured in orbit and that measured on Earth's surface would give important indication on the nature of the gravitational interaction between matter and antimatter as well as provide evidence for a quantum gravitational effect. In this paper we outline a new approach to the problem capable of providing a measurement taking advantage from the presence of energetic cosmic protons at the level of the LEO orbit around the Earth.

The rate of incoming protons in a LEO has, in fact, been measured [29] and [30], and when integrated on the target of the proposed orbiting detector, can reach as many as $2.2x10^4$ protons per second. The energy of the cosmic protons ranges from a few MeV to ∼200 GeV with the maximum flux around 1 GeV and several smaller local maxima at 5, 13, and 31 GeV.

We simulated the interaction of this spectrum of protons with several kinds of targets in order to extimate the KL and Ks production (Figures 1 and 2). We simulated the position of the decay vertex as well, downstream of a Wolframium target.

In the Figures 3 and 4 it is evident that a simple cut at the moment allows to depress the contamination of Ks in the measured events.

This means that the incident proton spectrum is energetic enough for the production of kaons allowing for a measurement of

$$R = \frac{\Gamma(\mathrm{K_L} \to \pi^+\pi^-)}{\Gamma(\mathrm{K_L} \to \pi^+\pi^-\pi^0)} \tag{1}$$

which is quadratic in $\varepsilon$. The mean gravitational field strength in a Low Earth Orbit (LEO) is about 10% less than the one on the Earth's surface. If CP violation depends linearly on the gravitational field [1] and [2] we expect a 10% effect to translate into a 20% effect on R. On a GEO orbit the intensity of the of the Earth's

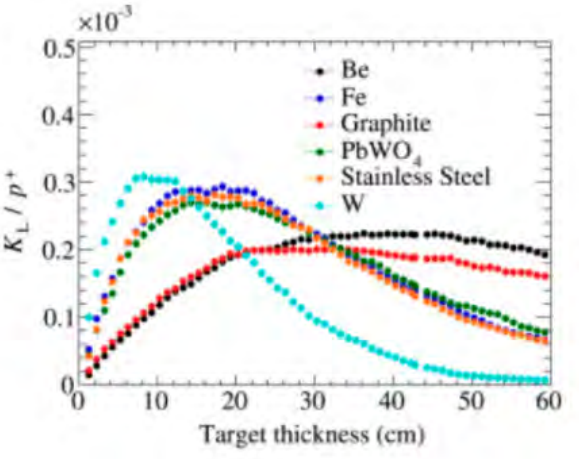



Fig. 1. KL production as simulated on different targets from cosmic proton spectrum as measured by AMS satellite[33,35].

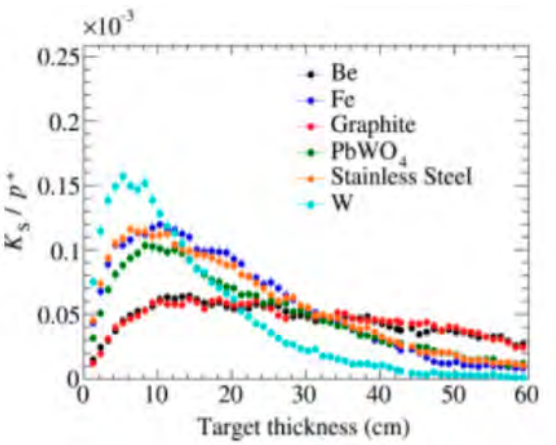



Fig. 2. KS production as simulated on different targets from cosmic proton spectrum as measured by AMS satellite[33,35].

gravitational field is of the order of few percent configuring an even better experimental condition. We suggest the use of a dedicated Satellite on a LEO orbit with a payload aimed at performing a particle physics experiment in orbit.

As can be seen in Figure 5 it includes active target, magnetic spectrometer, tracker and calorimeter. We estimate that on a square target of 70 cm of side, about $1.4x10^4$ protons per second will impact, and produce the neutral Kaons. The total number of K mesons decays over a space mission lifetime ($> 2$ years) will yield the required physical measurement of $12.5x10^5$ $K_L$ decays necessary to make a 5 $\sigma$ measurement of a 20% variation of R.

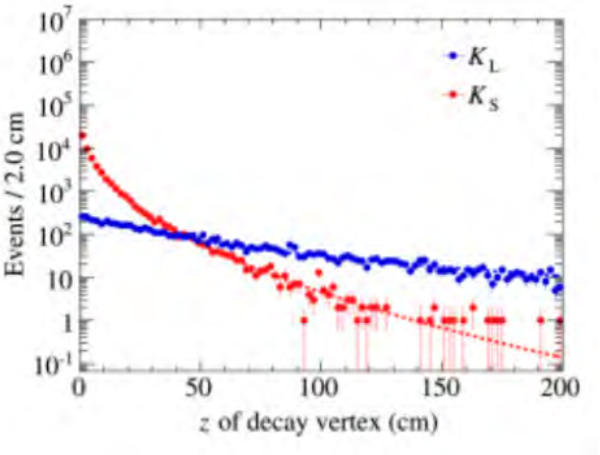



Fig. 3. Decay vertex z position for simulated production of KL and Ks by cosmic proton spectrum as measured by AMS satellite [35].

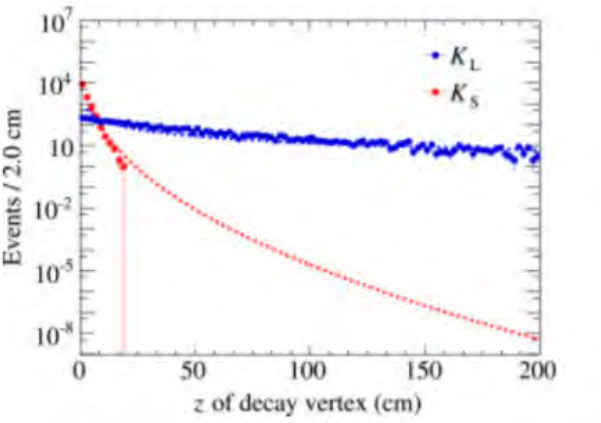



Fig. 4. Decay vertex z position for simulated production of KL and Ks by cosmic proton spectrum as measured by AMS satellite with the additional constraint $P_z < 0,5\ Gev/c$ [35].

## References

1. G. Chardin, Nuclear Phys. A **558**, 477c (1993).
2. G. Chardin, J.M. Rax, Phys. Lett. B **282**, 256-262 (1992).
3. A. Benoit-Lévy, G. Chardin, Astron. Astrophys. **537** A78 (2012).
4. A. Benoit-Lévy, G. Chardin, International Journal of Modern Physics: Conference Series. Antimatter and Gravity Conference (WAG 2013) **30**, 1460272 (2014).
5. D.S. Hajdukovic, Astrophys. Space Sci. **334**, 215-218 (2011).
6. D.S. Hajdukovic, Astrophys. Space Sci. **334**, 219-223 (2011).

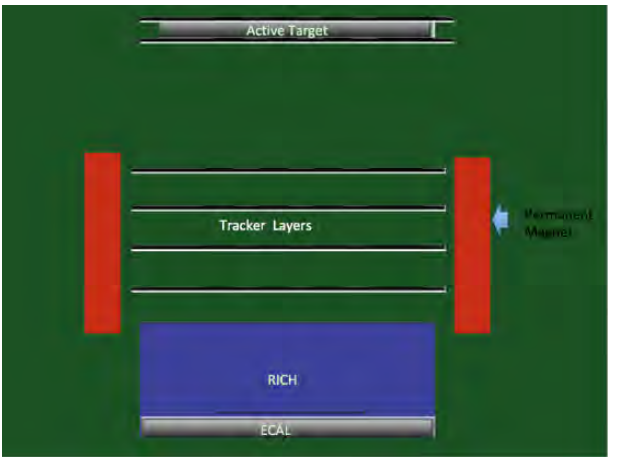



Fig. 5. Structure of the detector proposed as a payload of a satellite dedicate to the measurement of epsilon in a low gravity environment.


7. D.S. Hajdukovic, Space Sci. **337**, 9-14 (2012).
8. D.S. Hajdukovic, Astrophys. Space Sci. **339**, 1-5 (2012).
9. D.S. Hajdukovic, Phys. Dark Universe **3**, 34-40 (2014).
10. D.S. Hajdukovic, Proceedings of the 3rd International Workshop on Antimatter and Gravity, WAG (2015) https://hal.archives-ouvertes.fr/hal-01254678v2.
11. M. Villata, Astrophys. Space Sci. **345**, 1-9 (2013).
12. L. Blanchet, Classical Quantum Gravity **24**, 3529 (2007).
13. L. Blanchet, A. Le Tiec, Phys. Rev. D **78**, Article 024031 (2008).
14. L. Blanchet, A. Le Tiec, Phys. Rev. D **80**, 023524 (2009).
15. L. Bernard, L. Blanchet, Phys. Rev. D **91**, Article 103536 (2015).
16. M.M. Nieto, T. Goldman, Phys. Rep. **205** 221-281 (1991).
17. D.S.M. Alves, M. Jankowiak, P. Saraswat, *Experimental constraints on the free fall acceleration of antimatter*. arXiv:0907.4110v1.
18. M. Fischler, J. Lykken, T. Roberts, *Direct observation limits on antimatter gravitation*. arXiv:0808.3929v1.
19. A. Kellerbauer (AEgIS Collaboration) et al., Nucl. Instrum. Methods Phys. Res. B **266** 351 (2008). http://dx.doi.org/10.1016/j.nimb.2007.12.010.
20. A.E. Charman (ALPHA Collaboration) et al., Nature Comm. **4**, 1785 (2013).
21. G. Gabrielse (ATRAP Collaboration) et al., Phys. Rev. Lett. **108**, Article 113002 (2012).
22. G. Chardin, P. Grandemange, D. Lunney, et al., Tech. Rep. CERN-SPSC-2011-029 SPSC-P-342.
23. K. Kirsch, *Testing gravity with muonium*. arXiv:physics/0702143.
24. D.M. Kaplan, et al., *Measuring antimatter gravity with muonium*. arXiv:1308.0878.

25. M. Gai, A. Vecchiato, *Astrometric detection feasibility of gravitational effects of quantum vacuum.* arXiv:1406.3611.
26. CPLEAR Collaboration, Phys. Lett. B **452** 425-433 (1999).
27. G. Mambriani, L. Trentadue, *Testing CP conservation at KLOE.* arXiv:hep-ex/0007004.
28. M.L. Good Phys. Rev. **121** 311-313 (1961).
29. AMS Collaboration, Phys. Lett. B **427**, 251-226 (2000).
30. AMS Collaboration, Phys. Lett. B, **490**, 27-35 (2000).
31. S. Agostinelli et al., Nucl. Instrum. Methods. Phys. Res. A **506** 250-303 (2003).
32. Geant4 developments and applications, IEEE Trans. Nucl. Sci. **53**, (2006).
33. AMS-02 Collaboration, Phys. Rev. Lett. **110**, Article 141102 (2013).
34. Pamela Collaboration, Astrophysics Astropart. Phys. **27**, 296-315 (2007).
35. G.M. Piacentino, A. Palladino, G. Venanzoni, Physics of the Dark Universe **13**, 162-165 (2016).

# New expression of energy and the use of spatial parameters to achieve unification based on the ideas of Einstein and Wheeler

Arban Uka

*Faculty of Architecture and Engineering, Epoka University*
*Tirana, 1032, Albania*
*auka@epoka.edu.al*
*www.epoka.edu.al*

Shpetim Nazarko

*Independent Researcher*
*Tirana, 1031, Albania*
*shpetim.nazarko@gmail.com*

The first step to complete physical theories is to contemplate on the axiomatic character of the principles. This principle of relativity was formulated by Galileo and used by Newton to derive the laws of motion, and later was placed by Einstein at the center of the Special Theory and the General Relativity. This work considers the general principle of relativity as an expression of the inherent characteristic of the most important quantity in physics - the energy - that states that the energy is constant and is independent of the type or speed of movement of an object. The energy of an object is expressed as one entity that equals the sum of two energy expressions with different character. One of these two terms is denoted as the exposed energy with a kinetic character and represents the magnitude of the field of an object. This new definition of the energy offers several perspectives that are valid for both the classical and relativistic physics and provides insights beyond these two as: i) it leads to a coherent and complete model that relates the energy with the momentum and the force, ii) enables to construct the lagrangian that can be used to derive the equations of motion without artifices, iii) offers an original viewpoint on the dark matter and dark energy that is coherent with recent developments. Einstein was the first to deny the separate existence of gravitation and electromagnetism and has implied that the goal of a unified theory would be to explain the existence and to calculate the properties of matter. He revolutionized the use of the principles of symmetry in deriving the physical laws. It is known that the classical and quantum domain do not overlap and one would not expect the same governing symmetry principles in both domains. To calculate the properties of matter we employ spatial parameters that are related with the energy state of an object and the force that it exhibits. The mechanism assumes the common origin of different forces at the highest energy level. These spacetime parameters can be used to derive the electric and gravitational forces through a single mechanism that respects the LT dimensional analysis using neither the electrical charge nor the gravitational constant.

*Keywords*: Special relativity; exposed energy; gravitational force; non-gravitational forces; spacetime parameters; Planck scale.

## 1. Introduction

Newton and Einstein in his early efforts in formulating special relativity did not recognize the energy as an inherent characteristic of the object and this is the major reason why they postulated the expressions of the momentum in both the classical and relativistic physics. The aim of the construction of the fundamental theories is the complete understanding of the matter and its interactions. The steps to follow

in the formulation of fundamental theories would be: i) the matter, ii) fundamental quantities, iii) invariants, iv) the principle, v) mathematical apparatus, vi) the conservation laws. The matter is associated with the energy that is strongly related to the fundamental quantities of mass-space-time. The dynamic interaction of these quantities leads to consistent patterns, or constants of nature such as the constant of the speed of light c, Planck constant h, Sommerfeld constant $\alpha$ etc. These constants lead to the formulation of the principles that stand at the top of the theories. This approach allows to remove the axiomatic character of the principles, which according to the Method are considered simply as a representation of the character of the fundamental quantities. So, the fact that the speed of light is constant constitutes a postulate of the special theory, is just an unnecessary promotion to a undeserved rank. The speed of light being a constant is just a result of the fact that the ratio between the lengthscale and timescale parameters is always a constant. The energy dictates the behaviour of the space and time parameters, and their behaviour produces an important constant in physics: the constant of the speed of light. At this point we have arrived at the fourth step of the newly proposed scheme: the principle. The principle is associated by a mathematical apparatus that reflects the essence of the principle from a quantitative perspective. The scheme is concluded with the formulation of the conservation laws that determine the range of the applicability in practice of the principle that is put at the front of the theories. Placing the energy at the top of the hierarchical structure makes unnecessary the postulation of the expression of the physical quantities. Here we define the exposed energy that can be considered as the measurable energy of an object, or also as a term that is analoguous to the kinetic energy term. We define also the expression for the unexposed energy that is analoguous to the energy stored in the internal degrees of freedom of an object, or can be considered as the amount of the energy that cannot be measured experimentally. The exposed energy can be used to describe the overall interaction of an object and we can derive all the expressions of physical quantities from it. The expression for the momentum is not postulated but is directly obtained by taking the derivative of the exposed energy (of the kinetic term) with respect to the velocity for both the classical and relativistic physics. Taking the derivative of this quantity with respect to the velocity yields the expression of the momentum which for small velocities is the same as in the classical mechanics and relativistic physics.

Over time, through great efforts physicists have been able to construct elegant models of interaction of matter. Now the challenge is to unify these models as much as possible and to have a complete picture [1]. Sometimes, if not frequently, we may need to be more creative than analytical. Einstein was the first that used the principles of symmetry to derive the constraints to the dynamical laws. He wanted to include electromagnetism to the general relativity and was very persistent in the task. The major difficulty is to provide a common explanation to both gravity and quantum mechanics. The quantum effects become dominant only at smaller

lengthscales, whereas gravitation is dominant at large lengthscales and it would be unexpected to have a single symmetry principle at both domains.

De Broglie was the first that associated a space parameter to an object and on one of his works dating 1964 he calls for a reinterpretation of the wave mechanics and states that there is a need to postulate a relation between the wave and the corpuscle[2]. In a unique work that applies topological methods to cosmology, Bartini uses a unique model that a particle with a small inner radius (gravitational radius) can undergo an inversion to the larger outer radius (classical radius)[3]. He then proceeds to derive all the physical constants as a function of two parameters only.

Here in this work we present the following major points: (i) the association of an object with a field that is a function of the object's mass only, allows one to define spacetime parameters with a quantum character, and (ii) the spacetime parameters are used to determine all the parameters of an object including the potential energy that the object creates, from which the force emerges. Using these spatial parameters the force exerted by an object as a function of mass only is derived, thus dropping the gravitational constant, the Coulomb's constant and the concept of the charge. These spatial parameters are modified by the Lorentz factor and at the highest energy level are equal to the Planck length and it is at this specific point where the quantum foam-like medium is excited[4].

## 2. Definition of the Exposed Energy

The principles upon which the physical theories are constructed, rather than having an axiomatic character, are statements that describe the character of the most important fundamental quantity: the inherent character of the energy. The major principle should emphasize the constant character of energy and that is described as the sum of the internal energy (distributed among the internal degrees of freedom) and the exposed energy (defined here in this work and has kinetic character). The central idea emphasizes that as a force is exerted on an object, there is no change in its mass or its total energy but only a transformation of the internal energy into the exposed energy. The object changes only its attributes such as inertia, an attribute that should not be confused with the mass or the energy.

$$E_{exp} = mc^2 \left[ 1 - \sqrt{1 - \frac{v^2}{c^2}} \right] \tag{1}$$

An object with a defined mass and energy transforms its energy as the velocity changes. The part of the energy that is transformed is defined in this work as the exposed energy as in Eqn. (1), whereas the other part is called the unexposed energy and expressed as in Eqn. (2).

$$E_{unexp} = mc^2 \left[ \sqrt{1 - \frac{v^2}{c^2}} \right] \tag{2}$$

The sum of these two energies is a constant quantity which is called the total energy of an object, $E = E_{exp} + E_{unexp} = mc^2$. The exposed energy for small speeds converges to the expression of the kinetic energy in classical physics: $E = \frac{1}{2}mv^2$.

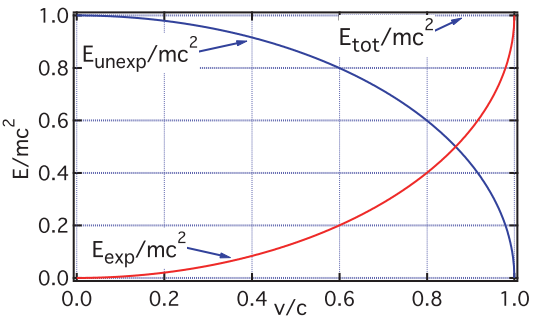



Fig. 1. The exposed (red) and unexposed (blue) energy as a function of velocity and their corresponding momentum.

The interpretation of the two newly defined energies is that the exposed energy has a kinetic character and relates to the experimentally measurable quantities. So if one massive object has a small speed then the effects that it exhibits are small, whereas for higher speed the measurable quantities increase. In this respect, the total percentage of the whole mass of the universe that is called the standard mass is very small because of the low exposed energy, or because of its relatively low kinetics. According to this interpretation, there is no need to include the dark matter as a constituent of the mass of the universe. Careful analysis of recent experimental data is raising serious questions on the existence of dark matter[5]. As the universe undergoes expansion the lengthscale increases and thus the energy scale decreases, and this decrease results in the lowering of the exposed energy. As the exposed energy decreases the experimentally measurable mass of the Universe decreases.

## 3. Derivation of the Equations of Motion

One can find the expression for the momentum by taking the derivative of the exposed energy – that represents the field of a material object as Einstein envisioned — with respect to the velocity and that is valid for relativistic and classical physics.

$$p = \frac{dE_{exp}}{dv} = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}} \tag{3}$$

If one takes the derivative of the momentum with time, one finds the expression of the force for both classical mechanics and relativistic mechanics.

$$F = \frac{dp_{exp}}{dv} = \frac{ma}{(1 - \frac{v^2}{c^2})^{3/2}} \tag{4}$$

Once we find this force, one can calculate the work done on an object by integrating the force with the distance and one finds: $W = \int F dx = (\gamma - 1)mc^2 = KE$.

The exposed energy can be used in the formulation of the Lagrangian from which the equations of motion for both classical and relativistic mechanics are derived.

$$L = mc^2 \left[ 1 - \sqrt{1 - \frac{v^2}{c^2}} \right] - V(x) \tag{5}$$

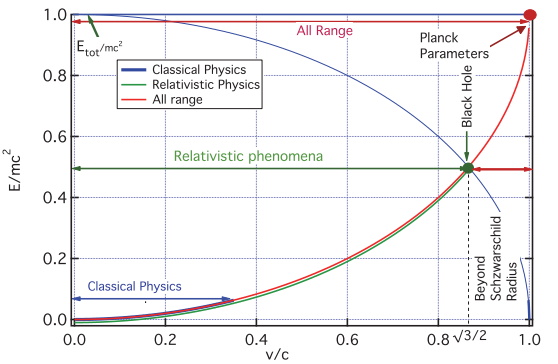



Fig. 2. The exposed and unexposed energy used to determine classical physics (blue), relativistic physics (green) and quantized geometrodynamics domain (red).

The expression for the energy used here in this work never diverges to infinity. The expression for the momentum that is postulated by Newton and then by Einstein is found here by taking the derivative of the energy. A hierarchical functional scale of the expressions of energy-momentum-force is constructed. The lagrangian described by Eqn. (5) is valid for the classical physics depicted in Figure 2 in blue (not to scale), and is also valid for the relativistic physics shown in green up to the intersection point of the exposed energy with the unexposed energy. In the literature one may find the Lagrangian expressed in the form as $L = -\frac{mc^2}{\gamma}$, where $\gamma = \frac{1}{\sqrt{1-\beta^2}}$. This expression may be criticized as the Lagrangian is commonly a relation of kinetic energy and this quantity cannot be a negative quantity. The Lagrangian reported as in Eqn (5) does not include a negative energy term and

the term is a bounded value that is between 0 and $mc^2$. The Lagrangian in polar coordinates can be expressed as below.

$$L = mc^2 \left[ 1 - \sqrt{1 - \frac{\dot{r}^2 + r^2\dot{\theta}^2}{c^2}} \right] - V(r) \tag{6}$$

The intersection point of the exposed energy with the unexposed energy is at $v = \frac{\sqrt{3}}{2} = 0.866c$ and this point determines the upper limit that an object can be described using general relativity. All the observed black holes have rotational speed smaller than this critical value. The highest speed observed is that of the black hole found in the center of the NGC 1365 galaxy with a reported value of 0.84c[6]. An object beyond this critical point cannot be analyzed without revealing its singularity. The intersection of the curves corresponds to the point where an object has its Schwarzschild radius or lengthscale parameter. After the intersection point, which would be beyond the black hole radius, all the space parameters have a quantum character and the calculations are done using the quantized geometrodynamics as formulated by Wheeler[4]. In the next section we will also show that beyond this limit an object can reach the highest energy level that would correspond to the Planck parameters.

The exposed energy increases as shown in the figure above and higher energy levels are associated with smaller lengthscales and at the highest energy level an object's spatial parameter correspond to the Planck parameters. This perspective is in line the idea of doubly-special relativity[7].

## 4. Wave Mechanics and Associating Space Parameters for Objects

Here we propose the use of mcr=$\alpha\hbar$ to find space parameters for any object. A similar formulation is the Compton wavelength, which plays an important role in determining the Compton shift and appears in Klein-Gordon equation. The equation allows parameterization of the minimum (corresponding to largest lengthscale) and the maximum energy state (smallest lengthscale) of objects in nano- and macroscale.

$$r^m_{max} = \frac{\alpha\hbar}{mc} \tag{7}$$

Using these space parameters, all the other parameters including the force can be determined. The equation to find the space parameters that determines the inertia of mass (gravitational radius) exhibits an inversion represented by replacing the mc by the $1/Mc$ term and resulting in $(1/Mc)r^M_{max} = 2\pi\alpha\hbar$. The equations for the minimum and maximum radii exhibit a symmetry with respect to a critical mass, such as the minimum radii for lighter masses (m) can be found using the same expression that finds the maximum radii for heavier masses (M).

$$(1/mc)r^m_{min} = 2\pi\alpha\hbar \tag{8}$$

Similarly we can write: $r_{min}^{M} = \frac{\alpha\hbar}{Mc}$. Quantum effects become evident in small dimensions and the parameter that determines whether the quantum effects are dominant or not, undergoes modifications as the mass of an object changes. The product of minimum and maximum radii for all the masses considered is equal to the square of a critical length scale that is equivalent to the Planck length.

The radius or the length scale found in Eqn. 7 is two orders smaller, or 1/137 of the Compton wavelength (scale where quantum effects become dominant). The length scales employed here can be thought as parameterizing tools for different energy scales.

The equations used for the lighter masses are equivalent to the representation of the real domain, whereas the equations for the heavier masses are equivalent to the representation of the reciprocal domain and this can be seen by the use of the inverse of 'mc' to derive $r_{max}$. An equivalent interpretation of the reciprocal space can be found in the literature denoted as the hologram approach[9]. In that interpretation, the reality is considered to be the projection of the three dimensional reality on a two dimensional screen with spatial frequency equal to the Planck length[10]. The imaging of reality on image planes is in the reciprocal domain whether done through a lens (requiring the Fourier optics analysis) or through a pinhole. An object exhibits its maximum gravitational radiation associated with the maximum force when it has its minimum length scale (with quantum character) and the product of these two quantities is equal to the total energy of an object denoted by $F_{max} r_{min} = mc^2$.

$$F_{max}^{m} = \frac{mc^2}{r_{min}} = \frac{c}{\alpha h} \tag{9}$$

This maximum value of the force is associated with a space parameter that is in the same order with the Planck lengthscale: $\lambda_{real}$. For heavier masses one finds: $F_{max}^{M} = \frac{Mc^2}{r_{min}^{M}} = \frac{M^2c^3}{\alpha\hbar}$. To calculate the force as a function of some real experimentally measurable distance ($r_{real} > \lambda_{real}$) one can use the inverse distance squared relation: $F_{max}\lambda_{real}^2 = F(r)r_{real}^2$. Finally we obtain one expressions for the force that is mass independent ($m < m_{crit}$) and one expression that is proportional to $m^2$ ($m > m_{crit}$). Using the expressions of $r_{min}^{m}$ and $r_{min}^{m}$, one can obtain space parameter $\lambda_{real}$ that is of the order of the Planck length: $r_{min}^{m} r_{max}^{m} = 2\lambda_{real}^2$. The mechanism that shows how the variables are linked is essential.

$$2\lambda_{real}^2 = r_{max} r_{min} = \frac{\alpha\hbar}{mc} mc(2\pi\alpha\hbar) \rightarrow \lambda_{real} = \sqrt{\pi}\alpha\hbar = (1.36)(10^{-36})m$$

If we control $r_{max}^{m}$ for the electron ($m_e < m_{crit}$), and $r_{max}^{M}$ for the black hole in the NGC 1365 galaxy ($m_{BH} > m_{crit}$)[6]:

$$r_{max}^{elec} = \frac{\alpha\hbar}{mc} = \frac{(1/137)(1.054)(10^{-}34)}{(9.1)(10^{-}31)(3)(10^{8})} = (2.8)(10^{-15})m$$

$$r_{max}^{BH} = mc(2\pi\alpha\hbar) = 2(10^{36})(3)(10^{8})\pi\frac{(1.054)(10^{-34})}{137} = (1.4)(10^{9})m$$

The maximum lengthscales that are found for the electron and for the black hole correspond to the lowest energy point. The radius would then decrease with the Lorentz factor as the energy is increased through the same factor. So, if we multiply the maximum radius of the Black hole with the Lorentz factor evaluated at $v = \sqrt{3}/2c$ one finds the Schwarzschild radius $r_{max}\sqrt{1 - v^2/c^2} = (0.7)(10^9)m$. This is the correct radius reported for this black hole confirming the use of the spatial parameters. The radius of the black hole will keep decreasing beyond the Schwarzschild radius as the speed increases and will converge to the Planck radius at the highest energy level. At this point of highest energy level, an object exerts its maximum force (see the expression for $F^m_{max}$ and $F^M_{max}$) on its surface assuming that all of its mass is contained within a Planck length. These units help one to express the gravitational field flux using $F_{max}\lambda^2_{real} = F(r)r^2_{real} = constant$, and the force decreases with the square of the distance. The determination of the maximum radius, minimum radius (equivalent to Planck length) and real physical parameter of an object makes the use classical radius and Schwarzschild radius redundant.

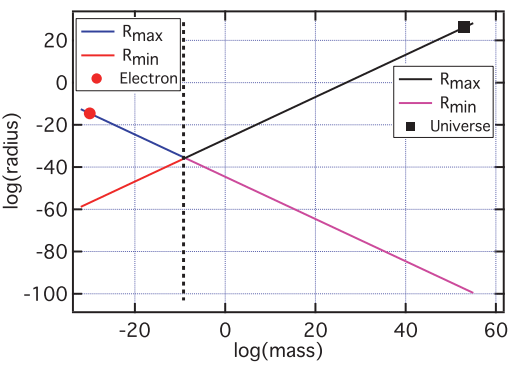



Fig. 3. Space parameters (radii) vs mass.

According to our approach, an object exhibits the maximum force (that is inversely proportional to the square of the physical distance) when it attains its maximum exposed energy and at that point the object has its smallest physical parameters, which are equal to the Planck length. We use the term physical distance so that it is not confused with the minimum and maximum space parameters that have a different character (scaling parameters) rather than a physical character. The space parameters that have the quantum character are used only to determine the energetic state of the object.

### 4.1. *Newton gravitational constant and electrostatic force*

Gravitational constant is the force between two objects each of a mass of 1 kg at a distance of 1 meter. At first we calculate the minimum radius of the object using

the expressions of $r^M_{min}$, $F^M_{max}$.

$$r_{min} = \frac{\alpha\hbar}{mc} = \frac{(7.297)(10^{-3})(1.054)(10^{-34})}{(1.3)(10^8)} = 2.563 \times 10^{-45} m$$

The second step relates the maximum force with the smallest lengthscale according to $F_{max} r_{min} = mc^2$ and resulting to $F_{max} = 3.5 \times 10^{61} kg\ m/sec^2$.

The maximum force corresponds to lengthscale of $(1.36)(10^{-36})$ m and decreases with the square of the distance. As the distance increases from $r_{min}$ to 1 meter, one finds:

$$F = F_{max} \frac{\lambda^2_{real}}{r^2}$$

The result is 98% of the gravitational constant G. If we use the expressions of $r^m_{min}$, $F^m_{max}$ for the electron ($m < m_{crit}$), to determine the force between two electrons at a distance of 1 meter we find a force that is equal numerically to the Coulomb force. $F = \frac{c}{\hbar\alpha} \frac{\lambda^2_{real}}{r^2} = \frac{3(10^8)(137)}{6.6(10^{-34})} \frac{(1.36\times10^{-36})^2}{1^2} = 1.14 \times 10^{-28}$N

$$ke^2 = 9(10^9)(1.6 \times 10^{-19})^2 = 2.3 \times 10^{-28} N$$

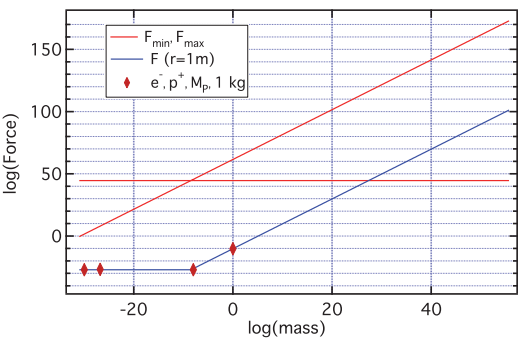



Fig. 4. $F_{min}$ and $F_{max}$ vs mass (in red). $F(r)$ at $r = 1m$ (in blue). Four markers (in red) are forces between two electrons, two protons, two objects with a mass equal to the Planck mass and two objects with mass of 1 kg (equal to gravitational constant $G = 6.53 \times 10^{-11}$).

The calculation of the forces presented here are based on principle that agree with LT dimensional analysis, and this way we show that both the gravitational and nongravitational forces can be derived using a single mechanism. We use spatial parameters with quantum character because they are equivalent to the Compton length and De Broglie wavelength. Below we show the relation between the force as a function of the mass only. An interesting observation is that the force is independent of the mass for objects lighter than the Planck's mass. Using the above formula one can calculate a force of around $10^2 N$ between two nucleons at a distance of $r \approx 10^{-15}$m. Using the model presented in this work that relates all the forces at

the smallest lengthscale one can find the gravitational and nongravitational force. It is known that gravitation coverges to the magnitude of the other forces at the smallest lengthscales. We found the forces using only the mass and did not use three parameters (k, e, G), thus eliminating the difference between the gravitational and nongravitational forces.

## 5. Conclusions

We established the idea that the correct expression of energy (the exposed energy) helps derive the expressions of other physical quantities (momentum and force). Using the exposed energy in the Lagrangian formalism the equations of motion for both the relativistic and classical physics were derived. This expression provides the insight that the experiments measure only the exposed energy and thus a lower mass. A lower mass as the system expands, which can be predicted using the exposed energy may be one of the alternatives to the dark matter and to the modified Newtonian dynamics (MOND)[8]. The major problem of MOND is that it cannot predict the lower mass in the system being analyzed. The exposed energy also provides a model on how to predict the upper limit of the spin of the black holes, and also the gravitational flux beyond the Schwarzschild radius.

Here in this work spatial parameters are used to derive important parameters for elementary particles and for celestial objects. These parameters exhibit an inversion at a mass that is analogous with the Planck mass. The calculation of the gravitational constant and of the electrostatic force between two elementary particles were provided as two examples to confirm the validity of the proposed mechanism. The same set of equations was used to determine the Schwarzschild radii of the black holes. The derivation of the gravitational force employs a mathematical step – denoted here as the reciprocal domain of the spatial domain – that is equivalent to the approach used by several researchers to describe the holographic representation of three dimensional reality. The need for such a set of spacetime parameters with a quantum character can help the researchers to measure other important parameters of matter. The rare equality between the physical and quantum spacetime parameters may have hindered our path to complete the full picture. For example, the accepted radius of the electron ($\approx 10^{-15}m$) is equal to $(1/137)\lambda_C$, where $\lambda_C$ is the Compton wavelength. In the case of the electron, classical radius corresponds to its physical size, but this is not always the case for the other elementary particles.

## References

1. G. Altarelli, "The Higgs: so simple yet so unnatural," EPJ Web Conf. **71** (2014), 00005 doi:10.1051/epjconf/20147100005
2. L. De. Broglie, "The thermodynamics of the isolated particle," Gauthier-Villars Editor, Paris, 1964.
3. R. O. di Bartini, "Relations between physical constants," Progress in Physics, 3(10):34-40, 2005.

4. J. A. Wheeler, "On the Nature of quantum geometrodynamics," Annals Phys. **2** (1957), 604-614 doi:10.1016/0003-4916(57)90050-7
5. E. Aprile *et al.* [XENON], "Dark Matter Search Results from a One Ton-Year Exposure of XENON1T," Phys. Rev. Lett. **121** (2018) no.11, 111302 doi:10.1103/PhysRevLett.121.111302 [arXiv:1805.12562 [astro-ph.CO]].
6. Risaliti, G. and Harrison, F. A. and Madsen, K. K. and Walton, D. J. and Boggs, S. E. and Christensen, F. E. and Craig, W. W. and Grefenstette, B. W. and Hailey, C. J. and Nardini, E. and et al., "A rapidly spinning supermassive black hole at the centre of NGC 1365", Nature, 494, 7438:449-451.
7. G. Amelino-Camelia, "Doubly special relativity: First results and key open problems," Int. J. Mod. Phys. D **11** (2002), 1643 doi:10.1142/S021827180200302X [arXiv:gr-qc/0210063 [gr-qc]].
8. L. D. Bekenstein, "Relativistic gravitation theory for the MOND paradigm," Phys. Rev. D **70** (2004), 083509 doi:10.1103/PhysRevD.70.083509.
9. L. Suskind, "The world as a hologram," Journal of Mathematical Physics, 36(11):6377-6396, (1995).
10. E. P. Verlinde, "On the Origin of Gravity and the Laws of Newton," JHEP **04** (2011), 029 doi:10.1007/JHEP04(2011)029

# Report on session AT7 of the 15th Marcel Grossmann Meeting — "Theories of gravity: alternatives to the cosmological and particle standard models"

Stefano Bellucci

*INFN-Laboratori Nazionali di Frascati, Frascati, Italy*
*E-mail: Stefano.Bellucci@lnf.infn.it*

Valerio Faraoni

*Dept. of Physics & Astronomy and STAR Research Cluster, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*
*E-mail: vfaraoni@ubishops.ca*

Orlando Luongo

*INFN-Laboratori Nazionali di Frascati, Frascati, Italy*
*E-mail: Orlando.Luongo@lnf.infn.it*

The parallel session AT7 of the 15th Marcel Grossmann Meeting held in Rome, Italy on July 1-7, 2018 is summarized. In particular, we briefly discuss each talk, highlighting the principal results and the most intriguing perspectives.

*Keywords*: Alternative theories of gravity; dark energy; dark matter; particle physics.

## 1. Introduction

The parallel session AT7, "Theories of gravity: alternatives to the cosmological and particle standard models" was devoted to alternative scenarios to the standard models of cosmology and particle physics. It run on Monday, July 2 and Tuesday, July 3, 2018 and included nineteen speakers. There was no poster session. There were several motivations for holding this session at the 15th Marcel Grossmann Meeting. To begin with, in the standard model of cosmology the $\Lambda$-Cold Dark Matter ($\Lambda$CDM) model based on Einstein's General Relativity, dark energy is introduced completely *ad hoc* in order to explain the present acceleration of the universe. This standard model requires also the introduction of dark matter dominating (by far) ordinary baryonic matter but yet undetected in the laboratory, and it suffers from phenomenological problems. The simplest explanation to account for the cosmic speed up lies on considering a vacuum energy cosmological constant, whose origin comes from quantum fluctuations at early times. Modifying gravity may give an interplay between cosmology and quantum field theory and represents a possible alternative, albeit a wide number of proposals have been presented in recent years. Likewise, the standard model of particle physics is unable to incorporate all of the current particle phenomenology and proposed dark matter candidates. Further, the model cannot match cosmological observations in terms of expected vacuum energy and it seems to be unable to predict the required cosmological constant or any dark energy terms. In this respect, cosmology and particle physics come together

in the early universe and, surprisingly, also in theories and models of the present, accelerating universe. The main target of present-time researches is to have a self-consistent scheme which incorporates quantum field theory and general relativity in a unique scenario.

Session AT7 was designed in a wide framework to include several topics related to these problems, and spanning alternative theories of gravity and cosmology, alternatives to the ΛCDM model, quantum field theory applied to gravity, extensions of the standard model of particle physics, and dark energy and dark matter from a particle physics point of view. This session represents the interplay between, and the efforts to match, particle physics and cosmology, giving particular emphasis to the role played by particle quantum field theory in the early and the late universe.

## 2. Cosmology and dark energy

Stefano Bellucci presented a contribution entitled *Constraints on chameleon dark energy model from SN Ia and BAO observations*. He considered a scalar field, known as chameleon, *i.e.*, a scalar field non-minimally coupled to matter. Such a chameleon field not only can be heavy enough in the laboratory tests so that the local gravity constraints are fulfilled but also it can be light enough at cosmological scales and considered as a dark energy candidate. In his talk Bellucci used the latest observational data to constrain the parameters of the tachyonic chameleon model of dark energy. Two important functions in the analysis were considered, *i.e.*, the tachyonic potential $V$ and the non-minimal coupling function $f$ between the scalar field and matter. In general Bellucci and collaborators considered two types of these functions, *i.e.*, power-law and exponential forms. They fitted data from Type Ia supernova (SN Ia) and Baryon Acoustic Oscillations (BAO) to obtain constraints on the present matter density parameter and the free parameters in the functional form of the functions $f$ and $V$, thus showing that the scenario is compatible with observations[1].

Hernando Quevedo discussed *geometrothermodynamic cosmology* by applying the formalism of geometrothermodynamics to derive a fundamental thermodynamic equation. In conjunction with the Friedmann equations, this was used to construct a cosmological model. The main features of this scenario are an early phase with thermodynamic interaction, corresponding to inflation, and a late phase without thermodynamic interaction, equivalent to the standard ΛCDM model[2].

Orlando Luongo presented *Speeding up the universe using dust with pressure*. He showed that the standard cosmological model can be *updated* assuming a mechanism which cancels vacuum energy out. This enables the universe to accelerate today without the need of some sort of "quantum fluctuating" cosmological constant. To do so, he suggested that standard matter, *i.e.*, baryons and cold dark matter, can show a non-vanishing pressure. Soon after a spontaneous symmetry breaking, one recovers two cases: the first, *i.e.*, the ΛCDM model, is disfavored with respect to the second case, *i.e.*, an effective dark fluid approach, in which the universe

accelerates with a fine-tuned emergent cosmological constant. To model the matter fluid able to realize the above mechanism, he introduced a Lagrangian composed of a Lagrange multiplier and a generalized kinetic term. Further, to account for the phase transition the model presumes a fourth order potential, in which the scalar field is the one associated to matter only. Evaluating the Helmoltz free energy, it is possible to constrain the sign of the pressure, instead of postulating exotic dark energy terms. Further, once the shift symmetry is invoked on the field, the associated Noether currents show that the entropy density is conserved and particle production does not contribute, in proximity of each minima. Luongo showed the net fluid is therefore irrotational invoking the Carter-Lichnerowicz equations. From small perturbations, after cumbersome algebra, he showed that the pressure turns out to be constant as the entropy shift suggests. In such a way, the adiabatic sound speed vanishes, in agreement with previous dark fluid approaches. As a byproduct, the numerical bounds on the pressure and matter densities fulfill modern observations and, under the hypothesis of thermal equilibrium, the dark matter candidate turns out to be a highly massive WIMP particle whose mass lies in the interval[3] $0.5 - 1.7$ TeV.

Behnaz Fazlpour talked about *Cosmological constraints on tachyonic teleparallel dark energy model*. She started from teleparallel gravity, reviewing that it is an equivalent formulation of general relativity in which instead of the Ricci scalar $R$, one uses the torsion scalar $T$ for the Lagrangian density. She showed that by adding a quintessence scalar field, and allowing a non-minimal coupling to gravity, the so-called teleparallel dark energy emerges as a natural fact. It has been proposed in hher slides, and it was found, that such a non-minimally coupled quintessence theory has a richer structure than other approaches based on general relativity. The work employs tachyonic teleparallel dark energy in which a tachyon scalar field is responsible for dark energy. She focused on constraining this model with the most recent observations of Type Ia supernova (SN Ia) and Baryon Acoustic Oscillations measurements and she showed that the scenario is compatible with observations[4].

Ali Banijamali presented *Dynamics and observational constraints on a scalar-tensor model with Gauss-Bonnet coupling*. He discussed the dynamics of a scalar-tensor model of dark energy in which the scalar field, playing the role of dark energy, is non-minimally coupled to the Gauss-Bonnet invariant in four dimensions. He showed how to use the dynamical system approach to extract the critical points, finding scaling attractor solutions with the property that the ratio of the dark energy and dark matter density parameters are comparable. These solutions may alleviate the coincidence problem in cosmology and may give an explanation of it. Further, constraints on the free parameters of the model have been computed using different sources of cosmological data. In this respect, the viability of the model has been explored combining different data sets, among all Type Ia supernova (SN Ia) and Baryon Acoustic Oscillations[5].

Antonino Marciano talked about *Emergent inflation from a Nambu-Jona-Lasinio mechanism in gravity with non-dynamical torsion*. He discussed a new scenario of early universe inflation emerging from a Nambu-Jona-Lasinio mechanism in the presence of non-dynamical torsion, which couples to Standard Model fermions. Without introducing new scalar fields, in this model inflation emerges from a four-fermion interaction induced by the torsion. The role of the inflaton field is played by a (nonminimally coupled) composite field of Standard Model particles arising from a curved space Nambu-Jona-Lasinio mechanism. Definite predictions for gravitational waves in this inflationary model were also presented[6].

## 3. $f(R)$, $f(\phi, R)$, and metric-affine gravity

The class of $f(\phi, R)$ theories of gravity (where $R$ is the Ricci scalar and $\phi$ is a gravitational scalar field) is very large. Fayçal Hammad discussed scalar and tensor perturbations in this general class, talking about *Density Perturbations in $f(R, \varphi)$-Gravity with an Application to the (Varying Power)-Law Model*. Within the specific model of varying power-law modified gravity, he computed tensor-to-scalar ratio, scalar tilt, and tensor tilt. In order to be compatible with current cosmological observations, this model requires inflation to occur at an energy scale less than the GUT energy scale. It was claimed that the observed density perturbations are obtained without fine tuning.

Restricting to the (still very broad) class of $f(R)$ theories of gravity, Rocco D'Agostino discussed a model-independent method to reconstruct the Lagrangian function $f(R)$ from cosmological observations, in both the metric and the Palatini formalisms in a talk entitled *Model-independent reconstructions of $f(R)$ gravity*. The method relies on building model-independent approximations of the luminosity distance by means of rational polynomials, in order to extend the convergence radius of the standard cosmographic series (Padé polynomials and ratios of Chebyshev polynomials are used)[7,8]. The function $f(R)$ is reconstructed from the Hubble function versus redshift $H(z)$ by first reconstructing $f(z)$ functions and then inverting the relation $R = R(z)$. The results of this method imply small deviations from the $\Lambda$CDM model and a time-evolving effective dark energy equation of state.

Roberto Giambò spoke about perfect fluids in $f(R)$ theories, presenting analytical results on the qualitative behavior of Friedmann-Lemaître-Robertson-Walker (FLRW) cosmologies in theories coupled nonminimally to matter in his talk: *Late Time Evolution of Perfect Fluids Universes in f(R) Theories of Gravity*. Assumptions on the scalar field potential arising from conformal compactification and on the fluid equation of state determine the late time behavior of the universe (big crunch if matter dominates over the scalar field, or perpetual expansion in the opposite situation)[9].

Again in the context of $f(R)$ gravity with a nonminimal coupling of $R$ to matter, Riccardo March reviewed models and Solar System constraints. In the talk *Nonminimally Coupled Curvature-Matter Gravity Models and Solar System constraints* that

he presented, an older perturbative approach due to Chiba *et al.*[10] was applied in[11] to nonminimally coupled $f(R) = R^{-n}$ gravity (with $n > 0$). It was claimed that, while pure $R^{-n}$ gravity violates Solar System constraints, the nonminimally coupled version of this theory does not. Another nonminimally coupled model admitting a Minkowski solution was found to possess a non-Newtonian Yukawa potential in the non-relativistic limit. A study of non-geodesic trajectories of test particles around a static spherical body was also reported.

In the broad context of metric-affine theories of gravity, concerning the contribution *Constraining metric-affine gravity with particle scattering*, Adriá Delhom I Latorre discussed how to probe non-metricity with particle scattering and how to constrain Ricci-based gravity from scattering data. Arguing that torsion and non-metricity could originate from new physics associated with defects in a hypothetical spacetime microstructure, he showed effects of non-metricity which are presumably observable in quantum fields in the form of 4-fermion contact interactions. Bhabha scattering then constrains the scale of non-metricity to be larger than 1 TeV (that is, orders of magnitude higher than previous experimental constraints)[12].

## 4. Symmetry violations

Deviations from the Standard Model of particle physics may take the form of symmetry violations. In a talk entitled *CP violation in space*, Giovanni Maria Piacentino discussed a new proposal to probe directly the gravitational acceleration of antimatter by measuring the branching fraction of CP-violating $K_L$ decays in space. He showed that, at the altitude of the International Space Station, gravitational effects could change the level of CP violation to the point that a $5-\sigma$ discrimination would be possible by collecting the $K_L$ produced by the cosmic proton flux within a few years.

Gravitational waves in the presence of string axion dark matter and the gravitational Chern-Simons coupling were instead the subject of Daiske Yoshida's contribution, *Exploring the string axiverse and parity violation in gravity with gravitational waves*. He reported on two recent papers[13,14] and discussed how parametric resonance of gravitational waves occurs due to coherent oscillation of the string axion and how circular polarization of gravitational waves is induced by the Chern-Simons coupling. The observational constraints set on the model by the failure to observe these two effects were also presented.

## 5. Quantum gravity-inspired scenarios

In a talk entitled *The gravity's rainbow cosmological constant*, Remo Garattini reported on his computation of the zero point energy due to quantum fluctuations around a fixed background using a reformulation of the Wheeler-DeWitt equation. Using a variational approach with Gaussian trial wave functionals, the first loop contribution of the graviton to the zero point energy was computed, controlling

ultraviolet divergences with a distorted gravitational field (the "gravity rainbow'). This finite zero point energy was identified with the induced cosmological constant.

Gabriele Gionti presented recent results of a Hamiltonian analysis of asymptotically safe quantum gravity in his talk entitled *Bouncing and Emergent Universes from Hamiltonian Analysis of Asymptotically Safe Quantum Gravity.* While renormalization group (RG) approaches to quantum gravity suggest that Newton's constant $G$ and the cosmological constant $\Lambda$ run with the characteristic energy scale, it is not clear how to embed this modification of Einstein theory in Dirac's theory of constrained systems. The Hamiltonian formalism for RG scale-dependent $G$ and $\Lambda$ were discussed with regard to Dirac's constraint analysis. It was reported that, under certain conditions, the algebra of Dirac constraints is closed. Brans-Dicke theory was also studied as a Dirac constrained system and compared with modified gravity via asymptotically safe quantum gravity. Assuming asymptotic safety, the results were applied to the early universe: in the minisuperspace with FLRW geometry, the RG-improved Friedmann equations exibit solutions describing bouncing and emergent universes. While, classically, emergent universes are only found in closed topology (curvature index $K = +1$), in the sub-Planckian regime they exist also for flat and open ($K = 0, -1$) topologies.

## 6. Miscellaneous

Sayantan Choudhury talked about *COSMOS-e-soft Higgsotic attractors.* He showed an algorithm to study the cosmological consequences from a huge class of quantum field theories (*i.e.*, superstring theory, supergravity, extra-dimensional theory, modified gravity, *etc.*), which are equivalently described by soft attractors in the effective field theory framework. He restricted the analysis to two scalar fields — dilaton and Higgsotic fields minimally coupled with Einstein gravity, which can be generalized to any arbitrary number of scalar field content with generalized non-canonical and non-minimal interactions. He used $R^2$ gravity to study the attractor and non-attractor phase and he presented also theoretical bounds on the amplitude, tilt and running of the primordial power spectrum and other quantities of interest. Additionally, three possible theoretical proposals have been presented to overcome the tachyonic instability at the time of late time acceleration. Last but not least, he provided also the bulk interpretation from the three- and four-point scalar correlation functions for completeness[15].

Andrea Addazi discussed aspects of black hole physics in extended theories of gravity. In his presentation, he focussed on classical evaporation and anti-evaporation instabilities, discussing their implications for black hole thermodynamics. He put particular emphasis on their connections with deformed energy conditions.

Aram Saharian spoke about *Current densities in braneworlds with compact dimensions.* There, the vacuum expectation values (VEVs) of the current densities for charged scalar and fermionic fields were investigated in braneworlds on AdS

spacetime with toroidally compact spatial dimensions and in the presence of a constant Abelian gauge field. Because of the non-trivial topology, the latter gives rise to an Aharonov-Bohm type effect for currents. The VEVs are thus decomposed into the purely AdS and brane-induced contributions. Both contributions are periodic functions of the magnetic flux enclosed by compact dimensions, with the period equal to the flux quantum. The asymptotic behavior of the current density along compact dimensions was investigated near the branes, near the AdS boundary, and near the horizon. In the talk, it was shown that, unlike the case of the Minkowski bulk, the total current density in supersymmetric models on an AdS bulk does not vanish.

Arban Uka in *New perspective in achieving the unification theory according to the ideas of Einstein and Wheeler* dealt with a philosophical conjecture between relativity and quantum physics.

## 7. Conclusions

The session AT7 represented the principal session in which different aspects of particle physics matched cosmology in a broad sense. The scope of the session was to stress the strong need of unifying quantum field theory with general relativity in a future context of a more universal theory. All the talks provided hints and represented first attempts toward this ambitious target.

## Acknowledgments

V.F. is supported by the Natural Sciences and Engineering Research Council of Canada (grant no. 2016- 03803). O.L. is supported in part by the Ministry of Education and Science of the Republic of Kazakhstan, Program 'Fundamental and applied studies in related fields of physics of terrestrial, near-earth and atmospheric processes and their practical application' IRN: BR05236494.

## References

1. A. Banijamali, S. Bellucci, B. Fazlpour, M. Solbi, Astrophys. Sp. Sci., **363**, 59, (2018).
2. C. Gruber, H. Quevedo, JCAP, **07**, 032, (2017).
3. O. Luongo, M. Muccino, Phys. Rev. D, **98**, 103520, (2018).
4. A. Banijamali, B. Fazlpour, Astrophys. Sp. Sci., **342**, 1, 229, (2012).
5. S. Bellucci *et al.*, arXiv:1807.04080
6. A. Addazi, P. Chen, A. Marciano, Eur. Phys. J. C, **79**, 4, 297, (2019).
7. S. Capozziello, R. D'Agostino, O. Luongo, JCAP, **05**, 008, 1805, (2018).
8. S. Capozziello, R. D'Agostino, O. Luongo, Mon. Not. Roy. Astron. Soc., **476**, 3, 3924, (2018).
9. R. Giambo, G. Magli, Class. Quant. Grav., **31**, 3, 035016, (2014).
10. T. Chiba, T.L. Smith, and A.L. Erickcek, Phys. Rev. D, **75**, 124014, (2007).

11. O. Bertolami, P. Frazão, and J. Pãramos, Phys. Rev. D, **81**, 104046, (2010).
12. A. Delhom, G. J. Olmo, M. Ronco, Phys. Lett. B, **780**, 294. (2018).
13. D. Yoshida and J. Soda, Int. J. Mod. Phys. D, **27**, 1850096, (2018).
14. J. Soda and D. Yoshida, Galaxies, **5**, 96, (2017).
15. S. Choudhury, Eur. Phys. J. C, **77**, 7, 469, (2017).

# Multi-band width gravitational-wave astronomy with intermediate-mass ratio inspirals

Pau Amaro-Seoane*

*Institute of Space Sciences (ICE, CSIC) & Institut d'Estudis Espacials de Catalunya (IEEC) at Campus UAB, Carrer de Can Magrans s/n 08193 Barcelona, Spain*
*Kavli Institute for Astronomy and Astrophysics, Beijing 100871, China*
*Institute of Applied Mathematics, Academy of Mathematics and Systems Science, CAS, Beijing 100190, China*
*Zentrum für Astronomie und Astrophysik, TU Berlin, Hardenbergstraße 36, 10623 Berlin, Germany*
**E-mail: pau@ice.cat*
*astro-gr.org*

Stellar-mass black holes come from the deaths of massive stars, and have been observed both in our own Galaxy and in other galaxies. Supermassive black holes have masses of up to 10,000,000,000 times the mass of the Sun and reside at the centers of galaxies. Intermediate-mass black holes (IMBHs), with masses in between these two regimes, are elusive and represent the missing link. When a compact object ventures too close the IMBH, it is captured because of the emission of gravitational waves, to be eventually swallowed whole when it crosses the event horizon. The process, an intermediate-mass ratio inspiral, represents a mapping of warped spacetime and can be observed with existing gravitational-wave ground-based detectors. Depending on the dynamical parameters, IMRIs can be detected jointly with space-borne ones, which will allow us to do multi-band width gravitational-wave astronomy.

*Keywords*: Gravitational waves.

## 1. Introduction

Contrary to the Newtonian description of gravity, in the theory of relativity compact objects can lose energy via the emission of gravitational radiation. A small compact object such as a stellar black hole, a neutron star, or a white dwarf, can therefore form a bound system with a massive black hole because of the emission of gravitational waves[1,2]. Since the binary loses energy, the semi-major axis shrinks which, combined with the precession and the change of inclination, represent a unique probe of spacetime around massive black holes in strong gravity[3]. The number of cycles that the small compact object will have to go until the plunge through the event horizon is roughly proportional to the mass ratio. Because this mass ratio is at least of five orders of magnitude, we call these systems extreme-mass ratio inspirals, EMRIs.

At such mass ratios, similar mass-ratio techniques to derive the waveforms are not useful, to say the least, because self-force is crucial[4]. But the problem is even more challenging than that, because the binary is not isolated. Surrounding it, the host environment can contain up to a few tens of millions of stars[2]. Since EMRI orbits are very eccentric, this poses another challenge at apocenter, where the bulk of the stellar system might have an impact at certain semi-major axes. This means that, in order to address the EMRI problem, we need to take into account not

"just" the relativistic two-body problem at these extreme mass ratios, but also the a-few-tens-of-millions body problem.

In spite of all of these problems, there has been impressive progress over the last two decades[2]. Even if we do not have a waveform which describes in precision this systems, there are good enough approaches that have proven to be accurate enough to ensure detection and, taking into account the restrictions of the different approaches, parameter extraction[5].

EMRIs are formed in galactic nuclei via two-body relaxation, which allows us to address this problem, within limitations, in an analytical way thanks to the analysis of the loss-cone[6–9]. This relies on the fact that the massive black hole is fixed at the centre of the system. Lighter-mass black holes, i.e. intermediate-mass black holes IMBHs, see[10,11] for reviews, with masses $\gtrsim 100\,M_\odot$ (but less massive than supermassive black holes) therefore pose a problem. We rely on numerical simulations to obtain a description of the process of formation and evolution of intermediate-mass ratio inspirals (IMRIs). This is unfortunate, because one numerical simulation of a cluster containing a realistic number of stars can last a few months of time to achieve a relevant amount of time[12]. The first numerical simulation of a globular cluster with an IMBH that formed an IMRI was performed by[13]. Their findings were corroborated by[14] and also the work of[15], which is almost a replica of their numerical setup. Also,[16] explore compact binaries forming with IMBHs, but in a lighter regime. These different numerical simulations show that IMRIs form at very small semi-major axes, $a \sim 10^{-5}$ pc, and very high eccentricities, $e \gtrapprox 0.999$, very likely (since the data writing does not capture the small timescale for this particular event) following the hyperbolic capture scenario described by[17], and see[18–22] for more recent works.

## 2. Cascade of harmonics

In Fig. (1) we show two typical examples of IMRIs forming at low-frequencies, within the LISA bandwidth, and merging in the range of ground-based detectors, such as the ET[23,24] or LIGO/Virgo. Because of the inspiral, it would be ideal to have a decihertz detector operating simultaneously to LISA and LIGO/Virgo such as the DECi-hertz Interferometer Gravitational Wave Observatory[25], the Superconducting Omni-directional Gravitational Radiation Observatory SOGRO, see[26,27] or the geocentric Tian Qin[28]. We can see that LISA can forewarn ground-based detectors up to five years in advance, and tenths of seconds of precision, so that ET (or LIGO/Virgo, depending on the parameters of the source) can optimise their sensitivity to optimise the detection.

## 3. Conclusions

IMRIs forming in globular clusters and have very high eccentricities and small semi-major axes, and are detectable both with space-borne observatories and ground-based detectors.

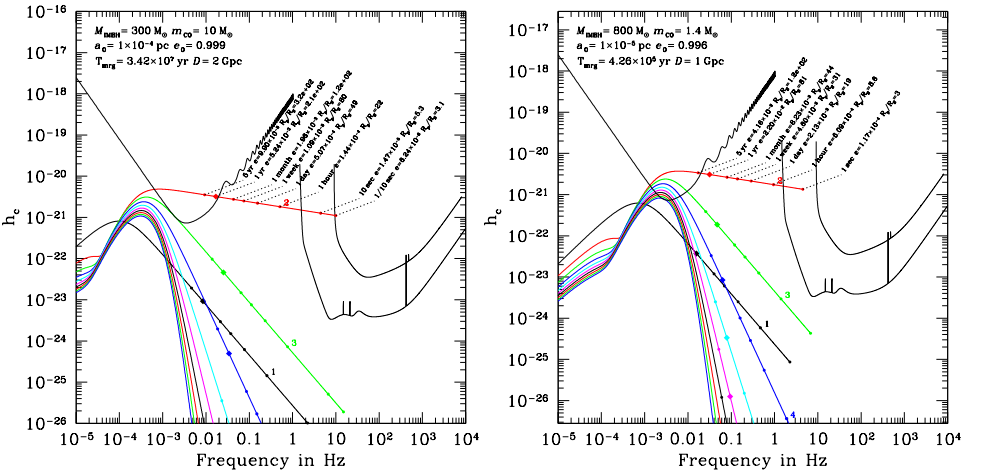



Fig. 1. Strain of the gravitational wave as a function of the frequency. The gravitational wave is approximated in the quadrupolar derivation of [29] in different harmonics, and we only show the first ten. The left curve corresponds to the LISA sensitivity curve and the lower-frequency, U-shaped curve to the Einstein Telescope, and the higher-frequency curve to Advanced LIGO. In each panel we give the initial parameters for the binary; namely the mass of the IMBH, $M_{\rm IMBH}$ in solar masses, the mass of the compact object, $m_{\rm CO}$, the initial semi-major axis, $a_0$ in pc, the initial eccentricity $e_0$, the distance to the source, $D$, as well as the timescale for merger, $T_{\rm mrg}$, in years. Along the second harmonic, which is the predominant one at higher frequencies, we show different values of the eccentricity and the periapsis in Schwarzschild radii, $R_{\rm S}$ at different moments in the evolution of the IMRI. The left panel corresponds to an IMBH of mass 300 $M_\odot$ and a stellar-mass black hole of 10 $M_\odot$ at a distance of 2 Gpc. The right panel depicts a binary of 800 $M_\odot$ and a neutron star of 1.4 $M_\odot$ at a distance of 1 Gpc.

The combined detection of a source of gravitational waves has been addressed in the literature previously [30–34], but with mass ratios not larger than 10. The joint detection of similar mass ratios for small compact objects, as motivated by the first detections of LIGO/Virgo, have been discussed by [35], and [36], who showed that a multi-bandwidth detection of this kind of binaries is only possible for circular binaries, so that the information about the formation channel (dynamical or via stellar evolution) cannot be obtained.

Conjunct IMRI detections from the space and the ground will allow us to place improved constraints on the physical parameters. A combined detection enables the possibility of splitting the different degeneracies as well. LISA can detect the inspiral with enough signal-to-noise ratio [37] and hence it can forewarn ground-based detectors years in advance with seconds and even tenths of seconds of precision. The ET and LIGO/Virgo can detect the merger with a very large SNR in the former case, of up to a few thousand, and LIGO/Virgo with more modest values, but still well above the threshold [37]. If LISA is operating at the same time as ET or LIGO/Virgo, one should think of them as a single detector. These results are discussed in more detail in the recent work of [37].

**Acknowledgments**

PAS acknowledges support from the Ramón y Cajal Programme of the Ministry of Economy, Industry and Competitiveness of Spain, as well as the COST Action GWverse CA16104. This work was supported by the National Key R&D Program of China (2016YFA0400702) and the National Science Foundation of China (11721303).

**References**

1. P. Amaro-Seoane, J. R. Gair, M. Freitag, M. C. Miller, I. Mandel, C. J. Cutler and S. Babak, Intermediate and extreme mass-ratio inspirals. Astrophysics, science applications and detection using LISA, *Classical and Quantum Gravity* **24**, 113 (September 2007).
2. P. Amaro-Seoane, Relativistic dynamics and extreme mass ratio inspirals, *Living Reviews in Relativity* **21**, p. 4 (May 2018).
3. P. Amaro-Seoane, J. R. Gair, A. Pound, S. A. Hughes and C. F. Sopuerta, Research Update on Extreme-Mass-Ratio Inspirals, *Journal of Physics Conference Series* **610**, p. 012002 (May 2015).
4. L. Barack and A. Pound, Self-force and radiation reaction in general relativity, *Reports on Progress in Physics* **82**, p. 016904 (Jan 2019).
5. S. Babak, J. G. Baker, M. J. Benacquista, N. J. Cornish, S. L. Larson, I. Mandel, S. T. McWilliams, A. Petiteau, E. K. Porter, E. L. Robinson, M. Vallisneri, A. Vecchio, t. M. L. Data Challenge Task Force, M. Adams, K. A. Arnaud, A. Błaut, M. Bridges, M. Cohen, C. Cutler, F. Feroz, J. R. Gair, P. Graff, M. Hobson, J. Shapiro Key, A. Królak, A. Lasenby, R. Prix, Y. Shang, M. Trias, J. Veitch, J. T. Whelan, The Mock LISA Data Challenges: from challenge 3 to challenge 4, *Classical and Quantum Gravity* **27**, 084009 (April 2010).
6. J. Frank and M. J. Rees, Effects of massive central black holes on dense stellar systems, *MNRAS* **176**, 633 (September 1976).
7. A. P. Lightman and S. L. Shapiro, The distribution and consumption rate of stars around a massive, collapsed object, *ApJ* **211**, 244 (January 1977).
8. T. Alexander and M. Livio, Tidal Scattering of Stars on Supermassive Black Holes in Galactic Centers, *ApJ Lett.* **560**, 143 (October 2001).
9. P. Amaro-Seoane and R. Spurzem, The loss-cone problem in dense nuclei, *MNRAS* **327**, 995 (November 2001).
10. M. Mezcua, Observational evidence for intermediate-mass black holes, *International Journal of Modern Physics D* **26**, p. 1730021 (2017).
11. N. Lützgendorf, M. Kissler-Patig, N. Neumayer, H. Baumgardt, E. Noyola, P. T. de Zeeuw, K. Gebhardt, B. Jalali and A. Feldmeier.
12. L. Wang, R. Spurzem, S. Aarseth, M. Giersz, A. Askar, P. Berczik, T. Naab, R. Schadow and M. B. N. Kouwenhoven, The DRAGON simulations: globular cluster evolution with a million stars, *MNRAS* **458**, 1450 (May 2016).

13. S. Konstantinidis, P. Amaro-Seoane and K. D. Kokkotas, Investigating the retention of intermediate-mass black holes in star clusters using N-body simulations, *A&A* **557**, p. A135 (September 2013).
14. N. W. C. Leigh, N. Lützgendorf, A. M. Geller, T. J. Maccarone, C. Heinke and A. Sesana, On the coexistence of stellar-mass and intermediate-mass black holes in globular clusters, *MNRAS* **444**, 29 (October 2014).
15. C.-J. Haster, F. Antonini, V. Kalogera and I. Mandel, N-Body Dynamics of Intermediate Mass-ratio Inspirals in Star Clusters, *ApJ* **832**, p. 192 (December 2016).
16. M. MacLeod, M. Trenti and E. Ramirez-Ruiz, The Close Stellar Companions to Intermediate-mass Black Holes, *ApJ* **819**, p. 70 (March 2016).
17. R. O. Hansen, Post-Newtonian Gravitational Radiation from Point Masses in a Hyperbolic Kepler Orbit, *Phys. Rev. D* **5**, 1021 (February 1972).
18. B. Kocsis, M. E. Gáspár and S. Márka, Detection Rate Estimates of Gravity Waves Emitted during Parabolic Encounters of Stellar Black Holes in Globular Clusters, *ApJ* **648**, 411 (September 2006).
19. I. Mandel, D. A. Brown, J. R. Gair and M. C. Miller, Rates and Characteristics of Intermediate Mass Ratio Inspirals Detectable by Advanced LIGO, *ApJ* **681**, 1431 (July 2008).
20. R. O'Leary, B. Kocsis and A. Loeb, Gravitational waves from scattering of stellar-mass black holes in galactic nuclei, *MNRAS* **395**, 2127 (June 2009).
21. W. H. Lee, E. Ramirez-Ruiz and G. van de Ven, Short Gamma-ray Bursts from Dynamically Assembled Compact Binaries in Globular Clusters: Pathways, Rates, Hydrodynamics, and Cosmological Setting, *ApJ* **720**, 953 (September 2010).
22. J. Hong and H. M. Lee, Black hole binaries in galactic nuclei and gravitational wave sources, *MNRAS* **448**, 754 (March 2015).
23. B. Sathyaprakash, M. Abernathy, F. Acernese, P. Ajith, B. Allen, P. Amaro-Seoane, N. Andersson, S. Aoudia, K. Arun, P. Astone and et al., Scientific objectives of Einstein Telescope, *Classical and Quantum Gravity* **29**, p. 124013 (June 2012).
24. S. Hild, M. Abernathy, F. Acernese, P. Amaro-Seoane and et al, Sensitivity studies for third-generation gravitational wave observatories, *Classical and Quantum Gravity* **28**, p. 094013 (May 2011).
25. S. Kawamura, M. Ando, N. Seto, S. Sato, T. Nakamura, K. Tsubono, N. Kanda, T. Tanaka, J. Yokoyama, I. Funaki, K. Numata, K. Ioka, T. Takashima, K. Agatsuma, T. Akutsu, K.-S. Aoyanagi, K. Arai, A. Araya, H. Asada, Y. Aso, D. Chen, T. Chiba, T. Ebisuzaki, Y. Ejiri, M. Enoki, Y. Eriguchi, M.-K. Fujimoto, R. Fujita, M. Fukushima, T. Futamase, T. Harada, T. Hashimoto, K. Hayama, W. Hikida, Y. Himemoto, H. Hirabayashi, T. Hiramatsu, F.-L. Hong, H. Horisawa, M. Hosokawa, K. Ichiki, T. Ikegami, K. T. Inoue, K. Ishidoshiro, H. Ishihara, T. Ishikawa, H. Ishizaki, H. Ito, Y. Itoh, K. Izumi,

I. Kawano, N. Kawashima, F. Kawazoe, N. Kishimoto, K. Kiuchi, S. Kobayashi, K. Kohri, H. Koizumi, Y. Kojima, K. Kokeyama, W. Kokuyama, K. Kotake, Y. Kozai, H. Kunimori, H. Kuninaka, K. Kuroda, S. Kuroyanagi, K.-I. Maeda, H. Matsuhara, N. Matsumoto, Y. Michimura, O. Miyakawa, U. Miyamoto, S. Miyoki, M. Y. Morimoto, T. Morisawa, S. Moriwaki, S. Mukohyama, M. Musha, S. Nagano, I. Naito, K. Nakamura, H. Nakano, K. Nakao, S. Nakasuka, Y. Nakayama, K. Nakazawa, E. Nishida, K. Nishiyama, A. Nishizawa, Y. Niwa, T. Noumi, Y. Obuchi, M. Ohashi, N. Ohishi, M. Ohkawa, K. Okada, N. Okada, K. Oohara, N. Sago, M. Saijo, R. Saito, M. Sakagami, S.-I. Sakai, S. Sakata, M. Sasaki, T. Sato, M. Shibata, H. Shinkai, A. Shoda, K. Somiya, H. Sotani, N. Sugiyama, Y. Suwa, R. Suzuki, H. Tagoshi, F. Takahashi, K. Takahashi, K. Takahashi, R. Takahashi, R. Takahashi, T. Takahashi, H. Takahashi, T. Akiteru, T. Takano, N. Tanaka, K. Taniguchi, A. Taruya, H. Tashiro, Y. Torii, M. Toyoshima, S. Tsujikawa, Y. Tsunesada, A. Ueda, K.-I. Ueda, M. Utashima, Y. Wakabayashi, K. Yagi, H. Yamakawa, K. Yamamoto, T. Yamazaki, C.-M. Yoo, S. Yoshida, T. Yoshino and K.-X. Sun, The Japanese space gravitational wave antenna: DECIGO, *Classical and Quantum Gravity* **28**, p. 094011 (May 2011).

26. H. J. Paik, C. E. Griggs, M. Vol Moody, K. Venkateswara, H. M. Lee, A. B. Nielsen, E. Majorana and J. Harms, Low-frequency terrestrial tensor gravitational-wave detector, *Classical and Quantum Gravity* **33**, p. 075003 (April 2016).
27. J. Harms and H. J. Paik, Newtonian-noise cancellation in full-tensor gravitational-wave detectors, *Phys. Rev. D.* **92**, p. 022001 (July 2015).
28. J. Luo, L.-S. Chen, H.-Z. Duan, Y.-G. Gong, S. Hu, J. Ji, Q. Liu, J. Mei, V. Milyukov, M. Sazhin, C.-G. Shao, V. T. Toth, H.-B. Tu, Y. Wang, Y. Wang, H.-C. Yeh, M.-S. Zhan, Y. Zhang, V. Zharov and Z.-B. Zhou, TianQin: a space-borne gravitational wave detector, *Classical and Quantum Gravity* **33**, p. 035010 (February 2016).
29. P. C. Peters and J. Mathews, Gravitational Radiation from Point Masses in a Keplerian Orbit, *Physical Review* **131**, 435 (July 1963).
30. P. Amaro-Seoane and M. Freitag, Intermediate-Mass Black Holes in Colliding Clusters: Implications for Lower Frequency Gravitational-Wave Astronomy, *ApJ Lett.* **653**, L53 (December 2006).
31. P. Amaro-Seoane, M. C. Miller and M. Freitag, Gravitational Waves from Eccentric Intermediate-Mass Black Hole Binaries, *ApJ Lett.* **692**, L50 (February 2009).
32. P. Amaro-Seoane, C. Eichhorn, E. K. Porter and R. Spurzem, Binaries of massive black holes in rotating clusters: dynamics, gravitational waves, detection and the role of eccentricity, *MNRAS* **401**, 2268 (February 2010).

33. P. Amaro-Seoane and L. Santamaría, Detection of IMBHs with Ground-based Gravitational Wave Observatories: A Biography of a Binary of Black Holes, from Birth to Death, *ApJ* **722**, 1197 (October 2010).
34. B. Kocsis and J. Levin, Repeated bursts from relativistic scattering of compact objects in galactic nuclei, *Phys. Rev. D* **85**, p. 123005 (June 2012).
35. A. Sesana, Prospects for Multiband Gravitational-Wave Astronomy after GW150914, *Physical Review Letters* **116**, p. 231102 (June 2016).
36. X. Chen and P. Amaro-Seoane, Revealing the Formation of Stellar-mass Black Hole Binaries: The Need for Deci-Hertz Gravitational-wave Observatories, *ApJ Lett.* **842**, p. L2 (June 2017).
37. P. Amaro-Seoane, Detecting intermediate-mass ratio inspirals from the ground and space, *Phys.Rev.D.* **98**, p. 063018 (September 2018).

# Using "enhanced quantization" to bound the cosmological constant and quantum number $n$ for production of 100 relic mini black holes

Andrew Walcott Beckwith

*Physics Department, Chongqing University, Chongqing 401331, People's Republic of China*
*E-mail: Rwill9955b@gmail.com; abeckwith@uh.edu*
*english.cqu.edu.cn*

We compare two action integrals and identify the Lagrangian multiplier to set up a constraint equation on cosmological expansion. This is a direct result of the fourth equation of our manuscript which unconventionally compares the general-relativity action integral with the second derived action integral. This leads to Eq. (3)–(5), a bound on the cosmological constant. We replace the Hamber quantum-gravity reference-based action integral with a result from Klauder's "Enhanced Quantization," while showing its relevance to early-universe black-hole production and the volume of space producing 100 black holes of $10^2$ Planck mass in a radius of $10^3$ Planck length for entropy of about 1000.

*Keywords*: Inflaton, Action Integral, Cosmological Constant, Penrose Cyclic Cosmology.

## 1. Basic Idea, Can Two First Integrals Give Equivalent Information?

Instead of using the Hamber result[1] for a first integral, we use what John Klauder's first integral[2] to make a one–one equivalence with the first integral associated with general relativity.[3,4] We have a one–one relationship between two first action integrals.[5] The the two integrals have a one–one and onto relationship by details.[6,7]

## 2. Now for the General Relativity First Integral[5]

We use Padmanabhan's first integral[7,8] (Eq. (1)), with a Ricci scalar[9] and usually the curvature, $\aleph$, set as extremely small, with general relativity[2,10]

$$S_1 = \tfrac{1}{2\kappa}\int \sqrt{-g}\,\mathrm{d}^4x(\mathfrak{R}-2\Lambda) \quad -g = -\det g_{\mathrm{uv}} \quad \mathfrak{R} = 6\left[\tfrac{\ddot{a}}{a} + \left(\tfrac{\dot{a}}{a}\right)^2 + \tfrac{\aleph}{a^2}\right]. \qquad (1)$$

Also, the variation of $\delta g_{tt} \approx a_{\min}^2\phi$[11,12] will have an inflaton, $\phi$,[5,7,7,9,12] which is combined into other procedures for a solution to the cosmological constant problem. Here, $a_{\min}$ is a minimum value of the scale factor.[5,9]

## 3. The Idea from Klauder

Next, a restricted quantum action principle[2] (p. 78) $S_2$ with one–one equivalence,[1,3,5]

$$S_2 = \int_0^T \mathrm{d}t[p(t)\dot{q}(t) - H_N(p(t),q(t))] \approx S_1 = \frac{1}{2\kappa}\int \sqrt{-g}\,\mathrm{d}^4x(\mathfrak{R}-2\Lambda). \qquad (2)$$

We assume $\Lambda$ is a constant. Hence, we assume the following approximation,[4,6]

$$\frac{p_0^2}{2} = \frac{p_0^2(N)}{2} + N; \quad \text{for } 0 < N \leq \infty \text{ and } q = q_0 \pm p_0 t \tag{3}$$

$$V_N = \begin{cases} 0; & \text{for } 0 < x < 1 \\ N; & \text{otherwise} \end{cases} \tag{4}$$

$$H_N(p(t), q(t)) = \frac{p_0^2}{2} + \frac{(\hbar\pi)^2}{2} + N; \quad \text{for } 0 < N \leq \infty. \tag{5}$$

Our innovation is to equate $q = q_0 \pm p_0 t \sim \phi$ and assume small time-step values:

$$\Lambda \approx \frac{-\left[\frac{V_0}{3\gamma-1} + 2N + \frac{\gamma(3\gamma-1)}{8\pi G\tilde{\imath}^2}\right]}{\frac{1}{\kappa}\int\sqrt{-g}\,\mathrm{d}^3x} + \left\{6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right]\right\}\Bigg|_{t=\tilde{\imath}}. \tag{6}$$

These terms within the bubble of space–time[7] use the same inflaton potential. The scale factor is presumed to obey the value of that in Camarade et al.[9]

## 4. Why This Is Linked to Gravity/Massive Gravitons

Klauder's program[2] is to embed via Eq. (6) as a quantum-mechanical well for a Pre-Planckian system for inflaton physics,[2] (Eq. (3)) as given in Klauder's treatment of the action integral[9] (p. 87) where Klauder talks of the weak correspondence principle in which an enhanced classical Hamiltonian is given one–one correspondence with quantum effects in a nonvanishing fashion. Thus, $m_g = \frac{\hbar\sqrt{\Lambda}}{c}$.[13]

## 5. Reviewing Penrose's Multiverse CCC Generalization and Suggestions as to a Uniform Bound to the Graviton

We extend Penrose's suggestion of cyclic universes, black-hole evaporation, and the embedding structure in which our universe is contained. This multiverse embeds black holes and may resolve what appears to be an impossible dichotomy. The following[14] has serious relevance to the final part of the conclusion. There are no fewer than $N$ universes undergoing Penrose "infinite expansion"[15] contained in a mega universe structure. Furthermore, each of the $N$ universes has black-hole evaporation, with Hawking radiation from decaying black holes. If each of the $N$ universes is defined by a partition function, called $\{\Xi_i\}_{i=N}^{i=1}$, then there exist an information ensemble of mixed minimum information correlated as about $10^7$–$10^8$ bits of information per partition function in the set $\{\Xi_i\}_{i=N}^{i=1}\big|_{\text{before}}$. So, minimum information is conserved between a set of partition functions per universe.[14]

$$\{\Xi_i\}_{i=N}^{i=1}\big|_{\text{before}} \equiv \{\Xi_i\}_{i=N}^{i=1}\big|_{\text{after}}. \tag{7}$$

However, there is nonuniqueness of information put into each partition function $\{\Xi_i\}_{i=N}^{i=1}$. Furthermore, Hawking radiation from the black holes is collated via a strange attractor collection in the mega universe structure to form a new big bang for

each of the $N$ universes represented by $\{\Xi_i\}_{i=N}^{i=1}$. Verification of this mega structure compression and expansion of information with a nonuniqueness of information placed in each of the $N$ universes favors ergodic mixing treatments of initial values for each universe expanding from a singularity beginning. The $n_f$, $S_{\text{entropy}} \sim n_f$.[16] Tying this energy in, as in Eq. (11), requires ig ignoring the formation of a nontrivial gravitational measure as a new big bang for each universe: $n(E_i)$. The density of states at a given energy $E_i$ for each partition function.[14]

$$\{\Xi_i\}_{i=N}^{i=1} \propto \left\{ \int_0^\infty \mathrm{d}E_i \cdot n(E_i) \cdot \mathrm{e}^{-E_i} \right\}_{i=1}^{i=N} . \tag{8}$$

Each $E_i$ identified with Eq. (8) iterates for $N$ universes.[14,15,17] Then, in the following two claims on the universe, as a mixed state, black holes playing a major part.

**Claim 1** Where subscipts bn and an refer to before and after nucleation,

$$\frac{1}{N} \cdot \sum_{j=1}^{N} \Xi_j\big|_{j\ \mathrm{bn}} \xrightarrow[\text{transfer}]{\text{vacuum nucleation}} \Xi_i\big|_{i\ \mathrm{an}} . \tag{9}$$

For $N$ universes, with each $\Xi_j\big|_{j\ \mathrm{bn}}$ for $j = 1$ to $N$ being the partition function of each universe just before the blend into the RHS of Eq. (12) for our present universe. Also, each of the independent universes given by $\Xi_j\big|_{j\ \mathrm{bn}}$ is constructed by the absorption of one to ten million black holes taking in energy.[14,15] Furthermore, the main point is in terms of general ergodic mixing.

**Claim 2**

$$\Xi_j\big|_{j\ \mathrm{bn}} \approx \sum_{k=1}^{\max} \widetilde{\Xi}_k\Big|_{\text{black holes in the } j\text{th universe}} . \tag{10}$$

In Claims 1 and 2, we show how a multidimensional black-hole physics enables continual space–time mixing[14,15] largely to avoid the anthropic principle and its preferred set of initial conditions. Claim 2 is CCC cosmology: First, have a big bang (initial expansion) for the universe, for say redshift $z = 10$. At ten billion years ago, SMBH formation starts. Matter–energy is vacuumed up by the SMBHs. Then,

$$\Lambda = c_1 \cdot T^{\beta}. \tag{11}$$

$c_1$ is constant. Penrose's proposal (Eq. (1)) evaluates a change in the metric $g_{ab}$ by a conformal mapping $\widehat{\Omega}$ to $\widehat{g}_{ab} = \widehat{\Omega}^2 g_{ab}$. Penrose suggested[14,15]

$$\widehat{\Omega} \xrightarrow[\mathrm{CCC}]{} \widehat{\Omega}^{-1}. \tag{12}$$

Infall into cosmic black holes has been the most useful mechanism for the recycling apparent in Eq. (12) with the caveat that $\hbar$ is kept constant from cycle to cycle as represented by $\hbar_{\text{old}} = \hbar_{\text{new}}$. We claim that the invariance of the Planck $\hbar$ combined with Eq. (10) gives a good indication of a uniform mass to a graviton, per cycle,

provided that $\hbar_{\text{old}} = \hbar_{\text{new}}$ holds. Note that all these above results should be compared with the initial Hamber-based results,[1] which led to an initial idea.[5] That is, we claim we have kept full fidelity with this program and improved on it.

We have what is known as a scale factor $a(t)$, which is nearly zero in the Pre-Planckian regime of space–time and equal to 1 in the present era. In addition, we will define, for the purpose of analysis of the integrals:

$$\int \mathrm{d}t\sqrt{g_{tt}}V_3(t) = V_4(t) \sim \tfrac{8\pi^2 r^4}{3} \quad V_3(t) = \tfrac{2\pi^2 a(t)^3}{3} \quad k_2 = 9(2\pi^2)^{\frac{2}{3}} \tag{13}$$

Here, the subscripts 3 and 4 in the volume refer to three- and four-dimensional spatial dimensions, and this will lead to us writing, a first integral,[1] if $G$ is the gravitational constant, a first integral defined by[1,5]

$$S_1 = \frac{1}{24\pi G}\cdot\left[\int \mathrm{d}t\sqrt{g_{tt}}\left(\frac{g_{tt}\dot{V}_3^2}{V_3(t)} + k_2 V_3^{\frac{1}{3}}(t) - \lambda V_3(t)\right)\right]. \tag{14}$$

Compare this to the Padmabhan first integral[7] of the form Eq. (1), with a Ricci scalar[3] and usually the curvature $\aleph$[3] set as extremely small.

Our presentation uses all this, and aligns it with the ideas of the Klauder's enhanced quantization[2] for what we think is a better extension of the same idea. We have improved upon this idea, and it is in full fidelity with the FFP 15 presentation.[5]

To obtain maximum results, we assume

$$\sqrt{g_{tt}}\left(\frac{g_{tt}\dot{V}_3^2}{V_3(t)} + k_2 V_3^{\frac{1}{3}}(t)\right) \sim \frac{1}{2\kappa}\int\sqrt{-g}\cdot\,\mathrm{d}^3x\cdot\left\{6\cdot\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2\right]\right\} \tag{15}$$

and

$$\sqrt{g_{tt}}\lambda V_3(t) \sim \frac{1}{2\kappa}\int\sqrt{-g}\cdot\,\mathrm{d}^3x\cdot(2\Lambda) \to \lambda \sim \frac{1}{\kappa}\sqrt{\frac{-g}{(\delta g_{tt} \approx a_{\min}^2\phi)}}\cdot\Lambda.^{5} \tag{16}$$

So, we argue that we have related "the Lagrangian multiplier and the cosmological constant."[5] We obtain the exact same physics for when we appeal to Eq. (7) as a bound to the enhanced quantization, hence we have extended our basic idea.[2,5]

## 6. Conclusion, Relevance to Black-Hole Production

Our assumption is that the Lagrangian multiplier is roughly equivalent to a mass of about $10^4$ times the mass of a Planck-sized black hole. That is, we have black holes initially produced which are say $10^2$ times the Planck mass. In Corda's recent work,[18] we have that $n$ is the quantum number where the Planck mass is normalized to 1. So, then, if there are $10^2$ black holes of mass $10^2$ times Planck mass $M = 10^2 mass_{\text{Planck}} = 100$. Here we simplify a horizon volume.

$$\Delta V_{n-1\to n}|_{\text{Total}} \approx 10^2 \times \Delta V_{n-1\to n} \approx 10^3 \times 16\pi\cdot\frac{n}{4\times 10^4}\cdot\left|\sqrt{1-\frac{n-1}{n}}\right|. \tag{17}$$

We assume $10^2$ mini black holes, each with a mass of $10^2$ times Planck mass. So, we find $n$ for quantum number.[18] Then we have an entropy count, on r.h.s. as

$$16\pi \cdot \frac{n}{4 \times 10^4} \cdot \left| \sqrt{1 - \frac{n-1}{n}} \right| \approx 10^6. \tag{18}$$

At 1000 Planck lengths, 100 mini black holes appear, each about 100 Planck masses, restricting $n$. Use Eq. (18) to choose a quantum value for quantization level used to penetrate beyond the shell creating the cosmological constant modeled in Eq. (6). The entire mass of initial black holes, $10^4$ Planck masses, would also scale to the right-hand side of Eq. (16). We use Corda's result[18] to delineate the quantum value of relic black holes and a quantum state commensurate with penetrating from Pre-Planckian to Planckian physics. Second, if there are, say, ten gravitons produced per relic black hole and 100 relic black holes, say, in a sphere of about 1000 times Planck distance, we can by Ng infinite quantum statistics,[16] as has been done by the author time and time again (entropy as a counting algorithm) of black holes creating entropy, use this above procedure to estimate an initially generated entropy of the order of $10^3$, in the immediate aftermath of black hole production, with $n$, as calculated by Eq. (18) as the production of initial entropy. We argue that all the above will, if we equate the nucleated black-hole mass of 100 relic black holes as proportional to the right-hand side of Eq. (16), lead to an integrated version of initial mass-energy. Note that the nonsingular start to the universe is justified.[4,9,19]

**Acknowledgements**

Supported in part by China National Natural Science grant No. 11375279.

**References**

1. H. Hamber, *Quantum Gravitation: The Feynman Path Integral Approach* (Springer-Verlag, Heidelberg, 2009).
2. J. Klauder, *Enhanced Quantization, Particles, Fields and Gravity* (World Scientific, Singapore, 2015).
3. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley and Sons, New York, 1972).
4. C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity* (Cambridge University Press, Cambridge, England, 2015).
5. A. Beckwith, Creating a (quantum?) constraint, in pre Planckian space–time early universe via the Einstein cosmological constant in a one to one and onto comparison between two action integrals (2017), Text of Talk for FFP 15, Spain November 30, 11 am-11:30 am, `http://vixra.org/abs/1711.0355`.
6. A. Beckwith, How to determine a jump in energy prior to a causal barrier, with an attendant current, for an effective initial magnetic field. in the pre Planckian to Planckian space–time, *J. High Energy Phys. Gravitation Cosmol.* **4**, 323 (2017), `https://doi.org/10.4236/jhepgc.2018.42021`.

7. T. Padmanabhan, *An Invitation to Astrophysics*, World Scientific Series in Astronomy and Astrophysics, Vol. 8 (World Scientific, Singapore, 2005), https://doi.org/10.1142/9789812700988_0007.
8. T. Padmanabhan, Understanding our universe: Current status and open issues, in *100 Years of Relativity, Space–Time, Structure: Einstein and Beyond*, ed. A. Ashtekar (World Scientific, Singapore, 2005) pp. 175–204. https://doi.org/10.1142/9789812700988_0007.
9. C. S. Camara, M. R. de Garcia Maia, J. C. de Carvalho and J. A. de Lima, Nonsingular FRW cosmology and non linear dynamics, *Phys. Rev. D* **69**, p. Article 063501 (2004).
10. A. Beckwith, Using "enhanced quantization" to bound the cosmological constant, (for a bound on graviton mass), by comparing two action integrals (one being from general relativity) at the start of inflation (2018), http://vixra.org/pdf/1802.0305v1.pdf.
11. A. Beckwith, Gedanken experiment for refining the Unruh metric tensor uncertainty principle via Schwarzschild geometry and Planckian space–time with initial nonzero entropy and applying the Riemannian–Penrose inequality and initial kinetic energy for a lower bound to graviton mass (massive gravity), *J. High Energy Phys. Gravitation Cosmol.* **2**, 106 (2016), https://doi.org/10.4236/jhepgc.2016.21012.
12. M. Giovannini, *A Primer on the Physics of the Cosmic Microwave Background* (World Scientific, Hackensack, NJ, 2008), https://doi.org/10.1142/6730.
13. M. Novello, The mass of the graviton and the cosmological constant puzzle (2005), https://arxiv.org/abs/astro-ph/0504505.
14. A. Beckwith, Analyzing black hole super-radiance emission of particles/energy from a black hole as a Gedanken experiment to get bounds on the mass of a graviton, *Adv. High Energy Phys.* **2014**, p. Article 230713 (2014), https://doi.org/10.1155/2014/230713.
15. R. Penrose, *Cycles of Time—An Extraordinary New View of the Universe* (Alfred A. Knopf, New York, 2011).
16. Y. J. Ng, Spacetime foam: From entropy and holography to infinite statistics and nonlocality, *Entropy* **10**, 441 (2008), https://doi.org/10.3390/e10040441.
17. R. Penrose, Before the big bang: An outrageous new perspective and its implications for particle physics, in *Proceedings of EPAC 2006*, eds. C. Biscari, H. Owen, C. Petit-Jean-Genaz, J. Poole and J. Thomason (Joint Accelerator Conferences Website, Geneva, 2006).
18. C. Corda, Space–time can be neither discrete nor continuous, *Mod. Phys. Lett. A* **33**, p. Article 1850069 (2018), https://doi.org/10.1142/S0217732318500694.
19. C. Corda and H. J. M. Cuesta, Inflation from $R^2$ gravity: A new approach using nonlinear electrodynamics, *Astropart. Phys.* **34**, 587 (2011).

# First observational indication of the gravitomagnetic monopole

Chandrachur Chakraborty*

*Kavli Institute for Astronomy and Astrophysics,*
*Peking University, Beijing 100871, China*
**E-mail: chandra@pku.edu.cn*
*http://kiaa.pku.edu.cn/*

Sudip Bhattacharyya

*Department of Astronomy and Astrophysics,*
*Tata Institute of Fundamental Research, Mumbai 400005, India*

The first observational indication of the gravitomagnetic monopole is reported, based on the X-ray observations of an astrophysical collapsed object: GRO J1655-40. Earlier, the three independent primary X-ray observational methods gave significantly different spin values for GRO J1655-40. We show that the inclusion of an extra parameter, corresponding to the gravitomagnetic monopole, makes the spin and other parameter values inferred from the three methods consistent with each other. In addition, our result weakly indicates that GRO J1655-40 could also be a naked singularity.

*Keywords*: Gravitomagnetic monopole, QPOs, Singularity, X-rays: binaries.

## 1. Introduction

In 1931, British physicist P. A. M. Dirac showed that if the magnetic monopoles exist in the universe, then all electric charge in the universe must be quantized[1,2]. When the analogy of Einstein's general relativity is drawn with electromagnetism, 'gravitoelectric charge' exists and is referred to as 'mass' which is an important constituent of the physical world. The next evident question that arises is 'does gravitomagnetic charge or the so-called gravitomagnetic monopole exist in nature?'[3] Historically, Newman, Tamburino and Unti discovered a stationary and spherically symmetric vacuum solution (which is now known as the NUT solution) of the Einstein's equation, that contains the gravitomagnetic monopole or the so-called NUT parameter[4]. Though this gravitomagnetic monopole is physically interpreted as "a linear source of pure angular momentum", i.e., "a massless rotating rod"[5–8], and is the proposed gravitational analogue of Dirac's magnetic monopole, an observational evidence of this aspect of fundamental physics was elusive[3].

Perhaps Lynden-Bell and Nouri-Zonoz[9] were the first to motivate such an investigation. They argued that the best place to look for the gravitomagnetic monopole is in the spectra of supernovae, quasars, or active galactic nuclei. We demonstrate that the X-ray observations of a black hole X-ray binary (BHXB), i.e., an accreting stellar-mass collapsed object, can provide a way to detect a non-zero NUT parameter or the gravitomagnetic monopole. A rotating astrophysical collapsed object is generally described by the Kerr spacetime[10], and is characterized by only two parameters : mass and spin. Depending on the spin value of the rotating astrophysical collapsed object, it could be a black hole, or a naked singularity (also sometimes

called a "superspinar") which is visible in principle to external observers. Now, if the Kerr spacetime contains the gravitomagnetic monopole or vice-versa, it is regarded as the Kerr-Taub-NUT (KTN) spacetime[11]. This is geometrically a stationary, axisymmetric vacuum solution of Einstein equation, and reduces to the Kerr spacetime if the NUT parameter is zero. Thus, if one can identify a collapsed object with the KTN spacetime, this could establish the existence of the gravitomagnetic monopole. Here, we demonstrate[3] that the X-ray observations of a black hole X-ray binary (BHXB), i.e., an accreting stellar-mass collapsed object, can provide a way to infer and measure the gravitomagnetic monopole. Although, the collapsed object is generally thought to be a black hole, i.e., a singularity covered by an event horizon due to the Cosmic Censorship Conjecture, we do not exclude the possibility that it could also be a naked singularity in some cases[12]. However, before going into detail, we first briefly recapitulate the KTN spacetime and discuss some important observables which are useful to explore the possibility of gravitomagnetic monopole in nature.

## 2. Kerr-Taub-NUT Spacetime and Fundamental Frequencies

The metric of the KTN spacetime can be expressed as[11]

$$ds^2 = -\frac{\Delta}{p^2}(dt - Ad\phi)^2 + \frac{p^2}{\Delta}dr^2 + p^2 d\theta^2 + \frac{1}{p^2}\sin^2\theta(adt - Bd\phi)^2 \tag{1}$$

with

$$\begin{aligned} &\Delta = r^2 - 2Mr + a^2 - n^2, \quad p^2 = r^2 + (n + a\cos\theta)^2, \\ &A = a\sin^2\theta - 2n\cos\theta, \quad B = r^2 + a^2 + n^2, \end{aligned} \tag{2}$$

where $M$ is the mass, $a/M$ is the Kerr parameter and $n/M$ is the NUT parameter.

Now, using the metric components ($g_{\mu\nu}$) of KTN spacetime one can obtain the three fundamental frequencies. The orbital frequency can be written as[13]

$$\Omega_\phi^{\rm KTN} = 2\pi\nu_\phi^{\rm KTN} = \pm\frac{m^{\frac{1}{2}}}{r^{\frac{1}{2}}\ (r^2 + n^2) \pm a\ m^{\frac{1}{2}}}, \tag{3}$$

where $m = M\ (r^2 - n^2) + 2\ n^2 r$. Here, the upper sign is applicable for the prograde orbits and the lower sign is applicable for the retrograde orbits. Similarly, the radial and vertical epicyclic frequencies are written as[3]:

$$\begin{aligned} \nu_r^{\rm KTN} = \frac{\nu_\phi^{\rm KTN}}{m^{\frac{1}{2}}\ (r^2 + n^2)}.\Big[&M(r^6 - n^6 + 15n^4r^2 - 15n^2r^4) - 2M^2r(3r^4 - 2n^2r^2 + 3n^4) \\ &-16n^4r^3 \pm 8ar^{\frac{3}{2}}m^{\frac{3}{2}} + a^2\left\{M(n^4 + 6n^2r^2 - 3r^4) - 8n^2r^3\right\}\Big]^{\frac{1}{2}} \end{aligned} \tag{4}$$

and

$$\nu_\theta^{\rm KTN} = \frac{\nu_\phi^{\rm KTN}}{m^{\frac{1}{2}}\ (r^2+n^2)}.\ \big[M(r^6-n^6+15n^4r^2-15n^2r^4)+2n^2r(3r^4-2n^2r^2+3n^4)$$
$$+16M^2n^2r^3 \mp 4ar^{\frac{1}{2}}m^{\frac{1}{2}}(n^2+Mr)(n^2+r^2)$$
$$-a^2\left\{M(n^4+6n^2r^2-3r^4)-8n^2r^3\right\}\big]^{\frac{1}{2}} \qquad (5)$$

respectively. Setting the square of Eq. (4) equal to zero (i.e., $[\nu_r^{\rm KTN}]^2 = 0$), we obtain the innermost stable circular orbit (ISCO) condition as follows[13,14]:

$$M(r^6-n^6+15n^4r^2-15n^2r^4)-2M^2r(3r^4-2n^2r^2+3n^4)$$
$$-16n^4r^3 \pm 8ar^{\frac{3}{2}}m^{\frac{3}{2}}+a^2\left\{M(n^4+6n^2r^2-3r^4)-8n^2r^3\right\}=0. \qquad (6)$$

The gravitational redshift in the KTN spacetime is expressed as[3]

$$Z^{\rm KTN} = \frac{r^{\frac{1}{2}}(r^2+n^2)+a\ m^{\frac{1}{2}}}{\left[(r^2+n^2)\ \left\{r(r^2-3n^2)+M(n^2-3r^2)+2a(mr)^{\frac{1}{2}}\right\}\right]^{\frac{1}{2}}}. \qquad (7)$$

## 3. Exploring the possibility of non-zero gravitomagnetic monopole

$a/M$ measurement can be very useful to probe the strong gravity regime and to characterize the collapsed object. Some of the X-ray spectral and timing features which originate from the accreted matter within a few gravitational radii of a collapsed object in a BHXB, is generally used to measure $a/M$[15–17]. There are two main spectral methods for $a/M$ estimation: (1) using continuum X-ray spectrum[16], and (2) using broad relativistic iron K$\alpha$ spectral emission line[15]. There is also a timing method based on the relativistic precession model (RPM) of quasi-periodic oscillations (QPOs) of X-ray intensity[18].

In continuum X-ray spectrum method, thermal spectral component from the accretion disk is fit with a relativistic thin-disk model, and this gives a measure of the $r_{\rm in}$[19]. Then from a known $M$ and $r_{\rm ISCO}/M$ values, $a/M$ is inferred assuming Kerr spacetime. In the second method, a broad relativistic iron K$\alpha$ spectral emission line in X-rays is observed from many BHXBs, which is believed to originate from the reflection of hard X-rays from the inner part of the geometrically thin accretion disk. The red wing of the line that determines the observed constraint on $a/M$[20], gives a measure of the gravitational redshift at the disk inner edge radius $r_{\rm in}$, and for $r_{\rm in} = r_{\rm ISCO}$ ($r_{\rm ISCO}$ is the innermost stable circular orbit (ISCO) radius), the $a/M$ value is inferred for a prograde accretion disk assuming the Kerr spacetime.

In the QPO-based timing method, three observed QPO frequencies are used to estimate $a/M$: (a) the upper high-frequency (HF) QPO, (b) the lower HFQPO, and (c) the type-C low-frequency (LF) QPO[18]. This method is based on the relativistic precession model (RPM)[21,22], in which the Type-C QPO frequency is identified with the LT precession frequency ($\nu_{\rm LT}$)[23–25], and the upper and lower HFQPO frequencies are identified with the orbital frequency ($\nu_\phi$) and the periastron precession

frequency ($\nu_{\rm per}$) respectively[18]. In Kerr spacetime, each of these three frequencies is a function of three parameters: $M$, $a/M$ and $r_{\rm qpo}$ ($r_{\rm qpo}$ is the location of origin of these QPOs). Therefore, one can obtain the numerical values of these three parameters[18] for a particular Kerr compact object using the RPM method.

The RPM method could be fully applied for GRO J1655–40, as the three above-mentioned QPOs could be simultaneously observed only from this BHXB[18]. The mass of GRO J1655–40 is either $(6.3 \pm 0.5)M_{\odot}$[26] or $(5.4 \pm 0.3)M_{\odot}$[27]. It is not known until now, which mass value is reliable[19]. The observed frequencies of the above mentioned three simultaneous QPOs imply $\nu_{\phi} = 440$ Hz, $\nu_{\rm per} = 300$ Hz and $\nu_{\rm LT} = 17$ Hz for GRO J1655–40. Using these frequencies, Motta et. al.[18] determined $a/M \approx 0.286 \pm 0.003$, $M \approx (5.31 \pm 0.07)M_{\odot}$ and $r_{\rm qpo} \approx (5.68 \pm 0.04)\ M$. Moreover, GRO J1655–40 is currently the only BHXB, for which all the three above mentioned $a/M$ estimation methods are available. Thus, this source provides a unique opportunity to test the reliability of these methods by comparing the three estimated $a/M$ values. The timing method gives $a/M \approx 0.286 \pm 0.003$[18], the line spectrum method gives $a/M \approx 0.90 - 0.99$[20], and the continuum spectrum method gives $a/M \approx 0.65 - 0.75$ for GRO J1655–40. Therefore, the $a/M$ value inferred from these three methods are grossly inconsistent with each other.

Are these three methods missing an essential ingredient, is it possible to make these methods reliable? We explore an possibility that the inclusion of NUT charge may make the results from three methods consistent, and it suggests that gravitomagnetic monopole exists in nature[3]. We have already calculated the important observables with non-zero NUT parameter values, i.e., by considering the KTN spacetime, instead of the previously used Kerr spacetime. Now, we can use those results directly to show the possibility of existence of non-zero gravitomagnetic monopole in GRO J1655–40.

As we have already discussed, we can apply the RPM method to GRO J1655–40 using KTN spacetime, instead of Kerr spacetime which was used in[18]:

$$\nu_{\phi}^{\rm KTN} = 440 \text{ Hz};\ \nu_{\rm per}^{\rm KTN} = 300 \text{ Hz};\quad \nu_{\rm LT}^{\rm KTN} = 17 \text{ Hz}. \tag{8}$$

Using Eqs. (3–5) we can solve Eqs.(8) for $a/M$, $M$ and the radius $r_{\rm qpo}$ of QPO origin for a given $n/M$ value[3]. For $n = 0$, one can easily recover the above-mentioned parameter values as described in[18]. As the three equations are also satisfied for other non-zero values of $n$, a new range of $n$ could be found[3], and this can give the same QPO values for GRO J1655–40 with a specific combinations of $M$, $a/M$ and $r_{qpo}$. One should note here that the range of $0 < a/M \leq 1$ represents a Kerr black hole, whereas $a/M > 1$ represents Kerr naked singularity[12,28]. Similarly, the KTN naked singularity is represented with the following range $a/M > \sqrt{1 + (n/M)^2}$[29] and KTN BH is represented with $0 < a/M \leq \sqrt{1 + (n/M)^2}$.

Using the line spectrum method[20], the $a/M$ range for GRO J1655–40 was estimated as $\approx 0.90 - 0.99$, assuming it as a Kerr black hole. Now, one can calculate the gravitational redshift range ($\approx 2.70 - 6.08$) from the reported $a/M$ range ($\approx 0.90 - 0.99$), and using this range ($\approx 2.70 - 6.08$) as the primary

observational constraint, and assuming GRO J1655–40 as the KTN spacetime, i.e., $Z^{\rm KTN}(\text{at } r_{\rm ISCO}) = 2.70 - 6.08$ (L.H.S. is given by Eqs. 6 and 7), one can solve for $M$, $a/M$, $n/M$, $r_{\rm ISCO}/M$ and $r_{\rm qpo}$ from Eq. (8). This solution gives the following constraints for GRO J1655–40, which are consistent with both the RPM method and the line spectrum method: $M \approx 6.76 - 6.83 M_{\odot}$, $a/M \approx 2.12 - 2.27$, $n/M \approx 1.86 - 1.97$ and $r_{\rm qpo}/M \approx 4.99 - 5.04$[3]. The $M$ range is consistent with an independently measured mass ($[6.3 \pm 0.5] M_{\odot}$;[26]) for GRO J1655–40, which provides a confirmation of the reliability of our method and results.

Using the continuum spectrum method, one can find the $a/M$ range as $\approx 0.65 - 0.75$[30] for GRO J1655–40. This was inferred from the $r_{\rm ISCO}$ range as $\approx 29.8 - 34.2$ km, using $M = 6.3 M_{\odot}$. Now, considering $r_{\rm ISCO} = 29.8 - 34.2$ km as the primary constraint and assuming the KTN spacetime (using Eq. 6), we solve Eq. (8) for $M$, $a/M$, $n/M$, $r_{\rm ISCO}/M$ and $r_{\rm qpo}/M$. Consequently, the following parameter constraints could be obtained: $M \approx 6.79 - 6.86 M_{\odot}$, $a/M \approx 2.04 - 2.21$, $n/M \approx 1.79 - 1.93$ and $r_{\rm qpo}/M \approx 4.96 - 5.02$[3]. It can be seen that these parameter ranges overlap with the parameter ranges obtained from the combined RPM and line spectrum methods.

## 4. Conclusion and Discussion

We have shown that $a/M \approx 2.12 - 2.21$ and $n/M \approx 1.86 - 1.93$ are consistent with all the three methods[3]. These ranges imply that the collapsed object in GRO J1655–40 could also be a naked singularity (see Fig. 1 of[3]). One can also realize from the above mentioned $n/M$ range ($\approx 1.86 - 1.93$) that the inferred non-zero $n/M$ values cannot be caused by the systematics, as is significantly away from $n/M = 0$. The earlier three independent primary X-ray observational methods provided significantly different spin values for GRO J1655-40, which were $0.286 \pm 0.003$, $0.65 - 0.75$ and $0.90 - 0.99$. Employing an interesting technique, we have demonstrated that the inclusion of one extra parameter (i.e., gravitomagnetic monopole) not only makes the spin and other parameter values inferred from the three methods consistent with each other, but also makes the inferred black hole mass consistent with an independently measured value[3]. This confirmation provides a new way to measure the NUT parameter, even when only two $a/M$ measurement methods are available for a BHXB. Here, one should note that the value of $n/M$ can be different for different objects like the value of $a/M$ and a high $n/M$ value inferred for one object does not mean that every object will have a high $n/M$ value. The value of $n/M$ can even be very close to zero for some objects[3]. However, the inferred significantly non-zero $n/M$ value for even one object, i.e., GRO J1655-40 could strongly suggest the existence of gravitomagnetic monopole in nature[3]. Finally, our inference of the first significant observational indication of the gravitomagnetic monopole, which, even though is not a direct detection, can have an exciting impact on fundamental physics and astrophysics. Moreover, as the accreting collapsed object GRO J1655-40 could be better described with the more general KTN spacetime, instead of the Kerr spacetime, this makes the KTN spacetime astrophysically relevant.

**Acknowledgments**

We thank S. Ramaswamy and P. Majumdar for useful discussions on this topic. C. C. gratefully acknowledges support from the National Natural Science Foundation of China (NSFC), Grant No. : 11750110410.

**References**

1. P. A. M. Dirac, *Proc. R. Soc. Lond. A* **133**, 60 (1931).
2. M. N. Saha, *Indian Journal of Physics* **X**, 141 (1936).
3. C. Chakraborty, S. Bhattacharyya, *Phys. Rev.* **D 98**, 043021 (2018).
4. E. Newman, L. Tamburino, T. Unti, *J. Math. Phys.* **4**, 915 (1963).
5. W. B. Bonnor, *Proc. Camb. Phil. Soc.* **66**, 145 (1969).
6. J. S. Dowker, *Gen. Rel. Grav.* **5**, 603 (1974).
7. S. Ramaswamy, A. Sen, *Phys. Rev. Lett.* **57**, 1088 (1986).
8. S. Ramaswamy, A. Sen, *J. Math. Phys. (N.Y.)* **22**, 2612 (1981).
9. D. Lynden-Bell, M. Nouri-Zonoz, *Rev. Mod. Phys.* **70**, 427 (1998).
10. R. Kerr, *Phys. Rev. Lett.* **11**, 237 (1963).
11. J. G. Miller, *J. Math. Phys.* **14**, 486 (1973).
12. C. Chakraborty et. al., *Phys. Rev.* **D 95**, 084024 (2017).
13. C. Chakraborty, Eur. Phys. J. C **74**, 2759 (2014).
14. C. Chakraborty, S. Bhattacharyya, arXiv:1901.04233 [astro-ph.HE](2019).
15. J. M. Miller, *ARAA* **45**, 441 (2007).
16. R. A., Remillard, J. E. McClintock, *ARAA* **44**, 49 (2006).
17. T. M. Belloni, L. Stella, *Space Science Reviews* **183**, 43 (2014).
18. S. E. Motta et. al., *MNRAS* **437**, 2554 (2014).
19. T. Fragos, J. E. McClintock, *Astrophys J.* **800**, 17 (2015).
20. R. C. Reis et. al., *MNRAS* **395**, 1257 (2009).
21. L. Stella, M. Vietri, *Astrophys J.* **492**, L59 (1998).
22. L. Stella, M. Vietri, *Phys. Rev. Lett.* **82**, 17 (1999).
23. A. Ingram et. al., *MNRAS* **461**, 1967 (2016).
24. C. Chakraborty, S. Bhattacharyya, *MNRAS* **469**, 3062 (2017).
25. S. Banerjee, C. Chakraborty, S. Bhattacharyya, *Astrophys J.* **870**, 95 (2019).
26. J. Greene, C. D. Bailyn, J. A. Orosz, *Astrophys. J.* **554**, 1290 (2001).
27. M. E. Beer, P. Podsiadlowski, *MNRAS* **331**, 351 (2002).
28. C. Chakraborty, P. Kocherlakota, P. S. Joshi, *Phys. Rev.* **D 95**, 044006 (2017).
29. S-W. Wei et. al., *JCAP* 10 (2012) 053.
30. R. Shafee et. al., *Astrophys. J.* **636**, L113 (2006).

# Kerr–Newman black hole in the formalism of isolated horizons

Aleš Flandera* and Martin Scholtz

*Institute of Theoretical Physics, Faculty of Mathematics and Physics, Charles University, Praha, 180 00, Czech Republic*
**E-mail: flandera.ales@utf.mff.cuni.cz*

Norman Gürlebeck

*ZARM, University of Bremen, Bremen, 28359, Germany*

We study the Kerr–Newman black hole in the formalism of isolated horizons using a general solution found by Krishnan in 2012.[1] It establishes the existence of a null tetrad which is tangent to the horizon and parallelly propagated off the horizon along a non-twisting null geodesic congruence. However, the explicit construction of such a tetrad for the Kerr–Newman metric was not given. We have formulated appropriate initial data and found the exact form (up to integrals) of the solution everywhere in the Kerr–Newman space-time.[2] In the near future we aim to apply this formulation to the study of distorted black holes by perturbing the initial data found in this work.

*Keywords*: Kerr–Newman black hole; isolated horizons; tetrad formalism.

## 1. Introduction

In 2012, Krishnan investigated the near horizon geometry in the vicinity of an isolated horizon in the Newman–Penrose formalism.[1] He found a solution of the Einstein field equations in the form of a power series of the Newman–Penrose tetrad. It was possible since the properties of weakly isolated horizons allow one to solve the constraint part equations in the Newman–Penrose formalism. The solution tells us a lot about how the Newman–Penrose tetrad shall look like. On the other hand, the approach does not provide an insight into the choice of the initial data for any particular space-time and, hence, it stayed unknown, including the Kerr–Newman space-time.

This particular space-time (in the formalism of isolated horizons) was interesting for us for the previous treatment of the Meissner effect in Ref. 3. They proved that the magnetic field is expelled from a general extremal axisymmetric stationary black hole. To be able to do so, one, however, needs to know the initial data. They found a first order approximation for the particular case. This solution represented a deformed Kerr solution, however, it was not clear, how it was deformed. The aim of our work was to find an approach which would lead to a general solution to this problem.

## 2. Weakly isolated horizons

A **non-expanding horizon** is a null hypersurface $\mathcal{H}$ with topology $\mathbb{R} \times \mathcal{S}^2$, where the Einstein equations hold and for any null normal of the horizon $l^a$ its expansion

is zero and if the normal is future pointing, also the vector $T_{ab}l^b$ is future pointing. The tensor $T_{ab}$ is the energy-momentum tensor.

The introduced expansion is the well-known optical scalar for null geodesic congruences[4] and its zeroness implies also vanishing shear and twist. Among other relations we can derive from the properties of non-expanding horizons there is one to be discussed here:

$$\mu = \bar{\mu}\,. \tag{1}$$

This tells us that $\mu$ is real and it shall be a key feature for us. For the discussion of all relevant relations, see Ref. 5.

A **weakly isolated horizon** is a pair $(\mathcal{H},\, [l^a])$ where $\mathcal{H}$ is a non-expanding horizon and $[l^a]$ is an equivalence class of normals satisfying

$$[\pounds_l, \mathcal{D}_a]l^b \doteq 0 \tag{2}$$

for any $l^a$ in this class. Here $\pounds$ is the Lie derivative and $\mathcal{D}_a$ is an induced covariant derivative on the horizon introduced by $X^a\,\mathcal{D}_aY^b \doteq X^a\,\nabla_aY^b$ for any tangent vectors $X^a$, $Y^a$. Eq. (2) ensures that the $l^a$-part of the induced connection is time independent.

## 3. Krishnan construction

In Ref. 1, it was described how a Newman–Penrose tetrad is constructed to reflect (and describe) the properties of isolated horizons and the initial value problem was formulated, though without stating the initial data. Here let us just summarize the key principles.

An arbitrary sphere $\mathcal{S}_0$ is chosen on the horizon and the vectors $m^a$ and $\bar{m}^a$ are chosen as the basis of its tangent space. They are Lie propagated along the null normal $l^a$ on the horizon. It is possible thanks to the following: the normal is null and therefore also tangent to the horizon, the normal is surface forming for it has vanishing twist,[6] the Lie transport preserves the scalar products in the case of a non-expanding horizon. In this manner, the triad is introduced on the entire horizon and thanks to the scalar products of the Newman–Penrose tetrad ($l^an_a = 1$, $m^a\bar{m}_a = -1$, others zero) the last vector $n^a$ is already fixed on the horizon. We are left with the propagation of the vectors off the horizon which is achieved by parallel transport of the triad ($l^a$, $m^a$, $\bar{m}^a$) alongside a geodesic in the direction of vector $n^a$ which is constructed first. The coordinates are propagated alike. The situation is schematically illustrated in Fig. 1.

The Bondi-like coordinates are introduced in the following way: $x^I$ are spherical coordinates on the sphere $\mathcal{S}^2$, $v$ is a time coordinate in the direction of $l^a$ and $r$ is a radial coordinate in the transversal direction of $n^a$.[a]

[a] The Newman–Penrose tetrad is a non-holonomic one and therefore there are no coordinates which could be connected to it. However, since some of the coordinates are connected to only some of

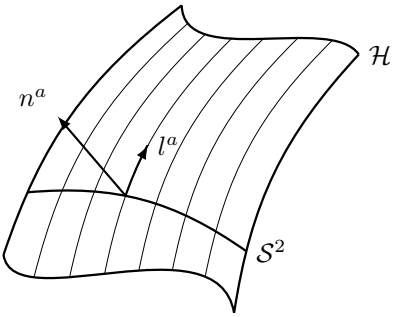



Fig. 1. Local topology of the space-time.

We have already discussed the vanishing optical scalars for the geodesic congruence in the direction of $l^a$. Let us now recall that $\mu$ spin coefficient has vanishing imaginary part, Eq. (1). Since the imaginary part of spin coefficient $\mu$ is the twist of a geodesic congruence in the direction of $n^a$, let us call such a tetrad *non-twisting* (in the direction of $n^a$). If there is an imaginary part of $\mu$, the tetrad is to be called *twisting*.

This tetrad is expanded in the radial parameter $r$

$$X = X^{(0)} + rX^{(1)} + \tfrac{1}{2}r^2X^{(2)} + \cdots \tag{3}$$

and the constraints are solved.[1] To do so, one employs the Ricci and Bianchi identities which, in the Newman–Penrose formalism, play the role of the field equations. In this manner Krishnan formulated the initial value problem which is displayed in Fig. 2. There are two characteristics: the horizon $\mathcal{H}$ and a transversal hypersurface $\mathcal{N}_0$ intersecting each other in the sphere $\mathcal{S}_0$ (both $l^a$ and $n^a$ are surface forming). On the sphere $\mathcal{S}_0$ are introduced the quantities: spin coefficients $\alpha$, $\pi$, $\mu$, $\lambda$, Maxwell scalar $\phi_1$ and the surface gravity $\kappa_{(l)}$. After solving all the constraints on $\mathcal{S}_0$ and obtaining so the Weyl scalars $\Psi_2$ and $\Psi_3$, the full geometry of the horizon can be reconstructed from the data on $\mathcal{S}_0$. Finally, in order to propagate the solution off the horizon, scalars $\phi_2$ and $\Psi_4$ must be prescribed on the entire $\mathcal{N}_0$.

## 4. Perturbative solution

Now we are ready to look for the appropriate tetrad connecting the Kerr–Newman space-time to the formalism of isolated horizons. We started with the Kinnersley tetrad[7] which is a Newman–Penrose tetrad, however, it is twisting. On the other hand, there must exist a sequence of Lorentz transformations of the tetrad which eliminates the imaginary part of the spin coefficient $\mu$.

the vectors, there is a way how to make a correspondence between the vectors of the tetrad and the coordinates. E.g. a gradient of vector $n^a$ is given by all the coordinates, but the $r$ coordinate is only present in this gradient.

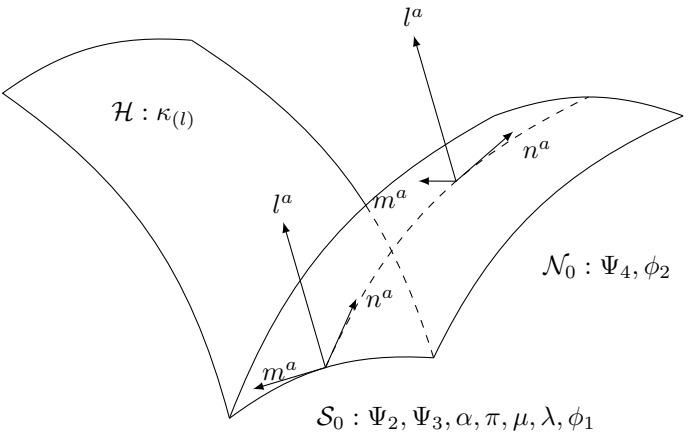



Fig. 2. The neighbourhood of the horizon and initial value surfaces. The scalars enlisted after the colons have to be introduced on the corresponding (hyper)surface.

In Ref. 5, we used a coordinate transformation, boost and rotation about $l^a$ (all valid on the horizon) and we ended up with a tetrad on the horizon:

$$\begin{aligned} l_{\mathrm{R}} &\doteq \partial_v \,, \\ n_{\mathrm{R}} &\doteq -\frac{a^2 \sin^2\theta}{2\left(r_+^2 + a^2\cos^2\theta\right)}\,\partial_v - \frac{a^2 + r_+^2}{r_+ + a^2\cos\theta}\,\partial_r - \frac{a}{a^2 + r_+^2}\,\partial_\phi \,, \\ m_{\mathrm{R}} &\doteq \frac{1}{\sqrt{2}\,(r_+ + \mathrm{i}a\cos\theta)}\,\partial_\theta + \frac{\mathrm{i}\,(r_+ - \mathrm{i}a\cos\theta)}{\sqrt{2}\left(a^2 + r_+^2\right)\sin\theta}\,\partial_\phi \,. \end{aligned} \tag{4}$$

The tetrad has all required properties on the horizon and to get it also off the horizon we used an expansion in the radial parameter $r$:

$$l^a = l^a_{(0)} + (r - r_+) l^a_{(1)} + \mathcal{O}(r - r_+)^2 \,, \qquad l^a_{(1)} = l^a_{(1)}(\theta, \phi) \,, \tag{5}$$

where the time-independence was inferred from the existence of the time-like Killing vector in Kerr–Newman space-time.

The subsequent orders of the tetrad were computed, using the parallel transport conditions, in Wolfram Mathematica. By this we get all quantities exactly on the horizon and up to an arbitrary order elsewhere.

## 5. Exact solution

To get the exact solution, it is needed to use all four types of Lorentz transformations: spin, boost, rotation about $l^a$ and rotation about $n^a$. Moreover, unlike for the expansion, they have to transform the tetrad properly in the entire space-time. We were unable to find the parameters directly, since the final step requires the solution of a difficult differential equation.

Instead, it is possible to write the vector $n^a$ in the form of a gradient:

$$n_\mu \mathrm{d}x^\mu \equiv E\,\mathrm{d}v - \frac{a^2}{P+\sqrt{R}}\,\mathrm{d}r + \sqrt{\Theta}\,\mathrm{d}\theta + L\,\mathrm{d}\varphi = \mathrm{d}\tilde{v} \tag{6}$$

where $P = aL + E(a^2 + r^2)$, $R = P^2 - \mathcal{K}\tilde{\Delta}$, $\Theta = \mathcal{K} - (L + a\,E)^2 + \left(a^2\,E^2 - L^2 \sin^{-2}\theta\right)\cos^2\theta$ and $\tilde{\Delta} = a^2 + r^2 - 2\,M\,r + Q^2$. Inspection of the integrability conditions for Eq. (6) implies the values $L = 0$ and $E = 1$,[8] then we also arrive at $K = a^2$.

With this simplification in hand, we can apply the Lorentz transformations. For the discussion see,[2] here we only briefly recapitulate them. We start with a boost followed by a rotation about $l^a$ to set the vector $n^a$ to the direction given by Eq. (6). The transformations are:

$$l^a \mapsto A^2 l^a\,, \qquad n^a \mapsto A^{-2} n^a\,, \qquad m^a \mapsto m^a\,, \tag{7}$$

$$l^a \mapsto l^a\,, \qquad n^a \mapsto n^a + c m^a + \bar{c}\bar{m}^a + |c|^2 l^a\,, \qquad m^a \mapsto m^a + \bar{c} l^a \tag{8}$$

and their parameters are (let us introduce also function $\rho = r + \mathrm{i}a\cos\theta$):

$$A^2 = \frac{2P}{P+\sqrt{R}}\,, \qquad c = -\frac{\mathrm{e}^{-\mathrm{i}\vartheta}}{\sqrt{2}\bar{\rho}}\,. \tag{9}$$

We proceed by the spin:

$$m^a \mapsto \mathrm{e}^{-\mathrm{i}\vartheta} m^a\,, \qquad \bar{m}^a \mapsto \mathrm{e}^{\mathrm{i}\vartheta}\bar{m}^a\,. \tag{10}$$

The remaining transformation is the rotation about $n^a$ and in order to find the corresponding parameter we have to solve the equation $n^a \nabla_a d = \tau$, where $\tau$ is a spin coefficient. This is however very difficult in the Kerr coordinates. In order to solve the problem, we perform a transformation to Bondi-like coordinates introduced by

$$\mathsf{v} = v - \int_{r_+}^{r} \frac{a^2\,\mathrm{d}r}{P+\sqrt{R}} + a\sin\theta\,, \qquad \sin\theta = \tanh\mathsf{X}\,, \qquad \tilde{\phi} = \varphi + J(r)\,, \tag{11}$$

where

$$\begin{aligned} \mathsf{X} &= \alpha(r) + \operatorname{arth}\sin\vartheta\,, \qquad \alpha(r) = \int_{r_+}^{r} \frac{a\,\mathrm{d}u}{\sqrt{u^4 + a^2\,u^2 + 2\,a^2\,M\,u - a^2\,Q^2}}\,, \\ J(r) &= -\int_{r_+}^{r} \frac{a}{P(u) + \sqrt{R(u)}}\left(1 + \frac{u^2}{\sqrt{R(u)}}\right)\mathrm{d}u\,. \end{aligned} \tag{12}$$

Then the equation for parameter $d$ is $-\partial_r d = \tau$ and can be integrated to get:

$$d(r,\vartheta) = -\int_{r_+}^{r} \frac{|\rho(u,\vartheta)|^2}{\sqrt{R(u)}}\,\tau(u,\vartheta)\,\mathrm{d}u\,. \tag{13}$$

Unfortunately, this cannot be evaluated explicitly and, hence, the final tetrad shall also be an implicit expression. To complete the construction of the Bondi-like coordinates, we define

$$\mathsf{r} = \int_{r_+}^{r} \frac{|\rho(u,\vartheta)|^2}{\sqrt{R(u)}}\,\mathrm{d}u\,, \qquad \phi = \tilde{\phi} - \frac{a\mathsf{v}}{a^2 + r_+^2}\,. \tag{14}$$

We arrive at the final tetrad, which is of the Krishnan type and has the form $(I = 2, 3)$:

$$l = \partial_v + U\,\partial_{\mathsf{r}} + X^I\,\partial_I\,, \qquad n = -\partial_{\mathsf{r}}\,, \qquad m = \Omega\,\partial_{\mathsf{r}} + \xi^I\,\partial_I\,. \tag{15}$$

The functions $U$, $X^I$, $\Omega$ and $\xi^I$ shall not be shown here for their length, see Ref. 2.

## 6. Conclusion

We constructed a parallelly propagated, Bondi-like tetrad in the Kerr–Newman space-time and, hence, we were able to formulate the initial data not only on the horizon, as in the previous works, but also on the transversal null hypersurface. Although the tetrad is not calculated explicitly, we present the results in the form suitable for both symbolic manipulations (e.g. in computer algebra systems) and numeric calculations, since all functions are expressed in terms of the integrals of one variable. These results were published in the paper.[2] In the near future we will apply this formalism to the study of deformed black holes. This is a promising approach since by varying data on the hypersurface $\mathcal{N}_0$ (see Fig. 2) one can generate a wide class of exact solutions describing black holes.

## Acknowledgments

Aleš Flandera acknowledges the support from grant GAUK 588217 of Charles University in Prague. Work of Martin Scholtz was financially supported by the grant GAČR 17-16260Y of the Czech Science Foundation. Norman Gürlebeck and MS gratefully acknowledge support from the DFG within the Research Training Group 1620 Models of Gravity. Partial support of NG comes also from NewCompStar, COST Action MP1304.

## References

1. B. Krishnan, The spacetime in the neighborhood of a general isolated black hole, *Classical and Quantum Gravity* **29**, p. 205006 (Oct 2012).
2. M. Scholtz, A. Flandera and N. Gürlebeck, Kerr-newman black hole in the formalism of isolated horizons, *Phys. Rev. D* **96**, p. 064024 (Sep 2017).
3. N. Gürlebeck and M. Scholtz, Meissner effect for weakly isolated horizons, *Phys. Rev. D* **95**, p. 064010 (Mar 2017).
4. E. Poisson, *A Relativist's Toolkit: The Mathematics of Black-Hole Mechanics* (Cambridge University Press, 2004).
5. A. Flandera, Geometry of isolated horizons, *ArXiv e-prints* (Nov 2016).
6. R. M. Wald, *General Relativity* (University of Chicago Press, 1984).
7. W. Kinnersley, Type d vacuum metrics, *J. Math. Phys.* **10**, 1195 (1969).
8. S. J. Fletcher and A. W. C. Lun, The Kerr spacetime in generalized Bondi–Sachs coordinates, *Class. Quantum Grav.* **20**, 4153 (2003).

# Testing the Kerr spacetime with X-ray reflection spectroscopy

Masoumeh Ghasemi-Nodehi

*National Astronomical Observatories, Chinese Academy of Sciences, Beijing 100101, China*
*E-mail: ghasemi@nao.cas.cn*

Astrophysical black holes are thought to be Kerr solution of general relativity but their Kerr geometry have not yet verified. Iron line is prominent feature in X-ray reflection spectra. Shape of iron line in X-ray reflection spectra of both stellar-mass and super massive black hole candidates are supposed to be strongly affected by spacetime geometry. This method is a powerful technique to prob strong gravity regime and deviations from Kerr spacetime. In this talk, I present iron line of some non-Kerr spacetimes. I also present data simulations of X-ray missions. The purpose is to understand whether X-ray reflection spectroscopy can distinguish these non-Kerr spacetimes from Kerr solution of general relativity.

*Keywords*: Tests of General Relativity, X-ray Reflection Spectroscopy.

## 1. Introduction

Final product of gravitational collapse is Kerr spacetime according to Einstein's theory of general relativity. General relativity is strongly constrained by weak field regime experiments. There have been efforts to test strong gravity regime by probing the metric around astrophysical candidates in the recent years[1]. The studies are based on properties of electromagnetic radiation emission of accretion disk. Kerr black holes are only characterized by mass and spin based on multipole moment expansions. Higher multipole moment expansions in a Kerr spacetime are only function of mass and spin. The compact object cannot be the Kerr black hole of general relativity in the case of measurement of independent higher multipole moments such as quadrupole moment by observational data.

Neutron stars have high density and strong gravity. They are fascinating laboratory to study possible deviation from prediction of general relativity, and in general, to study strong gravity regime. Multipole moment expansion can provide accurate approximation of spacetime surrounding a neutron star. Study of higher order multipole moments can reveal properties of neutron stars.

Approximate solution of exterior spacetime of neutron star is presented in reference[2]. The solution is based on Ernst formulation of general relativity up to five multipole moments, mass, angular momentum, mass quadrupole, spin octupole, and hexadecapole. The relation between these hairs and neutron star surrounding spacetime is also introduced in the reference. The hairs only depends on mass, spin, and quadrupole moment.

In present paper, I considered this metric based on Ernst formalism up to five multipole moments to study observational test of the neutron star exterior spacetime and distinguish it from Kerr spacetime. I employ iron line method. The iron line is prominent feature in X-ray reflection spectra. The emission is fluorescent narrow

lines from geometrically thin and optically thick accretion disk which produced by absorption of a hard X-ray photon of corona. This emission would be broadened and skewed in the inner region of the accretion disk due to relativistic effect of compact object there. In this paper, first, iron lines are simulated. Then, observational data are simulated and analyzed. The result is as follows. It is hard to constrain parameters of neutron star exterior spacetime using iron line method since both fast rotating and slow rotating case cannot be distinguished from Kerr solution of general relativity. The paper is structured as follows. In section 2, I present the spacetime metric. Section 3 is devoted to iron line method and data simulation and analysis is presented in section 3. Finally, section 4.1 is a short summary and conclusion. Throughout the paper, I employ units in which $G_N = c = 1$ and metric signature is $(-,+,+,+)$.

## 2. Theoretical Framework

A stationary and axially symmetric spacetime based on Ernst potential formalism to describe spacetime surrounding a neutron star is reported in [2]. The solution contains the first five relativistic multiple moments, the mass, $M$, angular momentum, $J$, the mass quadrupole, $M_2$, spin octupole, $S_3$, and the mass hexadecapole, $M_4$. The line element reads

$$ds^2 = -f\left(dt - \omega d\varphi\right)^2 + f^{-1}\left[e^{2\gamma}\left(d\rho^2 + dz^2\right) + \rho^2 d\varphi^2\right], \tag{1}$$

where metric functions $f$, $\omega$, and $\gamma$ are given as

$$\begin{aligned}
f(\rho,z) &= 1 - \frac{2M}{\sqrt{\rho^2+z^2}} + \frac{2M^2}{\rho^2+z^2} + \frac{\left(M_2 - M^3\right)\rho^2 - 2\left(M^3 + M_2\right)z^2}{\left(\rho^2+z^2\right)^{5/2}} \\
&\quad + \frac{2z^2\left(-J^2 + M^4 + 2M_2 M\right) - 2MM_2\rho^2}{\left(\rho^2+z^2\right)^3} \\
&\quad + \frac{A(\rho,z)}{28\left(\rho^2+z^2\right)^{9/2}} + \frac{B(\rho,z)}{14\left(\rho^2+z^2\right)^5}, \\
\omega(\rho,z) &= -\frac{2J\rho^2}{\left(\rho^2+z^2\right)^{3/2}} - \frac{2JM\rho^2}{\left(\rho^2+z^2\right)^2} + \frac{F(\rho,z)}{\left(\rho^2+z^2\right)^{7/2}} \\
&\quad + \frac{H(\rho,z)}{2\left(\rho^2+z^2\right)^4} + \frac{G(\rho,z)}{4\left(\rho^2+z^2\right)^{11/2}}, \\
\gamma(\rho,z) &= \frac{\rho^2\left(J^2\left(\rho^2-8z^2\right) + M\left(M^3+3M_2\right)\left(\rho^2-4z^2\right)\right)}{4\left(\rho^2+z^2\right)^4} - \frac{M^2\rho^2}{2\left(\rho^2+z^2\right)^2},
\end{aligned}$$

where

$$
\begin{aligned}
A(\rho,z) &= \left[8\rho^2z^2\left(24J^2M+17M^2M_2+21M_4\right)+\rho^4\left(-10J^2M+7M^5+32M_2M^2\right)\right.\\
&\quad \left.-21M_4+8z^4\left(20J^2M-7M^5-22M_2M^2-7M_4\right)\right],\\
B(\rho,z) &= \left[\rho^4\left(10J^2M^2+10M_2M^3+21M_4M+7M_2^2\right)+4z^4\left(-40J^2M^2-14JS_3\right.\right.\\
&\quad \left.+7M^6+30M_2M^3+14M_4M+7M_2^2\right)-4\rho^2z^2\left(27J^2M^2-21JS_3\right.\\
&\quad \left.\left.+7M^6+48M_2M^3+42M_4M+7M_2^2\right)\right],\\
H(\rho,z) &= \left[4\rho^2z^2\left(J\left(M_2-2M^3\right)-3MS_3\right)+\rho^4\left(JM_2+3MS_3\right)\right]\\
G(\rho,z) &= \left[\rho^2\left(J^3\left(-\left(\rho^4+8z^4-12\rho^2z^2\right)\right)+JM\left(\left(M^3+2M_2\right)\rho^4\right.\right.\\
&\quad -8\left(3M^3+2M_2\right)z^4+4\left(M^3+10M_2\right)\rho^2z^2\\
&\quad \left.\left.+M^2S_3\left(3\rho^4-40z^4+12\rho^2z^2\right)\right)\right]\\
F(\rho,z) &= \left[\rho^4\left(S_3-JM^2\right)-4\rho^2z^2\left(JM^2+S_3\right)\right].
\end{aligned}
$$

The first multipole moments can be expressed as follows to use for neutron star case

$$
\begin{aligned}
M_2 &= -\alpha j^2M^3\,,\\
S_3 &= -\beta j^3M^4\,,\\
M_4 &= \gamma j^4M^5\,,
\end{aligned}
\tag{2}
$$

where $M$ is mass and $j = J/M^2$ is spin parameter. The case $\alpha = \beta = \gamma = 1$ presents the Kerr case but these parameters can be large for neutron star case. The relation between $\alpha, \beta$, and $\gamma$ is[3,4]

$$
\begin{aligned}
y_1 &= -0.36 + 1.48\,x^{0.65}\\
y_2 &= -4.749 + 0.27613\,x^{1.5146} + 5.5168\,x^{0.22229}\,,
\end{aligned}
\tag{3}
$$

where $x = \sqrt{\alpha}\,, y_1 = \sqrt[3]{\beta}$, and $y_2 = \sqrt[4]{\gamma}$. Thus, one can determine the neutron star exterior spacetime by mass, spin, and parameter $\alpha$, ranges from $\sim 1.5$ for non-rotating case to 8 for cases with maximum value of spin parameter such as 0.5.

## 3. X-ray Reflection Spectroscopy

Iron line is prominent feature in X-ray reflection spectra of astrophysical black hole candidates. Iron atom absorbs hard X-ray photon of optically thin comptonized corona and emits fluorescence lines. This line is at 6.4 keV for weakly ionized or neutral atom and shifts up to 6.97 keV in the case of ionized H-like iron. This line would be broadened and skewed as the result of presence of strong gravity region and relativistic effects such as Doppler shift, gravitational redshift and light bending in the inner region of the accretion disk. This broad and skewed line provide us powerful tool to test strong gravity regime. This observational feature is determined by background metric, viewing angle, and emissivity profile of the disk. In my simulation, the inner edge of the disk is considered at the ISCO radius and

the outer edge is large enough that its impact is small. The emissivity is power-law, $1/r^q$, where $q$ is emissivity index. The assumption is $q = 3$ which corresponds to Newtonian limit at large radii. The code is described in [5,6].

The iron lines for metric 1 with two sets of spin parameter 0.5 and 0.2, and viewing angle 55° are presented in figures 1 and 2. The lines present slight difference with respect to Kerr one for case with spin 0.5 and the impact on the line profile with spin 0.2 is weak and maybe harder to constrain.

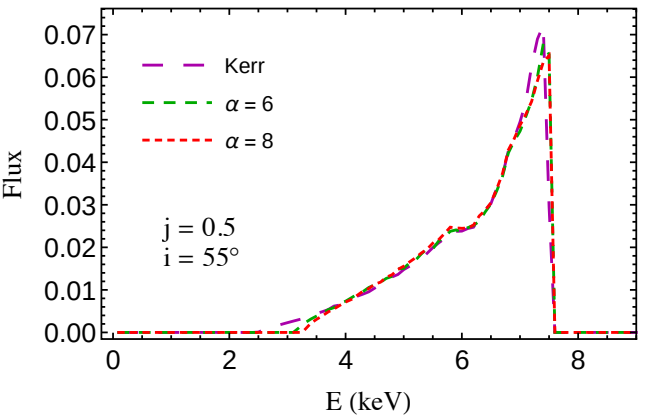



Fig. 1. Spin parameter value is 0.5 and viewing angle is 55°. The parameter values for $\alpha$ are 6 and 8. See text for more details.

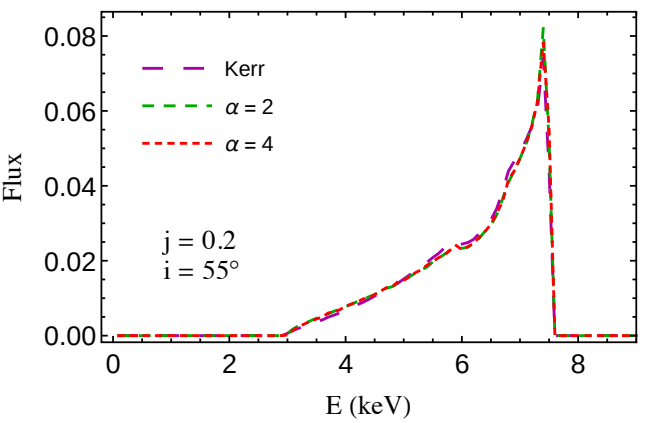



Fig. 2. This figure present the iron lines with spin parameter value 0.2 and viewing angle 55°. The parameter values for $\alpha$ are 2 and 4. See text for more details.

## 4. Simulating Data and Analysis

In this section, I simulate observed data of a neutron star and analyze them with a Kerr model. This tells us whether the Kerr model fit the data or not. I simulate data with a model consists of power-law and a single iron line of previous section. The considered flux is about $10^{-11}$ erg/cm$^2$/s in the range $2-10$ keV for neutron star case. The equivalent width is about 200 eV and photon index of power-law is 2. The simulation is done with *fakeit* command of XSPEC using RMF, ARF, and background file of LAD instrument with large area detector on board of enhanced X-ray Timing and Polarimetry (eXTP) China-Europe mission which is planned to launch before 2025. The LAD instrument has large effective area to provide more counts with less poisson noise. The data fitting is done with a power-law model and a RELINE model (a Kerr iron line). Figure 3 shows data to model ratio of best fit model for neutron star data simulation with spin parameter 0.5 and $\alpha = 8$ in the left panel. Right panel shows the case with spin value 0.2 and $\alpha = 2$. As we see there are slight differences for the case with spin 0.5 and there are not strong effect for spin value 0.2 which means it is hard to be constrained.

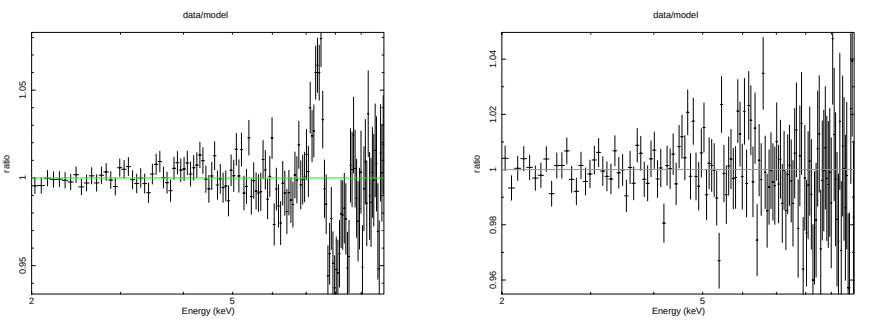



Fig. 3. Left panel: Data-to-model model ratio for simulation with iron line $j = 0.5$, $i = 55^\circ$ and $\alpha = 8$. The reduced $\chi^2$ is 2.1 and there are unresolved features. Right panel: The spin parameter for this figure is 0.5, $i = 55^\circ$ and $\alpha = 2$ for the iron line. The figure shows the data-to-model ratio. The reduced $\chi^2$ is 1.1 and there are no remarkable unresolved features, which makes this case harder constrain.

### 4.1. *Summary and Conclusion*

A solution of Einstein's field equations using Ernst potential formalism up to five multipole moments is introduced in the reference[2]. This method can accurately describe the exterior spacetime of neutron star. The iron line emitted from medium around neutron star is simulated in the present paper. Two different values of spin parameter 0.5 and 0.2 with different value of $\alpha$ are considered. The impact of iron lines with spin value 0.5 is not very strong and there is no impact on iron lines spin

value 0.2. The 100ks observation with LAD-eXTP for neutron star with different sets of parameter value are also presented. If the fit is not good, the neutron star exterior spacetime parameters and the deviations from the Kerr case can be constrained. The fit for higher value of spin parameter 0.5 and $\alpha = 8$ does not seem good. But the impact on iron lines are small and also there are uncertainties. Thus, it might be hard to constrain neutron star solution. The fit for spin value 0.2 and $\alpha = 2$ is good which indicates it is not possible to constrain deviations from Kerr solution.

## Acknowledgments

This work is partly supported by the National Key Program for Science and Technology Research and Development (Grant No. 2016YFA0400704), the National Natural Science Foundation of China under grant No. 11690024 and 11390372, and the Strategic Priority Program of the Chinese Academy of Sciences (Grant No. XDB 23040100). I also acknowledge support from the China Postdoctoral Science Foundation, Grant No. 2017LH021.

## References

1. J. Schee and Z. Stuchlik, Int. J. Mod. Phys. D **18**, 983 (2009), J. Schee and Z. Stuchlik, Gen. Rel. Grav. **41**, 1795 (2009), T. Harko, Z. Kovacs and F. S. N. Lobo, Phys. Rev. D **79**, 064001 (2009), T. Harko, Z. Kovacs and F. S. N. Lobo, Phys. Rev. D **80**, 044021 (2009), T. Harko, Z. Kovacs and F. S. N. Lobo, Class. Quant. Grav. **27**, 105010 (2010), C. Bambi and N. Yoshida, Phys. Rev. D **82**, 064002 (2010), T. Johannsen and D. Psaltis, Astrophys. J. **726**, 11 (2011), C. Bambi and E. Barausse, Astrophys. J. **731**, 121 (2011), T. Johannsen and D. Psaltis, Astrophys. J. **773**, 57 (2013), C. Bambi, JCAP **1209**, 014 (2012), N. Tsukamoto, Z. Li and C. Bambi, JCAP **1406**, 043 (2014), Z. Li and C. Bambi, Phys. Rev. D **90**, 024071 (2014), C. Bambi, J. Jiang and J. F. Steiner, Class. Quant. Grav. **33**, 064001 (2016), K. Yagi and L. C. Stein, Class. Quant. Grav. **33**, 054001 (2016), T. Johannsen, Class. Quant. Grav. **33**, 124001 (2016), M. Ghasemi-Nodehi and C. Bambi, Phys. Rev. D **94**, 104062 (2016), C. Bambi, Rev. Mod. Phys. **89**, 025001 (2017),
2. G. Pappas, Mon. Not. Roy. Astron. Soc. **466**, no. 4, 4381 (2017) [arXiv:1610.05370 [gr-qc]].
3. G. Pappas and T. A. Apostolatos, Phys. Rev. Lett. **112**, 121101 (2014) [arXiv:1311.5508 [gr-qc]].
4. K. Yagi, K. Kyutoku, G. Pappas, N. Yunes and T. A. Apostolatos, Phys. Rev. D **89**, no. 12, 124013 (2014) [arXiv:1403.6243 [gr-qc]].
5. C. Bambi, Astrophys. J. **761**, 174 (2012) [arXiv:1210.5679 [gr-qc]].
6. C. Bambi, Phys. Rev. D **87**, 023007 (2013) [arXiv:1211.2513 [gr-qc]].

# Super-massive black hole mass estimation from bright flares

Vladimír Karas,* Michal Bursa and Michal Dovčiak

*Astronomical Institute, Czech Academy of Sciences,*
*Boční II 1401, CZ-14100 Prague, Czech Republic*

Andreas Eckart,† Monika Valencia-S, Munawwar Khanduwala, and Michal Zajaček†‡

*I. Physikalisches Institut, Universität zu Köln,*
*Zülpicher Str. 77, D-50937 Cologne, Germany*

Super-Massive Black Holes reside in galactic nuclei, where they exhibit episodic bright flares due to accretion events. Taking into account relativistic effects, namely, the boosting and lensing of X-ray flares, we further examine the possibility to constraint the mass of the SMBH from the predicted profiles of the observed light curves. To this end, we have studied four bright flares from Sagittarius A*, which exhibit an asymmetric shape consistent with a combination of two intrinsically separate peaks that occur with a specific time delay with respect to each other. We have thus proposed (Karssen et al. 2017, Mon. Not. R. Astron. Soc. 472, 4422) that an interplay of relativistic effects could be responsible for the shape of the observed light curves and we tested the reliability of the method.

*Keywords*: Gravitation; Black Holes; Accretion; Galactic center.

## 1. Introduction

Astrophysical black holes influence their cosmic environs by gravitational and electromagnetic effects. They are described by a small number of free parameters, namely, the mass and angular momentum.[1] All evidence as well as theoretical arguments point to the fact that the effects of electric charge must be very small (at least in the steady-state limit), whereas the magnetic monopole and additional even more exotic parameters are thought to be negligible or they vanish completely to zero.[2–4] Several independent methods have been developed to measure the black hole mass as a primary parameter defining the black hole action on luminous matter in their neigborhood.[5–7] A new method of mass determination from optical polarization of broad emission lines has been proposed to constrain SMBH mass in active galactic nuclei.[8,9] In X-rays, also the possibility of measuring the black hole mass from the energy shift of narrow spectral lines has been discussed widely.[10–13] We describe a recently developed method that is based on the assumption that a fast orbital motion of radiating X-ray flares can be detected in time variability of accreting black holes.[14,15]

*E-mail: vladimir.karas@cuni.cz
†Also at Max-Planck-Institut für Radioastronomie, Bonn, Germany
‡New address: Center for Theoretical Physics, Polish Academy of Sciences, Warsaw, Poland

## 2. Bright flares from a source orbiting near a black hole

The time-scale of matter orbiting along $r$ = const circular trajectory around Kerr black hole of mass $M$ can be written

$$T_{\rm orb}(r;a) \simeq 310(r^{\frac{3}{2}} + a)M_7 \text{ sec.} \quad (1)$$

Hereafter we use dimensionless geometric units, where length is expressed in terms of gravitational radius, $R_g \equiv GM/c^2 \simeq 1.5 \times 10^{12} M_7$ cm and the spin parameter $-1 \leq a \leq 1$ (positive values for the prograde rotation with the black hole). The dimension-less angular momentum $a$ adopts values in the range $-1 \leq a \leq 1$. Positive values correspond to co-rotating motion, while negative values describe counter-rotation (many papers assume that the accretion disc co-rotates, although such an assumption may not be necessarily true). Circular orbits of free particles are possible above the innermost stable circular orbit (ISCO a.k.a. marginally stable orbit).[16] Gravitation governs the orbital motion near the horizon, and so the apparent variability associated with the motion can be scaled with the black hole mass.

The gravitational field is described by the metric of Kerr black hole[1]

$$ds^2 = -\frac{\Delta}{\Sigma}\Big(dt - a\sin^2\theta\, d\phi\Big)^2 + \frac{\Sigma}{\Delta}\, dr^2 + \Sigma\, d\theta^2 + \frac{\sin^2\theta}{\Sigma}\Big[a\, dt - \big(r^2 + a^2\big)\, d\phi\Big]^2 \quad (2)$$

in Boyer-Lindquist (spheroidal) coordinates $t$, $r$, $\theta$, $\phi$. The metric functions $\Delta(r)$ and $\Sigma(r,\theta)$ are known in an explicit form. The event horizon occurs at the roots of equation $\Delta(r) = 0$; the outer solution is found given by $r = R_+ = 1 + (1-a^2)^{1/2}$, which exists for $|a| < 1$. Once the black hole rotates ($a \neq 0$), particles and photons are pushed to co-rotate with the black hole due to frame-dragging effect. The co-rotation is obligatory within the ergosphere. Let us note that the Kerr spin parameter is limited to an equilibrium value by photon recapture from the disc ($a \simeq 0.998$) and by magnetic torques.[17,18]

The existence of ISCO is a remarkable feature of motion in strong gravity.[16] For a co-rotating equatorial disc,

$$R_{\rm ISCO} = 3 + z_2 - \big[\,(3-z_1)(3+z_1+2z_2)\,\big]^{\frac{1}{2}}, \quad (3)$$

where $z_1 = 1 + \alpha_+\alpha_-[\alpha_+ + \alpha_-]$, $\alpha_\pm = (1 \pm a)^{\frac{1}{3}}$, and $z_2 = (3a^2 + Z_1^2)^{\frac{1}{2}}$. The ISCO radius as function of spin spans the range of dimensionless radius from $R_{\rm ISCO} = 1$ (for $a = 1$, i.e. a maximally co-rotating Kerr black hole) to $R_{\rm ISCO} = 6$ (for $a = 0$, a static case of Schwarzschild black hole) to $R_{\rm ISCO} = 9$ ($a = -1$, case of maximum counter-rotation). The velocity of prograde Keplerian circular motion is

$$v^{(\phi)} = \Delta^{-1/2}\left(r^2 - 2ar^{1/2} + a^2\right)\left(r^{3/2} + a\right)^{-1} \quad (4)$$

with respect to locally non-rotating observers (LNRF). The corresponding angular velocity is $\Omega(r;a) = (r^{3/2} + a)^{-1}$. Finally, to derive the interval measured by

a distant observer, one needs to consider the Lorentz factor associated with the orbital motion,

$$\Gamma = \frac{\left(r^{3/2} + a\right) \Delta^{1/2}}{r^{1/4} \sqrt{r^{3/2} - 3r^{1/2} + 2a}\, \sqrt{r^3 + a^2 r + 2a^2}}\,. \tag{5}$$

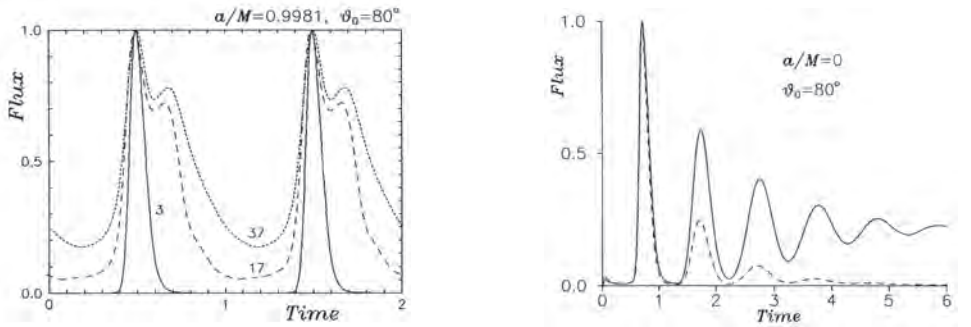



Fig. 1. Exemplary light curves of the signal from a spot orbiting on Keplerian circular trajectory near Kerr black hole. Effects of General Relativity have been taken into account, which allows us to constrain the parameters of the system. Left panel: radiation flux variation over a two revolutions. Parameters of the system are specified with the plot (radius of the orbits is given with each curve in units of the black hole gravitational radius, $R_g = GM/c^2$). Right panel: Two cases of tidally decaying spots due to differential rotation; (a) constant total emission from the spot (solid line) vs. (b) exponentially decaying intrinsic emission with time (dashed line), motivated by modelling the origin of the spot. Location of the spot is centered at $r = 6R_+$.[21]

Geodesic motion is determined by three constants of motion: the total energy $\mathcal{E}$, the azimuthal component of angular momentum $L_{\rm z}$, and Carter's constant $Q$. Null geodesics are relevant to describe propagation of photons, and for these the number of free constants can be reduced: $\xi = L_z/\mathcal{E}$, $\eta = Q/\mathcal{E}^2$. For photons propagating from the accretion disk towards a distant observer, the initial point is set at a given radius in the equatorial plane of the black hole, whereas the final point is at radial infinity, along the view angle of the observer. Relativistic effects become more prominent an high inclination angles.[12,19]

We employ geometrical optics and derive predicted light-curves from a source of light orbiting at a particular radius with the equatorial accretion disk. The equation for photon rays (null geodesics) relates an emission point $\mathcal{P}_0 = (r, \theta, \phi)_{\rm em}$ near the black hole with the terminal point $(\alpha, \beta)$ in observer's detector plane at spatial infinity.[20,21] The rays can be integrated in terms of elliptic integrals,

$$\int_{R_{\rm em}}^{R} \mathcal{R}(r')^{-1/2}\, dr' = \int_{\theta_{\rm em}}^{\theta} \Theta(\theta')^{-1/2}\, d\theta = \int_{\phi_{\rm em}}^{\phi} \mathcal{F}(\phi')^{-1}\, d\phi', \tag{6}$$

where

$$\mathcal{R}(r) = r^4 + (a^2 - \xi^2 - \eta)r^2 + 2[\eta + (\xi - a)^2]r - a^2\eta, \tag{7}$$
$$\Theta(\theta) = \eta + a^2\cos^2\theta - \xi^2\cot^2\theta, \tag{8}$$
$$\mathcal{F}(\phi) = [2ar + (\Sigma - 2r)\xi\csc^2\theta]\Delta^{-1}, \tag{9}$$
$$\Delta(r) = r^2 - 2r + a^2, \quad \Sigma = r^2 + a^2\cos^2\theta, \tag{10}$$
$$A(r,\theta) = (r^2 + a^2)^2 - \Delta(r)a^2\sin^2\theta. \tag{11}$$

The two constants of motion of the photon motion satisfy

$$\xi = A^{1/2}\Sigma^{-1/2}E^{-1}\sin\theta\sin\alpha\sin\beta_{|\mathcal{P}_0}\,, \tag{12}$$
$$\eta = \Delta^{-1}(r^2 + a^2 - a\xi)^2 - \Sigma E^{-2}\cos^2\alpha - \xi^2 + 2a\xi - a^2{}_{|\mathcal{P}_0} \tag{13}$$

with $EA^{1/2} = \Sigma^{1/2}\Delta^{1/2} + 2ar\Sigma^{-1/2}\sin\theta\,\sin\alpha\,\sin\beta$. Finally, time coordinate can be integrated in the form

$$t = \int_{t_0}^{t} \frac{A(r,\theta) - 2ar\xi}{\Sigma(r,\theta)\,\Delta(r)}\,d\tau. \tag{14}$$

To show exemplary solutions of the above-given equations and to illustrate the main effects of General Relativity, we plot several model light curves of an orbiting spot in Fig. 1. The main parameters are the radial position of the center of the spot, view angle of a distant observer, and the Kerr parameter of the black hole.[21–23] One can distinguish two local maxima that dominate the light curve variation over the revolution of the spot: (i) the Doppler boosting peak near the phase $\approx 0.7$ (on the approaching side of the trajectory with respect to the observer, with a certain time delay due to light-travel time), and (ii) the lensing peak near the phase 0.5 (due to light bending at the point of upper conjunction of the trajectory). Only one of the peaks may appear in some cases; the occurrence and the relative height of the two peaks depend on the model parameters. For spots on a close orbit around the black hole, the light rays are bent to the extent that the Doppler-peak follows immediately after the lensing peak, whereas for larger orbits the two peaks can be a quarter of an orbit apart from each other. We propose that the changing profile can be used to constrain the parameters of the system; in particular, this can allow us to determine the parameters of the central black hole.[14]

Solving the above-given equations usually requires to perform several steps numerically. Alternatively, let us mention that a number of semi-analytical approximations have been formulated that allow us to examine selected aspects (e.g. extremal values of the redshift function from an orbiting spot).[24–26] These can provide simple and practical estimates that are useful in analyzing the redshifted narrow lines, however, the approximation approaches assume additional limitations that reduce the accuracy.

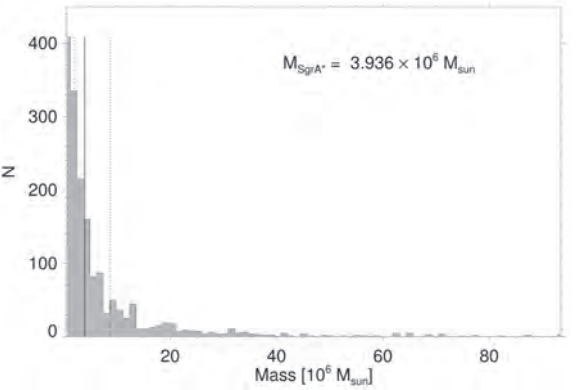



Fig. 2. Histogram of the predicted mass of Sgr A* supermassive black hole derived from a combination of the four flares taken into account. The most likely mass predicted by the method is the median value indicated by the solid vertical line; figure from Karssen et al. 2017.[14]

## 3. The Method and Results

Karssen et al. (2017) studied the light curves of the four bright flares of Sagittarius A* (Sgr A*) in the center of the Milky Way. They allowed the mass of the black to vary as a free parameter and employed the model profiles to fit the most likely value. The flare time-scale in periods is matched with the actual duration of the observed flare in physical units, which provides a way to constrain the black hole mass. The quality and the volume of data for Sgr A* SMBH is the best, however, the same approach can be employed also with other bright objects. In particular, the method has been tested with the Seyfert I galaxy RE J1034+396.[14]

The simulated light curves are normalized to the maximum flux of the light curve, which is set to be identical to the observed light curve. To conduct a time efficient fit of the models to the data, we introduce a time shift, a flux density scaling factor and a flux density offset. For each light curve the ratio is considered between the number of data points of the light curve which belong to the flaring period (defined by the requirement to on the flux to exceed 30 per cent of the maximum value) and the number of data points of the quiescent state. The main features of the flares are contained in the upper two-thirds of the flares. Every simulated light curve ratio is then compared to an observed one; quiescent state data points are removed from or added to the simulated data until the ratio is comparable.

By the fitting process the theoretical light curves are lined up with the observed ones and the conversion factor is determined between geometrized time units of a particular model and the physical time-scale of the observation. This relation depends on the mass of the black hole. By gauging the intrinsic clock of the black hole in its gravitational units to the clocks of the observations in seconds, the mass of the black hole can be estimated. Each fit of a particular simulated light curve to a particular observed flare then results in an estimate for $M$. The simulated data is

multiplied by a factor because in general the best-fit shape of the light curves does not depend on the initial normalization we inferred earlier. In Fig. 2 we show the resulting mass predicted from a combination of the four brightest Sgr A* flares by; the y-axis of the diagram represents the number of model results that have predicted the particular mass value. A peak in this diagram occurs where the most models predicting the corresponding value in the particular mass bin find the minimum of best-fit statistics. Let us note that the resulting distributions are not normal distributions, hence we use the medians as a measure for the most probable mass, as indicated by the vertical solid line in the graph.

## 4. Conclusions

The null geodesics were obtained by solving numerically the equations governing the photon propagation in Kerr spacetime. In this initial work we have not included additional complexities due to the intrinsic variation (decay) of the flares; such a generalization is postponed for a future study, however, the necessary methodology has been already developed; cf. the right panel in Fig. 1, where an exponentially fading signal of a tidally sheared spot has been studied.

We have outlined a method based on a comparison between the simulations with the bright X-ray flares from Sgr A*, which gives an estimate on the mass of the black hole. Let us emphasize that the result does not depend on the uncertainties about the object distance, which otherwise complicates approaches to the mass estimation by different methods. On the other hand, the hereby described method has other underlying assumptions and it can be applied only if the light curves of the flare events are dominated by the effects of relativistic motion of the source orbiting the black hole. This may be adequate for a limited sub-sample of targets; we have argued that the bright flares belong within the suitable category. For further details, see Karssen et al. (2017).[14]

## Acknowledgments

The authors acknowledge the Czech Science Foundation (GAČR) – German Research Foundation (DFG) collaboration project No. 19-01137J and the Czech Ministry of Education, Youth and Sports collaboration program Kontakt ref. LTAUSA 17095.

## References

1. Misner C. W., Thorne K. S. and Wheeler J. A., *Gravitation* (Freeman, San Francisco, 1973).
2. Kovář J., Kopáček O., Karas V. and Stuchlák Z., *Classical and Quantum Gravity* **27**, id. 135006 (2010).
3. Zajaček M., Tursunov A., Eckart A. and Britzen S., *Mon. Not. R. Astron. Soc.* **480**, 4408 (2018).

4. Zajaček M., Tursunov A., Eckart A., Britzen S., et al., in *Proc. of the Conference "FISICPAC-2018"* (University of Sharjah, UAE), *Journal of Physics: Conference Series*, submitted; arXiv:1812.03574 (2019).
5. Kormendy J. and Richstone J., *Ann. Rev. of Astron. and Astrophys.* **33**, 581 (1995).
6. Kormendy J. and Ho L., *Ann. Rev. of Astron. and Astrophys.* **51**, 511 (2013).
7. Czerny B. and Nikolajuk M., *Memorie della Società Astronomica Italiana* **81**, 281 (2010).
8. Afanasiev V. L. and Popović L. Č., *Astrophysical Journal* **800**, id. L35 (2015).
9. Savić D., Goosmann R., Popović L. Č., Marin F. and Afanasiev V. L., *The Astrophysical Journal* **614**, id. A120 (2018).
10. Dovčiak M., Bianchi S., Guainazzi M., Karas V. and Matt G., *Astronomy & Astrophysics* **350**, 745 (2004).
11. Pecháček T., Dovčiak M., Karas V. and Matt G., *Astronomy & Astrophysics* **441**, 855 (2005).
12. Karas V. and Sochora V., *The Astrophysical Journal* **725**, 1507 (2010).
13. Sochora V., Karas V., Svoboda J. and Dovčiak M., *Mon. Not. R. Astron. Soc.* **418, 276** (2011).
14. Karssen G. D., Bursa M., Eckart A., Valencia-S M., Dovčiak M., Karas V. and Horák, J., *Mon. Not. R. Astron. Soc.* **472** 4422 (2017).
15. Eckart A., Parsa M., Mossoux E. and Shahzamanian B., et al., in *Proc. of the Conference "Frontier Research in Astrophysics – III"* (Palermo, Italy); arXiv:1806.01096 (2018).
16. Bardeen J. M.; Press W. H. and Teukolsky S. A., *The Astrophysical Journal* **178**, 347 (1972).
17. Thorne K. S., *Astrophysical Journal* **191** 507 (1974).
18. Krolik J. H., Hawley J. F. and Hirose S., *The Astrophysical Journal* **622**, 1008 (2005).
19. Zakharov A. F. and Repin S. V., *Astronomy and Astrophysics* **406**, 7-13 (2003).
20. Carter B., *Physical Review* **174**, 1559 (1968).
21. Karas V., Vokrouhlický D. and Polnarev A. G., *Mon. Not. R. Astron. Soc.* **259**, 569 (1992).
22. Dovčiak M., Karas V., Martocchia A., Matt G. and Yaqoob, T., in *Proceedings of RAGtime 4/5 Workshops on black holes and neutron stars*, eds. S. Hledík and Z. Stuchlík (Silesian University in Opava, Czech Republic), pp. 33–73 (2004).
23. Dovčiak M., Karas V., Martocchia A., Matt G. and Yaqoob, T., *The Astrophysical Journal Supplement Series* **153**, 205 (2004).
24. Pecháček T., Dovčiak M., Karas V. and Matt, G., *Astronomy & Astrophysics* **441**, 855 (2005).
25. Semerák O., *The Astrophysical Journal* **800**, id. 77 (2015).
26. De Falco V., Falanga M. and Stella L., *Astronomy & Astrophysics* **595**, id.A38 (2016).

# Geodesics in a rotating black hole spacetime surrounded by quintessence

Hemwati Nandan* and Prateek Sharma

*Department of Physics, Gurukula Kangri Vishwavidyalaya, Haridwar, Uttarakhand 249 404, India*
**E-mail: hnandan@associates.iucaa.in*
*http://www.gkv.ac.in/*

Rashmi Uniyal

*Department of Physics, Government Degree College, Narendranagar , Uttarakhand 249 175, India*
*E-mail: rashmiuniyal001@gmail.com*

Philippe Jetzer

*Physik-Institut,University of Zürich, Winterthurerstrasse 190, 8057 Zurich, Switzerland*
*E-mail: jetzer@physik.uzh.ch*

The quintessence is one of the several candidates which represents the dark energy responsible for the acceleration of the universe. The quintessence is described by a canonical scalar field minimally coupled to gravity. We study the timelike geodesic congruences in the background of a rotating black hole spacetime surrounded by quintessence in equatorial plane. The effect of equation of state (EOS) parameter for quintessence and the normalization factor on geodesics is investigated in detail in view of the structure of possible orbits, including the innermost stable circular orbits (ISCOs). The structure of photon orbits is also studied for the different values of parameters involved in. The results obtained are also then compared with those of the Kerr black hole spacetime and Schwarzschild black hole spacetime in GR with or without quintessence.

*Keywords*: Geodesics, Black Holes, Dark Energy, Quintessence, ISCO.

## 1. Introduction

The Black Holes (BHs) are one of the numerous consequences of General Relativity (GR)[1]. A number of BH spacetimes emerge as the solution of Einstein field equations in GR and other alternative theories of gravity. The study of geodesic motion in the background of a given spacetime is one of the most an elegant ways to describe the underlying geometry of that particular spacetime[2]. Further, different observations from supernovae (Type Ia), cosmic microwave background radiation (CMBR), Baryon acoustic oscillations (BAO) and the Hubble measurements strongly advocates for an accelerated expansion of our universe which is believed to be due to the presence of some mysterious form of energy known as dark energy having a large negative pressure[2–4]. There are several candidates to represent the dark energy in our universe and *quintessence* is one among them. It would therefore be quite interesting and important to study the geodesic motion in the background of a rotating BH spacetime surrounded by quintessence to see the effect of dark energy, if any, on the universe locally.

## 2. The Effective Potential

The metric of a Rotating BH surrounded with quintessences is given below[2,5,6],

$$ds^2 = \frac{\left(\Delta - a^2 sin^2\theta\right) dt^2}{\Sigma} - \frac{\Sigma dr^2}{\Delta} + 2asin^2\theta \left[1 - \frac{\Delta - a^2 sin^2\theta}{\Sigma}\right] dtd\phi - \Sigma d\theta^2 - sin^2\theta \left[\Sigma + a^2 sin^2\theta \left(2 - \frac{\Delta - a^2 sin^2\theta}{\Sigma}\right)\right] d\phi^2, \quad (1)$$

with, $\Delta = r^2 + a^2 - 2Mr - \frac{\alpha}{\Sigma^{\frac{3\omega-1}{2}}}$ and $\Sigma = r^2 + a^2 cos^2\theta$ where $\alpha$ is normalization factor and $\omega$ is quintessential EOS parameter. For equatorial plane (i.e. $\theta = \frac{\pi}{2}$), the Lagrangian corresponding to the spacetime given by (1) can be written as[1],

$$2\mathcal{L} = (1 - f_r)\,\dot{t}^2 - \left(\frac{r^2}{r^2(1 - f_r) + a^2}\right)\dot{r}^2 + (2af_r)\,\dot{t}\dot{\phi} - \left(r^2 + a^2(1 + f_r)\right)\dot{\phi}^2, \quad (2)$$

where, $f_r = \frac{2M}{r} + \frac{\alpha}{r^{3\omega+1}}$. The generalized momentum may thus be written as,

$$\dot{t} = \frac{E\left(r^2 + a^2(1 + f_r)\right) - af_rL}{r^2(1 - f_r) + a^2}, \quad (3)$$

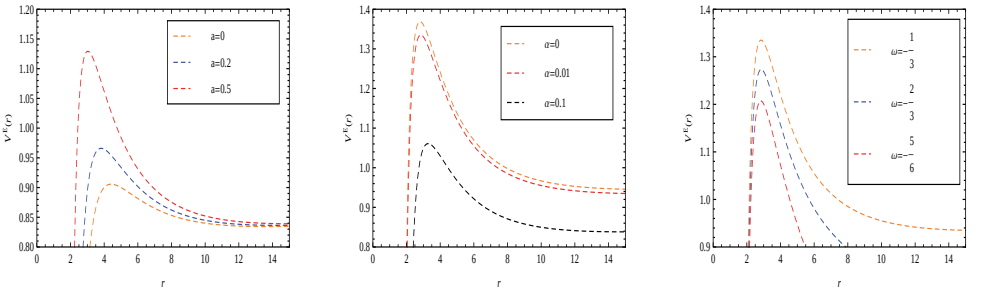
$$\dot{\phi} = \frac{af_rE + (1 - f_r)L}{r^2(1 - f_r) + a^2}, \quad (4)$$

$$\dot{r}^2 = E^2 + \frac{2M(L - aE)^2}{r^3} - \frac{\left(L^2 - a^2E^2\right)}{r^2} - \frac{r^2 + a^2 - 2Mr}{r^2} + f_r\frac{(L - aE)^2}{r^2} + f_r. \quad (5)$$

Here $E$ is the total energy per unit rest mass of test particle at infinity while $L$ is the angular momentum per unit rest mass of test particle. The effective potential $(V)$ with $\dot{r}^2 = E^2 - V^2$ then reads as[7–9],

$$V = \frac{\beta + \sqrt{\beta^2 - \delta\gamma}}{\delta}, \quad (6)$$

where, $\beta = \frac{La}{r^2}f_r$, $\delta = \frac{r^2 + a^2(1 + f_r)}{r^2}$ and $\gamma = \frac{L^2}{r^2}(f_r - 1) - 1 - \frac{a^2}{r^2} + f_r$.




(a) (b) (c)

Fig. 1. Effective Potential with M = 1, $L = \sqrt{20}$ for (a) $\alpha = 0.1$ $\omega = -1/3$ (b) a = 0.4 and $\omega = -1/3$ (c) $\alpha = 0.1$ and a = 0.4.

## 3. The ISCO Parameters

With $r = \frac{1}{u}, L - aE = x$ and the condition for Innermost Stable Circular Orbit (ISCO) $V = 0, V' = 0$ respectively further leads to [1,5,7,10–13] ,

$$E^2 + 2Mu^3x^2 - (x^2 + 2Eax)u^2 - (a^2u^2 - 2Mu + 1) + x^2u^2f_u + f_u = 0, \quad (7)$$

and

$$3Mx^2u^3 - (a^2u^2 - Mu) - (x^2 + 2Eax)u^2 + x^2u^2f_u + \left(x^2u^2 + 1\right)\left(\frac{3\omega + 1}{2}\right)f_u = 0, \quad (8)$$

here, $f_u = \alpha u^{3w+1}$. Solving equations (7) and (8) simultaneously,

$$x = -\frac{au \pm \sqrt{Mu + \left(\frac{3\omega + 1}{2}\right)f_u}}{uY}, \quad (9)$$

$$E = \frac{1 - 2Mu \mp au\sqrt{Mu + \left(\frac{3\omega + 1}{2}\right)f_u} - f_u}{Y}, \quad (10)$$

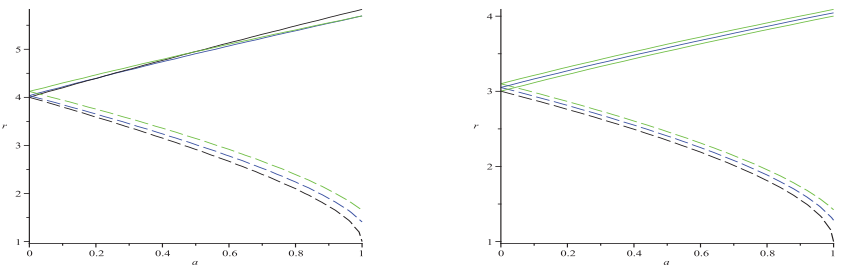
$$L = \frac{\mp\left(a^2u^2 + 1\right)\sqrt{Mu + \frac{(3\omega + 1)}{2}f_u} - 2Mau^2 - auf_u}{uY}, \quad (11)$$

with,

$$Y^2 = 1 - 3Mu - \left(\frac{3\omega + 3}{2}\right)f_u \mp 2au\sqrt{Mu + \left(\frac{3\omega + 1}{2}\right)f_u}, \quad (12)$$

where $\mp$ signature correspond to direct (i.e. co-rotating with $L > 0$) and retrograde (i.e. counter-rotating with $L < 0$) orbits respectively. The photon orbit is found to only exist if the denominator of equation (10) is real i.e.

$$1 - 3Mu - \left(\frac{3\omega + 3}{2}\right) f_u = \pm 2au\sqrt{Mu + \left(\frac{3\omega + 1}{2}\right) f_u}. \tag{13}$$




(a) (b)

Fig. 2. Variation of (a) Radii of marginally bound orbits (b) Radii of photon orbits. Here, black, blue and green colors correspond to $\alpha = 0$, 0.05, 0.1 at $\omega = -1/3$ as a function of specific angular momentum. Here, dash and solid lines represent direct and retrograde orbits respectively.

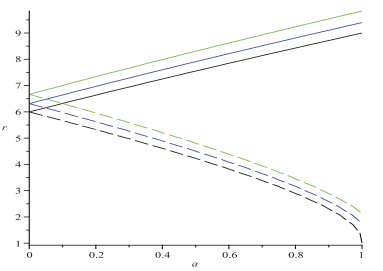



Fig. 3. Marginally stable orbit M = 1 for $\omega = -1/3$ here black, blue and green colors correspond to the cases $\alpha = 0$, $\alpha = 0.01$ and $\alpha$ =0.1 respectively. Here, dash and solid lines represent direct and retrograde orbit respectively.

Similarly for a marginally bound (MB) orbit, one can examine the variation of $r$ with $a$ and the condition for MB orbits just read as $E = \mu$. However, bound orbits are not necessarily stable and for the stability, we have the condition $V'' = 0$ [10,14,15]. For $\omega = -1/3$, the condition for marginally stable (MS) orbit is reduced as below,

$$3a^2 + 6Mr - r^2 \pm 8a\sqrt{Mr} + \alpha r^2 \leq 0. \tag{14}$$

The effective potential, ISCO parameters and radius of circular orbits (photon orbit, marginally bound orbit and ISCO orbit) are also visually presented (see figures for different sets of parameters involved in.

## 4. Conclusions

We have investigated the geodesic motion and structure of orbits in particular ISCO, MB and Photon orbits in the background of a rotating BH in presence of quintessence which is spherically symmetric solution of Einstein field equations with the quintessence as a source of energy-momentum. Some of the key results obtained are summarized below.

(i) The extrema of effective potential lies at lower value as compared to Kerr BH in GR when normalization factor ($\alpha$) increases and EOS parameter ($\omega$) for quintessence decreases which therefore indicates that unstable orbits can even be found at lower energy values.
(ii) The radius of direct photon orbit, MB orbit and MS orbit increases as normalization factor ($\alpha$) increases for the case of a non rotating BH.
(iii) The radius of photon orbit, MB orbit, and MS orbit are larger as compared to the Kerr BH when normalization factor increases, EOS parameter $\omega = -1/3$ and for finite value of spin parameter 'a' which in fact lies between 0 and M.
(iv) The direct photon orbit, direct MB orbit and direct MS orbit coincide at $a = M$ which is the same as for Kerr BH in GR which is basically due to the subtle nature of Boyer-Lindquist coordinates at $r = M$ for $a = M$.
(v) The radius of retrograde orbit for photons and retrograde MB orbits increases as normalization factor increases, for a fixed value of EOS parameter while spin parameter is small.
(vi) The radius of retrograde MS orbit increases with the increase in normalization factor at a particular value of EOS parameter i.e. $\omega = -1/3$.

In view of the current observations related to dark energy, the value of EOS parameter $\omega$ tends to $-1$ and similar investigations for the value $\omega = -1$ along with $\omega = -2/3$ are currently underway.

## Acknowledgments

The authors (HN, PS, RU) are thankful to Science and Engineering Research Board (SERB), Department of Science and Technology (DST), New Delhi, India, for providing the financial support through grant number **EMR/2017/000339**.

## References

1. Chandrasekhar, S, *The mathematical theory of black holes* (Oxford: Clarendon Press 1983).
2. R. Uniyal, and N. Chandrachani Devi, and H. Nandan, and K.D. Purohit, *Geodesic Motion in Schwarzschild Spacetime Surrounded by Quintessence*, (Gen. Rel. Grav. **47** 2, 2015 and references therein).
3. E. J. Copeland, M. Sami and S. Tsujikawa, *Dynamics of dark energy*, (International Journal of Modern Physics D, **15** 1753, 2004).

4. H.Nandan and R.Uniyal, *Geodesic Flows in a Charged Black Hole Spacetime with Quintessence*, (Eur. Phys. J. C **77**, no. 8, 552 2017).
5. B. Toshmatov, Z. Stuchlík, and B. Ahmedov, *Rotating black hole solutions with quintessential energy* (The European Physical Journal Plus **132** 98, 2017).
6. S. G. Ghosh, *Rotating black hole and quintessence*, (Eur. Phys. J. **C76** 222, 2016).
7. P. I. Jefremov, O. Yu. Tsupko, and G. S. Bisnovatyi-Kogan, *Innermost stable circular orbits of spinning test particles in Schwarzschild and Kerr space-times*, (Phys. Rev. **D91**, 124030, 2015).
8. R. Uniyal, H. Nandan and K. D. Purohit, *Null geodesics and observables around the KerrSen black hole*, (Class. Quant. Grav. **35**, no. 2, 025003 2018).
9. R. Uniyal, H. Nandan, A. Biswas and K. D. Purohit, *Geodesic motion in R-charged black hole spacetimes*, (Phys. Rev. D **92**, no. 8, 084023 2015).
10. J. M. Bardeen, W. H. Press and S. A. Teukolsky, *Rotating black holes: locally nonrotating frames, energy extraction, and scalar synchrotron radiation*, (The Astrophysical Journal **178** 347, 1972).
11. Yu-Peng Zhang, Shao-Wen Wei, Guo, Tao-Tao Sui and Yu-Xiao Liu, *Innermost stable circular orbit of spinning particle in charged spinning black hole background*, (Phys. Rev. **D97** 084056, 2018).
12. D. Pugliese, H. Quevedo and R. Ruffini, *Equatorial circular motion in Kerr spacetime*, (Phys. Rev. **D84** 044030, 2011).
13. Hyun-Chul Lee and Yong-Jin Han, *Innermost stable circular orbit of Kerr-MOG black hole*, (The European Physical Journal C **77** 655, 2017).
14. S. Beheshti and E. Gasperín, *Marginally stable circular orbits in stationary axisymmetric spacetimes*, (Phys. Rev. **D94** 024015, 2016).
15. S. Hod, *Marginally bound (critical) geodesics of rapidly rotating black holes*, (Phys. Rev. **D88** 087502, 2013).

# Maximal efficiency of collisional Penrose process with spinning particles

Kazumasa Okabayashi* and Kei-ichi Maeda

*Department of Physics, Waseda University, Shinjuku, Tokyo 169-8555, Japan*
** bayashioka@ruri.waseda.jp*
*maeda@waseda.jp*

Hirotada Okawa

*Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*
*Research Institute for Science and Engineering, Waseda University, Tokyo 169-8555, Japan*
*h.okawa@aoni.waseda.jp*

*Keywords*: Collisional Penrose process; energy extraction; spinning test particle.

## 1. Introduction

It is known that the center of mass energy can take arbitrarily large value if a collision occurs near the event horizon of an extreme rotating black hole[1]. However, it does not mean that an ejected particle will have very high energy at infinity because of red-shift. Hence, how large energy we will observe at infinity has been evaluated for various situations[2]. Most of those results were obtained for a point test particle. But a matter around a black hole normally has rotation. Since the equation of motion of a spinning particle is totally different from the geodesic equation, and the effect of spin is nontrivial, it is important to study the efficiency of the extracted energy ($\eta =$ (extracted energy)/(input energy)) in the collision of spinning particles. We focus on the collisional Penrose process of spinning particles near the horizon of an extreme Kerr black hole and evaluate the maximal value of the energy efficiency for various processes.

## 2. The orbit of a spinning particle

In the Kerr spacetime, there are two Killing vectors, and then we have two conserved quantities; the particle energy $E$ and the particle total angular momentum $J$. For simplicity, we assume that a spinning particle is orbiting on the equatorial plane ($\theta = \pi/2$) and the direction of a particle spin is parallel to the rotation of a black hole. The spin parameter $s$ is positive when a particle is parallelly spinning to the rotation of the black hole, while $s$ is negative when it is counter-rotating.

By this setting, we obtain the specific momentum $u^{(a)} = p^{(a)}/\mu$, where $p^{(a)}$ is the momentum and $\mu$ is the particle mass, as

$$u^{(0)} = \frac{\left[(r^3 + a(a+s)r + aMs)E - (ar + Ms)J\right]}{\mu r^2\sqrt{\Delta}\left(1 - \frac{Ms^2}{r^3}\right)}, \tag{1}$$

$$u^{(3)} = \frac{[J - (a+s)E]}{\mu r\left(1 - \frac{Ms^2}{r^3}\right)}, \tag{2}$$

where $M$ and $a$ are the black hole mass and its specific angular momentum, respectively. The suffix $(a)$ denotes the tetrad components. From the definition of the particle mass, we obtain the radial component $u^{(1)}$ as

$$u^{(1)} = \sigma\sqrt{(u^{(0)})^2 - (u^{(3)})^2 - 1}, \tag{3}$$

where $\sigma = \pm 1$ denotes the direction of the radial motion of the particle, i.e., a particle is moving outward for $\sigma = +1$ while it does inward for $\sigma = -1$.

Before the discussion of the collisional Penrose process, we should note a few important points as follows:
(1) Since the four-velocity of a spinning particle $v^{(a)} = dx^{(a)}/d\tau$ is not always parallel to the four-momentum $u^{(a)}$, we have to impose the timelike condition $v^{(a)}v_{(a)} < 0$ in our analysis.
(2) In order for a particle to reach the horizon, the condition $J_- \leq J \leq J_{cr}$ must be satisfied. Here, $J_{cr}(> 0)$ and $J_-(< 0)$ are the maximum and minimum values of the particle angular momentum, respectively. This condition is obtained from $(u^{(1)})^2 > 0$ for any radius of $r \geq r_H = M$.
(3) To obtain arbitrarily large center-of-mass energy, the collision must take a place near the horizon and one of collisional particles must have the critical angular momentum $J_{cr} = 2EM$, and the other particle must have sub-critical angular momentum $(J_- <)J < J_{cr}$.

## 3. Collisional Penrose process

We assume that two particles (the particle 1 [critical, $J_1 = 2E_1M$] and the particle 2 [sub-critical, $J_2 < J_{cr}$]) plunge from infinity, and collide at the point $r_c = r_H/(1-\epsilon)$ $(0 < \epsilon \ll 1)$, where $r_H = M$ is the horizon radius. After collision, the particle 3 (= the particle 1) is coming back to infinity, while the particle 4 (= the particle 2) will fall into a black hole.

At the collision, we have the following onservation laws:

$$E_1 + E_2 = E_3 + E_4, \quad J_1 + J_2 = J_3 + J_4, \tag{4}$$

$$s_1 + s_2 = s_3 + s_4, \quad p_1^{(1)} + p_2^{(1)} = p_3^{(1)} + p_4^{(1)}. \tag{5}$$

From these conservation laws, we find that particle 3 must have near-critical angular momentum $(J_3 = 2E_3M + O(\epsilon))$. We consider only the head-on collision with $\sigma_1 = \sigma_3 = +1$ $\sigma_2 = \sigma_4 = -1$, $s_2 = s_4$, and $s_1 = s_3$, because it gives the maximum efficiency.

We expand the above conservation equations in terms of $\epsilon$ and solve the energies $E_2$ and $E_3$ by use of the expansion coefficients. We then evaluate the energy efficiency $\eta = E_3/(E_1 + E_2)$. The detailed analysis is given in Ref. 4.

In Table 1, we show the maximal efficiencies and the input and output energies for the following three cases : [1] Collision of two massive particles, [2] "Compton" scattering (the particle 1: massless, the particle 2: massive), and [3] "Inverse Compton" scattering (the particle 1: massive, the particle 2: massless).

In the case of the collision of two massive particles, the maximal efficiency $\eta$ is about 15.01, while it is about 26.85 for the case of the "Compton" scattering in the limit of $E_1 \to \infty$. The maximal efficiency becomes the largest for the "Compton" scattering, which in fact is the same as the case for without spin. Compared with the spinless case, these maximal efficiencies become twice larger than the spinless case. We then conclude that a spin plays an important role in the context of energy extraction. Note that the efficiency does not change significantly in the case of the "inverse Compton" scattering because the absorbed massless particle is spinless and the escaped massive particle with large energy cannot have a large spin.

Table 1. The maximal efficiencies and energies for three cases of collisions of particles plunging from infinity. We include the result for the nonspinning case (Ref. 3) for comparison. The maximal efficiencies and maximal energies are enhanced when the spin effect is taken into account.

| collisional process | spin $(s_1, s_2)$ | input energy $(E_1, E_2)$ | output energy $(E_3)$ | maximal efficiency |
|---|---|---|---|---|
| Collision of two massive particles | non-spinning | $(\mu, \mu)$ | $12.66\mu$ | 6.328 |
| | $(0.01379\mu M, -0.2709\mu M)$ | | $30.02\mu$ | 15.01 |
| "Compton" scattering | non-spinning | $(+\infty, \mu)$ | $+\infty$ | 13.93 |
| | $(0, -0.2709\mu M)$ | | $+\infty$ | 26.85 |
| "Inverse Compton" scattering | non-spinning | $(\mu, 0)$ | $12.66\mu$ | 12.66 |
| | $(0.02679\mu M, 0)$ | | $15.64\mu$ | 15.64 |

The present setting may be very unlikely in more realistic astrophysical situations. Since there are many particles in an accretion disc around a black hole, it may be more natural for a particle plunging from infinity will collide particles in the innermost stable circular orbit (ISCO). The result in this case will be published elsewhere[5]. The ansatz of an extreme black hole is also not natural. The study for a non-extreme black hole is in progress.

## References


1. M. Banados, J. Silk and S. M. West, Phys. Rev. Lett. 103, 111102 (2009).
2. M. Bejger, T. Piran, M. Abramowicz and F. Hakanson, Phys. Rev. Lett. 109, 121101 (2012); T. Harada, H. Nemoto and U. Miyamoto, Phys. Rev. D86, 024027 (2012), [Erratum: Phys. Rev. D86, 069902 (2012)]; J. D. Schnittman, Phys. Rev. Lett. 113, 261102 (2014); K. Ogasawara, T. Harada and U. Miyamoto, Phys. Rev. D93, 044054 (2016); O. B. Zaslavskii, Phys. Rev. D94, 064048 (2016).
3. E. Leiderschneider and T. Piran, Phys. Rev. D93, 043015 (2016).
4. K. Maeda, K. Okabayashi and H. Okawa, Phys. Rev. D 98, 064027 (2018).
5. K. Okabayashi and K. Maeda, in preparation.

# Bound orbit domains in the phase space of the Kerr geometry

Prerna Rana[1,†] and A. Mangalam[1,‡]

[1] *Indian Institute of Astrophysics, Sarjapur Road, 2nd Block Koramangala, Bangalore, 560034, India*

*E-mail: prernarana@iiap.res.in*[†], *mangalam@iiap.res.in*[‡]

We derive the conditions for a non-equatorial eccentric bound orbit to exist around a Kerr black hole in two-parameter spaces: the energy, angular momentum of the test particle, spin of the black hole, and Carter's constant space ($E$, $L$, $a$, $Q$), and eccentricity, inverse-latus rectum space ($e$, $\mu$, $a$, $Q$). These conditions distribute various kinds of bound orbits in different regions of the ($E$, $L$) and ($e$, $\mu$) planes, depending on which pair of roots of the effective potential forms a bound orbit. We provide a prescription to select these parameters for bound orbits, which are useful inputs to study bound trajectory evolution in various astrophysical applications like simulations of gravitational wave emission from extreme-mass ratio inspirals, relativistic precession around black holes, and the study of gyroscope precession as a test of general relativity.

*Keywords*: Classical black holes; Relativity and Gravitation; Bound orbit trajectories.

## 1. Introduction

Bound trajectories in the Kerr geometry have been studied extensively, and some of the important results are discussed in a pioneering work by S. Chandrasekhar (Ref. 1). The general trajectory in the Kerr spacetime was first expressed in terms of quadratures in Ref. 2, while Ref. 3 discusses the essential conditions for bound spherical geodesics, and also horizon-skimming orbits. The quadrature form of the fundamental orbital frequencies for a general eccentric trajectory was first presented in Ref. 4. Later, to decouple the $r$ and $\theta$ motions, a parameter called Mino time, $\lambda$, was introduced in Ref. 5, which was then implemented to calculate a closed-form trajectory solution and orbital frequencies in Ref. 6. Recently, an alternate analytic solution was derived for the general bound and separatrix trajectories in a compact form using the transformation $1/r = \mu\,(1 + e\cos\chi)$ in Refs. (7, 8). The inputs to these integrals for calculating the trajectories are the constants of motion $E$, $L$, $Q$, and spin of the black hole, $a$. These parameters can also be translated to ($e$, $\mu$, $a$, $Q$) space, as derived in Refs. (7, 8). It is essential to find the canonical bound orbit conditions in these two parameter spaces to calculate the trajectory evolution.

We express the bound orbit conditions on ($E$, $L$, $a$, $Q$) parameters for the non-equatorial eccentric bound orbits around a Kerr black hole, and then find the analog of these conditions in the ($e$, $\mu$, $a$, $Q$) space. The regions of different bound orbits were graphically separated in the ($E$, $L$) plane in Ref. 9, according to the pair of roots of the effective potential spanning the radius of the bound orbit. It is essential to find the canonical bound orbit conditions in these two parameter spaces.

## 2. Conditions for bound trajectories around Kerr black hole

Now, we consider the radial motion of the bound trajectory which is described by the equation (Refs. 2, 4, 7, 8)

$$\frac{\left(E^2-1\right)}{2}=\frac{\rho^4}{2r^4}\left(\frac{\mathrm{d}r}{\mathrm{d}\tau}\right)^2-\frac{1}{r}+\frac{L^2-a^2\left(E^2-1\right)+Q}{2r^2}-\frac{\left(L-aE\right)^2+Q}{r^3}+\frac{a^2Q}{2r^4}, \quad (1)$$

to derive the conditions in $(E, L, a, Q)$ and $(e, \mu, a, Q)$ spaces for various bound orbit regions previously discussed in Ref. 9, where $\rho^2 = r^2 + a^2\cos^2\theta$, $\tau$ is the proper time, and $a = J/M^2$ which have their usual meanings; we use geometrical units throughout.

### 2.1. *Dynamical parameter space $(E, L, a, Q)$*

Next, we solve the quartic equation, Eq. (1), to find the four roots (that include the turning points of the bound orbit) of $\mathrm{d}r/\mathrm{d}\tau = 0$, which can be expressed as

$$r^4+\frac{2}{\left(E^2-1\right)}r^3+\frac{\left(a^2E^2-L^2-Q-a^2\right)}{\left(E^2-1\right)}r^2+\frac{2\left(x^2+Q\right)}{\left(E^2-1\right)}r-\frac{a^2Q}{\left(E^2-1\right)}=0, \quad (2)$$

where $x = L - aE$. Applying Ferrari's method (Ref. 10) to the above equation gives

$$r_1=\frac{1}{2\left(1-E^2\right)}+\frac{\sqrt{2z}}{2}+\frac{1}{2}\sqrt{D_1}, \quad (3a)$$

$$r_2=\frac{1}{2\left(1-E^2\right)}+\frac{\sqrt{2z}}{2}-\frac{1}{2}\sqrt{D_1}, \quad (3b)$$

$$r_3=\frac{1}{2\left(1-E^2\right)}-\frac{\sqrt{2z}}{2}+\frac{1}{2}\sqrt{D_2}, \quad (3c)$$

$$r_4=\frac{1}{2\left(1-E^2\right)}-\frac{\sqrt{2z}}{2}-\frac{1}{2}\sqrt{D_2}, \quad (3d)$$

where $r_1 > r_2 > r_3 > r_4$, and

$$D_1=-2G-2z-\frac{\sqrt{2}H}{\sqrt{z}}, \quad D_2=-2G-2z+\frac{\sqrt{2}H}{\sqrt{z}}, \quad (3e)$$

$$z=U+V-\frac{G}{3}, \quad U=\left(I+\sqrt{I^2+J^3}\right)^{1/3}, \quad V=\left(I-\sqrt{I^2+J^3}\right)^{1/3}, \quad (3f)$$

$$I=\frac{\left(2G^3+27H^2-72GT\right)}{432}, \quad J=-\frac{\left(G^2+12T\right)}{36}, \quad (3g)$$

$$G=\frac{\left[L^2-a^2\left(E^2-1\right)+Q\right]}{\left(1-E^2\right)}-\frac{3}{2\left(1-E^2\right)^2}, \quad (3h)$$

$$H=\frac{\left[L^2-a^2\left(E^2-1\right)+Q\right]}{\left(1-E^2\right)^2}-\frac{2\left(x^2+Q\right)}{\left(1-E^2\right)}-\frac{1}{\left(1-E^2\right)^3}, \quad (3i)$$

$$T=\frac{\left[L^2-a^2\left(E^2-1\right)+Q\right]}{4\left(1-E^2\right)^3}-\frac{3}{16\left(1-E^2\right)^4}-\frac{\left(x^2+Q\right)}{\left(1-E^2\right)^2}+\frac{a^2Q}{\left(1-E^2\right)}. \quad (3j)$$

The bound orbit regions were graphically classified in the $(E, L)$ plane in Ref. 9 on the basis of which pair of roots of quartic equation, Eq. (2), contains the bound orbit. We algebraically classify these regions using the expressions of roots, Eqs. (3a-3d), in the $(E, L, a, Q)$ parameter space as follows (Ref. 11):

(1) **Region $\Delta$**: Bound orbits exist between $(r_1, r_2)$, and between $(r_3, r_4)$: $D_1 > 0$, $D_2 > 0$, and $E < 1$.
(2) **Region $\varsigma$**: Bound orbit either exists between $r_1$ and $r_2$ if $(r_3, r_4)$ forms a complex pair i.e. $D_2 < 0$ or exists between $r_3$ and $r_4$ if $(r_1, r_2)$ forms a complex pair i.e. $D_1 < 0$: $(D_1 \cdot D_2) < 0$. This region exists for both $E < 1$ and $E > 1$.
(3) **Region $\Lambda$**: Bound orbit exists between $r_2$ and $r_3$ with $r_4 < r_3$ and $r_1 > r_2$: $D_1 > 0$, $D_2 > 0$, and $E > 1$.

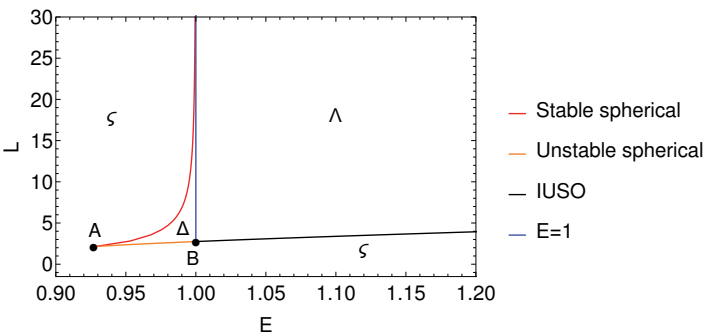



Fig. 1. The bound orbit regions $\Delta$, $\varsigma$, and $\Lambda$ are shown in the $(E, L)$ plane for $a = 0.5$, $Q = 5$, where the region $\Delta$ is bounded by stable and unstable (also corresponds to the separatrices) spherical orbits, and $E = 1$; the $\Lambda$ region is bounded by the inner unstable spherical orbits (IUSO). Points A and B represent the innermost stable spherical orbit (ISSO) and marginally stable spherical orbit (MBSO) respectively.

The classification of these regions in the $(E, L)$ plane is shown in Fig. 1. The bounding curves of these regions represent spherical orbits. The eccentricity and inverse latus-rectum of the bound orbit are defined as (Ref. 11)

$$e_{ij} = \frac{r_i - r_j}{r_i + r_j}, \quad \mu_{ij} = \frac{r_i + r_j}{2 r_i r_j}, \tag{4}$$

where we see that $\{e_{ij}, \mu_{ij}\}$ can be expressed in terms of $(E, L, a, Q)$ through roots, Eqs. (3). The details of derivations presented here will be provided in Ref. 11.

### 2.2. *Conic parameter space $(e, \mu, a, Q)$*

According to the definitions given for regions $\Delta$, $\varsigma$, and $\Lambda$ in §2.1, the convention adopted for $\{e, \mu\}$ is as follows: Region $\Delta$ given by $\{e_{12}, \mu_{12}\}$; Region $\Lambda$ given by $\{e_{23}, \mu_{23}\}$; Region $\varsigma$ given by $\{e_{12}, \mu_{12}\}$ or $\{e_{34}, \mu_{34}\}$, depending on which pair is

real. Now, we derive the defining conditions for $\Delta$, $\varsigma$, and $\Lambda$ regions in the ($e$, $\mu$, $a$, $Q$) space:

(i) **Region $\Delta$**: The defining conditions for this region was derived using the necessary constraints on the parameters of the Elliptic integrals involved in the trajectory solutions (Refs. 7, 8), which are

$$\left[\mu^3 a^2 Q\left(1+e\right)^2+\mu^2\left(\mu a^2 Q-x^2-Q\right)\left(3-e\right)\left(1+e\right)+1\right]>0, \tag{5a}$$

$$\left[\mu\left(1+e\right)\left(1+\sqrt{1-a^2}\right)\right]<1, \tag{5b}$$

$$E(e,\mu,a,Q)<1. \tag{5c}$$

(ii) **Region $\varsigma$**: The region $\varsigma$ is defined by two complex roots of $r$ or $u=1/r$ with a bound orbit existing between the two remaining real roots. We can write Eq. (2) for region $\varsigma$ in the form (Ref. 11)

$$\left[u-\mu\left(1-e\right)\right]\cdot\left[u-\mu\left(1+e\right)\right]\cdot\left(u^2+Au+B\right)=0, \tag{6a}$$

where bound orbit exists between $\mu\left(1+e\right)$ and $\mu\left(1-e\right)$ which is a real pair of the roots, and

$$A=-2\left[\frac{x^2+Q}{a^2Q}-\mu\right],\quad B=\frac{1-E^2}{\mu^2a^2Q\left(1-e^2\right)}. \tag{6b}$$

Hence, the remaining factor $\left(u^2+Au+B\right)$ of Eq. (6a) should have complex roots for the $\varsigma$ region, which reduces to the condition (Ref. 11)

$$\left\{\left(Q+x^2\right)^2\mu+a^4e^2Q^2\mu^3-a^2Q\left[1+\left(1+e^2\right)\left(Q+x^2\right)\mu^2\right]\right\}<0, \tag{7}$$

where we have substituted for the factor $(1-E^2)$ in terms of ($e$, $\mu$, $x$) using relations previously derived in Refs. (7, 8).

(iii) **Region $\Lambda$**: The region $\Lambda$ is defined by the condition that a bound orbit exists between $r_2$ and $r_3$ (or $u_2$ and $u_3$) with $r_1>r_2$ (or $u_1<u_2$) and $r_3>r_4$ (or $u_3<u_4$). We can express Eq. (2) for this region as

$$\left[u-\mu\left(1-e\right)\right]\cdot\left[u-\mu\left(1+e\right)\right]\cdot\left(u^2+Au+B\right)=0. \tag{8}$$

The remaining roots $u_1$ and $u_4$ can be derived from the factor $\left(u^2+Au+B\right)$, which can be substituted into the conditions $u_1<u_2$ and $u_3<u_4$ to obtain

$$\left[\mu^3 a^2 Q\left(1-e\right)^2+\mu^2\left(\mu a^2 Q-x^2-Q\right)\left(3+e\right)\left(1-e\right)+1\right]<0, \tag{9a}$$

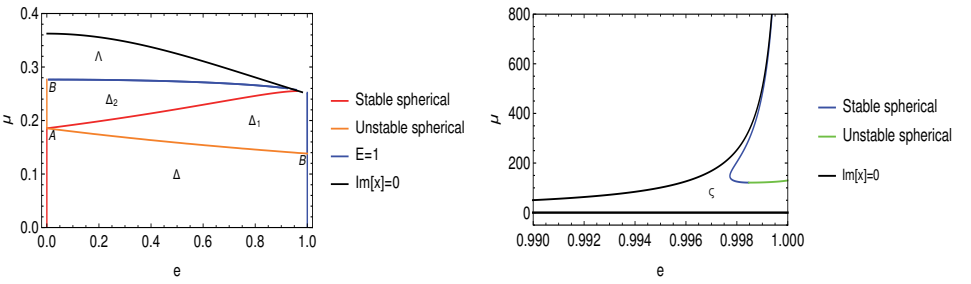
and

$$\left[\mu^3 a^2 Q\left(1+e\right)^2+\mu^2\left(\mu a^2 Q-x^2-Q\right)\left(3-e\right)\left(1+e\right)+1\right]<0, \tag{9b}$$

respectively. Next, as we see from Fig. 1, that the region $\Lambda$ corresponds to the orbits with $E>1$; this implies

$$\left[\mu^2\left(1-e^2\right)\left(\mu a^2 Q-Q-x^2\right)+1\right]<0. \tag{9c}$$

In effect, Eqs. (9a, 9b, 9c) together define $\Lambda$ region the ($e$, $\mu$) plane, and Fig. 2 shows all these regions in the ($e$, $\mu$) plane. The details of derivations of the above conditions will be provided in Ref. 11.




(a) (b)

Fig. 2. The bound orbit regions (a) $\Delta$ and $\Lambda$, and (b) $\varsigma$ in the $(e,\ \mu)$ plane for $(a = 0.2,\ Q = 2)$ are shown along with their defining bounding curves and end points. The regions $\Delta_1$ and $\Delta_2$ are replicas of region $\Delta$ when $(e_{13},\ \mu_{13})$ and $(e_{23},\ \mu_{23})$ are chosen respectively.

## 3. A prescription for selecting bound orbits

We present the scaling formulae specifying the parameters $(E,\ L)$ and $(e,\ \mu)$ in terms of the variables $(x_1,\ y_1)$ and $(x_2,\ y_2)$, where $x_i,\ y_i \in [0, 1]$, $i = 1, 2$, can be chosen to produce valid combinations of the parameters $(E,\ L)$ and $(e,\ \mu)$ for bound orbits. The formula for selecting $E$ for bound orbits written in terms of the variable $x_1$, $a$, and $Q$ for region $\Delta$ in Fig. 1 is (Ref. 11)

$$E\left(x_1, a, Q\right) = E_Z\left(a, Q\right) + x_1\left[E_Y\left(a, Q\right) - E_Z\left(a, Q\right)\right], \tag{10}$$

where $E_Z\left(a, Q\right)$ and $E_Y\left(a, Q\right)$ are the spherical orbit energies at ISSO and MBSO respectively, and where $Z(a, Q)$ and $Y(a, Q)$ are radii of ISSO and MBSO respectively (Eqs. (19), (20) in Refs. 7, 8).

Now, for a fixed $x_1$ and $a$, $y_1 \in [0, 1]$ defines the range of $L\left(r, a, Q\right)$ for bound orbits. The formula for selecting allowed $L$ ($\Delta$ in Fig. 1) can be written as (Ref. 11)

$$\frac{1}{L\left(x_1, y_1, a, Q\right)} = \frac{1}{L_-\left(x_1, a, Q\right)} - y_1\left[\frac{1}{L_-\left(x_1, a, Q\right)} - \frac{1}{L_+\left(x_1, a, Q\right)}\right], \tag{11a}$$

where $L_+\left(x_1, a, Q\right)$ and $L_-\left(x_1, a, Q\right)$ are end points of the $\Delta$ region in Fig. 1 given by (Ref. 11)

$$L_+\left(x_1, a, Q\right) = x(r_v\left(x_1, a, Q\right), a, Q) + a \cdot E(r_v\left(x_1, a, Q\right), a, Q), \tag{11b}$$

$$L_-\left(x_1, a, Q\right) = x(r_u\left(x_1, a, Q\right), a, Q) + a \cdot E(r_u\left(x_1, a, Q\right), a, Q), \tag{11c}$$

where $x\left(r_s, a, Q\right)$ and $E\left(r_s, a, Q\right)$ can be calculated using the spherical orbit formulae (derived in Refs. 7, 8) and where $r_v\left(x_1, a, Q\right)$ and $r_u\left(x_1, a, Q\right)$ are the two roots of $r_s$ in the equation

$$E\left(r_s, a, Q\right) = E_Z\left(a, Q\right) + x_1\left[E_Y\left(a, Q\right) - E_Z\left(a, Q\right)\right]. \tag{12}$$

The radii $r_v\left(x_1, a, Q\right)$ and $r_u\left(x_1, a, Q\right)$ obey $r_v\left(x_1, a, Q\right) > Z(a, Q)$ and $Y\left(a, Q\right) < r_u\left(x_1, a, Q\right) < Z\left(a, Q\right)$.

For the $(e, \mu)$ space, the corresponding formulae for the $\Delta$ region in Fig. 2 are given by

$$e\,(x_2) = \; x_2, \tag{13a}$$

$$\mu\,(y_2, r_x, a, Q) = \; y_2 \cdot \mu_s\,(r_x, a, Q)\,, \tag{13b}$$

where the allowed range of $\mu$ is $0 < \mu < \mu_s\,(r_x, a, Q)$ for a given $x_2$, and $\mu_s\,(r_x, a, Q)$ is the value of $\mu$ at the separatrix [Eq. (25b) in Refs. (7, 8)], and $r_x$ is a root of $r_s$ in the equation $x_2 = e_s\,(r_s, a, Q)$ and $e_s\,(r_s, a, Q)$ is the eccentricity value at the separatrix [Eq. (25a) in Refs. (7, 8)]. The radius $r_x$ lies between $Y\,(a, Q)$ and $Z\,(a, Q)$ for a given $a$ and $Q$. Hence, for a fixed $a$ and $Q$, $x_2$ and $y_2 \in [0, 1]$ which thereby defines the range of $(e, \mu)$.

## 4. Summary and discussion

We presented the algebraic conditions for non-equatorial bound trajectories in the $(E,\ L,\ a,\ Q)$ and $(e,\ \mu,\ a,\ Q)$ spaces and showed how these conditions classify the bound orbits into various regions, $\Delta$, $\varsigma$, and $\Lambda$, in the $(E,\ L)$ plane, which was previously discussed graphically in Ref. 9; see Fig. 1. In this article, we have also shown these bound orbit regions in the $(e,\ \mu)$ plane, Fig. 2, geometrically specified by their bound curves and vertices. For astrophysically relevant orbits, only the region $\Delta$ is applicable. We also provided a useful prescription to select the parameters $(E,\ L)$ and $(e,\ \mu)$ in the $\Delta$ region, which are canonical inputs to the trajectory solutions for studying their evolution in various applications like gravitational wave emission from extreme-mass ratio inspirals, relativistic precession around black holes, and the study of gyroscope precession as a test of general relativity.

We acknowledge the support from the SERB project CRG 2018/003415.

## References

1. S. Chandrasekhar, *The mathematical theory of black holes* 1983.
2. B. Carter, Global Structure of the Kerr Family of Gravitational Fields, *Physical Review* **174**, 1559 (October 1968).
3. D. C. Wilkins, Bound Geodesics in the Kerr Metric, *Phys. Rev. D* **5**, 814 (February 1972).
4. W. Schmidt, Celestial mechanics in Kerr spacetime, *Classical and Quantum Gravity* **19**, 2743 (May 2002).
5. Y. Mino, Perturbative approach to an orbital evolution around a supermassive black hole, *Phys. Rev. D* **67**, p. 084027 (April 2003).
6. R. Fujita and W. Hikida, Analytical solutions of bound timelike geodesic orbits in Kerr spacetime, *Classical and Quantum Gravity* **26**, p. 135002 (July 2009).
7. P. Rana and A. Mangalam, Astrophysically relevant bound trajectories around a Kerr black hole, *Classical and Quantum Gravity* **36**, p. 045009 (February 2019).

8. P. Rana and A. Mangalam, Astrophysically relevant bound trajectories around a Kerr black hole, *arXiv e-prints* , p. arXiv:1901.02730 (January 2019).
9. E. Hackmann, C. Lämmerzahl, V. Kagramanova and J. Kunz, Analytical solution of the geodesic equation in Kerr-(anti-) de Sitter space-times, *Phys. Rev. D* **81**, p. 044020 (February 2010).
10. G. . Cardano, *The great art; or, The rules of algebra. Translated and edited by T. Richard Witmer. With a foreword by Oystein Ore* (M.I.T. Press Cambridge, Mass, 1968).
11. P. Rana and A. Mangalam, (*in preparation,* 2020).

# Schwarzschild phase without a black hole

Sandipan Sengupta*

*Department of Physics and Centre for Theoretical Studies,*
*Indian Institute of Technology Kharagpur, Kharagpur-721302, India*
**E-mail: sandipan@phy.iitkgp.ac.in*

We present a smooth extension of the Schwarzschild exterior geometry, where the singular interior is superceded by a vacuum phase with vanishing metric determinant. Unlike the Kruskal-Szekeres continuation, this solution to the first order field equations in vacuum has no singularity in the curvature two-form fields, no horizon and no global time. The underlying non-analytic structure provides a distinct geometric realization of 'mass' in classical gravity. We also find that the negative mass Schwarzschild solution does not admit a similar extension within the first order theory. This is consistent with the general expectation that degenerate metric solutions associated with the Hilbert-Palatini Lagrangian formulation should satisfy the energy conditions.

*Keywords*: Curvature singularity, Schwarzschild, Degenerate metric, Naked singularity.

## 1. Introduction

In view of the imposing experimental success of Einstein's theory of general relativity, one could be tempted to accept the invertible metric phase ($\det g_{\mu\nu} \neq 0$) as a self-contained and complete framework in describing the classical dynamics of spacetimes. However, in general, (the first order formulation of) gravity theory is known to exhibit an additional phase based on non-invertible metrics[1–3]. From a general perspective, the Einsteinian theory is nothing more or less than the special phase of first order theory where the metric is invertible everywhere. It then seems natural to ask as to how certain robust features of Einsteinian solutions, such as singularities, get manifested in a more generic solution where both the phases could coexist.

In the spherically symmetric vacuum Einstein theory, the Schwarzschild spacetime happens to be the unique solution. For a positive (negative) mass, its singular interior represents a black hole with a horizon (a naked singularity without a horizon). However, in the presence of both the phases of first order gravity, there is no reason for Birkhoff's uniqueness theorem to be applicable, as it works only under the assumption of the invertibility of metric. Here we provide an explicit realization of such a scenario[4,5]. The solutions discussed are smooth extensions of the Schwarzschild exterior through a noninvertible metric phase, which supercedes the interior of the standard Schwarzschild solution of Einstein equations. These spacetimes are solutions to the first order equations of motion everywhere. The associated field-strength components are manifestly finite everywhere, unlike in the standard Schwarzschild case. Even though it is not possible to construct four dimensional curvature scalars in regions where the metric is degenerate, it is still possible to define effective lower-dimensional curvature scalars associated with the nondegenerate subspace of the four geometry in such regions. All such scalars are found to be finite, implying that the emergent three-geometry is regular.

Let us now elucidate our construction for the positive mass Schwarzschild geometry, followed by a brief discussion of the negative mass case.

## 2. A smooth extension of the Schwarzschild exterior

Our aim is to construct a smooth continuation of the Schwarzschild exterior, such that it satisfies the vacuum field equations of first order gravity[a]. These equations of motion are obtained through a variation of the Hilbert-Palatini Lagrangian density with respect to the tetrad and connection fields:

$$e^{[K}_{[\mu}D_{\nu}(\omega)e^{L]}_{\alpha]} = 0,\ e^{[J}_{[\nu}R_{\alpha\beta]}{}^{KL]}(\omega) = 0\ . \tag{1}$$

Although degenerate solutions to these equations may be associated with nontrivial torsion in general, we shall consider torsionless configurations which would be sufficient for our purpose here.

### 2.1. *Basic fields*

We define the two phases of the full spacetime through the following metric ($t \in (-\infty, \infty),\ u \in (-\infty, \infty),\ \theta \in [0, \pi],\ \phi \in [0, 2\pi]$):

$$\begin{aligned} ds^2 &= -\left[1 - \frac{2M}{f(u)}\right]dt^2 + \frac{f'^2(u)}{\left[1 - \frac{2M}{f(u)}\right]}du^2 + f^2(u)\left[d\theta^2 + \sin^2\theta d\phi^2\right]\ \ (u > u_0), \\ &= 0 + \sigma F^2(u)du^2 + H^2(u)\left[d\theta^2 + \sin^2\theta d\phi^2\right]\ \ (u \leq u_0). \end{aligned} \tag{2}$$

The smooth functions $f, F$ satisfy the following set of boundary conditions:

$$\begin{aligned} &f(u) \to 2M,\ f'(u) \to 0,\ F(u) \to 0 \text{ as } u \to u_0; \\ &f(u) \to \infty,\ f'(u) \to 1 \text{ as } u \to \infty. \end{aligned} \tag{3}$$

While the first line above ensures the continuity of the metric, the last line implies that the spacetime is flat as $u \to \infty$. The internal metric is given by $\eta_{IJ} = diag[-1, 1, 1, 1]$. The metric at $u > u_0$ may be brought to the Schwarzschild form through a reparametrization $u \to r = f(u)$. However, the degenerate metric at $u \leq u_0$ has no semblance to the Schwarzschild interior. Also, $u_0$ in this construction is not a new free parameter, but rather is dependent on the 'mass' parameter M. Although it is not necessary to adopt any specific $f(u)$, we choose it to be the following in order to be explicit:

$$f(u) = 2M\left[1 + \left|\frac{u}{u_0} - 1\right| e^{-\frac{u_0^2}{(u-u_0)^2}}\right].$$

This satisfies all the boundary conditions in (3) provided $u_0 = 2M$.

[a]The form of the metric defining the full spacetime here is the same as introduced in Ref. 5. However, the solutions there are based on fields that are continuous but not smooth and have nonvanishing torsion, unlike the case here.

The nonvanishing components of the associated (torsionless) spin-connection and the resulting field-strength are given below:

$$\omega_t{}^{01} = \frac{M}{f^2(u)},\ \omega_\phi{}^{23} = -\cos\theta,\ \omega_\phi{}^{31} = \left(1-\frac{2M}{f(u)}\right)^{\frac{1}{2}}\sin\theta = -\omega_\theta{}^{12}\sin\theta;$$
$$R_{tu}^{01} = \frac{2Mf'(u)}{f^3(u)},\ R_{t\phi}^{03} = -\frac{M}{f^2(u)}\left(1-\frac{2M}{f(u)}\right)^{\frac{1}{2}}\sin\theta = R_{t\theta}^{02}\sin\theta,$$
$$R_{\theta\phi}^{23} = \frac{2M}{f(u)}\sin\theta,\ R_{\phi u}^{31} = -\frac{Mf'(u)}{f^2(u)}\left(1-\frac{2M}{f(u)}\right)^{-\frac{1}{2}}\sin\theta = R_{u\theta}^{12}\sin\theta. \qquad (4)$$

At the region $u \le u_0$ with a degenerate phase, we choose a zero-torsion configuration, defined by the following connection fields along with the associated field-strength:

$$\hat{\omega}_\phi{}^{23} = -\cos\theta,\ \hat{\omega}_\phi{}^{31} = \frac{H'(u)}{\sqrt{\sigma}F(u)}\sin\theta = -\hat{\omega}_\theta{}^{12}\sin\theta;$$
$$\hat{R}_{\theta\phi}^{23} = \left[1-\sigma\left(\frac{H'(u)}{F(u)}\right)^2\right]\sin\theta,\ \hat{R}_{\phi u}^{31} = -\left[\frac{H'(u)}{\sqrt{\sigma}F(u)}\right]'\sin\theta = \hat{R}_{u\theta}^{12}\sin\theta, \qquad (5)$$

where we have displayed only the nonvanishing components. Note that even though some of the field components are imaginary for $\sigma = -1$, the physical fields, given by their $SO(3,1)$ gauge-invariant counterparts $\hat{R}_{\mu\nu\alpha\beta} = \hat{R}_{\mu\nu}{}^{IJ}\hat{e}_{\mu I}\hat{e}_{\nu J}$, are all real.

### 2.2. *Solving the field equations*

Since this configuration with degenerate tetrads have vanishing torsion by construction, the first of the set of equations of motion (1) is already satisfied. The remaining equation involving the curvature two-form is also satisfied provided the fields obey the constraint:

$$\left(\frac{H'^2(u)}{F^2(u)} - \sigma\right)F(u) + 2H(u)\left(\frac{H'(u)}{F(u)}\right)' = 0. \qquad (6)$$

A solution to the continuity conditions at $u = u_0$ and the constraint (6) is given by:

$$F(u) = -\frac{f'(u)}{\sqrt{\sigma}\left(1-\frac{2M}{f(u)}\right)^{\frac{1}{2}}},\ H(u) = f(u). \qquad (7)$$

Note that all the gauge covariant fields (tetrad and field-strength) are smooth across the phase boundary. The apparent discontinuity in the connection field ($\omega_t{}^{01} \neq \hat{\omega}_t{}^{01}$ at $u = u_0$) could be gauged away by a boost of the form:

$$\Lambda^I{}_J = \begin{pmatrix} \cosh\left[\frac{t}{4M}\right] & \sinh\left[\frac{t}{4M}\right] & 0 & 0 \\ \sinh\left[\frac{t}{4M}\right] & \cosh\left[\frac{t}{4M}\right] & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

### 2.3. *Distinct features*

We now summarize the main features of these new solutions of first order field equations:

a) The spacetime metric, along with all the gauge-covariant fields in this solution, are smooth everywhere.

b) This (nonanalytic) extension has no horizon; The two values $\sigma = \mp 1$ correspond to a Lorentzian and Euclidean three geometry which supercedes the standard Schwarzschild interior.

c) The curvature two-form fields are finite everywhere, in contrast to the case of a Schwarzschild interior.

d) The solution has a free parameter 'M', exactly as in the usual Schwarzschild case. However, there is no matter sourcing this mass, rather, its origin is purely geometric.

e) There is no horizon; Rather, the two-sphere at $u = u_0$ characterizes a minimal area surface $A_{min} = 16\pi M^2$ and is impenetrable, at least classically.

f) The method of defining the degenerate phase through $\hat{g}_{tt} = 0$ is unique, since it is not possible to obtain a nontrivial extension of the Schwarzschild exterior through any other kind of degeneracy (e.g. $\hat{g}_{uu} = 0$).

Finally, note that even though the four-curvature scalar polynomials cannot be defined everywhere due to the noninvertibility of tetrads, the spacetime at $u \leq u_0$ may be treated as an emergent three-geometry defined solely by the induced metric:

$$ds^2 = \sigma F^2(u) du^2 + H^2(u) \left[ d\theta^2 + \sin^2\theta d\phi^2 \right].$$

The three-scalar curvature polynomials built upon the torsionless connection for this three-geometry also turn out to be finite everywhere.

## 3. Naked singularity

Next, let us analyze the case of negative mass Schwarzschild geometry, the corresponding metric being obtained through a sign reversal $M \to -M$. This represents a naked singularity solution of vacuum Einstein equations.

Since there is no horizon in the original geometry ($g_{tt} \neq 0$), let us consider an extension in first order gravity through $g_{uu} = 0$ at $u \leq u_0$ is as follows:

$$\begin{aligned} ds^2 &= -\left[1 + \frac{2M}{f(u)}\right] dt^2 + \frac{f'^2(u)}{\left[1 + \frac{2M}{f(u)}\right]} du^2 + f^2(u) \left[ d\theta^2 + \sin^2\theta d\phi^2 \right] \ (u > u_0), \\ &= -\left[1 + \frac{2M}{f_0}\right] dt^2 + 0 + f_0^2 \left[ d\theta^2 + \sin^2\theta d\phi^2 \right] \ (u \leq u_0), \end{aligned} \tag{8}$$

where $u = u_0$ corresponds to the phase boundary hypersurface and $f_0 = f(u_0)$. Since the phase boundary must correspond to some finite but nonzero Schwarzschild radius ($r$), $f_0 > 0$. However, the noninvertible phase above just corresponds to the trivial restriction of the original negative mass solution at $u = u_0$ rather than a new

spacetime region. The same conclusion holds for a possible degeneracy through $g_{tt} = 0$ (based on a redefined 'time').

Thus, the singular curvature-two form fields associated with 'negative mass' solution cannot be regularized using a degenerate extension, in contrast to the 'positive mass' case discussed in the previous section.

## 4. Conclusions

The smooth extension of the exterior Schwarzschild geometry, as discussed here, involves a modification of the standard picture of a singular black hole interior in a purely classical setting[5,6]. These solutions could provide a fresh perspective into the information loss problem. Being rooted within a fairly conservative framework based on first order gravity, such an approach could complement (and perhaps supercede) some of the more exotic proposals (e.g. Firewall or Fuzzball programme) that have been put forth to resolve this celebrated paradox.

The cases of positive and negative mass curvature singularities of the Einsteinian theory are perceived very differently within this framework. While the singular interior of positive mass Schwarzschild geometry may be traded for a zero-determinant phase with regular field-strength components, the negative mass naked singularity admits no such smooth extension in first order gravity. This outcome, however, is consistent with the general expectation[7] that degenerate metric ($\det g_{\mu\nu} = 0$) solutions obtained within the Hilbert-Palatini Lagrangian formulation[1–3] should satisfy the energy conditions. The scenario here may be contrasted with the case of degenerate triad ($\det E^a_i = 0$) solutions obtained earlier within the complex SU(2) (Sen-Ashtekar) Hamiltonian framework, which are known to contain negative energy geometries[8].

Finally, let us note that in the vacuum solution constructed here, there is no matter sourcing the 'mass' M. Rather, its origin could be attributed to the time-nonorientability at the phase boundaries[5,6,9,10]. This is an interesting realization of mass through pure geometry. Further, this scenario is distinct from the remarkable ideas explored first by Einstein-Rosen[11] and later by Wheeler-Misner[12], through their respective attempts to generate 'mass' and 'charge' from a nontrivial geometry or topology. To emphasize, our analysis here does not involve wormholes (which are not solutions to the vacuum field equations in general) or quantum configurations such as geons. It would be natural to ask if there is a more interesting manifestation of such geometric 'mass' sourced by noninvertible phases in vacuum.

To conclude, given the number of intriguing features of these 'two-phased' solutions, one wonders if quantum gravity could be a more suitable arena for further explorations based on these.

## Acknowledgments

The support of (in part) the ECR award grant no. ECR/2016/000027 under the SERB, DST, Govt. of India and of (in part) the ISIRD grant 'RAQ' is gratefully acknowledged.

## References


1. A. A. Tseytlin, *J. Phys. A: Math. Gen.* **15**, L105 (1982).
2. R. K. Kaul and S. Sengupta, *Phys. Rev. D* **93**, 084026 (2016).
3. R. K. Kaul and S. Sengupta, *Phys. Rev. D* **94**, 104047 (2016).
4. I. Bengtsson, Class. Quantum Grav. 8 (1991) 1847-1858.
5. R. K. Kaul and S. Sengupta, *Phys. Rev. D* **96**, 104011 (2017).
6. S. Sengupta, *Phys. Rev. D* **96**, 104031 (2017).
7. J. Samuel, *Proc. Conf. of Physics at the Planck Scale (Puri), ed. J. Maharana* (1994).
8. M. Varadarajan, *Class. Quantum Grav.* **8**, L235 (1991).
9. S. Sengupta, *Phys. Rev. D* **97**, 124038 (2018).
10. S. Gera and S. Sengupta, *Phys. Rev. D* **99**, 124038 (2019).
11. A. Einstein and N. Rosen, *Phys. Rev.* **48**, 73 (1935).
12. J. A. Wheeler, *Phys. Rev.* **97**, 511, (1955); C. W. Misner and J. A. Wheeler, *Annals Phys.* **2**, 525 (1957); J. A. Wheeler, *Annals Phys.* **2**, 604 (1957); R. Sorkin, *J. Phys. A* **10**, 717 (1977).

# Scattering of Dirac fermions from spherically symmetric black holes: analytical phase shifts analysis

Ciprian A. Sporea*, Cosmin Crucean and Ion I. Cotaescu

*Faculty of Physics, West University of Timisoara,*
*Timisoara, 300223, Romania*
**E-mail: ciprian.sporea@e-uvt.ro*
*www.physics.uvt.ro*

The scattering of the Dirac fermions by the black-holes with spherical symmetry is studied by applying the method of partial wave analysis. The analytical expressions for the phase shifts and analytical formulas for the differential scattering cross section are written down for the fermion scattering from Schwarzschild, Reissner-Nordsrtöm and Bardeen black holes. A brief comment about the principal features of these cases is outlined.

*Keywords*: Black holes, fermions, scattering.

## 1. Introduction

In this paper we want to present a comparative study of the scattering of Dirac fermions by black-holes with spherical symmetry. For that we choose the Schwarzschild, the Reissner-Nordsrtöm and Bardeen black holes, which were each analysed separately in our previous works[1–4]. Other works that studied numerically the scattering of fermions by these three types of black holes can be found for example in Refs. 5–19.

The main steps for solving the Dirac equation in the geometries that describe black-holes with spherical symmetry are detailed. We obtain the scattering modes resulted solving the Dirac equation in the asymptotic zone of the black hole, which help us to obtain the phase shifts and the analytical expression of the differential scattering cross section. The differential scattering cross section will be analysed in terms of the relevant parameters such as the mass of the black hole and the charge of the black hole and scattering angle. From the analytical expression for the cross section the zero mass limit can be obtained.

We begin in the second section with the basic notions about the Dirac equation in geometries that describe black holes with spherical symmetry. The third section is dedicated to the phase shifts obtained using the partial wave analysis and in the fourth section we make a graphical analysis of the cross section in terms of scattering angle.

## 2. Dirac Equation and Black Holes

In this work we are discussing fermion scattering by black holes with spherical symmetry that have the following generic metric

$$ds^2 = h(r)dt^2 - \frac{dr^2}{h(r)} - r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right) \tag{1}$$

In the case of Schwarzschild, electrically charged Reissner-Nordsrtöm and (regular) Bardeen black holes the function $h(r)$ reads:

$$h(r)_S = 1 - \frac{2M}{r}, \qquad h(r)_{RN} = 1 - \frac{2M}{r} + \frac{Q^2}{r^2}, \qquad h(r)_B = 1 - \frac{2Mr^2}{\sqrt{(r^2+Q_m^2)^3}} \quad (2)$$

where $M$ is the mass of each black hole; $Q$ is the total electric charge and $Q_m$ is the nonlinear magnetic monopole charge[20].

It can be shown (for details see our previous papers[1,2,4,23]) that the Dirac equation, $i\gamma^a D_a\psi - m\psi = 0$, in the black holes geometries given by eq. (1) can be reduced to only a radial equation, namely

$$\begin{pmatrix} m\sqrt{h(r)} + V(r) & -h(r)\frac{d}{dr} + \frac{\kappa}{r}\sqrt{h(r)} \\ h(r)\frac{d}{dr} + \frac{\kappa}{r}\sqrt{h(r)} & -m\sqrt{h(r)} + V(r) \end{pmatrix} \begin{pmatrix} f^+_{E,\kappa}(r) \\ f^-_{E,\kappa}(r) \end{pmatrix} = E \begin{pmatrix} f^+_{E,\kappa}(r) \\ f^-_{E,\kappa}(r) \end{pmatrix} \quad (3)$$

where $f^{\pm}(r)$ are two radial wave function and $V(r) = eQ/r$. The angular part of the Dirac equation is contained into the standard 4-component angular spinors $\Phi^{\pm}_{m,\kappa}(\theta,\phi)$[21]. Thus the particle-like solutions with energy $E$ can be expressed as

$$\psi_{E,j,m,\kappa}(t,r,\theta,\phi) = \frac{e^{-iEt}}{r\,h(r)^{1/4}} \left\{ f^+_{E,\kappa}(r)\Phi^+_{m,\kappa}(\theta,\phi) + f^-_{E,\kappa}(r)\Phi^-_{m,\kappa}(\theta,\phi) \right\} \quad (4)$$

The separation of angular variables as in eq. (4) was possible by using the Cartesian gauge (see eqs. (1)–(4) in Ref. 2). In this gauge the Dirac equation is manifestly covariant under rotations.

The radial equation (3) can not be solved analytically for any of the three black holes (2) and up to now a combination of analytical and numerical methods was used to solve it and to compute numerical phase shifts as was done for ex. in Ref. 6 for Schwarzschild BH or in Ref. 19 for Bardeen BH.

### 2.1. *Scattering modes*

In the following we will briefly show how to obtain analytical phase shifts by finding (approximative) analytical scattering modes resulted from solving the Dirac equation in the asymptotic zone of the black hole as done in Refs. 1–4. One starts by changing the variable $r$ to a Novikov-like one, namely

$$x = \sqrt{\frac{z}{r_+} - 1}, \quad (5)$$

with $r_+$ the radius of the outer black hole horizon. For Schwarzschild BH $z = r$ and $r_+ = r_0 = 2M$; for Reissner-Nordsrtöm BH $z = r$ and $r_+ = M + \sqrt{M^2 - Q^2}$; while for Bardeen BH $z = \sqrt{r^2 + Q_M^2}$ and $r_+$ has a more complicated analytical expression not given here.

By rewriting eq. (3) in terms of the new variable $x$ and making a Taylor expansion with respect to $1/x$ one can show, after discarding the $O(1/x^2)$ terms and

higher, that in the asymptotic zone of the black hole the new equation becomes

$$\begin{pmatrix} \frac{1}{2}\frac{d}{dx}+\frac{\kappa}{x} & -x(\mu+\varepsilon)-\frac{1}{x}(\zeta+\beta) \\ x(\varepsilon-\mu)-\frac{1}{x}(\zeta-\beta) & \frac{1}{2}\frac{d}{dx}-\frac{\kappa}{x} \end{pmatrix} \begin{pmatrix} f^{+}_{E,\kappa}(x) \\ f^{-}_{E,\kappa}(x) \end{pmatrix} = 0 \tag{6}$$

where $\mu = r_+ m$, $\varepsilon = r_+ E$ and

- for Schwarzschild black hole one has:

$$\zeta = \frac{\mu}{2}, \quad \beta = \varepsilon, \quad r_+ = 2M \tag{7}$$

- for Reissner-Nordström black hole:

$$\zeta = \frac{1}{2}\,\mu\left(1-\frac{r_-}{r_+}\right), \quad \beta = \varepsilon - eQ, \quad r_\pm = M \pm \sqrt{M^2-Q^2} \tag{8}$$

- for regular Bardeen black hole:

$$\zeta = \mu\left(1-\frac{M}{r_+}\right), \quad \beta = \varepsilon, \tag{9}$$

The scattering mode solutions (with $\varepsilon > \mu$) of eq. (6) can be found in terms of Whittaker functions[1,2,4,23]

$$\begin{pmatrix} f^{+}(x) \\ f^{-}(x) \end{pmatrix} = \begin{pmatrix} -i\sqrt{\mu+\varepsilon} & i\sqrt{\mu+\varepsilon} \\ \sqrt{\varepsilon-\mu} & \sqrt{\epsilon-\mu} \end{pmatrix} \begin{pmatrix} C_1^{+}\,\frac{1}{x}M_{\rho_+,s}(2i\nu x^2) + C_2^{+}\,\frac{1}{x}W_{\rho_+,s}(2i\nu x^2) \\ C_1^{-}\,\frac{1}{x}M_{\rho_-,s}(2i\nu x^2) + C_2^{-}\,\frac{1}{x}W_{\rho_-,s}(2i\nu x^2) \end{pmatrix} \tag{10}$$

were $\nu = \sqrt{\varepsilon^2-\mu^2}$. The four integration constants $C_1^{\pm}$, $C_2^{\pm}$ are not independent, satisfying the following relations[1,23]:

$$\frac{C_1^{-}}{C_1^{+}} = \frac{s-iq}{\kappa-i\lambda}, \quad \frac{C_2^{-}}{C_2^{+}} = -\frac{1}{\kappa-i\lambda} \tag{11}$$

and were the following parameters are used:

$$s = \sqrt{\kappa^2+\zeta^2-\beta^2}, \quad \rho_\pm = \mp\frac{1}{2} - iq, \quad q = \frac{\beta\varepsilon-\zeta\mu}{\nu}, \quad \lambda = \frac{\beta\mu-\zeta\varepsilon}{\nu} \tag{12}$$

### 2.2. *Analytical Phase Shifts*

Applying the partial wave method on the scattering modes (10) as done in Refs. 1,2,4 we obtained for the first time in the literature analytical expressions for the phase shifts associated to the scattering of Dirac fermions by black holes, namely:

$$S_\kappa = e^{2i\delta_\kappa} = \left(\frac{\kappa-i\lambda}{s-iq}\right)\frac{\Gamma(1+s-iq)}{\Gamma(1+s+iq)}e^{i\pi(l-s)}\,. \tag{13}$$

Using these phase shifts one can now calculate analytical expressions for the differential scattering cross section

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2 + |g(\theta)|^2 \tag{14}$$

were the scalar amplitudes $f(\theta)$ and $g(\theta)$ are defined in terms of Legendre polynomials[22] and the phase shifts $e^{2i\delta_l}$

$$f(\theta) = \sum_{l=0}^{\infty} \frac{1}{2ip} \left[(l+1)(e^{2i\delta_{-l-1}} - 1) + l(e^{2i\delta_l} - 1)\right] P_l^0(\cos\theta)$$
$$g(\theta) = \sum_{l=1}^{\infty} \frac{1}{2ip} \left[e^{2i\delta_{-l-1}} - e^{2i\delta_l}\right] P_l^1(\cos\theta) \tag{15}$$

The above infinite series have a singularity present at $\theta = 0$ that can not be removed. This makes the scattering section to be divergent in the forward direction. However, by using a method first proposed in Ref. 24 one can define new reduced series for $f(\theta)$ and $g(\theta)$ (see eqs. (72)–(75) in Ref. 1 for more details) that converge more quickly.

## 3. Main Results and Brief Comments

In this section we perform a graphical analysis of the cross section in terms of the scattering angle. The relevant parameters in our analysis are the black hole mass $M$, the fermion mass $m$, fermion speed $v$, the charge of the black hole $Q$ and the scattering angle $\theta$.

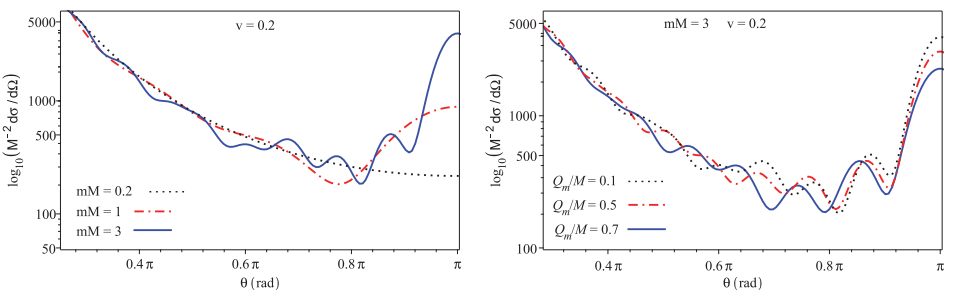



Fig. 1. Differential scattering cross section for fermion scattering by a Schwarzschild black hole (left panel), respectively for a Bardeen black hole (right panel).

In Fig. 1, we make a comparative study of the differential scattering cross section for fermion scattering by a Schwarzschild black hole and a Bardeen black hole. Our graphical results prove that the cross section has a oscillatory behaviour which become more pronounced as we increase the product of masses $mM$ in the case of scattering on Schwarzschild black hole. When the scattering angle approaches the value $\theta = 0$ the cross section increase proving that the forward scattering is dominant. In the case of Bardeen black hole the differential cross section depend on the ratio between the charge of the magnetic monopole and the black hole mass and we observe a variation of the differential cross section when we modify the ratio $Q_m/M$. In both cases the differential cross section have a maximum for $\theta = \pi$, which prove

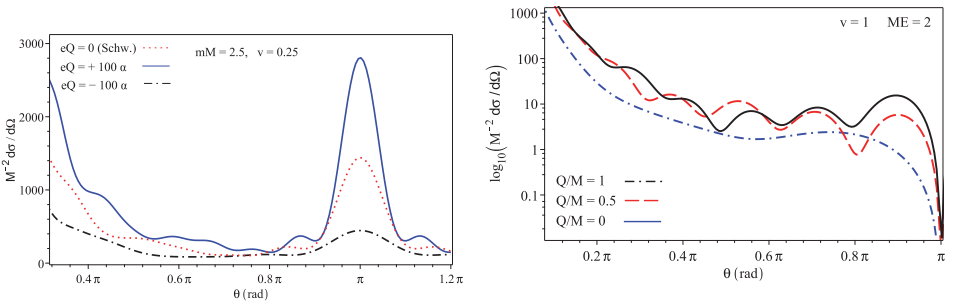



Fig. 2. Differential scattering cross section for fermion scattering by a Reisnner-Nordstrom black hole (left panel). Scattering of massless fermions (right panel).

that the backward scattering is also relevant. In the case of Schwarzschild black hole the backward scattering become important as the parameter $mM$ increase.

The results presented in Fig. 2, shows the dependence of the differential scattering cross section for fermions scattered by a Reisnner-Nordsrtöm black hole on the scattering angle. Here we analyse the case of massive fermions as well as the case when the mass of the fermion is zero. In the case of scattering of massive fermions we observe the differential scattering cross section variation in terms of the charge of the black hole and the charge of the fermion in comparison with the scattering on Schwarzschild black hole (left panel). When the charge of the black hole and the fermion charge have the same sign, the backward scattering become important because of the electrostatic repulsion between charges. For scattering of fermions with zero mass we observe the variation of the differential scattering cross section in terms of the ratio $Q/M$ and that in the backward direction the differential scattering cross section is vanishing.

## 4. Conclusions

The principal conclusion is that the scattering processes analyzed here can be studied by using exclusively analytical methods with suitable asymptotic conditions. In this manner we complete the previous studies which combine analytical and numerical methods obtaining new results and a synthetic overview over the considered processes, pointing out that these are described by similar formulas.

## Acknowledgments

C. A. Sporea was supported by a grant of Ministery of Research and Innovation, CNCS - UEFISCDI, project number PN-III-P1-1.1-PD-2016-0842, within PNCDI III.

## References


1. I. I. Cotăescu, C. Crucean and C. A. Sporea, Eur. Phys. J. C **76**:102 (2016).
2. I. I. Cotăescu, C. Crucean and C. A. Sporea, Eur. Phys. J. C **76**:413 (2016).
3. C. A. Sporea, Chinese Physics C, Vol. 41, No. 12 (2017) 123101.
4. C. A. Sporea, Chinese Physics C, Vol. 43, No. 3 (2019) 035104.
5. W. G. Unruh, Phys. Rev. D 14, 3251 (1976).
6. S. Dolan, C. Doran and A. Lasenby, Phys. Rev. D **74**, 064005 (2006).
7. C. Doran, A. Lasenby, S. Dolan and I. Hinder, Phys. Rev. D 71, 124020 (2005).
8. S. R. Das, G. Gibbons and S. D. Mathur, Phys. Rev. Lett. 78, 417 (1997).
9. W. M. Jin, Classical and Quantum Gravity 15, no. 10, 3163 (1998).
10. C. Doran and A. Lasenby, Phys. Rev, D 66, 024006 (2002).
11. E. Jung, S. H. Kim and D. K. Park, JHEP 09, 005 (2004).
12. L. C. B. Crispino, S. R. Dolan, A. Higuchi and E. S. de Oliveira, Phys. Rev. D 92, 084056 (2015).
13. C. F. B. Macedo and L. C. B. Crispino, The Fourteenth Marcel Grossmann Meeting, pp. 1697–1700 (2017).
14. H. Cho and Y. Lin, Classical and Quantum Gravity 22, no. 5, 775 (2005).
15. J. Jing, Phys. Rev. D 70, 065004 (2004).
16. I. I. Cotaescu and C. A. Sporea, Eur. Phys. J. C **79**:15 (2019).
17. A. B. Gaina and G. A. Chizhov, Moscow Univ. Phys. Bull. 38N2 (1983) 1–7.
18. A. B. Gaina, Moscow VINITI, No. 1970-80 Dep., 20 pp. (1980).
19. H. Huang, M. Jiang, J. Chen and Y. Wang, Gen. Relativ. Gravit. (2015) 47:8.
20. E. Ayon-Beato and A. Garcia, Phys. Lett B 493, 149 (2000).
21. B. Thaller, *The Dirac Equation*. Springer Verlag, Berlin Heidelberg, 1992.
22. V. B. Berestetski, E. M. Lifshitz and L. P. Pitaevski, *Quantum Electrodynamics*. Pergamon Press, Oxford, 1982.
23. I. I. Cotaescu, Mod. Phys. Lett. A 22, 2493 (2007).
24. D. R. Yennie, D. G. Ravenhall and R. N. Wilson, Phys. Rev. **95**, 500 (1954).

# Weak conservation law from hidden symmetry of spinning black holes

Vojtěch Witzany*

*Astronomical Institute of the Czech Academy of Sciences,*
*Boční II 1401, 141 00 Praha 4, Czechia*

*Center of Applied Space Technology and Microgravity (ZARM), University Bremen*
*Am Fallturm 2, 28359 Bremen, Germany*

**E-mail: vojtech.witzany@asu.cas.cz*

The Kerr space-time possesses a "hidden symmetry", which exhibits itself in an unexpected conserved quantity along geodesics known as the Carter constant. Here I report on a recent paper [1], where I answer the question whether this hidden symmetry can be used to formulate a conservation law for a general matter field. First, I show that there cannot exist a conserved sum of Carter constants for a matter system with momentum exchange between its components. Then, I derive a "weak" conservation law associated with the hidden symmetry and demonstrate its properties. Finally, I show how this weak conservation law and related identities can be used to detect violations of the evolution equations in numerical simulations of astrophysical processes near spinning black holes.

*Keywords*: Black holes; conservation laws; symmetries; numerical simulations.

## 1. Introduction

Accreting black holes are now widely believed to reside in many astrophysical systems including X-ray binaries and active galactic nuclei. Even though simple analytical models provide reasonable agreement with the observational appearance of such black holes, the detailed observational signature of their accretion disks are obtained from expensive numerical simulations that involve the non-linear and often turbulent interaction of magnetic fields, non-thermal radiation, and possibly non-thermal plasmas. This all takes place on the stage of the Kerr space-time, the gravitational field of a spinning black hole.

It is very much possible that errors occur in such computer simulations. This may come about either as an inherent error of the numerical method, or as an understandable mistake in the writing of such complex simulation programs. Hence, it is very useful to devise independent checks of the validity of the evolutions obtained with such numerical codes. This can be done either by checking the codes against known exact solutions [2–4], or by checking against conservation laws. For example, one may look for conservation laws that apply to general evolutions of particular matter models but also for conservation laws that apply only in situations with a given symmetry of the matter configuration, such as was done for the case of ideal magnetohydrodynamics [5]. However, both of these approaches will typically test the code only in very specific modes, and there are only two conservation laws that hold for completely general evolutions. These laws are the conservation of energy and azimuthal angular-momentum fluxes, which emerge, in fact, from the stationarity and axisymmetry of the Kerr background.

Nevertheless, the Kerr space-time has an additional geometric property that is often called a "hidden symmetry". The hidden symmetry follows from the existence of a Killing-Yano (KY) tensor, and it leads to conservation laws of an "angular momentum squared", also known as the Carter constant[6], for free test particles as well as linear wave equations (free test quantum particles) in Kerr space-time. It is natural to ask whether these conservation laws form a pattern which can be generalized to *any* test matter field evolving in the Kerr space-time. In other words, is there a conserved Carter-constant flux for general matter systems, even such as those needed to describe accretion disks near black holes? It turns out that the answer is *no* and that this can be easily proven. Nevertheless, there *does* exist a conserved *KY vorticity flux* which has the character of a differential constraint or a "weak conservation law". I briefly summarize these results in this contribution to the Proceedings, while further information can be found in Ref. 1.

## 2. Carter constant and colliding particles

I use the $G = c = 1$ geometrized units and the $-+++$ signature of the metric. Coordinate gradients are denoted by a comma preceding an index, and covariant gradients with a semi-colon preceding an index.

The Kerr metric is independent of time $t$ and azimuthal angle $\varphi$, which makes $\xi^\mu_{(t)} = \delta^\mu_t$ and $\xi^\mu_{(\varphi)} = \delta^\mu_\varphi$ Killing vectors with antisymmetric covariant gradients, $\xi_{\mu;\nu} = -\xi_{\nu;\mu}$. Consequently, geodesics with four-velocity $u^\mu$ have constants of motion $\xi^\mu u_\mu$,

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(\xi^\mu u_\mu) = \xi^\mu u_{\mu;\nu} u^\nu + \xi_{\mu;\nu} u^\mu u^\nu = 0\,. \tag{1}$$

A natural generalization of a Killing vector is a KY tensor, which fulfills $Y_{\mu\nu} = -Y_{\nu\mu}$, $Y_{\mu\nu;\kappa} = -Y_{\mu\kappa;\nu}$. Now a projection of four-velocity into a KY tensor is a vector that is parallel-transported along the geodesic,

$$\frac{\mathrm{D}}{\mathrm{d}\tau}(Y_{\mu\nu}u^\nu) = Y_{\mu\nu;\kappa}u^\nu u^\kappa + Y_{\mu\nu}u^\nu{}_{;\kappa}u^\kappa\,. \tag{2}$$

The square of a vector is conserved, so $C \equiv K_{\nu\lambda}u^\nu u^\lambda$, $K_{\nu\lambda} \equiv Y_{\mu\nu}Y^\mu{}_\lambda$ is an integral of motion of the geodesic.

Now consider a collisionless cloud of free test particles. This will have a stress-energy tensor and conserved current of Carter constants

$$T^{\mu\nu}(x^\lambda) = \sum_i m_{(i)} u^\mu_{(i)} u^\nu_{(i)} \int_{-\infty}^{\infty} \delta^4(x^\lambda - x^\lambda_{(i)}(\tau))\mathrm{d}\tau\,, \tag{3}$$

$$\mathcal{C}^\mu(x^\lambda) = \sum_i C_{(i)} m_{(i)} u^\mu_{(i)} \int_{-\infty}^{\infty} \delta^4(x^\lambda - x^\lambda_{(i)}(\tau))\mathrm{d}\tau\,, \tag{4}$$

where $(i)$ labels quantities associated with the $i$-th particle in the cloud. In other words, $\mathcal{C}^\mu_{\ ;\mu} = 0$ and the integral of $\mathcal{C}^t$ over the spatial extent of an isolated cloud will be conserved and equal to the sum of (mass-weighed) Carter constants at time $t$.

However, if we consider elastic collisions between particles, we notice that even though these do not violate the conservation of the sum of constants of motion *linear* in momenta, they will destroy the conservation of the sum of the constants of motion *quadratic* in momenta. Formally the elastic collision is characterized by

$$\sum_i p'^{\mu}_{(i)} - \sum_i p^{\mu}_{(i)} = 0\,, \tag{5}$$

where the primed and unprimed quantities always signify quantities right after and before the collision respectively.

As a consequence, one can contract the momentum conservation (5) with the KY tensor to see that the sums of angular momentum vectors associated with KY tensors are always conserved

$$\sum_i (Y_{\mu\kappa} p'^{\mu}_{(i)}) - \sum_i (Y_{\mu\kappa} p^{\mu}_{(i)}) = 0\,. \tag{6}$$

Similarly, if we contract (5) with a Killing vector $\xi_\mu$, we see that that the sum of $\xi_\mu p^{\mu}_{(i)}$ is conserved in collisions. However, as was noticed also by[7], if we try to see what happens to the sum of the quadratic Carter constants $C_{(i)} \equiv K_{\mu\nu} p^{\mu}_{(i)} p^{\nu}_{(i)}\,, K_{\mu\nu} \equiv Y_{\mu\kappa} Y^{\kappa}_{\ \nu}$, in momentum exchanges, we obtain

$$\sum_i C'_{(i)} - \sum_i C_{(i)} = \sum_i \sum_{j\neq i} K_{\mu\nu} p^{\mu}_{(i)} p^{\nu}_{(j)} - \sum_i \sum_{j\neq i} K_{\mu\nu} p'^{\mu}_{(i)} p'^{\nu}_{(j)}\,. \tag{7}$$

It can then be shown that this leads to $\mathcal{C}^{\mu}_{\ ;\mu} \neq 0$ whenever collisions occur within the particle cloud. The conservation of the angular momentum vectors (6) leads to the identity $(T^{\mu\nu} Y_{\nu}^{\ \kappa})_{;\mu} = 0$, but this is not a conservation law for a current. However, it is possible to create one by taking a divergence with respect to the redundant index and thus one obtains a conserved current $j^{\mu}_{\mathrm{KY}} = T^{\mu\nu;\kappa} Y_{\nu\kappa}$. I will call it the flux of KY vorticity, since it is proportional to anti-symmetric gradients of stress-energy.

As an interesting side note, the recent work[8,9] discusses that even though the collisions spoil the long-term evolution of the sum of Carter constant of a general matter flow, it is always possible to find a certain effective Carter constant that is conserved for sufficiently compact bodies for a short enough time. However, this fact seems to be useful only for the evolution of compact astrophysical objects such as neutron stars and black holes in extreme mass ratio inspirals.

## 3. Demonstration on dust in Kerr

Consider the Kerr space-time in Boyer-Lindquist coordinates $t, \varphi, r, \vartheta$, $\Delta \equiv r^2 - 2Mr + a^2$, $\Sigma = r^2 + a^2 \cos^2\vartheta$, and $M, a$ are the black hole mass and spin respectively.

We set up a field of dust $T^{\mu\nu} = \rho u^\mu u^\nu$ which is infalling into the Kerr black hole with a mass density $\rho$ and the four-velocity

$$u_\varphi = u_\vartheta = 0\,, \tag{8}$$

$$u_t = -1\,, \tag{9}$$

$$u_r = -\frac{\sqrt{2Mr(r^2+a^2)}}{\Delta} + \varepsilon\,, \tag{10}$$

where when the dimensionless constant $\varepsilon = 0$ then the velocity field fulfils the geodesic equation $a^\mu = 0$ but violates it for $\varepsilon \neq 0$ by introducing two non-zero components of acceleration

$$a_\vartheta = -\varepsilon\frac{2a^2\sin\vartheta\cos\vartheta\sqrt{2Mr(r^2+a^2)}}{\Sigma^2} + \mathcal{O}(\varepsilon^2)\,, \tag{11}$$

$$a_r = -\varepsilon\left(\frac{\sqrt{2Mr(r^2+a^2)}}{\Sigma}\right)_{,r} + \mathcal{O}(\varepsilon^2)\,. \tag{12}$$

Since the velocity field is purely radial, the continuity equation $(\rho u^\mu)_{;\mu} = 0$ has the simple solution

$$\rho = -\frac{\dot{M}(\vartheta)}{2\pi\Delta u_r}\,, \tag{13}$$

where $\dot{M}(\vartheta)$ is an arbitrary positive function of $\vartheta$ which represents the dust accretion rate through a $\vartheta = \text{const.}$ layer. In the examples here we choose $\dot{M} = const.$.

Now the $\varepsilon$ term will *not* induce deviations from the conservation of mass, energy, or angular momentum. Hence, it cannot be detected by existing tests. However, it *will* induce deviations from the KY conservation laws and relations given above. Specifically, I considered the integral formulation of the KY vorticity conservation $(j^\mu_{\rm KY})_{;\mu} = 0$, the KY-projected equations of motion $(T^{\mu\nu}Y_\nu{}^\kappa)_{;\mu} = 0$, and the equations of motion $T^{\mu\nu}{}_{;\nu} = 0$.

Now consider receiving the data about the flow on a finite numerical grid in the interval $r \in (3M, 6M)$. When evaluated numerically, the non-zero value of the right-hand sides of the equations above indicate an error in the evolution (in our case the dust infall). However, there will always be numerical noise in the right-hand sides. In particular, the evaluation of $j^\mu_{\rm KY}$ requires taking numerical derivatives of the data, which proved to be a strong source of noise. I have thus defined a "detection" value $\varepsilon_{\rm d}$ of $\varepsilon$ as such that the values of the right-hand sides are ten times that of the noise for the unperturbed "correct" flow. It is then instructive to see how sensitive is the conservation law as a function of the numerical grid spacing.

The results of the comparison are given in Fig. 1. It turns out that even though the KY vorticity is a "proper" conservation law, the numerical noise associated with the numerical derivatives makes it difficult to use for data with large grid spacings. Nevertheless, the "KY-projected equations of motion" represent an interesting alternative to just integrating the equations of motion themselves.

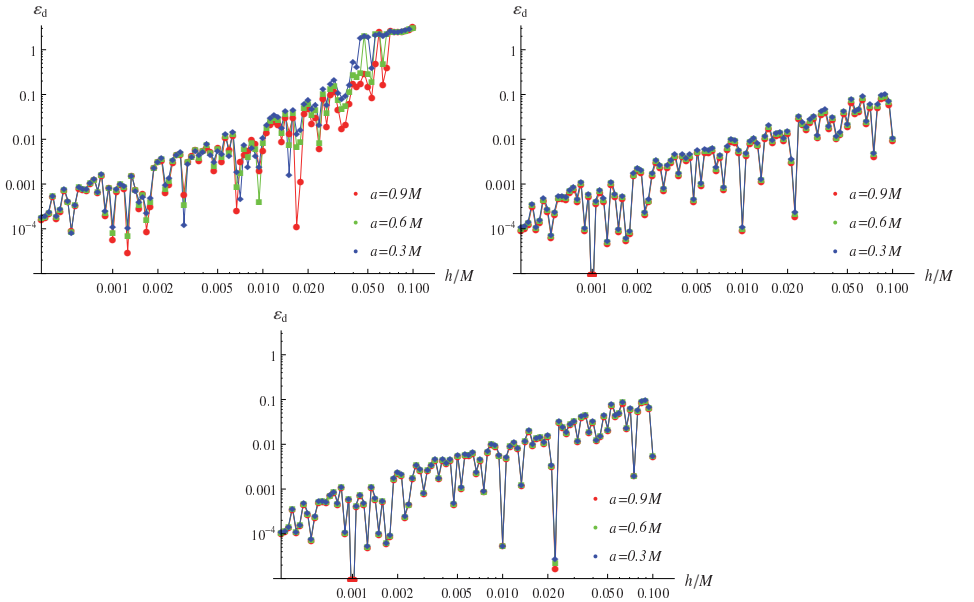



Fig. 1. The detection value of the perturbation $\varepsilon_d$ as a function of grid spacing $h$ for various choices of the black hole spin. The top left figure corresponds to the KY vorticity conservation, top right to KY-projected equations of motion, and the bottom to the plain equations of motion.

## Acknowledgments

I am grateful for the support from a Ph.D. grant of the German Research Foundation within its Research Training Group 1620 *Models of Gravity* and for the support from the Grant No. GACR-17-06962Y of the Czech Science Foundation.

## References

1. V. Witzany, Exploiting the hidden symmetry of spinning black holes: conservation laws and numerical tests, *Mon. Not. R. Astron. Soc.* **473**, 2434 (2017).
2. C. F. Gammie, J. C. McKinney and G. Tóth, Harm: a numerical scheme for general relativistic magnetohydrodynamics, *ApJ* **589**, p. 444 (2003).
3. S. S. Komissarov, Magnetized tori around Kerr black holes: analytic solutions with a toroidal magnetic field, *MNRAS* **368**, 993 (2006).
4. V. Witzany and P. Jefremov, New closed analytical solutions for geometrically thick fluid tori around black holes. Numerical evolution and the onset of the magneto-rotational instability, *Astron. Astrophys.* **614**, p. A75 (2018).
5. C. Markakis, K. Uryū, E. Gourgoulhon, J.-P. Nicolas, N. Andersson, A. Pouri and V. Witzany, Conservation laws and evolution schemes in geodesic, hydrodynamic, and magnetohydrodynamic flows, *Phys. Rev. D* **96**, p. 064019 (2017).

6. B. Carter, Global structure of the Kerr family of gravitational fields, *Phys. Rev.* **174**, p. 1559 (1968).
7. A. Grant and É. É. Flanagan, Non-conservation of carter in black hole spacetimes, *Class. Quant. Grav.* **32**, p. 157001 (2015).
8. V. Witzany, Spin-perturbed orbits near black holes (2019), Submitted to Phys. Rev. Lett., arXiv:1903.03649.
9. V. Witzany, Hamilton-Jacobi equation for spinning particles near black holes (2019), Submitted to Phys. Rev. D, arXiv:1903.03651.

# Black hole thermodynamics and Gravity's Rainbow

Remo Garattini

*Università degli Studi di Bergamo,*
*Dipartimento di Ingegneria e Scienze Applicate,*
*Viale Marconi,5 24044 Dalmine (Bergamo) Italy*
*I.N.F.N. - sezione di Milano, Milan, Italy*
** E-mail:remo.garattini@unibg.it*

We consider the effects of rotations on the calculation of some thermodynamical quantities like the free energy, internal energy and entropy. In ordinary gravity, when we evaluate the density of states of a scalar field close to a black hole horizon, we obtain a divergent result which can be kept under control with the help of some standard regularization and renormalization processes. We show that when we use the Gravity's Rainbow approach such regularization/renormalization processes can be avoided. A comparison between the calculation done in an inertial frame and in a comoving frame is presented.

## 1. Introduction

Gravity's Rainbow (GRw) is a modification of the space-time close to the Planck scale. It has been introduced for the first time by Magueijo and Smolin [1]. Basically one defines two unknown functions $g_1\left(E/E_P\right)$ and $g_2\left(E/E_P\right)$ having the following property

$$\lim_{E/E_P\to 0} g_1\left(E/E_P\right)=1 \qquad \text{and} \qquad \lim_{E/E_P\to 0} g_2\left(E/E_P\right)=1. \tag{1}$$

This property guarantees the recovery of ordinary General Relativity when sub-planckian physics is involved. In this formalism, the Einstein's field equations are replaced by a one parameter family of equations

$$G_{\mu\nu}\left(E\right)=8\pi G\left(E\right)T_{\mu\nu}\left(E\right)+g_{\mu\nu}\Lambda\left(E\right), \tag{2}$$

where $G\left(E\right)$ is an energy dependent Newton's constant and $\Lambda\left(E\right)$ is an energy dependent cosmological constant, respectively. They are defined so that $G\left(0\right)$ is the physical Newton's constant and $\Lambda\left(0\right)$ is the usual cosmological constant. In this context, the Schwarzschild solution of (2) becomes

$$ds^2\left(E\right)=-\left(1-\frac{2MG\left(0\right)}{r}\right)\frac{dt^2}{g_1^2\left(E/E_P\right)}+\frac{dr^2}{\left(1-\frac{2MG(0)}{r}\right)g_2^2\left(E/E_P\right)}+\frac{r^2}{g_2^2\left(E/E_P\right)}d\Omega^2. \tag{3}$$

An immediate generalization of the metric (3) is represented by the following line element

$$ds^2\left(E\right)=-\left(1-\frac{b\left(r\right)}{r}\right)\frac{\exp\left(-2\Phi\left(r\right)\right)}{g_1^2\left(E/E_P\right)}dt^2+\frac{dr^2}{\left(1-\frac{b(r)}{r}\right)g_2^2\left(E/E_P\right)}+\frac{r^2}{g_2^2\left(E/E_P\right)}d\Omega^2. \tag{4}$$

The function $b\left(r\right)$ will be referred to as the "shape function" and it may be thought of as specifying the shape of the spatial slices. The location of the horizon is

determined by the equation $b(r_H) = r_H$. On the other hand, $\Phi(r)$ will be referred to as the "redshift function" and describes how far the total gravitational redshift deviates from that implied by the shape function. The line element (4) describes any spherically symmetric space-time with a horizon: by definition, this is a black hole distorted by GRw. Note that a metric of the form

$$ds^2(E) = -\frac{\exp(-2\Phi(r))}{g_1^2(E/E_P)}dt^2 + \frac{dr^2}{\left(1-\frac{b(r)}{r}\right)g_2^2(E/E_P)} + \frac{r^2}{g_2^2(E/E_P)}d\Omega^2, \qquad (5)$$

describes a traversable wormhole modified by GRw if $\exp(-2\Phi(r))$ never vanishes[18]. The line element (5) has been used in a series of papers where to avoid any regularization/renormalization scheme which appear in conventional Quantum Field Theory calculations like one loop corrections to classical quantities[2]. On the other hand, the line element (5) has been considered for the computation of black hole entropy[3]. In this last case, the idea is to avoid to introduce a cut-off of Planckian size known as "*brick wall*"[4]. The "*brick wall*" appears when one uses a statistical mechanical approach to explain the famous Bekenstein-Hawking formula[5,6]

$$S_{BH} = \frac{1}{4}A/l_P^2, \qquad (6)$$

relating the entropy of a black hole and its area. Indeed, when one tries to adopt such an approach, one realizes that the density of energy levels of single-particle excitations is divergent near the horizon. Of course, several attempts have been done to avoid the introduction of the *brick wall*. For instance, without modifying gravity at any scale, it has been suggested that the *brick wall* could be absorbed in a renormalization of Newton's constant[7–9], while other authors approached the problem of the divergent brick wall using Pauli-Villars regularization[10–12]. Other than GRw other proposals have been made in the context of modified gravity. For instance, non-commutative geometry introduces a natural thickness of the horizon replacing the 't Hooft's brick wall[17] and *Generalized Uncertainty Principle* (GUP) modifies the Liouville measure[13–16]. Nevertheless we can wonder what happens when one introduces rotations. For instance, one could consider the free energy obtained for a real massless scalar field, rotating with an angular velocity $\Omega_0$ around the $z$ axis in Minkowski space

$$F = \frac{1}{\beta}\sum_m \int_0^\infty dg(E,m)\ln\left(1-e^{-\beta(E-m\Omega_0)}\right). \qquad (7)$$

For this case, it is better to work in cylindrical coordinates

$$ds^2 = -dt^2 + dr^2 + r^2 d\phi^2 + dz^2 \qquad (8)$$

and with the help of a WKB approximation it is possible to show that a divergence appear close to the speed-of-light (SOL) surface, defined as the surface where $r = \Omega_0^{-1}$[19]. This divergence can be taken under control with the help of GRw by

modifying the line element (8) in the following way [20]

$$ds^2 = -\frac{dt^2}{g_1^2\left(E/E_P\right)} + \frac{dr^2}{g_2^2\left(E/E_P\right)} + \frac{r^2 d\phi^2}{g_2^2\left(E/E_P\right)} + \frac{dz^2}{g_2^2\left(E/E_P\right)}, \tag{9}$$

where for simplicity we have replaced $ds^2\left(E\right)$ with $ds^2$. The same above thermodynamical system can be analyzed on a comoving frame rotating with the same angular velocity $\Omega_0$. By plugging $\phi' = \phi - \Omega_0 t$, from the line element (3) we obtain

$$ds^2 = -\frac{\left(1-\Omega_0^2 r^2\right)}{g_1^2\left(E/E_P\right)}dt^2 + \frac{2\Omega_0 r^2 d\phi' dt}{g_1\left(E/E_P\right)g_2\left(E/E_P\right)} + \frac{dr^2}{g_2^2\left(E/E_P\right)} + \frac{r^2 d\phi'^2}{g_2^2\left(E/E_P\right)} + \frac{dz^2}{g_2^2\left(E/E_P\right)}. \tag{10}$$

It is immediate to see that in this system of coordinates appears a fictitious horizon located at $r = \Omega_0^{-1}$, namely the SOL surface of the rotating heat bath introduced in (7). Note that a mixing between $g_1\left(E/E_P\right)$ and $g_2\left(E/E_P\right)$ appears. A similar mixing appears also in a Vaidya spacetime for

$$ds^2 = -\left(1 - \frac{2M\left(v\right)G}{r}\right)\frac{dv^2}{g_1^2\left(E/E_P\right)} + 2\frac{dvdr}{g_1\left(E/E_P\right)g_2\left(E/E_P\right)} + \frac{r^2 d\Omega^2}{g_2^2\left(E/E_P\right)}, \tag{11}$$

where $v$ is the advanced (ingoing) null coordinate and finally also for the Kerr metric which, in the context of GRw, becomes [22]

$$ds^2 = \frac{g_{tt}dt^2}{g_1^2\left(E/E_P\right)} + \frac{2g_{t\phi}dtd\phi}{g_1\left(E/E_P\right)g_2\left(E/E_P\right)} + \frac{g_{\phi\phi}d\phi^2}{g_2^2\left(E/E_P\right)} + \frac{g_{rr}dr^2}{g_2^2\left(E/E_P\right)} + \frac{g_{\theta\theta}d\theta^2}{g_2^2\left(E/E_P\right)}, \tag{12}$$

where

$$\begin{aligned} g_{tt} &= -\frac{\Delta - a^2\sin^2\theta}{\Sigma}, \qquad g_{t\phi} = -\frac{a\sin^2\theta\left(r^2+a^2-\Delta\right)}{\Sigma}, \\ g_{\phi\phi} &= \frac{\left(r^2+a^2\right)^2 - \Delta a^2\sin^2\theta}{\Sigma}\sin^2\theta, \qquad g_{rr} = \frac{\Sigma}{\Delta}, \qquad g_{\theta\theta} = \Sigma, \end{aligned} \tag{13}$$

and

$$\Delta = r^2 - 2MGr + a^2, \qquad \Sigma = r^2 + a^2\cos^2\theta. \tag{14}$$

Here $M$ and $a$ are mass and angular momentum per unit mass of the black hole, respectively. $\Delta$ vanishes when $r = r_\pm = MG \pm \sqrt{\left(MG\right)^2 - a^2}$, while $g_{tt}$ vanishes when $r = r_{S\pm} = MG \pm \sqrt{\left(MG\right)^2 - a^2\cos^2\theta}$: they are not modified by GRw and the outer horizon or simply horizon is located at $r_+ = r_H$. Units in which $\hbar = c = k = 1$ are used throughout the paper.

## 2. GRw Entropy for the Kerr Black Hole

To discuss the entropy for a Kerr black hole we have two options: we can use a rest observer at infinity (ROI) or we can use a Zero Angular Momentum Observer

(ZAMO)[19,23]. The ROI frame is described by the line element (12) and the appropriate form of the free energy is the following

$$F = \frac{1}{\beta}\int_0^\infty dn\,(E)\ln\left(1-e^{-\beta(E-m\Omega_0)}\right). \tag{15}$$

It is immediate to see that when we use a ROI, the problem of superradiance appears when the free energy (15) is computed in the range $0 < E < m\Omega_0$. On the other hand when a ZAMO is considered, the free energy (15) becomes similar to the one used for a Schwarzschild black hole. Basically this happens because near the horizon the metric becomes

$$ds^2 = -\frac{N^2dt^2}{g_1^2\,(E/E_P)} + g_{\phi\phi}\frac{d\phi^2}{g_2^2\,(E/E_P)} + g_{rr}\frac{dr^2}{g_2^2\,(E/E_P)} + g_{\theta\theta}\frac{d\theta^2}{g_2^2\,(E/E_P)} \tag{16}$$

and the mixing between $t$ and $\phi$ disappears. Moreover when we use a ZAMO frame, the superradiance does not come into play because there is no ergoregion. Indeed since we have defined

$$N^2 = g_{tt} - \frac{g_{t\phi}^2}{g_{\phi\phi}} = -\frac{1}{g^{tt}} = -\frac{\Delta\sin^2\theta}{g_{\phi\phi}}, \tag{17}$$

$N^2$ vanishes when $r \to r_H$. The number of modes with frequency less than $E$ is given approximately by

$$n(E) = \frac{1}{\pi}\int_0^{l_{max}}(2l+1)\int_{r_H}^{R}\sqrt{k^2(r,l,E)}drdl. \tag{18}$$

Here it is understood that the integration with respect to $r$ and $l$ is taken over those values which satisfy $r_H \leq r \leq R$ and $k^2(r,l,E) \geq 0$. Thus one finds

$$\frac{dn(E)}{dE} = \frac{1}{8\pi^2}\int d\theta d\bar{\phi}\int_{r_H}^{R}dr\left(-g^{tt}\right)^{\frac{3}{2}}\sqrt{g_{rr}g_{\theta\theta}g_{\phi\phi}}\frac{1}{3}\frac{d}{dE}\left(h^3\,(E/E_P)\,E^3\right). \tag{19}$$

In proximity of the horizon, the free energy can be approximated by

$$F_{r_H} = \frac{1}{8\pi^2\beta}\int d\theta d\bar{\phi}\int_0^\infty \ln\left(1-e^{-\beta E}\right)\frac{d}{dE}\left(\frac{1}{3}h^3\,(E/E_P)\,E^3\right)H\,(r_H,r_1)\,dE \tag{20}$$

where we have defined

$$H\,(r_H,r_1) = \int_{r_H}^{r_1}dr\left(-g^{tt}\right)^{\frac{3}{2}}\sqrt{g_{rr}g_{\theta\theta}g_{\phi\phi}}. \tag{21}$$

$F_{r_H}$ can be further reduced to

$$F_{r_H} \simeq \frac{C\,(r_H,\theta)}{8\pi^2\beta}\int_0^\infty\frac{\ln\left(1-e^{-\beta E}\right)}{\sigma\,(E/E_P)}\frac{d}{dE}\left(\frac{1}{3}h^3\,(E/E_P)\,E^3\right)dE, \tag{22}$$

where

$$C\,(r_H,\theta) = \int d\theta d\bar{\phi}\left[\frac{\left(r_H^2+a^2\right)^4\sin\theta}{r_H\left(r_H-r_-\right)^2\Sigma_H}\right] \tag{23}$$

and where we have assumed that, in proximity of the throat the brick wall can be written as $r_0\left(E/E_P\right)=r_H\sigma\left(E/E_P\right)$ with

$$\sigma\left(E/E_P\right)\to 0, \qquad E/E_P\to 0. \tag{24}$$

With an integration by parts one finds

$$\begin{aligned} F_{r_H} &= -\frac{C\left(r_H,\theta\right)}{24\pi^2\beta}\int_0^\infty \frac{E^3h^3\left(E/E_P\right)}{\sigma\left(E/E_P\right)} \\ &\times\left[\frac{\beta}{\left(\exp\left(\beta E\right)-1\right)}-\frac{\ln\left(1-e^{-\beta E}\right)}{E_P\sigma\left(E/E_P\right)}\sigma'\left(E/E_P\right)\right]dE. \end{aligned} \tag{25}$$

It is possible to show that

$$\begin{aligned} F_{r_H} &= -\frac{C\left(r_H,\theta\right)}{24\pi^2\beta}\int_0^\infty\left[\frac{\beta E e^{-3E/E_P}}{\left(\exp\left(\beta E\right)-1\right)}-2e^{-3E/E_P}\ln\left(1-e^{-\beta E}\right)\right]dE \\ &= -\frac{C\left(r_H,\theta\right)}{24\pi^2\beta}\left[\zeta\left(2,1+\frac{3}{\beta E_P}\right)+\frac{\beta E_P}{3}\left(\gamma+\Psi\left(1+\frac{3}{\beta E_P}\right)\right)\right], \end{aligned} \tag{26}$$

where $\zeta\left(s,\nu\right)$ is the Hurwitz zeta function, $\Gamma\left(x\right)$ is the gamma function and $\Psi\left(x\right)$ is the digamma function. In the limit where $\beta E_P\gg 1$, at the leading order, one finds that the entropy can be approximated by

$$S=\beta^2\frac{\partial F_{r_w}}{\partial\beta}=\frac{E_P^2}{36\beta}\int d\theta d\bar{\phi}\left[\frac{\left(r_H^2+a^2\right)^4\sin\theta}{r_H\left(r_H-r_-\right)^2\Sigma_H}\right] \tag{27}$$

and even when rotation is included, the "*brick wall*" does not appear. Of course the entropy (27) can always be cast in the familiar form

$$S=\frac{A_H}{4G}, \tag{28}$$

where $A_H$ is the horizon area. To summarize, we have shown that the ability of Gravity's Rainbow to keep under control the UV divergences applies also to rotations. However the connection between a ROI and a ZAMO has to be investigated with care[24]. Indeed in the ROI frame, the superradiance phenomenon appears, while in the ZAMO frame does not. Once the connection is established nothing forbids to extend this result to other rotating configuration like, for example, Kerr-Newman or Kerr-Newman-De Sitter (Anti-De Sitter).

## Acknowledgments

The author would like to thank MDPI for a partial financial support.

## References

1. J. Magueijo and L. Smolin, *Class. Quant. Grav.* **21**, 1725 (2004) [arXiv:gr-qc/0305055].

2. R. Garattini and G. Mandanici, *Phys. Rev.* **D 83**, 084021 (2011); arXiv:1102.3803 [gr-qc]. R. Garattini and G. Mandanici, *Phys. Rev.* **D 85**, 023507 (2012); arXiv:1109.6563 [gr-qc]. R. Garattini, *JCAP* **017** 1306, (2013); arXiv:1210.7760 [gr-qc]; R. Garattini and B. Majumder, *Nucl. Phys.* **B 884** 125, (2014); arXiv:1311.1747 [gr-qc]. R. Garattini and B. Majumder, *Nucl. Phys.* **B 883** 598, (2014); arXiv:1305.3390 [gr-qc]. R. Garattini and E. N. Saridakis, *Eur. Phys. J.* **C 75**, 343 (2015); ArXiv:1411.7257 [gr-qc].
3. R. Garattini, *Phys. Lett.* **B 685**, 329 (2010); arXiv:0902.3927 [gr-qc].
4. G. 't Hooft, *Nucl. Phys.* **B 256**, 727 (1985).
5. J. D. Bekenstein, *Phys. Rev.* **D 7**, 949 (1973).
6. S. W. Hawking, *Comm. Math. Phys.* **43**, 199 (1975).
7. L. Susskind and J. Uglum, *Phys. Rev.* **D 50**, 2700 (1994).
8. J. L. F. Barbon and R. Emparan, *Phys. Rev.* **D 52**, 4527 (1995), hep-th/9502155.
9. E. Winstanley, *Phys. Rev.* **D 63**, 084013 (2001), hep-th/0011176.
10. J.-G. Demers, R. Lafrance and R. C. Myers, *Phys. Rev.* **D 52**, 2245 (1995), gr-qc/9503003.
11. D. V. Fursaev and S. N. Solodukhin, *Phys. Lett.* **B 365**, 51 (1996), hep-th/9412020.
12. S. P. Kim, S. K. Kim, K.-S. Soh and Jae Hyung Yee, *Int. J. Mod. Phys.* **A 12**, 5223 (1997), gr-qc/9607019.
13. X. Li, *Phys. Lett.* **B 540**, 9 (2002), gr-qc/0204029.
14. Z. Ren, W. Yue-Qin and Z. Li-Chun, *Class. Quant. Grav.* **20** (2003), 4885.
15. G. Amelino-Camelia, *Class. Quant. Grav.* **23**, 2585 (2006), gr-qc/0506110. G. Amelino-Camelia, *Gen. Rel. Grav.* **33**, 2101 (2001), gr-qc/0106080.
16. E. C. Vagenas, A. F. Aliha, M. Hemeda and H. Alshal, *Eur. Phys. J.* **C 79** 398 (2019) no.5, arXiv:1903.08494 [hep-th].
17. H. Bai and M. L. Yan, JHEP **7** 58 (2003), gr-qc/0303006.
18. Visser M 1995 *Lorentzian Wormholes: From Einstein to Hawking* (American Institute of Physics, New York).
19. M. H. Lee and J. K. Kim, *Phys. Rev.* **D 54**, 3904 (1996); ArXiv:hep-th/9603055.
20. In preparation.
21. J. L. Cortes and J. Gamboa, *Phys. Rev.* **D 71**, 065015 (2005); hep-th/0405285
22. Y. Zhao and X. Liu, arXiv:1606.06285 [gr-qc].
23. E. Chang-Young, D. Lee and M. Yoon, *Class. Quant. Grav.* **26**, 155011 (2009) [arXiv:0811.3294 [hep-th]].
24. In preparation.

# Gravitational perturbation in holography between rotating black holes and conformal field theory

Masoud Ghezelbash

*Department of Physics and Engineering Physics,*
*University of Saskatchewan*
*Saskatoon, Saskatchewan, Canada*
*E-mail: amg142@campus.usask.ca*
*www.usask.ca*

The holography for rotating black holes provides a description of black hole physics in terms of a certain two-dimensional conformal field theory (CFT). The first realization of the holography for rotating black holes was proposed for extremal Kerr black hole. The proposal states that the near horizon geometry of the extremal Kerr black hole does have a dual holographic description in terms of a two-dimensional chiral CFT. The proposal has been studied and extended extensively for other extremal or near extremal rotating black holes. For all extremal and near-extremal black holes in different dimensions, the physical quantities associated to black hole (such as the entropy or scattering cross section) are in agreement with the corresponding microscopic dual quantities in CFT. In this talk, we find the explicit form of two-point function for the conformal spin-2 energy momentum operators on the near horizon of a near extremal Kerr black hole by variation of a proper boundary action. In this regard, we consider an appropriate boundary action for the gravitational perturbation of the Kerr black hole. We consider the boundary action and calculate it on-shell to find the two-point functions for the spin-2 energy momentum tensor fields, on the near horizon region of the Kerr black hole. We show that the variation of the boundary action with respect to the boundary fields yields the two-point function for the energy momentum tensor of a conformal field theory. We find agreement between the two-point function and the correlators of the dual conformal field theory to the Kerr black hole.*

*Keywords*: Gravitational perturbation, Kerr black hole, Kerr/CFT correspondence.

## 1. Introduction

Inspired by the fact that near horizon region of a near extremal Kerr black hole contains a two dimensional copy of Anti de Sitter[1], a field theoretic calculation can be done to find the two-point correlation functions of conformal fields with different spin. Explicitly, in Ref. 2, the two-point correlation functions of spinor fields were constructed by calculating the on-shell gravitational action for the spinor fields with appropriate boundary condition. Though the on-shell bulk action for the spinor fields vanishes, however one can add a boundary term to the gravitational action which does not vanish on-shell[3]. The variation of the boundary action yields the proper two-point correlation functions for the conformal operators. In a recent article, Ref. 4, the authors used an appropriate boundary action for the vector fields to calculate the two-point functions of conformal vector fields. In this article, we derive the two-point function for the energy-momentum tensor fields in the

*This presentation is entirely based on published paper in the journal *Gen. Relativ. Gravit.* **48**, 102 (2016) by A.M. Ghezelbash.

background of Kerr black hole. The first step is to find an appropriate boundary action for the spin-2 fields. In this regard, we write the Einstein-Hilbert action as the summation of two actions. The first action is a bulk action that unlike the Einstein-Hilbert action depends only on the metric tensor and its first derivative. Although the bulk action is not a scalar under the general coordinate transformations, however it leads to the same Einstein's field equations. The second action is a boundary action that depends on the metric tensor and the Christoffel symbols. We consider the boundary action and calculate it on-shell to find the two-point functions for the spin-2 energy momentum tensor fields, on the near horizon region of the Kerr black hole. We then compare the result of our calculation to the correlation functions for spin-2 conformal operators in a conformal field theory.

## 2. The Gravitational Perturbations and the Boundary Action for Them in the Rotating Black Hole

We consider the gravitational action for a general four-dimensional curved spacetime $\mathcal{M}$ with a boundary $\partial M$ as $S = S_{EH} + S_{GH}$, where $S_{EH} = \frac{1}{16\pi}\int_{\mathcal{M}} d^4x\sqrt{-g}R$ and $S_{GH} = \frac{1}{8\pi}\int_{\partial\mathcal{M}} d^3x\sqrt{k}K$ is the Gibbons-Hawking boundary action. $k_{\mu\nu}$ is the induced metric on the boundary $\partial\mathcal{M}$ and $K = \nabla_\mu N^\mu$ is the trace of extrinsic curvature and $N^\mu$ is the unit normal vector to $\partial\mathcal{M}$. We may use the following identity

$$S_{EH} = S_1 + \frac{1}{16\pi}\int_{\partial M} d^3x\sqrt{-g}N_\mu\{g^{\mu\nu}\Gamma^\sigma_{\nu\sigma} - g^{\sigma\nu}\Gamma^\mu_{\sigma\nu}\}, \tag{1}$$

where

$$S_1 = \frac{1}{16\pi}\int_{\mathcal{M}} d^4x\sqrt{-g}g^{\mu\nu}\{\Gamma^\tau_{\mu\nu}\Gamma^\sigma_{\tau\sigma} - \Gamma^\tau_{\mu\sigma}\Gamma^\sigma_{\tau\nu}\}, \tag{2}$$

to write $S_{EH}$ as the summation of a bulk action that depends only on the metric field and the first order derivatives of the metric field and a boundary term. In fact, The bulk action $S_1$ clearly depends on the gravitational field and its first derivatives and resembles to the action for a scalar field or spinor or electromagnetic field that only depends on dynamical field and its first derivatives. We consider the gravitational action $\mathcal{S}$ for the four-dimensional curved spacetime $\mathcal{M}$ with a boundary $\partial\mathcal{M}$ as[5]

$$\mathcal{S} = \frac{1}{16\pi}\int_{\mathcal{M}} d^4x\sqrt{-g}g^{\mu\nu}\{\Gamma^\tau_{\mu\nu}\Gamma^\sigma_{\tau\sigma} - \Gamma^\tau_{\mu\sigma}\Gamma^\sigma_{\tau\nu}\} + S_B, \tag{3}$$

where

$$S_B = \frac{1}{16\pi}\int_{\partial\mathcal{M}} d^3x\sqrt{-g}N_r\{g^{r\nu}\Gamma^\sigma_{\nu\sigma} - g^{\sigma\nu}\Gamma^r_{\sigma\nu}\}. \tag{4}$$

We consider $\mathcal{M}$ as the near extremal Kerr black hole and use the boundary action (4) to calculate the two-point functions of spin 2 gravitational perturbation fields $\tilde{h}_{\mu\nu}$ on the near-horizon geometry of the near extremal black hole where $g_{\mu\nu} = g^{(0)}_{\mu\nu} + \epsilon\tilde{h}_{\mu\nu}$

and $g^{(0)}_{\mu\nu}$ describes the Kerr black hole which the line element in Boyer-Lindquist coordinates is given by

$$ds^2 = -\frac{\Delta}{\rho^2}\left(dt - a\sin^2\theta d\phi\right)^2 + \frac{\rho^2}{\Delta}dr^2 + \rho^2 d\theta^2 + \frac{\sin^2\theta}{\rho^2}\left(adt - \left(r^2+a^2\right)d\phi\right)^2, \quad (5)$$

where $\rho^2 = \bar{\rho}\bar{\rho}^*$, $\bar{\rho} = r + ia\cos\theta$ and $\Delta = r^2 + a^2 - 2Mr$. We should note that in (4), $\partial\mathcal{M}$ is the hypersurface $r = r_B$ which is the near-horizon of near-NHEK geometry for the Kerr black hole. Moreover the field $\tilde{h}_{\mu\nu}$ is the spin 2 field which represents the gravitational perturbation of the Kerr black hole (5). The explicit representation of the different components of $\tilde{h}_{\mu\nu}$ is given in Ref. 5. In fact, a detailed analysis of the gravitational perturbation fields shows that we can represent the fields $\tilde{h}_{\mu\nu}$ as the real parts of $\mathbf{h}_{\mu\nu}$, where[5]

$$\mathbf{h}_{\mu\nu}(t, r, \theta, \phi) = h_{\mu\nu}(r, \theta)e^{-i\omega t + im\phi}, \quad (6)$$

and

$$h^{\mu\nu} = \sum_{i,j=+,-} f^{\mu\nu}_{ij}(r, \theta)P_i(r)S_j(\theta), \quad (7)$$

where

$$f^{\mu\nu}_{ij} = f^{\mu\nu(-1)}_{ij}\Delta^{-1} + f^{\mu\nu(0)}_{ij}\Delta^0 + \cdots. \quad (8)$$

In equation (7), $P_\pm(r)$ and $S_\pm(\theta)$ are related to the radial and angular Teukolsky functions, respectively[5]. We notice that the presence of $\rho$ in all components of the graviton $h^{\mu\nu}$, makes the results (8) completely non-separable in terms of coordinates $r$ and $\theta$. However, in calculation of the boundary action, we consider a spherical boundary with radius $r_B$ as the boundary of the near-NHEK geometry. We set the radial coordinate in $\rho$ equal to $r_B$ and so $\rho_B = \rho(r = r_B)$ depends only on $\theta$. As a result, we check out explicitly that any single term in the expansion (7) is completely separable in terms of coordinates $r$ and $\theta$. In other words, we can re-write (7) as

$$h^{\mu\nu} = \sum_{i,j=+,-} \chi_i \sum_{\aleph} f^{\overline{\mu\nu}}_{1ij\aleph}(r) f^{\overline{\mu\nu}}_{2ij\aleph}(\theta)R_i(r)S_j(\theta), \quad (9)$$

where $\chi_+ = \Delta_B$, $\chi_- = 1$, the over-line means that there is no summation over superscripts $\mu$ and $\nu$. The summation index $\aleph$ in (9) shows all the individual terms that appear in expansion (8). After a long calculation, we can show that the only relevant terms in the boundary action (4) that lead to the proper two-point function are given by[5]

$$S^{(2)}_B = \int_{r=r_B} d^3x\sqrt{-g^{(0)}}N_r \left\{ g^{(0)r\nu}\Gamma^{(2)\sigma}_{\nu\sigma} + \Gamma^{(1)\sigma}_{\nu\sigma}\left(\frac{1}{2}g^{(0)r\nu}\tilde{h} - \tilde{h}^{r\nu}\right) + \Gamma^{(0)\sigma}_{\nu\sigma}\left(-\frac{1}{2}\tilde{h}\tilde{h}^{r\nu} + \frac{g^{(0)r\nu}}{8}(\tilde{h}^2 - \tilde{h}_{\rho\lambda}\tilde{h}^{\rho\lambda})\right)\right\}, \quad (10)$$

where $\Gamma^{(i)\rho}_{\mu\nu}$, $i=0,1,2$ are the expansion coefficients of the Christoffel affine connection in terms of $\epsilon$, as $\Gamma^{\rho}_{\mu\nu}=\Gamma^{(0)\rho}_{\mu\nu}+\epsilon\Gamma^{(1)\rho}_{\mu\nu}+\epsilon^2\Gamma^{(2)\rho}_{\mu\nu}$. After some substitutions and showing the $(t,\phi)$-dependence of the boundary fields by $\mathcal{H}^{\alpha\beta}_{B+}(t,\phi)$ and rewriting the boundary field $h^{\alpha\beta}_{B+}$ as

$$h^{\alpha\beta}_{B+}(t,\theta,\phi)=\mathcal{H}^{\overline{\alpha\beta}}_{B+}(t,\phi)\Theta^{\overline{\alpha\beta}}_{\pm}(\theta), \tag{11}$$

where $\mathcal{H}^{\overline{\alpha\beta}}_{B+}(t,\phi)$ is equal to $\Delta_B e^{-i\omega t+im\phi}R_{B+}$ for $\alpha,\beta=t,r,\theta,\phi$ and

$$\Theta^{\overline{\alpha\beta}}_{\pm}=\sum_{\aleph}[f^{\overline{\alpha\beta}}_{1+\pm\aleph}(r_B)f^{\overline{\alpha\beta}}_{2+\pm\aleph}(\theta)]S_{\pm}(\theta), \tag{12}$$

we find that the boundary action (10) becomes

$$S^{(2)}_B=\frac{1}{\Delta^2_B}\int_{r=r_B}d^3x\sqrt{-g}N_r\Big\{[F^{-+*}_{++}(\theta)\Theta^{\overline{\alpha\beta}*}_{+}+F^{--*}_{+-}(\theta)\Theta^{\overline{\alpha\beta}*}_{-}](\Psi_B)_{\alpha\beta\gamma\delta}(r_B,\theta)$$
$$\times[F^{-+}_{++}(\theta)\Theta^{\overline{\gamma\delta}}_{+}+F^{--}_{+-}(\theta)\Theta^{\overline{\gamma\delta}}_{-}]\frac{R^*_-(r_B)R_-(r_B)}{R^*_{B+}R_{B+}}\mathcal{H}^{\overline{\alpha\beta}*}_{B+}(t,\phi)\mathcal{H}^{\overline{\gamma\delta}}_{B+}(t,\phi)+c.c.\Big\}. \tag{13}$$

In equation (13), $(\Psi_B)_{\alpha\beta\gamma\delta}(r_B,\theta)$ denotes

$$(\Psi_B)_{\alpha\beta\gamma\delta}(r_B,\theta)=\Big\{-\frac{g^{rr}}{8}\partial_r(g_{\alpha\gamma}g_{\beta\delta})+\frac{g^{rr}(\beta+3/2)}{8r_B}$$
$$\times\left(\frac{1}{2}g_{\alpha\beta}g_{\gamma\delta}-g_{\alpha\gamma}g_{\beta\delta}-\frac{g^r_\alpha g^r_\beta g_{\gamma\delta}}{g^{rr}}\right)+\frac{1}{8}\left(g^r_\alpha g^r_\beta-\frac{1}{2}g^{rr}g_{\alpha\beta}\right)g_{\delta\gamma,r}$$
$$-\frac{1}{16}g^{\sigma\lambda}g_{\sigma\lambda,r}g_{\alpha\beta}g^r_\gamma g^r_\delta+\frac{g^{rr}}{64}g^{\sigma\lambda}g_{\sigma\lambda,r}(g_{\alpha\beta}g_{\gamma\delta}-g_{\alpha\delta}g_{\beta\delta})\Big\}|_{r=r_B}, \tag{14}$$

and the $\theta$-dependent functions $F^{-+}_{++}$ and $F^{--}_{+-}$ are give by

$$F^{-+}_{++}(\theta)=\frac{\sum_{\aleph}f^{\overline{\mu\nu}}_{1-+\aleph}(r_B)f^{\overline{\mu\nu}}_{2-+\aleph}(\theta)}{\sum_{\aleph}f^{\overline{\mu\nu}}_{1++\aleph}(r_B)f^{\overline{\mu\nu}}_{2++\aleph}(\theta)}, \tag{15}$$

and

$$F^{--}_{+-}(\theta)=\frac{\sum_{\aleph}f^{\overline{\mu\nu}}_{1--\aleph}(r_B)f^{\overline{\mu\nu}}_{2--\aleph}(\theta)}{\sum_{\aleph}f^{\overline{\mu\nu}}_{1+-\aleph}(r_B)f^{\overline{\mu\nu}}_{2+-\aleph}(\theta)}, \tag{16}$$

respectively[5]. We can find now the two-point function of the boundary energy-momentum tensor operators $\mathcal{O}_{\alpha\beta}$ by taking the functional derivative of the action (10) with respect to the rescaled boundary gravitational field $\hat{\mathcal{H}}^{\alpha\beta}_{B+}$ as $<\mathcal{O}_{\alpha\beta}\mathcal{O}_{\gamma\delta}>=\frac{\delta^2S^{(2)}_B}{\delta\hat{\mathcal{H}}^{\alpha\beta}_{B+}\delta\hat{\mathcal{H}}^{\gamma\delta}_{B+}}$. We get the result as

$$<\mathcal{O}_{\alpha\beta}\mathcal{O}_{\gamma\delta}>=\frac{\delta^2S^{(2)}_B}{\delta\hat{\mathcal{H}}^{\alpha\beta}_{B+}\delta\hat{\mathcal{H}}^{\gamma\delta}_{B+}}=r_B^{2\beta-8}\frac{\left(R^B_-\right)^*R^B_-}{\left(R^B_+\right)^*R^B_+}\mathcal{F}_{\alpha\beta\gamma\delta}, \tag{17}$$

where all the factors are given explicitly in Ref. 5. Using the explicit form of Teukolsky radial functions, we find

$$< \mathcal{O}_{\alpha\beta}\mathcal{O}_{\gamma\delta} >= M^8 r_B^{2\beta-8}\mathcal{F}_{\alpha\beta\gamma\delta}\left\{1+M^{-2\beta}\frac{\mathcal{N}}{\mathcal{C}}G_R^*+M^{-2\beta}\frac{\mathcal{N}}{\mathcal{C}^*}G_R+\mathcal{N}M^{-4\beta}|G_R|^2\right\}, \tag{18}$$

where the function $G_R$ is given by[5]

$$G_R = T_R^{2\beta}\frac{\Gamma(-2\beta)\Gamma(\beta-3/2-im)\Gamma(1/2+\beta-i(n-m))}{\Gamma(2\beta)\Gamma(5/2-\beta-im)\Gamma(1/2-\beta-i(n-m))}, \tag{19}$$

and $\mathcal{C}=(\bar{\lambda}-2m^2-2im\beta)(\bar{\lambda}-2m^2-2im\beta+2)$. As we notice, the second and third terms in bracket in equation (18) are complex conjugate of each other and so we find a real-valued two-point function for the boundary conformal operators. However according to similar analysis in References 2 and 4, we can drop the second term in (18) to find a complex-valued two-point function for the boundary conformal operators. The main physical reason to drop the second term in (18) and so find a complex-valued two-point function, is that we can find the absorption cross section of spin 2 fields (using the optical theorem), only from a complex-valued two-point function (19). We also notice that the coefficient $\frac{\mathcal{N}}{\mathcal{C}^*}$ of term $G_R\mathcal{F}_{\alpha\beta\gamma\delta}$ in (18) depends on the momentum. However it is not a part of the two-point function (19). Moreover, similar momentum-dependent coefficients have been found for the two-point functions of other spins, such as spin-1/2 operators[2] as well as spin-1 operators[4]. Moreover, we note that the first term in bracket in equation (18) is just a constant contact term, compared to the third term that is proportional to $G_R$. In general, such contact terms can be cancelled by adding a proper counter term to the boundary action. The effect of adding the counter term is just to cancel out the constant contact term in (18) and we do not expect any changes in the two-point function $G_R\mathcal{F}_{\alpha\beta\gamma\delta}$. The last term in (18) is very small compared to the third term, since the ratio of the last term to the third term is proportional to $(\frac{\tau_H}{M})^{2\beta}$ where $\tau_H \to 0$[5]. Hence we conclude that the field theoretical two-point function for spin-2 conformal operators $\mathcal{O}_{\alpha\beta}$ on the boundary can be described by $G_R\mathcal{F}_{\alpha\beta\gamma\delta}$. The retarded Green's functions for fields with different spins were proposed in References 6 and 7. We note that $G_R$, as a part of two-point function $G_R\mathcal{F}_{\alpha\beta\gamma\delta}$ for the boundary conformal operators, is in perfect agreement with the proposed retarded Green's function for spin-2 fields. We also find that the absorption cross section of fields (with spin 2) is in exact agreement with the result of Reference 8 if we apply the optical theorem to the retarded Green's function (19).

The finite temperature two-point function of scalars[6,7], spin-1/2 fermions[2] as well as spin-1 fields[4] are given by

$$\langle\mathcal{O}\mathcal{O}\rangle \sim \left(\frac{\pi T_R}{\sinh\left(\pi T_R t_{12}^+\right)}\right)^{2h_R}\left(\frac{\pi T_L}{\sinh\left(\pi T_L t_{12}^-\right)}\right)^{2h_L}, \tag{20}$$

by proper identification of left and right conformal weights $h_L$ and $h_R$, where $1/T_L$ and $1/T_R$ are the circumferences of two fundamental circles of a torus. The Fourier

transform of the two-point function (20) is given by

$$\mathcal{F}(\langle\mathcal{O}\mathcal{O}\rangle) \sim {T_R}^{2\beta} \frac{\Gamma(1-2h_R)\,\Gamma(1-2h_L)}{\Gamma(1-h_R-in_R)\,\Gamma(1-h_R+in_R)\,\Gamma(1-h_L-in_L)\,\Gamma(1-h_L+in_L)}, \tag{21}$$

where $n_R$ and $n_L$ stand for $\frac{i\omega_R}{2\pi T_R}$ and $\frac{i\omega_L}{2\pi T_L}$ respectively. If we consider the conformal weights $h_R = \beta + 1/2$ and $h_L = \beta - 3/2$ for spin-2 operators, we then find

$$\mathcal{F}(\langle\mathcal{O}\mathcal{O}\rangle) \sim {T_R}^{2\beta} \frac{\Gamma(-2\beta)\,\Gamma(1/2+\beta-in_R)\,\Gamma(\beta-3/2-in_L)}{\Gamma(1/2-\beta-in_R)\,\Gamma(-\beta+5/2-in_L)\,\Gamma(2\beta)}. \tag{22}$$

Moreover, we consider $n_L = m$ and $n_R = n - m$ in (22) and find an exact agreement with the retarded Green's function (19) that we found as a part of two-point function for spin-2 operators.

## 3. Conclusions and Acknowledgements

The energy-momentum tensor operator of the CFT that is dual to the near extremal Kerr black hole, couples to the gravitational boundary fields with spin-2, on the near horizon of near extremal Kerr black hole. Hence we can find the two-point function of the energy-momentum tensor operator, by finding the explicit form of a boundary action that is bilinear in the boundary fields. In this article, we construct explicitly such a boundary action and vary the action with respect to the boundary fields. We derive the explicit expression for the two-point function and note that the two-point function factorizes into two terms. We show that the result for the two-point function is in agreement with correlation functions of spin-2 conformal operators in a dual CFT. We also note that one part of the result heavily depends on the gauge choice, that we made on fixing some of the Weyl scalars. Investigation about the gauge dependence of the two-point function would be an interesting project. Moreover, deriving the two-point function of the energy-momentum tensor operator with the primary fields of the dual CFT would be useful to establish a stronger field-theoretic approach to Kerr/CFT correspondence. The authoe would like to acknowledge the financial support by the Natural Sciences and Engineering Research Council of Canada.

## References

1. M. Guica, T. Hartman, W. Song and A. Strominger, *Phys. Rev. D* **80**, 124008 (2009).
2. M. Becker and W. Schulgin, *JHEP* **1204**, 063 (2012).
3. J.N. Laia and D. Tong, *JHEP* **1111**, 125 (2011).
4. A.M. Ghezelbash and H.M. Siahaan, *Phys. Rev. D* **89**, 024017 (2014).
5. A.M. Ghezelbash, *Gen. Relativ. Gravit.* **48**, 102 (2016).
6. B. Chen and C.S. Chu, *JHEP* **1005**, 004 (2010).
7. T. Hartman, W. Song and A. Strominger, *JHEP* **1003**, 118 (2010).
8. I. Bredberg, T. Hartman, W. Song and A. Strominger, *JHEP* **1004**, 019 (2010).

# Entropy bound for scalar fields in the near-horizon region

Kaushik Ghosh

*Vivekananda College (University of Calcutta),*
*269 Diamond Harbour Road,*
*Kolkata - 700063, India*
*ghosh_kaushik06@yahoo.co.in*

In this article, we will discuss a Lorentzian sector calculation of the entropy of a minimally coupled scalar field in the Schwarzschild black hole background using the brick wall model of G 't Hooft. In the original article, the Wentzel-Kramers-Brillouin (WKB) approximation was used. In this article, we will consider the entropy for a thin shell of matter field of a given thickness surrounding the black hole horizon. The thickness is chosen to be large compared with the Planck length and is of the order of the atomic scale. The corresponding leading-order entropy of the scalar field is found to be proportional to the area of the horizon and is logarithmically divergent. Thus, the entropy of a three-dimensional system in the near-horizon region is proportional to the boundary surface. The leading order entropy is a decreasing function of the the thickness of the thin shell. This is also valid for the sub-leading terms.

*Keywords*: Scwarzschild black hole, scalar field entropy, brick wall model, area law, logarithmic divergence, entropy bound.

## 1. Introduction

Since the four laws of black hole mechanics were formulated[1], there has been much effort to relate the laws of black hole mechanics to those of thermodynamics. The area theorem led Bekenstein to assign an entropy to a black hole[2]. Hawking established the thermodynamical aspects of black holes by showing that a black hole can radiate like a hot body at a temperature equal to a multiple of the surface gravity of the horizon[3]. The entropy of a black hole, considered as a thermodynamical system, was found to be $\frac{A}{4}$. Here $A$ is the area of the horizon.

We have to consider quantum field theory in curved spaces in most of the works related to black hole thermodynamics. In flat space, a uniformly accelerated observer detects a thermal spectrum when the field is in the Minkowski vacuum[4,5]. The temperature is dependent on the proper acceleration of the observer. A corresponding situation arises in the Schwarzschild black hole for a static observer outside the event horizon. The temperature is the same as that of the black hole if the static observer is at a large distance from the horizon[3,6]. In this article we will discuss the expression of the entropy of a minimally coupled massless scalar field in a Schwarzschild black hole in thermal equilibrium with the black hole. The expression is derived by us in a series of articles[7–9]. We have used a model proposed by G. 't Hooft known as the brick wall model and used the Wentzel−Kramers−Brillouin (WKB) approximation to count the states[10]. We can define a thermal equilibrium between a black hole and black body radiation of a massless scalar field surrounding the black hole at the same temperature as that of the black hole. The amount of radiation absorbed and emitted by the black hole is the same. Divergences appear

when we try to calculate the entropy of the scalar field[7,10]. This is associated with the continuous energy spectrum and unboundedness of the allowed angular quantum numbers in the near-horizon region. To regulate the divergence, 't Hooft proposed a boundary condition on the scalar field near the horizon[9]. He assumed the scalar field to be vanishing at a small distance away from the horizon. This radial parameter is known as the brick wall cut-off parameter. This boundary condition is a good model since in thermodynamic equilibrium there is no net interchange of particles between the black hole and the surrounding matter. The details can be found in[7,8]. The expression of the entropy is found to be logarithmically divergent in the brick wall cut-off parameter. The metric component $g_{rr}$ has a simple pole at the horizon. The WKB quantization rule indicates that the divergence is expected to be logarithmic. In the Schwarzschild black hole, we can also use the blueshift factor together with the form of the proper distance in the near-horizon region to understand this logarithmic divergence. The thickness of the thin shell takes care of the unbounded nature of the allowed angular quantum numbers. The expression of the entropy obtained by us is proportional to the area of the horizon when we take the temperature to be same as the Hawking temperature. The entropy is also found to be a decreasing function of the thickness of the thin shell. These aspects may be taken to be the semiclassical signs of a holography principle and entropy bound. Here the semiclassical theory refers to quantum field theory in a fixed background. The expressions obtained in this article are also significant in the context of entanglement entropy and the Euclidean approach to calculate the matter field entropy in a black hole background. We will discuss these aspects and the significance of the results obtained in this article in Sect. 2.

## 2. Entropy of a Thin Shell of Scalar Field in the Schwarzschild Black Hole

In this section we will consider the entropy of a minimally coupled scalar field in the Schwarzschild black hole. We will discuss the case of a massless field but the results can be easily extended for a massive case[7,8]. We will follow the method discussed in[7–9] and consider a thin shell of scalar field surrounding the horizon. We found earlier that in the spherical polar coordinates, the WKB approximation is suitable for a thin shell of scalar field[7,8]. Since the near-horizon region gives us interesting physics, we proceed to discuss the entropy of a thin shell of scalar field confined in the near-horizon region. The metric of the Schwaarzschild black hole is given by:

$$ds^2 = -\left(1 - \frac{2M}{r}\right) dt^2 + \frac{dr^2}{(1 - \frac{2M}{r})} + r^2(d\theta^2 + \sin^2\theta d\phi^2) \tag{1}$$

Here $M$ is the mass of the black hole defined with respect to an asymptotic observer. The wave equation of a minimally coupled massless scalar field in a curved space is given by the following expression:

$$(-g)^{-\frac{1}{2}} \partial_\mu [(-g)^{\frac{1}{2}} g^{\mu\nu} \partial_\nu \psi] = 0 \tag{2}$$

The wave equation is separable in the Schwarzschild black hole and we can take the basis function to be of the following form[7,8,10]:

$$f(l,m,E|x) = NR(r)P_l(\cos\theta)e^{im\phi}e^{-iEt} \tag{3}$$

Where, $l$ is the angular momentum quantum number and $E$ is the energy defined with respect to an asymptotic observer, $P_l(\cos\theta)$ are the Legendre polynomials and $R(r)$ satisfies the radial wave equation:

$$\left(1-\frac{2M}{r}\right)^{-1}E^2R(r)+\frac{1}{r^2}\partial_r[r(r-2M)\partial_rR(r)]-\frac{l(l+1)}{r^2}R(r)=0. \tag{4}$$

We use the WKB approximation to solve the above equation. This is a good approximation if we consider a thin shell of scalar field in the near-horrizon region[7–9]. We impose the brick wall boundary condition[10]:

$$R(2M+h)=0. \tag{5}$$

where, $h$ is a small radial parameter compared to $2M$ whose value will be discussed later. We can consider the scalar field to be confined in a half-infinite potential well in the near-horizon region[7,8]. The trial solutions are assumed to be of the form $\rho(r)\exp[iS(r)]$. The actual solutions are sinusoidal functions to satisfy the boundary conditions. In the near-horizon region, the radial wave function takes the following form:

$$R_l(r)\sim\exp\left[i\int k(r)dr\right] \tag{6}$$

Here, the $r$-dependent radial wave number is given by the following expression:

$$k(r)=\frac{1}{V(r)}\sqrt{E^2-\frac{V(r)}{r^2}l(l+1)} \tag{7}$$

To implement the thin shell boundary condition, we consider solutions that vanish at the brick wall and are stationary up to a certain value $'d'$ of the radial coordinate. We have the condition: $h << d << 2M$. The radial quantum number is given by the WKB quantization rule:

$$\pi n_d=\int_{2M+h}^{2M+d}dr\ k(r,l,E)\ . \tag{8}$$

To evaluate $n_d$, we have to ensure that the wave number is real throughout the range of the radial integral for each value of $E, l$ present in its argument. This gives an upper limit on the allowed values of $l$ given by[7,8]:

$$L(L+1)=\frac{E^2(2M+d+h)^3}{d+h}\approx\frac{8E^2M^3}{d} \tag{9}$$

The free energy of the thin shell can be obtained from the expression:

$$\pi\beta F = -\int_0^\infty \frac{\beta dE}{[\exp(\beta E)-1]} \int_{2M+h}^{2M+d} \frac{dr}{V(r)} \int_0^L dl(2l+1)\sqrt{E^2 - \frac{V(r)}{r^2}[l(l+1)]}. \tag{10}$$

The entropy of a scalar field confined within the near horizon region and as predicted by an observer at infinity is given by[8]:

$$S = \frac{1}{30}\ ln\left(\frac{\epsilon_d}{\epsilon_h}\right)\ \frac{A}{4\pi{\epsilon_d}^2}\ . \tag{11}$$

The corresponding expression in the Kerr black hole is given by the following expression[7]:

$$S' = \frac{1}{60}\ ln\left(\frac{\epsilon_d}{\epsilon_h}\right)\ \frac{A}{4\pi{\epsilon_d}^2}\ . \tag{12}$$

The temperature in both the cases are taken to be the Hawking temperature of the corresponding black holes. Here $\epsilon_x$ represents the proper radial variable: $\epsilon_x = \int_{r_h}^{r_h+x}\sqrt{g_{rr}}\ dr$, $r_h$ is the event horizon. 'A' represents the area of the horizon. The entropy of the scalar field in the Kerr black hole is half of the corresponding value in the Schwarzschild black hole. This is due to rotation and is consistent with the preferential emission of particles in the Kerr black hole with azimuthal angular momentum in the same direction as that of the black hole itself. The Lorentzian sector calculation of the scalar field entropy in the Kerr black hole is more significant since the corresponding Euclidean sector literature is not well formulated[11]. The expressions of the scalar field entropy contain two free parameters, the proper thickness of the thin shell and the proper brick wall cut-off parameter. We take the proper thickness of the thin shell to be of the order of atomic lengths. The other undetermined parameter is the proper brick wall cut-off. 't Hooft wanted to explain the black hole entropy in terms of the near horizon part of the matter field entropy and equated the two[9]. If we do so, the internal energy turns out to be finite: $U = \frac{3}{8}M$. This also determines the proper brick wall cut-off parameter. In the present case we then have:

$$\epsilon_h = (\epsilon_d)\exp\left[-30\pi\ \left(\frac{\epsilon_d}{l_p}\right)^2\right]. \tag{13}$$

Where, $l_p$ is the Plack length. If we take $\epsilon_d = 10^{-10}\ cm$, $\epsilon_h$ is given by the following expression in terms of the Planck length: $\epsilon_h \sim [10^{24}]\exp(-10^{50})(l_p)$. Thus, the brick wall is almost coincident with the horizon although the free energy is finite. This is expected since there exists solutions vanishing on the horizon[7,8]. These solutions are consistent with the brick wall boundary condition. We note that even in the extreme case of a thin shell of thickness comparable to the Planck length, the proper brick wall cut-off remains small compared to the thickness. However,

in this case the semiclassical theory of quantum fields in a fixed background may not be appropriate. The back reaction problem is discussed in detail by Candelas and is not much important in the present case[12]. These discussions may become significant in the context of an entropy bound discussed below.

The scalar field entropy in a Schwarzschild black hole with off mass shell temperature is given by:

$$S = \frac{32s\pi^3}{15{\epsilon_d}^2}(2M)^2\left(\frac{2M}{\beta}\right)^3 ln\left(\frac{\epsilon_d}{\epsilon_h}\right). \tag{14}$$

We find that the entropy is a decreasing function of the proper thickness even when the temperature is not given by the Hawking temperature. We now consider the subleading terms of the scalar field entropy. It can be seen that in the approximation $h << d << 2M$, the complete entropy is given by the following expression,[7]:

$$S = \frac{1}{30\pi{\epsilon_d}^2}\left[ln\left(\frac{\epsilon_d}{\epsilon_h}\right) - \chi\right]\frac{A}{4}, \tag{15}$$

where, $\chi = \frac{1}{4}[1 + \frac{1}{12} + ...] > 0$ and the temperature is given by the Hawking temperature. For $d >> h$, the second term is negligibly small compared with the first term. The complete expression is also a decreasing function of the proper thickness of the thin shell. The leading-order scalar field entropy in a Kerr black hole with off mass shell temperature is given by:

$$S = \frac{16\pi^3}{15}\ \frac{1}{\beta^3}\ \frac{({r_+}^2 + a^2)^4}{(r_+ - r_-)^3}\ \frac{1}{{\epsilon_d}^2}\ ln\left(\frac{\epsilon_d}{\epsilon_h}\right). \tag{16}$$

Where, $a$ is the angular momentum per unit mass, $r_+$ is the event horizon and $r_-$ is the inner horizon of the black hole. We again find that the entropy is a decreasing function of the proper thickness even when the temperature is not given by the Hawking temperature. The leading order entropy is a decreasing function of the proper thickness when the temperature is given by the corresponding Hawking temperature. This is obvious from Eq. (12).

We find that the entropy is a decreasing function of the proper thickness even when the temperature is not given by the Hawking temperature. Thus, we may conjecture that *the entropy of a scalar field in an asymptotically flat black hole background which is correlated with the horizon through the brick wall boundary condition has an upper bound and this upper bound is proportional to the horizon surface area when the temperature is taken to be the Hawking temperature.* It may require a quantum theory of gravity to fix the exact magnitude of the upper bound. In this article we have chosen the upper bound to be the same as the black hole entropy itself. This conclusion is expected to remain valid for the semiclassical theory of quantum fields in curved spaces where back reaction problem is not severe. We can further verify the validity of the present conclusion by considering the spectrum of the solutions of the wave equation defined throughout the manifold and satisfying the brick wall boundary condition. We can use numerical methods for this purpose. These discussions are also important to the holographic principle

in asymptotically flat spaces[13]. Similar discussions will remain valid in presence of the cosmological horizons. The event horizons are two dimensional fixed point sets of the Killing fields[14]. This aspect may be important to explain the results obtained in this article.

The logarithmic divergence in Eq. (11) and Eq. (12) is important to the Euclidean sector calculation of the scalar field entropy[11]. The logarithmic divergence is expected on physical grounds as discussed briefly in the introduction and in detail in[7,8]. The Euclidean sector method contains a parameter which can be fixed by comparing with the Lorentzian sector expression obtained in the present article. Lastly, we can consider the scalar field entropy to give quantum correction to the black hole entropy and use Eq. (11) and Eq. (12) to renormalize the gravitational constant[15].

## References


1. J. M. Bardeen, B. Carter and S. W. Hawking, Commun. Math. Phys. **31**, 161 (1973).
2. J. Bekenstein, Phys. Rev. D **7**, 2333 (1973).
3. S. W. Hawking, Commun. Math. Phys. **43**, 199 (1975).
4. S. A. Fulling, Phys. Rev. D **7**, 2850 (1973).
5. W. G. Unruh, Phys. Rev. D **14**, 870 (1976).
6. J. B. Hartle and S. W. Hawking, Phys. Rev. D **13**, 2188 (1976).
7. K. Ghosh, J. Phys. Soc. Japan **85**, 014101, (2016).
8. K. Ghosh, Prog. Theor. Exp. **2016**, 093E03
9. DAE
10. G. 't Hooft, Nucl. Phys. B **256**, 727 (1985).
11. S. N. Solodukhin, Living Rev. Relativity **14**, 8 (2011).
12. P. Candelas, Phys. Rev. D **21**, 2185 (1980).
13. R. Bousso, Rev. Mod. Phys. **74**, 2002.
14. G. W. Gibbons and S. W. Hawking, Commun. Math. Phys. **66**, 291 (1979).
15. L. Susskind and J. Uglum, Phys. Rev. D **50**, 2700 (1994).

# Thermodynamic analysis of a self-gravitating gas in astrophysical contexts

Christine Gruber*

*Hanse-Wissenschaftskolleg, Delmenhorst, Germany*
*and*
*Institute of Physics, Carl v. Ossietzky University of Oldenburg, Germany*
**E-mail: christine.gruber@uni-oldenburg.de*

The thermodynamics of a self-gravitating gas cloud of particles interacting only via their gravitational potential is an interesting problem with peculiarities arising due to the long-ranged nature of the gravitational interaction. Based on our recent work on the properties of such a configuration, we extend the system to contain a central gravitational field in which the particles are moving, to mimic the potential of a central compact object exerting an external force on the gas cloud. After an introduction to the general problem, including the aforementioned peculiarities and possible solutions, we will discuss the particular properties of the self-gravitating gas in a central field and its thermodynamic analysis.

*Keywords*: Thermodynamics, statistical mechanics; self-gravitating system; central potential.

## 1. Introduction

The topic of thermodynamic systems in the presence of gravity has been discussed in many occasions and forms[1–3], and has raised many questions on how to deal with the long-range effects of gravity in the thermodynamic analysis of systems, where concepts such as the isolation of a system in one or the other regard are important. Non-stationary equilibrium situations, negative heat capacities or simply divergences in the thermodynamic limit have been plaguing the analyses, and the conventional, very successful thermodynamic framework of Boltzmann-Gibbs statistics had to be adapted and modified in order to account for the peculiarities of the thermodynamics of a gravitational system.

Based on a Boltzmann-Gibbs analysis of the self-gravitating gas[4], the statistical analysis and subsequent calculation of thermodynamic properties have been carried out[5] assuming a generalized framework intended to describe a system with non-extensive properties, due to the presence of long-range forces such as gravity. The adopted generalized framework, i.e., Tsallis generalized $q$-statistics, has been developed in order to consider non-extensive effects, entailing an additional parameter $q$ in the statistical analysis.

This work is an extension of these previous investigations which generalizes the self-gravitating gas to a more realistic system featuring a centrally placed compact object, like e.g. a black hole, around which the gas is extending. Due to some peculiarities and open questions, we will not continue using the non-extensive $q$-statistics, but rather return to the conventional Boltzmann-Gibbs statistics, in order to get a first impression of the results. Other generalizations can be thought of, which will be commented on in the last section.

## 2. Statistical mechanics and thermodynamics of a self-gravitating gas

I will briefly review the most important steps in the analysis of a self-gravitating gas[4,5], from the system's properties to the peculiarities of the thermodynamic analysis and some of its outcomes.

The governing force of the self-gravitating gas is the gravitational attraction between its $N$ identical constituent particles which are otherwise moving freely, and thus the Hamiltonian of the system is

$$\mathcal{H} = \mathcal{T} + \mathcal{U} = \sum_{i=1}^{N} \frac{p_i^2}{2m} - Gm^2 \sum_{1 \le i < j \le N} \frac{1}{|\mathbf{q_i} - \mathbf{q_j}|_A}\,, \tag{1}$$

where $G$ is the gravitational constant, $m$ the mass of an individual particle, and $A$ represents a short-range cutoff imposed in order to avoid the unphysical collapse of the system to a point. This Hamiltonian is the basis for a thermodynamic analysis which can be done in principle in different ensembles, like the microcanonical one, where the energy of the system is kept constant, or the canonical one, where instead the temperature is fixed, and energy can be exchanged with a reservoir. In the microcanonical ensemble, the most important thermodynamic quantity from which everything is derived is the entropy, i.e., the logarithm of this sum over microstates $\Omega(E, V, N)$,

$$S = k_B \ln \Omega(E, V, N)\,, \tag{2}$$

where

$$\Omega = \frac{(2\pi m)^{3N/2}}{N!\, h^{3N} \Gamma\left(\frac{3N}{2} + 1\right)} \int d^{3N}q \, \left[E - \mathcal{U}\right]^{3N/2-1}. \tag{3}$$

From the entropy, you can obtain important thermodynamic quantities such as the temperature of the gas, or the equation of state, i.e., the relation between pressure, temperature and volume.

In the case of the canonical ensemble, the starting point is the partition function, defined as

$$Z = \frac{1}{N! h^{3N}} \int d^{3N}q \, d^{3N}p \, \exp_q\left(-\beta \mathcal{H}(\mathbf{p}, \mathbf{q})\right), \tag{4}$$

and everything else is derived from that quantity, like the equation of state. The temperature in this ensemble is fixed, so it cannot be calculated.

Following the definition of these basic thermodynamic functions, calculations can be simplified by the assumption of a weak gravitational interaction, i.e., the gravitational contribution can be treated as a small correction to the ideal gas, and results can be obtained analytically in this case.

Under this additional assumption, further quantities that are of thermodynamic interest can be calculated, like the heat capacity and other response functions of the system. Both equation of state and heat capacity have been calculated and

compared in the framework of Tsallis statistics[5], and the details can be found there.

An important point for further investigations is the question of the statistical framework, which is closely connected to the choice of thermodynamic limit. Tsallis' non-extensive statistics naturally features a modification of the thermodynamic limit, in which the thermodynamic state variables result in convergent functions. In the case the conventional Boltzmann-Gibbs statistics, another modification of the thermodynamic limit has to be adopted[4] in order to obtain convergent results. In the following, we will employ Boltzmann-Gibbs statistics with the modified thermodynamic limit.

## 3. Addition of the central gravitational potential

As a modification to the basic setup of a simple self-gravitating gas many complications can be thought of. The simplest gase perhaps is the addition of a central potential, to model the situation of a self-gravitating gas around a black hole. We will start with the assumption of an external gravitational field caused by a mass $M$ of size $r_S = 2GM/c^2$ in the center of the configuration, restricting the movement of the gas between the radius of the innermost stable circular orbit (ISCO) at $r_{ISCO} = 3r_S$ and infinity. This will make a difference in the integrals contained in the sum over microstates and the partition function, respectively. Moreover, the central potential will have its influence on every particle in the gas. The generalized Hamiltonian thus reads

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m} - Gm^2 \sum_{1\leq i<j\leq N} \frac{1}{|\mathbf{q_i} - \mathbf{q_j}|_A} - GmM \sum_{1\leq i<j\leq N} \frac{1}{|\mathbf{q_i} - \mathbf{r}|_A}\,, \qquad (5)$$

with $\mathbf{r}$ denoting the center of mass of the compact object. To simplify calculations, we choose $\mathbf{r} = 0$.

The computational procedures in order to extract the thermodynamic equation of state is fairly analogous to the case of a simple self-gravitating gas, and differs only in the restriction of the range of integration, due to the fact that we consider a ring-like structure, or even a flat two-dimensional disk shape. This restriction will manifest itself in the definition of the virial coefficients $b_i$, which will be slightly different.

The interesting question is whether the modification of the system will lead to differences in the thermodynamic limit, i.e., facilitate the calculation of otherwise divergent functions, or modify the qualitative dependence on the number of particles in the thermodynamic limit. Preliminary results indicate that this is not the case, and that modifications are limited to the virial coefficients of the problem.

## 4. Outlook

We have here discussed the thermodynamic properties of a self-gravitating gas under the influence of a central gravitational field caused by a heavy mass at the

center of the configuration. Basing on the analysis of a self-gravitating gas cloud consisting of ideal particles[4,5], an additional term accounting for the central gravitational potential was added to the analysis, and the resulting thermodynamic state variables were calculated. Preliminary results indicate slight modifications of the state variables, depending on the new parameter, the mass of the central object. The goal is to generalize the analysis of a simple self-gravitating gas to eventually be able to make predictions on the thermodynamic behavior of matter around a compact object, i.e., an accretion disk of sorts.

Besides the inclusion of a central compact object, the gas itself can be modified in its properties, e.g., by considering non-ideal interactions between the particles. This could be accounted for in an exact way by adding additional particle-particle interactions to the Hamiltonian, with the corresponding coupling constant, like for example an electromagnetic charge. The different strengths of gravity and any other interactions that may be added have to be weighed against each other, and approximations could be applied. Another possibility would be to include effective potentials which are used in condensed matter systems, Mie-type potential like the Lennard-Jones case or others, in order to describe different variations of the gas. Investigations in this direction would represent the first steps towards the description of non-ideal fluids in gravitational contexts - either gas clouds of interacting particles, or non-ideal fluids constituting accretion disks or clouds around a central compact object.

Further generalizations include rotation of the system, or charge of the central object. Importantly, these results should then be connected to results of other calculations on accretion disks, in particular the accretion of (charged) dust particles in a spherical shell or torus structure[6].

## Acknowledgments

C. G. was supported by funding from the DFG Research Training Group 1620 'Models of Gravity'.

## References

1. D. Lynden-Bell and R. Wood, *Mon. Not. Roy. Astron. Soc.* **138**, p. 495 (1968).
2. J. Oppenheim, *Phys. Rev. E* **68**, p. 016108 (2003).
3. A. Campa, T. Dauxois and S. Ruffo, *Phys. Rep.* **480**, 57 (2009).
4. H. J. de Vega and N. Sanchez, *Nuclear Physics B* **625**, 409 (2002).
5. L. F. Escamilla-Herrera, C. Gruber, V. Pineda-Reyes and H. Quevedo, *Submitted for publication* (2018).
6. K. Schroven, E. Hackmann and C. Lämmerzahl, *Phys. Rev. D* **96**, p. 063015 (2017).

# Hawking radiation as quantum tunneling phenomenon

Wajiha J. and Rimsha B.

*Department of Mathematics, Division of Science & Technology, University of Education, Lahore-54590, Pakistan*
*wajiha.javed@ue.edu.pk; wajihajaved84@yahoo.com; rimsha.babar10@gmail.com*

We explore the Hawking radiation phenomenon as a tunneling process of charged spin particles through event horizons of some particular black holes. Using the semi-classical WKB approximation to the general covariant wave equations of charged particles, we evaluate the tunneling probabilities of outgoing charged particles. The Hawking temperatures corresponding to these black holes are recovered. By considering the back-reaction effects of the emitted spin particles from black holes, We calculate their corresponding quantum corrections to the radiation spectrum. It is found that this radiation spectrum is not purely thermal due to energy and charge conservation but has some corrections.

*Keywords*: Hawking Radiation; Spin-1 particles; Quantum Corrections.

## 1. Introduction

A realm of space where gravity is so intense that nothing not even light can escape is known as Black Hole (BH). This idea laid by the astronomical pioneer John Michell (1784).[1] The outer surface of a BH is known as *event horizon*. Stephen Hawking (1974) evoked that BHs are not totally black, but they emit thermal radiation,[2] known as *Hawking radiation*. The mass and energy of BH decreases due to the continuous emission of these Hawking radiation, as a result eventually BH evaporates, this phenomenon is know as BH evaporation.

A number of techniques have been introduced to investigate the Hawking radiation spectrum. By taking into account the semi-classical effects, the quantum tunneling approach[3,4] is the most convenient strategy to study these BHs radiation. Quantum tunneling is a process, by which particles have finite probability to cross the event horizon. In order to understand this process their are two main approaches, one of them is *null geodesic technique* proposed by Kraus & Wilczek,[5] while the other one is *Hamilton-Jacobi technique* propounded by Angheben et al.[6] These techniques gives the tunneling probability, i.e., the estimated rate of emitted particles accompanying the classically forbidden trajectories from inside to the outside region of the BH through event horizon. The tunneling probability, $\Gamma$, of outgoing radiated particles can be defined as

$$\Gamma = \exp\left(-\frac{2}{\hbar} Im\mathcal{S}\right), \tag{1}$$

where $\mathcal{S}$ stands for the action of radiated particle and $\hbar$ represents the Planck's constant. Sharif and Javed[7] studied the Hawking radiation process through the quantum tunneling phenomenon for fermion particles from various types of BHs, also they analyzed quantum corrections for regular BHs. The quantum gravity

effects by using generalized uncertainty principle (GUP) have been discussed in literature.[8–14] Javed and Babar[15] investigated Hawking temperature and corrected Hawking temperature for fermions particles. Moreover, Javed et al.[16] analyzed the corrected temperature for spin-0 and spin-1/2 particles for cubic and quartic interactions. In the continuation of previous work, we are going to provide a general formulation of vector particles tunneling for general spherically symmetric rotating BH and investigate its original and corrected Hawking temperature.

## 2. Charged Vector Particles Tunneling

The metric can be defined as

$$ds^2 = -Adt^2 + \frac{1}{B}dr^2 + Cd\theta^2 + Dd\varphi^2 - 2Fdtd\varphi, \tag{2}$$

where the metric functions $A, B, C, D$ and $F$ are depend upon $r$ and $\theta$. One can obtain the event horizon $r_+$ of this BH by considering the horizon equation, $A(r_+, \theta) = 0$. The angular velocity of BH at event horizon is given by

$$\Omega_H = \frac{F(r_+, \theta)}{D(r_+, \theta)}. \tag{3}$$

In order to calculate the tunneling probability of charged spin-1 particles at horizon $r_+$, the required Proca equation can be expressed as[17]

$$\frac{1}{\sqrt{-g}}\partial_\mu\left(\sqrt{-g}\Psi^{\nu\mu}\right) + \frac{\iota}{\hbar}e\tilde{F}^{\nu\mu}\Psi_\mu + \frac{\iota}{\hbar}eA_\mu\Psi^{\nu\mu} + \frac{m^2}{\hbar^2}\Psi^\nu = 0, \tag{4}$$

here $g$ is the determinant of the coefficient matrix and $m$ represents the mass of emitted particle, the $\Psi^{\mu\nu}$ and $\tilde{F}^{\nu\mu}$ are anti-symmetric & electromagnetic field tensors, respectively, which are given by

$$\Psi_{\nu\mu} = \partial_\nu\Psi_\mu - \partial_\mu\Psi_\nu - \frac{\iota}{\hbar}eA_\mu\Psi_\nu + \frac{\iota}{\hbar}eA_\nu\Psi_\mu\ , \qquad \tilde{F}^{\nu\mu} = \nabla^\nu A^\mu - \nabla^\mu A^\nu, \tag{5}$$

where $A_\mu$ is electromagnetic vector potential of the given BH, $e$ denotes the charge of the radiated $W$-bosons and $\nabla_\mu$ represents the geometry of covariant derivative. The equation of motion is same for both outgoing $W^+$ and incoming $W^-$ bosons, therefore the tunneling rate for both cases is also similar. Utilizing, the following WKB approximation for the given BH[6]

$$\Psi_\nu = c_\nu \exp\left[\frac{\iota}{\hbar}\mathcal{S}(t, r, \theta, \varphi) + \Sigma\hbar^n\mathcal{S}_n(t, r, \theta, \varphi)\right], \tag{6}$$

and the Proca equation (4), as well as by considering only the leading order terms in $\hbar$, we obtain the following set of equations

$$\begin{aligned}
DB\ & [c_1(\partial_1\mathcal{S})((\partial_0\mathcal{S})+eA_0)-c_0(\partial_1\mathcal{S})^2]+FB[c_3(\partial_1\mathcal{S})^2-c_1(\partial_1\mathcal{S})((\partial_3\mathcal{S})\\
+\ & eA_3)]+DC^{-1}[c_2(\partial_2\mathcal{S})((\partial_0\mathcal{S})+eA_3)-c_0(\partial_2\mathcal{S})^2]+FC^{-1}[c_3(\partial_2\mathcal{S})^2\\
-\ & c_2(\partial_2\mathcal{S})((\partial_3\mathcal{S})+eA_3)]+[c_3(\partial_3\mathcal{S})((\partial_0\mathcal{S})+eA_0)-c_0(\partial_3\mathcal{S})((\partial_3\mathcal{S})\\
+\ & eA_3)]+eA_3[c_3(\partial_0\mathcal{S})+eA_3c_3-c_0(\partial_3\mathcal{S})-eA_3c_o]+m^2Dc_0\\
+\ & m^2Fc_3=0, \qquad (7)\\
DB\ & [c_1(\partial_0\mathcal{S})((\partial_0\mathcal{S})+eA_0)-c_0(\partial_1\mathcal{S})(\partial_0\mathcal{S})]+FB[c_3(\partial_1\mathcal{S})(\partial_0\mathcal{S})\\
-\ & c_1(\partial_0\mathcal{S})((\partial_3\mathcal{S})+eA_3)]+BC^{-1}(BD+F^2)[c_2(\partial_1\mathcal{S})(\partial_2\mathcal{S})-c_1(\partial_2\mathcal{S})^2]\\
+\ & AB[c_3(\partial_1\mathcal{S})(\partial_3\mathcal{S})-c_1(\partial_3\mathcal{S})((\partial_3\mathcal{S})+eA_3)]+m^2Bc_1(BD+F)\\
-\ & BF[c_1(\partial_3\mathcal{S})((\partial_0\mathcal{S})+eA_0)-c_0(\partial_1\mathcal{S})(\partial_3\mathcal{S})]+eA_0BD[c_1((\partial_0\mathcal{S})\\
+\ & eA_0)-c_0(\partial_1\mathcal{S})]+eA_0BF[c_3(\partial_1\mathcal{S})-c_1((\partial_3\mathcal{S})-eA_3)]\\
+\ & eA_3AB[c_3(\partial_1\mathcal{S})-c_1((\partial_3\mathcal{S})+eA_3)]-eA_3BF[c_1((\partial_0\mathcal{S})\\
+\ & eA_0)-c_0(\partial_1\mathcal{S})]=0, \qquad (8)\\
DC^{-1}\ & \left[c_2(\partial_0\mathcal{S})^2-c_0(\partial_2\mathcal{S})(\partial_0\mathcal{S})+eA_0(\partial_0\mathcal{S})c_2\right]+FC^{-1}[c_3(\partial_0\mathcal{S})(\partial_2\mathcal{S})\\
-\ & c_2(\partial_0\mathcal{S})(\partial_3\mathcal{S})-eA_3c_2(\partial_0\mathcal{S})]-m^2c_2C^{-1}(BD+F)\\
+\ & AC^{-1}[c_3(\partial_2\mathcal{S})(\partial_3\mathcal{S})-c_2(\partial_3\mathcal{S})^2-eA_3c_2(\partial_3\mathcal{S})]-FC^{-1}[(\partial_3\mathcal{S})(\partial_0\mathcal{S})c_2\\
-\ & c_0(\partial_2\mathcal{S})(\partial_3\mathcal{S})+c_2eA_0(\partial_3\mathcal{S})]+eA_0DC^{-1}[c_2(\partial_0\mathcal{S})-c_0(\partial_2\mathcal{S})+eA_0c_2]\\
+\ & eA_0FC^{-1}[c_3(\partial_2\mathcal{S})-c_2(\partial_3\mathcal{S})-c_2eA_3]+eA_3AC^{-1}[c_3(\partial_2\mathcal{S})-c_2(\partial_3\mathcal{S})\\
-\ & eA_3c_2]-eA_3FC^{-1}[c_2(\partial_0\mathcal{S})-c_0(\partial_2\mathcal{S})+eA_0c_2]\\
-\ & BC^{-1}(BD+F)[c_2(\partial_1\mathcal{S})^2-c_1(\partial_1\mathcal{S})(\partial_2\mathcal{S})]=0, \qquad (9)\\
BF\ & [c_1(\partial_0\mathcal{S})(\partial_1\mathcal{S})-c_0(\partial_1\mathcal{S})^2+eA_0c_1(\partial_1\mathcal{S})]+FC^{-1}[c_2(\partial_0\mathcal{S})(\partial_2\mathcal{S})\\
-\ & c_0(\partial_2\mathcal{S})^2+eA_0c_2(\partial_2\mathcal{S})]-AB[c_3(\partial_1\mathcal{S})^2-c_1(\partial_1\mathcal{S})(\partial_3\mathcal{S})\\
-\ & eA_3c_1(\partial_1\mathcal{S})]-AC^{-1}[c_3(\partial_2\mathcal{S})^2-c_2(\partial_2\mathcal{S})(\partial_3\mathcal{S})-eA_3c_2(\partial_2\mathcal{S})]\\
+\ & eA_0[c_3(\partial_0\mathcal{S})-c_0(\partial_3\mathcal{S})+eA_0c_3-eA_3c_0]+[c_3(\partial_0\mathcal{S})^2\\
-\ & c_0(\partial_3\mathcal{S})(\partial_0\mathcal{S})+eA_0c_3(\partial_0\mathcal{S})-eA_3c_0(\partial_0\mathcal{S})]+FC^{-1}[c_2(\partial_0\mathcal{S})(\partial_2\mathcal{S})\\
-\ & c_0(\partial_2\mathcal{S})^2+eA_0c_2(\partial_2\mathcal{S})]+m^2[Fc_0-c_3A]=0. \qquad (10)
\end{aligned}$$

By considering the separation of variables technique, one can write the particle's action in the following form

$$\mathcal{S}=-(E-j\Omega_H)t+\mathcal{R}(r)+j\varphi+\Theta(\theta), \qquad (11)$$

here $j$ and $E$ are the angular momentum and energy of radiated particle, respectively. The above set of Eqs. (7)–(10) can be describe in the form of following $4\times4$ matrix equation, i.e.,

$$\bigwedge(c_0,c_1,c_2,c_3)^t=0. \qquad (12)$$

In order to obtain the following non-trivial solution, we take $|\bigwedge| = 0$

$$\mathcal{R}^{\pm} = \pm\sqrt{\frac{(E - j\Omega_H - eA_0)^2 + Y}{AB}}dr. \tag{13}$$

The above equation implies the imaginary part of radial component of particle's action, i.e.,

$$Im\mathcal{R}^{\pm} = \pm\pi\frac{(E - j\Omega_H - eA_0)}{\sqrt{A_r(r_+,\theta)B_r(r_+,\theta)}}, \tag{14}$$

where the function $Y$ is given as follows

$$Y = -AN^2C^{-1} - m^2A + BF(E - j\Omega_H) - AB\dot{\Theta} - eA_0BF - eA_3AB.$$

The tunneling rate of charged spin-1 particles can be calculated as

$$\Gamma = \exp\left[-\frac{4}{\hbar}Im\mathcal{R}^{+}\right] = \exp\left[\frac{-4\pi(E - eA_0 - j\Omega_H)}{\sqrt{A_r(r_+,\theta)B_r(r_+,\theta)}}\right]. \tag{15}$$

By utilizing Boltzmann formula $\Gamma_B = \exp[(E - eA_0 - j\Omega_H)/T_H]$, the expected Hawking temperature $T_H$ can be obtained as

$$T_H = \left[\frac{\sqrt{A_r(r_+,\theta)B_r(r_+,\theta)}}{4\pi}\right]. \tag{16}$$

## 3. Quantum Corrections

The GUP-corrected equation of motion of massive uncharged $W$ bosons is given by[18]

$$\partial_\mu(\sqrt{-g}\Psi^{\mu\nu}) - \sqrt{-g}\frac{m_W^2}{\hbar^2}\Psi^{\nu} + \tilde{\beta}\hbar^2\partial_0\partial_0\partial_0(\sqrt{-g}g^{00}\Psi^{0\nu}) - \tilde{\beta}h^2\partial_i\partial_i\partial_i(\sqrt{-g}g^{ii}\Psi^{i\nu}) = 0, \tag{17}$$

where $m_W$ is vector particle's mass, the modified anti-symmetric tensor $\Psi^{\mu\nu}$ is defined below

$$\Psi^{\mu\nu} = (1 - \tilde{\beta}\hbar^2\partial_\mu^2)\partial_\mu\Psi_\nu - (1 - \tilde{\beta}\hbar^2\partial_\nu^2)\partial_\nu\Psi_\mu.$$

Using coordinate transformation, the metric can be rewritten as follows

$$ds^2 = -Adt^2 + \frac{1}{B}dr^2 + Cd\theta^2 + Dd\varphi^2. \tag{18}$$

The WKB approximation for the vector field $\Psi^{\mu}$ is given by

$$\Psi^{\mu} = c_\mu(t, r, \theta, \varphi)\exp\left[\frac{\iota}{\hbar}\mathcal{S}(t, r, \theta, \varphi)\right]. \tag{19}$$

Substituting values from Eqs. (18) and (19) in Eq. (17) and neglecting the higher order terms in $\hbar$, we obtain the following set of equations

$$B\,[c_0(\partial_r\mathcal{S})^2P_1^2 - c_1(\partial_r\mathcal{S})(\partial_t\mathcal{S})P_1P_0] + \frac{1}{C}[c_0(\partial_\theta\mathcal{S})^2P_2^2 - c_2(\partial_\theta\mathcal{S})(\partial_t\mathcal{S})P_2P_0]$$
$$+\ \frac{1}{D}[c_0(\partial_\varphi\mathcal{S})^2P_3^2 - c_3(\partial_\varphi\mathcal{S})(\partial_t\mathcal{S})P_3P_0] + c_0m_W^2 = 0, \tag{20}$$
$$-\ \frac{1}{A}[c_1(\partial_t\mathcal{S})^2P_0^2 - c_0(\partial_t\mathcal{S})(\partial_r\mathcal{S})P_0P_1] + \frac{1}{C}[c_1(\partial_\theta\mathcal{S})^2P_2^2 - c_2(\partial_\theta\mathcal{S})(\partial_r\mathcal{S})P_2P_1]$$
$$+\ \frac{1}{D}[c_1(\partial_\varphi\mathcal{S})^2P_3^2 - c_3(\partial_\varphi\mathcal{S})(\partial_r\mathcal{S})P_3P_1] + c_1m_W^2 = 0, \tag{21}$$
$$-\ \frac{1}{A}[c_2(\partial_t\mathcal{S})^2P_0^2 - c_0(\partial_t\mathcal{S})(\partial_\theta\mathcal{S})P_0P_2] + B[c_2(\partial_r\mathcal{S})^2P_1^2 - c_1(\partial_r\mathcal{S})(\partial_\theta\mathcal{S})P_1P_2]$$
$$+\ \frac{1}{D}[c_2(\partial_\varphi\mathcal{S})^2P_3^2 - c_3(\partial_\varphi\mathcal{S})(\partial_\theta\mathcal{S})P_3P_2] + c_2m_W^2 = 0, \tag{22}$$
$$-\ \frac{1}{A}[c_3(\partial_t\mathcal{S})^2P_0^2 - c_0(\partial_t\mathcal{S})(\partial_\varphi\mathcal{S})P_0P_3] + B[c_3(\partial_r\mathcal{S})^2P_1^2 - c_1(\partial_r\mathcal{S})(\partial_\varphi I_1)P_2P_3]$$
$$+\ \frac{1}{C}[c_3(\partial_\theta\mathcal{S})^2P_2^2 - c_2(\partial_\theta\mathcal{S})(\partial_\varphi\mathcal{S})P_2P_3] + c_3m_W^2 = 0, \tag{23}$$

where

$$P_0 = 1 + \tilde{\beta}\frac{1}{A}(\partial_t\mathcal{S})^2, \quad P_1 = 1 + \tilde{\beta}B(\partial_r\mathcal{S})^2,$$
$$P_2 = 1 + \tilde{\beta}\frac{1}{C}(\partial_\theta\mathcal{S})^2, \quad P_3 = 1 + \tilde{\beta}\frac{1}{D}(\partial_\varphi\mathcal{S})^2. \tag{24}$$

According to the Hamilton-Jacobi method, the separation of variables technique provides the following form of particle's action, i.e.,

$$\mathcal{S} = -(E - j\Omega_H)t + \mathcal{R}(r) + \Theta(\theta,\varphi). \tag{25}$$

Using Eq. (25) in Eqs. (20)–(23), we get the following $4\times4$ matrix equation, i.e.,

$$\bigwedge(c_0, c_1, c_2, c_3)^t = 0. \tag{26}$$

By considering only the leading order term in $\tilde{\beta}$ and taking $\det\bigwedge = 0$, we determine the following form of imaginary part of particle's action

$$\mathrm{Im}\mathcal{R}^{\pm}(r) = \pm\frac{\pi(E - j\Omega_H)}{\sqrt{A_r(r_+,\theta)B_r(r_+,\theta)}} \times (1 + \tilde{\beta}\Pi). \tag{27}$$

The tunneling rate of emitted particles at horizon can be obtained as

$$\Gamma = \exp\left[-\frac{4}{\hbar}Im\mathcal{R}^+\right] = \exp\left[-\frac{4\pi(E - j\Omega_H)}{\hbar\sqrt{A_r(r_+,\theta)B_r(r_+,\theta)}} \times (1 + \tilde{\beta}\Pi)\right]. \tag{28}$$

For $\hbar = 1$ and $\Gamma_B = \exp[(E - j\Omega_H)/T_H']$, the effective Hawking temperature is evaluated as

$$T_H' = \frac{\sqrt{A_r(r_+,\theta)B_r(r_+,\theta)}}{4\pi(1 + \tilde{\beta}\Xi)} = T_H(1 - \tilde{\beta}\Pi), \tag{29}$$

where $T_H$ is the original Hawking temperature defined by Eq. (16).

## 4. Conclusion

In this paper, we have studied the tunneling rate and Hawking temperature for vector particles. By considering the Hamilton-Jacobi method and applying the WKB approximation, we have considered Proca equation for massive charged vector field. Later, we have investigated the corresponding Hawking temperature of vector particles tunneled through horizon, which preserved charge and energy. Moreover, by utilizing GUP modified Proca equation, the corrected Hawking temperature has been computed by considering quantum gravity effects. We have concluded that the correction term decelerates the increase in Hawking temperature.

## References


1. G.E. Romero and G.S. Vila, *Introduction to Black Hole Astrophysics* (Springer, 2014).
2. S.W. Hawking, *Nature* **248**, 30(1974); *ibid. Commun. Math. Phys.* **43**, 199(1975); Erratum: *Commun. Math. Phys.* **46**, 206(1976).
3. R. Banerjee and B.R. Majhi, *JHEP* **0806**, 095(2008).
4. R. Kerner and R.B. Mann, *Phys. Rev.* **D73**, 104010(2006); *ibid. Phys. Rev.* **D75**, 084022(2007).
5. P. Kraus and F. Wilczek, *Phys. Lett.* **A9**, 3713(1994); *ibid. Phys. Lett.* **B437**, 231(1995).
6. M. Angheben, M. Nadalini, L. Vanzo and S. Zerbini, *J. High Energy Phys.* **05**, 014(2005).
7. M. Sharif and W. Javed, *J. Korean Phys. Soc.* **57**, 217(2010); *ibid. Can. J. Phys.* **90**, 903(2012); *ibid. J. Exp. Theor. Phys.* **115**, 782(2012); *ibid. Eur. Phys. J.* **C72**, 1997(2012); *ibid. Gen. Relativ. Gravit.* **45**, 1051(2013); *ibid. Can. J. Phys.* **91**, 43(2013); *ibid.* Proceedings of the 3rd Galileo–Xu Guangqi Meeting, *Int. J. Mod. Phys.*: Conference Series, **23**, 271(2013); *ibid.* Proceedings of the 13th Marcel Grossmann Meeting (Stockholm, 2012), *World Scientific*, **3**, 1950(2015).
8. A.F. Ali, *JHEP* **1209**, 067(2012).
9. B. Majumder, *Phys. Lett.* **B701**, 384(2011).
10. A. Bina, S. Jalalzadeh and A. Moslehi, *Phys. Rev.* **D81**, 023528(2010).
11. P. Chen, and R.J. Adler, *Nucl. Phys. Proc. Suppl.* **124**, 103(2003).
12. R.J. Adler, P. Chen and D.I. Santiago, *Gen. Relativ. Gravit.* **33**, 2101(2001).
13. L. Xiang and X.Q. Wen, *JHEP* **0910**, 046(2009).
14. W. Kim, E.J. Son and M. Yoon, *JHEP* **0801**, 035(2008).
15. W. Javed, and R. Babar, *Fermions Tunneling and Quantum Corrections for Quintessencial Kerr-Newman-AdS Black Hole*, arXiv:hep-th/1812.07937v1.
16. W. Javed, R. Babar and A. Övgün, *Hawking radiation from cubic and quartic black holes via tunneling of GUP corrected scalar and fermion particles*, arXiv:gen-ph/1808.09795v2.
17. W. Javed, G. Abbas and R. Ali, *Eur. Phys. J.* **C77**, 296(2017).
18. A. Övgün, W. Javed and R. Ali, *Adv. High Energy Phys.* **3131620**, 11(2018).

# Killing horizons, throats and bottlenecks in the ergoregion of the Kerr spacetime

D. Pugliese and H. Quevedo

*Institute of Physics, Faculty of Philosophy & Science, Silesian University in Opava, Bezručovo náměstí 13, CZ-74601 Opava, Czech Republic*
*Dipartimento di Fisica, Università di Roma "La Sapienza", I-00185 Roma, Italy Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, AP 70543, México, DF 04510, Mexico*
*E-mail:d.pugliese.physics@gmail.com*

The properties of Kerr black holes (**BHs**)and naked singularities (**NSs**) are investigated by using stationary observers and their limiting frequencies. We introduce the concept of **NS** Killing throats and bottlenecks for slowly spinning **NSs** to describe the frequency of stationary observers. In particular, we show the frequency on the horizon can be used to point out a connection between **BHs** and **NSs** and to interpret the horizon in terms of frequencies. The analysis is performed on the equatorial plane of the ergoregion.

*Keywords*: Black holes; Naked singularities; Killing horizons.

## 1. Introduction

We study the orbital angular frequencies of stationary observers in the Kerr spacetime. We introduce the concept of Killing throats that arise in the spacetime of **NSs** and can be interpreted as the "opening" and disappearance of Killing horizons. Killing bottlenecks are identified as "restrictions" of Killing throats that appear in the case of weak naked singularities[a] (**WNSs**) for which the spin-mass ratio is close to the value of the extreme **BH**. To explore these **NS** effects, and considering the dynamics of the zero angular momentum observers (**ZAMOs**), we introduce the concept of "metric bundles" and "extended planes". In this work, we limit the analysis of the equatorial plane of the Kerr spacetime. The generalization to the case of the Reissner-Nordström and Kerr-Newmann spacetimes is presented elsewhere[4].

We first analyze the behavior of the frequency of a stationary observer in terms of the radial distance and the spin parameter $a$ of the source. In this way, we find in the **NS** region a particular set of curves that we identify as the Killing throat. In the case of **WNSs**, for which $a/M \in ]1, 2]$, the Killing throats show "restrictions" identified as Killing bottlenecks. To explore the properties of the bottlenecks, we introduce the concept of extended plane which is a graph relating a particular characteristic of a spacetime in terms of the parameters entering the corresponding spacetime metric. Any curve on the extended plane represents, therefore, a metric bundle, i.e., a family of spacetimes defined by a characteristic photon orbital frequency $\omega$ and characterized by a particular relation between the metrics parameters. In the case of the Kerr spacetime, the extended plane turns out to establish a relation

[a]The concept of strong and weak **NSs**, defined through the values of the spin parameter, has been explored in several works[1–4]. However, they can also be defined as strong curvature singularities[5].

between **BHs** and **NSs**. As a consequence, **WNSs** turn out to be related to the appearance of (a portion of) the inner horizon, whereas strong naked singularities (**SNSs**) with $a > 2M$ are related to the outer horizon.

This work is organized as follows. In Sec. 2, we discuss the concept of stationary observers, introducing the concepts of Killing throats and Killing bottlenecks. A discussion of the significance of the Killing bottlenecks and their possible origin is presented in Sec. 2.0.1. In Sec. 3, we introduce a possible generalization of the Killing horizon definition in the extended plane. Finally, we discuss a possible connection between black holes and naked singularities, revisiting the definition of horizons and the role played by **NSs** in horizon construction. Final remarks follow in Sec. 4.

## 2. Stationary observers and light surfaces

Consider the Kerr spacetime in Boyer-Lindquist (BL) coordinates, $(t, r, \phi, \theta)$, with $M \geq 0$ as the mass parameter and $a \equiv J/M \geq 0$ as the *specific* angular momentum or *spin*, where $J$ is the *total* angular momentum of the gravitational source. The horizons and ergospheres radii are given by $r_\pm = M \pm \sqrt{M^2 - a^2}$ and $r_\epsilon^\pm = M \pm \sqrt{M^2 - a^2\cos^2\theta}$. *Stationary observers* are characterized by a four-velocity of the form

$$u^\alpha = \gamma(\xi_t^\alpha + \omega\xi_\phi^\alpha), \quad \gamma^{-2} \equiv -\kappa(\omega^2 g_{\phi\phi} + 2\omega g_{t\phi} + g_{tt}), \tag{1}$$

where $\gamma$ is a normalization factor with $\kappa = -g_{\alpha\beta}u^\alpha u^\beta$, $\xi_\phi$ is the rotational Killing field, $\xi_t$ is the time-translational Killing field, and $\omega$ is a uniform *angular velocity* (dimensionless quantity)[b]. In BL coordinates, in which the metric tensor depends on $(r, \theta)$ only; this means that $r$ and $\theta$ are constants along the worldline of each stationary observer (the observer does not see the spacetime changing along the trajectory). The spacetime causal structure of the Kerr spacetime can be also studied by considering stationary observers[6].

Timelike stationary particles have orbital frequencies in the range

$$\omega \in ]\omega_-, \omega_+[ \quad \text{where} \quad \omega_\pm \equiv \omega_Z \pm \sqrt{\omega_Z^2 - \omega_*^2}, \quad \omega_*^2 \equiv \frac{g_{tt}}{g_{\phi\phi}} = \frac{g^{tt}}{g^{\phi\phi}}, \quad \omega_Z \equiv -\frac{g_{\phi t}}{g_{\phi\phi}}, \tag{2}$$

where $\omega_Z$ is the orbital angular frequency of the zero angular momentum observers (ZAMOS) and $\omega_\pm$ are the limiting frequencies of photons orbits, which are solutions of the equation $g_{\alpha\beta}\mathcal{L}_\pm^\alpha\mathcal{L}_\pm^\beta = 0$ with $\mathcal{L}_\pm \equiv \xi_t + \omega_\pm\xi_\phi$. Killing vectors $\mathcal{L}_\pm$ are generators of Killing horizons. The Killing vector $\xi_t + \omega\xi_\phi$ becomes null at $r = r_+$, defining the frequency $\omega_+(r_+) = \omega_h$.

On the equatorial plane $\theta = \pi/2$ of the Kerr spacetime we have that

$$\omega_\pm \equiv \frac{2aM^2 \pm M\sqrt{r^2\Delta}}{r^3 + a^2(2M + r)}, \quad \Delta \equiv r^2 - 2Mr + a^2 \tag{3}$$

[b]The particular case $\omega = 0$ defines *static observers*; these observers cannot exist in the ergoregion.

with the asymptotic behavior

$$\lim_{r\to\infty}\omega_{\pm}=0,\quad \lim_{a\to\infty}\omega_{\pm}=0,\quad \lim_{r\to 0}\omega_{\pm}\equiv\omega_0=\frac{M}{a} \tag{4}$$

and the particular values

$$\omega_h\equiv\omega_{\pm}(r_+)=\omega_Z(r_+)=\frac{a}{2r_+},\quad \omega_\epsilon\equiv\omega_+(r_\epsilon^+)=\frac{aM}{2M^2+a^2}. \tag{5}$$

We can see that $\omega_-<0$ for $r>r_\epsilon^+$, and $\omega_->0$ inside the ergoregion, while $\omega_+>0$ everywhere. Moreover, since $\omega_+=\omega_-$ on the horizon, stationary observers cannot exist inside this surface. Therefore, $\omega_\pm$ are limiting angular velocities for physical observers. The behavior of the frequencies $\omega_\pm$ is depicted in Fig. 1. Notice that

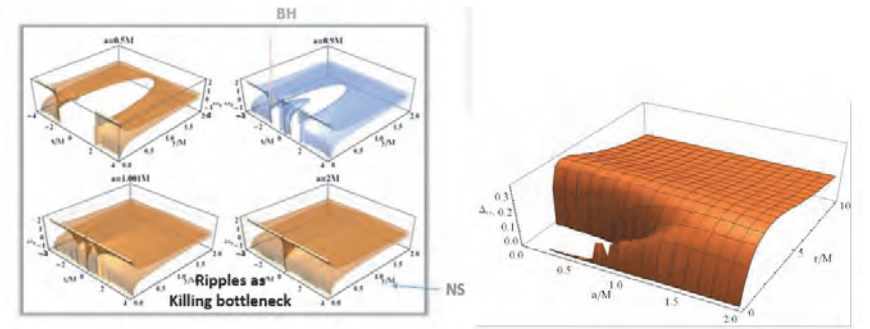



Fig. 1. Left panel: Plots of the frequency surfaces $\omega_\pm$ as functions of the radial distance $r$ in Cartesian coordinates $(x,y)$ for different spin values $a$, including **BH**s (upper plots) and **NS**s (bottom plots). Timelike stationary observers are defined in the region bounded by these planes. In the **BH** case, the horizons are clearly identified; as the spin increases, the horizons merge and, in the **NS** region, a rippled configuration (bottleneck) appears specially in the case of **WNSs** with $(a\in]M,2M[)$. Right panel: Plot of the frequency interval $\Delta\omega_\pm=\omega_+-\omega_-$ as a function of the radius $r/M$ and the spin $a/M$. The extrema $r_\Delta^\pm$ and $r_\blacksquare^\pm$ are solutions of the equations $\partial_r\Delta\omega_\pm=0$ and $\partial_a\Delta\omega_\pm=0$, respectively. Bottlenecks are shown explicitly[4].

instead of the spin parameter $a$, we could use the frequency $\omega_0$ or $\omega_h$ to parameterize the Kerr spacetime by using the relations

$$a=\frac{M}{\omega_0}\quad\text{or}\quad a=\frac{4M\omega_h}{1+4\omega_h^2}\ , \tag{6}$$

respectively, which can be prove to be $1-1$ relations. This simple observation has very important consequences. In fact, if we could measure all the values of $\omega_0\in]0,\infty[$ or $\omega_h\in]0,1/2]$, the entire Kerr family of spacetimes would be described by these frequencies. From a practical point of view, however, we can expect that the frequency on the (outer) horizon $\omega_h$ is a more suitable candidate for being measured. This, again, is an interesting fact that seems to be related to the well-known property that all the physical degrees of freedom of a black hole are encoded on the horizon.

The angular frequency $\omega_0$ ("frequency of the singularity") emerges as a relevant quantity in this analysis and, as it is clear from Fig. 1, it can be used to differentiate between low spin $a \in [0, M]$ sources, where the frequency gap $\Delta\omega_\pm = \omega_+ - \omega_-$ is larger, and **WNSs** and **SNSs**, where stationary observers are not defined (asymptotically, for large $a/M$, $\omega_0 \approx 0$ and $\Delta\omega_\pm \approx 0$ for a fixed orbit $r$).

The constant $\omega_h$ is also important in the context of **BH** thermodynamics. Indeed, from the fundamental Bekenstein-Hawking entropy equation, $S = \frac{1}{4}A_h$, where $A_h = A_h(r_+)$ is the area of the outer horizon, one can derive all the thermodynamic properties of a **BH**, including its temperature (surface gravity) and angular velocity at the horizon[c]. The frequency $\omega_h$ also determines the uniform (rigid) angular velocity on the horizon, representing the fact that the **BH** rotates rigidly. This quantity enters directly into the definition of the **BH** surface gravity and, consequently, into the formulation of the rigidity theorem. **BH** thermodynamics is established through the quantities $(r_+, \omega_h)$ and their variation (in $r > r_+$); consequently, from the point of view of the geometric laws of the **BH** thermodynamics considered here, the universe in the outer region, $r > r_+$, is regulated exclusively by the outer horizon $r_+$ (the region "inside" the black hole, i.e., $r < r_+$, does not directly enter in the outer region thermodynamics). This aspect will have a correspondence in the concept of inner horizon confinement (in the extended plane) that we will show in Sec. 3. The **BH** surface gravity is defined from the variation of the Killing field $\mathcal{L}$ norm on the outer **BH** horizon. In this analysis we study this norm and its variation on the equatorial plane of the Kerr geometries (i.e. $r > 0$ and $\theta = \pi/2$). We focus particularly on the ergoregion as the outer boundary of the ergosurface ($r_\epsilon^+ = 2M$) is a spin-independent quantity (in the static limiting metric, where $r_\epsilon^+ = 2M$, there is an event horizon associated to the Killing $\xi_t$–apparent, trapped, Killing and events horizons in stationary and static spacetimes coincide.). The relevance of this special feature of the outer ergosurface on the equatorial plane consists in the variation of the norm of the Killing field $\mathcal{L}$ with the dimensionless spacetime spin $a/M$. The off-equatorial case is instead considered in Ref. 7.

Since $S$ depends only on the radius of the outer horizon, which in turn depends only on $M$ and $a$, we can replace the spin parameter $a$ by the frequency parameter $\omega_h$, by using the $1-1$ relation (6). This means that $\omega_h$ can play the role of a thermodynamic variable for a **BH**. It would be interesting to investigate how the use of the frequency $\omega_h$, instead of the angular momentum $J = aM$, as a thermodynamic variable would affect the thermodynamics of the corresponding **BHs**.

**Killing vectors and Killing horizons**

It is convenient here to review some well-known facts about Killing horizons and Killing vectors. A **Killing horizon** is a *null* surface, $\mathcal{S}_0$, whose *null* generators

[c]In fact, one can write the Hawking temperature as $T_H = \hbar c\kappa/2\pi k_B$, where $k_B$ is the Boltzmann constant and $\kappa$ is the surface gravity. Temperature $T = \kappa/(2\pi)$; entropy $S = A/(4\hbar G)$, where $A =$ area of the horizon $A = 8\pi m r_+$; pressure $p = -\omega_h$; volume $V = GJ/c^2$ ($J = amc^3/G$); internal energy $U =$ GM ($M = c^2 m/G =$ mass) and $m$ is the mass.

coincide with the orbits of an one–parameter group of isometries (i.e., there is a Killing field $\mathcal{L}$ which is normal to $\mathcal{S}_0$). Therefore, it is a lightlike hypersurface (generated by the flow of a Killing vector) on which the norm of a Killing vector goes to zero. In static **BH** spacetimes, the event, apparent, and Killing horizons with respect to the Killing field $\xi_t$ coincide. In the Schwarzschild spacetime, therefore, $r = 2M$ is the Killing horizon with *respect* to the Killing vector $\partial_t$. The event horizons of a spinning **BH** are Killing horizons with respect to the Killing field $\mathcal{L}_h = \partial_t + \omega_h \partial_\phi$, where $\omega_h$ is defined as the angular velocity of the horizon (the event horizon of a stationary black hole must be a Killing horizon). In Ref. 7, we also consider the case of the Reissner–Nordström and Kerr-Newman spacetimes.

**Killing vectors and BH surface gravity**

The **surface gravity** of a **BH** may be defined as the *rate* at which the norm of the Killing vector vanishes from outside. The surface gravity for the Kerr **BH** metric, $\mathcal{SG}_{Kerr} = (r_+ - r_-)/2(r_+^2 + a^2)$, is a conformal invariant of the metric, but it re-scales with the conformal Killing vector. Therefore, it is not the same on all generators (but obviously it is constant along one specific generator because of the symmetries).

**Surface gravity and frequencies**

The Kerr **BH** surface gravity can be decomposed as $\kappa = \kappa_s - \gamma_a$, where $\kappa_s \equiv 1/4M$ is the Schwarzschild **BH** surface gravity, while $\gamma_a = M\omega_h^2$ is the contribution due to the additional component of the **BH** intrinsic spin; $\omega_h$ is, therefore, the angular velocity (in units of $1/M$) on the *event horizon*. The (strong) rigidity theorem connects then the event horizon with a Killing horizon, stating that, under suitable conditions, the event horizon of a stationary (asymptotically flat) solution satisfying suitable hyperbolic equations is a Killing horizon. The surface area of the **BH** event horizon is non-decreasing with time (which is the content of the second law of black hole thermodynamics). The **BH** event horizon of the stationary solution is a Killing horizon with constant surface gravity (zeroth law)[d]. Thus $\Lambda = \mathcal{L}^\alpha \mathcal{L}_\alpha$ is constant on the horizon. The surface gravity is then defined as the constant $\kappa : \nabla^\alpha \Lambda = -2\kappa \mathcal{L}^\alpha$ (on the outer horizon $r_+$). Alternatively, it holds that $\mathcal{L}^\beta \nabla_\alpha \mathcal{L}_\beta = -\kappa \mathcal{L}_\alpha$ and $L_{\mathcal{L}} \kappa = 0$, where $L_{\mathcal{L}}$ is the Lie derivative (a non affine geodesic equation). In other words, $\kappa$ is constant on the orbits of $\mathcal{L}$.

We here consider $\Lambda$ as a function of $(r, a)$ (as we are restricting the analysis to the equatorial plane). To analyze the dependence on $a$, we introduce the concept of the *extended plane* $\pi^+$ as the set of points $(a/M, \mathcal{Q})$, where $\mathcal{Q}$ is any quantity that characterizes the spacetime and depends on $a$. In general, the extended plane is an

[d]Progenitors, as stars or galaxies, have generally spin $a = J/(Mc)$ usually bigger than their mass $m = GM/c^2$. During the gravitational collapse, the body should lose mass and angular momentum, generating the Killing horizons, ending up as a black hole. Penrose Cosmic Censorship Hypothesis constrains the gravitational collapse from good (physically realistic) initial conditions for the progenitors to end up in a **BH**; this result, however, is still an hypothesis, strictly depending on the "physical" initial conditions.

$(n+1)$-dimensional surface, where $n$ is the number of independent parameters that enter $\mathcal{Q}$[7]. An example of an extended plane is plotted in Fig. 1, which corresponds to the 2D set of points $(a/M, \omega_\pm)$. The frequency $\omega_\pm$ shows "ripples", which emerge when the spin is within the interval $a \in ]M, 2M]$. The surfaces become connected and the two horizons disappear, leaving as "remnants" in the planes the "Killing bottlenecks". Those ripples appeared also in different analysis of the stationary observes[4].

2.0.1. *Light surfaces*

Light surfaces are determined by the solutions of the equation $g_{\alpha\beta}\mathcal{L}^\alpha_\pm \mathcal{L}^\beta_\pm = 0$, which in the case under consideration can be expressed as[4]

$$\frac{r_s^-}{M} \equiv \frac{2\beta_1 \sin\left(\frac{1}{3}\arcsin\beta_0\right)}{\sqrt{3}}, \quad \frac{r_s^+}{M} \equiv \frac{2\beta_1 \cos\left(\frac{1}{3}\arccos(-\beta_0)\right)}{\sqrt{3}} \tag{7}$$

$$\text{where} \quad \beta_1 \equiv \sqrt{\frac{1}{\omega^2} - \frac{1}{\omega_0^2}}, \quad \beta_0 \equiv \frac{3\sqrt{3}\beta_1\omega^2}{\left(\frac{\omega}{\omega_0}+1\right)^2}. \tag{8}$$

The surfaces $r_s = r_s^+ \cup r_s^-$ are represented for **BHs** and **NSs** in Fig. 2. In the **BH** case, we note the disconnection due to the presence of the horizons, which corresponds to the disconnection in the frequency planes of Fig. 1, where we limit the plots to the region $r_s < r_\epsilon^+$. Physical (timelike) observers are located between the inner cone and the outer cylinder surfaces of Fig. 2. As the spin increases into the **NS** regime, the surfaces merge at $a \geq M$, giving rise to a connected surface $r_s$ $(r_s^+ \cup r_s^-)$, the Killing throat (or tunnel). For spins $a \in ]0, 2M]$, the Killing bottleneck appears as restriction of the surface $r_s = r_s^+ \cup r_s^-$. We focus the analysis on the Killing bottleneck considering a $2D$ representation of the Killing throat versus the frequency $\omega$ at different spins, close to values where the bottleneck appears. Fig. 4 shows the behavior for **BHs** (gray region) and **NSs**. In the **BH** case, we do not consider the region inside the inner horizon (i.e. $r < r_-$). The outer horizon is the tangent point $r_+$ to the curve $r_s$ and the minimum (regular) point of the surface $r_s(\omega)$. For the extreme Kerr spacetime the horizon is a (non regular) cusp point in this representation. Each Killing throat is associated to one spacetime geometry.

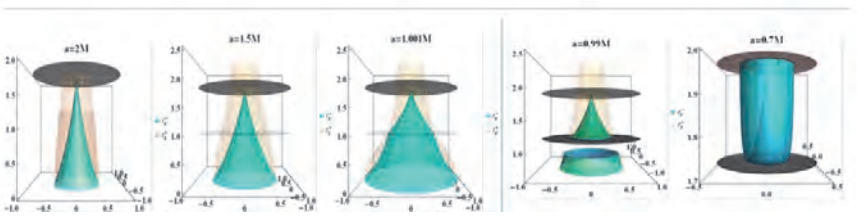


Fig. 2. Plots of the surfaces $r_s^\pm$ (in units of mass) versus the frequency $\omega$ for different spin values $a/M$, including **BH** and **NS** geometries. The Killing throat and bottleneck are plotted[4].

For example, the Killing throat defined for $a = 2M$ has a characteristic singularity frequency $\omega_0 = M/a = 0.5$, which is the limiting frequency at the singularity (note that the restriction of the throat approaching the singularity corresponds to a null frequency gap $\Delta\omega_\pm$, meaning that there are no physical stationary observers close to the singularity). At a fixed value $r_s(\omega)$, one orbit on a throat $a = 2M$, corresponds to two (positive) photon orbital frequencies $\omega_+$ and $\omega_-$ (a part on the singularity $r_s = 0$); viceversa, there can be two orbits $r_1(\omega) < r_2(\omega)$ on the surface $r_s$ (vertical lines in Fig. 2), where photons have equal orbital frequency (note that the photon orbital frequency $\omega_\pm$ are limiting frequencies for timelike particles; therefore, they determine also the range of possible values of $\omega$ for the physical stationary observers).

Killing bottleneck appears as restriction of the Killing throat in the light surfaces analysis and as ripples in the frequency planes of Fig. 1. We considered the possibility that these "horizons remnants" in **WNSs** were caused by failure of the **BL** coordinates "close" to the $a = M$, in the region near the horizon $r_+ = r_- = M$; nevertheless, Killing bottlenecks appear in the whole spin range $a \in ]M, 2M]$. To understand if the Killing bottlenecks are due to some known geometrical properties of the singularities, we focus our analysis on the spins $a \in ]M, 2M]$. Considering **ZAMOs** dynamical properties, we highlight some particular spin values of this range.

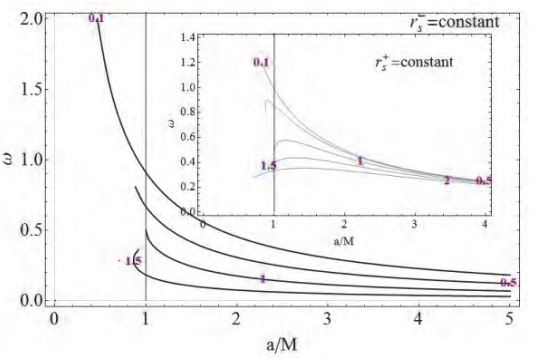



Fig. 3. Curves $r_s^-$ =constant and $r_s^+$ =constant (inside panel) in the plane $(\omega, a/M)$. The numbers denote the constant radii $r_s^\pm/M$ (light cylinders). The angular momentum and the velocity $(a, \omega)$ for $r_s^\pm(a, \omega) = 0$ are related by $\omega = M/a$.

**Zero Angular Momentum Observers**

Some properties of spacetimes with bottlenecks shown in Fig. 2 can be related to the dynamics of ZAMOs, which are defined by the condition

$$\mathcal{L} \equiv u_\alpha \xi^\alpha_{(\phi)} = g_{\alpha\beta}\xi^\alpha_\phi p^\beta = g_{t\phi}\dot{t} + g_{\phi\phi}\dot{\phi} = 0 \quad (d\phi/dt = -g_{\phi t}/g_{\phi\phi} \equiv \omega_Z = (\omega_+ + \omega_-)/2). \tag{9}$$

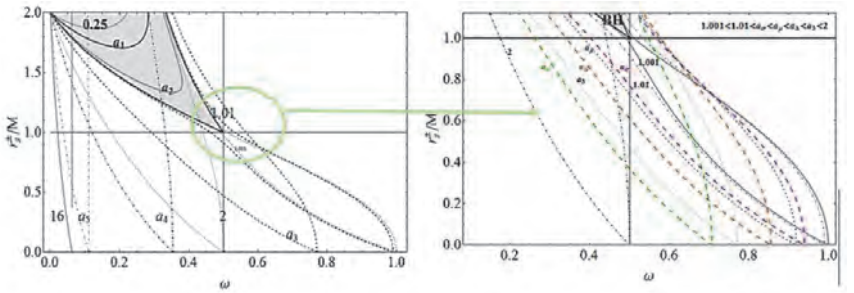



Fig. 4. The radii $r_s^{\pm}$ versus the frequency $\omega$ for different values of the spin $a/M$ (numbers close to the curves). The gray region is the only region allowed for the case of **BH** spacetimes. The surfaces $\hat{r}_{\pm}$ at $a = M$ (extreme-**BH**-case) are shown in black-thick (from Ref. 4).

The variation of the orbital frequencies of the ZAMOs with the spin, for different orbits, shows the existence of extreme orbits as shown in Fig. 6, $\partial_a \omega_Z|_{\pi/2} = 0$, $\omega_e \equiv \omega_Z(r_e)$, where

$$r_e \equiv \frac{\sqrt[3]{3}a^2 + \Upsilon^2}{3^{2/3}\Upsilon}, \quad \Upsilon \equiv \sqrt[3]{9Ma^2 + \sqrt{3}\sqrt{a^4\,(27M^2 - a^2)}}. \tag{10}$$

The relation between ZAMOs orbital frequency and light frequencies $\omega_{\pm}$ can be inferred from Eq. (2). In Figs. 5 and 6 some properties of the ZAMOs dynamics are shown, which follow from the analysis of the Killing bottleneck and the light surfaces, Fig. 4.

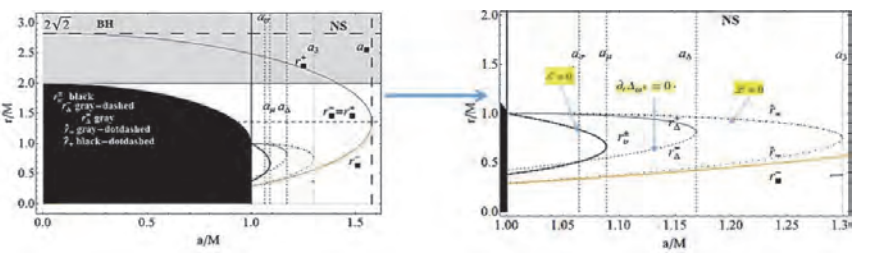



Fig. 5. The radius $r(a)$, solution of $\partial_r \Delta_{\omega_{\pm}} = 0$, i.e., it represents the critical points of the separation parameter $\Delta\omega_{\pm} \equiv (\omega_+ - \omega_-)|_{\pi/2}$ on the equatorial plane $\theta = \pi/2$. The radius $r_v^{\pm}$, where the orbital energy $\mathcal{E} = 0$, and the orbits $\hat{r}_{\pm}$, for which $\mathcal{L} = 0$, are also plotted. Dashed lines represent the spins $a_\sigma = 1.064306M$, $a_\mu = 4\sqrt{2/3}/3M \approx 1.08866M$, $a_\Delta = 1.16905M$ and $a_3 = 3\sqrt{3}/4M$. The black region corresponds to $r < r_+$. The radii $r_{\blacksquare}^{\pm} : \partial_a \Delta_{\omega_{\pm}} = 0$ are plotted as functions of $a/M$–from Ref. 4.

These spins are significant for the orbital properties of the ZAMOs which exist exclusively inside the ergoregion of **WNS** with $a \leq 1.31M$[e]. A more detailed

[e]The constant $\mathcal{L}$ and $\mathcal{E}$ shown in Figs. 5 and 6 are constant of motion associated respectively to $\xi_t$ and $\xi_\phi$ and compose the rotational version of the Killing fields i.e. the canonical vector fields

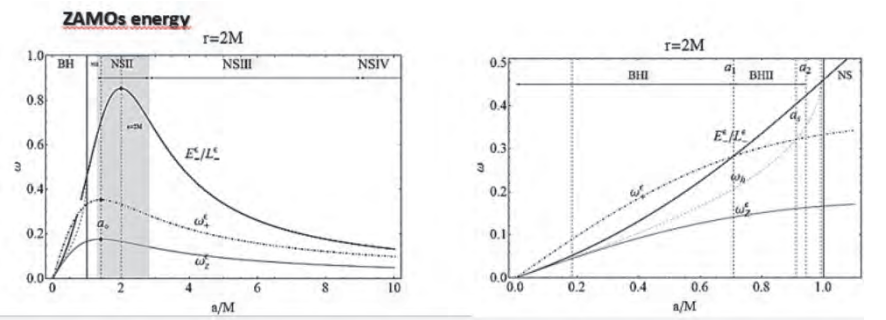



Fig. 6. Left panel: The ratio $\mathcal{E}^{\epsilon}_{-}/\mathcal{L}^{\epsilon}_{-}$ and the angular momentum of the ZAMOs $\omega^{\epsilon}_{Z}$ as a function of $a/M$ in the static limit $r = r^{+}_{\epsilon}$. The angular momentum $\omega^{\epsilon}_{+} \equiv \omega_{+}(r^{+}_{\epsilon})$ which is a boundary frequency for the stationary observer (outer light surface) is plotted (gray curve). The radius $r^{+}_{\epsilon}$ is defined by the condition $\omega_{-}(r^{+}_{\epsilon}) = 0$, $\omega_h$ is the ZAMOs angular velocity on $r = r_{+}$, i.e. $\omega_{\pm}(r_{\pm}) = \omega_h$. The maxima, related to the orbits $r_e(a)$, are denoted by points. A zoom of this plot in the **BH** region is in the right panel (from Ref. 4).

analysis is performed in Figs. 5 and 6. See also Ref. 4. From the analysis of Fig. 5 it follows that the radii connected with the the *frequency interval* $\Delta_{\omega_{\pm}} \equiv \omega_{+} - \omega_{-}$ are relevant to Killing bottlenecks:

$$
\begin{aligned}
& r^{+}_{\blacksquare} \equiv \eta \cos\left[\frac{1}{3}\arccos\left(-\frac{8a^2}{\eta^3}\right)\right], \quad r^{-}_{\blacksquare} \equiv \eta \sin\left[\frac{1}{3}\arcsin\left(\frac{8a^2}{\eta^3}\right)\right], \\
& \eta \equiv \frac{2\sqrt{8M^2 - a^2}}{\sqrt{3}}, \quad \text{for} \quad r \in ]0, 2\sqrt{2}M[ \\
& \text{where} \quad r^{\pm}_{\blacksquare} : \ \partial_a \Delta_{\omega_{\pm}}\big|_{r^{\pm}_{\blacksquare}} = 0 \quad (maximum \quad points).
\end{aligned} \tag{11}
$$

The plots of Fig. 5 show also the radii related to the frequency gap $\Delta\omega_{\pm}$. These radii are related to properties generally associated with so called "repulsive effects" of **NSs**. In this sense, the presence of ripples in the frequency sheets seem to be a related effect. It is also to be noted that the ripples are defined as a gap restriction in the frequency sheets; more precisely, they could be interpreted as due to the existence of two null orbits $r_1 \leq r_2$, where $\Delta\omega_{\pm}(r_1, r_2)$ is a minimum. That is, there is a pair of points $(r_1, r_2) \in r_s \times r_s$, on the light surface of a selected spacetime, where the photon orbital frequencies interval (range of possible timelike orbital frequency) is minimized. Clearly, the limiting case occurs for $r_1 = r_2 = r_* : \Delta\omega_{\pm}(r_*) = 0$, i.e., in $r_* \in \{0, r_{\pm}\}$. In this sense, the horizons and singularities can be interpreted

$\tilde{V} \equiv (r^2 + a^2)\partial_t + a\partial_\phi$ and $\tilde{W} \equiv \partial_\phi + a\sigma^2\partial_t$, $\sigma = \sin\theta$. The contraction of the geodesic four-velocity with $\tilde{W}$ leads to the non-conserved quantity $\mathcal{L} - \mathcal{E}a\sigma^2$, which on the equatorial plane reduces to the constant $\mathcal{L} - \mathcal{E}a$. Considering the principal null congruence $\gamma_{\pm} \equiv \pm\partial_r + \Delta^{-1}\tilde{V}$, there is the angular momentum $\mathcal{L} = a\sigma^2$, that is $\bar{\ell} = 1$ (and $\mathcal{E} = +1$, in proper units), every principal null geodesic is then characterized by $\bar{\ell} = 1$, with $\mathcal{L} = \mathcal{E} = 0$ on the horizon. In this analysis, the dimensionless radius $R \equiv r/a$ is relevant. For more discussions on the role of this ratio as $\bar{\mathcal{L}} = \mathcal{L}/a$ and $\bar{\ell} = \mathcal{L}/\mathcal{E}a$, see Ref. 4.

as the limiting cases of the Killing bottlenecks ("horizons remnants"). For more details see Ref. 7.

## 3. Horizon extension: Unveiling BH–NS connections

Figure 5 shows a bottleneck configuration for spins $a \in ]M, 1.31M]$. We now introduce the concept of metric bundle $g_\omega$ as the collection of Kerr metrics within the parameter range $a \in [a_0, a_g]$ with

$$a_\omega^\pm(r, \omega; M) \equiv \frac{2M^2\omega \pm \sqrt{r^2\omega^2\left[M^2 - r(r+2M)\omega^2\right]}}{(r+2M)\omega^2} \tag{12}$$

and a constant frequency $\omega$, which characterizes the bundle $g_\omega$. Any geometry of the bundle possesses two distinct lightlike orbits, $r_1 \leq r_2$, whose frequencies coincide with the characteristic frequency of the bundle, i.e., $\omega(r_1) = \omega(r_2) = \omega$, constrained within the limiting geometries with $a_0$ and $a_g$. Moreover, the orbital distance $(r_2 - r_1)$ reaches a maximum in the bundle and is null on the borders $a_0$ and $a_g$. *All and only* the geometries of that bundle share this property. From Fig. 7 we derive some general properties of the metric bundles: **1)** The vertical axis of the extended planes contains all the *origins* $a_0 \in ]0, \infty[$ of the metric bundles $g_\omega$. **2)** The curve $a_+$ represents the horizon in the extended plane $\pi^+$ and the boundary spin $a_g \in ]0, M]$ is associated to this curve. **3)** We can identify a correspondence between the inner horizon $r_-$ and outer horizon $r_+$ on the curve $a_+$, as shown in Fig. 7; we identify the **BH**-strip in the lower part of the panel and the **NS**-strip in the upper one as indicated. **4)** The metric bundles $g_\omega$ are closed as a consequence of the closing of the $\xi_\phi$ orbits. **5)** The metric bundles $g_\omega$ are tangent to the horizon. **6)** The metric bundle shown in Fig. 7 corresponds to the characteristic frequency $\omega = 0.5$ and origin $a_0 = 2M$, which is the upper limit of the spin region, where the Killing bottleneck appears. The bundle is tangent to the Killing horizon at the point $r_h = M$, corresponding to the spin $a_g = M$. This bundle is related also to the light surfaces with $\omega_0 = 0.5$ and $\omega_0 = 1$ as seen in Fig. 4. We note that this special metric bundle, corresponding to the spin $a_g = M$, is regular on the point

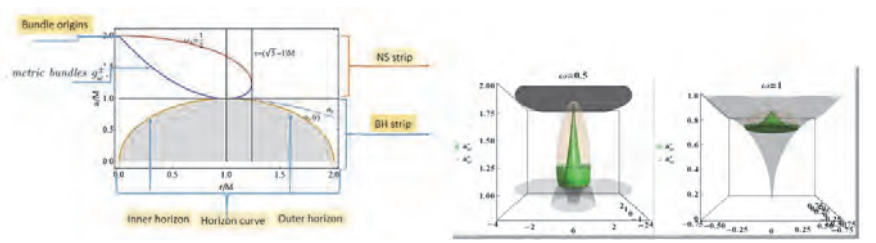



Fig. 7. The metric bundle $g_\omega$ for the spin $\omega_\pm = \omega_h = 1/2$ and bundle origin $a_0 = 2M$. The horizon curve $(a_+(r) \equiv \sqrt{r(2M-r)})$ and $a_\gamma$, solution of $r = r_\gamma$ where $r_\gamma \in \Sigma_\epsilon^+$ is the photon orbit in the ergoregion of the Kerr **BH**, are also plotted. The right panel shows a confinement of the metric bundle in the region bounded by the inner horizon and the bundle at $\omega = 0.5$.

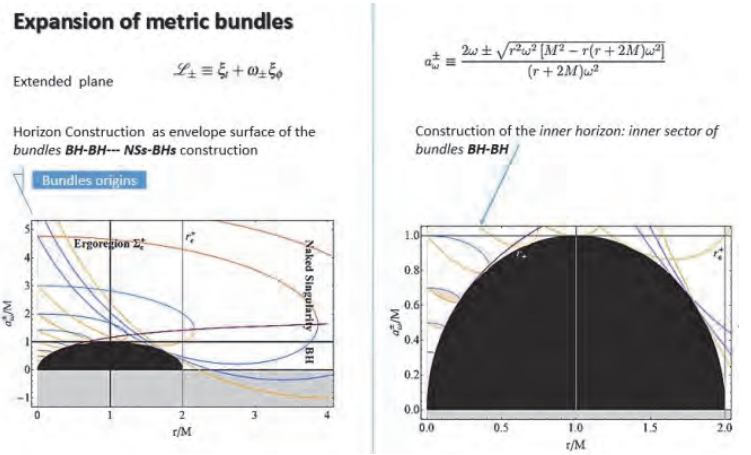



Fig. 8. The metric bundle $g_\omega$ with origins $a_0 > 0$. The horizon appears as the envelope surface.

on the horizon for the extreme Kerr **BH** and relates the spin-origin $a_0$ and the tangent point $r_h$ on the horizon with the spin $a_g$ defined by the tangent point. It is clear that the horizon curve on the extended plane corresponds to a unique horizon frequency $\omega_H$. It is possible to show that the correspondence $(a_0, a_g, \omega_H)$ can be set as a bijection.

Translating the bundle origin $a_0$ in all the range $a_0 > 0$ in Fig. 8, some relevant properties of the metric bundles and the Killing horizon appear: **i)** All the metric bundles are tangent to the horizon. **ii)** The metric bundles do not penetrate the horizon. **iii)** The space subtended by the horizon curve in the extended plane is not described by any metric of any bundle (in the sense of our analysis).

From the above properties, we can derive the following consequences: **1)** The frequencies $\omega$ of each bundle $g_\omega$ (and on every point $r$ of each bundle) are all and only those of the horizon frequency $\omega_H$. It is, therefore, sufficient to know the horizon frequency $\omega_H$ in the extended plane to fix the photon orbital frequencies (and therefore the physical observer frequency range) in each point of any **BH** or **NS** geometry of the Kerr spacetime family. **2)** The horizon arises as the *envelope surface* of all the metric bundles. **3)** The part of the horizon curve corresponding to the inner horizon in $\pi^+$ is built partially by metric bundles all contained in the region inside the inner horizon, i.e., confined in $a_g \in [0, M]$, $r \in [0, M]$ and $a_0 \in [0, M]$. These bundles have origins in $a_0 \in [0, M]$. However, these special bundles are not sufficient to construct, as envelope surface, the whole inner horizon in the extended plane. **4)** The bundles necessary for the construction of the other part of the the inner horizon in $\pi^+$ have origin in $a_0 \in [M, 2M]$, i.e., in the **WNS** geometries, where the Killing bottleneck appear. **5)** The Killing bottleneck appears to be related to the properties of these special metric bundles, which are involved in the construction of the inner horizon as envelope surface. **6)** The portion of the horizon curve corresponding to the outer Killing horizon on the equatorial plane is constructed by metric bundles with origins $a_0 > 2M$, which we identify as **SNS**.

We close this section noting that the whole set of photon limiting orbital frequencies $\omega_\pm$ (or alternatively the light surfaces $r_s^\pm$) of a single Kerr geometry with spin $\bar{a}$ (and therefore the range of orbital frequencies for the physical observers) in the extended plane, is given by the collection of points of all the metric bundles on the horizontal lines $\bar{a}$ =constant in Figs. 7 and 8. These frequencies are all and only those of the horizon frequency $\omega_H$. We discuss the interpretation of this result in the next conclusive section. Finally, we mentioned above that the internal bundles with origin in $a_0 \in [0, M]$ are all confined in the region of the inner horizon in $\pi^+$, i.e., in $a_g \in [0, M]$. This can be shown in several ways. In particular, considering again the horizon frequencies $\omega_h^\pm$, we introduce the radii $r_\mp^\mp$, defined as

$$r_-^- = \frac{1}{2}\left(\sqrt{\frac{32r_-}{a^2} - a^2 + 6\sqrt{1-a^2} - 22} - r_-\right) : \omega_-(r_-^-) = \omega_-(r_-) = \omega_h^- \quad (13)$$

$$r_+^+ = \frac{1}{2}\left(\sqrt{\frac{32r_+}{a^2} - a^2 - 6\sqrt{1-a^2} - 22} - r_+\right) : \omega_+(r_+^+) = \omega_+(r_+) = \omega_h^+, \quad (14)$$

$$\text{where} \quad (r_-^- < r_-) < (r_+ < r_+^+),$$

where we have used dimensionless units. The radii $(r_+^+, r_-^-)$ correspond to photon orbits with frequencies $\omega_\pm$ of the **BH** horizons $\omega_h^\pm$. The orbital frequency of the inner horizon has a replica on an orbit $r_-^- \in [0, r_-]$- "inner horizon frequency confinement". This result is in agreement with the **BHs** thermodynamic properties discussed in Sec. 2. Also, the relation $(a_g, a_0, r_h)$, where $r_h$ is obtained through $\omega_h$, has been mentioned to be bijective. We can show this through the relation between $(a_g, a_0)$ [7]

$$\forall\, a_0 > 0, \quad a_g \equiv \frac{4a_0M^2}{a_0^2 + 4M^2} \quad \text{where} \quad a_g \in [0, M] \quad \text{and} \quad (15)$$

$$\lim_{a_0\to 0} a_g = \lim_{a_0\to\infty} a_g = 0, \quad a_g(a_0 = 2M) = M. \quad (16)$$

This relation also allows us to formalize the **BH-BH** correspondence (construction of the inner horizon as an envelope surface in $\pi^+$), the **BH-WNS** relation (inner horizon construction) and the **BH-SNS** relation (construction of the outer horizon). Particularly, the horizon relates in the extended plane **BHs** with **NSs** (through the origins of the metrics bundles $a_0$ and $a_g$). In this sense, the **NSs** can be interpreted as necessary for the construction (as envelope surface) of the inner and outer horizons in the extended plane. Note that the static case of the Schwarzschild geometry, $a = 0$, can be seen as the limiting case in Figs. 7 and 8, where the horizon $a_g = 2M$ is in correspondence with the limiting **SNS** bundle with origin $a_0 = +\infty$. The plane $\pi^+$ in Figs. 7 and 8 have several symmetries. Note that in Fig. 8 negative values of $a_0$ and $a_g$ are possible in the metric bundles and are related to $\omega_\pm < 0$ frequencies, which are possible outside the ergoregion $(r > r_\epsilon^+)$. The quantity $\mathcal{A}_{r_\pm}^+ = \pi/2$ is the area of the region of $\pi^+$ bounded by the horizon (dimensionless quantities); $\mathcal{A}$ is the area of the regions in the extended plane $\pi^+$

bounded by the curves $a^{\pm}_{\omega}$, defining the metric bundles $g_{\omega}$. $\mathcal{A}$ is a decreasing function of the frequency $\omega$, shrinking at the origins $a_0 < M$, i.e. $\omega_0 = M/a > M$, where $g^{\pm}_{\omega}$ are all bounded by the inner horizons; viceversa, the region areas grow as the spin-mass ratio increases in the **NS** geometries. We repeated this analysis in the case of Reissner-Nordström and Kerr-Newman spacetimes and similar results are found.

## 4. Final remarks

In this work, we investigated the properties of stationary observers on the equatorial plane of the Kerr spacetime. The generalization to the off-equatorial case as well as the Reissner-Nordström and Kerr-Newman spacetimes is presented elsewhere[7]. We focused on the behavior of the frequency of stationary observers. To emphasize its peculiarities, we introduced the concept of Killing throats and bottlenecks. If we consider the frequency as a function of the spin, certain features appear that are better explained by introducing the concepts of extended planes and metric bundles. In the case of the Kerr metric on the equatorial plane, the extended plane is essentially equivalent to the function that relates the frequency with the spin.

Metric bundles and horizons remnants appear related to the concept of pre-horizon regimes. There is a pre-horizon regime in the spacetime when there are mechanical effects allowing circular orbit observers to recognize the close presence of an event horizon. This concept was introduced in Ref. 8 and detailed for the Kerr geometry in Refs. 13, 14, 9. The analysis of the pre-horizon structure led to the conclusion that a gyroscope would observe a memory of the static case in the Kerr metric. It is clear that this aspect could have an essential relevance in the investigation of the collapse[10–12]–see also[15].

In the extended plane of the Kerr metric, the frequency on the horizons determines a set of metrics, a metric bundle, describing in general **BHs** and **NSs**, and also the limiting frequencies for stationary observers, which correspond to light-like orbits. This fact can be interpreted as determining a connection between **BHs** and **NSs**. In the extended plane, the **NS** solutions have a clear meaning in relation to the construction as envelope surface of portions of the horizon in $\pi^+$. The inner **BH** horizon is connected to the origin of **BH** bundles and the outer horizon establishes **BHs-SNSs** correspondence. Moreover, the horizon in the extended plane can be interpreted as the envelope surface of all metric bundles. On the other hand, the metric bundles are all defined by all and only the frequency of the horizon. In this sense, the corresponding inner horizon in $\pi^+$ is partially constructed by **BHs** metric bundles. The inner horizon is completed by bundles including **BHs** and **WNSs**. This property appears related with the Killing bottlenecks appearing in the light surfaces. Interestingly, the outer horizon in $\pi^+$ is generated by **SNSs** metric bundles. It is then possible to argue that this result could be of interest for the investigation of the gravitational collapse. Indeed, suppose that the collapse is a quasi-stationary process in which each state is described by a Kerr spacetime. Since

rotating astrophysical compact objects are characterized by spin parameters, which correspond to **NS** configurations ($a/M > 1$), the formation of a **BH** ($a/M \leq 1$) would necessarily imply passing through a series of states with spin parameters in the **NS** regime. In this case, the extended plane of the Kerr metric as described above could contain the different states which are necessary for the formation of a **BH**. This fact has the interesting consequence that only horizon frequencies in extended plane determine the frequencies $\omega_{\pm}$ at each point, $r$, on the equatorial plane of a Kerr **BH** or **NS** geometry. All the frequencies $\omega_{\pm}(r)$ on the equatorial planes are only those of the horizon in $\pi^{+}$. Another relevant aspect connected with this fact is the confinement of the horizon in the sense of the frequencies given in Eq. (13). These aspects are currently under investigation.

*Acknowledgments*

D.P. acknowledges support from the Junior GACR grant of the Czech Science Foundation No:16-03564Y. This work was partially supported by UNAM-DGAPA-PAPIIT, Grant No. 111617. D.P. is grateful to Donato Bini, Fernando de Felice and Andrea Geralico for discussing many aspects of this work.

## References

1. D. Pugliese, H. Quevedo and R. Ruffini, Phys. Rev. D **84**, 044030 (2011).
2. D. Pugliese, H. Quevedo and R. Ruffini, Phys. Rev. D **88**, 2, 024042 (2013).
3. D. Pugliese & H. Quevedo, Eur. Phys. J. C **75**, 5, 234 (2015).
4. D. Pugliese & H. Quevedo, Eur. Phys. J. C **78**, 1, 69 (2018).
5. C. J. S. Clarke, F. De Felice, Gen. Rel. Grav., **16**, 2, 139-148 (1984).
6. D. B. Malament, J. Math. Phys., **18**, 1399 (1977).
7. D. Pugliese & H. Quevedo, Eur. Phys. J. C **79**, 3, 209 (2019).
8. F. de Felice, Mont. Notice R. Astr. Soc. **252**, 197-202 (1991).
9. F. de Felice & S. Usseglio-Tomasset, Class. Quantum Grav. **8**, 1871-1880 (1991).
10. F. de Felice & S. Usseglio-Tomasset, Gen. Rel. Grav. **24**, 10 (1992).
11. F. de Felice & L. Di G. Sigalotti Ap.J. **389**, 386-391 (1992).
12. F. de Felice & Y. Yunqiang, Class. Quantm Grav. **10**, 353-364 (1993).
13. F. de Felice, Class. Quantum Grav. **11**, 1283-1292, (1994).
14. F. de Felice & S. Usseglio-Tomasset, Gen. Rel. Grav. **28** 2 (1996).
15. C. Chakraborty, M. Patil, P. Kocherlakota, et al. Phys. Rev. D **95**, 8, 084024 (2017).

# A holographic approach to gravitational thermodynamics

Fil Simovic

*Perimeter Institute for Theoretical Physics,*
*Waterloo, Ontario N2L 6C1, Canada*

*Department of Physics and Astronomy, University of Waterloo,*
*Waterloo, Ontario N2L 1W6, Canada*

*E-mail: fil.simovic@gmail.com*

We formulate a description of 3+1 dimensional gravitational phenomena in terms of a relativistic fluid living on the 2+1 dimensional timelike boundary of an arbitrary bulk region of spacetime, called a *gravitational screen*. We establish a consistent dictionary between the geometric variables describing the evolution of the screen and the thermodynamic variables describing a relativistic viscous fluid, and discuss the interpretation. We also examine the construction of gravitational screens in different spacetimes and analyze the properties of the fluids they realize.

*Keywords*: Gravity; thermodynamics; quasilocal; holography; hydrodynamics.

## 1. Introduction

One of the greatest theoretical developments in modern physics has been the anti-de Sitter/conformal field theory correspondence (AdS/CFT), which conjectures an equivalence between a theory of gravity in a bulk region of spacetime, and a quantum field theory on the boundary of that spacetime[1]. AdS/CFT has illuminated many aspects of string and field theory, giving key insights into what a quantum theory of gravity might look like, having also provided us with powerful tools for studying a broad range of strongly coupled systems, such as fluids near quantum critical points and quark-gluon plasmas. This notion of a bulk/boundary duality is also manifest in the membrane paradigm[2], where the interior of a black hole is replaced with a relativistic fluid living near the event horizon, and the celebrated fluid/gravity correspondence[5], where the long wavelength limit of Einstein's equations are mapped to fluid dynamics on the boundary of AdS. These ideas are all, in essence, a statement of the holographic principle; the fundamental idea that all of the information contained in a bulk region of spacetime can be encoded in the boundary of that region.

These approaches offer a limited perspective however, since we are usually constrained to situations where knowledge of the boundary of space or the end of time is required. AdS/CFT and the fluid/gravity correspondence makes reference to the boundary of anti-de Sitter space at infinity, while standard formulations of black hole thermodynamics, as well as the membrane paradigm, relate quantities to the event horizon, which is teleological and can only be located by knowing the entire future history of the universe. From a practical point of view this is unsatisfactory; as local observers we are generally unable to access these types of boundaries.

Recent progress in moving beyond these restricted classes of boundaries has led to the concept of a *gravitational screen*[4], a 2+1 dimensional timelike hypersurface representing the time evolution of the 2D boundary of a bulk region of spacetime. In this paradigm, Einstein's equations in the bulk are mapped to equations governing the non-equilibrium thermodynamics of a viscous fluid living on the screen, with all of the gravitational dynamics present inside the screen being encoded in the boundary fluid. This quasilocal approach is reminiscent of the membrane paradigm and fluid/gravity correspondence, but can be formulated on arbitrary timelike surfaces, without the aforementioned restrictions. It also bears some resemblance to other quasilocal approaches, such as the original work of Brown and York[3] and rigid quasilocal frames[7], though no fluid interpretation is given in those.

In this work, we present the fully relativistic generalization of the gravitational screen[a]. Whereas the original work considered a correspondence between Einstein's equations in the bulk and the non-relativistic Navier-Stokes equations on the boundary, here we promote the boundary theory to a fully covariant one, placing it on equal footing with the bulk theory. We clarify the interpretation of the fluid pressure as the screen's normal acceleration, finding that the correct bulk/viscous splitting must include an expansion-dependent term. We also construct examples of gravitational screens both in Minkowski and Schwarzschild backgrounds and examine the properties of the corresponding holographic fluids. Finally, we examine how the fluid entropy can be linked to the curvature of spacetime, and remark on the salient features of the correspondence.

## 2. Screens as hypersurfaces

The gravitational screen $\Sigma$ is a 2+1 dimensional hypersurface which is the time evolution of a 2D boundary $\mathcal{S}$ of some 3D bulk connected region of interest. The screen is topologically $\mathbb{R}\times S^2$, admitting a smooth codimension-one foliation by the leaves $\mathcal{S}$, which are level sets of some time function $t(x)$ inherited from the bulk foliation. We consider spherically symmetric boundaries only, though the screen geometry can in principle be chosen arbitrarily. The screen is defined by $s^a$, the outward pointing spacelike unit normal vector to $\mathcal{S}$, and $u^a$, the timelike unit vector tangent to $\Sigma$. By construction $s_a s^a = 1$, $u_a u^a = -1$, and $s_a u^a = 0$.
The metrics on $\Sigma$ and $\mathcal{S}$ in terms of the bulk metric $g_{ab}$ are

$$\Sigma: \quad h_{ab} = g_{ab} - s_a s_b \qquad\qquad \mathcal{S}: \quad q_{ab} = h_{ab} + u_a u_b \ , \tag{1}$$

with the associated extrinsic curvatures $H_{ab} = h_a{}^c h_b{}^d \nabla_c s_d$ and $\Theta_{\mathbf{u}ab} \equiv q_a{}^c q_b{}^d \nabla_c u_d$. The junction conditions[6] then lead to a surface stress-energy tensor

$$\tilde{S}_{ab} = \frac{1}{8\pi G} S_{ab} = \frac{1}{8\pi G}\big([H]h_{ab} - [H_{ab}]\big) \ , \qquad H \equiv h^{ab} H_{ab} \ . \tag{2}$$

[a]These proceedings serve as a brief summary of the current work. A more detailed and complete presentation, along with numerous examples, will be available in a future publication.

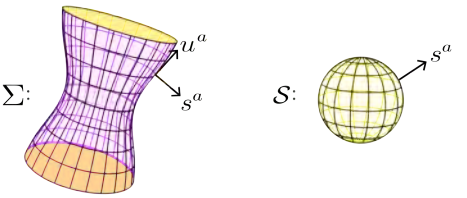



Fig. 1. A gravitational screen $\Sigma$ and the spatial boundary $\mathcal{S}$.

Square brackets represent the discontinuity of a quantity across the boundary. We will interpret this as the stress-energy tensor of a relativistic fluid living on the surface $\Sigma$. We adopt a holographic point of view, where the fluid stress-energy tensor (supported entirely on $\Sigma$) and the equations governing its evolution map to the geometry and dynamics within the screen. To this end, we let $[H_{ab}] \to H_{ab}$ and $[H] \to H$ giving

$$S_{ab} = Hh_{ab} - H_{ab} \ . \tag{3}$$

This is in the same spirit as the membrane paradigm [2], where the stretched horizon is taken to be the boundary of the spacetime and the stress energy tensor on the surface is chosen so that field lines terminate at the boundary. To arrive at a precise correspondence, we take Einstein's equations in the 4D space with zero cosmological constant, $R - \frac{1}{2}Rg_{ab} = T_{ab}$, and project them onto the timelike membrane $\Sigma$ via the Gauss-Codazzi equations, the first of which is the momentum constraint

$$D_b S^{ba} = T_{cb} s^c h^{ba} \tag{4}$$

where $D_a V_b = h_a{}^c h_b{}^d \nabla_c V_d$ is the covariant derivative on $\Sigma$. Equation (4) can further be projected into the spatial direction as $(D_b S^{ba}) q_{ac} = T_{\mathbf{s}c}$, where $T_{\mathbf{s}c} \equiv s^a T_{ac}$ and it is understood that the index $c$ represents components tangent to $S$. This equation expresses conservation of momentum on $\Sigma$, with $T_{\mathbf{s}c}$ representing the momentum flux density across the screen. Using (3) and inserting the factor $(8\pi G)$ gives:

$$(8\pi G) T_{\mathbf{s}c} = D_c(\gamma_{\mathbf{u}} + \tfrac{1}{2}\theta_{\mathbf{s}}) - \theta_{\mathbf{s}} \mathrm{a}_{\mathbf{u}c} - (\tfrac{3}{2}\theta_{\mathbf{u}} q_{bc} + \tilde{\Theta}_{\mathbf{u}bc} + \epsilon_{bc})\omega^b - D_{\mathbf{u}}\omega_c - D_b \tilde{\Theta}^b_{\mathbf{s}c}. \tag{5}$$

Here, $\mathrm{a}_{\mathbf{u}c} \equiv u^b \nabla_b u_c$ is the acceleration of screen observers, $\omega_a \equiv q_a{}^b (s_c \nabla_b u^c)$ is the normal one-form, $\gamma_u \equiv s_b u^a \nabla_a u^b$ is the normal acceleration, $\tilde{\Theta}_{\mathbf{u}bc}$ is the symmetric trace-free part of $\Theta_{\mathbf{u}bc}$, $\epsilon_{bc}$ is the antisymmetric part of $\Theta_{\mathbf{u}bc}$ (which vanishes if $u^a$ is hypersurface orthogonal) and $\theta$ is the expansion. Projecting (4) along $u^a$ instead gives conservation of energy on the screen

$$(8\pi G) T_{\mathbf{su}} = (D_u + \theta_{\mathbf{s}})\theta_{\mathbf{u}} - (\gamma_{\mathbf{u}} + \tfrac{1}{2}\theta_{\mathbf{s}})\theta_{\mathbf{u}} - \tilde{\Theta}^{ab}_{\mathbf{s}} \Theta_{\mathbf{s}ab} + (\mathrm{d}_a + 2\mathrm{a}_{\mathbf{u}a})\omega^a \ , \tag{6}$$

where $T_{\mathbf{su}} \equiv s^a u^b T_{ab}$ represents the energy flux density flowing across the screen and we have defined the covariant derivative on $\mathcal{S}$ as $\mathrm{d}_a V_b = q^c{}_a q^b{}_d \nabla_c V_d$.

## 3. Relativistic hydrodynamics

We consider now the equations governing a relativistic viscous fluid living on a 2+1 dimensional surface with topology $S^2 \times \mathbb{R}$. The conservation equations are

$$\nabla_a T^{ab} = 0 \ , \tag{7}$$

where in the general case we consider non-perfect (non-equilibrium) fluids with stress-energy tensors of the form

$$T^{ab} = e u^a u^b + (p+\pi) q^{ab} + u^a q^b + u^b q^a + \Pi^{ab} \ . \tag{8}$$

$u^a$ is the fluid 4-velocity, $e$ is the internal energy density of the fluid, $p$ is the isotropic pressure, $q^{ab}$ is the metric on $\mathcal{S}^2$, $\Pi^{ab}$ is the anisotropic stress tensor, $q^a$ is the heat flux, and $\pi$ is the dynamic pressure (the difference between the total pressure and pressure at equilibrium). $\Pi^{ab}$, $q^a$, and $\pi$ are thermodynamic fluxes that capture the deviations from a perfect fluid, and are given in this case by

$$\Pi^{ab} = q^a{}_c q^b{}_d T^{cd} - (p+\pi) q^{ab} \ , \qquad q^a = -q^a{}_c T^{cb} u_b \ , \qquad p + \pi = \frac{1}{2} q_{ab} T^{ab} \ . \tag{9}$$

As in the gravity picture, we project (7) in the directions parallel and orthogonal to $u^a$, giving the analogue of the Gauss-Codazzi equations but with $T_{ab}$ playing the role of $S_{ab}$. Projecting in directions orthogonal to $u^a$ gives conservation of momentum,

$$0 = \dot{u}_c (e + p + \pi) + \mathrm{d}_c (p + \pi) + (\mathrm{d}_a + \dot{u}_a) \Pi^a{}_c + q_{cb} \dot{q}^b + \left( \omega_{ac} + \sigma_{ac} + \tfrac{3}{2} \theta q_{ac} \right) q^a \ , \tag{10}$$

while the projection in the direction of $u^a$ gives conservation of energy,

$$0 = -\dot{e} - (e + p + \pi) \theta - (\mathrm{d}_a + 2 \dot{u}_a) q^a + \Pi^{ab} \sigma_{ab} \ . \tag{11}$$

## 4. The dictionary

We can now construct the map between the geometric variables describing the boundary evolution, and the thermodynamic variables governing the evolution of the relativistic fluid. Remarkably, this can be done consistently in the relativistic case just as in the Newtonian case considered in Ref. 4. Comparing the conservation of momentum (10) and energy (11) equations for the fluid to the gravity equations (5) and (6) leads to the following identifications:

$$e = -\frac{\theta_{\mathbf{s}}}{8\pi G} \ , \qquad p + \pi = \frac{\gamma_{\mathbf{u}} + \frac{1}{2}\theta_{\mathbf{s}}}{8\pi G} \ , \qquad \theta = \theta_{\mathbf{u}} \ , \qquad \omega_{ab} = \epsilon_{ab} \tag{12}$$

$$\Pi_{ab} = -\frac{\tilde{\Theta}_{\mathbf{s}ab}}{8\pi G} \ , \qquad \sigma_{ab} = \Theta_{\mathbf{u}ab} \ , \qquad q_a = -\frac{\omega_a}{8\pi G} . \tag{13}$$

We initially took the hydrodynamic conservation laws (7) to be source-free, resulting in the vanishing left hand side of (10) and (11). It is clear now that a non-zero $T_{\mathbf{su}}$ or $T_{\mathbf{s}c}$ manifest themselves as non-zero source terms in (7).

In this picture, the fluid energy density $e$ is related to the rate of expansion of outgoing radial null geodesics at the boundary. Our choice of sign leads to a negative energy density which increases with the mass of the screened region. Positivity of entropy and temperature for the screen fluid requires the presence of a negative chemical potential $\mu$ for the fluid. The 2D fluid pressure $p$ is now identified with $\gamma_{\mathbf{u}}+\frac{1}{2}\theta_{\mathbf{s}}$, in contrast to both the membrane paradigm and the non-relativistic screen formalism, where it is simply $\gamma_{\mathbf{u}}$. The pressure can thus vanish for non-trivial screen geometries leading to complex thermodynamic behaviour. The fluid expansion $\theta$ is directly related to the expansion in the time direction $\theta_t$ of the screen. The intuition is clear; the fluid expands as the 2D volume of the boundary increases with time.

A new feature appearing in our work is the fluid twist or kinematic vorticity $\epsilon_{ab}$ which maps to the twist $\omega_{ab}$ of the screen observers. The fluid twist measures rigid rotations of fluid lines with respect to the local inertial rest frame. Traditionally, quasilocal approaches to studying gravitational thermodynamics use shell observers whose 4-velocities are hypersurface orthogonal [7], and therefore have vanishing twist. In hydrodynamics however, $\epsilon_{ab}$ plays an important role in establishing the Kelvin-Helmholtz and Bernoulli theorems [8], as well as modelling turbulence, and is therefore a critical ingredient in understanding the full non-linear dynamics of gravity in hydrodynamic terms. With this dictionary in place, well-known results from relativistic hydrodynamics as well as the laws of thermodynamics can be used to study gravitational phenomena.

## 5. Thermodynamics of screen fluids

The thermodynamic system governed by (6) has more degrees of freedom than constraints, requiring additional information to complete. In hydrodynamics, this comes in the form of an equation of state $e(p)$ which characterizes the fluid. From the gravity point of view, $e \sim \theta_u$ and $p \sim \gamma_s$, so it is the screen evolution/geometry which fixes the equation of state of the fluid. We are also interested in relating the temperature $T$ and entropy $S$ of the boundary fluid to the internal geometry. These quantities are determined by assuming quasistatic equilibrium and employing the first law and Euler relation for a (charge-free) relativistic fluid,

$$de = Tds + \mu dn \ , \qquad e + p = Ts + \mu n \ . \tag{14}$$

A natural question that arises is "Which screen evolutions give rise to physical equations of state?". As an example, one can construct[b] a static, spherically symmetric screen in Minkowski space, whose boundary fluid has the simple equation of state

$$e(p) = -2p \ , \tag{15}$$

which has the form $p(e) = \omega e$ with $\omega = -1/2$. Such equations of state appear in cosmology as models of dark energy fluids, which satisfy the strong energy condition

[b] We omit the details of the construction for brevity. All of the relevant quantities can be calculated from the definitions given in Section 2.

but do not support classical perturbations. This fluid requires a negative chemical potential in order to have a positive entropy and temperature, suggesting that the underlying microscopic constituents are bosonic in nature.

Another example is a static, spherically symmetric screen centered in a Schwarzschild background, lying outside of the event horizon. In this case the fluid has an equation of state

$$e(p) = \frac{-4\,p\left(1 - 32\pi mp + \sqrt{1 - 32\pi mp}\right)}{\left(1 + \sqrt{1 - 32\pi mp}\right)^2} \ , \tag{16}$$

where $m$ is the ADM mass. The fluid retains its barotropic nature, but can now support classical perturbations (the speed of sound is positive) and has an entropy which increases monotonically with the screen radius, allowing one to map the fluid entropy to the curvature of the space enclosed by the screen.

These simple examples of screens demonstrate some of the salient features of the dictionary, and illustrate the subtleties involved in constructing screens which have interpretations in terms of physical fluids. It is also interesting to consider fixing the properties of the boundary fluid first, and ask what dynamics are required of the screen to realize that fluid. Indeed, we would eventually like to classify all possible physical screen evolutions and boundary fluids. Moving forward, we aim to further apply this formalism to dynamic backgrounds (which generically lead to time-dependent equations of state) and better understand how phenomena like gravitational waves in the bulk manifest in the boundary fluid. We would also like to leverage the power of the bulk/boundary map to shed light on dissipative phenomena in relativistic hydrodynamics, where exact solutions are generally unavailable.

## Acknowledgements

We would like to thank Laurent Freidel and Richard J. Epp for their many useful comments and insights throughout the course of this work.

## References

1. J. M. Maldacena, *Int. J. Theor. Phys.* **38** 1113–1133 (1999).
2. K. S. Thorne, R. H. Price, D. A. Macdonald, and S. Detweiler, *Phys. Today* **41** 74 (1988).
3. J. D. Brown and J. W. York, *Phys. Rev. D* **47** 4 (1993).
4. L. Freidel and Y. Yokokura, *Class. Quant. Grav.* **32** 215002 (2015).
5. V. E. Hubeny, *Class. Quant. Grav.* **28** 114007 (2011).
6. W. Israel, *Nuovo. Cim. B10*, **48** 463 (1967).
7. R. J. Epp, R. B. Mann, and P. L. McGrath, *Class. Quant. Grav.* **26** 035015 (2009).
8. L. Rezzolla and O. Zanotti, *Oxford Univ. Press* (2013).

# Trapped surfaces, energy conditions, and horizon avoidance in spherically-symmetric collapse

Valentina Baccetti

*Centre for Quantum Computation and Communication Technology, School of Science, RMIT University, Melbourne, VIC 3000, Australia*

Robert B. Mann

*Department of Physics and Astronomy, University of Waterloo, Waterloo, ON, N2L 3G1, Canada*

*Perimeter Institute for Theoretical Physics, Waterloo, ON, N2L 6B9, Canada*

Daniel R. Terno*

*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*
*E-mail: daniel.terno@mq.edu.au*

We consider spherically-symmetric black holes in semiclassical gravity. For a collapsing radiating thin shell we derive a sufficient condition on the exterior geometry that ensures that a black hole is not formed. This is also a sufficient condition for an infalling test particle to avoid the apparent horizon of an existing black hole and approach it only within a certain minimal distance. Taking the presence of a trapped region and its outer boundary — the apparent horizon — as the defining feature of black holes, we explore the consequences of their finite time of formation according to a distant observer. Assuming regularity of the apparent horizon we obtain the limiting form of the metric and the energy-momentum tensor in its vicinity that violates the null energy condition (NEC). The metric does not satisfy the sufficient condition for horizon avoidance: a thin shell collapses to form a black hole and test particles (unless too slow) cross into it in finite time. However, there may be difficulty in maintaining the expected range of the NEC violation, and stability against perturbations is not assured. On the other hand, expansion of a trapped region that was formed in a finite time of a distant observer leads to a firewall that contradicts the quantum energy inequality.

*Keywords*: Black holes; null energy condition; thin shells.

## 1. Introduction

Event horizon — the null surface that bounds the spacetime region from which signals cannot escape — is the defining feature of black holes in classical general relativity[1–3]. Nevertheless, it is a global teleological entity that is in principle unobservable[4,5]. Theoretical, numerical and observational studies therefore focus on other characteristic features of black holes[3,6]. A local expression of the idea of the absence of communications with the outside world is provided by the notion of a trapped region. It is a domain where both outgoing and ingoing future-directed null geodesics emanating from a spacelike two-dimensional surface with spherical topology have negative expansion[1,3,6,7]. The apparent horizon is defined as the

evolving outer boundary of the trapped region[1,6]. According to classical general relativity the apparent horizon is located inside the event horizons if the matter satisfies energy conditions[1,8].

However, quantum states can violate energy conditions[8]. Black hole evaporation proceeds precisely because $T_{\mu\nu} = \langle \hat{T}_{\mu\nu} \rangle$ violates the null energy condition (NEC): there is a null vector $k^\mu$ such that $T_{\mu\nu}k^\mu k^\nu < 0$. In this case the apparent horizon is outside the event horizon. In fact, the very existence of the latter is uncertain[9,10]. While existence of spacetime singularities is no longer prescribed, their appearance without the horizon cover ("naked") is not excluded either. This situation motivated the introduction of many models of non-singular ultra-compact objects[5].

We consider the formation of a trapped region with a non-singular boundary in the finite time of a distant observer as a definition of existence of a black hole. Working in the framework of semiclassical gravity with spherical symmetry[11] we derive the condition that allows an infalling observer to avoid horizon crossing. We then show that if the apparent horizon is regular, the energy-momentum tensor and the metric in its neighbourhood are determined by the Schwarzschild radius $r_g$ and its rate of change. Finally, we discuss intriguing implications of these results.

## 2. Spherical symmetry. Sufficient condition for horizon avoidance

A general spherically symmetric metric in Schwarzschild coordinates is given by

$$ds^2 = -e^{2h(t,r)} f(t,r) dt^2 + f(t,r)^{-1} dr^2 + r^2 d\Omega. \tag{1}$$

The function $f(t,r) = 1 - C(t,r)/r$ is coordinate-independent[12], where the function $C(t,r)$ is the Misner-Sharp mass[3,13]. In an asymptotically flat spacetime $t$ is the physical time of a distant observer.

Trapped regions exist only if the equation $f(t,r) = 0$ has a root[7]. This root (or, if there are several, the largest one) is the Schwarzschild radius $r_g(t)$. Apparent horizons are in general observer-dependent entities. However they are unambiguously defined in the spherically symmetric case for all spherical-symmetry preserving foliations[13]. In this case the apparent horizon is located at $r_g$. In the Schwarzschild spacetime $C(t,r) = 2M$ and $h = 0$, hence $r_g = 2M$.

Radial infall of non-gravitating massive or massless point particle and thin shell collapse[2] are the simplest models that elucidate the near-horizon geometry and its dynamics. In thin shell models the geometry inside the shell is given by the flat Minkowski metric. The matter content of the shell is given by the surface energy-momentum tensor. The trajectory of a massive shell or particle is parametrized by its proper time $\tau$ and expresses as $\big(T(\tau), R(\tau)\big)$ in the exterior Schwarzschild coordinates. Initially the shell is located outside its gravitational radius, $R(0) > r_g$. Its dynamics is obtained by using the so-called junction conditions[2,14].

The first junction condition is the statement that the induced metric $h_{ab}$ on the shell $\Sigma$ is the same on both sides $\Sigma^\pm$, $ds^2_\Sigma = h_{ab} dy^a dy^b = -d\tau^2 + R^2 d\Omega$. Since for

massive particles the four-velocity $u^\mu$ satisfies $u_\mu u^\mu = -1$, their trajectory obeys

$$\dot{T} = \frac{\sqrt{F + \dot{R}^2}}{e^H F}, \tag{2}$$

where $\dot{A} = dA/d\tau$, $H = h(T, R)$, $F = f(T, R)$. This condition is used to identify the radial coordinate of the shell in interior and exterior coordinates, $R_- \equiv R$.

Discontinuity of the extrinsic curvature $K_{ab}$ is described by the second junction condition[2,14] that relates it to the surface energy-momentum tensor. Given the exterior metric the junction conditions result in the equations of motion for the shell. The classical collapse in vacuum, and the resulting equation for $R(\tau)$, is simple enough to have an analytic solution $\tau(R)$, leading to the finite proper time $\tau(r_g)$ and infinite time $T(r_g)$.

This equation of motion is modified for a general exterior metric. Its solution for the collapsing shell and for a particle that falls into a formed black hole has some remarkable features[15–17]. Here we focus on the possibility of crossing the Schwarzschild sphere of an evaporating black hole ($dr/dt = r_g'(t) < 0$) in finite proper time. For a finite evaporation time $t_E$ the finite proper crossing time is equivalent to having a finite time $t_g$ of a distant observer. By monitoring the gap between the shell and the Schwarzschild radius[11,18],

$$X(\tau) := R(\tau) - r_g\big(T(\tau)\big), \tag{3}$$

we discover the sufficient condition for a thin shell to never cross its Schwarzschild radius. The same condition applies to the study of an infalling test particle into an existing black hole. The analysis is generalized to null shells and test particles[16].

The rate of approach to the Schwarzschild radius behaves as $\dot{X} = \dot{R} - r_g'(T)\dot{T}$. Close to the Schwarzschild radius we have $\dot{T} \approx -\dot{R}e^{-H}/F$, and hence

$$\dot{X} \approx \dot{R}(1 - |r_g'|e^{-H}/F). \tag{4}$$

If for a fixed $t$ the function $\exp(h)f$ goes to zero as $x := r - r_g \to 0$, then there is a stopping scale $\epsilon_*(\tau)$. If the shell comes to the Schwarzschild radius closer than $\epsilon_*$ the gap has to increase, $\dot{X} > 0$, evidently indicating in this case that the shell never collapses to a black hole.

The collapse outcome depends on the metric choice: the outgoing Vaidya metric with decreasing mass (and thus positive energy density) prevents formation of the horizon, but the shell either becomes superluminal[15] or develops a surface pressure[16]. On the other hand, the metric functions of Eqs. (5) and (6) below lead to horizon formation in finite time according to both infalling and distant observers.

## 3. Metric outside an apparent horizon

Using only one additional assumption it is possible to obtain the explicit form of the metric near $r_g$. In fact this metric satisfies the sufficient condition for the horizon avoidance. We consider an evaporating black hole that is formed at some distant

observer's finite time, i.e. its apparent horizon radius $r_g(t)$ is a decreasing function of time. In addition we assume that the horizon is regular (the standard "no drama at the horizon" postulate[19], where the established regularity of the classical results is assumed to hold in the quantum-dominated regime). The regularity is expressed by finite values of the curvature scalars that can be directly expressed in terms of the energy-momentum tensor[20], $\mathrm{T} := T^{\mu}{}_{\mu}$ and $\mathfrak{T} := T^{\mu\nu}T_{\mu\nu}$.

The existence of an apparent horizon and regularity assumptions strongly constrain the energy-momentum tensor, and consistency with the known results on the background of an eternal black hole[21,22] specify its limiting form uniquely. The leading terms in the $(tr)$ block of the energy-momentum tensor are determined by a single function $a(t)$ (that for evaporating black holes is identified below), and the functions $C(t)$ and $h(t)$ take the following form[20],

$$C = r_g(t) - a(t)\sqrt{x} + \frac{1}{3}x \ldots . \tag{5}$$

and

$$h = -\ln\frac{\sqrt{x}}{\xi_0(t)} + \frac{4}{3a}\sqrt{x} + \ldots , \tag{6}$$

where $x = r - r_g$. The function of time $\xi_0(t)$ is determined by the choice of the time variable and the asymptotic behaviour of the solutions of the Einstein equations. The metric takes a particularly simple form in ingoing Vaidya coordinates[20].

The energy-momentum tensor that corresponds to this metric violates the null energy condition in the vicinity of the apparent horizon[20]. The comoving density and pressure at the apparent horizon of an evaporating black hole are negative,

$$\rho = p = -\frac{r_g'^2}{16\pi\dot{R}^2 r_g^2}, \tag{7}$$

where $R$ is the radial coordinate of the comoving observer[17]. We focus on the question of horizon avoidance. Return first to a collapsing thin shell problem where the exterior metric is given now by Eq. (1) with the metric functions that are given in Eqs. (5) and (6) above. Expanding Eqs. (5) and (6) for $X \to 0$ the rate of approach is

$$\dot{X} = -\left(\frac{2|\dot{R}|}{a} - \frac{a}{2|\dot{R}|r_g}\right)\sqrt{X} + \ldots , \tag{8}$$

which is independent of the function $\xi_0$. Hence if a test particle is in the vicinity of the apparent horizon, $X \ll a^2$, it will cross the horizon unless $|\dot{R}| < a/(2\sqrt{r_g})$.

For the evaporating case we match the $r_g'$ that is obtained as a consistency requirement on the functions of $C$ and $h$ from the Einstein equations[20], with the known results[21] for the quasi-static mass loss $r_g' = -\kappa/r_g^2$. In Planck units $\kappa \sim 10^{-3} - 10^{-4}$, and we obtain

$$\xi_0 \approx \sqrt{\frac{\kappa}{2r_g}} \approx \frac{a}{2}. \tag{9}$$

We have seen that the violation of the NEC is necessary for the existence of a black hole. Such violations, however, are bounded by quantum energy inequalities (QEI). Adapting the QEI of Ref. 24 we obtain that the region where the NEC is violated is bounded by

$$x_{\max} < \frac{11}{960\pi |r_g'| r_g^2} \sim 1, \tag{10}$$

that is obtained by ignoring the sub-Planckian features of the bound[11]. Even this estimate fails short of the conventional estimate $x_{\max} \sim r_g$. Our results indicate either that the required negative energy density for having a Schwarzschild sphere at finite time $t_S$ cannot be maintained or the trapped regions forms at a much later stage of the collapse. Alternatively, the domain of validity of the self-consistent metric is much narrower than that of the approximate metric that is obtained by taking the backreaction into account[21]. If the black hole was formed in the finite time of a distant observer and is subsequently accreting, the situation is would more radical: on approaching the apparent horizon the comoving density, pressure and flux diverge as $1/X$, while the curvature invariants remain finite[17]. We see that growth of the trapped region leads a falling-in observer to encounter a firewall at its boundary that, moreover, violates the NEC.

Both scenarios may indicate that the semiclassical approximation and its associated classical notions are modified already at the horizon or larger scales. A rigourous analysis of this situation is in order.

Another question results from two properties we discussed above. On the one hand, existence of the apparent horizon requires violation of the NEC. Test particles cross it in finite time unless moving too slow. It is not clear how the apparent horizon that requires NEC violation for its existence reacts to a perturbation by infalling normal matter. Given that collapsing thin shells are known to cross the apparent horizon and form a black hole with nearly all their rest mass intact[23], it is possible that the answer will involve considerable adjustment of our concept of black holes.

**Acknowledgements**

The work of RBM was supported in part by the Natural Sciences and Engineering Research Council of Canada and the Distinguished Visitor Program of the Faculty of Science and Engineering of the Macquarie University. We thank Sabine Hossenfelder, Jorma Luoko, Pady Padmanabhan, Amos Ori, Don Page, Bill Unruh, Matt Visser, and Mark Wardle for useful discussions and critical comments.

**References**

1. S. W. Hawking and G. F. R. Ellis, *The Large Scale Structure of the Space-Time*, (Cambrdge University Press, 1973).
2. E. Poisson, *A Relativist's Toolkit*, (Cambridge University Press, Cambridge, 2004).

3. C. Bambi, *Black Holes: a Laboratory for Testing Strong Gravity* (Springer, Singapore, 2017).
4. M. Visser, *Phys. Rev. D* **90**, 127502 (2014).
5. V. Cardoso and P. Pani, *Nature Astr.* **1**, 586 (2017).
6. V. Faraoni, *Cosmological and Black Hole Apparent Horizons*, (Springer, Heidelberg, 2015).
7. B. Krishnam, *Quasi-local horizons*, in A. Ashtekar and V. Petkov, (eds.), *Springer Handbook of Spacetime*, (Springer, NY, 2014), p. 527.
8. P. Martín-Moruno and M. Visser, *Classical and Semi-classical Energy Conditions*, in F. S. N. Lobo (ed), *Wormholes, Warp Drives and Energy Conditions*, (Springer, 2017), p. 193; C. J. Fewster, *ibid.*, p. 215.
9. A. Ashtekar and M. Bojowald, *Class. Quant. Grav.* **22**, 3349 (2005); S. A. Hayward, *Phys. Rev. Lett.* **96,** 031103 (2006).
10. V. P. Frolov and G. A. Vilkovisky, *Phys. Lett. B* **106**, 307 (1981); T. A. Roman and P. G. Bergmann, *Phys. Rev. D* **28**, 1265 (1983); S. W. Hawking, `arXiv:1401.5761` (2014).
11. V. Baccetti, R. B. Mann, and D. R. Terno, *Class. Quant. Grav.* **35**, 185005 (2018).
12. J. M. Bardeen, *Phys. Rev. Lett.* 46, 382-385 (1981).
13. V. Faraoni, G. F. R. Ellis, J. T. Firouzjaee, A. Helou, and I. Musco, *Phys. Rev. D* **95**, 024008 (2017).
14. W. Israel, *Nuovo Cimento* **44B**, 1 (1966).
15. B. Arderucio-Costa, W. G. Unruh, *Phys. Rev. D* **97**, 024005 (2018).
16. R. B. Mann, I. Nagle, and D. R. Terno, *Nucl. Phys. B* **936**, 19 (2018).
17. D. R. Terno, *Phys. Rev. D* **100**, 124025 (2019).
18. H. Kawai, Y. Matsuo, and Y. Yokokura, *Int. J. Mod. Phys. A* **28**, 1350050 (2013).
19. R. B. Mann, *Black Holes: Thermodynamics, Information, and Firewalls* (Springer, New York, 2015).
20. V. Baccetti, R. B. Mann, S. Murk, and D. R. Terno, *Phys. Rev. D* **99**, 124014 (2019).
21. R. Brout, S. Massar, R. Parentani, P. Spindel, *Phys. Rep.* **260**, 329 (1995).
22. A. Levi and A. Ori, *Phys. Rev. Lett.* **117**, 231101 (2016).
23. V. Baccetti, S. Murk, and D. R. Terno, *Phys. Rev. D* **100**, 064054 (2019).
24. C. J. Fewster, and M. J. Pfenning J. Math. Phys. **47**, 082303 (2006).

# Statistical description of black hole entropy in terms of trapped gravitons and its physical consequences

Stefano Viaggiu

*Dipartimento di Matematica, Universietà di Roma "Tor Vergata",*
*Via della Ricerca Scientifica, 1, I-00133 Roma, Italy*
*E-mail: viaggiu@axp.mat.uniroma2.it,*
*Universitá Guglielmo Marconi, Via Plinio 44, I-00193 Rome,*
*INFN, Sezione di Napoli, Complesso Universitario di Monte S. Angelo,*
*Via Cintia Edificio 6, 80126 Napoli, Italy*

The well known Schwarzschild black hole entropy is depicted in terms of trapped gravitons within the event horizon. A discrete spectrum for the so trapped gravitons is obtained and used to calculate thermodynamic quantities. The semi-classical expression for the black hole entropy is obtained with a temperature proportional to the usual Bekenstein-Hawking one and as a result a pressure term arises in the first law of thermodynamic. My approach is an attempt to obtain the semi-classical black hole entropy in terms of degrees of freedom stored inside (but 'near') the event horizon. Moreover, it is also shown that by modifying the internal energy by a term motivated by quantum Planckian fluctuations, a phase transition emerges during evaporation process with a vanishing specific heat at Planckian scales, thus representing the end of the evaporation process.

*Keywords*: Black Hole Thermodynamics; Black Hole Entropy; Trapped Gravitons.

## 1. Black Hole Entropy from Trapped Gravitons

As well known[1], black holes in asymptotically flat spacetimes are equipped with a non vanishing entropy given by $S_{BH} = \frac{A_h}{4L_P^2}$ and temperature $T_h = \frac{c\hbar}{4\pi K_B R}$ $(r \to \infty)$. Open questions are:

- What is the physical origin of the degrees of freedom implying $S_h$?
- Where is the lost information during evaporation?
- Where is the $PdV$ term in the first law?
- How can we obtain the log. corrections to the BH entropy?

In usual approaches (String, Loop quantum gravity) one quantize the area of the black hole: nevertheless the physical (statistical) origin of the BH entropy degrees of freedom is still obscure.

Moreover, what is the fate of the matter-radiation falling inside the event horizon?

In[2,3], I have proposed a possible description of the Black Hole entropy in terms of massless excitations provided by gravitons trapped inside the event horizon located at proper areal radius $R = 2GM/c^2$, with $M$ the ADM mass of the Black Hole.

The Bekenstein-Hawking (B-H) entropy is a semi-classical formula obtained with arguments involving classical general relativity with ordinary quantum mechanics. In this regard, the B-H entropy is 'phenomenological' and universal since it does not assume an underlying quantum gravity proposal, and for this reason it is expected to

hold in any physically sound quantum gravity theory. In particular, it is expected in low energy regime of the quantum gravity theory. Moreover, logarithmic corrections are expected to arise when quantum Planckian fluctuations in a full quantum gravity regime come into action.

To star with, we must obtain a suitable approximate formula for trapped gravitons within a box of radiaus $R$. Consider a gravitational wave in the vacuum and then put this in a spherical box, not a simple task!

Fortunately, we can use the well known perturbation formalism of the general relativity (Regge-Wheeler 1957, Zerilli $\sim$ 1970, Chandrasekhar 1984) on a Minkowskian spacetime $\eta_{ik} = diag(-1, +1, +1, +1)$ (Cartesian coordinates), perturbed by $h_{ik}$:

$$g_{ik} = \eta_{ik} + h_{ik}, \quad |h_{ik}| << |\eta_{ik}|.$$

After imposing Dirichelet boundary conditions on the Regge-Wheeler or Zerilli (In a Minkowski spacetime equations are the same) equation

$$Z^{(a,p)}_{\ell m,r,r} + \omega^2 Z^{(a,p)}_{\ell m} = \frac{\ell(\ell+1)}{r^2} Z^{(a,p)}_{\ell m},$$

at $r = R$ we get the following approximate formula for the allowed angular frequencies:

$$\omega_{\ell n} \simeq \frac{c}{2R}\,(2 + \ell + 2n)\,\pi, \;\; \ell \geq 2, \;\; n \in \mathbb{N}, \tag{1}$$

where $n$ is the radial quantum number and $\ell$ is interpreted as a species index (quadrupolar, sextupolar...gravitons). With the (1), we can build the ordinary statistical mechanics by calculating the partition function for $N$ oscillators[2]: we obtain the correct equation of state expected for radiation, i.e. $PV = \frac{U}{3}$, where $P$ is the pressure and $U$ the internal energy. In this way the pressure, missing in the usual treatment with respect to an external observer, emerges from the interior of the black hole. The first law of thermodynamics requires that $T_i = \alpha T_h = \frac{\alpha c\hbar}{4\pi K_B R}$, $\alpha \in (0, \infty)$. For a radiation field from the integrated first law we expect $\alpha = 2$. For the entropy we thus obtain the expression:

$$S = K_B Y(\alpha) \frac{A_h}{4L_P^2}, \;\; A_h = 4\pi R^2,$$

$$Y(\alpha) = \frac{b}{\alpha\pi^2\left(3 + e^{\frac{2\pi^2}{\alpha}}\right)} \tag{2}$$

$$b = -\alpha e^{\frac{4\pi^2}{\alpha}} \ln\left(1 - e^{-\frac{2\pi^2}{\alpha}}\right) - \alpha e^{\frac{4\pi^2}{\alpha}} \ln\left(1 - e^{-\frac{4\pi^2}{\alpha}}\right) + 6\pi^2 +$$

$$+ 2\pi^2 e^{\frac{2\pi^2}{\alpha}} + \alpha \ln\left(1 - e^{-\frac{2\pi^2}{\alpha}}\right) + \alpha \ln\left(1 - e^{-\frac{4\pi^2}{\alpha}}\right). \tag{3}$$

Hence $S = S_h$ iff. $Y(\alpha) = 1$ with a unique solution $\alpha \simeq 2.2$. If a summation over the azimuthal index $m$ is performed, we obtain $\alpha \simeq 8.48 \rightarrow$ that must be ruled out because too large with respect to the approximations made.

In[4] also the logarithmic corrections have been obtained by considering quantum Planckian fluctuations motivated by a quantum spacetime[6–8]. It is also interesting to note that, by taking the internal energy 'dressed' with Planckian fluctuations, i.e. $U(R) = \frac{c^4}{2G}R + \frac{C_1 c^4 L_P^2}{2GR}$, $\{C_1\} \in \mathbf{R}^+$, evaporation stops[3] at the radius $R = R_0 > R_c = \sqrt{C_1}L_P$: as expected, Planckian fluctuations act as a repulsive force thus opposing to the gravitational one.

## 2. Remarks

We have obtained a statistical description of the black hole entropy in terms of trapped gravitons: why gravitons and not photons? First of all, the B-H bentropy has been obtained with semi-classical considerations concerning general relativity and quantum mechanics. Since the Schwarzschild metric is a vacuum solutions of Einstein's equations, we expect that, thanks to the equivalence principle, only gravitons can survive inside the horizon. In fact, any other matter-energy is equipped with a non vanishing energy momentum tensor $T_{\mu\nu}$, and in a quantum context it will appear the mean value of $T_{\mu\nu}$ in some quantum state, i.e. $< T_{\mu\nu} >$, in the field equations, and as a consequence a non vacuum solution emerges.

It should be noticed that we have obtained the B-H entropy only with two reasonable assumptions. First of all, we have assumed for the internal energy the expression $U_h = U_i = Mc^2$, where $M$ is the ADM mass. This is a rather natural assumption. In practice, from the interior point of view the black hole is equipped with the same mass-energy as seen by an external observer with the entropy $S_h = \frac{K_B A_h}{4L_P^2}$. This entropy is interpreted in different ways. The external observer measures the Hawking radiation at spatial infinity with temperature $T_h$ and thus associates a non vanishing entropy to this radiation: the external observer cannot see the interior of the black hole and thus cannot have direct experience of the degrees of freedom giving the black hole entropy. Conversely, from the interior point of view, gravitons are stored inside and provide the degrees of freedom leading to the B-H entropy. In some sense, the point of view of the interior is 'dual' to the one present from the exterior.

The second assumption is the validity of the first law of thermodynamics. The application of the first law implies that the internal temperature is proportional to the one at the exterior $T_h$ by a factor 2: in this way a pressure term naturally arises in the first law.

We have obtained for $\alpha$ a numerical value, $\sim 2.2$, that is near the true value $\alpha = 2$ within the approximations made. To this purpose, note that the function $Y(\alpha)$ in[3] is strongly sensitive to $\alpha$ and therefore our consistent check is far from trivial.

An interesting property of our construction is that we have not used an underlying quantum gravity theory and thus our approach, according to the aim of the Hawking drivation, is 'phenomenological' and thus it is expected to be physically sound, since we used only sound arguments of general relativity, quantum mechanics

and statistical mechanics. In this regard is rather intriguing the fact that, as a consequence of the two assumptions depicted above, we obtain that $N \sim A$. This means that we obtain a result according to the holographic principle without assuming it from the onset: this results is a simple exact consequence of the fact that the mass of the black hole is the ADM one and the validity of the first law, no other assumptions are necessary.

In[5] we have extended the results in[2–4] for any massless excitations. This open the doors to a physical description of the cosmological constant.

## References


1. S. Hawking, *Commun. Math. Phys.* **43**, 199 (1975).
2. S. Viaggiu, *Physica A* **473**, 412 (2017.
3. S. Viaggiu, *Physica A* **488**, 72 (2017).
4. S. Viaggiu, *Int. J. Mod. Phys. D* **27**, 1850061 (2018).
5. S. Viaggiu, *Int. J. Mod. Phys. A* **33**, 1850074 (2018).
6. S. Doplicher, K. Fredenhagen and J. E. Roberts, *Comm. Math. Phys.* **172**, 187 (1995).
7. L. Tomassini and S. Viaggiu, *Class. Quantum Grav.* **28**, 075001 (2011).
8. L. Tomassini and S. Viaggiu, *Class. Quantum Grav.* **31**, 185;001 (2014).

# Squashed and magnetized black holes in five dimensional minimal gauged supergravity

Jose Luis Blázquez-Salcedo* and Jutta Kunz

*Institut für Physik, Universität Oldenburg*
*Postfach 2503, D-26111 Oldenburg, Germany*
**E-mail: jose.blazquez.salcedo@uni-oldenburg.de*

Francisco Navarro-Lérida,

*Departamento de Física Teórica and IPARCOS, Ciencias Físicas,*
*Universidad Complutense de Madrid, E-28040 Madrid, Spain*

Eugen Radu

*Departamento de Física da Universidade de Aveiro and CIDMA,*
*Campus de Santiago, 3810-183 Aveiro, Portugal*

*Keywords*: Black holes; Solitons; Supersymmetric solutions; Supergravity; Einstein-Maxwell-Chern-Simons theory; AdS/CFT correspondence.

## 1. Introduction

Black holes in gravity models with a negative cosmological constant are of considerable interest because of their fundamental role in the Anti-de Sitter/Conformal Field Theory (AdS/CFT) correspondence[1,2], in which an equivalence between classical gravitational solutions with AdS asymptotics in $D-$dimensions and field theory strongly coupled states in $(D-1)-$dimensions is proposed.

In this context the solutions of five dimensional $\mathcal{N}=4$ $SO(6)$ gauged supergravity are of special interest, since this theory is thought to be a consistent truncation of type IIB supergravity on $AdS_5 \times S^5$ [3,4]. The simplest version of the model corresponds to the bosonic sector of the theory, which results in Einstein-Maxwell (EM) theory with a Chern-Simons (CS) term and a negative cosmological constant. The action for $D=5$ minimal gauged supergravity is the following

$$I=-\frac{1}{16\pi}\int_{\mathcal{M}} d^5\mathrm{x}\sqrt{-g}\left[R+\frac{12}{L^2}-F_{\mu\nu}F^{\mu\nu}-\frac{2\lambda}{3\sqrt{3}}\varepsilon^{\mu\nu\alpha\beta\gamma}A_\mu F_{\nu\alpha}F_{\beta\gamma}\right]. \quad (1)$$

Here $R$ is the curvature scalar, $L$ is the AdS length scale (related with the cosmological constant $\Lambda=-6/L^2$) and $A_\mu$ is the $U(1)$ gauge potential with corresponding field strength tensor $F_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$. Note that $\lambda$ is the CS coupling constant, but in the minimal gauged supergravity case this is fixed to $\lambda=1$.

The field equations of this model is a generalized version of the Einstein-Maxwell equations,

$$R_{\mu\nu}-\frac{1}{2}Rg_{\mu\nu}-\frac{6}{L^2}g_{\mu\nu}=2F_{\mu\rho}F^{\rho}{}_{\nu}-\frac{1}{2}F^2g_{\mu\nu}\ ,\ \ \nabla_\nu F^{\mu\nu}=\frac{-\lambda}{2\sqrt{3}}\varepsilon^{\mu\nu\alpha\beta\gamma}F_{\nu\alpha}F_{\beta\gamma}\ . \quad (2)$$

This model has a number of interesting solutions. For instance, static and spherically symmetric black holes (the Schwarzschild-Tangherlini solution), stationary vacuum black holes (the Myers-Perry-AdS solution, with two different angular momenta in general), or the static, spherically symmetric and electrically charged configuration (the Reissner-Nordström-AdS black hole). More interesting are the stationary and electrically charged configurations (the Chong-Cvetic-Lü-Pope solution). These solutions possess four global charges: the mass, two angular momenta, and the electric charge[5]. The symmetry can be enhanced when the two angular momenta are equal in magnitude. This particular case was obtained by Cvetič, Lü and Pope (CLP)[6,7]. They possess a regular extremal limit, which contains a subset of solutions that preserves a fraction of supersymmetry[8]. Several generalizations of the CLP solution have been presented in[16–18].

In this paper we will review a particular family of solutions that generalize the CLP black hole. We will see that these new solutions, in addition to the mass, angular momentum and charge, are characterized by some non-trivial asymptotics (magnetization and squashing of the AdS boundary). These solutions are motivated by similar configurations studied in simpler settings, such as in Einstein-Maxwell-AdS theory in 4 dimensions[9–11], or static configurations (black holes with a solitonic limit) that are supported by a purely magnetic gauge field[12]. The box-like behaviour of AdS space-time, when combined with some nontrivial asymptotical multipolar behavior of the electromagnetic field at the boundary can be used to obtain new classes of regular solitons, that can be combined with a horizon to construct black holes. We will present the properties of the magnetized configurations with $AdS_5$ asymptotics[13], and of the more general squashed and magnetized solutions[14,15], which possess only a locally-AdS boundary, but allow for the existence of a new family of supersymmetric black holes in the extremal limit.

## 2. Ansatz, asymptotic behavior and charges

Since we are interested in configurations with equal angular momenta, spherical topology of the horizon, and asymptotical AdS behavior, we can propose the following Ansatz for the line element,

$$ds^2 = -f(r)\left(1+\frac{r^2}{L^2}\right)dt^2 + \frac{m(r)dr^2}{f(r)\left(1+\frac{r^2}{L^2}\right)} + \frac{r^2 m(r)}{4f(r)}(d\theta^2+\sin^2\theta d\phi^2) + \frac{r^2 n(r)}{4f(r)}\left(d\psi+\cos d\theta d\phi - \frac{2\omega(r)}{r}dt\right)^2 , \qquad (3)$$

where $r$ and $t$ are the radial and temporal coordinates, and $\theta$, $\phi$, $\psi$ are the Euler angles on $S^3$. The gauge field compatible with this line element (3) contains in general an electric and a magnetic potential, $A = a_0(r)dt + a_k(r)(d\psi+\cos d\theta d\phi)/2$.

A first requirement to the solutions we are going to consider is the existence of a non-trivial magnetic field on the AdS boundary[13], $\Phi_m = \frac{1}{4\pi}\int_{S^2_\infty} F = -c_m/2$ which can be understood as the magnetic flux passing through the base space $S^2$ of the $S^1$ fibration.

In addition, we can relax the asymptotical behaviour of the line element so that the boundary space is no longer a global $AdS_5$ manifold, but a space with only the local AdS symmetries. We refer to this as a squashed AdS boundary.

These two requirements on the asymptotical behaviour of the configuration impose the following expansion on the functions when close to the boundary:

$$f(r) = 1 + \frac{4}{9}(1-v^2)\left(\frac{L}{r}\right)^2 + \dots \ , \quad m(r) = 1 - \frac{1}{9}(1-v^2)\left(\frac{L}{r}\right)^2 + \dots \ ,$$
$$n(r) = v^2\left(1 + \frac{17}{9}(1-v^2)\left(\frac{L}{r}\right)^2\right) + \dots \ , \quad w(r) = \hat{j}\frac{1}{r^3} + \dots \ ,$$
$$a_0(r) = -\frac{q}{r^2} + \dots \ , \quad a_k(r) = c_m + \left(\mu - 2c_m L^2 v^2 \log\left(\frac{r}{L}\right)\right)\frac{1}{r^2} + \dots \ . \quad (4)$$

The full asymptotic expansion can be found in [14,15]. It depends on the squashing parameter $v$ (where $v = 1$ for the standard $AdS_5$ asymptotics) and the magnetization parameter $c_m$, apart from other standard parameters that are related with the angular momentum ($\hat{j}$), electric charge ($q$) and magnetic moment ($\mu$) (the parameters related with the mass appear only in the next order of the expansion, which is not shown here for simplicity).

This expansion has important consequences for the total charges of the configuration [14,15]. For instance, now we have to distinguish between the Page charge and the holographic charge, calculated with the boundary flux formalism. It is possible to see that

$$Q^{(p_0)} = \frac{1}{4\pi}\int_{S^3_\infty} d\Sigma_3(\sqrt{-g}F^{rt} - p_0\frac{\lambda}{\sqrt{3}}\varepsilon^{abc}A_a F_{bc}) = -\pi\left(qv - p_0\frac{2\lambda}{\sqrt{3}}c_m^2\right) \ , \quad (5)$$

corresponds to the Page charge for $p_0 = 1$ (whose conservation is implied by the field equations (2)), and to the holographic charge for $p_0 = 2/3$ (which is calculated using the boundary counterterms of the action (1) and the corresponding stress-energy tensor at the boundary [14,15]). Note that both charges coincide in the absence of magnetic flux at the boundary.

In addition to this far-field expansion, it is possible to study perturbatively the properties of solutions with a horizon (in which case the horizon charges can be computed, such as the area, horizon mass, horizon angular momentum, etc). It is also possible to find perturbatively regular solitonic solutions, but only in the case that the Page charge of such solutions vanishes [14,15]).

Despite the existence of these perturbative analytical solutions, the full construction of the black holes and the solitons extending up to the AdS boundary has to be performed numerically. In the next sections we will discuss the properties of such solutions.

## 3. Magnetized and squashed configurations

Let us start with the simplest case, in which the squashing parameter is $v = 1$, but the solutions possess some non-trivial magnetization $c_m \neq 0$.

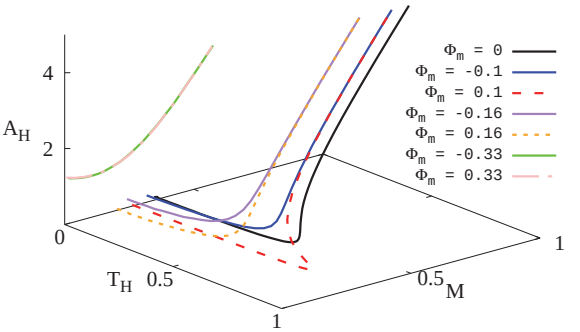



Fig. 1. Area ($A_H$) versus horizon temperature $T_H$ and total mass $M$ for the magnetized black holes. In black, the family CLP solutions with angular momentum $J = 0.003$, and (holographic) charge $Q = 0.044$. In colors, families of magnetized solutions with the same angular momentum and holographic charge, but different values of the magnetic flux at the $AdS_5$ boundary.

In Figure 1 we show the relation between the horizon area ($A_H$), horizon temperature ($T_H$) and total mass ($M$) of the magnetized black holes, for several values of the magnetization at the boundary ($\Phi_m$). Note that in general, the behavior of the magnetized solutions is very similar to the standard CLP black hole: one can see that there is a regular extremal limit (zero temperature), and that depending on the ratio between the different parameters, the families can possess a local maximum of the temperature (a phase transition between small and large black holes).

Apart from the previous families which possess an extremal limit, it is also possible to obtain families of black holes that asymptotically deform to magnetic solitons. To do so, one has to require the Page charge to vanish, and only then it is possible to continuously decrease the horizon size (as compared to the AdS length). In such a case the horizon shrinks and vanishes. We construct numerically the full solitonic configuration and compare this solution to the asymptotical behavior of the shrinking black holes, obtaining a perfect match. A more detailed discussion of all these configurations can be found in here[13].

It is possible to obtain numerically the squashed generalizations of the previous magnetized CLP black holes, by introducing the new parameter $v$ in the asymptotical behaviour of the metric functions. Such solutions were first presented in[14,15]. The space of solutions has very similar properties to what we show in Figure 1, but now with an extra parameter (the squashing $v$). Similar solitonic solutions can be constructed, and the black holes possess a regular extremal limit as well.

## 4. Supersymmetric solutions

More interesting is the existence of supersymmetric solutions contained in certain limits of the previous squashed and magnetized configurations. Some of them are well known (the Gutowski-Reall black hole[8], which is contained in the extremal limit of the CLP black hole, and hence it is unmagnetized and the AdS boundary is

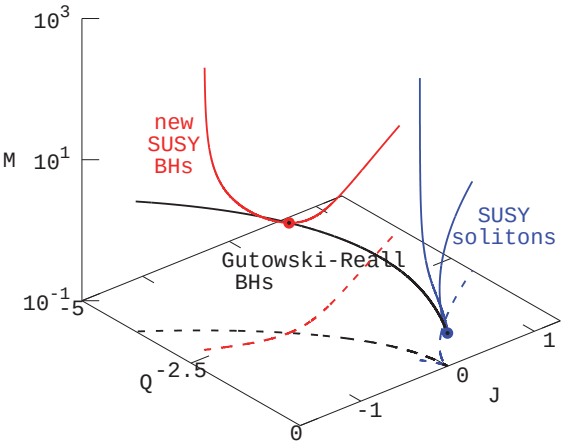



Fig. 2. Mass ($M$) versus holographic charge $Q$ and angular momentum $J$ for three families of supersymmetric solutions: in black, the Gutowski-Reall black holes; in blue the supersymmetric magnetized solitons with squashing of the boundary; in red the new squashed and magnetized supersymmetric black holes.

not deformed), or supersymmetric solitons[19] (which are both squashed and magnetized). In references[14,15] it was shown that an additional family of supersymmetric solutions, both squashed and magnetized, exist and branches off from a particular Gutowski-Reall configuration. In fact, all the supersymmetric solutions have a fixed relation between magnetization and squashing. The black holes can be connected to the supersymmetric solitons of Cassini and Martelli[19] only by increasing the black hole temperature, meaning they are connected by a family of non-extremal black holes (and hence non-supersymmetric).

In Figure 2 we show the mass, holographic charge and angular momentum for these three families of supersymmetric solutions. Interestingly, the new black holes (in red) have a 'frozen' horizon: although they form a one parameter family of solutions as it can be seen in the Figure, they all share the same near-horizon geometry (for instance, the same as the particular Gutowski-Reall black hole from which they branch off). They also share the same Page charge, but not the holographic charge, because of the non-trivial magnetization of the boundary.

## 5. Conclusions

We have presented here a summary of the properties of the new classes of magnetized and squashed black holes that generalize the CLP black hole by allowing for non-trivial asymptotics. In addition, the new black holes can be deformed continuously into solitons, without horizon. The black holes possess a regular extremal limit, and a particular subset of these extremal black holes turn out to be supersymmetric. Recently, further generalizations of these black hole solutions have been studied in more general settings[20–23].

**Acknowledgments**

JLBS and JK would like to acknowledge support by the DFG Research Training Group 1620 Models of Gravity. JLBS would like to acknowledge support from the DFG project BL 1553. The work of E.R. was supported by Funda¸c˜ao para a Ciˆencia e a Tecnologia (FCT), within project UID/MAT/04106/2019 (CIDMA), and by the FCT grant PTDC/FIS-OUT/28407/2017. JLBS, JK and ER would also like to acknowledge networking support by the COST Action CA16104 GWverse.

**References**

1. E. Witten, Adv. Theor. Math. Phys. **2** (1998) 253
2. J. M. Maldacena, Adv. Theor. Math. Phys. **2** (1998) 231 [Int. J. Theor. Phys. **38** (1999) 1113]
3. M. Cvetic *et al.*, Nucl. Phys. B **558** (1999) 96
4. J. P. Gauntlett, E. O Colgain and O. Varela, JHEP **0702** (2007) 049
5. Z. W. Chong, M. Cvetic, H. Lu and C. N. Pope, Phys. Rev. Lett. **95** (2005) 161301
6. M. Cvetic, H. Lu and C. N. Pope, Phys. Lett. B **598** (2004) 273
7. Z. W. Chong, M. Cvetic, H. Lu and C. N. Pope, Phys. Lett. B **644** (2007) 192 Z. W. Chong, M. Cvetic, H. Lu and C. N. Pope, Phys. Rev. D **72** (2005) 041901 M. Cvetic, H. Lu and C. N. Pope, Phys. Rev. D **70** (2004) 081502
8. J. B. Gutowski and H. S. Reall, JHEP **0402** (2004) 006
9. C. Herdeiro and E. Radu, Phys. Lett. B **749** (2015) 393
10. C. Herdeiro and E. Radu, Phys. Lett. B **757** (2016) 268
11. C. A. R. Herdeiro and E. Radu, Phys. Rev. Lett. **117** (2016), 221102
12. J. L. Blázquez-Salcedo, J. Kunz, F. Navarro-Lérida and E. Radu, Entropy **18** (2016) 438
13. J. L. Blázquez-Salcedo, J. Kunz, F. Navarro-Lérida and E. Radu, Phys. Lett. B **771** (2017) 52
14. J. L. Blázquez-Salcedo, J. Kunz, F. Navarro-Lérida and E. Radu, JHEP **1802**, 061 (2018)
15. J. L. Blázquez-Salcedo, J. Kunz, F. Navarro-Lérida and E. Radu, Phys. Rev. D **97**, no. 8, 081502 (2018)
16. J. L. Blázquez-Salcedo, Entropy **19**, 567 (2017)
17. J. Kunz, J. L. Blázquez-Salcedo, F. Navarro-Lérida and E. Radu, J. Phys. Conf. Ser. **942**, no. 1, 012003 (2017)
18. J. L. Blázquez-Salcedo, J. Kunz, F. Navarro-Lérida and E. Radu, Phys. Rev. D **95**, no. 6, 064018 (2017)
19. D. Cassani and D. Martelli, JHEP **1408** (2014) 044
20. A. Bombini and L. Papini, arXiv:1903.00021 [hep-th]
21. D. Cassani and L. Papini, JHEP **1812**, 037 (2018)
22. H. L. Dao and P. Karndumri, Eur. Phys. J. C **79**, no. 3, 247 (2019)
23. U. Gran, J. Gutowski and G. Papadopoulos, Phys. Rept. **794**, 1 (2019)

# Non-linear perturbation of black branes at large $D$

Umpei Miyamoto

*RECCS, Akita Prefectural University, Akita 015-0055, Japan*
*E-mail: umpei@akita-pu.ac.jp*

The Einstein equations describing the black-brane dynamics both in Minkowski and AdS background were recently recast in the form of coupled diffusion equations in the large-$D$(imension) limit. Using such results in the literature, we formulate a higher-order perturbation theory of black branes in time domain and present the general form of solutions for arbitrary initial conditions. For illustrative purposes, the solutions up to the first or second order are explicitly written down for several kind of initial conditions, such as a Gaussian wave packet, shock wave, and rather general superposed sinusoidal waves. These could be the first examples describing the non-trivial evolution of black-brane horizons in time domain. In particular, we learn some interesting aspects of black-brane dynamics such as the Gregory-Laflamme (GL) instability and non-equilibrium steady state (NESS). The formalism presented here would be applicable to the analysis of various black branes and their holographically dual field theories.

*Keywords*: Black holes in higher dimensions; Gregory-Laflamme instability; AdS/CFT correspondence.

## 1. Introduction

It is important to understand the dynamics of higher-dimensional black objects, since it tells us much about the nature of higher-dimensional gravitational theories and their holographically dual quantum field theories. The strong non-linearity of gravity, however, usually prevents us from understanding the dynamical properties of black objects beyond the linear-perturbation regime without highly sophisticated skills of numerical computation.

The Gregory-Laflamme (GL) instability[1], which is a universal instability of higher-dimensional black objects, is a good example to see the above situation. Though the GL instability in the non-linear regime is quite interesting, its analysis needs sophisticated skills of numerical relativity[2]. While there exists a semi-analytic higher-order perturbation method[3], it seems applicable only to static problems.

Recently, Emparan, Suzuki, and Tanabe showed that the Einstein equations describing the horizon dynamics of black branes in both Minkowski and Anti-de Sitter (AdS) background are recast in the form of coupled non-linear diffusion-type equations when the number of spatial dimensions is large[4]. This result provides us with a unique approach to the non-linear dynamics of black objects in higher dimensions. The authors indeed showed that the unstable black strings converge to non-uniform black strings (NUBSs), which had been predicted to happen above a critical dimension[5], by solving the diffusion equations numerically with a few lines of *Mathematica* code.

Once the simple diffusion equations were obtained[4], it is natural to ask if the non-linear properties of black-brane dynamics can be understood analytically. In this work, we develop a systematic non-linear perturbation theory of asymptotically

flat and AdS black branes, allowing the perturbations to be dynamical. Using the Fourier and Laplace transformation to solve the partial differential equations (PDEs), the perturbation equations are solved order by order for given arbitrary initial conditions up to the integration associated with the inverse transformation.

While the formulation is so general that it would be applicable to various problems, we pick up several examples as the initial conditions, which are a Gaussian wave packet, a step-function like shock configuration, and quite general discretely superposed sinusoidal waves. For these examples, the integration associated with the inverse transformation is completed up to the first or second order, and the properties of solutions are examined. Through these examples, one will see the validity of formalism itself and some unknown, or yet-to-be-confirmed, non-linear properties of black-brane dynamics. In particular, in the case of asymptotically flat black branes, an interesting non-liner property of GL instability resulting from the mode-mode coupling is unveiled at the second order. In the case of shock propagation on asymptotically AdS black branes, the analytic description of non-equilibrium steady state (NESS), which was recently discussed in the Riemann problem of relativistic fluid mechanics and field theories[6], is presented.

In this article, we present the formulation and a few results in the asymptotically flat case. In the original paper Ref. 7 and talk, the results in the asymptotically AdS case are presented.

## 2. Perturbation equations and general form of solutions

In the large-$D$(imension) approach, the horizon dynamics of vacuum black branes without a cosmological constant are described by two functions, $m(t,z)$ and $p(t,z)$, where $t$ is time and $z$ is the spatial coordinate along which the horizon extends[4]. $m$ and $p$ represent the mass and momentum distributions along the horizon, respectively. $m \to +0$ corresponds to the pinching off of the horizon. The equations of motion for these quantities take form of coupled non-linear diffusion equations,

$$(\partial_t - \partial_z^2)m + \partial_z p = 0, \quad (\partial_t - \partial_z^2)p - \partial_z m = -\partial_z \left(\frac{p^2}{m}\right). \tag{1}$$

A uniform black-brane solution corresponds to $m(t,z) \equiv 1$ and $p(t,z) \equiv 0$. Since we are interested in the dynamical deformation of such a uniform solution, we introduce one-parameter families of $m(t,z)$ and $p(t,z)$, and expand them around the uniform black-brane solution,

$$m(t,z;\epsilon) = 1 + \sum_{\ell=1}^{\infty} m_\ell(t,z)\epsilon^\ell, \quad p(t,z;\epsilon) = \sum_{\ell=1}^{\infty} p_\ell(t,z)\epsilon^\ell, \tag{2}$$

where $\epsilon$ is a constant parameterizing the families. Substituting these expansions into Eq. (1), we obtain the equations of motion at $O(\epsilon^\ell)$ ($\ell \in \mathbb{N}$),

$$\dot{m}_\ell - m_\ell'' + p_\ell' = 0, \quad \dot{p}_\ell - p_\ell'' - m_\ell' = \psi_\ell, \tag{3}$$

where the dot and prime denote the derivatives with respect to $t$ and $z$, respectively. The right-hand side of right equation in (3), $\psi_\ell(t,z)$, which we call a source term, is a polynomial of the lower-order perturbations and their first spatial derivatives. For example, the source terms for $\ell=2$ and $\ell=3$ are given by

$$\psi_2 = -2p_1p_1', \quad \psi_3 = 2m_1p_1p_1' + m_1'p_1^2 - 2p_1p_2' - 2p_1'p_2. \tag{4}$$

We are looking for the general form of solutions to the perturbation equations (3), combining the Fourier and Laplace transformations. A similar technique is found to be used in Ref. 8 to analyze the higher-order perturbation of surface-diffusion equation, which is a single non-linear PDE. Before starting to solve Eq. (3), let us introduce the notations associated with the Fourier and Laplace transformations. For a given function, say $f(t,z)$, we shall denote its Fourier transformation with respect to $z$ by $\bar{f}(t,k)$, and its Laplace transformation with respect to $t$ by the corresponding capital letter $F(s,z)$. Namely,

$$\bar{f}(t,k) := \mathcal{F}[f(t,z)] = \int_{-\infty}^{\infty} f(t,z)e^{-ikz}dz, \tag{5}$$

$$F(s,z) := \mathcal{L}[f(t,z)] = \int_{0}^{\infty} f(t,z)e^{-st}dt. \tag{6}$$

With the notations introduced above, the Fourier-Laplace transformed version of Eq. (3) is written as coupled algebraic equations in a matrix form,

$$\boldsymbol{A}\begin{pmatrix}\bar{M}_\ell(s,k)\\ \bar{P}_\ell(s,k)\end{pmatrix} = \begin{pmatrix}\bar{m}_\ell(0,k)\\ \bar{p}_\ell(0,k)+\bar{\Psi}_\ell(s,k)\end{pmatrix}, \tag{7}$$

$$\boldsymbol{A} := \begin{pmatrix}s+k^2 & ik\\ -ik & s+k^2\end{pmatrix}. \tag{8}$$

The solution to Eq. (3) is obtained after multiplying Eq. (7) by $\boldsymbol{A}^{-1}$ from left and inversely transforming it,

$$\begin{pmatrix}m_\ell(t,z)\\ p_\ell(t,z)\end{pmatrix} = (\mathcal{F}^{-1}\circ\mathcal{L}^{-1})\left[\boldsymbol{A}^{-1}\begin{pmatrix}\bar{m}_\ell(0,k)\\ \bar{p}_\ell(0,k)+\bar{\Psi}_\ell(s,k)\end{pmatrix}\right]. \tag{9}$$

By simple algebra, the inverse matrix $\boldsymbol{A}^{-1}$ is found to be decomposed into two parts,

$$\boldsymbol{A}^{-1} = \sum_{\sigma=+,-}\frac{1}{s-s_\sigma(k)}\boldsymbol{B}_\sigma, \tag{10}$$

$$\boldsymbol{B}_\sigma := \frac{1}{2}\begin{pmatrix}1 & -\sigma i\\ \sigma i & 1\end{pmatrix}, \quad s_\sigma(k) := k(\sigma 1 - k). \tag{11}$$

After this decomposition, one can perform the inverse Laplace transformation $\mathcal{L}^{-1}$ in Eq. (9) to obtain

$$\begin{pmatrix}m_\ell(t,z)\\ p_\ell(t,z)\end{pmatrix} = \sum_{\sigma=+,-}\boldsymbol{B}_\sigma\begin{pmatrix}\mathcal{F}^{-1}[e^{s_\sigma(k)t}\bar{m}_\ell(0,k)]\\ \mathcal{F}^{-1}[e^{s_\sigma(k)t}\bar{p}_\ell(0,k)+e^{s_\sigma(k)t}*\bar{\psi}_\ell(t,k)]\end{pmatrix}, \tag{12}$$

where $*$ denotes the convolution with respect to time.

Equations (12) is exactly what we wanted, namely, the general form of solutions to the perturbation equations (3) for given arbitrary initial conditions, $m_\ell(0,z)$ and $p_\ell(0,z)$ ($\ell \in \mathbb{N}$).

## 3. Example of initial conditions: Superposed sinusoidal waves

As an example, we consider the situation where the black brane is initially given an $O(\epsilon)$ perturbation being a superposition of an arbitrary number of sinusoidal waves. This example is simple but interesting enough to see what happens in the non-linear regimes. We set the following initial conditions,

$$m_1(0,z) = \sum_{n=1}^{N} a_n \cos(k_n z + \varphi_n), \quad p_1(0,z) = m_1'(0,z), \tag{13}$$

$$m_\ell(0,z) = p_\ell(0,z) = 0 \quad \text{for all} \quad \ell \geq 2, \tag{14}$$

where $a_n$, $k_n$, and $\varphi_n$ ($n = 1, 2, \cdots, N$) are real constants. We shall compute the right-hand side of Eq. (9) order by order for these initial conditions.

Since we have no source term at $O(\epsilon)$, $\psi_1 \equiv 0$, we see from Eq. (9) that what to compute is only the inverse Fourier transformation of the initial spectra, $\bar{m}_1(0,k)$ and $\bar{q}_1(0,k)$, multiplied by $e^{s_\sigma(k)t}$. Then, we obtain the first-order solutions,

$$m_1(t,z) = \frac{1}{2}\sum_{n=1}^{N} a_n[(1+k_n)e^{s_+(k_n)t} + (1-k_n)e^{s_-(k_n)t}]\cos(k_n z + \varphi_n), \tag{15}$$

$$p_1(t,z) = -\frac{1}{2}\sum_{n=1}^{N} a_n[(1+k_n)e^{s_+(k_n)t} - (1-k_n)e^{s_-(k_n)t}]\sin(k_n z + \varphi_n). \tag{16}$$

Since we are considering linear equations of motion, there is no mode-mode coupling appearing in non-linear regime, and therefore Eqs. (15) and (16) allow simple interpretation. The factor of $\cos(k_n z + \varphi_n)$ in Eq. (15) represents the time-dependent amplitude of the initially given mode $\cos(k_n z + \varphi_n)$. Each mode evolves independently according to its growth or damping rate determined by $e^{s_+(k_n)t}$ and $e^{s_-(k_n)t}$. From the concrete form of $s_\pm(k)$ in Eq. (11), one can see that if $k_n \in (-1, 0)$ (resp. $k_n \in (0, 1)$), such a mode grows exponentially due to $e^{s_-(k_n)t}$ (resp. $e^{s_+(k_n)t}$), which represents the GL instability.

Since we assume that the initial perturbations vanish at $O(\epsilon^2)$, $m_2(0,z) = p_2(0,z) = 0$, we see from Eq. (9) that what to compute at the second order is the inverse Fourier transformation of the convolution between $e^{s_\sigma(k)t}$ and the Fourier spectrum of source term $\bar{\psi}_2(t,k)$.

Although the second-order solutions can be written down analytically, let us omit writing them down here since the space is restricted (see the original paper Ref. 7 for the explicit form of solutions). At the second order, the non-linear source term involves the mode-mode coupling, which is absent at the linear order. This coupling

excites the terms of $\cos[(k_n \pm k_{n'})z]$ and $\sin[(k_n \pm k_{n'})z]$. For example, one can find that both $\cos[(k_n + k_{n'})z]$ and $\cos[(k_n - k_{n'})z]$ terms involve the following three kind of time dependence,

$$e^{s_+(k_n+k_{n'})t}, \quad e^{s_-(k_n+k_{n'})t} \quad \text{and} \quad e^{s_+(k_n-k_{n'})t}, \quad e^{s_-(k_n-k_{n'})t}. \tag{17}$$

Thus, the second-order solutions exhibit a variety of dispersion given by the exponents of (17).

Let us consider the meaning to investigate the higher-order perturbations from the stability point of view. The asymptotically flat black brane we consider here is essentially unstable. Namely, if the initial perturbation contains any mode of which wave number $k_n \in (-1, 1) \setminus \{0\}$, such a mode grows unboundedly. However, we consider the black brane in the large-$D$ limit, namely, above the critical dimension[5]. Thus, the GL instability initially grows but it gradually damps in non-linear regimes, and eventually the horizon converges to non-uniform configuration[4].

A black brane that is linearly stable can become unstable at the second order. If the initial perturbation does not contain any unstable mode, the initial perturbation will damp exponentially at linear level. However, the second-order perturbation involves various time dependence as seen in Eq. (17). In order to see directly this situation, let us focus on a simple case as follow.

Suppose that the initial perturbation is the superposition of two modes $k_1$ and $k_2$ both of which are stable modes, $k_1 > k_2 > 1$. In addition, assume that their difference is smaller than unity, $k_1 - k_2 \in (0, 1)$. In this case, the term of $C_{21}^{(+)(-)}(t)\cos[(k_1 - k_2)z + (\varphi_1 - \varphi_2)]$ in $m_2(t, z)$ includes terms having growing factor $e^{s_+(k_1-k_2)t}$. Thus, the perturbation does not grow at $O(\epsilon)$ but does at $O(\epsilon^2)$.

The above phenomenon is an interesting aspect of the GL instability, which was revealed for the first time by the present non-linear perturbation theory in time domain. It is intuitively understandable. The superposition of the two modes forms the beat. For simplicity, assume $a_1 = a_2$ ($\neq 0$), then the superposed wave is written as

$$2a_1 \cos\left[\frac{(k_1 + k_2)z + (\varphi_1 + \varphi_2)}{2}\right] \cos\left[\frac{(k_1 - k_2)z + (\varphi_1 - \varphi_2)}{2}\right]. \tag{18}$$

This exhibits the fast spatial oscillation with the large wave number $\frac{k_1+k_2}{2}$ which is enveloped by the slow oscillation with the small wave number $\frac{k_1-k_2}{2}$, which is called the beat phenomenon especially when the difference of the wave numbers is rather small $k_1 - k_2 \ll k_1 + k_2$. This slow oscillation is nothing but the origin of the GL instability at the second order. One can observe that the beat formed by the superposition of two modes at $t = 0$. As soon as the dynamics starts, such an initial wave rapidly damps as predicted by the first-order perturbation. As the time proceeds, however, the waves of which scale is the same order as that of the beat begin to grow and eventually diverge.

**Acknowledgments**

The author would like to thank R. Emparan, R. Suzuki, and K. Tanabe for useful discussion and comments during The Spanish-Portuguese Relativity Meetings held at Lisbon (12–15th, Sep. 2016), and T. Torii during his stay at Akita (12–17th, Feb. 2017). This work was supported by JSPS KAKENHI Grant Number 15K05086 and 18K03652.

**References**

1. R. Gregory and R. Laflamme, Phys. Rev. Lett. **70** (1993) 2837.
2. L. Lehner and F. Pretorius, Phys. Rev. Lett. **105** (2010) 101102.
3. S. S. Gubser, Class. Quant. Grav. **19** (2002) 4825.
4. R. Emparan, R. Suzuki and K. Tanabe, Phys. Rev. Lett. **115** (2015) 9, 091102.
5. E. Sorkin, Phys. Rev. Lett. **93** (2004) 031601.
6. C. P. Herzog, M. Spillane and A. Yarom, JHEP **1608**, 120 (2016).
7. U. Miyamoto, JHEP **1706** (2017) 033.
8. U. Miyamoto, Phys. Rev. D **78** (2008) 026001.

# Unitary evaporation via modified Regge-Wheeler coordinate

Aizhan Myrzakul* and Michael R.R. Good†

*Physics Department,*
*Nazarbayev University,*
*Astana, Kazakhstan*

Constructing an exact correspondence between a black hole model, formed from the most simple solution of Einstein's equations, and a particular moving mirror trajectory, we investigate a new model that preserves unitarity. The Bogoliubov coefficients in 1+1 dimensions are computed analytically. The key modification limits the origin of coordinates (moving mirror) to sub-light asymptotic speed. Effective continuity across the metric ensures that there is no information loss. The black hole emits thermal radiation and the total evaporation energy emitted is finite without backreaction, consistent with the conservation of energy.

## 1. Introduction

The Hawking effect [1] has an exact correspondence [2] with the moving mirror. [3,4] The specific and analytically known accelerated boundary condition on the quantum field, $\psi$, located at the origin of coordinates, $r = 0$, corresponding to the location of the black hole singularity, depends explicitly on the form of the tortoise coordinate, $r^*$. The moving mirror perfectly reflects the field modes and accelerates with the precise trajectory, $t(x)$.

The physical effect of the mirror is that it arouses quantum field fluctuations reflecting virtual particles into real ones. The black mirror,[2] which is the aforementioned tortoise coordinate associated boundary condition, extracts energy indefinitely and does not preserve unitarity. However, we present a summary of a modified model [5] — "a drifting black mirror" — which resolves these two problems. It was demonstrated recently [6] that the new corresponding coordinate to $r^*$ is the generalized or giant tortoise coordinate (GTC). The giant tortoise coordinate results in unitarity preservation and finite energy emission of the black hole during evaporation.

In this MG15 proceedings contribution, first, the usual tortoise coordinate, $r^*$, and its relation to the moving mirror, $t(x)$, is briefly considered. Then we generalize this coordinate to the giant tortoise coordinate and investigate the correspondence between the black hole and the moving mirror in the context of the GTC. We find no information loss, finite evaporation energy, thermal equilibrium, analytical beta coefficients, and a left-over [7] remnant.

*aizhan.myrzakul@nu.edu.kz
†michael.good@nu.edu.kz

## 2. The Tortoise Coordinate and the Black Mirror

In this section the textbook matching solution[8] for the outside and inside of the black hole over the shock wave is derived.

Let us start from the usual tortoise coordinate, (the Regge-Wheeler $r^*$):

$$r^* \equiv r + 2M \ln\left(\frac{r}{2M} - 1\right). \tag{1}$$

Requiring the metric to be the same on both sides of the shock wave, $v_0$:

$$r(v_0, u_{\rm in}) = r(v_0, u_{\rm out}), \tag{2}$$

where

$$r(v_0, u_{\rm in}) = \frac{v_0 - u_{\rm in}}{2}, \qquad \text{and} \qquad r^*(v_0, u_{\rm out}) = \frac{v_0 - u_{\rm out}}{2}, \tag{3}$$

the tortoise coordinate, Eq. (1), can be rewritten as:

$$r(v_0, u_{\rm out}) + 2M \ln\left(\frac{r(v_0, u_{\rm out})}{2M} - 1\right) = \frac{v_0 - u_{\rm out}}{2}. \tag{4}$$

Solving this for the red-shift function $u_{\rm out}$ yields:

$$u_{\rm out} = u_{\rm in} - 4M \ln\frac{|v_H - u_{\rm in}|}{4M}. \tag{5}$$

The $u_{\rm out} = +\infty$ limit corresponds to the formation of a black hole event horizon location, $v_H \equiv v_0 - 4M$. Eq. (5) is exactly the matching solution[8] for the Eddington-Finkelstein background (exterior) to the Minkowski background (interior) with a strict event horizon.

Substituting $u_{\rm out} \equiv t(x) - x$ and $u_{\rm in} \equiv t(x) + x$ into Eq. (5), and solving for $t(x)$ gives the time-space trajectory of the black mirror[a]:

$$t(x) = v_H - x - 4Me^{x/2M}, \tag{6}$$

as investigated[b] in Good-Anderson-Evans (2016).[2] The range of the coordinates are: $0 < r < \infty$ and $-\infty < x < \infty$. Eq. (6) is transcendentally invertible and is an analytic relation between the black hole matching solution, Eq. (5), and the space-time trajectory of its moving mirror, $x(t)$.

## 3. The Giant Tortoise Coordinate and the Drifting Black Mirror

We impose a strong restriction on the maximum speed of the black mirror: it must always travel slower than light, even asymptotically.[c] That is, in any coordinate system the origin of the black hole, $r = 0$, should not speed away to null-future

[a] The black mirror, which is not eternally thermal,[9] is called Omex for short, due to the **Om**ega constant, $\Omega e^{\Omega} = 1$, and **ex**ponent argument.[2]

[b] See also the proceedings of the MG14 meeting[10,11] and the 2nd LeCosPA Symposium.[12]

[c] An exception in a different model can give finite energy and preserve information if the acceleration asymptotes to zero sufficiently fast.[13]

infinity, $\mathscr{I}^+$. This is highly restrictive and gives a new time-space trajectory[5] of the origin of coordinates (an asymptotically drifting mirror):

$$t(x,\xi) = v_H - \frac{x}{\xi} - 4Me^{\frac{x}{2M\xi}}. \tag{7}$$

Here $\xi$ is the asymptotic drifting speed which lies in the range $0 < \xi < 1$. The corresponding matching condition can be easily derived and is:

$$u_{\rm out} = u_{\rm in} - 4M\xi \ln\left[\frac{1-\xi}{2}\mathcal{W}\left(\frac{2e^{\frac{v_H - u_{\rm in}}{2M(1-\xi)}}}{1-\xi}\right)\right], \tag{8}$$

where $\mathcal{W}$ is the product log. When $\xi \to 1$, the matching solution Eq. (8) is equivalent to Eq. (5), meaning that one has operative formation of an effective event horizon. Thus the modest modification in Eq. (7) safeguards the formation of a black hole[d] that occur with Eq. (6).

Eq. (8) represents the world line of the origin. It is easy to see that this origin is a perfectly reflecting boundary as nothing can go behind $r = 0$ into the $r < 0$ space. Therefore we use the generalized matching solution Eq. (8) in order to find the giant tortoise coordinate (analogous to going backwards from Eq. (5) to Eq. (1)),

$$r^*(\xi) \equiv r + 2M\xi \ln\left[\frac{1-\xi}{2}\mathcal{W}\left(\frac{2e^{\frac{r-2M}{M(1-\xi)}}}{1-\xi}\right)\right]. \tag{9}$$

It is a crucial fact that these two coordinates, $r^*$ and $r^*(\xi)$, are effectively the same when $\xi \approx 1$. The distinction is that there is no singularity at $r = 2M$ in Eq. (9) as in Eq. (1) when $\xi \neq 1$: $r^*(\xi)_{r=2M} = 2M\left[1 - \xi\mathcal{W}(2/\epsilon)\right]$, where $\epsilon \equiv 1 - \xi$. The user may define the free parameter $\xi$ as close to $\xi \approx 1$ (for effective continuity) as is arbitrarily desired as long as strictly, $\xi < 1$ (for unitarity).

## 4. Unitarity: Finite Asymptotic Entanglement Entropy

Qualitatively, the black mirror correspondence demonstrates information loss by the acceleration horizon which prohibits some left-movers from hitting the mirror and becoming right-movers. We can say these modes are lost forever in the black hole. However, the drifting black mirror has an asymptotic approach to time-like future infinity, $i^+$, rather than null future infinity, $\mathscr{I}_L^+$ and all the left-movers hit the mirror and become right-movers, preserving information to an observer at $\mathscr{I}_R^+$.

Quantitatively, we see this result by the von-Neumann entanglement entropy[17] for the black mirror:[5]

$$S(t) = \frac{1}{6}\tanh^{-1}\left(\frac{1}{\mathcal{W}(2e^{2\kappa(v_H - t)}) + 1}\right), \tag{10}$$

[d]For more information on whether any type of horizon is formed during gravitational collapse taking into account quantum effects see, e.g. Ref. 14 (for horizonless models see Refs. 15, 16).

whose limit in the far future diverges:

$$S_f \equiv \lim_{t\to\infty} S(t) = \infty, \tag{11}$$

signaling information loss (see e.g. the entropy divergence in Ref. 18). Here $\kappa \equiv 1/4M$, the surface gravity for Schwarzschild background. In contrast, the drifting black mirror has entropy,

$$S_\xi(t) = \frac{1}{6}\tanh^{-1}\left(\frac{\xi}{\mathcal{W}(2e^{2\kappa(v_H-t)})+1}\right), \tag{12}$$

whose limit is

$$S_f \equiv \lim_{t\to\infty} S_\xi(t) = \frac{1}{6}\tanh^{-1}(\xi) = \frac{\eta}{6} \neq \infty. \tag{13}$$

The final asymptotic entropy is the drifting rapidity and its measure as a finite constant signals information preservation. To achieve effective equilibrium ($\xi \approx 1$), the final asymptotic entropy will be very large ($\eta \ggg 1$), but finite.

## 5. Finite Evaporation Energy

A prime advantage of the GTC, Eq. (9), is that during the collapse the global energy emission of the black hole is finite and analytic. The consistency of the result with the analytically computed beta Bogolubov coefficient can be shown via a numerical verification of the stress-energy tensor[5,19] whose total energy production is,

$$E = \frac{1}{96\pi M}\left(\gamma^2 + \frac{\eta}{\xi}\right), \tag{14}$$

where $\gamma \equiv 1/\sqrt{1-\xi^2}$ is the final drifting Lorentz factor, $\eta \equiv \tanh^{-1}\xi$ is the final drifting rapidity, with $\xi < 1$ corresponding to the final drifting speed. For high drifting speed (thermality), $\xi \approx 1$, then $\gamma^2 \gg \eta/\xi$, and:

$$E = \frac{\gamma^2}{96\pi M}. \tag{15}$$

One immediately sees that the energy diverges as the origin moves to the speed of light, (i.e. mirror drift, $\xi \to 1$ and $\gamma \to \infty$). Eq. (15) is the final expression for the energy emission of the thermal black hole which is finite and consistent with the conservation of energy. Finite energy is anticipated to be a primary result of backreaction,[8] yet, we have obtained consistency by restricting origin speed, $\xi < 1$.

## 6. Temperature and the Giant Tortoise Coordinate

Using the GTC, a constant energy flux plateau is apparent (high drifting speeds, $\xi \approx 1$). During equilibrium, $F = \pi T^2/12$. Expanding the temperature as a function of maximum energy flux (where the radiation is closest to equilibrium) gives the temperature of the black hole:[5]

$$T(\epsilon, M) = \frac{1}{8\pi M}\left[1 - 3\left(\frac{3}{4}\right)^{1/3}\epsilon^{2/3} + O(\epsilon)\right], \tag{16}$$

to lowest order in $\epsilon$ where $\epsilon \equiv 1 - \xi$. The first term in this expression corresponds to the equilibrium temperature of the unmodified black mirror model which uses the usual tortoise coordinate. The other terms correspond to the deviation due to sub-light speed drift, $\xi < 1$, which are negligible for small $\epsilon$. The modification of Eq. (6) to Eq. (7) still results in a constant energy flux plateau and effective long-term thermal equilibrium of Eq. (16) for $\xi \approx 1$. This confirms the robustness of the GTC model for describing Planckian distributed particle creation from a black hole.

## 7. The Beta Bogoliubov Coefficient

The particle spectrum,

$$\langle N_\omega \rangle \equiv \langle 0_{in} | N_\omega^{out} | 0_{in} \rangle = \int_0^\infty |\beta_{\omega\omega'}|^2 d\omega', \tag{17}$$

requires knowing the GTC beta Bogoliubov coefficient which is a simple integral to compute analytically[5] with result:

$$\beta_{\omega\omega'}(\xi) = -\frac{\xi\sqrt{\omega\omega'}}{2\pi\kappa\omega_p}\left(\frac{i\kappa}{\omega_p}\right)^A \Gamma(A), \tag{18}$$

where $\Gamma(x)$ is the gamma function, $A \equiv \frac{i}{2\kappa}[(1+\xi)\omega + (1-\xi)\omega']$ and $\omega_p \equiv \omega + \omega'$. The integrand[5] of Eq. (17) is Planckian using Eq. (18) with $\xi \approx 1$ and $\omega' \gg \omega$, which is consistent with Eq. (16).

## 8. Summary

| Quantity | Tortoise | Giant Tortoise |
|---|---|---|
| $r^*$ | $r^* \equiv r + 2M\ln\left(\frac{r}{2M} - 1\right)$ | $r^*(\xi) \equiv r + 2M\xi\ln\left[\frac{\epsilon}{2}\mathcal{W}\left(\frac{2}{\epsilon}e^{\frac{r-2M}{M\epsilon}}\right)\right]$ |
| $t(x)$ | $t = -x - 4Me^{x/2M}$ | $t(\xi) = -x/\xi - 4Me^{x/2M\xi}$ |
| $u_{\rm out}$ | $u_{\rm out} = u_{\rm in} - 4M\ln\frac{\lvert v_H - u_{\rm in}\rvert}{4M}$ | $u_{\rm out}(\xi) = u_{\rm in} - 4M\xi\ln\left[\frac{\epsilon}{2}\mathcal{W}\left(\frac{2}{\epsilon}e^{\frac{v_H - u_{\rm in}}{2M\epsilon}}\right)\right]$ |
| $S_f$ | $\infty$ | $S_f(\xi) = \eta/6$ |
| $E$ | $\infty$ | $E(\xi) = \frac{1}{96\pi M}\left(\gamma^2 + \frac{\eta}{\xi}\right)$ |
| $T$ | $T = \frac{1}{8\pi M}$ | $T(\xi) = \frac{1}{8\pi M} + \mathcal{O}(1-\xi)$, for $\xi \approx 1$ |
| $\beta$ | $\beta_{\omega\omega'} = -\frac{\sqrt{\omega\omega'}}{2\pi\kappa\omega_p}\left(\frac{i\kappa}{\omega_p}\right)^{\frac{i\omega}{\kappa}}\Gamma(\frac{i\omega}{\kappa})$ | $\beta_{\omega\omega'}(\xi) = -\frac{\xi\sqrt{\omega\omega'}}{2\pi\kappa\omega_p}\left(\frac{i\kappa}{\omega_p}\right)^A\Gamma(A)$ |

We have presented an overview[20] of a unitary black hole evaporation model that, without backreaction, manages to produce a finite total energy emission. Exact

analytic results for several important corresponding quantities are found: the beta Bogoliubov coefficients, the finite total energy emission, the matching condition for the modes and the generalized tortoise coordinate.

The model[e] relaxes uncompromising continuity across the shock wave in the metric in exchange for preserving information.[f] Arbitrary precision in continuity is permitted with arbitrary fast sub-light drifting speeds for the origin of coordinates (the moving mirror must travel at speeds less than light). As long as this requirement is met, the information is preserved and the energy emitted is finite. With ultra-relativistic, sub-light speeds the black hole emits particles in a Planck distribution with constant energy flux at equilibrium temperature.

## References


1. S. W. Hawking, Commun. Math. Phys. **43** (1975) 199.
2. M. R. R. Good, P. R. Anderson, C. R. Evans, Phys. Rev. D **94** (2016) 065010, [1605.06635].
3. S. A. Fulling, P. C. W. Davies, Proc. Roy. Soc. Lond. A **348** (1976) 393.
4. P. C. W. Davies, S. A. Fulling, Proc. Roy. Soc. Lond. A **356** (1977) 237.
5. M. R. R. Good, K. Yelshibekov, Y. C. Ong, JHEP **1703** (2017) 013, [1611.00809].
6. M. R. R. Good, Y. C. Ong, A. Myrzakul, K. Yelshibekov, [1801.08020].
7. F. Wilczek, IASSNS-HEP-93-012, [9302096].
8. A. Fabbri, J. Navarro-Salas, London, UK: Imp. Coll. Pr., 2005.
9. M. R. R. Good, Int. J. Mod. Phys. A **28** (2013) 1350008, [1205.0881].
10. P. R. Anderson, M. R. R. Good, C. R. Evans, MG14, [1507.03489].
11. M. R. R. Good, P. R. Anderson, C. R. Evans, MG14, [1507.05048].
12. M. R. R. Good, 2nd LeCoSPA Meeting, [1602.00683].
13. M. R. R. Good, World Scientific, Singapore (2018), [1612.02459].
14. R. B. Mann, I. Nagle, D. R. Terno, Nuc. Phys. B **936**, (2018) 19 [1801.01981].
15. M. R. R. Good, E. V. Linder, Phys. Rev. D **96**, (2017) 125010, [1707.03670].
16. M. R. R. Good, E. V. Linder, Phys. Rev. D **97**, (2018) 065006, [1711.09922].
17. P. Chen and D. H. Yeom, Phys. Rev. D **96**, (2017) 025016, [1704.08613].
18. M. R. R. Good, E. V. Linder, Phys. Rev. D **99**, (2019) 025009, [1807.08632].
19. M. R. R. Good, P. R. Anderson, C. R. Evans, Phys. Rev. D **88** (2013) 025023, [1303.6756].
20. M. R. R. Good, Universe **4**, no. 11, 122 (2018).
21. S. A. Fulling and J. H. Wilson, Phys. Scr. **94** (2019) 014004, [1805.01013].
22. C. Holzhey, F. Larsen and F. Wilczek, Nucl. Phys. B **424**, 443 (1994) [9403108].


[e]The analytically solvable theory exploits the special dual properties of conformal flatness/invariance.[21]

[f]Information as contained and expressed via the renormalized entanglement entropy.[22]

# Equal-spin limit of the Kerr–NUT–(A)dS spacetime

Eliška Polášková and Pavel Krtouš

*Institute of Theoretical Physics, Faculty of Mathematics and Physics, Charles University, V Holešovičkách 2, 180 00 Prague 8, Czech Republic*
**E-mail: eli.polaskova@email.cz, Pavel.Krtous@utf.mff.cuni.cz*

We summarize properties of the Kerr–NUT–(A)dS spacetime in a general dimension. We also investigate a limit when several rotational parameters in the metric are set equal and study geometry of the resulting spacetime. In dimension $D = 6$, we found a suitable Killing vector basis, whose algebraic structure proves that symmetries of the spacetime after the limit are further enhanced.

*Keywords*: Black holes; higher dimensions; rotation; NUT charges; symmetries.

## 1. Introduction

In four dimensions, the Plebański–Demiański metric[1] represents a large family of algebraic type D vacuum spacetimes. For example, it includes the rotating Kerr black hole, the Taub–NUT solution with one NUT parameter as well as accelerating black holes described by the C-metric.

In higher dimensions, the most general spacetime metric known so far is called the *Kerr–NUT–(A)dS metric*[2]. It includes rotational and NUT parameters as well as the cosmological constant, however, a solution with acceleration and charge is yet to be discovered.

The Kerr–NUT–(A)dS metric can describe various geometries, such as maximally symmetric spaces, a so-called Euclidean instanton and black hole solutions[3]. It also admits explicit and hidden symmetries that can be generated from a single object — a closed conformal Killing–Yano tensor, which we call the principal tensor[4,5]. Moreover, the principal tensor uniquely determines canonical coordinates in which the Hamilton–Jacobi, Klein–Gordon[6] and Dirac[7] equations are fully separable, and therefore the geodesic motion is completely integrable[8].

The main focus of our work is a limit of the metric when several rotational parameters are set equal. Other limit cases have already been studied, such as the limit with some of the black hole's rotations switched off[9] or the limit when particular roots of the metric function $X_\mu$ degenerate[10]. These papers have demonstrated that not only can performing various limits of the general metric shed light on the role of NUT charges, but it can also result in new interesting geometries.

## 2. Kerr–NUT–(A)dS Spacetime

This section introduces the general Kerr–NUT–(A)dS spacetime and describes some of its interesting properties. Firstly, the metric in its general form is discussed and subsequently, we show how to obtain a black hole geometry from the general metric as a special case.

For simplicity, we restrict ourselves to even dimensions $D = 2N$. Note that the generalization to odd dimensions is straightforward — a corresponding term is added to the metric.

### 2.1. *Metric*

A metric describing the Kerr–NUT–(A)dS geometry can be written in the form[2]

$$\boldsymbol{g} = \sum_{\mu} \left[ \frac{U_\mu}{X_\mu} \mathbf{d}x_\mu^2 + \frac{X_\mu}{U_\mu} \left( \sum_k A_\mu^{(k)} \mathbf{d}\psi_k \right)^2 \right] , \tag{1}$$

with greek and latin indices going over slightly different ranges

$$\begin{aligned} \mu, \nu &= 1, \dots, N \,, \\ k, l &= 0, \dots, N-1 \,, \end{aligned}$$

unless stated otherwise explicitly in the sum.

The functions $U_\mu$ and $A_\mu^{(k)}$ that appear in the metric are defined as polynomials in coordinates $x_\mu$

$$U_\mu = \prod_{\substack{\nu \\ \nu \neq \mu}} \left( x_\nu^2 - x_\mu^2 \right) , \qquad A_\mu^{(k)} = \sum_{\substack{\nu_1, \dots, \nu_k \\ \nu_1 < \dots < \nu_k \\ \nu_i \neq \mu}} x_{\nu_1}^2 \dots x_{\nu_k}^2 \,. \tag{2}$$

Each function $X_\mu = X_\mu(x_\mu)$ is dependent on a single coordinate $x_\mu$ and if it is left unspecified, the metric is then referred to as *off-shell*. On the other hand, if the function $X_\mu$ is also defined as a polynomial, namely

$$X_\mu = \lambda \mathcal{J}\big(x_\mu^2\big) - 2b_\mu x_\mu \,, \tag{3}$$

where $\mathcal{J}\big(x_\mu^2\big)$ reads

$$\mathcal{J}\big(x_\mu^2\big) = \prod_{\nu} \left( a_\nu^2 - x_\mu^2 \right) , \tag{4}$$

then the metric (1) satisfies the Einstein equations in vacuum and we refer to it as *on-shell*.

The coordinates we used are divided into two sets. In the black hole case, which is discussed in detail in Sec. 2.2, the coordinates $x_\mu$ represent radius and latitudinal angles whereas the coordinates $\psi_k$ represent time and longitudinal angles. Moreover, since the metric functions are independent of $\psi_k$, they are also the *Killing coordinates*.

The on-shell metric is described by several parameters, but only $\lambda$ can be interpreted easily — it is related to the cosmological constant $\Lambda$ as

$$\Lambda = (2N-1)(N-1)\lambda \,. \tag{5}$$

Regarding the others, in general we can say that the parameters $a_\mu$ somehow describe rotations and the parameters $b_\mu$ encode mass and NUT charges. However,

their interpretation strongly depends on specific other choices that can be made. This will be demonstrated in the next subsection where a black hole geometry that can be obtained from the general metric is discussed.

Instead of $\psi_k$, we can also use a more convenient set of angular coordinates $\phi_\alpha$ defined as

$$\phi_\alpha = \lambda a_\alpha \sum_k \mathcal{A}^{(k)}_\alpha \psi_k \,, \qquad \psi_k = \sum_\alpha \frac{\left(-a_\alpha^2\right)^{N-1-k}}{\mathcal{U}_\alpha} \frac{\phi_\alpha}{\lambda a_\alpha} \,, \tag{6}$$

where the functions $\mathcal{U}_\mu$ and $\mathcal{A}^{(k)}_\mu$ are defined similarly to $U_\mu$ and $A^{(k)}_\mu$ in (2), only $x_\mu$ is replaced with $a_\mu$

$$\mathcal{U}_\mu = \prod_{\substack{\nu \\ \nu\neq\mu}} \left(a_\nu^2 - a_\mu^2\right) \,, \qquad \mathcal{A}^{(k)}_\mu = \sum_{\substack{\nu_1,\dots,\nu_k \\ \nu_1<\dots<\nu_k \\ \nu_i\neq\mu}} a^2_{\nu_1} \dots a^2_{\nu_k} \,. \tag{7}$$

The metric (1) then obtains the form

$$\boldsymbol{g} = \sum_\mu \left[ \frac{U_\mu}{X_\mu} \mathbf{d}x_\mu^2 + \frac{X_\mu}{U_\mu} \left( \sum_\alpha \frac{J_\mu\left(a_\alpha^2\right)}{\mathcal{U}_\alpha} \frac{1}{\lambda a_\alpha} \mathbf{d}\phi_\alpha \right)^2 \right] \,, \tag{8}$$

where $J_\mu\left(a_\alpha^2\right)$ reads

$$J_\mu\left(a_\alpha^2\right) = \prod_{\substack{\nu \\ \nu\neq\mu}} \left(x_\nu^2 - a_\alpha^2\right) \,. \tag{9}$$

Since $\phi_\alpha$ are simply linear combinations of $\psi_k$, they are the Killing coordinates as well.

### 2.2. *Black hole*

The metric (8) can have both Euclidean and Lorentzian signatures, depending on the coordinate ranges and values of the parameters. Detailed discussion of geometries with the Euclidean signature, such as maximally symmetric spaces and Euclidean instantons, can be found in Ref. 3.

In order to obtain the Lorentzian signature, some of the coordinates and parameters need to be Wick-rotated, namely

$$x_N = ir \,, \qquad \phi_N = \lambda a_N t \,, \qquad b_N = iM \,. \tag{10}$$

Moreover, we use a one-parametric gauge freedom in rescaling parameters and set

$$a_N^2 = -\frac{1}{\lambda} \,. \tag{11}$$

We also assume that $a_{\bar\mu}$ are ordered

$$0 < a_1 < \ldots < a_{\bar N - 1} < a_{\bar N} \,, \tag{12}$$

where $\bar{N} = N - 1$ and barred indices go over the ranges

$$\bar{\mu}, \bar{\nu} = 1, \dots, \bar{N}\,,$$
$$\bar{k}, \bar{l} = 0, \dots, \bar{N} - 1\,.$$

This notation is used to separate the temporal and radial coordinates from the angular ones.

The metric (8) then becomes

$$\begin{aligned}
\boldsymbol{g} = &- \frac{\Delta_r}{\Sigma}\left(\prod_{\bar{\nu}}\frac{1+\lambda x_{\bar{\nu}}^2}{1+\lambda a_{\bar{\nu}}^2}\mathbf{d}t - \sum_{\bar{\nu}}\frac{\bar{J}\big(a_{\bar{\nu}}^2\big)}{a_{\bar{\nu}}\,(1+\lambda a_{\bar{\nu}}^2)\,\bar{\mathcal{U}}_{\bar{\nu}}}\mathbf{d}\phi_{\bar{\nu}}\right)^2 \\
&+ \frac{\Sigma}{\Delta_r}\mathbf{d}r^2 + \sum_{\bar{\mu}}\frac{\big(r^2+x_{\bar{\mu}}^2\big)\,\bar{U}_{\bar{\mu}}}{\Delta_{\bar{\mu}}}\mathbf{d}x_{\bar{\mu}}^2 \\
&+ \sum_{\bar{\mu}}\frac{\Delta_{\bar{\mu}}}{\big(r^2+x_{\bar{\mu}}^2\big)\,\bar{U}_{\bar{\mu}}}\left[\frac{1-\lambda r^2}{1+\lambda x_{\bar{\mu}}^2}\prod_{\bar{\nu}}\frac{1+\lambda x_{\bar{\nu}}^2}{1+\lambda a_{\bar{\nu}}^2}\mathbf{d}t + \sum_{\bar{\nu}}\frac{\big(r^2+a_{\bar{\nu}}^2\big)\,\bar{J}_{\bar{\mu}}\big(a_{\bar{\nu}}^2\big)}{a_{\bar{\nu}}\,(1+\lambda a_{\bar{\nu}}^2)\,\bar{\mathcal{U}}_{\bar{\nu}}}\mathbf{d}\phi_{\bar{\nu}}\right]^2,
\end{aligned} \tag{13}$$

with the metric functions $\Delta_r$, $\Delta_{\bar{\mu}}$, $\Sigma$ and $\bar{J}\big(a_{\bar{\mu}}^2\big)$ defined as

$$\begin{aligned}
&\Delta_r = -X_N = \big(1-\lambda r^2\big)\prod_{\bar{\nu}}\big(r^2+a_{\bar{\nu}}^2\big) - 2Mr\,, \qquad && \Sigma = U_N = \prod_{\bar{\nu}}\big(r^2+x_{\bar{\nu}}^2\big)\,, \\
&\Delta_{\bar{\mu}} = -X_{\bar{\mu}} = \big(1+\lambda x_{\bar{\mu}}^2\big)\,\bar{\mathcal{J}}\big(x_{\bar{\mu}}^2\big) + 2b_{\bar{\mu}}x_{\bar{\mu}}\,, && \bar{J}\big(a_{\bar{\mu}}^2\big) = \prod_{\bar{\nu}}\big(x_{\bar{\nu}}^2 - a_{\bar{\mu}}^2\big)
\end{aligned} \tag{14}$$

and the functions $\bar{U}_{\bar{\mu}}$, $\bar{\mathcal{U}}_{\bar{\mu}}$, $\bar{J}_{\bar{\mu}}\big(a_{\bar{\nu}}^2\big)$ and $\bar{\mathcal{J}}\big(x_{\bar{\mu}}^2\big)$ defined in the same way as their unbarred counterparts in Sec. 2.1, only with modified sets of coordinates and parameters (i.e. without $x_N$ and $a_N$).

Let us now discuss suitable ranges of the coordinates $x_{\bar{\mu}}$. For vanishing NUT parameters (and non-zero mass), i.e. $b_{\bar{\mu}} = 0$, the metric is non-singular and has the desired signature if we assume that the coordinates are ordered as

$$-a_1 < x_1 < a_1 < x_2 < a_2 < \ldots < x_{\bar{N}} < a_{\bar{N}}\,. \tag{15}$$

For non-vanishing NUT charges, the coordinate ranges change only slightly — we assume that they are restricted by

$$a_{\bar{\mu}-1} < {}^{-}x_{\bar{\mu}} < x_{\bar{\mu}} < {}^{+}x_{\bar{\mu}} < a_{\bar{\mu}} \tag{16}$$

with the only exception being ${}^{-}x_1$ since ${}^{-}x_1 < -a_1$. Figure 1 shows the discussed ranges.

The horizon structure is given by the roots of the metric function $\Delta_r$ defined in Eq. (14). For $\lambda \leq 0$, there exist at most two horizons — an inner and an outer one. The two horizons can also coincide, thus forming a single extremal horizon, or we can obtain a naked singularity in case there are no horizons. For $\lambda > 0$, there exists one additional cosmological horizon.

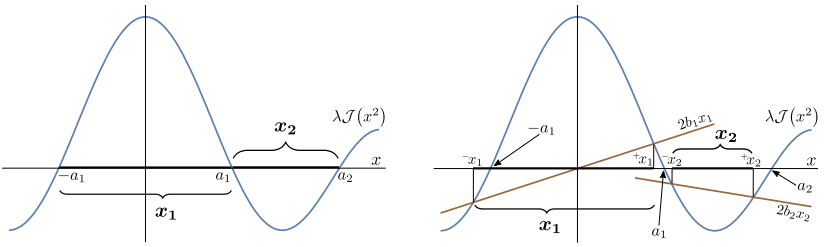



Fig. 1. Ranges of the coordinates $x_{\bar{\mu}}$ for (a) zero and (b) non-zero NUT parameters $b_{\bar{\mu}}$.

As for the interpretation of the metric parameters, for zero NUT charges $b_{\bar{\mu}} = 0$, $a_{\bar{\mu}}$ can be identified with the rotational parameters of the black hole. On the other hand, when the NUT parameters are non-trivial, then both $a_{\bar{\mu}}$ and $b_{\bar{\mu}}$ deform the geometry, however, their exact role remains elusive. This is where studying various limit cases might prove useful as they can help clarify this issue.

## 3. Equal-spin limit in $D = 6$

In general, performing the limiting procedure is not trivial — since some of the coordinate ranges and metric parameters degenerate, it is necessary to rescale them using an appropriate parametrization. Moreover, we expect the symmetry of the resulting spacetime to be enhanced, therefore we also need to find new Killing vectors or another proof of such enhanced symmetry.

For simplicity, we assume that the cosmological constant and NUT charges vanish, i.e. $\lambda = 0 = b_{\bar{\mu}}$. We are interested in the black hole form of the metric, which has only two rotational parameters in $D = 6$. Therefore, we will perform the limit $a_2 \to a_1$ using the parametrization

$$\begin{aligned} a_2 &= a_1 + 2\delta a_2 \varepsilon \,, \\ x_2 &= a_1 + (1 + \xi_2)\, \delta a_2 \varepsilon \,, \end{aligned} \tag{17}$$

where $\varepsilon \ll 1$ is a small parameter and $\delta a_2 > 0$ controls how fast $a_2$ approaches $a_1$. Instead of $x_2 \in (a_1, a_2)$ whose range becomes degenerate, we have introduced a rescaled coordinate $\xi_2 \in (-1, 1)$.

Using a more convenient set of angular coordinates $\phi_\pm = \phi_1 \pm \phi_2$, the resulting metric is

$$\begin{aligned} \boldsymbol{g} \approx -\mathbf{d}t^2 + \frac{2Mr}{\Sigma}\left[\mathbf{d}t - \frac{x_1^2 - a_1^2}{2a_1}\left(\mathbf{d}\phi_+ + \xi_2\, \mathbf{d}\phi_-\right)\right]^2 + \frac{\Sigma}{\Delta_r}\mathbf{d}r^2 - \frac{r^2 + x_1^2}{x_1^2 - a_1^2}\mathbf{d}x_1^2 \\ - \frac{\left(r^2 + a_1^2\right)\left(x_1^2 - a_1^2\right)}{4a_1^2}\left(\frac{\mathbf{d}\xi_2^2}{1 - \xi_2^2} + \mathbf{d}\phi_+^2 + \mathbf{d}\phi_-^2 + 2\xi_2\, \mathbf{d}\phi_+ \mathbf{d}\phi_-\right) \end{aligned} \tag{18}$$

with the metric functions in the form

$$\begin{aligned} \Delta_r &\approx \left(r^2 + a_1^2\right)^2 - 2Mr\,, \\ \Sigma &\approx \left(r^2 + a_1^2\right)\left(r^2 + x_1^2\right)\,. \end{aligned} \tag{19}$$

Most importantly, in addition to the original Killing vectors $\boldsymbol{l}_+$ and $\boldsymbol{l}_-$, we have obtained two new ones after the limit — here denoted by $\boldsymbol{m}$ and $\boldsymbol{n}$. These new vectors are independent of the original Killing vectors and they Lie-preserve the full spacetime metric (18). The complete set of Killing vectors regarding the spatial part of the metric is therefore

$$\begin{aligned} \boldsymbol{l}_+ &= \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\phi_+}\,, \qquad & \boldsymbol{m} &= \cos\phi_- \sqrt{1-\xi_2^2}\,\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\xi_2} - \frac{\sin\phi_-}{\sqrt{1-\xi_2^2}}\left(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\phi_+} - \xi_2\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\phi_-}\right)\,, \\ \boldsymbol{l}_- &= \frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\phi_-}\,, \qquad & \boldsymbol{n} &= -\sin\phi_- \sqrt{1-\xi_2^2}\,\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\xi_2} - \frac{\cos\phi_-}{\sqrt{1-\xi_2^2}}\left(\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\phi_+} - \xi_2\frac{\boldsymbol{\partial}}{\boldsymbol{\partial}\phi_-}\right)\,. \end{aligned} \tag{20}$$

Their Lie brackets suggest that the symmetry of the metric (18) is indeed further enhanced as three of the Killing vectors above generate the algebra of an $SO(3)$ group

$$[\boldsymbol{l}_-, \boldsymbol{m}] = \boldsymbol{n}\,, \qquad [\boldsymbol{m}, \boldsymbol{n}] = \boldsymbol{l}_-\,, \qquad [\boldsymbol{n}, \boldsymbol{l}_-] = \boldsymbol{m}\,. \tag{21}$$

The vector $\boldsymbol{l}_+$ commutes with all the other vectors. Therefore, the symmetry group in the spatial part of the spacetime decouples from the original $U(1) \times U(1)$ into $SO(3) \times U(1)$ after the limit.

## Acknowledgments

This research was supported by Charles University, project GAUK No. 514218, and by the Czech Science Foundation, project GAČR No. 19-01850S.

## References

1. J.F. Plebański and M. Demiański, *Ann. Phys. (N. Y).* **98**, 98 (1976).
2. W. Chen, H. Lü and C.N. Pope, *Class. Quantum Gravity* **23**, 5323 (2006).
3. V.P. Frolov, P. Krtouš and D. Kubizňák, *Living Rev Relativ* **20**, 6 (2017).
4. D. Kubizňák and V.P. Frolov, *Class. Quantum Gravity* **24**, F1 (2007).
5. P. Krtouš, D. Kubizňák, D.N. Page and V.P. Frolov, *J. High Energy Phys.* **02**, 004 (2007).
6. V.P. Frolov, P. Krtouš and D. Kubizňák, *J. High Energy Phys.* **02**, 005 (2007).
7. T. Oota and Y. Yasui, *Phys. Lett. B* **659**, 688 (2008).
8. D.N. Page, D. Kubizňák, M. Vasudevan and P. Krtouš, *Phys. Rev. Lett.* **98**, 061102 (2007).
9. P. Krtouš, D. Kubizňák, V.P. Frolov and I. Kolář, *Class. Quantum Gravity* **33**, 115016 (2016).
10. I. Kolář and P. Krtouš, *Phys. Rev. D* **95**, 1 (2017).

# Static spherically symmetric black holes in quadratic gravity

Podolský, Jiří° and Švarc, Robert*

*Institute of Theoretical Physics, Charles University, Prague,*
*Faculty of Mathematics and Physics, V Holešovičkách 2,*
*180 00 Prague 8, Czech Republic*
*° E-mail: podolsky@mbox.troja.mff.cuni.cz*
** E-mail: robert.svarc@mff.cuni.cz*

Pravda, Vojtěch° and Pravdová, Alena*

*Institute of Mathematics of the Czech Academy of Sciences,*
*Žitná 25, 115 67 Prague 1, Czech Republic*
*° E-mail: pravda@math.cas.cz*
** E-mail: pravdova@math.cas.cz*

We study static, spherically symmetric spacetimes in quadratic gravity. We show that using a conformal-to-Kundt ansatz leads to a considerable simplification of the vacuum field equations. This simplification allows us to find Schwarzschild-Bach black hole in the form of a power series expansion with coefficients given by a recurrent formula. The Schwarzschild-Bach solution is specified by two parameters, the horizon position and an additional "non-Schwarzschild parameter" $b$ that encodes the value of the Bach tensor on the horizon. For vanishing "Bach" parameter, the Bach tensor vanishes as well and the Schwarzschild metric is recovered. The new form of the metric enables us to investigate the geometrical and physical properties of these black holes, such as tidal effects on test particles and thermodynamical quantities.

*Keywords*: Quadratic gravity; static spherically symmetric solutions; black hole; Bach tensor.

## 1. Introduction

In quadratic gravity, terms quadratic in curvature are added to the usual Einstein-Hilbert action. In the case of four dimensions and a vanishing cosmological constant, the action can be expressed as

$$S = \int \mathrm{d}^4x\,\sqrt{-g}\,\Big(\gamma\,R + \beta R^2 - \alpha\,C_{abcd}\,C^{abcd}\Big), \tag{1}$$

where $\gamma = 1/G$ ($G$ is the Newtonian constant) and $\alpha, \beta$ are additional parameters.

Recently, new static spherically symmetric solutions, representing in particular a non-Schwarzschild black hole, to quadratic gravity have been found in Ref. 1 and further analyzed in Refs. 2, 3, 4, mainly by using numerical methods.

Instead of using the standard spherically symmetric coordinates, we study [5,6] static spherically symmetric solutions in the conformal-to-Kundt form [7] which leads to a considerable simplification of the vacuum field equations allowing us to construct a solution describing the Schwarzschild-Bach black hole in the form of a power series expansion. The coefficients in the expansion are determined by a recurrent formula. We also study physical properties of the Schwarzschild-Bach black hole such as tidal effects on test particles and thermodynamical quantities. The

Schwarzschild-Bach black hole can be also generalized to include a non-vanishing cosmological constant, see Ref. 8.

## 2. The Field Equations

We study the case $R = 0$ (in the case of Einstein–Weyl gravity corresponding to $\beta = 0$, this directly follows from the trace of the field equations). Then the field equations reduce to

$$R_{ab} = 4k\, B_{ab}\,, \tag{2}$$

where $k \equiv \alpha\, G$ and $B_{ab}$ is the Bach tensor

$$B_{ab} \equiv \big(\nabla^c\nabla^d + \tfrac{1}{2}R^{cd}\big)C_{acbd}\,, \tag{3}$$

which is traceless, symmetric and conserved ($g^{ab}B_{ab} = 0$, $B_{ab} = B_{ba}$, $B_{ab}{}^{;b} = 0$).

Instead of using the standard static spherically symmetric metric

$$\mathrm{d}s^2 = -h(\bar{r})\,\mathrm{d}t^2 + \frac{\mathrm{d}\bar{r}^2}{f(\bar{r})} + \bar{r}^2(\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2)\,, \tag{4}$$

we employ a conformal-to-type D-direct-product-Kundt form [7] of a static spherically symmetric spacetime

$$\mathrm{d}s^2 = \Omega^2(r)\mathrm{d}s^2_{\mathrm{Kundt}} = \Omega^2(r)\big[\,\mathrm{d}\theta^2 + \sin^2\theta\,\mathrm{d}\phi^2 - 2\,\mathrm{d}u\,\mathrm{d}r + \mathcal{H}(r)\,\mathrm{d}u^2\,\big], \tag{5}$$

where

$$h(\bar{r}) = -\Omega^2\,\mathcal{H}\,, \qquad f(\bar{r}) = -(\Omega'/\Omega)^2\,\mathcal{H}\,, \tag{6}$$

$\Omega'$ denotes the derivative of $\Omega$ with respect to $r$. The metrics (5) and (4) are related via the transformation

$$\bar{r} = \Omega(r)\,, \qquad t = u - \int \mathcal{H}(r)^{-1}\mathrm{d}r\,. \tag{7}$$

We use the time scaling $t \to \sigma^{-1}\,t$ (implying $h \to \sigma^2\,h$) to set $h = 1$ at spatial infinity for asymptotically flat solutions.

The Killing horizon associated with the Killing vector $\partial_t = \partial_u$ is located at $r_h$:

$$\mathcal{H}\big|_{r=r_h} = 0\,, \tag{8}$$

which implies that also $h(\bar{r}_h) = 0 = f(\bar{r}_h)$ (see Eq. (6)).

The main reason behind the considerable simplification of the field equations in the Kundt coordinates is that higher-order corrections to the Einstein theory are in quadratic gravity represented by the Bach tensor (see (2)) which is conformally well behaved $B_{ab} = \Omega^{-2}\,B^{\mathrm{Kundt}}_{ab}$.

It can be shown that in Kundt coordinates, the field equations (2) for the metric functions $\Omega(r)$ and $\mathcal{H}(r)$ reduce to [5,6]

$$\Omega\Omega'' - 2\Omega'^2 = \tfrac{1}{3}k\,\mathcal{B}_1\mathcal{H}^{-1}\,, \tag{9}$$

$$\Omega\Omega'\mathcal{H}' + 3\Omega'^2\mathcal{H} + \Omega^2 = \tfrac{1}{3}k\,\mathcal{B}_2\,, \tag{10}$$

where 2 independent components of the Bach tensor are

$$\mathcal{B}_1 \equiv \mathcal{H}\mathcal{H}'''', \qquad \mathcal{B}_2 \equiv \mathcal{H}'\mathcal{H}''' - \tfrac{1}{2}\mathcal{H}''^2 + 2\,. \tag{11}$$

$\mathcal{B}_1$ always vanishes on the horizon (8).

The Bach and Weyl scalar curvature invariants read

$$B_{ab}\,B^{ab} = \tfrac{1}{72}\,\Omega^{-8}\left[(\mathcal{B}_1)^2 + 2(\mathcal{B}_1+\mathcal{B}_2)^2\right], \tag{12}$$

$$C_{abcd}\,C^{abcd} = \tfrac{1}{3}\,\Omega^{-4}\,(\mathcal{H}''+2)^2\,. \tag{13}$$

Thus, $C_{abcd}\,C^{abcd}=0$ implies $B_{ab}=0$. Moreover, $B_{ab}=0$ if and only if $B_{ab}\,B^{ab}=0$. The invariant (12) implies that there are two geometrically distinct classes of solutions to (9), (10), depending on the Bach tensor, one with $B_{ab}=0$ and another one with $B_{ab}\neq 0$.

In the case of vanishing Bach tensor $\mathcal{B}_1=0=\mathcal{B}_2$, the only solution of Eqs. (9)–(10) is the Schwarzschild metric in accordance with the Birkhoff theorem

$$\Omega(r) = -\frac{1}{r}\,, \qquad \mathcal{H}(r) = -r^2 - 2m\,r^3\,, \tag{14}$$

where a coordinate freedom $r\to\lambda\,r+\nu,\ u\to\lambda^{-1}u$ of the metric (5) was used. Eqs. (7), (6) give $r=-1/\bar{r}$, $f(\bar{r})=1-2m/\bar{r}=h(\bar{r})$, respectively, where $\bar{r}>0$ corresponds to $r<0$ ($r$ increases with $\bar{r}$).

For non-vanishing Bach tensor, Eqs. (9), (10) form an autonomous system (they do not explicitly depend on the variable $r$) and thus its solutions can be found as expansions in the powers of $r$ to any order around any fixed value $r_0$,

$$\Omega(r) = \Delta^n \sum_{i=0}^{\infty} a_i\,\Delta^i\,, \qquad \mathcal{H}(r) = \Delta^p \sum_{i=0}^{\infty} c_i\,\Delta^i\,, \tag{15}$$

where $\Delta\equiv r-r_0,\ n,p\in\mathbb{R}$, and $a_0,c_0\neq 0$. Eqs. (9) and (10) restrict possible dominant powers $n,p$ in (15), see Refs. 5, 6. It turns out that there are only four classes of solutions allowed, namely $[n,p]=[-1,2],\ [0,1],\ [0,0],\ [1,0]$. The class $[-1,2]$ corresponds to the Schwarzschild black hole (14), the class $[1,0]$ is equivalent to the $(s,t)=(2,2)$ class of Refs. 9 and 2 corresponding to a naked singularity, while the Schwarzschild–Bach black hole is contained in the classes $[0,1]$ and $[0,0]$, see Refs. 5, 6.

## 3. Class [0, 1]: Schwa–Bach Black Hole Expressed Around the Horizon $r_h$

Since in the $[0,1]$ class, $r_0$ is the root of $\mathcal{H}$ it corresponds to the horizon $r_h$, see Eq. (8). This class of solutions of quadratic gravity includes non-Schwarzschild black holes with $B_{ab}\neq 0$, see Refs. 5, 6

$$\Omega(r) = -\frac{1}{r} - \frac{b}{r_h}\sum_{i=1}^{\infty}\alpha_i\Big(1-\frac{r}{r_h}\Big)^i, \tag{16}$$

$$\mathcal{H}(r) = (r-r_h)\left[\frac{r^2}{r_h} + 3b\,r_h\sum_{i=1}^{\infty}\gamma_i\Big(\frac{r}{r_h}-1\Big)^i\right], \tag{17}$$

where the initial coefficients are

$$\alpha_0 \equiv 0\,, \quad \alpha_1 = 1\,, \quad \gamma_1 = 1\,, \quad \gamma_2 = \frac{1}{3}\Big(4 - \frac{1}{2kr_h^2} + 3b\Big)\,, \tag{18}$$

and $\alpha_l, \gamma_{l+1}$ for $l \geq 2$ are given by the recurrent relations

$$\alpha_l = \frac{1}{l^2}\Big[\alpha_{l-1}(2l^2 - 2l + 1) - \alpha_{l-2}(l-1)^2 - 3\sum_{i=1}^{l}(-1)^i\,\gamma_i\,(1 + b\,\alpha_{l-i})\big(l(l-i) + \tfrac{1}{6}i(i+1)\big)\Big],$$

$$\gamma_{l+1} = \frac{(-1)^l}{kr_h^2\,(l+2)(l+1)l(l-1)}\sum_{i=0}^{l-1}\big(\alpha_i + \alpha_{l-i}(1 + b\,\alpha_i)\big)(l-i)(l-1-3i)\,. \tag{19}$$

This family of spherically symmetric black holes depends on two parameters:

- The horizon position $r_h$, $r = r_h$ is the root of $\mathcal{H}$ given by (17).
- The Bach parameter $b \neq 0$ for the Schwarzschild–Bach (or Schwa–Bach) black hole with non-zero Bach tensor, while $b = 0$ for the Schwarzschild solution.

On the horizon $r_h$, Eq. (11) gives

$$\mathcal{B}_1(r_h) = 0\,, \quad \mathcal{B}_2(r_h) = -\frac{3}{kr_h^2}\,b\,. \tag{20}$$

The invariants (12), (13) are $B_{ab}\,B^{ab}(r_h) = \frac{r_h^4}{4k^2}\,b^2$ and $C_{abcd}\,C^{abcd}(r_h) = 12\,r_h^4\,(1+b)^2$, respectively.

Fig. 1 shows the rapid convergence in the near-horizon region of the function $h(\bar{r})$ of the metric (4) expressed from $\Omega$ and $\mathcal{H}$ using Eq. (6). The first 20 (red), 40 (orange), 60 (green), 80 (blue), and 100 (violet) terms in the series are compared with the numerical solution of Ref. 1 (black). The horizon is located at $\bar{r}_h = 1$ (i.e., $r_h = -1$), $k = 0.5$ and the value $b = 0.3633018769168$ corresponds to the asymptotically flat case (see Ref. 1). We employed the scaling freedom with $\sigma^2 \approx 2.18$ to obtain $h \to 1$ asymptotically.

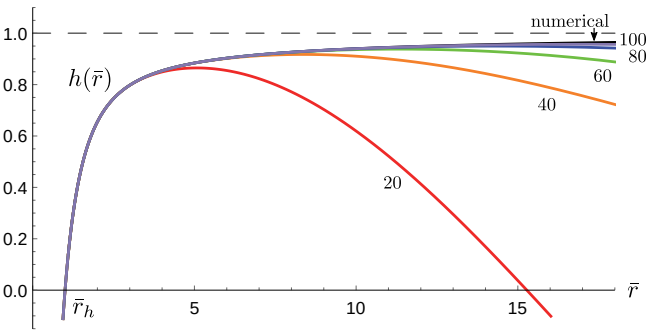



Fig. 1. Rapid convergence of the function function $h(\bar{r})$.

## 4. Class [0, 0]: Schwa–Bach Black Hole Expressed Around Any Point $r_0 \neq r_h$

The solution to Eqs. (9), (10) of the form (15) with $n = 0 = p$ is given by the power series with

$$a_{l+1} = \frac{-1}{l(l+1)\,c_0}\left[\tfrac{1}{3}\,a_{l-1} + \sum_{i=1}^{l+1} c_i\,a_{l+1-i}\left(l(l+1-i) + \tfrac{1}{6}i(i-1)\right)\right],$$

$$c_{l+3} = \frac{3}{k\,(l+3)(l+2)(l+1)l}\sum_{i=0}^{l} a_i\,a_{l+1-i}(l+1-i)(l-3i)\,, \tag{21}$$

for any $l \geq 1$, where $a_0, a_1, c_0, c_1, c_2$ are five free parameters and

$$c_3 = \frac{1}{6kc_1}\left[3a_0(a_0 + a_1c_1) + 9a_1^2c_0 + 2k(c_2^2 - 1)\right],$$

see Refs. 5, 6.

To identify the Schwa–Bach black hole (16), (17), previously expressed around the horizon $r_h$ in the class $[0, 1]$, we have to uniquely determine the five free parameters by evaluating the functions (16), (17) and their derivatives at $r = r_0$. For $B_{ab} = 0$, both classes $[0, 0]$ and $[0, 1]$ reduce to the Schwarzschild black hole. However, the $[0, 0]$ class is a larger family of solutions than the Schwa-Bach black hole since it admits one more parameter.

Moreover, the power series (15) with integer exponents transform in some cases to series with non-integer exponents in the coordinate $\bar{r}$. For example, a new class $(w, t) = (4/3, 0)$ in the notation of Ref. 2 also belongs to our $[0, 0]$ class, see Refs. 5, 6.

## 5. Observable Effects Caused By The Schwa–Bach Black Hole

The Bach tensor, namely two independent parts of the Bach invariant (12), $\mathcal{B}_1, \mathcal{B}_2$, could produce measurable affects on test particles described by the geodetic deviation equation[10]. Apart from the classical Newtonian tidal deformation caused by the Weyl curvature, there are two additional effects due to the Bach tensor, one in the transverse components of the acceleration, the second in the radial component along $\partial_{\bar{r}}$, with amplitudes given by $\mathcal{B}_1, \mathcal{B}_2$ defined in (11). On the horizon, there is only the radial effect caused by $\mathcal{B}_2$ since $\mathcal{B}_1(r_h) = 0$, see (20).

## 6. Thermodynamical Properties: Horizon Area, Temperature, Entropy

The horizon of the Schwarzschild–Bach black hole is generated by the (rescaled) null Killing vector $\ell \equiv \sigma\partial_t = \sigma\partial_u$ and is located at $\mathcal{H} = 0$, i.e., at $r = r_h$, see (8), (17). Its area (using (16)), surface gravity ($\kappa^2 \equiv -\frac{1}{2}\,\ell_{\mu;\nu}\,\ell^{\mu;\nu}$, see Ref. 11), and the black-hole horizon temperature[12] read

$$\mathcal{A} = 4\pi\,\bar{r}_h^2 = 4\pi\,\Omega^2(r_h) = 4\pi\,r_h^{-2}\,, \tag{22}$$

$$\kappa/\sigma = -\tfrac{1}{2}\,\mathcal{H}'(r_h) = -\tfrac{1}{2}\,r_h = \tfrac{1}{2}\,\bar{r}_h^{-1}\,, \tag{23}$$

$$T/\sigma = \tfrac{1}{2\pi}\,\kappa/\sigma = -\tfrac{1}{4\pi}\,r_h = \tfrac{1}{4\pi}\,\bar{r}_h^{-1}\,, \tag{24}$$

respectively. The value of the scaling factor $\sigma$ is fixed by the condition that $h = -\Omega^2\mathcal{H} \to 1$ asymptotically as $\bar{r} = \Omega(r) \to \infty$.

Using the generalized definition of entropy $S = (2\pi/\kappa) \oint \mathbf{Q}$ (see Ref. 13), where the Noether charge 2-form on the horizon is $\mathbf{Q} = -\frac{\Omega^2\,\mathcal{H}'}{16\pi}\left[1+\frac{4}{3}k^2\,\frac{\mathcal{B}_1+\mathcal{B}_2}{\Omega^4}\right]\Big|_{r=r_h} \sin\theta\,\mathrm{d}\theta\wedge\mathrm{d}\phi\,,$ and using Eqs. (20), (22), and (23) the entropy becomes

$$S = \tfrac{1}{4}\mathcal{A}\left(1 - 4k\,r_h^2\,b\right) = \tfrac{1}{4}\mathcal{A}\left(1 - 4k\,\bar{r}_h^{\,-2}\,b\right), \tag{25}$$

which agrees with the results of Ref. 1, with the identification $k = \alpha$, $b = \delta^*$. For the Schwarzschild black hole ($b = 0$) or in Einstein's theory ($k = 0$), we recover the standard expression. For smaller Schwa–Bach black holes (smaller $\bar{r}_h$), the deviations from $S = \frac{1}{4}\mathcal{A}$ are larger, analogously to Ref. 14. $S > 0$ implies $4kb < \bar{r}_h^2$.

Combining Eqs. (22), (24), and (25), gives a relation between the temperature and the entropy

$$T = \tfrac{1}{4}\sigma\left(\pi\,S + 4\pi^2 k\,b\right)^{-1/2}, \tag{26}$$

generalizing $T = \frac{1}{4}(\pi S)^{-1/2}$ for the Schwarzschild black hole.

## Acknowledgments

We acknowledge the support by the Czech Science Foundation Grant No. GAČR 17-01625S (JP, RŠ), GAČR 19-09659S (VP, AP) and the Research Plan RVO: 67985840 (VP, AP).

## References

1. H. Lü, A. Perkins, C. N. Pope, and K. S. Stelle, *Phys. Rev. Lett.* **114**, 171601 (2015).
2. H. Lü, A. Perkins, C. N. Pope, and K. S. Stelle, *Phys. Rev. D* **92**, 124019 (2015).
3. B. Holdom and J. Ren, *Phys. Rev. D* **95**, 084034 (2017).
4. K. D. Kokkotas, R. A. Konoplya, and A. Zhidenko *Phys. Rev. D* **96**, 064007 (2017).
5. J. Podolský, R. Švarc, V. Pravda, and A. Pravdová, *Phys. Rev. D* **98**, 021502(R) (2018).
6. J. Podolský, R. Švarc, V. Pravda, and A. Pravdová, (in preparation) (2019).
7. V. Pravda, A. Pravdová, J. Podolský, and R. Švarc, *Phys. Rev. D* **95**, 084025 (2017).
8. R. Švarc, J. Podolský, V. Pravda, and A. Pravdová, *Phys. Rev. Lett.* **121**, 231104 (2018).
9. K. S. Stelle, *Gen. Relativ. Gravit.* **9**, 353 (1978).
10. J. Podolský and R. Švarc, *Phys. Rev. D* **85**, 044057 (2012).
11. R. M. Wald, *General Relativity* (Chicago: University of Chicago Press, 1984).
12. Zhong-Ying Fan, H. Lü, *Phys. Rev. D* **91**, 064009 (2015).
13. R. M. Wald, *Phys. Rev. D* **48**, R3427(R) (1993).
14. M. Nozawa and H. Maeda, *Class. Quantum Grav.* **25**, 055009 (2008).

# Binary population synthesis calculations of Type Ia supernovae

Zheng-Wei Liu[1,2]

[1] *Yunnan Observatories, Key Laboratory for the Structure and Evolution of Celestial Objects, Chinese Academy of Sciences (CAS), Kunming 650216, China*
[2] *Center for Astronomical Mega-Science, CAS, Beijing, China*
[*] *E-mail: zwliu@ynao.ac.cn*

The nature of the progenitors of Type Ia supernovae (SNe Ia) remains a mystery. By performing binary population synthesis calculations, theoretical predictions for different SN Ia progenitor models, such as supernova rates, delay times, pre-explosion companion properties and early ultraviolet signatures from ejecta-companion interaction have been presented. The results are then compared with the observations to place constraints on the possible progenitor models of SNe Ia.

*Keywords*: Binaries: close – supernovae: general.

## 1. Introduction

There is consensus that Type Ia supernovae (SNe Ia) arise from thermonuclear explosions of white dwarfs (WDs) in binary systems[1]. However, the nature of their progenitors and the physics of explosion are still unknown[2]. Depending on the nature of the companion star, progenitor scenarios of SNe Ia fall into three general categories: (i) the single-degenerate (SD) scenario[3], in which a carbon-oxygen (CO) WD accretes matter from a non-degenerate companion, potentially a main-sequence (MS), subgiant, red giant (RG), or even helium (He) star, to trigger an explosion when its mass reaches the Chandrasekhar-mass (Ch-mass) limit, (ii) the double-degenerate (DD) scenario[4], in which explosions are caused by the merger of two WDs, (iii) the WD accretes material from an He-burning star or an He WD to lead to a sub-Chandrasekhar-mass (sub-Ch) explosion when the He-shell accumulation reaches a critical value of about $0.02 - 0.2\,M_{\odot}$ [5,6].

Comparing the expected rates and delay times of SNe from binary population synthesis (BPS) calculations with those inferred from the observations can place constraints on progenitor systems of SNe Ia[7]. It is expected that the companion stars in the SD and sub_Ch scenario will be detectable in pre-explosion observations[8]. Also, in the SD scenario, the collision of the SN Ia ejecta with its companion star is expected to produce detectable ultraviolet (UV) emission in the first few days after the SN explosion[9]. To constrain the possible progenitor models of SNe Ia by their pre-explosion and early UV observations, it is required to preform BPS calculations to provide theoretical predictions for the pre-explosion companion properties and early UV signatures in the SD scenario.

## 2. Method and Results

In this work, the Cambridge stellar evolution code STARS[11] is used to trace the detailed binary evolution of a set of systems consisting of a WD and a MS or SG,

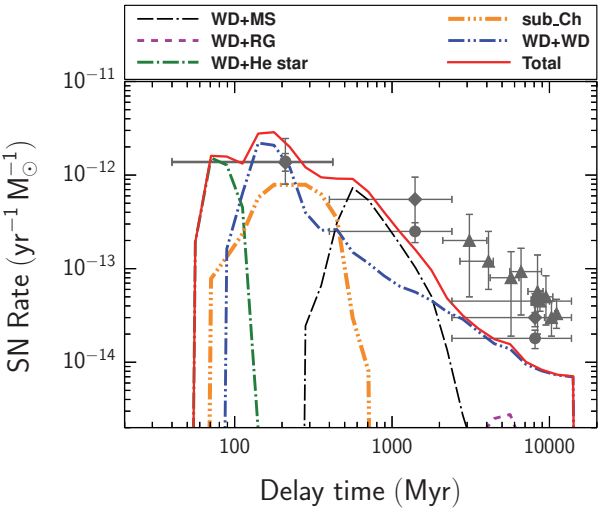



Fig. 1. The delay-time distribution (DTD) of different progenitor scenarios in our calculations. The results are compared to the observed DTD (gray dots[10]).

a RG, or a He star companion. The WDs are treated as point masses. The mass transfer occurs though Roche-lobe overflow (RLOF). In the SD scenario, once the WD accretes the transferred companion material to increase its mass to near the Ch-mass limit ($\sim 1.4\,M_\odot$), an SN Ia explosion is assumed to occur. In the sub_Ch scenario (i.e., for the low mass-transfer-rate He-accreting case), a thick layer of helium is believed to grow on the surface of the WD. As a result, once a He-shell accumulation reaches a critical value ($\sim 0.05\,M_\odot$), the sub-Ch mass thermonuclear explosion is assumed to trigger.

With a series of binary evolution calculations for various close WD binary systems, we determine the initial parameter space leading to SNe Ia in the orbital period–secondary mass (i.e. $\log_{10} P^i$–$M_2^i$ ) plane for various initial CO WD masses ($M_{\rm WD}^i$) in the MS, RG, and He star donor channel. Furthermore, a rapid population synthesis code[13] is used to perform a detailed Monte Carlo simulation. When a binary system evolves to a CO WD + MS (SG, RG, or He star) system, assuming that it will lead to an SN Ia explosion if it is located in the plane of ($\log_{10} P^i$, $M_2^i$) for its $M_{\rm WD}^i$ based on the detailed binary evolution. For the DD scenario in this BPS calculation, SNe Ia are assumed to arise from the merging of two close CO WDs that have a combined mass larger than or equal to the Ch-mass (Webbink 1984). We refer to some previous works for a detailed description about the method and basic assumptions used in this work[7,12,14].

### 2.1. *Rates and delay times*

Figure 1 shows a comparison between the delay-time distribution (DTD[7,14]) of different progenitor scenarios in this calculations and the observed DTD[10]. It is

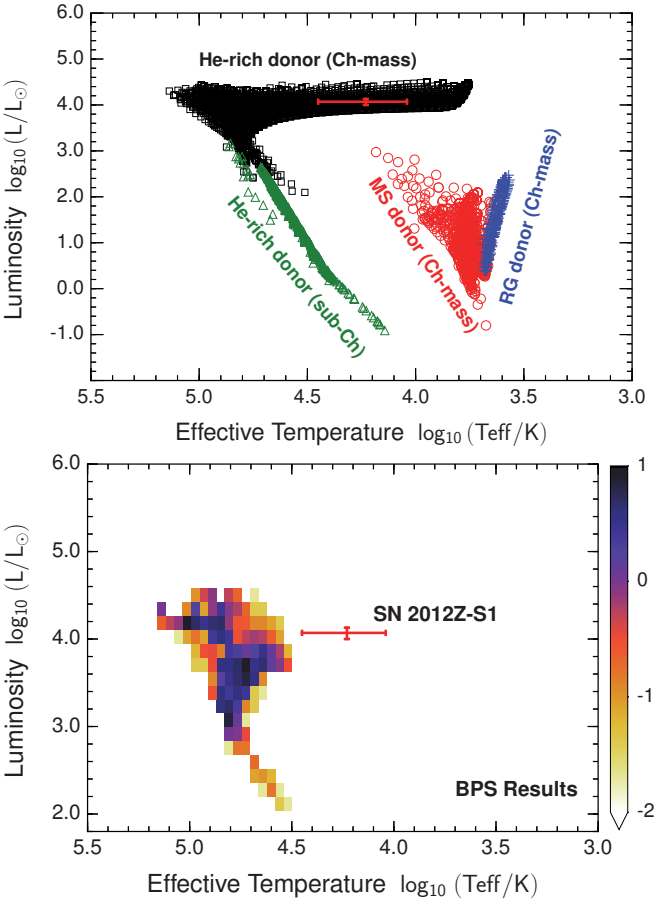



Fig. 2. Top panel: H-R diagram of pre-explosion companions in different SN Ia progenitor scenarios from binary evolution calculations[12]. Bottom panel: the distributions of companions in the the plane of ($\log_{10} T_{\rm eff}$, $\log_{10} L$), which are obtained from BPS calculations assuming a constant star-formation rate of $3.5\,{\rm M}_{\odot}\,{\rm yr}^{-1}$. Here, only the He donor Ch-mass channel is considered. The pre-explosion observation of the SN 2012Z-S1 is also included (the error in red[8]).

found that the DD scenario has a wide distribution of the delay times. The SD scenario seems to be difficult to produce old SNe Ia with the delay times of $>$ 3000 Myr. Also, it is really hard to explain the observed rates and delay times with a single progenitor model, which may indicate that a combination of current models might be an alternative scenario for SNe Ia. Otherwise, a new progenitor paradigm may be needed if SNe Ia are indeed generated from the same origin.

### 2.2. *Pre-explosion companion stars*

The WD can only be observed directly in our own Milky Way and several very nearby galaxies as the WD would be faint. However, the non-degenerate compan-

ions in SD and sub_Ch scenario will be detectable in pre-explosion observations. Therefore, analyzing pre-explosion images at the SN position provides a direct way to constrain SN progenitor systems. With performing detailed binary evolution calculations, the pre-explosion properties of companions in the SD and sub_Ch scenario are shown in Fig. 2. This will be very helpful for constraining progenitor systems of SNe Ia from their pre-explosion observations. Recently, a luminous source (i.e., SN 2012Z-S1) has been discovered from pre-explosion HST images of an SN Ia SN 2012Z, and it is further suggested to be the probable progenitor system of SN 2012Z[8]. In this work, it is found that the detailed binary evolution calculations for the WD+He star Ch-mass channel can contribute significantly to the number of systems in the vicinity of SN 2012Z-S1 (Fig. 2). However, further BPS calculations suggest that the WD+He star channel produces companions as cool as the SN 2012Z-S1 only at very low probability (see Fig. 2).

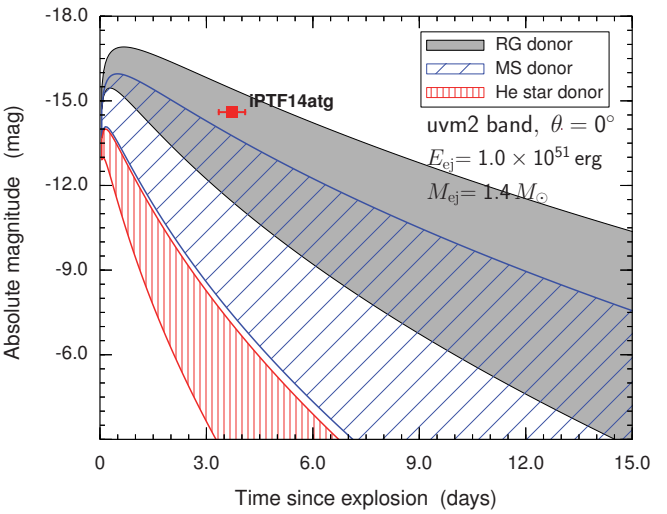



Fig. 3. Distributions of theoretical light curves predicted from the ejecta-companion interaction for different SD progenitor models. Here, only a viewing angle $\theta = 0^\circ$ and $uvm2$ band are considered. The ejecta mass ($M_{\rm ej} = 1.4\,M_\odot$) and explosion energy ($E_{\rm ej} = 10^{51}$ erg) are adopted.

### 2.3. *Early UV emission of ejecta-companion interaction*

In the Kasen's model[9], SN ejecta collides with its companion star, the impacting layers are re-shocked and the kinetic energy partially is dissipated, causing UV flux emission from the shock-heated region. This emission dominates the early-time (few days after SN explosion) light curves within certain viewing angles[9]. By combining our BPS results into the Kasen's analytical model[12], the distributions of early-UV emission from ejecta-companion in the different SD progenitor models are presented in Fig. 3. This theoretical predictions will be helpful for future early-time observations of SNe Ia to constrain their possible progenitors.

A strong UV pulse in the early light curves of an SN Ia iPTF14atg has been detected recently[15]. It has been further suggested that this strong UV pulse supports that some SNe Ia arise from the SD Ch-mass scenario. Here, the predictions from our BPS calculations are compared to the early UV flash of iPTF14atg to put a more stringent on the possible progenitor of this event. It is found that it is unlikely that the detected early UV flash of iPTF14atg arises from the interaction with a MS or He star companion by considering that the actual viewing angle may be larger than $\theta = 0^\circ$ and taking the uncertainties on the explosion parameters into account. However, the results from the RG donor scenario can explain the early UV flash of iPTF14atg (Fig. 3). Therefore, it can be concluded that the progenitor system of iPTF14atg would be most likely a red-giant donor binary system if we assume the observed strong UV pulse of iPTF14atg was indeed produced by the SN ejecta-companion interaction.

## 3. Conclusion and Summary

By performing the detailed binary evolution and population synthesis calculations, some helpful theoretical predictions for different progenitor scenario of SNe Ia have been presented. The results are also compared to current observations to place constraints on the possible SN Ia progentiors. The results and conclusions can be summarized as follows.

i) The DD scenario has a very wide DTD and the SD scenario seems to be difficult to produce old SNe Ia ($> 3000$ Myr). It is hard to explain the observed rates and delay times with a single progenitor model. A combination of different progenitor models seems to provide a better explanation.

ii) Pre-explosion companion properties in the SD and sub_Ch scenario have been presented. This will be helpful for constraining progenitor systems of SNe Ia from their pre-explosion observations. It is found that the He companions in SD scenario can explain the observations of SN 2012Z-S1 at a low probability.

iii) Population synthesis models of early UV emission from ejecta-companion interaction in the SD scenario are presented. The results are useful for comparing with early UV observations of SNe Ia to put constraints on their progenitor models. It is also found that the progenitor system of iPTF14atg would be most likely to have a red-giant donor if we assume its observed UV pulse was indeed produced by the ejecta-companion interaction.

## Acknowledgments

The work is supported by the National Natural Science Foundation of China (NSFC, No. 11873016).

## References


1. F. Hoyle and W. A. Fowler, Nucleosynthesis in Supernovae, *ApJ* **132**, p. 565 (November 1960).
2. W. Hillebrandt and J. C. Niemeyer, Type IA Supernova Explosion Models, *ARA&A* **38**, 191 (2000).
3. J. Whelan and I. Iben, Jr., Binaries and Supernovae of Type I, *ApJ* **186**, 1007 (December 1973).
4. I. Iben, Jr. and A. V. Tutukov, The evolution of low-mass close binaries influenced by the radiation of gravitational waves and by a magnetic stellar wind, *ApJ* **284**, 719 (September 1984).
5. E. Livne, Successive detonations in accreting white dwarfs as an alternative mechanism for type I supernovae, *ApJ* **354**, L53 (May 1990).
6. S. E. Woosley and T. A. Weaver, Sub-Chandrasekhar mass models for Type IA supernovae, *ApJ* **423**, 371 (March 1994).
7. Z.-W. Liu and R. J. Stancliffe, Rates and delay times of Type Ia supernovae in the helium-enriched main-sequence donor scenario, *MNRAS* **475**, 5257 (April 2018).
8. C. McCully, S. W. Jha, R. J. Foley, L. Bildsten, W.-F. Fong, R. P. Kirshner, G. H. Marion, A. G. Riess and M. D. Stritzinger, A luminous, blue progenitor system for the type Iax supernova 2012Z, *Nature* **512**, 54 (August 2014).
9. D. Kasen, Seeing the Collision of a Supernova with Its Companion Star, *ApJ* **708**, 1025 (January 2010).
10. D. Maoz and O. Graur, Star Formation, Supernovae, Iron, and $\alpha$: Consistent Cosmic and Galactic Histories, *ApJ* **848**, p. 25 (October 2017).
11. P. P. Eggleton, The evolution of low mass stars, *MNRAS* **151**, 351 (1971).
12. Z.-W. Liu, R. J. Stancliffe, C. Abate and B. Wang, Pre-explosion Companion Stars in Type Iax Supernovae, *ApJ* **808**, p. 138 (August 2015).
13. J. R. Hurley, O. R. Pols and C. A. Tout, Comprehensive analytic formulae for stellar evolution as a function of mass and metallicity, *MNRAS* **315**, 543 (July 2000).
14. Z.-W. Liu, T. J. Moriya, R. J. Stancliffe and B. Wang, Constraints on single-degenerate Chandrasekhar mass progenitors of Type Iax supernovae, *A&A* **574**, p. A12 (February 2015).
15. Y. Cao, S. R. Kulkarni, D. A. Howell, A. Gal-Yam, M. M. Kasliwal, S. Valenti, J. Johansson, R. Amanullah, A. Goobar, J. Sollerman, F. Taddia, A. Horesh, I. Sagiv, S. B. Cenko, P. E. Nugent, I. Arcavi, J. Surace, P. R. Woźniak, D. I. Moody, U. D. Rebbapragada, B. D. Bue and N. Gehrels, A strong ultraviolet pulse from a newborn type Ia supernova, *Nature* **521**, 328 (May 2015).

# Type Ia supernovae triggered by helium detonation

Keiichi Maeda[1,*], Jian Jiang[2], Mamoru Doi[2,3] and Toshikazu Shigeyama[3]

[1] *Department of Astronomy, Kyoto University, Kitashirakawa-Oiwake-cho, Sakyo-ku, Kyoto 606-8502, Japan*

[2] *Institute of Astronomy, Graduate School of Science, The University of Tokyo, 2-21-1 Osawa, Mitaka, Tokyo 181-0015, Japan*

[3] *Research Center for the Early Universe, Graduate School of Science, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan*

[*] *E-mail: keiichi.maeda@kusastro.kyoto-u.ac.jp*

There has recently been an increasing interest in a possible population of type Ia supernovae (SNe Ia) triggered by helium detonation on the surface of a massive white dwarf. In this paper, we first summarize possible observational signatures of the He detonation-triggered SNe Ia, emphasizing the new diagnostics of the He detonation mode potentially seen in the SN light within the first few days since the explosion. We then argue that observational properties of a peculiar SN Ia, MUSSES1604D as discovered by the Hyper Suprime-Cam (HSC) attached with the Subaru telescope, are best explained by the He-detonation scenario. We then discuss possible origins, including the He detonation scenario, of the diversity seen in the photometric properties of SNe Ia in the first few days. While the He detonation could reproduce observational properties of a fraction of SNe Ia showing the excessive emission in the first few days, it is likely that a bulk of them are linked to a different explosion mechanism where the early excess would arise due to an extensive mixing of $^{56}$Ni during the explosion. A combined analysis of the very early phase observations and the maximum-phase observations will be key in mapping the diverse SN Ia zoo into different populations reflecting different progenitors and/or different explosion modes.

*Keywords*: Supernovae; binary stars; nucleosynthesis.

## 1. SNe Ia in the first few days

It has been widely accepted that the type Ia supernova (SN Ia) is a thermonuclear explosion of a massive C+O white dwarf (WD) in a binary system. Details are however yet to be clarified on the nature(s) of the progenitor system(s) and the explosion mode(s) (see Ref. 1 for a review). There are two popular progenitor scenarios; the single degenerate (SD) scenario involves an accreting WD from a non-degenerate companion star,[2,3] while the double degenerate (DD) scenario involves a binary merger of two WDs.[4,5] There is also a variety in the possible mode of the explosion, especially in a way how the explosion is initiated. In the popular delayed-detonation model, the carbon deflagration flame is first triggered near the center of a massive C+O WD close to the Chandrasekhar limiting mass, then the transition to the detonation wave is hypothesized.[6–8]

In the last few years, there has been an increasing interest in a specific mode of the explosion, which involves accretion of helium (He) on the surface of a primary C+O WD.[9–11] If the He accretion is sufficiently rapid, a detonation is initiated within the He layer. The shock wave thus generated would penetrate into the C+O WD, triggering the carbon detonation near the center of the WD. This scenario does not require a Chandrasekhar-mass progenitor; the progenitor WD mass

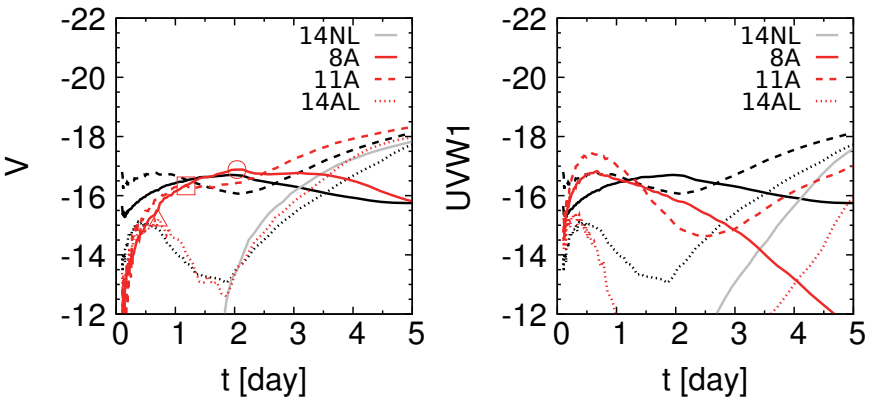



Fig. 1. Examples of the simulated early-phase light curves for the He detonation models, in the $V$-band (left) and in the SWIFT $UVW1$-band (right).[21] Three models with different masses of the primary WD ($0.8, 1.1, 1.4 M_\odot$ for 8A, 11A, 14AL, respectively) are shown by the red curves. The bolometric light curves of the same models are shown by the black curves. The model corresponding to model 14AL but without the He detonation ash (14NL) is shown by the gray curve. Reproduced from the following publication: K. Maeda, et al., The Astrophysical Journal, 861, 78 (2018).[21]

is phenomenologically limited by the capability of synthesizing $^{56}$Ni mass to be consistent with observed SN Ia luminosities.[12,13] To produce $\sim 0.6 M_\odot$ of $^{56}$Ni to account for normal SNe Ia, the WD mass ($M_{\rm WD}$) is $\gtrsim 1 M_\odot$. The primary WD with $M_{\rm WD} \sim 0.8 M_\odot$ will lead to a low luminosity SN, thus at most applicable to sub-luminous sub-classes of SNe Ia. The He-detonation model has been first considered within the SD scenario (the 'double detonation' model), i.e., a C+O WD accreting from a non-degenerate He star, while it has been recently discussed in a context of the DD scenario as well, i.e., a C+O WD merging with another C+O WD with a massive He layer or even with a He WD.[14,15]

Infracting opportunities of the time domain science now allow SNe and transient objects to be studied from the very beginning of their outbursts. The discovery and quick follow-up observations within the first few days since the SN explosions are becoming common in the last few years. For SNe Ia, it has been suggested that an additional blue (UV-strong) and bright emission should be associated with the SD scenario; an interaction of the ejecta with a companion star creates an additional thermal energy, which is then radiated away in a few days, superimposed to the 'red' light diffusing out of the main body of the SN ejecta as powered by the decay chain $^{56}$Ni/Co/Fe.[16] A similar excessive light in the first few days is expected for the interaction between the ejecta and circumstellar media (CSM).[17] As yet another possibility, mixing of $^{56}$Ni towards the outer region of the ejecta may also contribute to the early emission.[18] The first attempt to search for such a signal generally resulted in non-detection; however, as the observational technique and strategy are

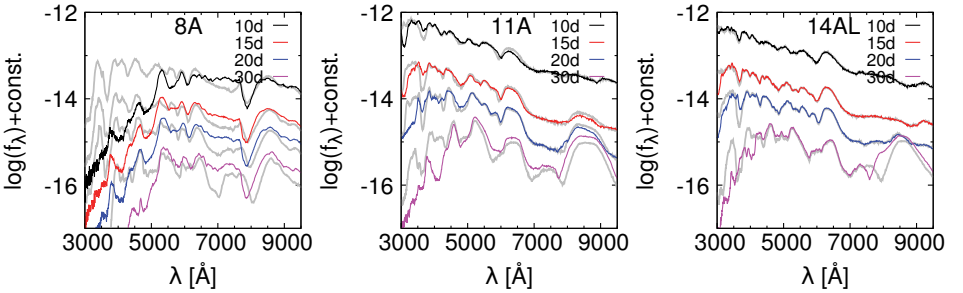



Fig. 2. Synthetic spectra of the He detonation models around maximum light (8A, 11A, 14AL; see the Fig. 1 caption). The spectra of the corresponding models without the He detonation products at the same epochs are also shown by gray lines. Reproduced from the following publication: K. Maeda, et al., The Astrophysical Journal, 861, 78 (2018).[21]

being matured, there is an increasing number of reports for SNe showing the early excess in the first few days (see Ref. 1 for a review).

## 2. He detonation: Expected observational features

Recently, an additional mechanism has been suggested which would create a bright emission in the first few days of SNe Ia; it is the He detonation.[19–21] The He detonation creates either $^{56}$Ni or other radioactive isotopes ($^{52}$Fe which decays into Mn and then Cr, or $^{48}$Cr which decays into V and then Ti). The newly synthesized radioactive isotopes are contained in the thin outermost layer of expanding SN ejecta. Given the small diffusion time there, it creates a bright emission in the first few days, then it is eventually overwhelmed by the main SN emission. The mass of the He layer would be between $0.004M_{\odot}$ (for $M_{\rm WD} \sim 1.4M_{\odot}$) to $\sim 0.13M_{\odot}$ (for $M_{\rm WD} \sim 0.8M_{\odot}$) for the SD scenario, but it is highly uncertain (or can be quite diverse) in the DD scenario. The amount of the surface radioactivity, which determines both the luminosity and duration of the early flash, can be diverse but can create optical signals similar to the predictions from the companion/CSM interaction scenarios. An important difference between the He detonation and the interaction scenarios is the UV-optical color; it is much redder in the He detonation model. Examples of synthetic observables from the He detonation models are shown in Figure 1.[21]

The He detonation would result in highly red spectra around the maximum phases,[13,21,22] as the UV portion of the light experiences substantial line absorptions by the He detonation ash (Figure 2). This feature has indeed been a main argument against the He detonation model to account for a main population of SNe Ia. We however note that this prediction is highly sensitive to the combination of the SN luminosity (which determines ionization) and the mass of the He shell (which determines the mass of the absorbers); either a very thin shell and/or a high luminosity would erase the potential signature of the UV absorptions.[21] This

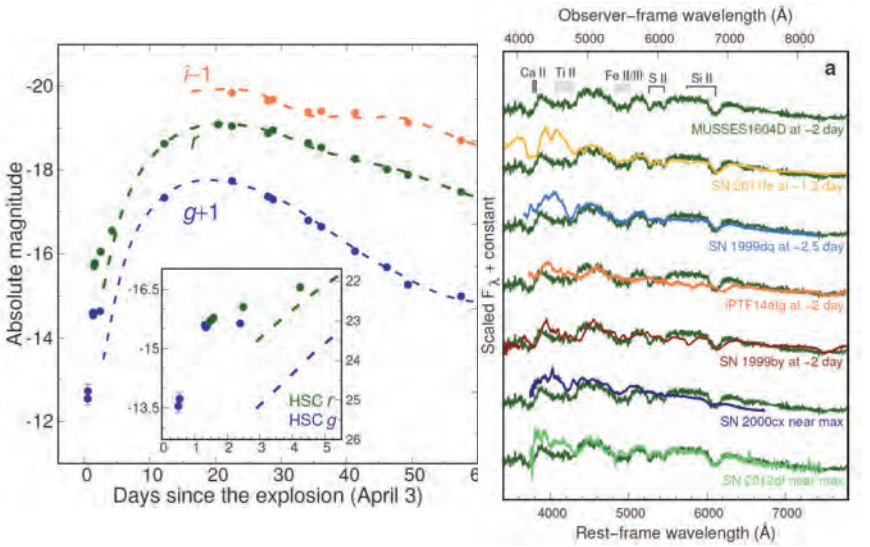



Fig. 3. Multi-band light curves (left) and a spectrum around the maximum light (right) of MUSSES1604D. The right panel shows comparisons to various SNe Ia; the spectrum of MUSSES1604D is shown by the dark green line, while the comparison SNe are shown in different colors. Reproduced from the following publication: J. Jiang, et al., Nature, 550, 80 (2017).[19]

feature adds a predicting power to the He detonation scenario: the properties in the early flash and those in the maximum phases are coupled, and therefore the combined analysis provides a strong diagnostics for the He detonation scenario.

## 3. MUSSES1604D: The first robust candidate of the He detonation

MUSSES1604D is a peculiar SN Ia discovered by a survey using the Hyper Suprime-Cam (HSC) attached with the Subaru telescope.[19] The survey is designed to make the best use of the wide-field capability of the HSC and the deep sensitivity of the 8.2m Subaru telescope, in order to discover relatively high-z SNe Ia in the first day of the explosion.

MUSSE1604D shows an excess of the light in the first few days, as compared to a simple extrapolation from the template SN Ia light curve (Fig. 3). Indeed, except for this early excessive emission, the light curve is largely consistent with normal SNe Ia. Around the maximum phase, it shows clear peculiarities in its spectral features. While it generally resembles normal SNe Ia, additional absorption features are seen in the blue part ($\lesssim 4,500$Å). These absorption features are identified as Ti II and Fe-peaks, which are seen in sub-luminous SNe Ia but not in normal SNe Ia.

A straightforward interpretation is the He detonation-triggered explosion.[19,21] The scenario predicts both the emission in the earliest phase (due to the radioactive power) and the absorptions around the maximum phase (due to the decayed Fe-peak elements and Ti). As a demonstration, radiation transfer simulations are

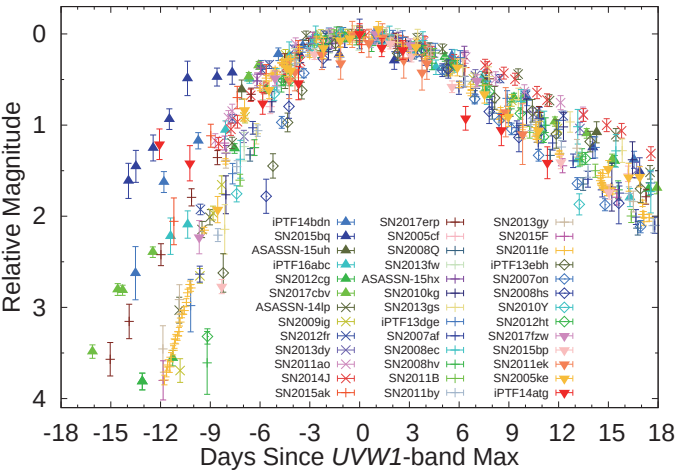



Fig. 4. Light curves of SNe Ia in the SWIFT $UVW1$-band. Reproduced from the following publication: J. Jiang, et al., The Astrophysical Journal, 865, 149 (2018).[23]

performed on a grid of simplified kinematic models which are constructed for a set of parameters; the ejecta mass (which is translated to the mass of the primary WD), the mass of $^{56}$Ni in the main SN ejecta, and the mass of the He shell. The best match is found for $1.28 - 1.38 M_{\odot}$ for the WD mass, $\sim 0.43 M_{\odot}$ for the $^{56}$Ni mass, and $\sim 0.02 - 0.03 M_{\odot}$ for the He shell mass. We note that this combination in the WD mass and the $^{56}$Ni mass deviates from the expectation from the carbon detonation, and this issue is yet to be clarified. The mass of the He shell is largely consistent with the classical double detonation model within the SD scenario.

The rate of MUSSES1604D-like SNe Ia is small; only $\sim 0.5\%$ of SNe Ia show the combination of normal peak luminosity and the peculiar maximum spectra.[19,23] We however note that the properties of the maximum-phase spectra are highly sensitive to the combination of the luminosity and the He shell mass; the MUSSES1604D-like objects may represent a tip of icebergs of the 'He detonation population'.[21] Indeed, the second robust candidate of the He detonation-triggered SN Ia 2018byg has different characteristics than MUSSES1604D, but the variation is understandable within the He detonation model sequence.[24]

## 4. Diversity of SNe Ia in the earliest phases

There are some reports of the detection of the early-phas excess for individual SNe Ia (see Ref. 23 for a review). A systematic study of existing data with early-phase photometric observations well before the maximum light (Fig. 4), which include the objects which were not reported previously, show an interesting trend;[23,25] most, if not all, of SNe Ia showing the early-phase excess are classified as bright SN 1991T/1999aa-like SNe Ia (or the 'shallow Si class'). The exceptions

are iPTF14atg[26] and MUSSES1604D, for which the companion interaction and He detonation, respectively, have been suggested for the mechanism responsible for the early excess.

Furthermore, all the bright SNe Ia with sufficient data in the early phase show the excess. Given that the companion interaction predicts a substantial viewing direction dependence on the strength of the early flash,[16,27] this possibility is immediately rejected to explain the early-phase excess seen in these bright SNe Ia. Further, the duration is long ($\sim$ 10 days) and the color is red, which are inconsistent with the CSM interaction scenario. These properties suggest that the mechanism to create the early excess for the bright SNe Ia is likely the mixing of $^{56}$Ni outward during the explosion.[21,23] Indeed, modeling the spectral sequence for the bright SNe Ia suggests that the $^{56}$NI distribution is extended toward the outer layer,[28] which may support the interpretation. The strength of the signal is diverse; this may indicate that a different degree of the mixing is realized in these SNe Ia.

## 5. Concluding remarks

In this paper, we have summarized our ongoing activity to understand the progenitors and explosion mechanisms of SNe Ia, highlighting one particular mode of the explosion, i.e., the explosion triggered by the He detonation on the surface of a C+O WD. Our conclusions are summarized as follows:

- The He detonation model offers a new mechanism to create the excessive emission from SNe Ia in the first few days, adding to the previously suggested mechanisms (companion interaction, CSM interaction, and $^{56}$Ni mixing).
- The He detonation has a unique prediction; the properties seen in the earliest phase are connected to those in the maximum phases, through the absorption by the He detonation ash.
- Properties of MUSSES1604D, as discovered by the Subaru/HSC high-cadence survey likely within the day of the explosion, are best explained by the He detonation scenario.
- Except for two SNe Ia (iPTF14atg and MUSSES1604D), SNe with the early excess show strong preference to a bright SN 1991T/1999aa-like subclass. The most likely mechanism to create the early excessive emission for these bright SNe is the mixing of $^{56}$Ni towards the outer layers during the explosion.
- A combined analysis of the very early phase observations and the maximum-phase observations will be key in mapping the diverse SN Ia zoo into different populations reflecting different progenitors and/or different explosion modes.

**Acknowledgments**

The authors thank the organizers of the successful MG15 meeting, and all the participants in the BN4 session for stimulating discussion. This work has been partly supported by JSPS KAKENHI (18H04585, 18H05223, 17H02864.) from the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT).

**References**

1. K. Maeda, Y. Terada, *International Journal of Modern Physics D* **25**, 1630024 (2016).
2. J. Whelan, I. Iben, Jr., *ApJ* **186**, 1007 (1973).
3. K. Nomoto, *ApJ* **253**, 798 (1992).
4. I. Iben, Jr., A. V. Tutukov, *ApJ* **284**, 719 (1984).
5. R. F. Webbink, *ApJ* **277**, 355 (1984).
6. A. M. Khokhlov, *A&A* **245**, 114 (1991).
7. D. Arnett, E. Livne, *ApJ* **427**, 315 (1994).
8. K. Iwamoto, et al., *ApJS* **125**, 439 (1999).
9. E. Livne, A. S. Glasner, *ApJ* **361**, 244 (1990).
10. S. E. Woosley, T. A. Weaver, *ApJ* **423**, 371(1994).
11. K. J. Shen, L. Bildsten, *MNRAS* **699**, 1365 (2009).
12. M. Fink, et al., *A&A* **514**, 53 (2010).
13. S. E. Woosley, D. Kasen, *ApJ* **734**, 38 (2011).
14. K. J. Shen, et al., *ApJ* **865**, 15 (2018).
15. A. Tanikawa, K. Nomoto, N. Nakasato, *ApJ* **868**, 90 (2018).
16. D. Kasen, *ApJ* **708**, 1025 (2010).
17. A. L. Piro, V. S. Morozova, *ApJ* **826**, 96 (2016).
18. A. L. Piro, *ApJ* **759**, 83 (2012).
19. J. Jiang, et al., *Nature* **550**, 80 (2017).
20. U. M. Naebauer, et al., *MNRAS* **463**, 2972 (2016).
21. K. Maeda, J. Jiang, M. Doi, T. Shigeyama, *ApJ* **861**, 78 (2018).
22. M. Kromer, et al., *ApJ* **719**, 1067 (2010).
23. J. Jiang, M. Doi, K. Maeda, T. Shigeyama, *ApJ* **865**, 149 (2018).
24. K. De, et al., *ApJ* **873**, L18 (2019).
25. M. D. Stritzinger, et al., *ApJ* **864**, L35 (2018).
26. Y. Cao, et al., *Nature* **521**, 328 (2015).
27. M. Kutsuna, T. Shigeyama, *PASJ* **67**, 54 (2015).
28. J. Zhang, et al., *ApJ* **817**, 114 (2016).

# Spin flip-flops from secular dynamics of compact binaries

Keresztes, Zoltán, Tápai, Márton and Gergely, László Á.

*Institute of Physics, University of Szeged, Dóm tér 9, Szeged, H-6720, Hungary*
*E-mail: zkeresztes.zk@gmail.com,laszlo.a.gergely@gmail.com*

Compact binary systems are investigated on timescales where the gravitational radiation backreaction is negligible but which is much longer than the orbital period. On this conservative timescale we use averaged equations over one orbital period removing the short time-scale modulations but keeping the secular variations. The effects of the eccentricity on the spin flip-flop is studied for equal mass binary systems. We find that eccentricity does not affect the flip-flop magnitude but it has a significant influence on its period.

## 1. Introduction

The first direct detection of gravitational waves was announced in 2016 by the LIGO Scientific Collaboration and Virgo Collaboration, one century after Einsten proposed their existence. Ten additional events have emerged since the first detection. All observed gravitational waves originated from compact binary systems,[1–8] and in two cases the spin of the merging black holes was identified with high significance.[2,8,9] Therefore the accurate description of the orbital and spin dynamics of compact binary systems is important. Effects due to the spin evolution can also occur before the merger. The precession of the dominant spin (of the larger mass supermassive black hole of the binary) was identified for the first time through its jet characteristics from VLBI radio data spanning over 18 years.[10]

The evolution of the compact binary system consists of three consecutive phases: i) inspiral, ii) merger and iii) ringdown. In the first regime the dynamics can be described by the post-Newtonian (PN) formalism. On the conservative timescale where the gravitational radiation backreaction is negligible the spin effects appear throught the spin–orbit (SO), the spin–spin (SS) and the quadrupole[a]–monopole (QM) couplings, giving contributions to the relative acceleration of the binary components and leading to spin precessional equations.[11,12] The conservative timescale contains two additional subscales. The shorter is given by the orbital period, while the longer by the precessional period of the orbital plane and spins. On the orbital timescale the spin effects on the orbit were analyzed setting up a generalized Kepler equation.[13–16] On the precessional timescale the effects occurring on the shortest timescale can be swept away by averaging methods. Then a simplified system of equations is obtained, governing the longtime spin evolution. The averaged dynamics for circular orbits with SO and SS coupling terms[17] and for generic bounded orbits with SO couplings[18,19] were derived. With the inclusion of the SO, SS and QM couplings Racine identified a new constant of motion for black holes which have specific quadrupole parameter and solved analytically the averaged spin equations

[a]If the quadrupole moment is spin induced.

for circular orbits.[20] By an additional averaging over the precession timescale the evolution of the magnitude of total angular momentum during the inspiral was investigated.[21,22] These efforts were exploited for building waveforms.[23]

Several interesting spin related behaviours were identified in compact binary dynamics. When the orbital angular momentum and the total spin are anti-aligned and nearly cancel each other the orbital plane changes significantly during a relatively short-lived transitional precession period[17]. In binaries having components with significantly different masses, the direction of dominant spin relative to the total angular momentum can change significantly during the inspiral.[24] This spin-flip may explain the formation of X-shape radio galaxies.[25] When the spin vector of the larger black hole $\mathbf{S_1}$ (dominant spin) approximately lies in the plane of motion, while the smaller spin $\mathbf{S_2}$ is closely aligned with the orbital angular momentum $\mathbf{L}_N$, the smaller spin slowly evolves to be anti-aligned with $\mathbf{L}_N$, then periodically changes back and forth on a timescale shorter than the gravitational radiation reaction timescale.[26] This effect, dubbed spin flip-flop was investigated in a wider parameter range for binaries moving on circular orbits.[27–29] Recently a parameter range has been identified where the flip-flop happens on relatively short timescales, and dubbed as wide precession.[30] Then over the period during which the magnitude of the total spin changes from its minimum to its maximum and back to the minimum value, one of two spins evolves from complete alignment with $\mathbf{L}_N$ to complete anti-alignment.

Here we investigate the effects of the eccentricity on the magnitude and the period of spin flip-flop on conservative timescale. For this purpose we use averaged equations over one orbital period keeping only the secular effects.

## 2. Secular spin evolution of compact binaries moving on eccentric orbits

The orbital and spin dynamics of the compact binary systems up to 2PN orders with SO, SS and QM couplings is given by equations (36)–(42)[31] in terms of dimensionless osculating orbital elements and spin angles. This system of equations describes the motion of an object with reduced mass $\mu$ around another component with total mass $m$. There are two variables which characterize the shape of the reduced mass' orbit:

$$\mathfrak{l}_r = \frac{cL_N}{Gm\mu} \ , \tag{1}$$

and the orbital eccentricity

$$e_r = \frac{A_N}{Gm\mu} \ , \tag{2}$$

with speed of light $c$, gravitational constant $G$, magnitudes $L_N$ and $A_N$ of the orbital angular momentum $\mathbf{L}_N$ and of the Laplace-Runge-Lenz vector $\mathbf{A}_N$. There are three Euler angles in the system of equations: the inclination $\alpha = \arccos\left(\hat{\mathbf{J}} \cdot \hat{\mathbf{L}}_N\right)$,

the longitude of the ascending node $\phi_n = \arccos\left(\hat{\mathbf{x}} \cdot \hat{\mathbf{l}}\right)$ and the argument of the periastron $\psi_p = \arccos\left(\hat{\mathbf{l}} \cdot \hat{\mathbf{A}}_N\right)$. Here $\mathbf{J}$ is the total angular momentum, conserved in the 2PN dynamics, $\hat{\mathbf{l}} = \hat{\mathbf{L}}_\mathbf{N} \times \hat{\mathbf{J}}$, $\hat{\mathbf{x}}$ denotes a fixed direction perpendicular to $\hat{\mathbf{J}}$ and the hat denotes unit vectors. The magnitudes of the spin vectors are constant and they are characterized by the dimensionless quantities:

$$\chi_i = S_i \frac{c}{G} m_i^{-2} \ . \tag{3}$$

The polar angles of the spin vectors are

$$\kappa_i = \arccos\left(\hat{\mathbf{S}}_i \cdot \hat{\mathbf{L}}_\mathbf{N}\right) \ , \tag{4}$$

and their azimuthal angles

$$\zeta_i = \arccos\left(\hat{\mathbf{S}}_i^{\perp} \cdot \hat{\mathbf{A}}_N\right) \ , \tag{5}$$

with $\mathbf{S}_i^{\perp} = \mathbf{S}_i - \hat{\mathbf{L}}_\mathbf{N}\left(\mathbf{S}_i \cdot \hat{\mathbf{L}}_\mathbf{N}\right)$.

We investigate the secular effects which occur over many orbital periods. The average of any function $f$ over one orbital period is

$$T\bar{f} = \int_0^T f\left(t\right) dt = \int_0^{2\pi} f\left(\chi_p\right) \frac{1}{\dot{\chi}_p} d\chi_p \ , \tag{6}$$

where $\chi_p$ is the true anomaly, the angle between $\hat{\mathbf{A}}_N$ and the position of the reduced mass. The distance $r$ between the binary components changes periodically, thus

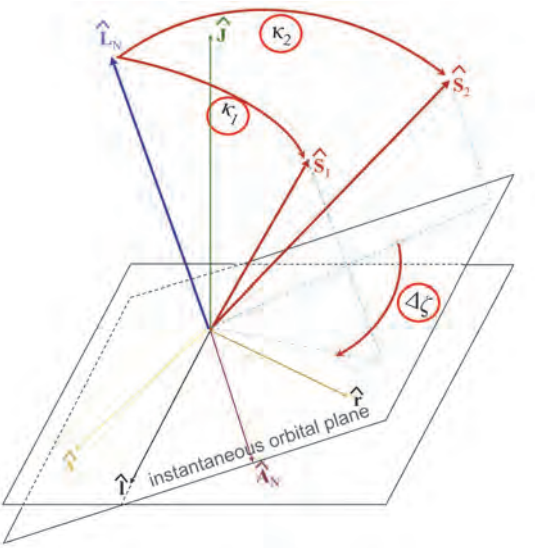



Fig. 1. The spin angles $\kappa_i$ and $\Delta\zeta = \zeta_1 - \zeta_2$ are shown.

there are turning points, where $\dot{r} = 0$. The parameter $\chi_p$ is chosen such that $\dot{r}(\chi_p = k\pi) = 0$, where $k$ is an integer number. The orbital period $T$ is the time spent between $\chi_p = 0$ and $\chi_p = 2\pi$. Applying Eq. (6) for the time derivatives $f = \left\{\dot{\iota}_r, \dot{e}_r, \dot{\alpha}, \dot{\phi}_n, \dot{\psi}_p, \dot{\kappa}_i, \dot{\zeta}_i\right\}$ leads to the secular dynamics. These equations are given with 1PN corrections and SO coupling.[32] The averaged equations valid up to 2PN orders will be presented.[33] We note that the shape variables are constants of motion $\dot{\bar{\iota}}_r = \dot{\bar{e}}_r = 0$ on the secular time-scale (where the gravitational radiation backreaction is negligible). In addition, the equations for the variables $\kappa_i$ and $\Delta\zeta = \zeta_1 - \zeta_2$ (shown on Fig. 1) are decoupled from the set of equations governing the evolutions of $\alpha, \phi_n, \psi_p$ and $\zeta_1 + \zeta_2$.

On Fig. 2 the color map highlights the values of the difference between the extremal values of $\kappa_2$ occuring during the evolution, denoted $\Delta\kappa_2$, for an equal mass binary system in the parameter space span by $\kappa_1(0)$ and $\kappa_2(0)$. The orbital eccentricity and the initial value of $\Delta\zeta$ were chosen as $\bar{e}_r = 0.1$ and $\Delta\zeta(0) = \pi/2$. On the left panel $\chi_1 \ll \chi_2$. Large flip-flops with $\Delta\kappa_2 \approx \pi$ occur at $\kappa_1(0) \approx \pi/2$, originally identified for circular orbits.[26] For $\chi_1 = \chi_2 = 0.95$ the right panel shows that large flip-flops appear in completely different regions than for $\chi_1 \ll \chi_2$.

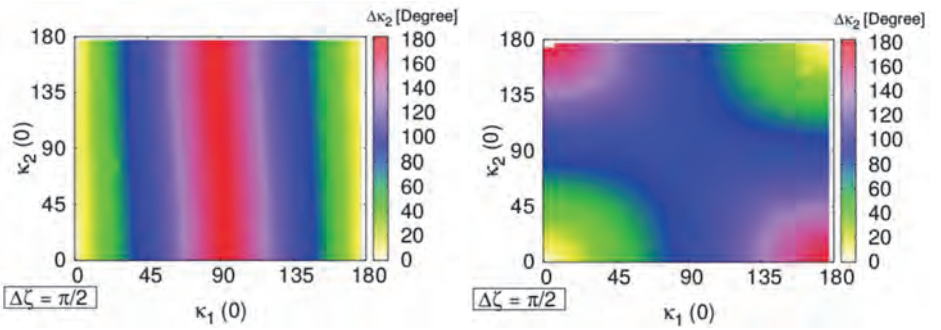



Fig. 2. The color map represents $\Delta\kappa_2$ (the difference between the extremal values of $\kappa_2$ occuring during the evolution) as a function of $\kappa_1(0)$ and $\kappa_2(0)$. On the left panel $\chi_1 = 0.95$ and $\chi_2 = 0.05$ corresponding to $\chi_2 \ll \chi_1$ while on the right panel $\chi_1 = \chi_2 = 0.95$ (in this case $\Delta\kappa_2 = \Delta\kappa_1$). The additional fixed parameters are $\Delta\zeta(0) = \pi/2$, $m_1 = m_2$, $m = 20M_\odot$, $\bar{e}_r = 0.1$ and $\bar{\varepsilon} = 0.0001$.

The period of the flip-flop during which $\kappa_2$ changes from its minimum to its maximum and back to its minimum is shown on Fig. 3 as a function of eccentricity. The flip-flop period (a dimensionless number) is given in units of $Gm/c^3$. On the left panel $\chi_1 \ll \chi_2$ while on the right panel $\chi_1 = \chi_2 = 0.95$, and $\Delta\zeta(0) = \pi/2$ on both panels. At the orange, white and black contour lines $\Delta\kappa_2$ has the values $\pi/3$, $\pi/2$ and $2\pi/3$, respectively. Since these lines are horizontal, $\Delta\kappa_2$ does not depend on the eccentricity. We have confirmed this property in a wide parameter range for binaries with equal mass.

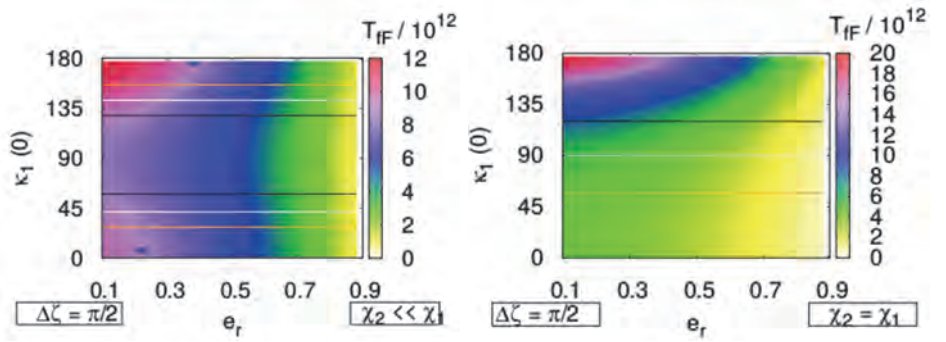



Fig. 3. The color map shows the flip-flop periods in dimensionless units as a function of $\bar{e}_r$. At the horizontal orange, white and black contour lines $\Delta\kappa_2$ has the values $\pi/3$, $\pi/2$ and $2\pi/3$, respectively. On the left panel $\chi_1 = 0.95$ and $\chi_2 = 0.05$ corresponding to $\chi_2 << \chi_1$ while on the right panel $\chi_1 = \chi_2 = 0.95$ (in this case $\Delta\kappa_2 = \Delta\kappa_1$). Common parameters for both panels are $\kappa_2\,(0) = \pi/10$, $\Delta\zeta\,(0) = \pi/2$, $m_1 = m_2$, $m = 10^5 M_\odot$ and $\bar{\varepsilon} = 0.0001$.

## 3. Conclusions

Spin flip-flop effects were considered for equal mass binaries moving on eccentric orbits. When one of the binary components is slowly spinning with respect to the other, the largest variation in the polar angle of the smaller spin occurs when the dominant spin is near perpendicular to the orbital angular momentum. However the magnitude of the flip-flop strongly depends on the ratio of the spin parameters. For binaries with same identical parameters the flip-flop magnitude is largest in the near regions of the two corner points ($\kappa_1\,(0) = 0$, $\kappa_2\,(0) = \pi$) and ($\kappa_1\,(0) = \pi$, $\kappa_2\,(0) = 0$) (see Fig. 2).

We have found that the magnitude of the flip-flop is unchanged for different orbital eccentricities. However the eccentricity significantly influences the period of the flip-flop. As the eccentricity increases, the flip-flop period decreases. Comparing the low ($\bar{e}_r \approx 0.1$) and high ($\bar{e}_r \approx 0.9$) eccentric orbits, the period decreases by one order of magnitude in certain cases, for fixed $\kappa_1\,(0)$.

## Acknowledgments

This work was supported by the Hungarian National Research Development and Innovation Office (NKFI) in the form of the grant 123996. The work of Z. K. was also supported by the János Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the UNKP-18-4 New National Excellence Program of the Ministry of Human Capacities.

## References

1. B. P. Abbott et al., *Phys. Rev. Lett.* **116**, 061102 (2016).
2. B. P. Abbott et al., *Phys. Rev. Lett.* **116**, 241103 (2016).

3. B. P. Abbott et al., *Phys. Rev. Lett.* **118**, 221101 (2017).
4. B. P. Abbott et al., *Phys. Rev. Lett.* **119**, 141101 (2017).
5. B. P. Abbott et al., *Phys. Rev. Lett.* **119**, 161101 (2017).
6. B. P. Abbott et al., *Astrophys. J. Lett.* **848**, L13 (2017).
7. B. P. Abbott et al., *Astrophys. J.* **851**, L35 (2017).
8. B. P. Abbott et al., `arXiv:1811.12907` (2018).
9. C. Kimball, C. P. L. Berry and V. Kalogera, `arXiv:1903.07813` (2019).
10. E. Kun, K. É. Gabányi, M. Karouzos, S. Britzen, L. Á. Gergely, *Mon. Not. Royal Astron. Soc.* **445** (2), 1370 (2014).
11. B. M. Barker and R. F. O'Connell, *Phys. Rev. D* **2**, 1428 (1970).
12. B. M. Barker and R. F. O'Connell, *Phys. Rev. D* **12**, 329 (1975).
13. N. Wex, *Class. Quantum Grav.* **12**, 983 (1995).
14. C. Königsdörffer and A. Gopakumar, *Phys. Rev. D* **71**, 024039 (2005).
15. Z. Keresztes, B. Mikóczi and L. Á. Gergely, *Phys. Rev. D* **71**, 124043 (2005).
16. C. Königsdörffer and A. Gopakumar, *Phys. Rev. D* **73**, 044011 (2006).
17. T. A. Apostolatos, C. Cutler, G. J. Sussman and K. S. Thorne, *Phys. Rev. D* **49**, 6274 (1994).
18. R. Rieth and G. Schäfer, *Class. Quantum Grav.* **14**, 2357 (1997).
19. L. Á. Gergely, Z. I. Perjés and M. Vasúth, *Phys. Rev. D* **58**, 124001 (1998).
20. É. Racine, *Phys. Rev. D* **78**, 044021 (2008).
21. M. Kesden, D. Gerosa, R. O'Shaughnessy, E. Berti and U. Sperhake, *Phys. Rev. Lett.* **114**, 081103 (2015).
22. D. Gerosa, M. Kesden, U. Sperhake, E. Berti and R. O'Shaughnessy, *Phys. Rev. D* **92**, 064016 (2015).
23. K. Chatziioannou, A. Klein, N. Cornish and N. Yunes, *Phys. Rev. Lett.* **118**, 051101 (2017).
24. L. Á. Gergely and P. L. Biermann, *Astrophys. J.* **697**, 1621, (2009).
25. Gopal-Krishna, P. L. Biermann, L. Á. Gergely and P. J. Wiita, *Res. Astron. Astrophys.* **12**, 127 (2012).
26. C. O. Lousto and J. Healy, *Phys. Rev. Lett.* **114**, 141101 (2015).
27. D. Gerosa, M. Kesden, R. O'Shaughnessy, A. Klein, E. Berti, U. Sperhake and D. Trifirn, *Phys. Rev. Lett.* **115**, 141102 (2015).
28. C. O. Lousto and J. Healy, *Phys. Rev. D* **93**, 124074 (2016).
29. C. O. Lousto, J. Healy and H. Nakano, *Phys. Rev. D* **93**, 044031 (2016).
30. D. Gerosa, A. Lima, E. Berti, U. Sperhake, M. Kesden and R. O'Shaughnessy, `arXiv:1811.05979` (2018).
31. L. Á. Gergely and Z. Keresztes, *Phys. Rev. D* **91**, 024012 (2015).
32. L. Á. Gergely, Z. Keresztes and M. Tápai, *Universe* **4**, 40 (2018).
33. Z. Keresztes, M. Tápai and L. Á. Gergely, in preparation (2019).

# A public framework for Feynman calculations and post-Newtonian gravity

Michele Levi

*Institut de Physique Théorique, CEA & CNRS, Université Paris-Saclay*
*91191 Gif-sur-Yvette, France*
*michele.levi@ipht.fr*

We report here on a line of work that has played a key role in formally establishing and going beyond the state of the art in the effective field theory (EFT) approach and in post-Newtonian (PN) gravity. We also outline here how this comprehensive framework in fact forms the outset of a prospective rich research program, building on the public Feynman and PN technology developed.

*Keywords*: Effective field theories; Post-Newtonian theory; Gravitational waves; High precision calculations.

## 1. Introduction

In recent years significant progress was made in a line of research, based on an Effective Field Theory (EFT) approach for the treatment of gravitational waves (GWs), emitted from coalescing binaries, which was put forward by Goldberger et al.[1–3]. Such binaries contain compact components, which still orbit in non-relativistic (NR) velocities in their long inspiral phase, and are therefore analytically studied via the post-Newtonian (PN) approximation in gravity[4]. Here we report in particular on a series of works, which focused on the treatment of the challenging and realistic case, where the constituents of the binary are rotating[5–13], and on making the general EFT framework widely accessible and available to the community, including a public automated computation tool[14–16]. This line of work has played a key role in formally establishing and going beyond the state of the art in the EFT approach and in PN gravity, where we also outline here how this comprehensive framework in fact forms the outset of a prospective rich research program, building on the public Feynman and PN technology developed. For a general overview presentation of the main unique advances in the field of EFTs of PN gravity, and the various prospects of using field theory to study gravity theories at all scales, see Ref. 17.

## 2. Effective Field Theories of Post-Newtonian Gravity

Multi-scale problems are prevalent in physics, and as it turns out the compact binary inspiral is also of this kind. Such problems are amenable to a systematic treatment using the framework of EFTs, originally developed in the context of quantum field theories (QFTs) and particle physics. The compact binary inspiral involves indeed a hierarchy of scales, controlled by the NR orbital velocity, $v$, which constitutes the small parameter of the effective theories at each of the intermediate scales in the problem. There are three widely separated characteristic scales in the binary inspiral problem: The scale of the single compact object, $r_s$, the scale of the orbital

separation of the binary, $r$, and the scale of radiation, emitted from the binary in GWs, with a wavelength $\lambda$. These scales satisfy $r_s \sim Gm \sim rv^2 \sim \lambda v^3$, where $m$ is the typical mass of the single compact object, $v \ll 1$, and in the PN approximation $n\text{PN} \equiv v^{2n}$ order corrections to Newtonian gravity are found. Therefore one needs to proceed to construct a tower of EFTs in stages, where in each stage an effective theory is devised to remove the corresponding intermediate characteristic scale.

The setup of EFTs is then universal, and it follows one of the two complementary procedures to construct an EFT, which are referred to as "bottom-up" and "top-down". Both of these approaches are also used in the binary inspiral problem. At the first stage, where the scale of the single object is to be eliminated, the full theory of gravity is considered as General Relativity (GR), and is given, e.g., in terms of the following Einstein-Hilbert action:

$$S[g_{\mu\nu}] = -\frac{1}{16\pi G}\int d^4x\sqrt{g}\,R[g_{\mu\nu}], \tag{1}$$

with $g_{\mu\nu}$ for the gravitational field. Then, in the bottom-up approach we construct the effective action by identifying the relevant degrees of freedom (DOFs) and the symmetries at the desirable scale. In this case, this means that the remaining bulk action, after decomposing the gravitational field into $g_{\mu\nu} \equiv g^s_{\mu\nu} + \bar{g}_{\mu\nu}$, where $g^s_{\mu\nu}$ represents the strong field modes to be removed in the effective action, is augmented with a worldline action, which contains new worldline DOFs of a point particle. The effective action is then of the following form:

$$\begin{aligned} S_{\text{eff}}[\bar{g}_{\mu\nu}(x), y^\mu(\sigma), e^\mu_A(\sigma)] &= S[\bar{g}_{\mu\nu}] + S_{\text{pp}}[\bar{g}_{\mu\nu}(y), y^\mu(\sigma), e^\mu_A(\sigma)] \\ &= -\frac{1}{16\pi G}\int d^4x\sqrt{\bar{g}}\,R[\,\bar{g}_{\mu\nu}] + \sum_{i=1}^{\infty} C_i(r_s)\int d\sigma\,\mathcal{O}_i(\sigma), \end{aligned} \tag{2}$$

where the worldline DOFs depend on the worldline parameter $\sigma$, $\mathcal{O}_i$ are the operators evaluated on the worldline, and $C_i$ denote the "Wilson coefficients", which encapsulate the UV physics suppressed, and depend only on the scale $r_s$.

At the second stage, the field component is further decomposed into components with definite scaling in the velocity, i.e. $\bar{g}_{\mu\nu} \equiv \eta_{\mu\nu} + H_{\mu\nu} + \widetilde{h}_{\mu\nu}$, where $H_{\mu\nu}$ denotes the orbital field modes, which we want to eliminate at this stage, and $\widetilde{h}_{\mu\nu}$ denotes the radiation modes. We then proceed with the top-down approach to integrate out the orbital scale by starting from the following two-particle action, obtained from the previous stage:

$$S_{\text{eff}}[\bar{g}_{\mu\nu}, y^\mu_1, y^\mu_2, e^{\ \mu}_{1\,A}, e^{\ \mu}_{2\,A}] = S[\bar{g}_{\mu\nu}] + \sum_{a=1}^{2} S_{pp}[\bar{g}_{\mu\nu}(y_a), (y_a)^\mu, (e_a)^{\ \mu}_{A}](\sigma_a). \tag{3}$$

We integrate out the orbital field modes explicitly, using a standard diagrammatic expansion with the following definition for the new effective action:

$$e^{iS_{\text{eff}}[\widetilde{h}_{\mu\nu},(y_c)^\mu,(e_c)^{\ \mu}_{A}]} \equiv \int \mathcal{D}H_{\mu\nu}\, e^{iS_{\text{eff}}[\bar{g}_{\mu\nu},y^\mu_1,y^\mu_2,e^{\ \mu}_{1\,A},e^{\ \mu}_{2\,A}]}, \tag{4}$$

where now there are new worldline DOFs, with the subscript 'c', of a composite particle, that is the binary system.

Finally, when the conservative sector is concerned, no radiation modes are present, and one should arrive at an effective action, which no longer contains field modes, or another scale to remove.

## 3. Spinning Particles in Gravity

Gravitating spinning objects have been considered as particles in an action approach both in special and general relativity already more than 40 years ago, see e.g. Refs. 18, 19, yet they were first tackled in the context of the EFT approach for PN computations of the binary inspiral in Ref. 20. Spin in Relativity necessarily implies having a non-vanishing finite size[a], i.e. a deviation from the point particle view, which is also adopted in the EFT approach, and this poses a challenge for the construction of an EFT for a gravitating spinning particle. Moreover, once we leave Newtonian physics, a unique notion of a "center of mass" is lost, and hence there is an ambiguity as to the choice of the variables describing the rotation of the object.

Nevertheless, as we noted above the way to proceed is to construct an EFT for a gravitating spinning particle, following a bottom-up procedure, i.e. by identifying the relevant DOFS, and coupling them in all possible ways allowed by the proper symmetries at this scale. After this first EFT is obtained, we can proceed to explicitly compute, using the top-down procedure starting from the two-particle EFT, the PN binding interactions of the binary system, i.e. the second EFT of a composite particle with spins. The formulation of these EFTs was indeed introduced in Ref. 9 (see also Ref. 16 for a comprehensive pedagogic presentation), and allowed in particular to constrain spin-induced finite size effects to all orders in spin, and to arrive at equations of motion (EOMs) and Hamiltonians in a straightforward manner.

### 3.1. *A gravitating spinning particle*

Let us then go over the relevant DOFs and symmetries for the EFT at the first stage[9]. First, it is assumed that the isolated object has no intrinsic permanent multipole moments beyond the spin dipole.

Then, for a spinning object, there are two types of worldline DOFs, added to the gravitational field DOFs. The crucial points to highlight for these three types of DOFs, specific to the spinning case, are: 1. The tetrad field, $\eta^{ab}\tilde{e}_a{}^\mu(x)\tilde{e}_b{}^\nu(x) = g^{\mu\nu}(x)$, represents the gravitational field DOFs, rather than simply the metric, $g_{\mu\nu}(x)$, in order to couple the field to the rotational particle DOFs; 2. The spinning particle position in the worldline coordinates, $y^\mu(\sigma)$, does not necessarily represent the rotating object's "center"; 3. The worldline tetrad, $\eta^{AB}e_A{}^\mu(\sigma)e_B{}^\nu(\sigma) = g^{\mu\nu}$,

[a]Otherwise, the rotational velocity of the object would surpass the speed of light.

is considered for the worldline rotational DOFs, and enables to define the worldline angular velocity, $\Omega^{\mu\nu}(\sigma) \equiv e_A^\mu \frac{De^{A\nu}}{D\sigma}$, so that its conjugate, the worldline spin, $S_{\mu\nu}(\sigma)$, is considered as a further rotational DOF in the action.

Similarly, when considering the spinning case, further symmetries should be taken into account in addition to general covariance and worldline reparametrization invariance, which arise already in the point mass case. First, there are parity invariance, and Lorentz invariance of the tetrad field. For the worldline tetrad, there is the $SO(3)$ invariance of the spatial triad, and on the other hand, an invariance under the choice of completion of the spatial triad to a tetrad through a timelike vector, to which we refer as "spin gauge invariance". The latter is a gauge of the rotational variables, namely of the tetrad, as well as of the spin.

The point particle action in Eq. 2 for a rotating object can then be written in the following form[9,18,19]:

$$S_{\rm pp} = \int d\sigma \left[ -m\sqrt{u^2} - \frac{1}{2} S_{\mu\nu}\Omega^{\mu\nu} + L_{\rm NMC}\left[u^\mu, S_{\mu\nu}, \bar{g}_{\mu\nu}\left(y^\mu\right)\right] \right], \tag{5}$$

where $u^\mu \equiv dy^\mu/d\sigma$, $S_{\mu\nu} \equiv -2\frac{\partial L}{\partial \Omega^{\mu\nu}}$, and $L_{\rm NMC}$ denotes the non-minimal coupling part, which only contains mass/spin-induced higher order terms. Here, the minimal coupling part is fixed from considering solely general covariance and reparametrization invariance, and it is implicitly assumed that the covariant gauge[21] is used for the rotational variables. Yet, the minimal coupling rotational term should be generalized, in the same vein of Stueckelberg action, to manifestly display the gauge freedom related with the choice of rotational variables[9]. The non-minimal coupling part of the action should also be constrained, using parity invariance and other symmetries and properties of the problem[9].

The minimal coupling rotational term is then transformed as follows[9]:

$$\frac{1}{2} S_{\mu\nu}\Omega^{\mu\nu} = \frac{1}{2}\hat{S}_{\mu\nu}\hat{\Omega}^{\mu\nu} + \frac{\hat{S}^{\mu\nu}p_\nu}{p^2}\frac{Dp_\mu}{D\sigma}, \tag{6}$$

where the hatted variables on the right hand side are the generic rotational variables, $p_\mu$ is the linear momentum, and an extra term with an acceleration emerges in the action. This extra term is not preceded by any Wilson coefficient, as it originates from minimal coupling, yet it contributes to finite size effects, and just accounts for the fact that a gravitating spinning object must have some finite size. Moreover, the leading order (LO) spin-induced non-minimal worldline couplings are recovered to all orders in the spin, and can be written in the following concise form[9]:

$$\begin{aligned} L_{\rm NMC} = &\sum_{n=1}^{\infty} \frac{(-1)^n}{(2n)!}\frac{C_{ES^{2n}}}{m^{2n-1}} D_{\mu_{2n}}\cdots D_{\mu_3}\frac{E_{\mu_1\mu_2}}{\sqrt{u^2}} S^{\mu_1}S^{\mu_2}\cdots S^{\mu_{2n-1}}S^{\mu_{2n}} \\ &+ \sum_{n=1}^{\infty} \frac{(-1)^n}{(2n+1)!}\frac{C_{BS^{2n+1}}}{m^{2n}} D_{\mu_{2n+1}}\cdots D_{\mu_3}\frac{B_{\mu_1\mu_2}}{\sqrt{u^2}} S^{\mu_1}S^{\mu_2}\cdots S^{\mu_{2n-1}}S^{\mu_{2n}}S^{\mu_{2n+1}}, \end{aligned} \tag{7}$$

where new Wilson coefficients precede each of the spin-induced non-minimal coupling terms. These terms are composed from the electric or magnetic curvature components, $E_{\mu\nu}$ or $B_{\mu\nu}$, respectively, and their covariant derivatives, and the spin vector, $S^\mu$, dual to the antisymmetric spin tensor, $S_{\mu\nu}$. Of this infinite set of operators, the first three, namely the quadrupole, octupole, and hexadecapole couplings, contribute up to the 4PN order for rapidly rotating objects[9,10,12].

### 3.2. *A gravitating composite particle with spins*

The obtainment of the EFT at the second stage is in principle more straightforward, as the one-particle EFT is already at our disposal to begin with, and so the computation from the perturbative expansion of the functional integral over the two-particle EFT is expected to be automatic. Yet, when spins are concerned, the situation is less than obvious. First, the field DOFs and the particle DOFs are entangled in the coupling of spin to gravity, e.g., in Eqs. 5, 6. For this to be remedied, it is necessary to also fix the gauge of the rotational variables of the particle at the level of the action of the spinning particles, as was first advocated in Ref. 5. Indeed, in an action approach one can proceed to directly insert the gauge of the rotational variables at any stage. Only once the field DOFs at the orbital scale are cleanly separated, one can proceed to integrate them out, using standard perturbative techniques.

Hence, first we transform to new rotational variables: the worldline Lorentz matrices, contained in the locally flat angular velocity, $\hat{\Omega}^{ab}_{\rm LocFla} = \hat{\Lambda}^{Aa}\frac{d\hat{\Lambda}_A{}^b}{d\sigma}$, and the spin projected to the local frame, $\hat{S}_{ab} = \tilde{e}^\mu_a \tilde{e}^\nu_b \hat{S}_{\mu\nu}$. The minimal coupling term in Eq. 6 can then be rewritten in the following form[6,9]:

$$\frac{1}{2}\hat{S}_{\mu\nu}\hat{\Omega}^{\mu\nu} = \frac{1}{2}\hat{S}_{ab}\hat{\Omega}^{ab}_{\rm LocFla} + \frac{1}{2}\hat{S}_{ab}\omega_\mu{}^{ab}u^\mu, \tag{8}$$

using the Ricci rotation coefficients, $\omega_\mu{}^{ab} \equiv \tilde{e}^b{}_\nu D_\mu \tilde{e}^{a\nu}$, defined from the tetrad field. The gauge of the rotational variables is then fixed to what we formulate as the "canonical" gauge. Other than that, we should also fix the gauge of the field DOFs, where we recall that the Kaluza-Klein (KK) decomposition is beneficially applied to the NR space+time metric[22,23]. Then, the gauge of the tetrad field is fixed to the Schwinger's time gauge, which corresponds to the KK decomposition. Eventually, all unphysical DOFs are eliminated from the action, and one can proceed to compute the EFT using standard QFT multi-loop techniques, such as integration by parts (IBP), and known loop master integrals.

Finally, it was shown how to derive directly from the resulting effective actions the corresponding EOMs and the Hamiltonians.

## 4. Applications and State of the Art

The formulation of the tower of EFTs with spins for the conservative sector, presented in the previous section, built on Refs. 5, 6, 7, 8, which handled the sector

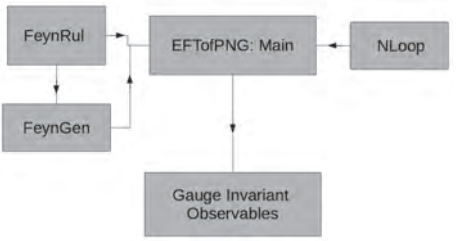



Fig. 1. An outline of the public "EFTofPNG" package version 1.0, which is both a high precision Feynman computation tool and a PN theory code[14]. The design is modular, where the flow among the independent units is shown.

linear in the individual spins, up to the two-loop level, and clarified some crucial points, which led up to the complete framework introduced in Ref. 9. In the latter, all the conservative sector with spins was derived up to the 3PN order for rapidly rotating objects, including the NLO spin-squared sector, which contains spin-induced quadrupole effects, and was shown to agree with the corresponding results derived within the canonical Hamiltonian approach, see e.g., Ref. 24. Moreover, higher order results were derived at the 3.5PN order, for the LO cubic-in-spin sector, which contains up to the spin-induced octupole[10], and for the NNLO spin-orbit sector[11]. Further new higher order results for the quartic-in-spin sector, with up to the spin-induced hexadecapole[10], and for the NNLO spin-squared sector[12], were completed at the 4PN order. All in all, this line of work completed the conservative sector with spins to the 4PN order accuracy, on par with the level of accuracy attained in the non-spinning case[4,16].

### 4.1. *A public Feynman and post-Newtonian code*

The entire framework presented here was also put into a fully open automated computation tool, "EFTofPNG", that is approachable and encouraging engagement within the community[14]. The comprehensive and self-contained "EFTofPNG" code can be found in the public repository in `https://github.com/miche-levi/pncbc-eftofpng`, where the current version handles the point mass sector, and all the spin sectors up to the state of the art. The design of the code is modular, see Fig. 1, so that various units of the code could be efficiently used and developed independently for different purposes in various generic contexts in physics. In particular, since the Mathematica code builds on the "xTensor" package for efficient computer tensor algebra, the treatment of Wick contractions is made significantly more efficient, compared to other common Feynman codes, by considering $n$-point functions as tensors of rank $n$. On the other hand, this is the first code made public in PN theory, and is also geared towards the needs of the GW community.

At this time when the influx of GW data is expected to multiply, it is essential to publicly share the up to date analytic output of PN theory, required for

the improvement of theoretical gravitational waveform templates[25–27]. We strongly advocate public development of the code to improve its efficiency, and to extend it, innovating towards higher orders of PN accuracy, and incorporating extensions of the theory to cover more sectors, especially non-conservative ones. We maintain that community development will also help push a number of eventually necessary research directions. Considering the independent units of the code, as seen in Fig. 1, let us list then just some of the diverse collective prospects for development: The "FeynRul" unit can be expanded to implement various alternative theories of gravity rather than just the theory of GR; The "FeynGen" unit can be modified to efficiently generate Wick contractions and Feynman graphs for any generic perturbative scheme; The "NLoop" unit, which handles and manipulates loop master integrals and IBP relations, useful in any context of amplitudes computations, can be extended analytically to higher loop orders, and to improve in computational efficiency; Finally, the "Gauge Invariant Observables" unit, which is aimed to deliver the final output of PN theory, in terms of useful gauge invariant quantities, is to be progressively developed along with the advances in PN theory, such that these would be directly available to the GW community for the waveform modeling.

## Acknowledgments

I am grateful to John Joseph Carrasco for his pleasant and significant support. I would like to thank Luc Blanchet, who chaired the session, and invited me to present the EFTs of PN theory with spins. I would like to especially acknowledge my collaborator on this line of research, Jan Steinhoff, who also created with me the "EFTofPNG" code. It is with great pleasure that I extend my gratitude to Donato Bini for the warm hospitality throughout the MG15 meeting. My work is supported by the European Research Council under the European Union's Horizon 2020 Framework Programme FP8/2014-2020, "preQFT" grant no. 639729, "Strategic Predictions for Quantum Field Theories" project.

## References

1. W. D. Goldberger and I. Z. Rothstein, *An Effective field theory of gravity for extended objects*, *Phys. Rev.* **D73** (2006) 104029.
2. W. D. Goldberger, *Les Houches lectures on effective field theories and gravitational radiation*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*, 2007.
3. W. D. Goldberger and A. Ross, *Gravitational radiative corrections from effective field theory*, *Phys. Rev.* **D81** (2010) 124015.
4. L. Blanchet, *Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*, *Living Rev. Rel.* **17** (2014).
5. M. Levi, *Next to Leading Order gravitational Spin1-Spin2 coupling with Kaluza-Klein reduction*, *Phys. Rev.* **D82** (2010) 064029.

6. M. Levi, *Next to Leading Order gravitational Spin-Orbit coupling in an Effective Field Theory approach*, *Phys. Rev.* **D82** (2010) 104004.
7. M. Levi, *Binary dynamics from spin1-spin2 coupling at fourth post-Newtonian order*, *Phys. Rev.* **D85** (2012) 064043.
8. M. Levi and J. Steinhoff, *Equivalence of ADM Hamiltonian and Effective Field Theory approaches at next-to-next-to-leading order spin1-spin2 coupling of binary inspirals*, *JCAP* **1412** (2014) 003.
9. M. Levi and J. Steinhoff, *Spinning gravitating objects in the effective field theory in the post-Newtonian scheme*, *JHEP* **09** (2015) 219.
10. M. Levi and J. Steinhoff, *Leading order finite size effects with spins for inspiralling compact binaries*, *JHEP* **06** (2015) 059.
11. M. Levi and J. Steinhoff, *Next-to-next-to-leading order gravitational spin-orbit coupling via the effective field theory for spinning objects in the post-Newtonian scheme*, *JCAP* **1601** (2016) 011.
12. M. Levi and J. Steinhoff, *Next-to-next-to-leading order gravitational spin-squared potential via the effective field theory for spinning objects in the post-Newtonian scheme*, *JCAP* **1601** (2016) 008.
13. M. Levi and J. Steinhoff, `arXiv:1607.04252` *Complete conservative dynamics for inspiralling compact binaries with spins at fourth post-Newtonian order*, 2016.
14. M. Levi and J. Steinhoff, *EFTofPNG: A package for high precision computation with the Effective Field Theory of Post-Newtonian Gravity*, *Class. Quant. Grav.* **34** (2017) 244001.
15. M. Levi, *Effective Field Theory of Post-Newtonian Gravity Including Spins*, in *52nd Rencontres de Moriond on Gravitation (Moriond Gravitation 2017): La Thuile, Italy, March 25-April 1, 2017*, 2017.
16. M. Levi, `arXiv:1807.01699` *Effective Field Theories of Post-Newtonian Gravity: A comprehensive review*, 2018.
17. M. Levi, `arXiv:1901.01282` *Field theory for gravity at all scales*, 2019.
18. A. J. Hanson and T. Regge, *The Relativistic Spherical Top*, *Annals Phys.* **87** (1974) 498.
19. I. Bailey and W. Israel, *Lagrangian Dynamics of Spinning Particles and Polarized Media in General Relativity*, *Commun. Math. Phys.* **42** (1975) 65–82.
20. R. A. Porto, *Post-Newtonian corrections to the motion of spinning bodies in NRGR*, *Phys. Rev.* **D73** (2006) 104031.
21. W. Tulczyjew, *Motion of multipole particles in general relativity theory*, *Acta Phys. Polon.* **18** (1959) 393.
22. B. Kol and M. Smolkin, *Non-Relativistic Gravitation: From Newton to Einstein and Back*, *Class. Quant. Grav.* **25** (2008) 145011.
23. B. Kol, M. Levi, and M. Smolkin, *Comparing space+time decompositions in the post-Newtonian limit*, *Class. Quant. Grav.* **28** (2011) 145021.

24. G. Schäfer and P. Jaranowski, *Hamiltonian formulation of general relativity and post-Newtonian dynamics of compact binaries*, *Living Rev. Rel.* **21** (2018).
25. A. Buonanno and T. Damour, *Effective one-body approach to general relativistic two-body dynamics*, *Phys. Rev.* **D59** (1999) 084006.
26. T. Damour, *The General Relativistic Two Body Problem and the Effective One Body Formalism*, *Fundam. Theor. Phys.* **177** (2014) 111–145.
27. A. Le Tiec, *The Overlap of Numerical Relativity, Perturbation Theory and Post-Newtonian Theory in the Binary Black Hole Problem*, *Int. J. Mod. Phys.* **D23** (2014) 1430022.

# Search for high-energy neutrinos from binary neutron star mergers

Nora Linn Strotjohann* for the IceCube Collaboration

*DESY Zeuthen,*
*15738 Zeuthen, Germany*
**E-mail: nora.linn.strotjohann@desy.de*

To search for transient astrophysical neutrino sources, IceCube's optical and X-ray follow-up program is triggered by two or more neutrino candidates arriving from a similar direction within 100 s. However, the rate of such neutrino multiplets was found to be consistent with the expected background of chance coincidences, such that the data does not provide indications for the existence of short-lived transient neutrino sources. Upper limits on the neutrino flux of transient source populations are presented in Aartsen et al. (2019)[1] and we show here how these limits apply to the predicted neutrino emission from binary neutron star mergers.

*Keywords*: Neutrino astronomy; binary neutron star mergers; short GRBs.

## 1. Introduction

The first (and so far only) detected binary neutron star merger is the object GW 170817[2] with the electromagnetic counterpart GRB 170817A.[3] The detected gamma-ray and multiwavelength emission indicates that a relativistic jet was launched during the merger.[4] It has been suggest that cosmic rays could be accelerated within this jet and an associated flux of high-energy neutrinos has been predicted.[5,6] A search for neutrino emission from the position of GRB 170817A was carried out by the ANTARES, IceCube and Pierre Auger observatories, but no signal was detected.[7]

Here, we present limits on the neutrino emission of short-lived transients,[1] which also apply to binary neutron star mergers and are independent of the limits derived for GRB 170817A. The results are based on data from the IceCube neutrino observatory, an ice-Cherenkov detector located at the geographic South pole.[8] An array of 5 160 optical sensors instruments a volume of $\sim 1\,\mathrm{km}^3$ and allows to detect neutrino interactions in the energy range between $\sim 100\,\mathrm{GeV}$ and a few PeV.[1]

IceCube's optical and X-ray follow-up program (OFU program)[1,9–11] is taylored to search for neutrino emission in connection with gamma-ray bursts (GRBs) or core-collapse supernovae (CCSNe) with choked jets. An alert is defined as two or more tracklike events which are detected within 100 s with a maximal angular separation of 3.5°. The OFU program currently only uses neutrino candidates from the northern sky to avoid the large background of downgoing atmospheric muons. The aim of the program is to identify astrophysical neutrino sources through optical and X-ray follow-up observations, which can be triggered within $\sim 1\,\mathrm{min}$.[12,13] Examples of follow-up campaigns are presented in Aarten et al. (2015)[10], Aartsen et al. (2017)[11], and Evans et al. (2015)[13], but the observations so far did not identify a likely neutrino source and the number of neutrino multiplets is consistent with the expected number of chance coincidences of atmospheric neutrinos or muons.[1]

## 2. Generic Limits on Short-Lived Neutrino Transients

Upper limits on the neutrino flux of short-lived transients have been published in Aartsen et al. (2019)[1] and are summarized in this section. The analysis is based on 1648.1 days of IceCube data (collected between September 2011 and May 2016). Only one alert consisting of three neutrino candidates was detected during this time.[11] This is consistent with the estimated background of 0.34 chance coincidences and the resulting 90% upper limit on the expected number of astrophysical neutrino multiplets (alerts consisting of more than two neutrino candidates) is $< 4.0$.

To calculate a limit on the median neutrino luminosity of short-lived transients, an extragalactic source population is simulated and those scenarios which yield more than 4 astrophysical neutrino multiplets are disfavored at 90% confidence level. The source properties are chosen to be similar to long GRBs or CCSNe and several assumptions are tested to quantify their impact on the resulting limit. The assumed population properties are:

(1) **redshift distribution**: star formation rate for CCSN-like sources[14] and the distribution observed by *Swift* for GRB-like sources.[15]
(2) **neutrino luminosity function**: small fluctuations (one astronomical magnitude) for CCSN-like sources, observed gamma-ray luminosity function for GRB-like sources.[15]
(3) **neutrino spectrum**: an $E^{-2.13}$ or $E^{-2.5}$ spectrum between 100 GeV and 10 PeV as measured for the astrophysical neutrino flux at $\sim 100$ TeV.[16,17]
(4) **transient duration**: duration of long GRBs observed by the *Swift* BAT.[1]

The rate density of transients and the population's neutrino flux are free parameters in the calculation. The resulting upper limits are shown as orange and gray bands in Fig. 1 and correspond to the 90% c.l. upper limits on the energy emitted in neutrinos by the median source in the population. The limit for the GRB-like population is stronger mainly due to the steeply falling luminosity function of long GRBs.[15] The lower edge of the band shows the limit for an $E^{-2.13}$ spectrum, while the upper edge is for an $E^{-2.5}$ spectrum. The dashed diagonal lines show the source energy for which the population would produce the entire astrophysical neutrino flux assuming an $E^{-2.5}$ spectrum. The corresponding lines for an $E^{-2.13}$ spectrum are lower by a factor of 13, but are not shown here for clarity.

Rare sources, such as GRBs with a rate of $\sim 4 \times 10^{-10}\,\mathrm{Mpc}^{-3}\,\mathrm{yr}^{-1}$,[18] can only produce a small fraction of the astrophysical neutrino flux without producing more than 4 neutrino multiplets. This confirms the non-detections from the stacked IceCube searches for neutrinos from GRBs.[19,20] If, on the other hand, many faint sources contribute to the flux, no neutrino multiplets are expected as even the brightest sources only yield one or two detected events. The limits are purely based on neutrino data and hence also apply to short-lived transients that commonly remain undetected, such as binary neutron star mergers or low-luminosity GRBs. They are therefore more general than stacked searches.

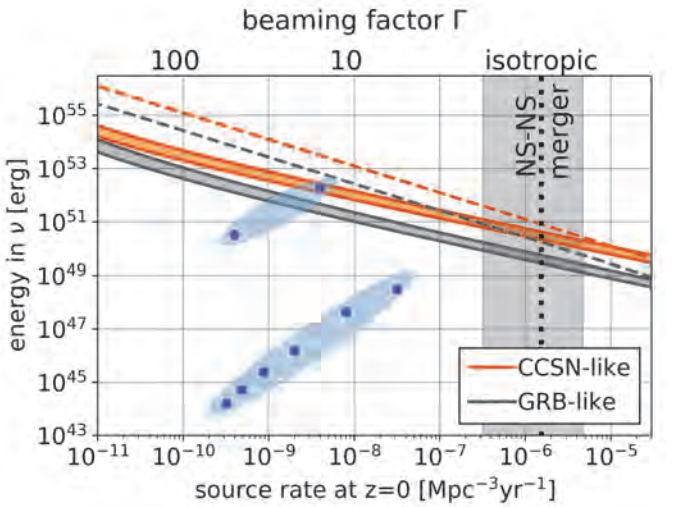



Fig. 1. Upper limits (orange and gray bands) on the median energy emitted by a transient in neutrinos (between 100 GeV and 10 PeV; for three neutrino flavors; compare Ref. 1). The dashed lines correspond to the source energy for which the population would produce the entire astrophysical neutrino flux for an $E^{-2.5}$ spectrum. The blue dots show the predicted high-energy neutrino emission from binary neutron star mergers, where the upper dots are for the Kimura extended emission (EE) model[5] and the lower ones for the Biehl prompt flux model.[6]

## 3. Limits on the Neutrino Flux of Binary Neutron Star Mergers

While the results presented in Aartsen et al. (2019)[1] (and summarized in the previous section) were calculated for a generic population of neutrino sources, the limits are now compared to specific models for the high-energy neutrino emission from binary neutron star mergers. Such models have for example been published by Kimura et al. (2017)[5] and Biehl et al. (2018)[6]. There are considerable uncertainties on the model parameters, such as the beaming factor or the baryonic loading, which result in large differences between the precited neutrino emission shown in Fig. 2. We test several models, listed in Table 1, that predict neutrino emission within 100 s and yield a considerably higher neutrino flux between 100 GeV and 10 PeV.

To quantify the sensitivity of the OFU program to the predicted neutrino fluxes, the spectral shapes are multiplied with the effective area of the OFU event selection. This yields the expected number of neutrino events from a merger in the northern sky at a distance of 40 Mpc which is given in the fourth column of Table 1. The fifth column lists within which distance three or more detected neutrinos are expected from a merger. It is neglected here that an object with fewer expected events might nevertheless produce a neutrino triplet due to Poisson fluctuations.[21]

Contrary to the assumptions for the limit calculation (summarized in Sect. 2), the spectra shown in Fig. 2 do not follow a power law. To allow a comparison an effective source energy is calculated. It corresponds to the energy of a source with an $E^{-2.13}$ spectrum between 100 GeV and 10 PeV for which the same number

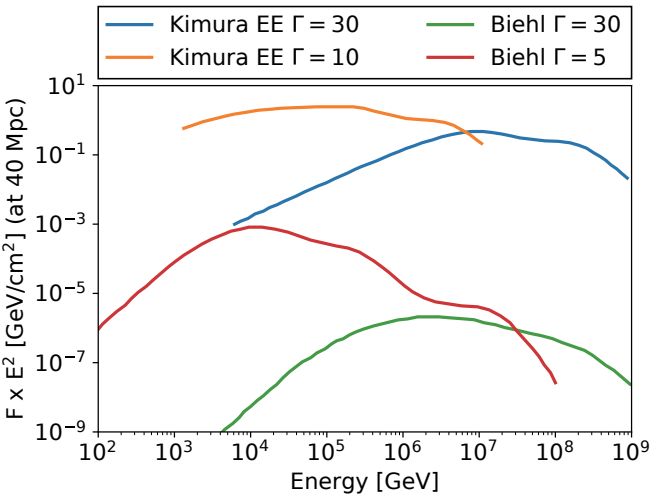



Fig. 2. Predicted neutrino spectra of binary neutron star mergers model according to the Kimura extended emission model[5] and the Biehl prompt emission model[6].

Table 1. Predicted high-energy neutrino emission from binary neutron star mergers. The beaming factor $\Gamma$ determines the rate of detectable transients. The expected number of $\nu_\mu$ events (fourth column) is for the OFU event selection for sources in the northern sky. The next column lists the distance within which three or more events are expected a merger in the northern sky. The two last columns show the energy emitted in neutrinos (for the sum of all three neutrino flavors) between 100 GeV and 10 PeV and the effective source neutrino energy, the energy of a source with an $E^{-2.13}$ spectrum with the same expected number of OFU events.

| Model | $\Gamma$ | rate [$\mathrm{Mpc^{-3}\,yr^{-1}}$] | $\nu_\mu$ events at 40 Mpc | max. dist. [Mpc] | energy in $\nu$ [erg] | eff. energy in $\nu$ [erg] |
|---|---|---|---|---|---|---|
| Kimura EE opt. | 10 | $4 \times 10^{-9}$ | 90 | $< 220$ | $1.2 \times 10^{52}$ | $1.9 \times 10^{52}$ |
| Kimura EE mod. | 30 | $4 \times 10^{-10}$ | 1.5 | $< 30$ | $8 \times 10^{50}$ | $3 \times 10^{50}$ |
| Biehl optimistic | 5 | $3 \times 10^{-8}$ | 0.02 | $< 10$ | $1.6 \times 10^{48}$ | $3 \times 10^{48}$ |
| Biehl moderate | 30 | $1 \times 10^{-9}$ | $1 \times 10^{-5}$ | $< 0.1$ | $4 \times 10^{45}$ | $2 \times 10^{45}$ |

of events is expected. The effective source energy is given in the last column of Table 1. The actual source energy is calculated by integrating the spectra in Fig. 2 and is listed in the second last column of the table. The difference is less than a factor of three for all considered neutrino spectra.

The flux upper limits moreover depend on the rate density of transients at $z = 0$. As shown in Fig. 1, the limits get stronger for larger rates, because it becomes more likely that a source happens within the distance where it is detectable (listed in the fifth column of Table 1). The rate of binary neutron star mergers was measured based on the detected event and the gray, vertical band in Fig. 1 shows the rate within its $1\sigma$ uncertainty. The considered models, however, predict collimated neutrino emission which reduces the rate of observable events by a factor of $1/(2\Gamma)$. The collimation factors and resulting rates are listed in the second and third column

of Table 1. Extended gamma-ray emission is moreover only expected for every second short GRB[5] (and was not detected for GRB 170817A) which reduces the rate accordingly. Potential neutrino emission from misaligned mergers[5,6] is not considered for simplicity. The redshift distribution only has a minor influence on the upper limits. When using the observed evolution of short GRBs[22] the limits become 13% stronger compared to the star formation rate used for the CCSN-like population (see appendix of Aartsen et al. (2019)[1]).

The predicted rates and the source energies for binary neutron star mergers are shown as blue dots in Fig. 1. The upper dots are for the Kimura EE model[5] and the lower ones for the Biehl prompt emission model[6]. The Kimura prompt emission model is not shown: It assumes a beaming factor of 1000 and has an effective neutrino energy of $5 \times 10^{43}$ erg.[5] Figure 1 shows that the optimistic EE model is within reach of the OFU program. If the luminosity fluctuations between individual mergers are as large as the fluctuations observed for long GRBs in gamma rays, the lower edge of the gray band represents the relevant upper limit and the model would be disfavored at 90% confidence limit. The models for prompt neutrino emission, however, predict fluxes and rates that are many orders of magnitude below the IceCube sensitivity.

## 4. Discussion

The observed gamma-ray emission of GRB 170817A indicates the presence of a relativistic jet which might be able to accelerate cosmic rays and yield a neutrino flux. No neutrino emission was detected from GRB 170817A which is expected since the jet was likely misaligned by 30°.[4] Moreover, no extended gamma-ray emission was observed from the central engine.[3] An additional difficulty was that the source was located above the horizon for both the ANTARES and the IceCube detectors which reduces the sensitivity.[7]

Independent upper limits on the neutrino emission of 100 s-long transients are provided by IceCube's OFU program. The sensitivity of the OFU program to a merger located at 40 Mpc was evaluated for several models and the results are listed in Table 1. For the most optimistic model the OFU program would likely be triggered if the merger happens within $\sim$ 220 Mpc in the northern sky. To estimate how likely a nearby merger occurs during the $\sim$ 5 years of analyzed data, we calculate the rate of mergers which have a jet pointed at Earth. The most optimistic model predicts a relatively large rate of bright sources, such that one or several neutrino multiplets would have been detected by the OFU program.

Most prompt models, however, predict that the flux from the population of binary neutron star mergers only makes up a small fraction of the astrophysical neutrino flux. In this case, IceCube might not be sensitive enough to detect the neutrino emission from a merger. With only one detected binary neutron star merger, the uncertainties on the physical properties of these objects are still large. Models might become more precise, once more observational information is available.

Figure 2 shows that the predicted neutrino flux of several models peaks at PeV energies. Since April 2016, IceCube announces events with energies $\gtrsim 100\,\mathrm{TeV}$ to the public.[12] Follow-up observations of these alerts could hence provide an alternative way to find binary neutron star mergers. The source of a single event is, however, on average located at a large distance (the median distance is 3 Gpc assuming no source evolution)[21]. The horizon distance of the Ligo-Virgo detector was $\sim 200\,\mathrm{Mpc}$ in run O2 and might increase to 300 Mpc in run O3.[2,23] The optical peak magnitude of GRB 170817A was $-15.5$ in the $R$-band.[3] This would allow a detection within 140 Mpc (300 Mpc) by a telescope with a limiting magnitude of 20 (22). If a single PeV neutrino is detected from a binary neutron star merger, the source is likely in most cases not detectable by gravitational wave detectors or electromagnetic follow-up telescopes.

**Acknowledgment**

I would like to thank Marek Kowalski, Anna Franckowiak, Vladimir Lipunov, Markus Ackermann, Daniel Biehl, Shigeo Kimura, Ignacio Taboada, and the IceCube collaboration for scientific discussions, support and inspiration.

**References**

1. M. G. Aartsen et al., *Phys. Rev. Lett.* **122**, 051102 (2019)
2. B. P. Abbott et al., *Phys. Rev. Lett.* **119**, 16 (2017)
3. M. M. Kasliwal et al., *Science* **358**, 1559 (2017)
4. D. Lazzati et al., *Phys. Rev. Lett.* **120**, 24 (2018)
5. S. S. Kimura, K. Murase, P. Mészáros, and K. Kiuchi, *ApJL* **848**, L4 (2017)
6. D. Biehl, J. Heinze and W. Winter *Mon. Not. Roy. Ast. Soc.* **476**, 1191 (2018)
7. A. Albert et al., *ApJL* **850**, L35 (2017)
8. M. G. Aartsen et al., *Journal of Instrumentation* **12**, P03012 (2017)
9. R. Abbasi et al., *A&A* **539**, A60 (2012)
10. M. G. Aartsen et al., *ApJ* **811**, 52 (2015)
11. M. G. Aartsen et al., *A&A* **607**, A115 (2017)
12. M. G. Aartsen et al., *Astroparticle Physics* **92**, 30 (2017)
13. P. A. Evans et al., *Mon. Not. Roy. Ast. Soc.* **448**, 2210 (2015)
14. P. Madau and M. Dickinson, *Mon. Not. Roy. Ast. Soc.* **52**, 415 (2014)
15. D. Wanderman and T. Piran, *Mon. Not. Roy. Ast. Soc.* **406**, 1944 (2010)
16. M. G. Aartsen et al., *ApJ* **833**, 3 (2016)
17. M. G. Aartsen et al., *ApJ* **809**, 98 (2015)
18. A. Lien et al., *ApJ* **783**, 24 (2014)
19. M. G. Aartsen et al., *ApJ* **843**, 112 (2017)
20. M. G. Aartsen et al., *ApJ* **824**, 115 (2016)
21. N. L. Strotjohann, M. Kowalski and A. Franckowiak, *A&A* **622**, L9 (2019)
22. D. Wanderman and T. Piran, *Mon. Not. Roy. Ast. Soc.* **448**, 3026 (2015)
23. H.-Y. Chen et al., *ArXiv e-prints* 1709.08079 (2017)

# Hyperbolic scattering in a two-body system

Donato Bini* and Andrea Geralico**

*Istituto per le Applicazioni del Calcolo "M. Picone," CNR I-00185 Rome, Italy*
**E-mail: donato.bini@gmail.com*
***E-mail: andrea.geralico@gmail.com*

We review some known results obtained in the context of hyperbolic scattering in a two-body system, stressing the state of the art, the main actual challenges and the related difficulties.

## 1. Introduction

The gravitational interaction between two (massive, spinning or with a 2−polar structure, . . . , with $n$−polar structure) bodies is a fundamental problem which has been largely investigated in many different contexts, all aiming at a complete description of the underlying gravitational phenomena but limited by the range of validity of the considered modeling used in the description of the interaction itself. The problem should be discussed in exact general relativity (GR), allowing for all the necessary information concerning the (most important) strong field regime and the effects of the nonlinearity of the gravitational interaction as follows from the Einstein's equations. Unfortunately, no such exact solutions are available up to now in a genuine GR context and one is forced to cope with some approximate treatment, ranging from weak-field and slow-motion (Post-Newtonian (PN) approximation[1]), weak-field (Post-Minkowskian (PM) approximation[2,3]), extreme-mass-ratios (i.e., the condition in which the mass of one of the two bodies is much larger of the other, discussed with perturbation theories and gravitational self-force (GSF)[4]), numerical relativity (NR)[5]. In between, there are various effective field theory formalism, like the effective-one body[6,7], which have shown over the years the capability to simply absorb information from any such context and re-use properly and efficiently.

Given an initial state configuration of the two bodies, the attractive nature of gravitation imply that they necessarily approach each other and two possible scenarios are available: capture (the more massive of the two attracts and then swallows the other) or escape (the two bodies can be close enough but they have enough energy to resist the attraction). The first situation is characterized by elliptic-like motions, the second by hyperbolic-like or parabolic-like motions. Clearly, the description becomes more and more difficult as soon as the strong field regime is reached and any approximation scheme becomes of poor utility, if not even useless.

Anyway, both cases are of particular importance in view of the possibility to detect gravitational wave signals from Earth-based interferometers operating now (Ligo[8] and Virgo[9]) and also (more likely) by future, forthcoming satellite missions involving space-based interferometers.

Since elliptic-like motions have been largely studied, we will review here the scattering situation, focusing on the known results for the scattering angle, and witnessing an increasing interest on this topic in the recent literature.

## 2. Scattering of a test particle by a black hole

Let us consider first the simplest case of a test particle scattered by a black hole, e.g., a Kerr black hole with line element written in standard Boyer-Lindquist coordinates $(t, r, \theta, \phi)$ given by

$$ds^2 = -\left(1 - \frac{2Mr}{\Sigma}\right) dt^2 - \frac{4aMr}{\Sigma} \sin^2\theta dt d\phi + \frac{\Sigma}{\Delta} dr^2 + \Sigma d\theta^2 + \frac{A}{\Sigma} \sin^2\theta d\phi^2 \,. \quad (1)$$

Here $\Delta = r^2 - 2Mr + a^2$, $\Sigma = r^2 + a^2\cos^2\theta$ and $A = (r^2 + a^2)^2 - \Delta a^2 \sin^2\theta$, $M$ and $a$ denoting the total mass and the specific angular momentum, respectively, so that the quantity $\hat{a} = a/M$ is dimensionless. The inner and outer horizons are located at $r_\pm = M \pm \sqrt{M^2 - a^2}$.

The geodesic equations on the equatorial plane lead to the following equation for the orbit

$$\left(\frac{du}{d\phi}\right)^2 = \frac{\hat{a}^4}{2} \frac{(u - u_+)^2 (u - u_-)^2}{(u_4 - u)^2} (u - u_1)(u - u_2)(u - u_3) \,, \quad (2)$$

where $u = M/r$ and $u_1 < u_2 < u_3$ are the ordered roots of the equation

$$u^3 - (\hat{x}^2 + 2\hat{a}\hat{E}\hat{x} + \hat{a}^2) \frac{u^2}{2\hat{x}^2} + \frac{u}{\hat{x}^2} + \frac{\hat{E}^2 - 1}{2\hat{x}^2} = 0 \,, \quad (3)$$

the constants $\hat{E}$ and $\hat{L}$ denoting the conserved energy and azimuthal angular momentum per unit mass of the particle, respectively, with $x = M\hat{x} = \hat{L} - a\hat{E}$, while

$$u_\pm = \frac{M}{r_\pm} \,, \quad u_4 = \frac{\hat{L}}{2M\hat{x}} = \frac{1}{2}\left(1 + \frac{\hat{a}\hat{E}}{\hat{x}}\right) . \quad (4)$$

For hyperbolic orbits we have $u_1 < 0 < u \leq u_2 < u_3$, with $u_2$ corresponding to the closest approach distance.

The ultrarelativistic limit ($\hat{E} \gg 1$) of Eq. (2) has been studied in Ref. 10 for fixed values of the impact parameter $b = M\hat{b} = \hat{L}/\sqrt{\hat{E}^2 - 1}$. The deflection angle $\delta(\hat{a}, \alpha) = 2\phi(0) - \pi$ associated with the whole scattering process has been computed there. For large values of the impact parameter, i.e., for small values of $\alpha = \hat{b}^{-1}$, the first few terms of its expansion are given by

$$\begin{aligned} \delta(\hat{a}, \alpha) = {} & 4\alpha + \left(\frac{15}{4}\pi - 4\hat{a}\right)\alpha^2 + \left(\frac{128}{3} - 10\pi\hat{a} + 4\hat{a}^2\right)\alpha^3 \\ & + \left(\frac{3465}{64}\pi - 192\hat{a} + \frac{285}{16}\pi\hat{a}^2 - 4\hat{a}^3\right)\alpha^4 \\ & + \left(\frac{3584}{5} - \frac{693}{2}\pi\hat{a} + 512\hat{a}^2 - 27\pi\hat{a}^3 + 4\hat{a}^4\right)\alpha^5 + O(\alpha^6) \,. \end{aligned} \quad (5)$$

Ref. 10 has also compared this result for a test particle without structure to the high-energy scattering of an extended body with spin-induced quadrupolar structure (characterized by the dimensionless spin parameter $\hat{s}$ and the polarizability constant $C_Q$) by a Kerr black hole as well as of a pointlike particle in the field of a (slowly) rotating (slightly) deformed source endowed with a mass quadrupole moment (characterized by the dimensionless intrinsic angular momentum $\hat{a}$ and the quadrupole parameter $q$), as described by the (approximate) Hartle-Thorne solution[11]. In all these situations the scattering angle can be formally written as

$$\delta = \delta_{\rm geo} + \delta_{\rm struct}\,, \tag{6}$$

where $\delta_{\rm geo}(q,\hat{a},\alpha)$ contains information on the background source only, whereas

$$\delta_{\rm struct}(\hat{s},C_Q;\hat{a},\alpha) = \delta_{\rm spin}(\hat{s};\hat{a},\alpha) + \delta_{\rm spin^2}(\hat{s}^2,C_Q;\hat{a},\alpha)\,, \tag{7}$$

also depends on the constitutive parameters of the scattered extended body. The first few terms of the corresponding expansion are given by

$$\begin{aligned}
\delta_{\rm spin}(\hat{s};\hat{a},\alpha) &= \hat{s}\left[-4\alpha^2 + \left(8\hat{a} - \frac{15}{2}\pi\right)\alpha^3 \right.\\
&\left. +(-128 + 30\pi\hat{a} - 12\hat{a}^2)\alpha^4 + O(\alpha^5)\right],\\
\delta_{\rm spin^2}(\hat{s}^2,C_Q;\hat{a},\alpha) &= \hat{s}^2\left\{4C_Q\alpha^3 + \left[\frac{45}{4}\pi - 12\hat{a}\right.\right.\\
&\left.\left. + \left(\frac{135}{16}\pi - 12\hat{a}\right)(C_Q-1)\right]\alpha^4 + O(\alpha^5)\right\}.
\end{aligned} \tag{8}$$

Therefore, if the spin is known, measuring $\delta_{\rm struct}$ can be used for instance to constrain the equation of state of the extended body undergoing the scattering process.

## 3. Two-body scattering in a Post-Newtonian framework

The present knowledge of the PN orbital Hamiltonian for a two-body system is at the 4th PN order, with several interesting new features with respect to the previous 3PN level which include the presence of a non-local term. Up to the 4PN level the scattering angle is given by the sum of a local part $\chi_{\rm loc}$ and a tail part $\chi_{\rm tail}$, the latter starting at the 4PN level, i.e.,

$$\begin{aligned}
\chi_{\rm loc}(E,L) &= \sum_{n=0}^{4}\frac{1}{c^{2n}}\chi^{(\rm nPN)}(E,L) + O\left(\frac{1}{c^{10}}\right),\\
\chi_{\rm tail}(E,L) &= \frac{1}{c^8}\chi^{(\rm 4PN)}_{\rm tail}(E,L) + O\left(\frac{1}{c^{10}}\right).
\end{aligned} \tag{9}$$

To show some explicit expression (see Ref. 12 for additional details) it is convenient to introduce the notation

$$\alpha \equiv \frac{1}{\sqrt{2\bar{E}j^2}}\,, \qquad B(\alpha) \equiv \arctan\alpha + \frac{\pi}{2}\,, \tag{10}$$

where $j$ is the dimensionless angular momentum $L$ and $\bar{E}$ is a Newtonian-like energy variable (see Eq. (32) in Ref. 12). A part from the Newtonian contribution

$$\frac{1}{2}\chi^{(\mathrm{N})}(\bar{E},j) = \arctan\alpha = B(\alpha) - \frac{\pi}{2}\,, \tag{11}$$

the form of the various PN terms $\chi^{(\mathrm{nPN})}$ (including the local part of the 4PN term) is of the type

$$\frac{1}{2}\chi^{(\mathrm{nPN})}(\bar{E},j) = \frac{1}{j^{2n}}\left[C^{B}_{(\mathrm{nPN})}B(\alpha) + C^{0}_{(\mathrm{nPN})}\right] \tag{12}$$

with the coefficients $C^{B}_{(\mathrm{nPN})}$ and $C^{0}_{(\mathrm{nPN})}$ rational functions of $\alpha$. For example

$$\begin{aligned}
\frac{1}{2}\chi^{(1\mathrm{PN})}(\bar{E},j) &= \frac{1}{j^{2}}\left[3B(\alpha) + \frac{(3\alpha^2+2)}{\alpha(1+\alpha^2)}\right] \\
\frac{1}{2}\chi^{(2\mathrm{PN})}(\bar{E},j) &= \frac{1}{j^{4}}\left[C^{B}_{(2\mathrm{PN})}B(\alpha) + C^{0}_{(2\mathrm{PN})}\right]\,,
\end{aligned} \tag{13}$$

with

$$\begin{aligned}
C^{B}_{(2\mathrm{PN})} &= -\frac{3[2(5\alpha^2+1)\nu - 5(1+7\alpha^2)]}{4\alpha^2}\,, \\
C^{0}_{(2\mathrm{PN})} &= -\frac{[2(15\alpha^4+28\alpha^2+13)\nu - 105\alpha^4 - 190\alpha^2 - 81]}{4\alpha(1+\alpha^2)^2}\,.
\end{aligned} \tag{14}$$

In $\chi^{(4\mathrm{PN})}_{\mathrm{loc}}(\bar{E},L)$, however, it enters the following (Hadamard-regularized) integral

$$I_\chi \equiv -\frac{16j\nu}{15}\,\mathrm{Pf}\int_0^{u_{(\max)}} \frac{u^4\ln(u)(-74\bar{E}+37j^2u^2-62u)}{(2\bar{E}-j^2u^2+2u)^{3/2}}du\,, \tag{15}$$

whose value can be anyway given in an approximated form, $u_{(\max)}$ being the (Newtonian) minimum approach dimensionless inverse distance.

## 4. Two-body scattering in a Post-Minkowskian framework

The scattering angle is known since much time at 2PM[3,13,14]

$$\chi = \frac{2(2\gamma^2-1)}{j(\gamma^2-1)^{1/2}} + \frac{3}{4}\pi\frac{(5\gamma^2-1)}{j^2h}\,, \tag{16}$$

where $\gamma$ denotes the rescaled effective energy of the system $\gamma = \frac{E-Mc^2}{\mu c^2}$ and

$$h = \sqrt{1+2\nu(\gamma-1)}\,. \tag{17}$$

The connection between the ultrarelativistic limit of this result and the ultra high-energy quantum scattering results of Amati, Ciafaloni and Veneziano[15,16] has been recently exploited in Ref. 17. It has been shown there that the scattering process between the two (nonrotating) bodies is equivalently described in terms of the scattering of a massless particle in a modified Schwarzschild metric of the form

$$ds^2 = -f_t dt^2 + f_r^{-1}dr^2 + r^2(d\theta^2 + \sin^2\theta d\phi^2)\,, \tag{18}$$

with

$$f_t(u) = (1-2u)\left(1 + \frac{15}{2}u^2 - 18u^3 + \frac{1845}{16}u^4 + \dots\right),$$
$$f_r(u) = (1-2u)\,. \tag{19}$$

The scattering angle then turns out to be

$$\delta(\alpha) = 4\alpha + \frac{56}{3}\alpha^3 + O(\alpha^5)\,, \tag{20}$$

with vanishing both the coefficients at $O(\alpha^2)$ and $O(\alpha^4)$.

Very recently, after the MG meeting, in the literature has appeared the 3PM scattering angle, derived straightforwardly from re-summation of scattering amplitudes and quantum field theory approaches[18,19]. The high-energy limit of the latter result does not seem to reproduce previous results by Amati, Ciafaloni and Veneziano, even if both approaches use rather involved techniques, making difficult independent checks. The ongoing discussion will hopefully clarify this issue soon.

## 5. Difficulties for a GSF approach

The main difficulty in a GSF approach is related to the spectrum of radiation emitted. In fact, differently from the case of a particle orbiting a black hole moving along a circular trajectory for which the spectrum is peaked around the circular motion frequency, in the case of a hyperbolic orbits the spectrum is infinite (all frequencies are excited) and this complicates matters.

## References


1. L. Blanchet, "Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries," Living Rev. Rel. **17**, 2 (2014) doi:10.12942/lrr-2014-2 [arXiv:1310.1528 [gr-qc]].
2. L. Bel, T. Damour, N. Deruelle, J. Ibanez and J. Martin, "Poincaré-invariant gravitational field and equations of motion of two pointlike objects: The post-linear approximation of general relativity," Gen. Rel. Grav. **13**, 963 (1981). doi:10.1007/BF00756073
3. K. Westpfahl, "High-Speed Scattering of Charged and Uncharged Particles in General Relativity," Fortsch. Phys. **33**, 417 (1985). doi:10.1002/prop.2190330802
4. S. L. Detweiler, "Perspective on gravitational self-force analyses," Class. Quant. Grav. **22**, S681 (2005) doi:10.1088/0264-9381/22/15/006 [gr-qc/0501004].
5. F. Pretorius, "Numerical relativity using a generalized harmonic decomposition," Class. Quant. Grav. **22**, 425 (2005) doi:10.1088/0264-9381/22/2/014 [gr-qc/0407110].
6. A. Buonanno and T. Damour, "Effective one-body approach to general relativistic two-body dynamics," Phys. Rev. D **59**, 084006 (1999) [gr-qc/9811091].

7. A. Buonanno and T. Damour, "Transition from inspiral to plunge in binary black hole coalescences," Phys. Rev. D **62**, 064015 (2000) [gr-qc/0001013].
8. See the LIGO website at http://www.ligo.org
9. See the VIRGO website at http://www.virgo-gw.eu
10. D. Bini and A. Geralico, "High-energy hyperbolic scattering by neutron stars and black holes," Phys. Rev. D **98**, no. 2, 024049 (2018) doi:10.1103/PhysRevD.98.024049 [arXiv:1806.02085 [gr-qc]].
11. J. B. Hartle and K. S. Thorne, "Slowly Rotating Relativistic Stars. II. Models for Neutron Stars and Supermassive Stars," Astrophys. J. **153**, 807 (1968). doi:10.1086/149707
12. D. Bini and T. Damour, "Gravitational scattering of two black holes at the fourth post-Newtonian approximation," Phys. Rev. D **96**, no. 6, 064021 (2017) doi:10.1103/PhysRevD.96.064021 [arXiv:1706.06877 [gr-qc]].
13. T. Damour, "Gravitational scattering, post-Minkowskian approximation and Effective One-Body theory," Phys. Rev. D **94**, no. 10, 104015 (2016) doi:10.1103/PhysRevD.94.104015 [arXiv:1609.00354 [gr-qc]].
14. D. Bini and T. Damour, "Gravitational spin-orbit coupling in binary systems at the second post-Minkowskian approximation," Phys. Rev. D **98**, no. 4, 044036 (2018) doi:10.1103/PhysRevD.98.044036 [arXiv:1805.10809 [gr-qc]].
15. D. Amati, M. Ciafaloni and G. Veneziano, "Superstring Collisions at Planckian Energies," Phys. Lett. B **197**, 81 (1987). doi:10.1016/0370-2693(87)90346-7
16. D. Amati, M. Ciafaloni and G. Veneziano, "Classical and Quantum Gravity Effects from Planckian Energy Superstring Collisions," Int. J. Mod. Phys. A **3**, 1615 (1988). doi:10.1142/S0217751X88000710
17. T. Damour, "High-energy gravitational scattering and the general relativistic two-body problem," Phys. Rev. D **97**, no. 4, 044038 (2018) doi:10.1103/PhysRevD.97.044038 [arXiv:1710.10599 [gr-qc]].
18. Z. Bern, C. Cheung, R. Roiban, C. H. Shen, M. P. Solon and M. Zeng, "Scattering Amplitudes and the Conservative Hamiltonian for Binary Systems at Third Post-Minkowskian Order," Phys. Rev. Lett. **122**, no. 20, 201603 (2019) doi:10.1103/PhysRevLett.122.201603 [arXiv:1901.04424 [hep-th]].
19. A. Antonelli, A. Buonanno, J. Steinhoff, M. van de Meent and J. Vines, "Energetics of two-body Hamiltonians in post-Minkowskian gravity," Phys. Rev. D **99**, no. 10, 104004 (2019) doi:10.1103/PhysRevD.99.104004 [arXiv:1901.07102 [gr-qc]].

# Extended bodies in perturbed black hole spacetimes: the gravitational self-force approach

Andrea Geralico* and Donato Bini†

*Istituto per le Applicazioni del Calcolo "M. Picone," CNR, Rome, I-00185, Italy*

**E-mail: andrea.geralico@gmail.com*

*†E-mail: donato.bini@gmail.com*

We review the necessary framework to study self-force effects induced by an extended body endowed with internal structure up to the quadrupole moving in a Schwarzschild black hole spacetime. The motion is described according to the Mathisson-Papapetrou-Dixon model. The metric perturbations are computed by using the Regge-Wheeler gauge. We apply this formalism to the case of Detweiler's redshift invariant for spinning particles along circular equatorial orbits.

*Keywords*: Black hole perturbations; extended bodies; gravitational self-force.

## 1. Introduction

After the detection of the first gravitational wave signal emitted by a binary black hole system,[1] further improving our analytical modelling of the general relativistic two-body gravitational interaction has become mandatory. Many methods have been actively pursued over the years, including post-Newtonian (PN) theory,[2] gravitational self-force (GSF) theory[3] and the Effective-One-Body (EOB) formalism.[4–6] The synergy between all these approaches is currently successfully used to construct semi-analytical templates which are able to describe the complete waveform emitted by coalescing binary black holes.

GSF theory applies when the mass ratio of the two bodies is very small, so that the whole dynamics can be treated as a perturbation induced by the smaller body with mass $m_1$ on the background gravitational field of the bigger one with mass $m_2$. The goal is to compute gauge-invariant quantities, like redshift and spin precession invariants, which can be eventually converted into useful information for other approaches, e.g., the EOB model. Many analytical results have been obtained in the last few years for a particle moving along circular as well as slightly eccentric equatorial orbits around both non-rotating (Schwarzschild) and rotating (Kerr) black holes (see, e.g., Ref. 7 for a recent review). To first order in the mass ratio, the small body is considered as following a geodesic orbit in a suitably regularized perturbed spacetime. However, if the small body has a non-negligible internal structure (e.g., spin, quadrupole moment and higher multipoles) the orbit will be no longer geodesic, but accelerated due to the coupling between curvature and multipolar structure of the body as described by the Mathisson-Papapetrou-Dixon (MPD) model.[8–10] The first GSF computations have been recently performed in the case of redshift invariant and tidal invariants for spinning particles orbiting a Schwarzschild black hole, to linear order in spin.[11,12]

## 2. Motion of extended bodies in a given background spacetime

The MPD equations governing the motion of an extended body endowed with a quadrupole moment in a given gravitational background read

$$\begin{aligned}\frac{DP^\mu}{d\tau} &= -\frac{1}{2}R^\mu{}_{\nu\alpha\beta}U^\nu S^{\alpha\beta} - \frac{1}{6}J^{\alpha\beta\gamma\delta}\nabla^\mu R_{\alpha\beta\gamma\delta}\\ \frac{DS^{\mu\nu}}{d\tau} &= 2P^{[\mu}U^{\nu]} + \frac{4}{3}J^{\alpha\beta\gamma[\mu}R^{\nu]}{}_{\gamma\alpha\beta}\,,\end{aligned} \tag{1}$$

where $P = mu$, with $u \cdot u = -1$ and $||P||^2 = -m^2$, is the (timelike) generalized 4-momentum of the body with effective mass $m$ (which does not coincide with the bare mass $\mu$, but depends on its structure) and $U^\alpha = dx^\alpha/d\tau$ is the (timelike) unit tangent vector to the particle's world line, $\mathcal{C}$, used to make the multipole reduction. The two four vectors $U$ and $u$ are in general not aligned, and $P$ (i.e., $u$) has support only along $\mathcal{C}$. The (antisymmetric) spin tensor $S^{\mu\nu}$ (with support only along $\mathcal{C}$, like $P$) is assumed to satisfy the supplementary conditions

$$S^{\mu\nu}u_\nu = 0\,. \tag{2}$$

As standard, the spin vector (orthogonal to $u$) associated with the spin tensor is given by

$$S(u)^\alpha = \frac{1}{2}\eta(u)^{\alpha\beta\gamma}S_{\beta\gamma}\,, \qquad \eta(u)_{\alpha\beta\gamma} = u^\mu\eta_{\mu\alpha\beta\gamma}\,. \tag{3}$$

Its magnitude $s$ is such that

$$s^2 = S(u)\cdot S(u) = \frac{1}{2}S_{\mu\nu}S^{\mu\nu} = -\frac{1}{2}\mathrm{Tr}[S^2]\,, \qquad [S^2]^\alpha{}_\beta = S^{\alpha\mu}S_{\mu\beta}\,. \tag{4}$$

Finally, $J^{\alpha\beta\gamma\delta}$ is the quadrupole tensor of the particle, again with support only along $\mathcal{C}$. For instance, in the case of a spin-induced quadrupole tensor we have[13]

$$J^{\alpha\beta\gamma\delta} = 4u^{[\alpha}\tilde{\chi}(u)^{\beta][\gamma}u^{\delta]}\,, \qquad \tilde{\chi}(u) = \frac{3}{4}\frac{C_Q}{m}[S^2]^{\mathrm{STF}}\,, \tag{5}$$

where

$$[S^2]^{\mathrm{STF}\,\alpha\beta} = [S^2]^{\alpha\beta} - \frac{1}{3}P(u)^{\alpha\beta}\mathrm{Tr}[S^2] = [S(u)\otimes S(u)]^{\mathrm{TF}}\,, \tag{6}$$

$P(u) = g + u\otimes u$ projecting orthogonally to $u$, represents the symmetric-tracefree part (STF) of the spin tensor (or, equivalently, of the spin vector). The constant parameter $C_Q$ is equal to unity in the case of a black hole, whereas for neutron stars it depends on the equation of state.

When the background spacetime has Killing vectors, there are conserved quantities along the motion. For example, in the case of stationary axisymmetric spacetimes with coordinates adapted to the spacetime symmetries, $\xi = \partial_t$ is the timelike Killing vector and $\eta = \partial_\phi$ is the azimuthal Killing vector. The corresponding conserved quantities are the energy $E$ and the angular momentum $J$ of the particle,

namely

$$E = -\xi_\alpha P^\alpha + \frac{1}{2} S^{\alpha\beta} \nabla_\beta \xi_\alpha \,,$$
$$J = \eta_\alpha P^\alpha - \frac{1}{2} S^{\alpha\beta} \nabla_\beta \eta_\alpha \,. \tag{7}$$

Finally, the energy momentum tensor of a quadrupolar particle is given by

$$T^{\alpha\beta} = \int d\tau \frac{1}{\sqrt{-g}} \mathcal{T}^{\alpha\beta} \,, \tag{8}$$

where

$$\mathcal{T}^{\alpha\beta} = \left( U^{(\alpha} P^{\beta)} + \frac{1}{3} R_{\gamma\delta\epsilon}{}^{(\alpha} J^{\beta)\epsilon\delta\gamma} \right) \delta^4 - \nabla_\gamma \left( S^{\gamma(\alpha} U^{\beta)} \delta^4 \right)$$
$$- \frac{2}{3} \nabla_\delta \nabla_\gamma \left( J^{\delta(\alpha\beta)\gamma} \delta^4 \right) \,, \tag{9}$$

$\delta^4$ denoting the 4-dimensional delta function centered on the particle's worldline, i.e., $\delta^4 \equiv \delta^4(x^\alpha - x^\alpha(\tau))$.

## 3. GSF program

Let $g^{\rm R}_{\alpha\beta} = \bar{g}_{\alpha\beta} + q h^{\rm R}_{\alpha\beta}$ with $q = m_1/m_2 \ll 1$ be the regularized perturbed metric, where

$$\bar{g}_{\alpha\beta} dx^\alpha dx^\beta = -f dt^2 + f^{-1} dr^2 + r^2 (d\theta^2 + \sin^2\theta d\phi^2) \,, \tag{10}$$

with $f = 1 - 2m_2/r$, is the background Schwarzschild spacetime written in standard coordinates $(t, r, \theta, \phi)$, and $h^{\rm R}_{\alpha\beta}$ is the regularized metric perturbation due to the extended body. The main steps to analytically compute first-order metric perturbations in a Schwarzschild spacetime by using the Regge-Wheeler-Zerilli (RWZ) approach[14,15] are recalled below.

Decompose first both the perturbed metric $h_{\mu\nu}$ and the energy momentum tensor $T_{\mu\nu}$ associated with the perturbing body in tensor harmonics in the RW gauge. Einstein's field equations then lead to two different sets of equations for even-parity and odd parity quantities. Odd-parity waves satisfy the simple RW equation with odd-parity source terms; even-parity waves satisfy, instead, the more complicated Zerilli equation with even-parity source terms. However, the Zerilli equation can be mapped onto a RW equation with different source terms, so that one eventually has to solve a single RW equation, with appropriate (odd-parity or even-parity) source terms, i.e.,

$$\mathcal{L}^{(r)}_{\rm (RW)}[R^{\rm (even/odd)}_{lm\omega}] = S^{\rm (even/odd)}_{lm\omega}(r) \,. \tag{11}$$

Here $\mathcal{L}^{(r)}_{\rm (RW)}$ denotes the RW operator

$$\mathcal{L}^{(r)}_{\rm (RW)} = f(r)^2 \frac{d^2}{dr^2} + \frac{2m_2}{r^2} f(r) \frac{d}{dr} + [\omega^2 - V_{\rm (RW)}(r)]$$
$$= \frac{d^2}{dr_*^2} + [\omega^2 - V_{\rm (RW)}(r)] \,, \tag{12}$$

with $d/dr_* = f(r)d/dr$, and the RW potential

$$V_{(\mathrm{RW})}(r) = \frac{f(r)}{r^2}\left[l(l+1) - \frac{6m_2}{r}\right]. \tag{13}$$

The source terms are singular at the location $r = r_0(t)$ of the particle, namely

$$S^{(\mathrm{even/odd})}_{lm\omega}(r) = \sum_k s^{(\mathrm{even/odd})}_k(t)\delta^{(k)}(r - r_0(t))\,, \tag{14}$$

with $\delta^{(k)}(r - r_0(t))$ denoting the $kth$-order derivative of the Dirac delta function.

Next one needs the solutions to the homogeneous RW equation for $l \geq 2$, which are obtained by using different methods: PN approximation (generic $l$, weak-field and slow motion), Wentzel-Kramers-Brillouin (WKB) approximation (large $l$), Mano, Suzuki and Takasugi (MST) technique[16] (specified values of $l$). The solutions to the inhomogeneous RW equation are then computed through the Green's function method for both even and odd-parity source terms, i.e.,

$$R^{(\mathrm{even/odd})}_{lm\omega}(r) = \int dr' \,\frac{G(r,r')}{f(r')} S^{(\mathrm{even/odd})}_{lm\omega}(r')\,, \tag{15}$$

where the (retarded) Green's function

$$G(r,r') = \frac{1}{W}\Big[X^{\mathrm{in}}_{l\omega}(r)X^{\mathrm{up}}_{l\omega}(r')H(r'-r) + X^{\mathrm{in}}_{l\omega}(r')X^{\mathrm{up}}_{l\omega}(r)H(r-r')\Big]\,, \tag{16}$$

is expressed in terms of the two independent homogeneous solutions $X^{\mathrm{in}}_{l\omega}$ and $X^{\mathrm{up}}_{l\omega}$ of the RW operator. Here $W_{l\omega}$ denotes the (constant) Wronskian

$$W_{l\omega} = f(r)\left[X^{\mathrm{in}}_{l\omega}(r)\frac{dX^{\mathrm{up}}_{l\omega}(r)}{dr} - \frac{dX^{\mathrm{in}}_{l\omega}(r)}{dr}X^{\mathrm{up}}_{l\omega}(r)\right] = \mathrm{const.} \tag{17}$$

and $H(x)$ is the Heaviside step function.

Once the radial function is known for both parities, the perturbed metric components are then computed by Fourier anti-transforming. At this point one can construct the gauge-invariant quantity one is interested in, which still depends on the azimuthal number $m$ and the orbital angular momentum number $l$, and evaluate it at the source location, where the perturbed metric is singular. After multiplying by the angular part, the sum over $m \in [-l, l]$ is done by using spherical harmonic identities, whereas the sum over $l \in [0, \infty)$ requires a suitable regularization procedure to converge. The contribution of the non-radiative multipoles $l = 0, 1$ has to be evaluated separately. In fact, these are purely gauge modes which carry into the hole the energy $E$ and the angular momentum $J$ of the particle (which are both $O(q)$), leading to a linearized Kerr solution with total mass $m_2 + E$ and angular momentum $J$.

## 4. Detweiler's redshift invariant for spinning particles along circular orbits

Let us consider the simplest case of a spinning particle moving along a circular orbit on the equatorial plane with spin vector orthogonal to it. By neglecting higher multipoles the MPD equations imply that $P$ and $U$ are aligned to first order

in spin, the spin vector is parallel transported and both the particle's mass and the spin magnitude are constants of motion. It is convenient to introduce the spin parameter $\hat{s} \equiv s/(m_1 m_2)$ such that $|\hat{s}| \ll 1$.

Detweiler's redshift invariant is defined as the linear-in-mass-ratio change in the time component of the particle's 4-velocity. The latter is given by $U = u^t k$, where $k = \partial_t + \Omega \partial_\phi$ is assumed to be a Killing vector field of the perturbed metric. The normalization condition $U \cdot U = -1$ gives

$$-(u^t)^{-2} = k \cdot k = -f + \Omega^2 r^2 + h_{kk} \,, \tag{18}$$

where

$$h_{kk} = h_{\alpha\beta} k^\alpha k^\beta = h^{(0)}_{kk} + \hat{s} h^{\hat{s}}_{kk} \,. \tag{19}$$

The MPD equations of motion then imply the following solution for the angular velocity $\Omega$

$$m_2 \Omega = u^{3/2} \left[ 1 - \frac{3}{2} \hat{s} u^{3/2} + q(\tilde{\Omega}_1 + \hat{s} \tilde{\Omega}_{1\hat{s}}) \right] , \tag{20}$$

where

$$\tilde{\Omega}_1 = -\frac{m_2}{4u^2} \partial_r h^{(0)}_{kk} \,, \tag{21}$$

as well as the spin correction

$$\tilde{\Omega}_{1\hat{s}} = \tilde{\Omega}^{(h)}_{1\hat{s}} + \tilde{\Omega}^{(\partial h)}_{1\hat{s}} + \tilde{\Omega}^{(\partial^2 h)}_{1\hat{s}} \,, \tag{22}$$

both depend on the metric components $h_{\alpha\beta} = h^{(0)}_{\alpha\beta} + \hat{s} h^{(1)}_{\alpha\beta}$ and their derivatives evaluated at $r = m_2/u$ (see Eq. (2.15) in Ref. 12).

It is useful to introduce the dimensionless frequency parameter $y = (m_2 \Omega)^{2/3}$, which once substituted into Eq. (18) in place of both $\Omega$ and $u$ leads to the following definition of the Detweiler's inverse redshift invariant $z_1(y) = [u^t(y)]^{-1}$

$$\begin{aligned} z_1(y) &= \sqrt{1-3y} - \frac{q}{2\sqrt{1-3y}} \left[ h^{(0)}_{kk}(y) + \hat{s} \left( h^{\hat{s}}_{kk}(y) + m_2 y^{1/2} [\partial_r h^{(0)}_{kk}](y) \right) \right] \\ &\equiv z_1^{(0)}(y) + q \left[ z_1^{(1)\hat{s}^0}(y) + \hat{s} z_1^{(1)\hat{s}^1}(y) \right] . \end{aligned} \tag{23}$$

Therefore, in order to determine the spin dependent term $z_1^{(1)\hat{s}^1}(y)$ one needs two compute $h^{\hat{s}}_{kk}(y)$ and $[\partial_r h^{(0)}_{kk}](y)$ separately, following the procedure outlined in the previous section (all necessary details can be found in Ref. 11). The final result is

$$\begin{aligned} z_1^{(1)\hat{s}^1}(y) = {} & y^{7/2} - 3y^{9/2} - \frac{15}{2} y^{11/2} \\ & + \left( -\frac{6277}{30} + \frac{20471}{1024}\pi^2 - 16\gamma - \frac{496}{15}\ln(2) - 8\ln(y) \right) y^{13/2} \\ & + \left( \frac{653629}{2048}\pi^2 - \frac{87055}{28} - \frac{729}{14}\ln(3) + \frac{3772}{105}\ln(2) - \frac{52}{5}\gamma - \frac{26}{5}\ln(y) \right) y^{15/2} \\ & + O(y^8) \,, \end{aligned} \tag{24}$$

where only the first few terms have been shown, for simplicity.

This is the first GSF analytical computation of a gauge-invariant quantity which takes into account the internal structure of the perturbing body, simply given by its spin in this case. Tidal invariants in the same context has been studied in Ref. 12. These works can be further generalized by adding higher multipoles to the structure of the body (e.g., quadrupole moment), or by considering different orbits (e.g., eccentric), or even a rotating background field (Kerr black hole). In these more general cases the redshift variable should be properly modified in order to be a gauge-invariant quantity. This is still an open problem which requires further investigation.

## Acknowledgments

We thank Prof. T. Damour and Dr. C. Kavanagh for useful discussions.

## References

1. B. P. Abbott et al. [LIGO Scientific and Virgo Collaborations], *Phys. Rev. Lett.* **116**, 061102 (2016).
2. L. Blanchet, *Living Rev. Rel.* **17**, 2 (2014).
3. L. Barack, *Class. Quant. Grav.* **26**, 213001 (2009).
4. A. Buonanno and T. Damour, *Phys. Rev. D* **59**, 084006 (1999)
5. A. Buonanno and T. Damour, *Phys. Rev. D* **62**, 064015 (2000).
6. T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
7. D. Bini and A. Geralico, *Found. Phys.* **48**, 1349 (2018).
8. M. Mathisson, *Acta Phys. Polon.* **6**, 163 (1937).
9. A. Papapetrou, *Proc. Roy. Soc. Lond. A* **209**, 248 (1951).
10. W. G. Dixon, *Proc. Roy. Soc. Lond. A* **314**, 499 (1970).
11. D. Bini, T. Damour, A. Geralico and C. Kavanagh, *Phys. Rev. D* **97**, 104022 (2018).
12. D. Bini and A. Geralico, *Phys. Rev. D* **98**, 084021 (2018).
13. D. Bini, G. Faye and A. Geralico, *Phys. Rev. D* **92**, 104003 (2015).
14. T. Regge and J. A. Wheeler, *Phys. Rev.* **108**, 1063 (1957).
15. F. J. Zerilli, *Phys. Rev. D* **2**, 2141 (1970).
16. S. Mano, H. Suzuki and E. Takasugi, *Prog. Theor. Phys.* **96**, 549 (1996).

# Field theory for gravity at all scales

Michele Levi

*Institut de Physique Théorique, CEA & CNRS, Université Paris-Saclay*
*91191 Gif-sur-Yvette, France*
*michele.levi@ipht.fr*

We review here the main advances made by using effective field theories (EFTs) in classical gravity, with notable focus on those unique to the EFTs of post-Newtonian (PN) gravity. We then proceed to overview the various prospects of using field theory to study the real-world gravitational wave (GW) data, as well as to ameliorate our fundamental understanding of gravity at all scales, by going from the EFTs of PN gravity to modern advances in scattering amplitudes, including computational techniques and intriguing duality relations between gauge and gravity theories.

*Keywords*: Effective field theories; Theories of gravity, Duality of gauge and gravity theories; Scattering amplitudes; High precision calculations; Gravitational waves.

## 1. Introduction

In recent years there has been a growing recognition in the all-encompassing relevance of the field theoretic framework beyond the quantum realm of particle physics, in which much of it has been conceived. Most notably, for long, in the absence of a viable quantum theory of gravity, the overlap between field theory and classical gravity has been largely deemed nonexistent. In particular, the conceptual framework of effective field theories (EFTs), which is intimately tied with the notion of renormalization, was widely regarded as uniquely ingrained in the domain of quantum field theories (QFTs). In order to dispel the latter fallacy, however, suffice it to acknowledge that 'renormalization' can in fact be taken as synonymous with 'resolution of physical scales'. Once this realization is made, it becomes clear that the framework of EFTs is evidently a universal one, and that it can be applied on any wide range of physical scales. Moreover, the EFT framework is a solid and powerful one, providing an inherent setup strategy and a robust toolbox, geared towards yielding high precision predictions to a desirable accuracy.

Indeed, just over a decade ago, an EFT approach to handle gravitational waves (GWs), emitted from inspiralling compact binaries, which are analytically treated with post-Newtonian (PN) gravity [1], was put forward by Goldberger et al. [2–4]. This novel approach has resulted, in turn, further applications of EFTs in classical gravity. Notably, significant progress was demonstrated in the application of EFTs to higher dimensional gravity in the context of large extra dimensions [5–9]. Further, also in the context of GWs, an investigation of an EFT treatment for the extreme mass ratio inspiral (EMRI) case of compact binaries, was initiated in Ref. 10. Moreover, within a decade the initial EFT approach to PN gravity has been able – in important sectors of the theory – to catch up and go beyond the spectacular state of the art [11], already accomplished within the traditional framework of General Relativity (GR).

In what follows we highlight the main meaningful unique advances made in the EFT approach to PN gravity, yet more importantly, we point to the main prospects on a broader scope of using field theory to study gravity, stretching from delivering highly demanding accurate predictions for real-world GW data, to confronting our fundamental grasp of QFTs and gravity theories at all scales.

## 2. From Gravitational Waves to Gravity at all Scales

Significant progress was made via the EFT approach to PN gravity, which is concerned with the analytical prediction of the inspiral portion of the GW signal. While this progress is evident and encouraging, it crucially serves to illustrate the great potential, still awaiting, in the use of field theory to study gravity, with many implications at various levels. First, there are clear anticipated developments still within the EFTs of PN gravity and the theoretical modeling of the GW signal. Further, even if still only in the context of GWs, modern advances in scattering amplitudes hold the promise to possibly enable an analytical treatment of the GW signal in its entirety, that is also in the strong field regime, beyond the inspiral phase. This may significantly improve upon the current phenomenological GW modeling, which relies heavily on the effective-one-body (EOB) formulation[12], by allowing for a smooth analytical model of the whole signal. Yet, more broadly, these advances from scattering amplitudes also importantly allow us to study gauge and gravity theories at a fundamental level, across the classical and quantum regimes.

Let us first review the main unique advancements accomplished within the EFTs of PN gravity so far. The most notable and extensive progress was realized in the treatment of the spinning sector, with several new higher order PN corrections. Next-to-leading order (NLO) effects in the conservative sector were first tackled in Refs. 13, 14, 15, following Refs. 16, 14, and further resolved and extended to the current state of the art, at the fourth PN (4PN) order for rapidly rotating compact objects, in Refs. 17, 18, 19, 20, 21, 22, 23, following Refs. 17, 24, 20, 18. Further, partial results at NLO in the radiative sector were presented in Ref. 25. Moreover, following Ref. 26, which also led to a new finite size correction of radiation reaction in classical electrodynamics[27], and Refs. 28, 29, a leading order effect of radiation reaction with spins was obtained in Ref. 30. All in all, beyond the specific new PN results obtained with spins, there has been an improvement in the understanding of classical spins in gravity.

Another remarkable finding in the field was the uncovering of classical renormalization group (RG) flows of the Wilson coefficients, which characterize the effective theories, both at the one-particle level[31], and at the level of the composite system, for the multipole moments of the binary[4,32]. These give rise to higher order PN logarithm corrections, which constitute unique predictions of the EFT framework. In addition, the 2PN order correction for a generic $n$-body problem was considered in Ref. 33, where automated computations were first advocated. Finally, the 'EFTofPNG' code for high precision Feynman calculations in PN gravity was

created[34], comprising the first comprehensive code in PN theory made public. The code consists of independent modules for easy adaptation and future development, where some prospects of building on this state of the art PN technology for various research purposes, were outlined in Ref. 35.

Within the methodology of the EFTs of PN gravity, there are several important directions, where development is required. The first is the solution of higher order PN equations of motion (EOMs) for accurate orbital dynamics. One interesting related way to find such closed form solutions, is the dynamical RG method, suggested in Ref. 36. Further, the treatment of non-conservative sectors with the EFTs of PN gravity has been rather limited, and so these various sectors should be tackled with a proper EFT formulation, and implementation for new PN results. Another objective to follow is the possible improvement or extension of the EFT formulation for spinning objects. For example, an alternative EFT formulation, which is ideally suited for the treatment of slowly rotating objects, based on a coset construction, was presented in Ref. 37. Further investigation is required in order to see whether these independent EFT formulations may possibly be integrated, and entail an even better understanding of classical spins in gravity. Next, we note the analysis of the mass- and spin-induced Wilson coefficients, and of the binary multipole coefficients, which also mostly remains to be approached via formal EFT matching. Finally, the continual public development of the 'EFTofPNG' code, or similar open codes, is required in order to keep up to date with the actual progress in PN theory. In that regard, it is important to note that classical theories of gravity, which modify GR in the IR, motivated by the cosmological constant problem and the dark matter puzzle[38–41], can be incorporated in a rather straightforward manner into the 'EFTofPNG' code, in order to study their analytical predictions for the GW signal from the compact binary inspiral, see, e.g., Refs. 42, 43, 44.

### 2.1. *Advances and prospects for gravity in field theory*

Ultimately, we would like to invoke the correspondence – at any level – between gauge and gravity theories, in order to push beyond the state of the art in high precision computation for concrete analytical predictions of the real-world GW signal, as well as to directly ameliorate our fundamental understanding of the foundations of these theories, in particular of gravity. These objectives can be considerably facilitated by turning to modern field theory advances in the domain of scattering amplitudes, see, e.g., Ref. 45. In fact, the field of EFTs in PN gravity, on which we have focused hitherto, and of scattering amplitudes, which are both directly concerned with baldly tackling demanding high precision computation driven by experiment, share more profound parallels beyond the obvious robust technical tools they entail, see, e.g., in Ref. 46. Both of these fields push to the exposure of the universal commonalities across classical and quantum field theories, and drive us to confront our fundamental grasp of the underlying foundations of these theories. One crucial difference to note, though, between these fields, is that where in

scattering amplitudes the computational outcome is directly related with the physical observables, the objects which are commonly computed in classical gravity are coordinate/gauge dependent quantities.

Let us then note, more specifically, some prospective avenues to deploy this broad correspondence of gauge theories and gravity. First, there is the standard working knowledge, which can be exchanged between the theories, e.g. multi-loop techniques in QFT, such as integration by parts (IBP), and high loop master integrals, along with other Feynman calculus and technology[34,35,46]. This was nicely demonstrated, e.g., in Refs. 47, 48, where an analytic evaluation of a four-loop master integral was provided for the first time from the classical gravity context. Furthermore, modern scattering amplitudes advances, such as the BCFW on-shell recursion relations[49], and generalized unitarity methods, which imply that tree level data encodes all multiplicity at the integrand level, were put forward to extract classical higher order loop results for gravitational scattering[50–53]. Such a scattering treatment may also enable to analytically tackle the strong field regime of the GW signal, and hence smoothly model analytically the entire signal, as we noted above.

Yet, of particular interest are the novel intriguing color-kinematics or BCJ duality relations, and the related double copy correspondence[54,55], see, e.g., review in Ref. 56, which were discovered in the context of high loop computations of amplitudes in supersymmetry and supergravity theories. Such relations were already known from string theory to hold at tree level, but formulated in terms of the novel generic double copy correspondence, these relations have been successfully used to study UV divergences of supergravity theories. This recent new perspective on gravity, viewing gravitons as double copies of gluons, suggests that even the simplest gauge theories and GR are in fact frameworks, which are intimately connected. Though this correspondence was uncovered in the perturbative context, it was also found, remarkably, that particular classical *exact* solutions in GR, are in fact double copies of exact "single copy" counterparts in corresponding gauge theories. Such classical solutions include all vacuum stationary solutions, most notably, Kerr black holes, as well as their higher dimensional generalizations, see Ref. 57 and references therein. Yet, the underlying origin of these perturbative relations, as well as their essential connection to the particular classical exact correspondence revealed, are yet to be uncovered.

In conjunction with the recent novel methods from scattering amplitudes, it is expected that this double copy correspondence can be used to advance the analytical calculations for the theoretical prediction of the GW signal, which may also help to shed more light on the nature and origin of this correspondence. Related with that end, an important time dependent case was studied in terms of the exact double copy of the Kerr-Schild form, of an arbitrarily accelerating point source[57]. In this case the radiation current is double copied to the radiation stress-energy tensor in Fourier space, and both are related to the corresponding scattering amplitudes, which can be obtained from the amputated currents. Hence, the work in Ref. 57

made a first explicit connection between the classical double copy in an exact form, to that in the perturbative scattering amplitudes context. At this point it should also be stressed, that from the scattering amplitudes context, it is already known that gluon amplitudes can double copy to arbitrary combinations of amplitudes for gravitons, and the additional unobserved dilaton and B fields, depending on the choice of the polarization states in the gauge theory amplitudes. This was indeed an ambiguous issue in Ref. 58, which built on Ref. 57, and that subsequently Ref. 59 set out to address, with Ref. 60 following up successfully at NLO.

Finally, we note the topic of soft graviton theorems for scattering amplitudes, which were recently demonstrated to be equivalent to gravitational memory effects, as part of a triple equivalence of the IR structure of gauge and gravity theories among soft theorems, asymptotic symmetry, and memory effects[61]. The latter may have observable signatures on the GW signal.

## Acknowledgments

I am grateful to John Joseph Carrasco for his meaningful encouragement. I would like to thank Donato Bini for the warm hospitality throughout the MG15 meeting, and in particular on the session, where this review was presented. It is also a delightful pleasure to acknowledge Roy Kerr who graced us with his inspiring presence. My work is supported by the European Research Council under the European Union's Horizon 2020 Framework Programme FP8/2014-2020 "preQFT" grant no. 639729, "Strategic Predictions for Quantum Field Theories" project.

## References

1. L. Blanchet, *Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries*, *Living Rev. Rel.* **17** (2014).
2. W. D. Goldberger and I. Z. Rothstein, *An Effective field theory of gravity for extended objects*, *Phys. Rev.* **D73** (2006) 104029.
3. W. D. Goldberger, *Les Houches lectures on effective field theories and gravitational radiation*, in *Les Houches Summer School - Session 86: Particle Physics and Cosmology: The Fabric of Spacetime Les Houches, France, July 31-August 25, 2006*, 2007.
4. W. D. Goldberger and A. Ross, *Gravitational radiative corrections from effective field theory*, *Phys. Rev.* **D81** (2010) 124015.
5. Y.-Z. Chu, W. D. Goldberger, and I. Z. Rothstein, *Asymptotics of d-dimensional Kaluza-Klein black holes: Beyond the Newtonian approximation*, *JHEP* **03** (2006) 013.
6. B. Kol and M. Smolkin, *Classical Effective Field Theory and Caged Black Holes*, *Phys. Rev.* **D77** (2008) 064033.
7. R. Emparan, T. Harmark, V. Niarchos, and N. A. Obers, *World-Volume Effective Theory for Higher-Dimensional Black Holes*, *Phys. Rev. Lett.* **102** (2009) 191301.

8. J. B. Gilmore, A. Ross, and M. Smolkin, *Caged black hole thermodynamics: Charge, the extremal limit, and finite size effects*, *JHEP* **09** (2009) 104.
9. R. Emparan, T. Harmark, V. Niarchos, and N. A. Obers, *Essentials of Blackfold Dynamics*, *JHEP* **03** (2010) 063.
10. C. R. Galley and B. Hu, *Self-force on extreme mass ratio inspirals via curved spacetime effective field theory*, *Phys. Rev.* **D79** (2009) 064002.
11. M. Levi, `arXiv:1807.01699` *Effective Field Theories of Post-Newtonian Gravity: A comprehensive review*, 2018.
12. A. Buonanno and T. Damour, *Effective one-body approach to general relativistic two-body dynamics*, *Phys. Rev.* **D59** (1999) 084006.
13. R. A. Porto and I. Z. Rothstein, *The Hyperfine Einstein-Infeld-Hoffmann potential*, *Phys. Rev. Lett.* **97** (2006) 021101.
14. R. A. Porto and I. Z. Rothstein, *Spin(1)Spin(2) Effects in the Motion of Inspiralling Compact Binaries at Third Order in the Post-Newtonian Expansion*, *Phys. Rev.* **D78** (2008) 044012, [*Erratum-ibid.* D**81** (2010) 029904].
15. R. A. Porto and I. Z. Rothstein, *Next to Leading Order Spin(1)Spin(1) Effects in the Motion of Inspiralling Compact Binaries*, *Phys. Rev.* **D78** (2008) 044013, [*Erratum-ibid.* D**81** (2010) 029905].
16. R. A. Porto, *Post-Newtonian corrections to the motion of spinning bodies in NRGR*, *Phys. Rev.* **D73** (2006) 104031.
17. M. Levi, *Next to Leading Order gravitational Spin1-Spin2 coupling with Kaluza-Klein reduction*, *Phys. Rev.* **D82** (2010) 064029.
18. M. Levi and J. Steinhoff, *Spinning gravitating objects in the effective field theory in the post-Newtonian scheme*, *JHEP* **09** (2015) 219.
19. M. Levi, *Binary dynamics from spin1-spin2 coupling at fourth post-Newtonian order*, *Phys. Rev.* **D85** (2012) 064043.
20. M. Levi and J. Steinhoff, *Equivalence of ADM Hamiltonian and Effective Field Theory approaches at next-to-next-to-leading order spin1-spin2 coupling of binary inspirals*, *JCAP* **1412** (2014) 003.
21. M. Levi and J. Steinhoff, *Leading order finite size effects with spins for inspiralling compact binaries*, *JHEP* **06** (2015) 059.
22. M. Levi and J. Steinhoff, *Next-to-next-to-leading order gravitational spin-squared potential via the effective field theory for spinning objects in the post-Newtonian scheme*, *JCAP* **1601** (2016) 008.
23. M. Levi and J. Steinhoff, `arXiv:1607.04252` *Complete conservative dynamics for inspiralling compact binaries with spins at fourth post-Newtonian order*, 2016.
24. M. Levi, *Next to Leading Order gravitational Spin-Orbit coupling in an Effective Field Theory approach*, *Phys. Rev.* **D82** (2010) 104004.
25. R. A. Porto, A. Ross, and I. Z. Rothstein, *Spin induced multipole moments for the gravitational wave flux from binary inspirals to third Post-Newtonian order*, *JCAP* **1103** (2011) 009.

26. C. R. Galley and M. Tiglio, *Radiation reaction and gravitational waves in the effective field theory approach*, *Phys. Rev.* **D79** (2009) 124027.
27. C. R. Galley, A. K. Leibovich, and I. Z. Rothstein, *Finite size corrections to the radiation reaction force in classical electrodynamics*, *Phys. Rev. Lett.* **105** (2010) 094802.
28. C. R. Galley, *Classical Mechanics of Nonconservative Systems*, *Phys. Rev. Lett.* **110** (2013) 174301.
29. C. R. Galley, D. Tsang, and L. C. Stein, `arXiv:1412.3082` *The principle of stationary nonconservative action for classical mechanics and field theories*, 2014.
30. N. T. Maia, C. R. Galley, A. K. Leibovich, and R. A. Porto, *Radiation reaction for spinning bodies in effective field theory II: Spin-spin effects*, *Phys. Rev.* **D96** (2017) 084065.
31. B. Kol and M. Smolkin, *Black hole stereotyping: Induced gravito-static polarization*, *JHEP* **02** (2012) 010.
32. W. D. Goldberger, A. Ross, and I. Z. Rothstein, *Black hole mass dynamics and renormalization group evolution*, *Phys. Rev.* **D89** (2014) 124033.
33. Y.-Z. Chu, *The n-body problem in General Relativity up to the second post-Newtonian order from perturbative field theory*, *Phys. Rev.* **D79** (2009) 044031.
34. M. Levi and J. Steinhoff, *EFTofPNG: A package for high precision computation with the Effective Field Theory of Post-Newtonian Gravity*, *Class. Quant. Grav.* **34** (2017) 244001.
35. M. Levi, `arXiv:1811.12401` *A public framework for Feynman calculations and post-Newtonian gravity*, 2018.
36. C. R. Galley and I. Z. Rothstein, *Deriving analytic solutions for compact binary inspirals without recourse to adiabatic approximations*, *Phys. Rev.* **D95** (2017) 104054.
37. L. V. Delacrétaz, S. Endlich, A. Monin, R. Penco, and F. Riva, *(Re-)Inventing the Relativistic Wheel: Gravity, Cosets, and Spinning Objects*, *JHEP* **11** (2014) 008.
38. T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, *Modified Gravity and Cosmology*, *Phys. Rept.* **513** (2012) 1–189.
39. C. Deffayet and D. A. Steer, *A formal introduction to Horndeski and Galileon theories and their generalizations*, *Class. Quant. Grav.* **30** (2013) 214006.
40. C. de Rham, *Massive Gravity*, *Living Rev. Rel.* **17** (2014) 7.
41. A. Joyce, B. Jain, J. Khoury, and M. Trodden, *Beyond the Cosmological Standard Model*, *Phys. Rept.* **568** (2015) 1–98.
42. Y.-Z. Chu and M. Trodden, *Retarded Green's function of a Vainshtein system and Galileon waves*, *Phys. Rev.* **D87** (2013) 024011.
43. C. de Rham, A. J. Tolley, and D. H. Wesley, *Vainshtein Mechanism in Binary Pulsars*, *Phys. Rev.* **D87** (2013) 044025.

44. C. de Rham, A. Matas, and A. J. Tolley, *Galileon Radiation from Binary Systems*, *Phys. Rev.* **D87** (2013) 064024.
45. H. Elvang and Y. T. Huang, *Scattering Amplitudes in Gauge Theory and Gravity* (Cambridge University Press, 2015).
46. V. A. Smirnov, *Analytic tools for Feynman integrals,* Springer Tracts Mod. Phys. **250**, 1 (2012).
47. S. Foffa, P. Mastrolia, R. Sturani, and C. Sturm, *Effective field theory approach to the gravitational two-body dynamics, at fourth post-Newtonian order and quintic in the Newton constant*, *Phys. Rev.* **D95** (2017) 104009.
48. T. Damour and P. Jaranowski, *Four-loop static contribution to the gravitational interaction potential of two point masses*, *Phys. Rev.* **D95** (2017) 084005.
49. R. Britto, F. Cachazo, and B. Feng, *New recursion relations for tree amplitudes of gluons*, *Nucl. Phys.* **B715** (2005) 499–522.
50. N. E. J. Bjerrum-Bohr, J. F. Donoghue, and P. Vanhove, *On-shell Techniques and Universal Results in Quantum Gravity*, *JHEP* **02** (2014) 111.
51. F. Cachazo and A. Guevara, `arXiv:1705.10262` *Leading Singularities and Classical Gravitational Scattering*, 2017.
52. A. Guevara, `arXiv:1706.02314` *Holomorphic Classical Limit for Spin Effects in Gravitational and Electromagnetic Scattering*, 2017.
53. N. E. J. Bjerrum-Bohr, P. H. Damgaard, G. Festuccia, L. Planté, and P. Vanhove, *General Relativity from Scattering Amplitudes*, *Phys. Rev. Lett.* **121** (2018) 171601.
54. Z. Bern, J. J. M. Carrasco, and H. Johansson, *New Relations for Gauge-Theory Amplitudes*, *Phys. Rev.* **D78** (2008) 085011.
55. Z. Bern, J. J. M. Carrasco, and H. Johansson, *Perturbative Quantum Gravity as a Double Copy of Gauge Theory*, *Phys. Rev. Lett.* **105** (2010) 061602.
56. J. J. M. Carrasco, *Gauge and Gravity Amplitude Relations*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders (TASI 2014): Boulder, Colorado, June 2-27, 2014*, pp. 477–557, WSP, 2015.
57. A. Luna, R. Monteiro, I. Nicholson, D. O'Connell, and C. D. White, *The double copy: Bremsstrahlung and accelerating black holes*, *JHEP* **06** (2016) 023.
58. W. D. Goldberger and A. K. Ridgway, *Radiation and the classical double copy for color charges*, *Phys. Rev.* **D95** (2017) 125010.
59. A. Luna, I. Nicholson, D. O'Connell, and C. D. White, *Inelastic Black Hole Scattering from Charged Scalar Amplitudes*, *JHEP* **03** (2018) 044.
60. C.-H. Shen, *Gravitational Radiation from Color-Kinematics Duality*, *JHEP* **11** (2018) 162.
61. A. Strominger, `arXiv:1703.05448` *Lectures on the Infrared Structure of Gravity and Gauge Theory*, 2017.

# Gravitational waves from spinning binary black holes at the leading post-Newtonian orders at all orders in spin

Nils Siemonsen[1*], Jan Steinhoff[2] and Justin Vines[2]

[1] *Institute for Theoretical Physics, ETH Zürich*
*Wolfgang-Pauli-Strasse 27, 8093 Zürich, Switzerland*

[2] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute)*
*Am Mühlberg 1, 14476 Potsdam, Germany*
[*] *E-mail: sinils@student.ethz.ch*

We determine the binding energy, the total gravitational wave energy flux, and the gravitational wave modes for a binary of rapidly spinning black holes, working in linearized gravity and at leading orders in the orbital velocity, but to all orders in the black holes' spins. Though the spins are treated nonperturbatively, surprisingly, the binding energy and the flux are given by simple analytical expressions which are finite (respectively third- and fifth-order) polynomials in the spins. Our final results are restricted to the important case of quasi-circular orbits with the black holes' spins aligned with the orbital angular momentum.

*Keywords*: Binary black holes; Gravitational waves; Nonperturbative method; High precision calculations; Spinning point particles; Post-Newtonian approximation.

## 1. Introduction

The general relativistic two-body problem, in particular the description of compact binaries, as one of the most important sources of gravitational waves (GW) detectable from Earth, poses a large challenge for analytical as well as numerical calculations. Accurate waveform models are an essential ingredient for the detection of GW signals. The detections of such signals[1] have demonstrated the success of the approaches, but the need for more accurate and more general descriptions of binary dynamics in general relativity (GR) persists and will grow with future more sensitive GW detectors. Several (perturbative) analytic approaches to the two-body problem in GR have been developed. Among others, the post-Newtonian (PN) expansion[2]. It is an analytic approach for arbitrary mass ratio compact binaries in the weak-field, slow-motion (characterized by $\epsilon_{\rm PN} \sim v^2/c^2 \sim Gm/rc^2$) and small spin regime (determined by the expansion parameter $\epsilon_{\rm spin} \sim Gm\chi/rc^2$)[a]. Analytic results from the PN approximation, together with numerical descriptions of the late inspiral, plunge, and merger, can be combined synergetically in effective-one-body (EOB) models[3,4].

Previously, PN results for binary black holes (BBHs), in particular for the black holes' (BH) spins, were obtained order by order in the perturbative series, since for BBHs: $\epsilon_{\rm PN} \sim \epsilon_{\rm spin}$. This way, the conservative dynamics of spinning BBHs was computed up to 4PN order[5]; for the leading orders at various orders in spin see Refs. 6–13. For spin-dependent leading orders in the radiative sector, see Refs. 14–17.

[a] With $cS/Gm^2 = \chi \in [0, 1)$, the dimensionless spin parameter.

However, by treating both expansion parameters separately, we are able to derive compact analytic expressions for gauge-invariant quantities (GW energy flux, GW modes, conserved energy etc.) for a BBH at the leading PN orders to all orders in the BHs' spins – treating the spins nonperturbatively.

## 2. Effective action

In a harmonic-gauge framework, with a linearized metric perturbation $h_{\mu\nu} = g_{\mu\nu} - \eta_{\mu\nu} \sim \mathcal{O}(G)$, following the derivations in Refs. 19 and 20, we implement an effective action of a two point-particle system with a spin-induced multipole structure (for each particle) and respective multipole couplings to the gravitational field. In $G = c = 1$ units, this effective description is split into particle- and spin-kinetic terms $S_{\rm kin}[\mathbf{T}_A]$, interaction terms $S_{\rm int}[\mathbf{h}, \mathbf{T}_A]$ of the $A$th black hole and a term, $S_G[\mathbf{h}]$, containing the gravitational field's dynamics at linear order in $h_{\mu\nu}$, so that

$$S^{\rm BBH}_{\rm eff}[\mathbf{h}, \mathbf{T}_1, \mathbf{T}_2] = S_G[\mathbf{h}] + \{S_{\rm kin}[\mathbf{T}_1] + S_{\rm int}[\mathbf{h}, \mathbf{T}_1] + (1 \leftrightarrow 2)\}.$$

The Wilson coefficients in Ref. 18 were set to unity through a matching procedure to Kerr spacetime, as outlined in Ref. 20. This way, all spin-induced multipoles of the black holes are coupling to the linear metric perturbation $h_{\mu\nu}$ by (still containing nonlinear velocity contributions) [18,19]

$$S_{\rm int}[\mathbf{h}, \mathbf{T}_A] = \int dt \left\{ \sum_{\ell=0}^{\infty} \frac{m U^\mu_A U^\nu_A}{2\gamma_A\, \ell!} {\rm Re}\left[ i^\ell a^L_A \partial_L h_{\mu\nu} + i^{\ell-1} a^\rho_A \epsilon_{\nu\rho}{}^{\alpha\beta} a^{L-1}_A \partial_\alpha \partial_{L-1} h_{\beta\mu} \right] \right\},$$

where $a^\mu_A$ is the spin 4-vector, $U^\mu_A$ the velocity 4-vector (parameterized by coordinate time $t$) and $\gamma_A$ the Lorentz factor of the $A$th BH respectively. $\mathbf{h}$ encompasses the gravitational degrees of freedom and $\mathbf{T}$ encodes the multipolar degrees of freedom of the individual black holes. Note, we used the multi-index notation $L := \mu_1 \dots \mu_\ell$. This interaction term arises from considering all possible combinations of the vacuum Riemann tensor, multipole moments and particle's velocities at linear order in the metric perturbation, under the restriction of reparameterization invariance [18].

## 3. Conservative dynamics

From the above action principle, the linearized field equations in the near zone of the binary can be deduced (to linear order the velocities) and solved to all orders in spin. A Fokker-type action with Lagrangian (including all spin-spin couplings in this linearized framework with up to linear velocity contributions)

$$\begin{aligned} \mathcal{L} = &\left[ - m_1 + \frac{m_1}{2} v_1^2 + \frac{1}{2} \boldsymbol{S}_1 (\boldsymbol{v}_1 \times \dot{\boldsymbol{v}}_1) + \boldsymbol{S}_1 \cdot \boldsymbol{\Omega}_1 + (1 \leftrightarrow 2) \right] \\ &+ m_1 m_2 \sum_{\ell=0}^{\infty} \frac{(-1)^\ell}{(2\ell)!} \left[ a_0^{2L} + \frac{2 v^i a_0^j \epsilon_{ij}{}^k a_0^{2L}}{(2\ell+1)} \frac{\partial}{\partial r^k} \right] \frac{\partial r^{-1}}{\partial r^{2L}} \end{aligned} \tag{1}$$

in the center-of-mass frame, is obtained. We introduced the BHs' masses $m_A$, 3-velocities $\boldsymbol{v}_A(t)$, spin 3-vectors $\boldsymbol{a}_A = \boldsymbol{S}_A/m_A$ and combination $\boldsymbol{a}_0 = \boldsymbol{a}_1 + \boldsymbol{a}_2$, as

well as the BHs' separation vector $\boldsymbol{r}$ with $r = |\boldsymbol{r}|$ and the angular velocity vector $\boldsymbol{\Omega}_A$. The equations of motion for the BBH, in a quasi-circular limit and for spin vectors aligned with the orbital angular momentum of the system, can be resummed and solved for the angular frequency $\omega$ of the orbits in compact fashion. To linear order in the metric perturbation $h_{\mu\nu}$ and velocities $\boldsymbol{v}$ (i.e., at leading PN order), but to all orders in the spins $\boldsymbol{a}_A$, the angular frequency is given by $\ddot{\boldsymbol{r}} = -\omega^2 \boldsymbol{r}$, with $\omega^2 = M/r[r - v(2a_0 + \sigma^*)](r^2 - a_0^2)^{-3/2}$. Here we defined $M = m_1 + m_2$ and $\sigma^* = |\boldsymbol{\sigma}^*| = |m_2\boldsymbol{a}_1 + m_1\boldsymbol{a}_2|/M$.

Gauge-invariant quantities (e.g., Noether charges or the GW modes) are conveniently expanded in the commonly used expansion parameter $x = (M\omega)^{2/3}$, such that $\epsilon_{\rm PN} \sim x$ and $\epsilon_{\rm spin} \sim x\chi$. In the following $\chi$ serves as book keeping parameter for the order in spin considered (e.g., $a_A \sim \mathcal{O}(\chi)$, $a_A^2 \sim \mathcal{O}(\chi^2)$, ...). At each order in $\chi$, we consider the lowest order in $x$, i.e., the leading PN order at each order in spin. Notice that this is different from the traditional PN order counting in $1/c^2$ (since we treated $\epsilon_{\rm spin}$ and $\epsilon_{\rm PN}$ as independent).

The conserved energy and total angular momentum of the two-body system can be obtained as the associated Noether charges of (1). Recasting the angular frequency $\omega$ in $x$ yields, together with the respective Noether charges as functions of $r$, the binding energy $E(x)$ and total angular momentum $J(x) = L(x) + m_1 a_1 + m_2 a_2$ perpendicular to the orbital plane of the BBH at the leading PN orders at all orders in spin

$$E(x) = -\frac{\mu x}{2}\Big\{1 + \frac{x^{3/2}}{3M}(7a_0 + \delta a_-) - \frac{x^2 a_0^2}{M^2} - \frac{x^{7/2} a_0^2}{M^3}(a_0 - \delta a_-)\Big\}, \text{ and}$$
$$L(x) = \mu x^{-1/2}\Big\{M - \frac{5}{12}x^{3/2}(7a_0 + \delta a_-) + \frac{x^2 a_0^2}{M} + \frac{3x^{7/2}}{4M^2}a_0^2(a_0 - \delta a_-)\Big\}.$$

Where the symmetric mass ratio $\nu = \mu/M$, the antisymmetric mass ratio $\delta = (m_1 - m_2)/M$, and $a_-$ being the projection of $\boldsymbol{a}_- = \boldsymbol{a}_1 - \boldsymbol{a}_2$ orthogonal to the orbital plane.

Remarkably, the spin-expansions of the binding energy $E$ and orbital angular momentum $L$ terminate after cubic-in-spin contributions. Note that this polynomial structure in spin hinges on the use of $x$ (or $\omega$) as the variable in the conserved quantities.

## 4. Radiative sector

The far zone dynamics, i.e., the gravitational effects at future null infinity, are related with the near zone dynamics through a matching of the PN solution obtained above and the far zone post-Minkowskian expansion[2]. The matching procedure yields a relation between the source multipole moments and the emitted GW modes and total GW energy flux at future null infinity. The GW polarization waveform $h_+ - ih_\times = \sum_{\ell,m} {}_{-2}Y_{\ell m} h_{\ell m}$ is projected onto a basis of spin weighted spherical harmonics ${}_{-s}Y_{\ell m}$. The GW modes $h_{\ell m}$ are explicitly given in eq. (84) of

Ref. 20. The spin expansion of the even-$m$ modes terminates at a finite order. The odd-$m$ modes have contributions at all orders in spin, though they, are resumed in a compact form, i.e., terms like $\sqrt{M^2+x^2a_0^2}$ appear. Expanding these term, $\sqrt{M^2+x^2a_0^2} = M + x^2a_0^2/(2M) - x^4a_0^4/(8M^3) + \mathcal{O}(x^6\chi^6)$, yields the contributions at the leading PN order at each order in spin.

In an adiabatic approximation, at the leading PN orders at each order in the BHs' spin, we find the total GW energy flux

$$\mathcal{F} = \frac{\mu^2 x^5}{M^2}\Bigg[\frac{32}{5} - \frac{8x^{3/2}}{5M}\Big\{8a_0 + 3\delta a_-\Big\} + \frac{2x^2}{5M^2}\Big\{32a_0^2 + a_-^2\Big\} - \frac{4x^{7/2}}{15M^3}\Big\{16a_0^3 + 2a_0a_-^2 + 52\delta a_0^2a_- + \delta a_-^3\Big\} + \frac{2x^4a_0^2}{5M^4}\Big\{16a_0^2 + a_-^2\Big\} + \frac{2a_0^2x^{11/2}}{15M^5}\Big\{64a_0^3 + a_0a_-^2 - 68\delta a_0^2a_- - 3\delta a_-^3\Big\}\Bigg].$$

Again, the infinite sets of spin-induced multipolar interactions of the two BHs remarkably cancel out at higher than quintic-in-spin contributions. Hence, the total energy flux conveys full information about the spin effects at leading PN order in the first five terms of the spin expansion.

## 5. Conclusion

We determined the binding energy, the gravitational wave modes and total energy flux emitted by a spinning nonprecessing binary black hole in quasi-circular motion at leading post-Newtonian orders at all orders in spin. Our results include contributions of arbitrarily large PN order, counting in $1/c^2$. In particular, we obtained for the first time the quartic-in-spin contributions to the 4PN waveform and total energy flux, along with all higher-order-in-spin contributions at the corresponding leading PN orders. Remarkably, the binding energy, the total energy flux, as well as the even-in-$m$ gravitational wave modes only contain a finite number of nonzero contributions in their spin expansions at leading post-Newtonian order.

Conversely, the modes where all powers in spin appear are nevertheless rather compact, which can be used to improve the resummation of modes, e.g., in the synergetic EOB waveform model[21]. Though our results are only valid for aligned spins, they can still be used to approximate waveforms from precessing binaries[22].

## References


1. B. P. Abbott *et. al.* (Virgo, LIGO Scientific), *Phys. Rev. Lett.* **116**, 24113 (2016).
2. L. Blanchet, *Living Rev. Relativity* **17**, 2 (2014).
3. A. Buonanno and T. Damour, *Phys. Rev. D* **59**, 084006 (1999).
4. A. Buonanno and T. Damour, *Phys. Rev. D* **62**, 064015 (2000).
5. M. Levi and J. Steinhoff, (2016), arXiv:1607.04252 [gr-qc].

6. W. Tulczyjew, *Acta Phys. Polon* **18**, 37 (1959).
7. B. M. Barker and R. F. O'Connell, *Phys. Rev. D* **2**, 1428 (1970).
8. P. D. D'Eath, *Phys. Rev. D* **12**, 2183 (1975).
9. K. S. Thorne and J. B. Hartle, *Phys. Rev. D* **31**, 1815 (1985).
10. E. Poisson, *Phys. Rev. D* **57**, 5287 (1998).
11. T. Damour, *Phys. Rev. D* **64**, 124013 (2001).
12. S. Hergt and G. Schäfer, *Phys. Rev. D* **77**, 104001 (2008).
13. M. Levi and J. Steinhoff, *JHEP* **06**, 059 (2015).
14. S. Marsat, *Class. Quant. Grav.* **32**, 085008 (2015).
15. L. E. Kidder, C. M. Will and A. G. Wisemann, *Phys. Rev. D* **47**, R4183 (1993).
16. L. E. Kidder, *Phys. Rev. D* **52**, 821 (1995).
17. A. Buonanno, G. Faye and T. Hinderer, *Phys. Rev. D* **87**, 044009 (2013).
18. M. Levi and J. Steinhoff, *JHEP* **09**, 219 (2015).
19. J. Vines and J. Steinhoff, *Phys. Rev. D* **97**, 064010 (2018).
20. N. Siemonsen, J. Steinhoff and J. Vines, *Phys. Rev. D* **97**, 124046 (2018).
21. Y. Pan, A. Buonanno, R. Fujita, E. Racine and H. Tagoshi *Phys. Rev. D* **83**, 064003 (2011).
22. Y. Pan, A. Buonanno, A. Taracchini, L. E. Kidder, A. H. Mroué, H. P. Pfeiffer, M. A. Scheel and B. Szilágyi, *Phys. Rev. D* **89**, 084006 (2014).

# Quantitative analysis of the inflationary model: reheating process

Massimo Carinci[1a] and Giovanni Montani[1b]

[1] *Department of Physics, La Sapienza, Rome, Italy*
[a] *massimo.carinci@uniroma1.it;* [b] *giovanni.montani@enea.it*

There are various models that were proposed as scalar fields for inflation, some of these with an 'attractor solutions' to which the system evolves. For each of these models, we have a different evolution of the energy density for relativistic matter and dark matter, which were created during the reheating period, at the end of inflation in the early Universe, as a result of the 'inflaton' decay. Furthermore, the temperature at which the reheating takes place also depends on the type of scalar field. In this work, we consider the polynomial scalar field (chaotic inflation) of the sixth degree, and we study it with numerical methods varying the 'free parameters' a, b, c, obtaining the attractor behavior in particular conditions. Finally, we calculate the reheating temperature, establishing, for this case, that we are in the regime of *good reheating*.

## 1. Scalar Field Dynamics

During inflation, the Universe is dominated by a scalar field '$\varphi$', the **inflaton**, which causes the accelerated expansion. The field behavior is given by the potential $V(\phi)$. The field 'rolls' along the potential (the so-called **slow-roll inflation**, see Figure 1 on the right) and stops in the minimum, after that inflation ends.

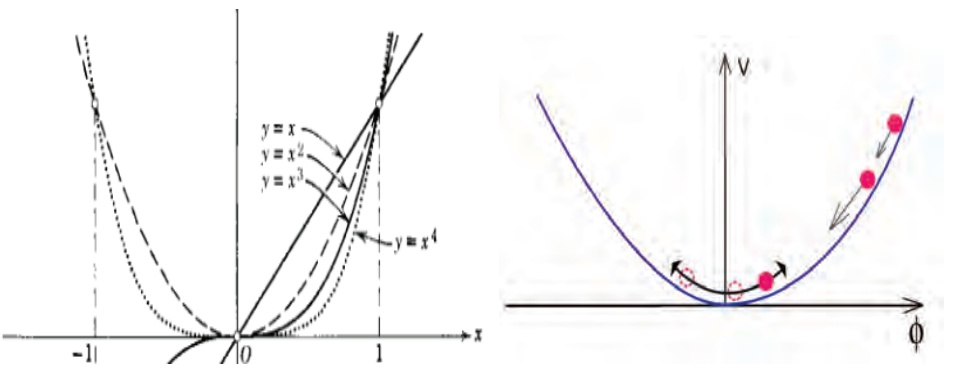



Fig. 1. On the left, form of the chaotic potential varying the exponent of the power scalar field $\alpha$. On the right, the quadratic potential: the circles show the inflaton field at various stages of its evolution. When the circle reaches the minimum, the inflation ends. So the circle starts to oscillate around the minimum and the reheating phase begins.

The simplest potentials which describe the inflation are (see Figure 1 on the left):

$$V = a \times \phi^{\alpha} \quad a = \text{const}, \quad \alpha = 2, 4, 6$$

This kind of potential, proposed by Linde, is said '*chaotic inflation*'. It is symmetric and has a minimum for $\phi = 0$, where $V = 0$.

Fixed $\alpha$, the coefficient a can be calculated with the following formula[3]:

$$\frac{g}{M_{Pl}^{4-n}} = \frac{3}{8\pi}\left(\frac{n}{4}\right)^2 \cdot \left(\frac{4\pi}{nN_e}\right)^{(n+2)/2} \cdot \Delta_{\mathcal{R}}^2.$$

where $M_{pl}$ is the Planck mass, $\Delta_{\Re}$ is the amplitude of the scalar perturbation, whose value is $5 \times 10^{-5}$ and $\mathrm{N}_e$ is the e-folding number. So, assuming $\mathrm{N}_e = 60$, we obtain:

| | |
|---|---|
| ✓ $a \leq 10^{-13}$ | *if* $\alpha = 4$; |
| ✓ $a \leq 10^{-15} M_{pl}^{-2}$ | *if* $\alpha = 6$. |

## 2. Attractor Solution

In the phase space, these potentials have attractor solution to which its solutions must converge over time (see Figure 2 on the left, for the quadratic potential). If we consider a potential with all the three terms:

$$\boxed{\mathrm{V} = \mathrm{a} \times \varphi^2 + \mathrm{b} \times \varphi^4 + \mathrm{c} \times \varphi^6} \tag{1}$$

When $\mathrm{a} > \mathrm{b}, \mathrm{c}$ we have no changes (Figure 2 on the right, in red). But if b and/or c is of the same order of a, we have a pick (in blue). In fact, in this case the potential is less inclined (Figure 1, left), and the solution, at first, converges more slowly.

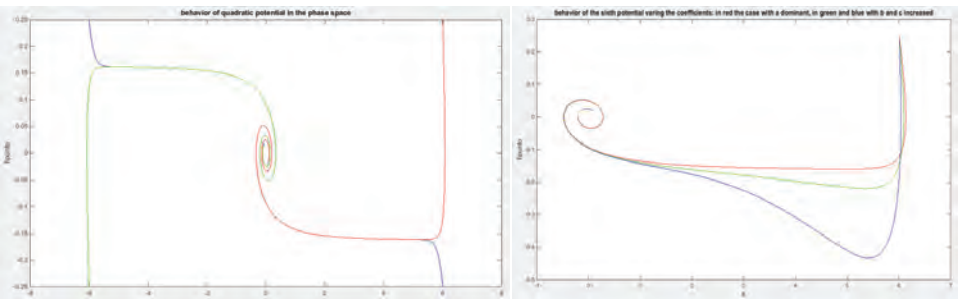


Fig. 2. Phase portrait in the $(\varphi, \dot{\varphi})$-plane of the dynamics of a scalar field for a quadratic potential (left) and for a sixth grade potential (right). For the last case, we plot the solutions for different values of the coefficients. It's clearly visible the pick, putting appropriate coefficients. This illustrates the mechanism of attraction toward the slow-roll solution.

## 3. Reheating Phase

During inflation, all the energy density is concentrated in the scalar field. This energy density can be used to be transformed in other form, for example radiation. This process of conversion is called **reheating**, while the temperature at which the process takes place is called **reheating temperature**. Let's see what happen during this process. Just before of the end of inflation the Universe is cold and

hibernated in a state of low entropy: in fact, $T(t) \sim a^{-1}(t)$, where $T(t)$ is the temperature as a function of the time. The scalar field starts to move quickly along the Hubble time-scale, and starts to oscillate around the minimum of the potential (see Figure 1 on the right). During this process, the inflaton can decay, producing light energetic particles. So, the interaction of the inflaton provides a *probability decay* $\Gamma$. This phase end when all the energy is in thermal equilibrium with the radiation.

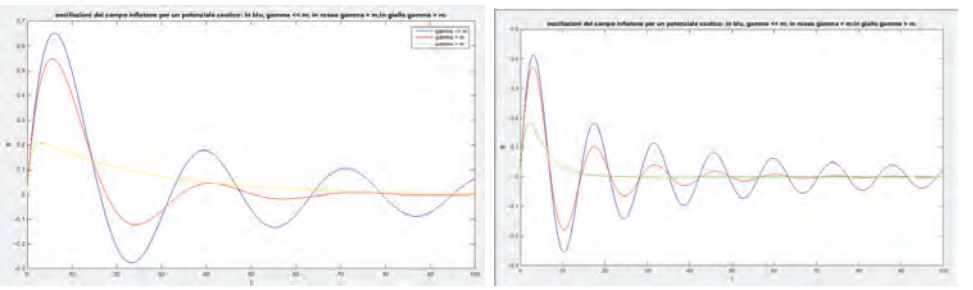


Fig. 3. Oscillations of the field f after inflation for quadratic potential (left) and for sixth grade potential (right) varying $\Gamma$. For the last case, we have a more oscillations because of the potential form.

The Figure 3 on the left shows the behavior of $\phi$ during reheating for quadratic potential. At the end of inflation $\Gamma$ is small, but later it increases, damping the oscillations. When $\Gamma \sim$ H, the reheating end. If $\Gamma <$ 3H, the reheating can't start, and we have a slow-roll phase. If we consider the sixth grade potential (eq. 1), we have the same behavior, but with a greater number of oscillations and a smaller amplitude (see Figure 3 on the right). This happens because of the potential form: it's less curved and wider in the minimum (Figure 1 on the left), so that the scalar field, having a less space, oscillates with greater frequency.

Now, we investigate the matter creation process during the reheating period. For this purpose, we consider the Universe as an open thermodynamics system, restricting our analysis to the case of a flat FRW metric. During the reheating, the constituents of the Universe are the inflationary scalar field, the ordinary matter, and the dark matter, assumed to be perfect fluid. So, the total energy density $\rho$ is:

$$\rho = \rho_\phi + \rho_{\mathrm{DM}} + \rho_{\mathrm{m}}$$

Let's see what are the equations which describe the reheating process. First of all, we have got to consider the dynamics of the scalar field and the matter. From the respective components of the *energy-momentum tensor* in a flat space-time, and applying them the energy conservation law, we have:

$$\dot{\phi}\ddot{\phi} = 3H\dot{\phi}^2 + V'(\phi)\dot{\phi} + \dot{\rho}_m + 3H(\rho_m + p_m) = 0.$$

In order to describe the transition process between the scalar field and the matter component, we introduce the friction term $\Gamma$, so we can decompose the previous equation in two equations:

$$\ddot{\phi} = 3H\dot{\phi} + V'(\phi) + \Gamma\dot{\phi} = 0 \quad \dot{\rho}_m + 3H(\rho_m + p_m) - \Gamma\dot{\phi}^2 = 0. \tag{2}$$

For the dark matter, we can consider a similar equation. Moreover, we need to consider the *balance equation of the particle flux vector* in the FRW geometry:

$$\Psi_m = \dot{n}_m + 3Hn_m$$

We assume that the function $\Psi$ is proportional to the energy density of the inflationary scalar field $\rho_\varphi$, and introducing the friction term $\Gamma$, we get:

$$\dot{n}_{\mathrm{m}} + 3Hn_{\mathrm{m}} = \frac{\Gamma_{\mathrm{m}}}{m_{\mathrm{m}}}\rho_\phi \tag{3}$$

We also consider the *Einstein gravitational field equations* in a flat FRW Universe, which, with inflationary scalar field, ordinary matter and dark matter fluids, becomes:

$$3H^2 = \frac{1}{m_{Pl}^2}(\rho_\phi + \rho_{\mathrm{DM}} + \rho_{\mathrm{m}}) \tag{4}$$

The particle creation is essentially determined by the energy density $\rho_\varphi$ and pressure $p_\varphi$ of the scalar field $\varphi$.

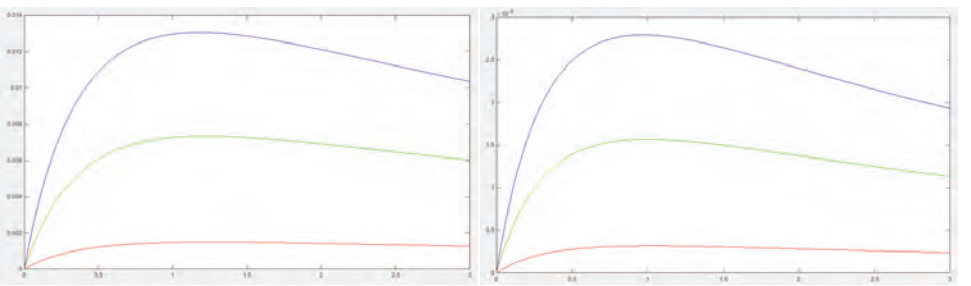


Fig. 4. Time variation of the dark matter energy density (left) and of the radiation (right) during the reheating period in the presence of a quadratic potential, varying the rate decay $\Gamma$.

The equations (2), (3) and (4) form a system of differential equations, which can be solved numerically. The Figure 4 shows the behavior of dark matter (on the left) and ordinary matter (on the right) for different values of $\Gamma$, while the scalar field decreases to zero. In fact, matter is created at the expense of the field. As we can see, it's possible to distinguish two phases: in the first, the matter energy density increases from zero to a maximum value; in the second, the matter energy density begins to decrease because of the expansion of the Universe, which becomes the dominant force. On the other hand, the energy density of the scalar field decreases

during the reheating phase, and tend to zero in the large time limit. During this phase, the Universe continues to accelerate because of the negative creation pressure, which induces an additional acceleration.

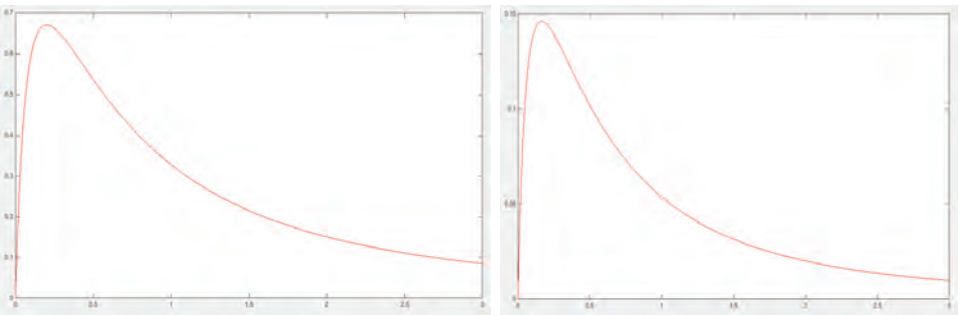


Fig. 5. The same as Figure 4, but with a sixth grade potential, with fixed Γ.

On the other hand, considering the sixth grade potential (see Figure 5, left and right), we observe a pick at the beginning of reheating and, after that, the density rapidly decreases. Also in this case, we can attribute this behavior to the potential form: there is a large particles production when reheating starts, but after the scalar field decays.

Finally, we must calculate the reheating temperature. At the end of inflation, $T_{reh}$ must have an upper limit because of the Hubble parameter, which decreases in time:

$$\frac{T_{max}}{M_{pl}} < \left(\frac{90}{8\pi^3 g_*}\right)^{1/4} \left(\frac{H_e}{M_{pl}}\right)^{1/2}$$

where $g$ is the number of the degrees of freedom (its value is ∼100). In fact, at the end of inflation, the slow-roll parameter $\varepsilon \sim 1$, from which we can estimate $H_e$ for different potentials $V(\varphi_e)$. Then, we obtain*:

| | |
|---|---|
| ✓ $T_{reh} \leq 2 \times 10^{15}$ GeV | *if* $\alpha = 2$, |
| ✓ $T_{reh} \leq 10^{15}$ GeV | *if* $\alpha = 4$, |
| ✓ $T_{reh} \leq 8 \times 10^{14}$ GeV | *if* $\alpha = 6$ |

So, the model of potential that warms the Universe better, when the reheating phase begins, is the quadratic potential. For comparison, the maximum temperature, independent from the kind of potential considered, is: $T_{max} \leq 10^{16}$ GeV.

*In this proceeding, we not consider the self-coupling of the inflaton. In the case of weakly-coupled inflaton, the Temperature is lower: $T_{reh} \leq 10^{12}$ GeV.

## 4. Conclusion

Summarizing, during the reheating phase the production of particles takes place, due to the decay of the inflationary scalar field. We got a plot describing the evolution of the matter in the time for the sixth grade potential using the Einstein equations and the components of the energy-momentum tensor, and we compared it with the case of the quadratic potential. So, we can say that a different function of the scalar field have some consequences on the inflation theory. The next goal is to understand what is the best potential for inflation. I mean, is the quadratic potential a good approximation to understood inflation (and the other terms of higher order are only little corrections)? Or do we need to add other terms yet? In fact, we can see the sixth grade potential as a *Taylor expansion of the hyperbolic cosine*.

## References


1. Montani G., "*Primordial cosmology*", World Scientific, 2011.
2. Montani G., "*Canonical quantum gravity*", World Scientific, 2014.
3. Gorbunov D. S., "*Introduction to the theory of the early Universe*", World Scientific, 2011.
4. Remmen G., "*Attractor solutions in scalar field cosmology*", Physical Review D 88, 083518, 2013.
5. Amin M., "*Nonperturbative dynamics of reheating after inflation: a review*", arxive 1410.3808v3, 2014.
6. Kofman L., "*Towards the theory of reheating after inflation*", arxive hep-ph/9704452v2, 1997.
7. Kofman L., "*Reheating after Inflation*", arxive hep-th/9405187v2, 1994.
8. Mukhanov V., "*Physical foundations of cosmology*", Cambridge, 2005.
9. Peter P., "*Primordial cosmology*", Oxford, 2009.
10. Linde A., "*Particle physics and inflationary cosmology*", CRC Press, 1989.

# Black holes with synchronised hair: connecting Kerr black holes with Bose Einstein condensates

Carlos A. R. Herdeiro*

*Centro de Astrofísica e Gravitação - CENTRA,*
*Departamento de Física, Instituto Superior Técnico - IST, Universidade de Lisboa - UL,*
*Avenida Rovisco Pais 1, 1049-001 Lisboa, Portugal*
** E-mail: carlosherdeiro@tecnico.ulisboa.pt*

I provide a brief summary of the work done on asymptotically flat black holes with synchronised hair. These black holes form families of solutions that interpolates between vacuum Kerr black holes and Bose-Einstein condensates of massive bosonic fields.

*Keywords*: Black holes, fundamental fields.

## 1. Testing the Kerr hypothesis

The *Kerr hypothesis* is the conjecture that astrophysical black holes, when near equilibrium, are approximately described by the Kerr metric.[1] This hypothesis is motivated by the electro-vacuum uniqueness theorems,[2] but it is a conjecture when considering arbitrary sorts of matter or fundamental fields.[3]

Testing, theoretically, the Kerr hypothesis amounts to keep some healthy skepticism about this idea, which, albeit not in tension with any observation at present,[4] is far from proved in any rigorous theoretical or observational sense. In particular one may ask: 1) are all astrophysical black holes in equilibrium *exactly* represented by the Kerr metric? or 2) Are *all* these black hole of the same type? or, allowing for an even more dramatic possibility, 3) Are all these astrophysical black hole *candidates* really black holes?

## 2. Good models

The Kerr solution is not just the unique theoretical vacuum black hole model; it is a *good* model. One may postulate some broad theoretical criteria for a good theoretical model of a compact object: 1) It should appear in a well motivated and consistent physical model (at least as an effective field theory); 2) It should have a dynamical formation mechanism; 3) It should be *sufficiently* stable. Here the key word is "sufficiently". Indeed absolute stability is not mandatory for physical relevance, that always hinges on a competition of timescales for any putative instability as compared to other relevant astrophysical processes. As an extreme example, an astrophysical object that has a slowly growing instability, with an e-folding time comparable to the Hubble time, is, for all astrophysical processes, effectively stable.[5] For Kerr, 1) is General Relativity (GR), 2) is gravitational collapse and 3) has been established at the level of mode stability.[6]

To match observations, moreover, the good model must give rise to the right phenomenology, including all electromagnetic observables (X-ray spectrum, shadows, QPOs, star orbits, *etc*), for which, however, current data is not particularly

constraining and the need for astrophysical modelling often yields degeneracies that prevent significant constraints. In this respect the gravitational wave channel, being more pristine, is also more promising. But at the moment we can only expect data for events with a limited range of black hole masses (say, between 1-100 solar masses). LISA will, of course, improve this prospect.

As mentioned above, there is no clear tension between the Kerr model and current observations. Still, given the precision and spread of observations will increase in the next decades, it is certainly worth asking if there are other theoretically sound models that have different "good" black holes, and how different is their phenomenology.

## 3. No-hair theorems

There are two meanings for "black hole no-hair theorems" in the literature. Firstly, the uniqueness theorems[2] are often called no-hair theorems. This falls under the rationale that the uniqueness theorems show electro-vacuum, physical, single black hole spacetimes have essentially no structure/no multipolar freedom/no hair. Despite the fact that Kerr black holes have a non-trivial multipolar structure,[7] this structure is completely determined by the mass and angular momentum, the two only macroscopic degrees of freedom they are allowed to have. Secondly, in models *beyond* electrovacumm (*i.e.* with more general matter content and/or modified gravitational field equations) one can establish (typically partial) results on the inexistence of new black hole solutions with other degrees of freedom rather than those associated to Gauss laws, and hence gauge symmetries, such as total mass, angular momentum (and electric-type charges). Many such no-hair theorems, in models with scalar fields for instance, have been established.[8] The scalar field case is particularly appealing as a test ground, since scalar fields are technically simple and are well motivated in models beyond the standard model of particle physics (string theory, supersymmetry, multi-Higgs, Kaluza-Klein, *etc*). But many examples of "hairy" black holes have also been found, even in models where no-hair theorems exist, by circumventing one or more of the assumptions. Typically these models violate either explicitly or effectively energy conditions. A different type of mechanism that has been found to allow for black hole hair is the *synchronisation mechanism*.

## 4. The synchronisation mechanism

In 2014, a new type of asymptotically flat, black hole spacetimes with a new macroscopic degree of freedom was constructed.[9] The model was Einstein's gravity minimally coupled to a massive complex scalar field. The scalar field has an associated global charge, a Noether charge, but since the scalar field is not gauged, this charge does not have an associated Gauss law. Thus, it provides the black hole with *hair*. Physically, if the scalar field falls into the black hole, this charge vanishes (unlike,

say, electric charge, due to the Gauss law). Thus, the scalar field needs to be in a non-linear equilibrium with the black hole, for this Noether charge to be non-trivial.

This example of hairy black holes occurs in GR, not modified gravity. The matter obeys all relevant energy conditions. The black holes are asymptotically flat, regular on and outside an event horizon, they continuously connect to the Kerr solution, to which they reduce in the limit where the scalar field vanishes. They also connect continuously to relativistic Bose-Einstein condensates called boson stars,[10] to which they reduce in the limite where the horizon size is reduced to zero. They have an independent scalar charge which is primary hair.[8] They can form dynamically and be sufficiently long lived (see below). Although these solutions have been constructed numerically, an existence proof has been provided.[11] Moreover, more than an example, these solutions provide a mechanisms for endowing black holes with hair. Generalisations with self interacting scalar hair,[12] with Proca hair,[13] Kerr-Newman black holes with synchronised hair,[14] higher dimensional Myers-Perry black holes with this sort of synchronised hair,[15,16] also in AdS,[17] excited states[18] and generalisations with higher azimuthal winding numbers[19] are just some of the examples considered so far.

## 5. The synchronisation mechanism: physics rationale

The physical understanding for the existence of Kerr black holes in equilibrium with a stationary scalar field configuration can be obtained by a linear analysis. One considers the massive Klein-Gordon equation in the background of a Kerr black hole. Separation of variables can be achieved in Boyer-Lindquist coordinates $(t, r, \theta, \phi)$[20], by taking a scalar field ansatz of the sort

$$\Phi = f(r, \theta) e^{-i(\omega t + m\phi)} \ , \tag{1}$$

where $\omega$ is the frequency of the mode and $m$ its azimuthal winding number, and one ends up with solving a radial equation which can be cast in a Schrödinger-like form. Unlike the Hydrogen atom problem of elementary quantum mechanics, however, the generic problem in the Kerr black hole does not admit bound states solutions with a real frequency $\omega$. This is because the boundary condition at the event horizon (a purely ingoing wave) makes the Hamiltonian non-hermitian. In this generic case, only *quasi*-bound states are possible, with a complex frequency, wherein the imaginary part, $\mathcal{I}(\omega)$ is determining the decay (for $\mathcal{I}(\omega) < 0$) or growth rate (for $\mathcal{I}(\omega) > 0$) of the mode. But for the special case in which the black hole's horizon angular velocity $\Omega_H$ is synchronous with the phase angular velocity of the mode $\omega/m$, that is

$$\omega = m\Omega_H, \tag{2}$$

which is the *synchronisation condition*, $\mathcal{I}(\omega) = 0$ and true bound states, in the sense of quantum mechanics, exist. These are stationary clouds around Kerr black holes. It was S. Hod who first considered these clouds by specialising to the extremal

Kerr background,[21] in which case the radial equation becomes *precisely* the same confluent hypergeometric equation that arises in the case of the Hydrogen atom (without spin) in the Coulomb potential. This yields a beautiful and perfect analogy between atomic orbitals and stationary scalar clouds around Kerr black holes. For the non-extremal case the qualitative analogy remains.[9,22–25]

For $\omega > m\Omega_H$, $\mathcal{I}(\omega) < 0$; this is the decaying regime (the only one which is present in the Schwarzschild case). For $\omega < m\Omega_H$, $\mathcal{I}(\omega) > 0$; this is the growing or *superradiant regime*, in which the mode extracts rotational energy from the Kerr black hole.[26] Thus the stationary clouds exist as zero modes of the superradiant instability of Kerr black holes. Kerr black holes with scalar hair can be seen as the non-linear continuation of these scalar clouds, when they are made heavy and backreact.

How to construct such heavy scalar clouds in a rotating black hole background? It turns out that self-gravitating scalar field configurations of a scalar of the type (1) were already known in the literature: *boson stars*. So, these hairy black holes can be seen as a rotating boson star in non-linear equilibrium with a central black hole horizon.[27]

Let us remark on how the existence of these black holes with synchronised hair is compatible with well known no-scalar hair theorems. Symmetry non-inheritance between the matter field and the geometry allows them to circumvent the powerful Bekenstein's theorem[28] and rotation allows them to circumvent the Peña-Sudarksy theorem[29] that applies to spherical symmetry non-inheriting solutions.

## 6. Black holes with synchronised hair: phenomenological features

Since Kerr black holes with synchronised hair interpolate between the vacuum Kerr solution and boson stars, their phenomenology will be Kerr-like in the vicinity of the Kerr limit, and Kerr-unlike in the vicinity of the boson star limit. This is what has been observed, for instance, at the level of their shadows,[30,31] K$\alpha$-line in the $X$-ray spectrum,[32] star orbits and their possible relation to observable quasi-periodic-oscillations[33], or the morphology of accretion disks around them.[34] Consequently, current observations that are compatible with the Kerr model, can only rule out regions in the parameter space wherein deviations are large, but not regions in the parameter space where deviations are small, sufficiently close to the Kerr limit.

## 7. Black holes with synchronised hair: dynamical issues

It has been known since the 1970s[35–37] that in the presence of ultra-light massive bosonic fields, astrophysical Kerr black holes would be unstable against superradiance. The end point of this instability has been a lingering question. Recent numerical evidence suggests that the instability saturates when the synchronisation

condition is met,[38] and that the resulting solutions are Kerr black holes with synchronised hair.[39,40] As pointed out before these developments, these configurations could not be the end point of the instability, since they are themselves prone to superradiant instabilities.[41] Recent progress managed to compute the time scale of the superradiant instability of a sample of hairy black hole solutions, all of them very close to the Kerr limit.[42] These results motivated a bound for the timescale of the leading superradiant instability of Kerr black holes with synchronised hair.[5] For regions of the parameter space this timescale can be larger than the Hubble scale. Moreover, the instability becomes weaker as the hairiness of the black holes increase, when a proper analysis is performed.[5] The upshot is that these solutions could form dynamically from the superradiant instability of Kerr and be sufficiently long lived, at least as superradiance goes.

## 8. Open issues and further remarks

The three theoretical broad criteria given above for a good model of compact object are met by Kerr black holes with syncronised hair: 1) They appear in a well motivated and consistent physical model, which is GR minimally coupled to innocuous, albeit exotic, matter; 2) They have a dynamical formation mechanism (superradiance); 3) They can be sufficiently stable, with an instability timescale for the only known instability – superradiance – above the Hubble time.

Of course, it remains to understand what the ultralight scalar (or vector) field could be from the view point of fundamental particle physics. Here, the issue of fuzzy dark matter may provide an interesting connection.[43,44] Or if there are other more catastrophic instabilities. In any case, these models illustrate that there exist fairly reasonable theoretical possibilities (albeit with admittedly exotic physics) of alternative models to the Kerr paradigm. Their study is a high risk, high gain endeavour: producing detailed phenomenology will constrain the model and the corresponding exotic physics or, in the best case scenario, provide a smoking gun to this new physics. But above all, testing the Kerr hypothesis by looking for alternative models represents scientific open mindedness.

## Acknowledgments

I extremely grateful to all my co-authors in the works reported herein, first and foremost Eugen Radu. This work is supported by CENTRA (FCT) strategic project UID/FIS/00099/2013, the project PTDC/FIS-OUT/28407/2017, by the European Union's Horizon 2020 research and innovation (RISE) programmes H2020-MSCA-RISE-2015 Grant No. StronGrHEP-690904 and H2020-MSCA-RISE-2017 Grant No. FunFiCO-777740. The authors would like to acknowledge networking support by the COST Action CA16104.

## References


1. R. P. Kerr, Phys. Rev. Lett. **11** (1963) 237.
2. P. T. Chrusciel, J. Lopes Costa and M. Heusler, Living Rev. Rel. **15** (2012) 7 [arXiv:1205.6112 [gr-qc]].
3. R. Ruffini and J. A. Wheeler, Phys. Today **24** (1971) no.1, 30.
4. E. Berti *et al.*, Class. Quant. Grav. **32** (2015) 243001 [arXiv:1501.07274 [gr-qc]].
5. J. C. Degollado, C. A. R. Herdeiro and E. Radu, Phys. Lett. B **781** (2018) 651 [arXiv:1802.07266 [gr-qc]].
6. B. F. Whiting, J. Math. Phys. **30** (1989) 1301.
7. R. O. Hansen, J. Math. Phys. **15** (1974) 46.
8. C. A. R. Herdeiro and E. Radu, Int. J. Mod. Phys. D **24** (2015) no.09, 1542014 [arXiv:1504.08209 [gr-qc]].
9. C. A. R. Herdeiro and E. Radu, Phys. Rev. Lett. **112** (2014) 221101 [arXiv:1403.2757 [gr-qc]].
10. F. E. Schunck and E. W. Mielke, Class. Quant. Grav. **20** (2003) R301 [arXiv:0801.0307 [astro-ph]].
11. O. Chodosh and Y. Shlapentokh-Rothman, Commun. Math. Phys. **356** (2017) no.3, 1155 [arXiv:1510.08025 [gr-qc]].
12. C. A. R. Herdeiro, E. Radu and H. Rúnarsson, Phys. Rev. D **92** (2015) no.8, 084059 [arXiv:1509.02923 [gr-qc]].
13. C. Herdeiro, E. Radu and H. Rúnarsson, Class. Quant. Grav. **33** (2016) no.15, 154001 [arXiv:1603.02687 [gr-qc]].
14. J. F. M. Delgado, C. A. R. Herdeiro, E. Radu and H. Runarsson, Phys. Lett. B **761** (2016) 234 doi:10.1016/j.physletb.2016.08.032 [arXiv:1608.00631 [gr-qc]].
15. Y. Brihaye, C. Herdeiro and E. Radu, Phys. Lett. B **739** (2014) 1 [arXiv:1408.5581 [gr-qc]].
16. C. Herdeiro, J. Kunz, E. Radu and B. Subagyo, Phys. Lett. B **748** (2015) 30 [arXiv:1505.02407 [gr-qc]].
17. O. J. C. Dias, G. T. Horowitz and J. E. Santos, JHEP **1107** (2011) 115 [arXiv:1105.4167 [hep-th]].
18. Y. Q. Wang, Y. X. Liu and S. W. Wei, arXiv:1811.08795 [gr-qc].
19. J. F. M. Delgado, C. A. R. Herdeiro and E. Radu, arXiv:1903.01488 [gr-qc].
20. D. R. Brill, P. L. Chrzanowski, C. Martin Pereira, E. D. Fackerell and J. R. Ipser, Phys. Rev. D **5** (1972) 1913.
21. S. Hod, Phys. Rev. D **86** (2012) 104026 Erratum: [Phys. Rev. D **86** (2012) 129902] [arXiv:1211.3202 [gr-qc]].
22. S. Hod, Eur. Phys. J. C **73** (2013) no.4, 2378 [arXiv:1311.5298 [gr-qc]].
23. S. Hod, Phys. Rev. D **90** (2014) no.2, 024051 [arXiv:1406.1179 [gr-qc]].
24. C. L. Benone, L. C. B. Crispino, C. Herdeiro and E. Radu, Phys. Rev. D **90** (2014) no.10, 104024 [arXiv:1409.1593 [gr-qc]].
25. S. Hod, JHEP **1701** (2017) 030 [arXiv:1612.00014 [hep-th]].

26. R. Brito, V. Cardoso and P. Pani, Lect. Notes Phys. **906** (2015) pp.1 [arXiv:1501.06570 [gr-qc]].
27. C. A. R. Herdeiro and E. Radu, Int. J. Mod. Phys. D **23** (2014) no.12, 1442014 [arXiv:1405.3696 [gr-qc]].
28. J. D. Bekenstein, Phys. Rev. Lett. **28** (1972) 452.
29. I. Pena and D. Sudarsky, Class. Quant. Grav. **14** (1997) 3131.
30. P. V. P. Cunha, C. A. R. Herdeiro, E. Radu and H. F. Runarsson, Phys. Rev. Lett. **115** (2015) no.21, 211102 [arXiv:1509.00021 [gr-qc]].
31. P. V. P. Cunha, J. Grover, C. Herdeiro, E. Radu, H. Runarsson and A. Wittig, Phys. Rev. D **94** (2016) no.10, 104023 [arXiv:1609.01340 [gr-qc]].
32. Y. Ni, M. Zhou, A. Cardenas-Avendano, C. Bambi, C. A. R. Herdeiro and E. Radu, JCAP **1607** (2016) no.07, 049 [arXiv:1606.04654 [gr-qc]].
33. N. Franchini, P. Pani, A. Maselli, L. Gualtieri, C. A. R. Herdeiro, E. Radu and V. Ferrari, Phys. Rev. D **95** (2017) no.12, 124025 [arXiv:1612.00038 [astro-ph.HE]].
34. S. Gimeno-Soler, J. A. Font, C. Herdeiro and E. Radu, Phys. Rev. D **99** (2019) no.4, 043002 [arXiv:1811.11492 [gr-qc]].
35. W. H. Press and S. A. Teukolsky, Nature **238** (1972) 211.
36. T. J. M. Zouros and D. M. Eardley, Annals Phys. **118** (1979) 139.
37. S. L. Detweiler, Phys. Rev. D **22** (1980) 2323.
38. W. E. East and F. Pretorius, Phys. Rev. Lett. **119** (2017) no.4, 041101 [arXiv:1704.04791 [gr-qc]].
39. C. A. R. Herdeiro and E. Radu, Phys. Rev. Lett. **119** (2017) no.26, 261101 [arXiv:1706.06597 [gr-qc]].
40. S. Dolan, APS Physics **10** (2017) 83.
41. C. Herdeiro and E. Radu, Phys. Rev. D **89** (2014) no.12, 124018 [arXiv:1406.1225 [gr-qc]].
42. B. Ganchev and J. E. Santos, Phys. Rev. Lett. **120** (2018) no.17, 171101 [arXiv:1711.08464 [gr-qc]].
43. A. Surez, V. H. Robles and T. Matos, Astrophys. Space Sci. Proc. **38** (2014) 107 [arXiv:1302.0903 [astro-ph.CO]].
44. L. Hui, J. P. Ostriker, S. Tremaine and E. Witten, Phys. Rev. D **95** (2017) no.4, 043541 [arXiv:1610.08297 [astro-ph.CO]].

# Adiabatic perturbations are not always conserved: the case of global adiabaticity

Antonio Enea Romano

*Instituto de Fisica, Universidad de Antioquia, A.A.1226, Medellin, Colombia*
*and*
*Center for Gravitational Physics, Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*

In the context of single-field inflation, the conservation of the curvature perturbation on comoving slices, $\mathcal{R}_c$, on super-horizon scales is one of the assumptions necessary to derive the consistency condition between the squeezed limit of the bispectrum and the spectrum of the primordial curvature perturbation. However, the conservation of $\mathcal{R}_c$ holds only after the perturbation has reached the adiabatic limit where the constant mode of $\mathcal{R}_c$ dominates over the other (usually decaying) mode. In this case, the non-adiabatic pressure perturbation defined in the thermodynamic sense, $\delta P_{nad} \equiv \delta P - c_w^2 \delta\rho$ where $c_w^2 = \dot{P}/\dot{\rho}$, usually becomes also negligible on superhorizon scales. Therefore one might think that the adiabatic limit is the same as thermodynamic adiabaticity. This is in fact not true. In other words, thermodynamic adiabaticity is not a sufficient condition for the conservation of $\mathcal{R}_c$ on super-horizon scales. In this paper, we consider models that satisfy $\delta P_{nad} = 0$ on all scales, which we call global adiabaticity (GA), which is guaranteed if $c_w^2 = c_s^2$, where $c_s$ is the phase velocity of the propagation of the perturbation. A known example is the case of ultra-slow-roll(USR) inflation in which $c_w^2 = c_s^2 = 1$. In order to generalize USR we develop a method to find the Lagrangian of GA K-inflation models from the behavior of background quantities as functions of the scale factor. Applying this method we show that there indeed exists a wide class of GA models with $c_w^2 = c_s^2$, which allows $\mathcal{R}_c$ to grow on superhorizon scales, and hence violates the non-Gaussianity consistency condition.

## 1. Introduction

A period of accelerated expansion during the early stages of the evolution of the Universe, called inflation,[1–3] is able to account for several otherwise difficult to explain features of the observed Universe such as the high level of isotropy of the CMB[4] radiation and the small value of the curvature. Some of the simplest inflationary models are based on a single slowly-rolling scalar field, and they are in good agreement with observations. It is commonly assumed in slow-roll models that adiabaticity in the thermodynamic sense, $\delta P_{nad} \equiv \delta P - c_w^2 \delta\rho = 0$ where $c_w^2 = \dot{P}/\dot{\rho}$, implies the conservation of the curvature perturbation on uniform density slices $\zeta$, and hence the conservation of the curvature perturbation on comoving slices $\mathcal{R}_c$, on super-horizon scales.

In[5] it was shown that there can be important exceptions, i.e. in some cases thermodynamic adiabaticity does not necessarily imply the super-horizon conservation of $\mathcal{R}_c$ and $\zeta$, and that they can differ from each other. This can happen even for models in which $c_w^2 = c_s^2$. Here $c_s$ is the speed of propagation of the curvature perturbation. It turns out that it may be defined as $c_s^2 \equiv (\delta P/\delta\rho)_c$, where the suffix "c" means a quantity evaluated on comoving slices defined by $\delta T_0^i = 0$ (or equivalently slices on which the scalar field is homogeneous). An example is ultra-slow-roll

(USR) inflation , in which the flat potential $V(\phi) = V_0$ yields exact adiabaticity $\delta P_{nad} = 0$ on all scales. USR inflation could in principle last for 60 e-folds, but then it would be difficult to make it consistent with observation. Alternatively, one can study models in which a USR phase is followed by a conventional slow-roll phase , at which stage $\mathcal{R}_c$ becomes conserved. In USR inflation, both $\mathcal{R}_c$ and $\zeta$ exhibit super-horizon growth but their behavior is very different from each other. As it has been stressed in , the non-freezing of $\mathcal{R}_c$ has important phenomenological consequences. Since the freezing of $\mathcal{R}_c$ on superhorizon scales is a necessary ingredient for Maldacena's consistency relation to hold, models that do not conserve $\mathcal{R}_c$ can actually violate that consistency condition. We note as well that another consequence of superhorizon growth of $\mathcal{R}_c$ is that it should be evaluated at the end of inflation, instead of just after horizon crossing .

In this paper focusing on K-inflation, i.e., Einstein-scalar models with a general kinetic term, we explore in a general way other single field models which have $c_w^2 = c_s^2$, hence satisfy $\delta P_{nad} = 0$ on all scales which we call globally adiabatic (GA), but which may not conserve $\mathcal{R}_c$. We find a generalization of the USR model.

The method we adopt is based on establishing a general condition for the non-conservation of $\mathcal{R}_c$ in terms of the dependence of the background quantities, in particular the slow-roll parameter $\epsilon \equiv -\dot{H}/H^2$ and the sound velocity $c_s$, on the scale factor $a$.

We first derive the necessary condition for the comoving curvature perturbation $\mathcal{R}_c$ to grow on superhorizon scales. Next we determine $\rho(a)$ and $P(a)$ by solving the continuity equation. Then using the equivalence between barotropic fluids and $K$-inflationary models which satisfy the condition $c_w^2 = c_s^2$ , we determine the corresponding Lagrangian for the equivalent scalar field model. Using this method we obtain a new class of GA scalar field models which do not conserve $\mathcal{R}_c$.

Throughout the paper we denote the proper-time derivative by a dot ($\dot{} = d/dt$), the conformal-time derivative by a prime ($\prime = d/d\eta = a\, d/dt$) and the Hubble expansion rates in proper and conformal times by $H = \dot{a}/a$ and $\mathcal{H} = a'/a$, respectively. We also use the terminology "adiabaticity" for thermodynamic adiabaticity $\delta P_{nad} = 0$ throughout the paper.

## 2. Conservation of $\mathcal{R}_c$ and global adiabaticity

We set the perturbed metric as

$$ds^2 = a^2\Big[-(1+2A)d\eta^2 + 2\partial_j B dx^j d\eta + \big\{\delta_{ij}(1+2\mathcal{R}) + 2\partial_i\partial_j E\big\}dx^i dx^j\big\}\Big]\,. \tag{1}$$

In[5] it was shown that independently of the gravity theory and for generic matter the energy-momentum conservation equations imply

$$\delta P_{nad} = \left[\left(\frac{c_w}{c_s}\right)^2 - 1\right](\rho + P)A_c\,. \tag{2}$$

In the case of general relativity, the additional relation $A_c = \dot{\mathcal{R}}_c/H$ gives an important relation for the time derivative of $\mathcal{R}_c$

$$\delta P_{nad} = \left[\left(\frac{c_w}{c_s}\right)^2 - 1\right](\rho + P)\frac{\dot{\mathcal{R}}_c}{H}. \tag{3}$$

The non-adiabatic pressure perturbation is given according to its thermodynamics definition

$$\delta P_{nad} \equiv \delta P - c_w^2 \delta\rho. \tag{4}$$

This definition of $\delta P_{nad}$ is important because it is gauge invariant and $\delta P_{nad} = \delta P_{ud}$, where $\delta P_{ud}$ is the pressure perturbation on uniform density ($\delta\rho = 0$) slices. It appears in the equation for the curvature perturbation on uniform density slices $\zeta \equiv \mathcal{R}_{ud}$ obtained from the energy conservation law,

$$\zeta' = -\frac{\mathcal{H}\delta P_{nad}}{(\rho + P)} + \frac{1}{3}\overset{(3)}{\Delta}(v - E')_{ud} \tag{5}$$

where $v$ is the 3-velocity potential ($v = \delta\phi/\phi'$ for a scalar field). In general, the curvature perturbations on uniform density and comoving slices are related as

$$\zeta = \mathcal{R}_c + \frac{\delta P_{nad}}{3(\rho + P)(c_s^2 - c_w^2)}\,. \tag{6}$$

A common interpretation of these equations is that when when $\delta P_{nad} \approx 0$ with $c_w^2 \neq c_s^2$, $\zeta$ and $\mathcal{R}_c$ are approximately equal because of eq. (6), and they are both approximately conserved on super-horizon scales because of eq. (3). This is in agreement with the well-known coincidence of $\zeta$ and $\mathcal{R}_c$ on super-horizon scales for slow roll-models in general relativity, since in this case $c_s \neq c_w$ and $\delta P_{nad} \approx 0$ on superhorizon scales.

The equation (3) is the key relation to understand how $\mathcal{R}_c$ depends on the non-adiabatic pressure $\delta P_{nad}$. First of all let us note that this equation is valid on any scale. The advantage of it with respect to eq. (5) is that it does not involve gradient terms, so it allows us to directly relate $\delta P_{\rm nad}$ to $\dot{\mathcal{R}}_c$ if $c_w^2 \neq c_s^2$, while in eq. (5) $\dot{\zeta}$ depends on spatial gradients, which in the case of USR are not negligible on super-horizon scales[5]. This explains why in USR in which $c_w^2 = c_s^2 = 1$, both $\mathcal{R}_c$ and $\zeta$ are not conserved despite $\delta P_{nad} = 0$.

It should be noted here that for slow-roll attractor models $c_w^2 \neq c_s^2$ in general, and $\mathcal{R}_c$ is time-varying on sub-horizon scales. This implies that the non-adiabatic pressure perturbation $\delta P_{nad}$ on sub-horizon scales is not zero. In other words, the attractor models are adiabatic only on super-horizon scales, and we call these models super-horizon adiabatic (SHA).

From eq. (3) we can immediately deduce that in general relativity there are two possible scenarios for the non-conservation of $\mathcal{R}_c$,

$$\begin{aligned} &(1) \quad c_s^2 = c_w^2\,, \quad \delta P_{nad} = 0\,, \\ &(2) \quad c_s^2 \neq c_w^2\,, \quad \delta P_{nad} \neq 0\,. \end{aligned} \tag{7}$$

Here we focus on the first case. It is trivial to see that because of the gauge invariance of $\delta P_{nad}$ the condition $c_w^2 = c_s^2$ automatically implies $\delta P_{nad} = 0$. The models satisfying the condition $c_s^2 - c_w^2 = \delta P_{nad} = 0$ are adiabatic on any scale, and because of this we call them globally adiabatic (GA). In GA models an explicit calculation can reveal the super-horizon behavior of $\mathcal{R}_c$, and $\zeta$, as was shown in[5] in the case of USR. Below, we develop an inversion method to find a new class of models that violate the conservation of $\mathcal{R}_c$ without solving the perturbations equations.

## 3. Globally adiabatic K-essence models

The condition $c_w^2 = c_s^2$ has been studied in the context of K-inflation[6] described by the action ($X \equiv -g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi/2$)

$$S = \frac{1}{2}\int d^4x\sqrt{-g}\left[M_{Pl}^2 R + 2P(X,\phi)\right] , \tag{8}$$

and it was shown that it is satisfied by scalar field models with the Lagrangian of the form,

$$P(X,\phi) = u(Xg(\phi)) \equiv u(Y) \tag{9}$$

where $u$ and $g$ are arbitrary functions. These models are equivalent to a barotropic perfect fluid, i.e. a fluid with equation of state $P(\rho)$. We note again that these models *are adiabatic on any scale* (GA), contrary to the slow-roll attractor models, which are adiabatic *only on super-horizon scales* (SHA). The fact that they are mutually exclusive can be readily seen by considering the hypothetical case of $\delta P_{nad} = 0$ *and* $c_w^2 \neq c_s^2$. In this case eq. (3) which is valid on any scale would mean $\mathcal{R}_c$ should be *frozen on all scales*. In contrast, the condition $c_w^2 = c_s^2$ allows for the curvature perturbation to *evolve* both on *sub-horizon* and *super-horizon* scales.

In[6] it was shown that is possible to associate any barotropic perfect fluid with an equivalent K-inflation model according to

$$2\int^P \frac{du}{F(u)} = \log(Y) , \tag{10}$$

where $F(P) = \rho(P) + P$ and $Y = g(\phi)X$. These models are the ones which could violate the conservation of $\mathcal{R}_c$ for adiabatic perturbations, since they satisfy $c_w^2 = c_s^2$. It is noted of course that the global adiabaticity is not the sufficient condition for the non-conservation of $\mathcal{R}_c$. Not all GA models violate the conservation of $\mathcal{R}_c$ on super-horizon scales.

## 4. General conditions for super-horizon growth of $\mathcal{R}_c$

From the equation for the curvature perturbation on comoving slices,

$$\frac{\partial}{\partial t}\left(\frac{a^3\epsilon}{c_s^2}\frac{\partial}{\partial t}\mathcal{R}_c\right) - a\epsilon\Delta\mathcal{R}_c = 0 , \tag{11}$$

we can deduce, after re-expressing the time derivative in terms of the derivative respect to the scale factor $a$, that on superhorizon scales there is (apart from a constant solution) a solution of the form,

$$\mathcal{R}_c \propto \int^a \frac{da}{a} f(a)\,; \quad f(a) \equiv \frac{c_s^2(a)}{Ha^3\epsilon(a)}\,, \tag{12}$$

where we have introduced the function $f(a)$ for later convenience. In conventional slow-roll inflation $c_s^2$ and $\epsilon$ are both slowly varying, hence the integral rapidly approaches a constant, rendering $\mathcal{R}_c$ conserved. The time dependent part of the above solution corresponds to the decaying mode.

The necessary and sufficient condition for super-horizon freezing is that there exists some $\delta > 0$ for which

$$\lim_{a\to\infty} a^\delta f(a) = 0. \tag{13}$$

By definition of inflation, $H$ must be sufficiently slowly varying; $\epsilon = -\dot{H}/H^2 \ll 1$. So we may neglect the time dependence of $H$ in eq. (12) at leading order, while $\epsilon$ and $c_s^2$ may vary rapidly in time. For models for which $\epsilon \approx a^{-n}$ and $c_s^2 \approx a^q$ we get

$$f \propto a^{q+n-3}\,, \tag{14}$$

hence the condition for freezing is

$$q + n - 3 < 0\,. \tag{15}$$

If this condition is violated, i.e. $q + n - 3 \geq 0$, then the solution (12) will grow on super-horizon scales. This happens for example in USR, which corresponds to $c_s^2 = 1$ and $\epsilon \propto a^{-6}$, i.e. $q = 0$, and $n = 6$. In general, we expect that $q$ would not become very large. This implies $\epsilon$ should decrease sufficiently rapidly. Conversely, if $\epsilon$ decreases sufficiently rapidly, then the growth of $\mathcal{R}_c$ on superhorizon scales will follow.

## 5. Conclusions

In conventional slow-roll models, one has $c_s^2 \neq c_w^2$ and the superhorizon freezing of $\mathcal{R}_c$ can be understood as a result of $\delta P_{\rm nad} \approx 0$ on superhorizon scales. When $c_s^2 = c_w^2$, one has $\delta P_{\rm nad} = 0$ on all scales, but following eq. (3) this does not constrain the superhorizon behaviour of $\mathcal{R}_c$ anymore. This behaviour now follows from $\mathcal{R}_c$'s equation of motion given in eq. (11), and the condition for superhorizon freezing is given in eq. (13). Violation of this condition leads to superhorizon growth of $\mathcal{R}_c$.

We have developed a method to construct the Lagrangian of a K-essence globally adiabatic (GA) model by specifying the behavior of background quantities such as $\epsilon\rho$ where $\epsilon$ is the slow-roll parameter, using the equivalence between barotropic fluids and GA K-essence models. We have applied the method to find the equations of state of the fluids and derive the Lagrangian of the equivalent single scalar field models. Interestingly, we have found that the requirement to avoid the gradient

instability, i.e., $c_s^2 > 0$ is almost identical to the condition for the non-conservation on superhorizon scales.

The advantage of our approach is that we did not have to solve any perturbation equation explicitly. We have begun from requiring some behaviour for $\epsilon$, or for $b \equiv 2\epsilon\rho$, and have then used our inversion method to find the Lagrangian that produces that behaviour.

We have shown that the main difference between attractor models and GA models is that the latter are adiabatic on all scales, while attractor models are approximately adiabatic in the sense of $\delta P_{nad} = 0$ only on super-horizon scales and $c_w^2 \neq c_s^2$.

The detailed study of the new models found in this paper will be done in a separate upcoming work but we can already predict that they can be compatible with observational constraints on the spectral index thanks to the extra parameter $n$ which is not present in USR. Furthermore they can violate the Maldacena's consistency condition and consequently produce large local shape non-Gaussianity.

In the future it will be interesting to apply the inversion method we have developed to other problems related to primordial curvature perturbations, or to develop a similar method for the adiabatic sound speed as function of the scale factor.

## Acknowledgements

This work was supported by the Dedicacion exclusica and Sostenibilidad programs at UDEA, the UDEA CODI project IN10219CE and 2015-4044, and Colciencias mobility project COSOMOLOGY AFTER BICEP.

## References

1. A. D. Linde, "A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems," Phys. Lett. B **108**, 389 (1982).
2. A. Albrecht and P. J. Steinhardt, "Cosmology for Grand Unified Theories with Radiatively Induced Symmetry Breaking," Phys. Rev. Lett. **48**, 1220 (1982).
3. A. A. Starobinsky, "Dynamics of Phase Transition in the New Inflationary Universe Scenario and Generation of Perturbations," Phys. Lett. B **117**, 175 (1982).
4. P. A. R. Ade *et al.* [Planck Collaboration], "Planck 2015 results. XX. Constraints on inflation," arXiv:1502.02114 [astro-ph.CO].
5. A. E. Romano, S. Mooij and M. Sasaki, "Adiabaticity and gravity theory independent conservation laws for cosmological perturbations," Phys. Lett. B **755**, 464 (2016) [arXiv:1512.05757 [gr-qc]].
6. F. Arroja and M. Sasaki, "A note on the equivalence of a barotropic perfect fluid with a K-essence scalar field," Phys. Rev. D **81**, 107301 (2010) [arXiv:1002.1376 [astro-ph.CO]].

# No-Go theorems for ekpyrosis from ten-dimensional supergravity

K. Uzawa

*Department of Physics, School of Science and Technology, Kwansei Gakuin University, Sanda, Hyogo 669-1337, Japan*
*E-mail: kunihito.uzawa@gmail.com*

In this work, we discuss whether the new ekpyrotic scenario can be embedded into ten-dimensional supergravity. We use that the scalar potential obtained from flux compactifications of type II supergravity with sources has a universal scaling with respect to the dilaton and the volume mode. Similar to the investigation of inflationary models, we find very strong constraints ruling out ekpyrosis from analysing the fast-roll conditions. We conclude that flux compactifications tend to provide potentials that are neither too flat and positive (inflation) nor too steep and negative (ekpyrosis).

*Keywords*: Ekpyrosis; Type II string theory; Flux compactifications; Moduli.

## 1. Introduction

The strong no-go theorems which exclude tree-level de Sitter compactifications under a few simple assumptions with or without negative tension objects such as orientifold planes have been much explored because of the possible cosmological and phenomenological interests. However, the no-go theorems for ekpyrotic scenario which is alternative to inflation model in string theory is much less extensive. One motivation for the present work is to improve this situation. Since orientifold planes are a common ingredient in phenomenologically interesting type II string theory, it seems natural to explore the possibility of treelevel de Sitter vacua or inflation models in type II string theory with orientifolds. On the other hand, there is no constraint for ekpyrotic scenario in type II string theory at present.

The ekpyrosis inspired by string theory and brane world model suggests alternative solutions to the early universe puzzles such as inflation and dark energy, and assumes that two four-dimensional boundary branes which are located at the endpoints of orbifold in the higher-dimensional bulk spacetime[1–6]. For brane world picture, all forces except for gravity are localized on the branes while gravity can propagate freely in the bulk. When we assume that there is an attractive force between two branes, these branes approach to each other, which gives big bang. Since the big bang is described as a collision of branes there is not the beginning of time in the ekpyrotic scenario. Although two branes move through each other at once after collision, we can get a model so that the branes become closer again.

The motion of branes is described by the potential of scalar field in a four-dimensional effective theory. Non-perturbative effects result in a potential which attracts one brane towards the other brane[1]. In order to resolve a horizon and flatness problem, the potential during a period of slow contraction before the big bang is negative and steeply falling. There was plenty of time before the big bang for the universe to be in causal contact over large regions, and in this way the horizon and flatness problem is automatically solved[4]. Then, the usual statement

of the ekpyrosis is that the universe slowly contracts before the big bang with the equation of state $w = p/\rho \gg 1$. The scalar potential is expected to turn up towards zero at large negative value of scalar field in the ekpyrotic or cyclic models.

As two branes approach each other, the branes are rippled because of generating quantum fluctuations. Since the collision of branes cannot be happened at exactly the same time, the branes collide slightly earlier or later in some places. Our universe thus has a little bit more time to cool or hot. When we consider the curvature perturbation in the ekpyrotic scenario, it occurs a strong blueshift, which is in sharp contradiction with the small redshift of the scalar fluctuation in the CMB[7–15]. The occurrence of blueshift scalar perturbation is nonetheless a minor flaw of ekpyrotic theory which can be easily corrected. Hence, new ekpyrotic theory considers multiple scalar field and successfully generated a scalar spectrum which is scale invariant and slightly redshifted[16–20]. Producing primordial gravitational waves sourced by the gauge field in the ekpyrotic scenario was also studied in Ref. 21.

It is the purpose of this work to give a No-Go theorem of the ekpyrotic scenario in a ten-dimensional supergravity model which is low energy limit of a string theory. We study the dynamics of two scalar fields in the four-dimensional effective theory after a compactification in string theory. There are a dilaton and the volume modulus of the internal manifold in the effective theory, which is a four-dimensional theory of gravity minimally coupled to two-scalar fields. We will derive the two moduli fields with negative exponential potentials. Since this potential is steep, the scalar field should be satisfied by "fast-roll" condition instead of slow-roll parameter in inflation model[22,23]. Although the terms coupling the scalar fields to the scalar curvature of the internal space and orientifold plane contribute the negative value to the potential in the string theory, we find that the potential does not satisfy the fast-roll condition in general. Therefore, it is not possible for us to realize the ekpyrotic phase in a string theory.

Section 2 describes the potential of scalar field for ekpyrotic scenario and the way it derives the four-dimensional effective theory. We discuss the approach to the effective action in more detail. The No-Go theorem of the ekpyrosis thus given by the string theory is discussed. We also investigate the detailed properties of these models, their embedding in a string theory and their viability. For simplicity, we do not consider D-branes and the associated moduli except for the volume modulus (breathing mode) of internal space although the analysis would not be different. Finally, section 3 provides a brief summary and an outlook to future developments.

## 2. No-Go theorem of the ekpyrotic scenario in the type II theory

In this section, we consider compactifications of the type II theory to four-dimensional spacetime on compact manifold Y. The ten-dimensional low-energy

effective action for the type II theory takes the form[24]

$$S = \frac{1}{2\bar{\kappa}^2}\int d^{10}x\sqrt{-g}\left[\mathrm{e}^{-2\phi}\left(R + 4g^{MN}\partial_M\phi\partial_N\phi - \frac{1}{2}\left|H\right|^2\right) - \frac{1}{2}\sum_p \left|F_p\right|^2\right] - \sum_p \left(T_{\mathrm{D}p} + T_{\mathrm{O}p}\right)\int d^{p+1}x\sqrt{-g_{p+1}}\,\mathrm{e}^{-\phi}\,, \tag{1}$$

where $\bar{\kappa}^2$ is the ten-dimensional gravitational constant, $g$, $R$ denote the determinant, the Ricci scalar with respect to the ten-dimensional metric $g_{MN}$, respectively, $\phi$ is the scalar field, $H$ is the NS-NS 3-form field strength, $F_p$ are the R-R $p$-form field strengths ($p = 0, 2, 4, 6, 8$ for type IIA, and $p = 1, 3, 5, 7, 9$ for type IIB) that are sourced by D-branes, and $T_{\mathrm{D}p}$ ($T_{\mathrm{O}p}$) is the D$p$-brane (O$p$-plane) charge and tension. Although there are Chern-Simons terms in the ten-dimensional action, these are essentially independent of the dilaton and the scale of the background metric. Hence, we will not consider them.

To compactify the theory to four dimensions, we consider the a metric ansatz of the form[24]

$$ds^2 = q_{\mu\nu}dx^\mu dx^\nu + g_{ij}dy^i dy^j = q_{\mu\nu}dx^\mu dx^\nu + \rho\, u_{ij}(\mathrm{Y})dy^i dy^j\,, \tag{2}$$

where $\rho$ is breathing mode (volume modulus of the compact space), $x^\mu$ denote the coordinates of four-dimensional spacetime, $y^i$ are local coordinates on the internal space Y, and $q_{\mu\nu}$, $u_{ij}(\mathrm{Y})$ are the metrices of four-dimensional spacetime, six-dimensional internal space, respectively. We assume that $q_{\mu\nu}$, $u_{ij}(\mathrm{Y})$ depend only on the coordinates $x^\mu$, $y^i$, respectively. Since we factored out the overall volume modulus $\rho$ of the internal space in the ten-dimensional metric (2), the modulus $\rho$ is related to the total physical volume of the internal space $v_6$ and the volume of Y space $v(\mathrm{Y})$ as

$$\rho = \left[\frac{v_6}{v(\mathrm{Y})}\right]^{1/3}, \qquad v_6 = \int d^6y\sqrt{g_6}\,, \qquad v(\mathrm{Y}) = \int d^6y\sqrt{u}\,, \tag{3}$$

where $g_6$ and $u$ denote the determinant of the metric $g_{ij}$ and $u_{ij}(\mathrm{Y})$, respectively. The volume modulus $\rho$ is chosen such that the metric $u_{ij}(\mathrm{Y})$ of the internal space is normalized $v(\mathrm{Y}) = 1$.

Now let us consider the four-dimensional effective action $S_{\mathrm{E}}$ in the Einstein frame after we integrate over the internal space Y. It is especially interesting to understand the dynamics of moduli at negative potential energy. If there are non-trivial fluxes in the background, one notes that these make uplifting the moduli potential to positive energy. In this work, we consider the moduli potential without field strengths and D-branes. Then we obtain the four-dimensional effective action in the Einstein frame:

$$S_{\mathrm{E}} = \int d^4x\sqrt{-\bar{q}}\left[\frac{1}{2\kappa^2}\bar{R} - \frac{1}{2}\bar{q}^{\mu\nu}\partial_\mu\bar{\tau}\,\partial_\nu\bar{\tau} - \frac{1}{2}\bar{q}^{\mu\nu}\partial_\mu\bar{\rho}\,\partial_\nu\bar{\rho} - V(\bar{\tau}\,,\bar{\rho})\right], \tag{4}$$

where $\kappa^2$ is the four-dimensional gravitational constant, $V(\bar{\tau}\,,\bar{\rho})$ denotes the moduli potential, and we have defined the dilaton modulus[24]

$$\tau = \mathrm{e}^{-\phi}\rho^{3/2}\,, \qquad \bar{\rho} = \sqrt{\frac{3}{2}}\kappa^{-1}\ln\rho\,, \qquad \bar{\tau} = \sqrt{2}\kappa^{-1}\ln\tau\,. \tag{5}$$

We have performed a conformal transformation on the four-dimensional metric $q_{\mu\nu} = (\bar{\kappa}/\tau\kappa)^2\bar{q}_{\mu\nu}$, where $\bar{q}_{\mu\nu}$ is the four-dimensional metric in the Einstein frame. $\bar{q}$ and $\bar{R}$ in the four-dimensional action (4) are the Ricci scalar and the determinant constructed from the metric $\bar{q}_{\mu\nu}$, respectively. Orientifold planes occupy $(p-3)$-dimensional internal space due to extending our four-dimensional universe. Then, the contribution of O$p$-plane ($p \geq 3$) to moduli potential survives. The moduli potential of four-dimensional effective theory is given by

$$\begin{aligned} V(\bar{\tau}\,,\bar{\rho}) = V_{\mathrm{Y}} + \sum_p V_{\mathrm{O}p} = &-A_{\mathrm{Y}}\exp\left[-\kappa\left(\sqrt{2}\bar{\tau} + \frac{\sqrt{6}}{3}\bar{\rho}\right)\right]R(\mathrm{Y}) \\ &- \sum_p A_{\mathrm{O}p}\exp\left[-\kappa\left\{\frac{3\sqrt{2}}{2}\bar{\tau} + \frac{\sqrt{6}}{6}(6-p)\bar{\rho}\right\}\right]\int d^{p-3}x\sqrt{g_{p-3}}\,, \end{aligned} \tag{6}$$

where $R(\mathrm{Y})$ denotes the Ricci scalar constructed from the metric $u_{ij}(\mathrm{Y})$ and $A_{\mathrm{Y}}$, $A_{\mathrm{O}p}$ are coefficients to scale with numbers of O$p$-planes. These coefficients in general depend on the function of the moduli of the internal space $\mathrm{Y}$.

When the potential form for the ekpyrotic scenario gives the negative and steep, the fast-roll parameters for the ekpyrosis have to obey[23]

$$\varepsilon_{\mathrm{f}} \equiv \kappa^2\frac{V^2}{\left(\partial_{\bar{\tau}}V\right)^2 + \left(\partial_{\bar{\rho}}V\right)^2} \ll 1\,, \qquad |\eta_{\mathrm{f}}| \equiv \left|1 - \frac{V\left(\partial_{\bar{\tau}}^2V + \partial_{\bar{\rho}}^2V\right)}{\left(\partial_{\bar{\tau}}V\right)^2 + \left(\partial_{\bar{\rho}}V\right)^2}\right| \ll 1\,. \tag{7}$$

This is analogy with the standard slow-roll parameters in inflation. The potential form satisfying the condition (7) gives the ekpyrotic period of slow contraction before the big bang.

Now we consider IIA compactifications on an internal space (2), namely positive curvature and Ricci flat spaces, involving orientifold planes, and discuss the No-Go theorem for ekpyrotic scenario. The analysis will focus on the behavior of the moduli potential in the volume modulus and dilaton. In order to present the no-go theorem using these fields, we have to still make sure that there are no steep directions of the scalar potential in the $(\bar{\tau}\,,\bar{\rho})$-plane. In such cases one can then study directions involving $\bar{\tau}\,,\bar{\rho}$, and finds that the scalar potential satisfies

$$\varepsilon_{\mathrm{f}} = \frac{V^2}{2}\left[\left(V_{\mathrm{Y}} + \frac{3}{2}\sum_p V_{\mathrm{O}p}\right)^2 + \frac{\left(V_{\mathrm{Y}} + V_{\mathrm{O4}} - V_{\mathrm{O8}}\right)^2}{3}\right]^{-1} > \frac{6}{31}\,, \tag{8}$$

where $p = 4, 6, 8$. The fast-roll parameter $\varepsilon_{\mathrm{f}}$ has the bound $\varepsilon_{\mathrm{f}} > 6/31$. This result does not depend on the choice of coefficients $A_{\mathrm{Y}}$ and $A_{\mathrm{O}p}$. The value of parameter $\varepsilon_{\mathrm{f}}$ is not much less than one, which is the contradiction with the fast-roll condition for ekpyrosis (7). Hence, ekpyrosis is not allowed[25,26].

On the other hand, for type IIB compactifications in the ekpyrotic model, we have also seen that it is possible to obtain simple no-go theorems in the $(\bar{\tau}\,,\,\bar{\rho})$-plane if one includes orientifold planes and the curvature of the internal space. From the Eq. (7), we find a constraint of the fast-roll parameter $\varepsilon_{\rm f}$

$$\varepsilon_{\rm f} = \frac{V^2}{2}\left[\left(V_{\rm Y} + \frac{3}{2}\sum_p V_{{\rm O}p}\right)^2 + \frac{\left(2V_{\rm Y} + 3V_{\rm O3} + V_{\rm O5} - V_{\rm O7} - 3V_{\rm O9}\right)^2}{12}\right]^{-1} > \frac{1}{6}\,, \quad (9)$$

where $p = 3, 5, 7, 9$. Unfortunately, the form moduli potential is not steep again as $\varepsilon_{\rm f}$ parameter turns out to be large value[26]. Just as in the IIA analogue, one obtains the bound $\varepsilon_{\rm f} > 1/6$. If we choose different values for $A_{\rm Y}$ and $A_{{\rm O}p}$ in the moduli potential (6), we can find the same bound.

## 3. Discussions

In this work, we have studied the No-Go theorem of the ekpyrosis for string theory in a spacetime of ten dimensions. We gave a potential of the scalar fields in four-dimensional effective theory, in terms of the compactification with smooth manifold. The effective potential of two scalar fields can be constructed by postulating suitable emergent gravity, orientifold planes, and vanishing fluxes on the ten-dimensional background. The scalar potential depends only on two moduli: $\bar{\tau}$ and $\bar{\rho}$.

In such a simple setting, one can show that $\varepsilon_{\rm f} > 6/31$ for IIA theory and $\varepsilon_{\rm f} > 1/6$ for IIB theory whenever $V(\bar{\tau}\,,\bar{\rho}) < 0$. It has been known for some time that the effective potential of scalar fields requires the fast-roll parameter to be small during the ekpyrotic phase. However, with the help of the tools developed in section 2, this is prohibited in a string theory with a compactification we have considered. Hence, the explicit nature of the dynamics has made it impossible to realize the ekpyrotic phase in the present study. This is consistent with results in Ref. 25.

Let us comment about the No-Go theorem of ekpyrotic scenario in heterotic M-theory. The original ekpyrotic model was inspired by heterotic M-theory which is obtained by going from eleven to five dimensions first, and only at even lower energies to 4 dimensions. It arose precisely from the realization that in heterotic M-theory a realistic phenomenology is obtained by introducing non-perturbative effects in the scalar field potential. Otherwise, in the Heterotic theory with a simple Calabi-Yau compactification we have considered in this work, the four-dimensional effective potential form of scalar fields $\varphi_i$ is given by $V(\varphi_i) \propto {\rm e}^{c_i\varphi_i}$, where $c_i < \sqrt{6}$. Then, it is impossible for us to construct the ekpyrotic scenario in the heterotic M-theory without non-perturbative correction. Unfortunately, no one has obtained an exact calculation of the potential involving all non-perturbative effects at present.

We have not considered the warped compactification because it does not make the modification of fast roll condition even if it couples to matter field in the ten-dimensional theory. The effect of warped factor gives only the change of coefficients for each term in the four-dimensional effective action after integration of internal space.

In order to embed ekpyrotic or cyclic models in a ten-dimensional supergravity we have investigated in this work, we may consider some ingredients, for instance, the dynamics of remaining moduli, higher curvature correction other than orientifold and flux. We have not attempted an explicit construction here, since that will take us beyond the scope of this study. A lot of study remains to be done in string theory before a cosmologically realistic case is treated.

## Acknowledgments

We would like to thank Thomas Van Riet for numerous valuable discussions and comments. This work is supported by Grants-in-Aid from the Scientific Research Fund of the Japan Society for the Promotion of Science, under Contracts No. 16K05364.

## References

1. J. Khoury, *et al.*, *Phys. Rev. D* **64**, 123522 (2001).
2. P. J. Steinhardt and N. Turok, *Science* **296**, 1436 (2002).
3. J. K. Erickson, *et al.*, *Phys. Rev. D* **75**, 123507 (2007).
4. J. L. Lehners, *Phys. Rept.* **465**, 223 (2008).
5. J. L. Lehners, *et al.*, *Int. J. Mod. Phys. D* **18**, 2231 (2009).
6. J. Khoury and P. J. Steinhardt, *Phys. Rev. Lett.* **104**, 091301 (2010).
7. D. H. Lyth, *Phys. Lett. B* **524**, 1 (2002).
8. R. Brandenberger and F. Finelli, *JHEP* **0111**, 056 (2001).
9. J. C. Hwang, *Phys. Rev. D* **65**, 063514 (2002).
10. J. Khoury, *et al.*, *Phys. Rev. D* **66**, 046005 (2002).
11. D. H. Lyth, *Phys. Lett. B* **526**, 173 (2002).
12. S. Tsujikawa, *Phys. Lett. B* **526**, 179 (2002).
13. A. Notari and A. Riotto, *Nucl. Phys. B* **644**, 371 (2002).
14. S. Tsujikawa, R. Brandenberger and F. Finelli, *Phys. Rev. D* **66**, 083513 (2002).
15. J. L. Lehners, *et al.*, *Phys. Rev. D* **76**, 103501 (2007).
16. E. I. Buchbinder, J. Khoury and B. A. Ovrut, *Phys. Rev. D* **76**, 123503 (2007).
17. K. Koyama, S. Mizuno and D. Wands, *Class. Quant. Grav.* **24**, 3919 (2007).
18. K. Koyama, *et al.*, *JCAP* **0711**, 024 (2007).
19. E. I. Buchbinder, *et al.*, *Phys. Rev. Lett.* **100**, 171302 (2008).
20. A. Fertig and J. L. Lehners, *JCAP* **1601**, no.01, 026 (2016).
21. A. Ito and J. Soda, *Phys. Lett. B* **771**, 415 (2017).
22. S. Gratton, *et al.*, *Phys. Rev. D* **69**, 103505 (2004).
23. J. Khoury, P. J. Steinhardt and N. Turok, *Phys. Rev. Lett.* **92**, 031302 (2004).
24. M. P. Hertzberg, *et al.*, *JHEP* **0712**, 095 (2007).
25. Eline Meeus, *A nogo-theorem for Ekpyrosis from 10D supergravity*, master's thesis, KU Leuven (2016).
26. K. Uzawa, *JHEP* **1806**, 041 (2018).

# The effects of anisotropy and non-adiabaticity on the evolution of the curvature perturbation

Sergio Andrés Vallejo Peña[1,†], Antonio Enea Romano[1,2],
Atsushi Naruko[3] and Misao Sasaki[4,5,6]

[1] *Instituto de Fisica, Universidad de Antioquia, A.A.1226, Medellin, Colombia*
[2] *Theoretical Physics Department, CERN, CH-1211 Geneva 23, Switzerland*
[3] *Frontier Research Institute for Interdisciplinary Sciences & Department of Physics, Tohoku University, Sendai 980-8578, Japan*
[4] *Kavli Institute for the Physics and Mathematics of the Universe, University of Tokyo, Kashiwa 277-8583, Japan*
[5] *Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*
[6] *Leung Center for Cosmology and Particle Astrophysics, National Taiwan University, Taipei 10617, Taiwan*
[†] *E-mail: sergio.vallejo@udea.edu.co*

We derive a general equation for the evolution of the curvature perturbation on comoving slices $\mathcal{R}_c$ in the presence of anisotropy and non-adiabaticity in the energy-momentum tensor of matter fields. The equation is obtained by manipulating the perturbed Einstein equations in the comoving slicing. It could be used to study the evolution of perturbations for a system with an anisotropic energy-momentum tensor, such as in the presence of a vector field or in the presence of non-adiabaticity, such as in a multi-field system.

*Keywords*: Curvature perturbation, anisotropic stress, non-adiabatic pressure.

## 1. Introduction

The theory of cosmological perturbations is very useful to study the early stages of the Universe, especially during inflation, that is, an exponential expansion phase which the standard cosmological model hypothesizes to explain observations such as anisotropies in the cosmic microwave background radiation (CMB). One quantity which is particularly important in this context is the curvature perturbation on comoving slices, $\mathcal{R}_c$. In slow-roll single field inflationary models this quantity is conserved on super-horizon scales,[1,2] which has important implications on the relation between primordial perturbations and late-time observables such as CMB anisotropies. For a globally adiabatic system in a single field model this quantity may not be conserved.[3] Other possible causes of super-horizon evolution could be the presence of anisotropy or non-adiabaticity in the energy-momentum tensor. We derive the equations for the curvature perturbation on comoving slices, $\mathcal{R}_c$, including these two terms showing that they act, as expected, as source terms which are relevant also on super-horizon scales. Our approach is quite generic and can be applied to any system which can be described by an energy-momentum tensor of the form we use, not only to a multi-scalar fields system.

The derivation is based on manipulating the Einstein equations in order to obtain an equation involving only $\mathcal{R}_c$, the anisotropy and non-adiabatic pressure and background quantities. The equation can be used to study phenomenologically the effects of anisotropy and non-adiabaticity without assuming any specific model.

One useful application could be to study models which violate the non-Gaussianity consistency relation[4] that was derived in fact based on the assumption of the conservation of the comoving curvatue perturbation on superhorizon scales.

## 2. Evolution of comoving curvature perturbations

The Einstein equations in a spatially flat Friedmann-Lemaître-Robertson-Walker (FLRW) background are

$$3\mathcal{H}^2 = a^2\,\rho\,, \tag{1}$$

$$2(\mathcal{H}' - \mathcal{H}^2) = -a^2\,(\rho + P)\,. \tag{2}$$

Here a prime denotes a derivative with respect to the conformal time $\eta$ and $\mathcal{H}$ stands for the conformal Hubble parameter defined by $\mathcal{H} = a'/a$. $\rho$ and $P$ represent the background energy density and pressure of the matter field respectively. We use the units in which $8\pi G = c = 1$.

Scalar perturbations on a spatially flat FLRW metric can be written as

$$ds^2 = a^2\left[-(1+2A)\mathrm{d}\eta^2 + 2\partial_i B\mathrm{d}x^i\mathrm{d}\eta + \left\{(1+2C)\delta_{ij} + 2\partial_i\partial_j E\right\}\mathrm{d}x^i\mathrm{d}x^j\right], \tag{3}$$

where the Latin indices run from 1 to 3. The corresponding energy-momentum tensor takes the form:

$$T^0{}_0 = -(\rho + \delta\rho)\,, \quad T^0{}_i = \frac{\rho + P}{a}u_i\,, \quad T^i{}_j = (P + \delta P)\delta^i{}_j + \Pi^i{}_j\,, \tag{4}$$

where

$$u_i = a\,\partial_i(v + B)\,, \quad \Pi^i{}_j = \delta^{ik}\partial_k\partial_j\Pi - \frac{1}{3}\overset{(3)}{\Delta}\Pi\delta^i{}_j\,, \quad \Pi^i{}_i = 0\,. \tag{5}$$

In the above equations $\Pi^i{}_j$ is the anisotropic stress component of the energy-momentum tensor, $v$ is the velocity potential, $\Pi$ is the anisotropy potential and we have defined $\overset{(3)}{\Delta} \equiv \delta^{ij}\partial_i\partial_j$.

The curvature perturbation on comoving slices $\mathcal{R}_c$ is a gauge-invariant quantity defined as the curvature perturbation $C$ evaluated on the hypersurfaces in which $v + B$ vanish. The spatial Fourier expansion of the linearly perturbed Einstein equations on comoving slices[5] takes the form :

$$2k^2(\mathcal{R}_c - \mathcal{H}\sigma_c) = a^2\delta\rho_c\,, \tag{6}$$

$$\mathcal{R}_c' - \mathcal{H}A_c = 0\,, \tag{7}$$

$$2(\mathcal{H}' - \mathcal{H}^2)A_c = a^2\left[\delta P_c - (2k^2/3)\Pi_c\right], \tag{8}$$

$$\sigma_c' + 2\mathcal{H}\sigma_c - A_c - \mathcal{R}_c = a^2\Pi_c\,, \tag{9}$$

where $\sigma = E' - B$ is the scalar shear.

In general we can decompose the pressure perturbation as

$$\delta P_c = c_s^2(\eta)\delta\rho_c + \Gamma_c\,, \tag{10}$$

where we can interpret $c_s$ and $\Gamma_c$ as the adiabatic sound speed and the non-adiabatic part of the pressure respectively. For a minimally coupled scalar field model $c_s = 1$ and $\Gamma_c$ is zero, but in general one would expect that $\Gamma_c$ could be non-vanishing. Our goal is to derive an equation for $\mathcal{R}_c$ in the presence of both anisotropic stress $\Pi^i{}_j$ and non-adiabatic pressure $\Gamma_c$.

First we use Eq. (7) to express $A_c$ in terms of $\mathcal{R}_c$,

$$A_c = \frac{\mathcal{R}_c'}{\mathcal{H}}\,. \tag{11}$$

We substitute this $A_c$ and $\delta P_c$ given in Eq. (10) into Eq. (8), and solve it for $\delta\rho_c$ :

$$\delta\rho_c = \frac{\mathcal{H}\left(2k^2\Pi_c - 3\Gamma_c\right) - 3(\rho+P)\mathcal{R}_c'}{3\mathcal{H}c_s^2}\,. \tag{12}$$

We then insert this into Eq. (6) to get an expression for $\sigma_c$ :

$$\sigma_c = \frac{\mathcal{R}_c}{\mathcal{H}} - \frac{a^2\left[\mathcal{H}(2k^2\Pi_c - 3\Gamma_c) - 3(\rho+P)\mathcal{R}_c'\right]}{6k^2\mathcal{H}^2c_s^2}\,. \tag{13}$$

Finally we substitute $A_c$ and $\sigma_c$ given by Eqs. (11) and (13), respectively, into Eq. (9) to obtain

$$\mathcal{R}_c'' + \frac{(z^2)'}{z^2}\mathcal{R}_c' - c_s^2\,\overset{(3)}{\Delta}\,\mathcal{R}_c + \frac{\mathcal{H}}{\rho+P}Y_c = 0\,, \tag{14}$$

where we have defined

$$z^2 \equiv \frac{a^4(\rho+P)}{c_s^2\mathcal{H}^2}\,, \tag{15}$$

$$Y_c \equiv \left[\log\left(\frac{a^4}{\mathcal{H}c_s^2}\right)\right]'\left(\frac{2}{3}\,\overset{(3)}{\Delta}\,\Pi_c + \Gamma_c\right) + 2\mathcal{H}c_s^2\,\overset{(3)}{\Delta}\,\Pi_c + \frac{2}{3}\,\overset{(3)}{\Delta}\,\Pi_c' + \Gamma_c'\,. \tag{16}$$

As expected, for adiabatic ($\Gamma_c = 0$) and isotropic perturbations ($\Pi_c = 0$) the above equation takes the well-known form :

$$\mathcal{R}_c'' + \frac{(z^2)'}{z^2}\mathcal{R}_c' - c_s^2\,\overset{(3)}{\Delta}\,\mathcal{R}_c = 0\,. \tag{17}$$

## 3. Curvature perturbation for scalar fields

Given the generality of the form of the energy momentum tensor used in the derivation of Eq. (14) it can be applied to a wide class of physical scenarios, including multi-field systems. Let us consider the case of two minimally coupled scalar fields with Lagrangian

$$L = -\sum_{n=1}^{2} X_n - 2V(\Phi_1, \Phi_2)\,, \tag{18}$$

where $X_n = g^{\mu\nu}\partial_\mu\Phi_n\partial_\nu\Phi_n$ and $\Phi_n(x^\mu) = \phi_n(\eta) + \delta\phi_n(x^\mu)$. The perturbed energy-momentum tensor, without gauge fixing, is given by

$$\begin{aligned}
\delta\rho &= \frac{\phi_1'\delta\phi_1' + \phi_2'\delta\phi_2' - A(\phi_1'^{\,2} + \phi_2'^{\,2})}{a^2} + V_1\delta\phi_1 + V_2\delta\phi_2\,,\\
\delta P &= \frac{\phi_1'\delta\phi_1' + \phi_2'\delta\phi_2' - A(\phi_1'^{\,2} + \phi_2'^{\,2})}{a^2} - V_1\delta\phi_1 - V_2\delta\phi_2\,,\\
\Pi = 0 \quad &, \quad \delta T^0{}_i = \partial_i\left(-\frac{\phi_1'\delta\phi_1 + \phi_2'\delta\phi_2}{a^2}\right),
\end{aligned} \tag{19}$$

where we denote the partial derivatives as $V_n = (\partial V/\partial\Phi_n)(\phi_1, \phi_2)$.

The field perturbations transform under an infinitesimal time translation $\eta \to \eta + \delta\eta$

$$\widetilde{\delta\phi_1} = \delta\phi_1 - \phi_1'\delta\eta \quad , \quad \widetilde{\delta\phi_2} = \delta\phi_2 - \phi_2'\delta\eta\,. \tag{20}$$

The time translation $\delta\eta_c$ necessary to define the comoving slices can be found by imposing the condition $(\delta T^0{}_i)_c \propto \phi_1'\widetilde{\delta\phi_1} + \phi_2'\widetilde{\delta\phi_2} = 0$, giving

$$\delta\eta_c = \frac{\phi_1'\delta\phi_1 + \phi_2'\delta\phi_2}{\phi_1'^{\,2} + \phi_2'^{\,2}}\,. \tag{21}$$

The comoving curvature perturbation for the two scalar fields system is given by

$$\mathcal{R}_c = C - \mathcal{H}\delta\eta_c = C - \mathcal{H}\frac{\phi_1'\delta\phi_1 + \phi_2'\delta\phi_2}{\phi_1'^{\,2} + \phi_2'^{\,2}}\,. \tag{22}$$

The gauge invariant field perturbations in the comoving slices now can be defined

$$U_1 = \delta\phi_1 - \phi_1'\frac{\phi_1'\delta\phi_1 + \phi_2'\delta\phi_2}{\phi_1'^{\,2} + \phi_2'^{\,2}}\,, \tag{23}$$

$$U_2 = \delta\phi_2 - \phi_2'\frac{\phi_1'\delta\phi_1 + \phi_2'\delta\phi_2}{\phi_1'^{\,2} + \phi_2'^{\,2}}\,, \tag{24}$$

and similarly for the pressure and energy perturbations in the comoving slices we get

$$\delta\rho_c = \frac{\phi_1'U_1' + \phi_2'U_2' - A_c(\phi_1'^{\,2} + \phi_2'^{\,2})}{a^2} + V_1U_1 + V_2U_2\,, \tag{25}$$

$$\delta P_c = \frac{\phi_1'U_1' + \phi_2'U_2' - A_c(\phi_1'^{\,2} + \phi_2'^{\,2})}{a^2} - V_1U_1 - V_2U_2\,. \tag{26}$$

Note that all the above quantities are gauge invariant by construction.

Combining Eqs. (23), (24) and the background field equations of motion we find

$$\frac{\phi_1'U_1' + \phi_2'U_2'}{a^2} = V_1U_1 + V_2U_2 = -\frac{\phi_1'^{\,2} + \phi_2'^{\,2}}{4a^2}\Theta\,, \tag{27}$$

where we have defined the function $\Theta$ according to

$$\Theta = \left[\frac{\partial}{\partial\eta}\left(\frac{\phi_1'^{\,2} - \phi_2'^{\,2}}{\phi_1'^{\,2} + \phi_2'^{\,2}}\right)\right]\left(\frac{\delta\phi_1}{\phi_1'} - \frac{\delta\phi_2}{\phi_2'}\right). \tag{28}$$

Assuming a classical field trajectory parameterized as $\phi_2 = \phi_2(\phi_1)$ we can write $\Theta$ in this form

$$\Theta = -4\frac{\mathrm{d}^2\phi_2}{\mathrm{d}\phi_1^2}\left[\left(\frac{\mathrm{d}\phi_2}{\mathrm{d}\phi_1}\right)^2 + 1\right]^{-2}\left(\frac{\mathrm{d}\phi_2}{\mathrm{d}\phi_1}\delta\phi_1 - \delta\phi_2\right). \tag{29}$$

From the above expression we can see that in order for $\Theta$ to be different from zero the trajectory has to have non vanishing second derivative, i.e. there must be some turn in the field space.

After replacing Eqs. (7) and (27) into Eqs. (25) and (26) we get

$$\delta\rho_c = -\frac{\phi_1'^2 + \phi_2'^2}{a^2\mathcal{H}}\mathcal{R}_c' - \frac{\phi_1'^2 + \phi_2'^2}{2a^2}\Theta\,, \tag{30}$$

$$\delta P_c = -\frac{\phi_1'^2 + \phi_2'^2}{a^2\mathcal{H}}\mathcal{R}_c'\,. \tag{31}$$

It follows from Eqs. (30) and (31) that

$$\delta P_c = \delta\rho_c + \frac{\phi_1'^2 + \phi_2'^2}{2a^2}\Theta\,, \tag{32}$$

and comparing this with Eq. (10) we obtain the sound speed and the entropy perturbations

$$c_s^2(\eta) = 1\,, \qquad \Gamma_c = \frac{\phi_1'^2 + \phi_2'^2}{2a^2}\Theta\,. \tag{33}$$

From these relations we get

$$\mathcal{R}_c'' + 2\frac{z'}{z}\mathcal{R}_c' - \overset{(3)}{\Delta}\mathcal{R}_c = -\frac{a^2\mathcal{H}}{(\phi_1'^2 + \phi_2'^2)}Y_c\,, \tag{34}$$

where

$$z^2 = \frac{a^2(\phi_1'^2 + \phi_2'^2)}{\mathcal{H}^2}\,, \tag{35}$$

$$Y_c = \left[\log\left(\frac{a^4}{\mathcal{H}}\right)\right]'\Gamma_c + \Gamma_c'\,. \tag{36}$$

Eq. (34) is in agreement with Ref. [6], confirming that Eq. (14) is general and can also be applied to multi-field systems once the entropy has been appropriately defined. In general in order to use Eq. (14) it is first necessary to compute the energy momentum tensor in the comoving slices with a procedure similar to the one shown above for two fields.

## 4. Conclusions

We have derived a general equation for the evolution of the comoving curvature perturbation by taking into account the effects of anisotropy and non-adiabaticity in the energy-momentum tensor. The equation can be applied to multi-field systems. This approach does not require the decomposition of field perturbations in

components parallel and perpendicular to the classical field trajectory, but is based just on the fundamental definition of non-adiabatic pressure. In the future it will be interesting to apply the equation to more complex systems where both entropy and anisotropy are present.

## References


1. A. E. Romano, S. Mooij, and M. Sasaki, *Phys. Lett. B* **755**, 464 (2016), arXiv:gr-qc/1512.05757.
2. D. Wands, K. A. Malik, D. H. Lyth, A. R. Liddle, *Phys. Rev. D* **62**, 043527 (2000), arXiv:astro-ph/0003278.
3. A. E. Romano, S. Mooij, and M. Sasaki, *Phys. Lett. B* **761**, 119 (2016), arXiv:gr-qc/1606.04906.
4. J. M. Maldacena, *JHEP* **05**, 013 (2003), arXiv:astro-ph/0210603.
5. K. A. Malik and D. Wands, *Physics Reports* **475**, 1 (2009), arXiv:0809.4944.
6. J. García-Bellido and D. Wands, *Phys. Rev. D* **52**, 6739 (1995), arXiv:gr-qc/9506050.

# Horndeski charged boson stars and black holes

Yosef Verbin* and Yves Brihaye**

*Department of Natural Sciences, The Open University of Israel, Raanana 43107, Israel

**Physique Théorique et Mathématiques, Université de Mons, Place du Parc, B–7000 Mons, Belgique

We find two distinct families of new charged boson star solutions of a complex scalar field coupled non-minimally to gravity by a "John-type" term of Horndeski theory. The end points of one of the branches are extremal Reissner-Nordström black hole solutions.

## 1. Introduction and general considerations

Although the main motivation for the resurrection[1] of Horndeski scalar-tensor theory[2] was cosmological, a large number of investigations of its small scale aspects was also carried out. To mention just a few, black hole solutions[3,4] were discovered in the "John"-type version of Horndeski theory which contains a non-minimal coupling term of the scalar kinetic term to the Einstein tensor, $\mathcal{L}_{NM} \sim G^{\mu\nu}\partial_\mu\phi\partial_\nu\phi$ and in another version known as the Sotiriou-Zhou model[5] which has a $\phi\Re_{GB}$ term in the Lagrangian where $\Re_{GB}$ is the Gauss-Bonnet term. Black hole solutions were found[6] also in a "biscalar extension" of the theory and boson stars were found[7] in a different biscalar version where the scalar field becomes complex. On the other hand, the real scalar system was also studied in the presence of a radial electric field[8] and black holes were found in that case too. These latter black hole solutions are essentially a "superposition" of the pure scalar black holes with the electric field of a point charge at their centers. Further study[9] of the real scalar system produced magnetic black holes as well.

In some of these studies the scalar field ceases to play a role of a gravitational degree of freedom, and the Horndeski theory is used as an inspiration for new kinds of couplings between scalar matter and gravity. These scalar fields may be interpreted as dark matter candidates. This is the attitude we take in this work where we analyze the gauged version of the "John" Lagrangian thus producing the conventional minimal coupling between the scalar and Abelian vector fields. The dynamics of such a system is determined by the following action which is a simple generalization of that for self-gravitating solitons like boson stars or Q-stars[10,11]:

$$S = \int d^4x\sqrt{-g}\left(\frac{R}{2\kappa} - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}|D_\mu\Phi|^2 - U(|\Phi|) - \beta G^{\mu\nu}D_\mu\Phi^* D_\nu\Phi\right). \quad (1)$$

The non-minimal coupling term to the Einstein tensor is known to introduce in the field equations of the pure scalar case derivatives of order not higher than two.[2] It is easy to show that the gauged version still has this property.

Now we assume a static spherically symmetric metric and a radial "electric" configuration for the scalar and vector fields:

$$ds^2 = A(r)^2dt^2 - B(r)^2dr^2 - r^2d\Omega_2^2, \qquad \Phi = f(r)e^{i\omega t}, \qquad A_\mu dx^\mu = a_0(r)dt. \quad (2)$$

We take the simplest potential of a mass term only and obtain 4 independent field equations for the 4 functions $f(r)$, $a_0(r)$, $A(r)$ and $B(r)$. We do not present them here due to lack of space, but refer the readers instead to Ref. 12 where a more detailed study of this system can be found.

We will characterize the solutions by their total mass, electric charge and particle number. We will define the mass function $\mathcal{M}(r)$ by $1/B^2(r) = 1 - 2\mathcal{M}(r)/r$ so the ADM mass may be read of the asymptotic decay of the metric, i.e. $\mathcal{G}M = \mathcal{M}(\infty)$.

The local $U(1)$ charge is readily obtained from the time component of the conserved total current:

$$\mathcal{J}^\nu = J^\nu - 2\beta G^{\mu\nu} J_\mu, \quad J_\nu = -\frac{ie}{2}\left(\Phi^* D_\nu \Phi - \Phi D_\nu \Phi^*\right) \tag{3}$$

which in the spherically-symmetric case yields the following expression:

$$Q = 4\pi e \int_0^\infty dr \frac{r^2 B}{A}(\omega - e a_0) f^2 \left(1 + \frac{4\beta \mathcal{M}'}{r^2}\right). \tag{4}$$

Without loss of generality we will take $\omega > 0$, so $Q > 0$ as well.

These two quantities, mass and charge determine the asymptotic behavior of the localized solution that we seek, i.e. the exterior Reissner-Nordström solutions:

$$A(r)^2 = 1/B(r)^2 = \left(1 - \frac{\kappa M}{4\pi r} + \frac{\kappa Q^2}{2(4\pi)^2 r^2}\right), \;\; a_0 = \frac{Q}{4\pi r}. \tag{5}$$

So, we may extract also the charge from the asymptotic behavior of the solutions.

The total electric charge is carried in our case by $N$ particles of mass $m$ and charge $e$. We can therefore define a particle density and particle number by

$$\rho_N = (\omega - e a_0) f^2 \left(1 + \frac{4\beta \mathcal{M}'}{r^2}\right), \quad N = 4\pi \int_0^\infty dr \frac{r^2 B}{A} \rho_N. \tag{6}$$

Thus, we can use the mass to particle number ratio, or more precisely, $M/mN$ in order to study the stability of these solutions against "fission" into a number of smaller stable structures: this ratio $M/mN$ needs to be less than 1 in order for the solutions to be stable against this kind of processes. The condition $M/mN < 1$ is however not sufficient to guarantee the stability under linear perturbations[11]. We will postpone this issue to a future work.

The field equations[12] depend a priori on $m$, $e$, $\beta$ and Newton's constant $\mathcal{G}$. The equations can be written in a dimensionless form which reveals a dependence on two independent parameters only: $\overline{\beta} = \beta m^2$ and $\overline{e} = e/\sqrt{\kappa m^2}$. The rescaled and original versions of the charge, particle number and mass are related to one another by $\bar{M} = Mm/\mu^2 = \kappa m M$, $\bar{Q} = Qm^2/\mu^2 = Q\kappa m^2$ and $\bar{N} = Nm^2/\mu^2 = N\kappa m^2$. The stability ratio may be calculated both ways: $M/mN = \bar{M}/\bar{N}$.

The mass function $\mathcal{M}(r)$ which we defined above has a dimension of length and is rescaled therefore similarly as $\bar{\mathcal{M}} = m\mathcal{M}$. The total mass may therefore be expressed in terms of $\bar{\mathcal{M}}(\infty)$ or $\bar{M}$ as $M = 8\pi\bar{\mathcal{M}}(\infty) m_{Pl}^2/m = \bar{M} m_{Pl}^2/m$, where we use the Planck mass $m_{Pl} = 1/\sqrt{\kappa}$. Similarly the rescaled particle number $\bar{N}$ may be used to express the total particle number as $N = 8\pi\bar{N}(m_{Pl}/m)^2$.

## 2. Solutions and their properties

Although in some special cases of this theory solutions were obtained explicitly, we found it unlikely to achieve this in the present system we study here. We therefore revert to numerical analysis in order to construct our solutions.

Within our ansatz (2), the unknown functions $B, A, f, a_0$ appear with maximal derivatives $B', A'', f'', a_0''$ in the system of field equations. These maximal derivatives appear linearly, so that the equations are symbolically written as

$$M(B', A'', f'', a_0'')^T = K \tag{7}$$

where $M$ represents a *non-diagonal* $4\times4$ matrix and $K$ a $4\times1$ matrix (vector) whose entries depend on the lower derivatives. The quantity $\Delta(r) \equiv \det M$ plays a major role in the existence of regular solutions over space-time. Indeed zeroes of $\Delta(r)$ correspond to singular points of the system. In particular, the occurrence of such a point renders impossible the construction of numerical solutions on $r \in [0, \infty]$.

For a given set of $\beta$ and $e$, the solutions comprise a one-parameter set parametrized by $f(0)$ or alternatively by $\omega$ or $A(0)$, although in some regions of parameter space the correspondence is not one to one as will become evident shortly.

In other areas of parameter space we find a second branch of solutions with significantly different behavior with respect to the first ("ordinary") branch. Fig. 1-Left shows the general structure in the $(f(0), A(0))$ plane for $e = 0.7$ and five values of $\beta$: $\beta = \pm1, \pm0.5$ and the minimally coupled case $\beta = 0$. The two-branch structure appears for $\beta > 0$ only and above a certain value of $e$.

In order to understand the situation better, we calculate the main properties of the solutions: mass, particle number and their ratio which is in one to one correspondence with the binding energy per particle $1 - M/N$. We study their

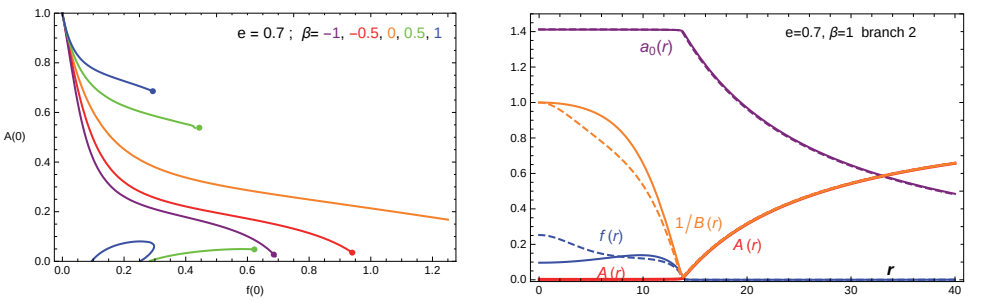



Fig. 1. Left panel: Families of solutions for $e = 0.7$ and 5 values of $\beta$. Note the two branch structure for $\beta > 0$ and the different behavior of branch 2 families appearing at the lower part. The "bullets" mark the end point solutions without back-bending, i.e. $\Delta = 0$ solutions. The other end points at $A(0) = 0$ have almost extremal Reissner-Nordström behavior - Right panel: The two end-point solutions for $e = 0.7$ and $\beta = 1.0$. Both solutions have $N = 348.2$ and $M = 344.7$ and the same external structure (with $A(r)$ and $1/B(r)$ almost identical), but very different internal structure: $f(0) = 0.096255$ and $f''(0) > 0$ - solid; $f(0) = 0.2525$ and $f''(0) < 0$ - dashed.

dependence on the parameters of the theory $\beta$ and $e$ as well as their behavior within a given family corresponding to specific $(\beta, e)$ and parametrized by $f(0)$.

Fig. 2 contains some of these aspects: the mass as a function of the central scalar field and the dependence of binding energy per particle on the particle number.

First we observe from our results that the mass and particle number decrease with $\beta$ which is in accord with the fact that gravity becomes stronger with increasing $\beta$, so a smaller number of particles can form bound states.

The dependence on the coupling constant $e$ is expected to cause the opposite effect since increasing its value amounts to a stronger electrostatic repulsion. This is indeed the general behavior as we will see in what follows. The electrostatic repulsion is also responsible for a critical value of $e$ which in terms of the dimensionless parameter reads $\overline{e}_{cr} = 1/\sqrt{2}$. Actually, this bound was obtained[11] for minimal coupling ($\beta = 0$), but it still holds here with a weak dependence on $\beta$.

Branch 1 of the solutions is quite similar to that of charged ordinary (i.e. $\beta = 0$) boson stars for both signs of $\beta$: it starts at $f(0) = 0$ with $M = N = 0$, then the mass and particle number grow steeply, attain (together) a maximum and then decrease. However, unlike the case of ordinary charged boson stars, the solutions hit an end point beyond which no solutions exist. This end point has a similar origin as reported for the ungauged counterpart[7] where the function $\Delta(r)$ was found to develop a zero at the end point excluding solutions beyond that point.

Branch 2 solutions exists only in restricted regions of parameter space: for a given $e$, $\beta$ should be larger than a certain value. For given $e$ and $\beta$, branch 2 starts at a minimal non-zero $f(0)$ and $A(0) = 0$ and has also a maximal $f(0)$. This maximal $f(0)$ point can be either: (i) An end point of a similar kind of branch 1 where solutions cease to exist because $\Delta(r)$ gets a zero. In this case $A(0) > 0$ at that point. (ii) A turning point in which back-bending occurs and solutions still exist down to some final $f(0)$ which is an end point because $A(0) = 0$.

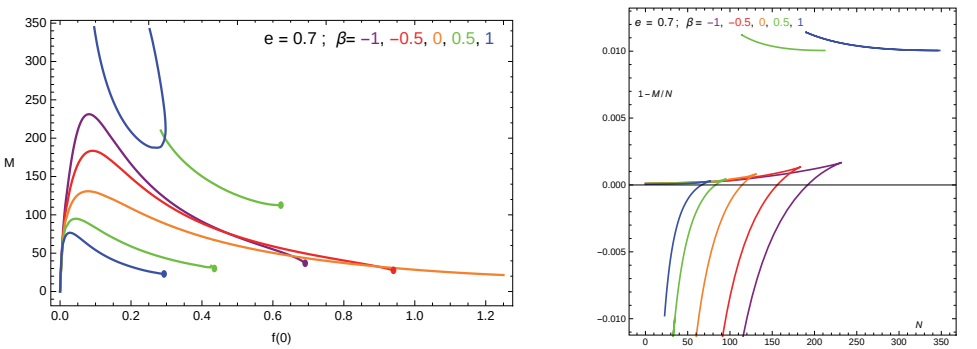



Fig. 2. Mass and binding energy per particle for $e = 0.7$ and 5 values of $\beta$: $\beta = \pm 1, \pm 0.5$ and $\beta = 0$. Left panel: Mass versus the central value of the scalar field; Right panel: Binding energy per particle as a function of particle number. The branch 2 curves are the upper ones at both figures. The "un-bulleted" ends of the branch 2 curves in the left represent the extremal Reissner-Nordström solutions. These correspond to the $A(0) = 0$ points of Fig. 1.

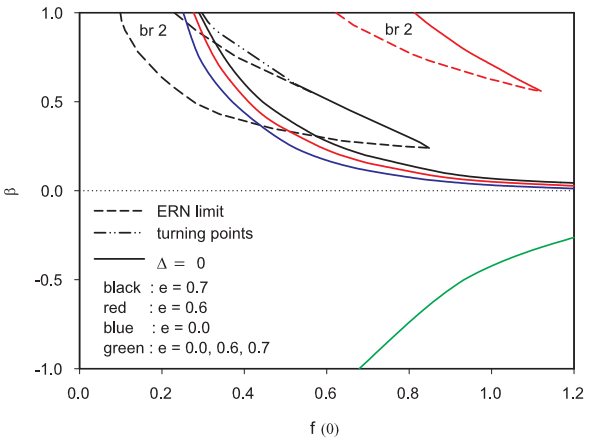



Fig. 3. The branch structure of parameter space for $e = 0$, 0.6, 0.7 and $-1 \leq \beta \leq 1$. Note that branch 1 extends from the vertical axis at $f(0) = 0$ up to the curves. Branch 2 regions are indicated explicitly. For further explanation - see text.

To sum up, there are two kinds of end points of curves in parameter space as is seen clearly in Fig. 1 and 2: The first kind are the "bulleted" which correspond to the $\Delta = 0$ termination condition. These appear on the $\beta \neq 0$ branch 1 curves and the higher $f(0)$ end point of the $\beta = -0.5$ branch 2 curve. End points of the second kind are the two "un-bulleted" end points of the $\beta = -1$ branch 2 curve and the lower $f(0)$ end point of the $\beta = -0.5$ branch 2. These correspond to the $A(0) = 0$ termination condition.

At the end points of branch 2 where $A(0) \to 0$ and the scalar field becomes confined within a ball with a sharp boundary at $r = r_S$, we find that the solutions approach extremal Reissner-Nordström black hole solutions (ERN or ERNS in the following) which have identical mass and particle number, but correspond to very different internal structure. In our parametrization the ERNS have the characteristic relation $M = \sqrt{2}Q$ or in terms of particle number, $M = \sqrt{2}eN$. The right panel of Fig. 1 shows for $e = 0.7$ and $\beta = 1.0$ the identical exterior solutions of the end-point solutions that correspond to the two different internal structures. The $M$ and $N$ values of these solutions $M = 344.7$ and $N = 348.2$ satisfy this relation very nicely. Another characteristic of the ERNS, is the equality $A(r) = \omega - ea_0(r)$ which we do not show in our plots, but was checked to hold very well.

As a summary of the branch structure of this theory we present Fig. 3 which illustrates the structure of parameter space in the $(f(0), \beta)$ plane for $e = 0.6$, 0.7 and for comparison with previous work[7] also $e = 0$. Note that the line for the $\beta < 0$ solutions is the line of end points for any value of $e$: the maximal $f(0)$ depends on $\beta$ but not on $e$.

## 3. Conclusion

In this work we studied the parameter space of a new kind of charged boson stars with a non-minimal coupling to gravity of the "John-type" of Horndeski theory. We found two distinct families according to their position in parameter space. More widespread are those of branch 1 existing for a finite interval of the central value of the scalar field, $f(0)$ starting from zero and ending at some finite maximal value. This branch contains as a special case the charged boson stars of the minimally coupled theory. In some regions of parameter space we found a new second branch of solutions which are more massive and more stable than those of branch 1. This second branch exists also in a finite interval of $f(0)$, but its end points (either both or in some cases only one) are extremal Reissner-Nordström solutions.

It would be interesting to study whether the intriguing solutions of branch 2 could exist in different versions of the tensor scalar gravity as well. As a first step in this direction, one could check whether a self-interacting potential would lead to branch 2-type of solutions with smaller values of the non-minimal parameter $\beta$.

The existence of hairy black holes in some modification of this system is also of much interest. It might be sufficient to introduce a bare cosmological constant as was done previously[4] for the neutral system.

## References


1. C. Deffayet *et al.*, Phys. Rev. D **84**, 064039 (2011); T. Kobayashi *et al.*, Prog. Theor. Phys. **126**, 511 (2011); C. Charmousis *et al.*, Phys. Rev. Lett. **108**, 051101 (2012); C. de Rham and L. Heisenberg, Phys. Rev. D **84**, 043503 (2011); C. Charmousis *et al.*, Phys. Rev. D **85**, 104040 (2012).
2. G. W. Horndeski, Int. J. Theor. Phys. **10**, 363 (1974).
3. M. Rinaldi, Phys. Rev. D **86**, 084048 (2012); E. Babichev and C. Charmousis, JHEP **1408**, 106 (2014).
4. A. Anabalon, A. Cisterna and J. Oliva, Phys. Rev. D **89**, 084050 (2014); M. Minamitsuji, Phys. Rev. D **89**, 064017 (2014).
5. T. P. Sotiriou and S. Y. Zhou, Phys. Rev. Lett. **112**, 251102 (2014); T. P. Sotiriou and S. Y. Zhou, Phys. Rev. D **90**, 124063 (2014).
6. C. Charmousis *et al.*, JHEP **1407**, 085 (2014).
7. Y. Brihaye, A. Cisterna and C. Erices, Phys. Rev. D **93**, 124057 (2016).
8. T. Kolyvaris *et al.*, Class. Quant. Grav. **29**, 205011 (2012); A. Cisterna and C. Erices, Phys. Rev. D **89**, 084038 (2014).
9. E. Babichev, C. Charmousis and M. Hassaine, JCAP **1505**, 031 (2015).
10. R. Friedberg, T. D. Lee and A. Sirlin, Phys. Rev. D **13**, 2739 (1976); S. R. Coleman, Nucl. Phys. B **262**, 263 (1985)+**269**, 744 (1986) [Erratum]; B. W. Lynn, Nucl. Phys. B **321**, 465 (1989); T. D. Lee and Y. Pang, Phys. Rept. **221**, 251 (1992); F. E. Schunck and E. W. Mielke, Class. Quant. Grav. **20**, R301 (2003).
11. P. Jetzer, Phys. Rept. **220**, 163 (1992).
12. Y. Verbin and Y. Brihaye, Phys. Rev. D **97**, 044046 (2018).

# 21-cm signal of neutral hydrogen from high cosmological redshifts

Anastasia Fialkov

*Astronomy Centre, Department of Physics and Astronomy, University of Sussex, Brighton BN1 9QH, UK*

The high-redshift 21-cm signal of neutral hydrogen is expected to be a sensitive probe of primordial star formation as well as of the heating and ionization histories of the Universe. The recent tentative detection by the EDGES Low-Band of the cosmological 21-cm signal, if confirmed, is the first ever signature from the dawn of star formation. However, the magnitude and the shape of this signal are incompatible with standard astrophysical predictions, requiring either colder than expected gas, or an excess radio background above the Cosmic Microwave Background (CMB) radiation. Here we provide a brief overview of 21-cm cosmology, the status of the global 21-cm experiments and proposed theoretical explanations for the EDGES detection.

## 1. Introduction

The dawn of star formation at redshifts $z \sim 15 - 30$ remains veiled in mystery being out of reach of existing observations. Large scale structure surveys map cosmic structure only out to redshift of $z \sim 3$. Observations of lensed fields with the Hubble Space Telescope (HST) reach further and have revealed galaxies at the Epoch of Reionization (EoR, $z \sim 6 - 10$) with the most distant galaxy detected at redshift $z \sim 11.1$[31]; the Atacama Large Millimeter Array (ALMA) sees dusty bright galaxies at $z \sim 9.1$[21]; while the record-holder high-redshift quasar is located at $z = 7.54$[2]. Existence of so massive and metal-rich objects when the Universe was less than a billion years old suggests an early formation of the first stars (e.g., as early as $z \sim 15$ in the case of the metal-rich galaxy at $z = 9.1$[21]). However, even with the most powerful telescopes, only the brightest out of the high-redshift sources are currently within the reach, while more typical objects – those that reheated and reionized the Universe – remain unobserved. As a result, the process of heating and ionization of the intergalactic medium (IGM) during the first few hundred million years after the Big Bang, as well as the processes of the primordial star formation, remain poorly understood. The 21-cm signal of neutral hydrogen, reviewed here, provides a unique way to fill in the gap.

## 2. Expected 21-cm Signal: Dependence on Astrophysical Processes

As the Universe evolves, astrophysical processes strongly affect the thermal and ionization states of the IGM. Right after recombination high fraction of free residual electrons mediates thermal coupling between the gas and the Cosmic Microwave Background (CMB) radiation. The IGM and the CMB are cooling together until the coupling becomes inefficient at around $z_{\rm th} \sim 200$ as a result of the cosmic expansion. At lower redshifts, and prior to the formation of the first astrophysical heating sources, gas is cooling faster than the CMB. This period in cosmic history, prior to star formation, is called the dark ages and is characterized by cold and

neutral IGM. Emergence of first stars and first X-ray sources marks the onset of cosmic dawn, completely changing the high-redshift environment. During this epoch stars produce Ly-$\alpha$ background, and the IGM is reheated by X-ray photons to temperatures higher than that of the CMB. At that epoch, the heating rate of the IGM depends on the properties of the X-ray sources, such as their spectral energy distribution (SED) and bolometric luminosity[13,14,32]. The process of reionization happens slower than heating. This is because the first small galaxies are inefficient in ionizing the gas. The appearance of massive galaxies in dark matter halos of mass $\mathrm{M_h} \gtrsim 10^9\ \mathrm{M_\odot}$ leads to a more efficient process of reionization which, as observations suggest, starts not long before $z = 7.5$ and terminates at $z \sim 6$[19,20]. This sequence of cosmic events is imprinted in the neutral hydrogen signal.

The lowest energy level of the hydrogen atom is split due to the interaction between proton and electron spins (hyper-fine splitting) which gives rise to the rest-frame $\nu_{21} = 1.42$ GHz radio signal with the equivalent wavelength of 21 cm (see Barkana (2016)[3] for a recent review). The cumulative signal of neutral IGM at redshift $z$ observed against the background radiation depends on the processes of cosmic heating and ionization and can be written as

$$T_{21} \propto x_{\mathrm{HI}} \left(1 - T_{\mathrm{rad}} T_{\mathrm{S}}^{-1}\right), \tag{1}$$

where $x_{\mathrm{HI}}$ is the fraction of neutral hydrogen, $T_{\mathrm{rad}}$ is the brightness temperature of the background radiation at 1.42 GHz (which is usually assumed to be the temperature of the CMB[1] at redshift $z$) and $T_{\mathrm{S}}$ is the spin temperature of the transition which encodes complex astrophysical dependence. The spin temperature can be written as $T_{\mathrm{S}} = (1 + x_{\mathrm{tot}}) \left(T_{\mathrm{rad}}^{-1} + x_{\mathrm{tot}} T_{\mathrm{K}}^{-1}\right)^{-1}$, where $x_{\mathrm{tot}} = x_C + x_\alpha$ (see Eqs. 46 and 48 of Barkana (2016)[3]), $x_\alpha$ is the Wouthhausen-Field (WF[17,37]) coupling coefficient which depends on the intensity of the Ly-$\alpha$ background radiation produced by stars, $x_C$ is the collisional coupling coefficient, and $T_{\mathrm{K}}$ is the kinetic temperature of the neutral IGM. $x_C$ is important at high redshifts when the Universe was dense and intra-atomic collisions were frequent enough to couple $T_{\mathrm{S}}$ to $T_{\mathrm{K}}$. This effect became sub-dominant as the Universe expanded and the gas rarefied. $x_\alpha$, which is negligible before the onset of star formation, became important at cosmic dawn. Ly-$\alpha$ photons produced by stars coupled $T_S$ to the color temperature of the Ly-$\alpha$ radiation, which at cosmic dawn is close to $T_{\mathrm{K}}$. Finally, ionizing radiation depletes the amount of neutral gas in the IGM, reducing $x_{\mathrm{HI}}$. Owing to its dependence on the underlying astrophysics and cosmology, the 21-cm signal is a powerful tool to characterize the formation and the evolution of the first populations of astrophysical sources and, potentially, properties of dark matter, across cosmic time.

Figure 1 shows a cartoon of the evolution of the global (sky-averaged) 21-cm signal and corresponding power spectrum of the 21-cm fluctuations (shown at a fixed comoving angular scale) as a function of redshift. The shapes of the global signal and the power spectrum reflect the thermal and ionization histories, and are affected by properties of the sources. In particular, the deep absorption trough in the global 21-cm signal (around $z \sim 20$ in the figure) is visible only thanks to

the stellar Ly-$\alpha$ photons and is a direct evidence of the initially cold IGM which was gradually heated by X-rays. Features of the absorption trough depend on the properties of stars and X-ray sources and can be used to put constraints on the first luminous populations. Our poor understanding of high-redshift astrophysics translates into a vast landscape of plausible 21-cm signals in addition to what is shown in Figure 1. A large compilation of possible scenarios was presented recently by Cohen et al. (2017)[8] and Cohen et al. (2018)[9]. Here we focus on the global signal as the probe of the high-redshift environment and star formation.

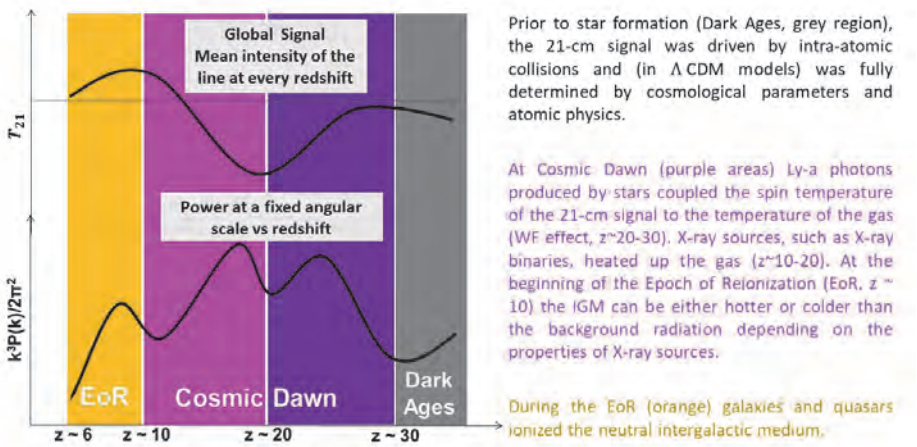



Fig. 1. Cartoon of the evolution of the global 21-cm signal (top) and of the power spectrum (bottom, shown at a fixed comoving angular scale) as a function of redshift. Cosmic epochs are shown with different colors including the Epoch of Reionization (EoR, orange), Cosmic Dawn (purple/magenta), and Dark Ages (grey). At each epoch, different types of sources had the dominant impact on the 21-cm signal (see the text on the right of the plot).

## 3. Global Signal Observations

The global 21-cm signal (computed by averaging Eq. (1) over the entire sky) is targeted by several observational teams around the world including the EDGES collaboration with the observational site located in Western Australia, SARAS based in India, LEDA observing from Owens Valley (California), PRIZM located on Marion Island. All these telescopes feature a very wide beam that measures the total sky radio signal composed of the cosmic microwave background which carries the imprint of neutral hydrogen at the intrinsic 21 cm wavelength, foreground emission of the Galaxy, integrated contribution of numerous radio sources as well as radio frequency interference. The signal is distorted by ionosphere and is convolved with

the instrumental beam. Extracting the cosmological contribution is very challenging and requires superb understanding of each component (e.g., calibration of the instrument[25]).

### 3.1. *Probing the EoR*

Observations with the EDGES High-Band antenna[24,26,27] and SARAS2[33,34] above 100 MHz yield consistent upper limits on the 21-cm signal from the EoR. In the framework of standard astrophysical modelling, the non-detections suggest that stars formed in small dark matter halos[27] and rule out extremely non-efficient X-ray heating sources[27,33,34]. As we have pointed out in Monsalve et al. (2019)[27], out of all existing observations (including the CMB optical depth[1] and constraints on the neutral fraction from high-redshift quasars[2,19,20]), properties of the high-redshift X-ray sources are uniquely constrained by the 21-cm data.

### 3.2. *Cosmic Dawn Signal*

Recently, the first detection of the sky-averaged 21-cm signal from $z \sim 13 - 27$ has been claimed based on two years of observations with the EDGES Low-Band antenna in the $50 - 100$ MHz frequency range[6]. The reported cosmological signal is centred at $\nu = 78.2$ MHz corresponding to $z = 17.2$, and features an absorption trough of $T_{21} = -500^{+200}_{-500}$ mK with the errors corresponding to 99% confidence including both thermal and systematic noise. If confirmed, this is the first (and only) observational evidence of the primordial star formation at $z \sim 20$ ($\sim 180$ million years after the Big Bang) and of the early X-ray heating at $z \sim 17$.

The frequency and width of the detected feature are in agreement with theoretical expectations requiring star formation in relatively small halos and subsequent efficient X-ray heating. Therefore, the location and width of the detected signal in the Low Band is consistent with the non-detections by EDGES High-Band and SARAS2.

However, the EDGES Low signal is at least twice as strong as predicted in standard astrophysical scenarios (based on the assumption of ΛCDM cosmology and hierarchical structure formation and assuming the CMB as the background radiation) where the strongest possible feature at $z = 17$ is $-209$ mK, with $T_{\rm CMB} = 49.5$ K and the coldest possible temperature of the IGM of $\sim 7.2$ K in the absence of X-ray heating sources. The observed $T_{21} < -300$ mK requires either the gas to be much colder (around 5.2 K) or the background radiation to be much stronger (around 67.2 K). Both requirements are impossible to satisfy by changing astrophysical modelling.

To explain the depth of the reported feature, exotic mechanisms have to be invoked. The amplitude of the observed 21-cm profile can be explained by non-gravitational interaction between dark matter and baryons, e.g., via Rutherford-like scattering, first proposed by Barkana (2018)[4] as a solution to the EDGES Low-Band anomaly. Because dark matter is cold, scattering could drain excess energy

from the gas lowering its kinetic temperature[11,28]. Even though this scenario is strongly constrained by observations, it is not completely ruled out and is still plausible for a narrow range of parameters including the dark matter mass, electric charge of dark matter particle and cross-section[5,23]. A smoking gun signature of the baryon-dark matter scattering with a velocity-dependent cross-section was shown to be an enhanced pattern of Baryon Acoustic Oscillations (BAO) in the 21-cm power spectrum[15,30]. The power spectrum itself could be boosted by as much as three orders of magnitude calling for verifications of the EDGES Low signal by interferometers.

An excess radio background above the CMB at the rest-frame 1.42 GHz at $z = 17$, even though is not expected in most theories, would also explain the detected feature[6,12]. Interestingly, evidence of excess radio background at low radio frequencies was reported by ARCADE2 at $3 - 90$ GHz[18] and recently confirmed by LWA1 at $40 - 80$ MHz[10]. However, it is still not clear what part of the observed excess is extragalactic[36]. The effect of an extra radio background is to deepen and widen the absorption trough. In our exploration[16] we found that the reported global signal is consistent with values of the amplitude of the excess radio background of $1.9 - 418$ relative to the CMB at the 78 MHz reference frequency (corresponding to $0.1 - 22\%$ of the CMB at 1.42 GHz). With the additional radio background, the EDGES Low-Band detection requires efficient star formation occurring in small dark matter halos of $M_h \sim 7.8 \times 10^8$ $M_\odot$ at $z = 17$ (or lower), which is consistent with the non-detections by High-Band instruments. If exists, the excess radio background is also expected to boost the 21-cm power spectrum compared to the standard astrophysical scenarios[16]. However, compared to the case of baryon-dark matter scattering, the power spectrum in the case of the extra radio background has a smooth shape (no BAO) and is not as strongly enhanced on large scales, as can be seen in Figure 5 of Fialkov & Barkana (2019)[16].

### 3.3. *Follow-ups and Concerns*

The reported global 21-cm signal is stronger than expected, implying that fluctuations in the signal might be stronger than predicted within standard astrophysical modelling. Therefore, low frequency observations with interferometers, such as LOFAR, which is currently conducting observations in the Low Band within the framework of the ACE proposal (1000 hours), and, once operational, the SKA, HERA and NenuFAR, will be able to verify the discovery. Interferometers can assess the nature of the enhanced 21-cm signal (if confirmed) by measuring the amplitude of the BAO. Verification with independent global signal experiments is also on the way using EDGES Mid-Band antenna, SARAS3, LEDA, and PRIZM. However, part of the community remains sceptical regarding the interpretation of the EDGES Low data, and concerns have been voiced in the literature[7,22,35].

### 3.4. *Conclusions*

With the first observational constraints at high radio frequencies (above 100 MHz) by EDGES High-Band and SARAS2, and the first claimed detection from $z \sim 17$, this is a truly exciting time for 21-cm cosmology. More than a year after the EDGES team claimed discovery of the global 21-cm signal from Cosmic Dawn, the nature and origin of this signal are still mysterious. If this signal is genuine, it is the first evidence of primordial star formation some 200 million years after the Big Bang.

**Acknowledgments**

Currently AF is a University Research Fellow at Sussex University and is supported by the Royal Society.

**References**

1. Planck Collaboration: Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., et al., 2018, arXiv:1807.06209.
2. Bañados, E., Venemans, B. P., Mazzucchelli, C., Farina, E. P., Walter, F., et al., 2018, Nature, 553, 473.
3. Barkana, R., 2016, PR, 645, 1.
4. Barkana, R., 2018, Nature, 555, 71.
5. Berlin, A., Hooper, D., Krnjaic, G., McDermott, S. D., 2018, PRL, 121, 1102.
6. Bowman, J. D., Rogers, A. E. E., Monsalve, R. A., Mozdzen, T. J., Mahesh, N., 2018, Nature, 555, 67.
7. Bradley, R. F., Tauscher, K., Rapetti, D., Burns, J. O., 2019, ApJ, 874, 153.
8. Cohen, A., Fialkov, A., Barkana, R., Lotem, M., 2017, MNRAS, 472, 1915.
9. Cohen, A., Fialkov, A., Barkana, R., 2018, MNRAS, 478, 2193.
10. Dowell, J., Taylor, G. B., 2018, ApJL, 858, 9.
11. Dvorkin, C., Blum, K., Kamionkowski, M., 2014, PRD, 89, 3519.
12. Feng, C., Holder, G., 2018, ApJL, 858, 17.
13. Fialkov, A., Barkana, R., 2014, MNRAS, 445, 213.
14. Fialkov, A., Barkana, R., Visbal, E., 2014, Nature, 506, 197.
15. Fialkov, A., Barkana, R., Cohen, A., 2018, PRL, 121, 1.
16. Fialkov, A., Barkana, R., 2019, arXiv:1902.02438.
17. Field, G. B., 1958, Proc. IRE 46 (1958) 240.
18. Fixsen, D. J., Kogut, A., Levin, S., Limon, M., Lubin, P., Mirel, P., Seiffert, M., Singal, J., Wollack, E., Villela, T., Wuensche, C. A., 2011, ApJ, 734, 5.
19. Greig, B., Mesinger, A., 2017, MNRAS, 465, 4838.
20. Greig, B., Mesinger, A., Bañados, E., 2018, arXiv:1807.01593.
21. Hashimoto, T., Laporte, N., Mawatari, K., Ellis, R. S., Inoue, A. K., et al., 2018, Nature, 557, 392.
22. Hill, J. C., Baxter, E. J., 2018, arXiv:1803.07555.

23. Kovetz, E. D., Poulin, V., Gluscevic, V., Boddy, K. K., Barkana, R., Kamionkowski, M., 2018, PRD, 98, 10.
24. Monsalve, R. A., Rogers, A. E. E., Bowman, J. D., Mozdzen, T. J., 2017, ApJ, 847, 64.
25. Monsalve, R. A., Rogers, A. E. E., Bowman, J. D., Mozdzen, T. J., 2017, ApJ, 835, 49.
26. Monsalve, R. A., Greig, B., Bowman, J. D., Mesinger, A., Rogers, A. E. E., Mozdzen, T. J., Kern, N. S., Mahesh, N., 2018, ApJ, 863, 11.
27. Monsalve, R. A., Fialkov, A., Bowman, J. D., Rogers, A. E. E., Mozdzen, T. J., Cohen, A., Barkana, R, Mahesh, N., 2019, ApJ.
28. Muñoz, J. B., Kovetz, E. D., Ali-Ha´imoud, Y., 2015, PRD, 92, 3528.
29. Muñoz, J. B., Loeb, A., 2018, Nature, 557, 684.
30. Muñoz, J. B., Dvorkin, C., Loeb, A., 2018, PRL, 121l1301.
31. Oesch, P., Brammer, G., van Dokkum, P. G., Illingworth, G. D., Bouwens, R. J., Labbé, I., Franx, M., Momcheva, I., Ashby, M. L. N., Fazio, G. G., Gonzalez, V., Holden, B., Magee, D., Skelton, R. E., Smit, R., Spitler, L. R., Trenti, M., Willner, S. P., 2016, ApJ, 819, 129.
32. Pacucci, F., Mesinger, A., Mineo, S., Ferrara, A., 2014, MNRAS, 443, 678.
33. Singh, S., Subrahmanyan, R., Udaya Shankar, N., Sathyanarayana Rao, M., Fialkov, A., Cohen, A., Barkana, R., Girish, B. S., Raghunathan, A., Somashekar, R., Srivani, K. S., 2017, ApJL, 845, 12.
34. Singh, S., Subrahmanyan, R., Udaya Shankar, N., Sathyanarayana Rao, M., Fialkov, A., Cohen, A., Barkana, R., Girish, B. S., Raghunathan, A., Somashekar, R., Srivani, K. S., 2018, ApJ, 858, 54.
35. Singh, S., Subrahmanyan, R., 2019, submitted.
36. Subrahmanyan, R., & Cowsik, R., 2013, ApJ, 776, 42.
37. Wouthuysen, S. A., 1952, Astron. J. 57 32.

# Virial clouds to explain rotational asymmetry in galactic halos

Asghar Qadir*

*School of Natural Sciences, National University of Science and Technology, H-12 Islamabad, 46000, Pakistan*
**E-mail: asgharqadir46@gmail.com*

Noraiz Tahir

*School of Natural Sciences, National University of Science and Technology, H-12 Islamabad, 46000, Pakistan*
*E-mail: noraiztahir78637@gmail.com*

Francesco De Palois

*Department of Mathematics and Physics "E. De Giorgi", University of Salento, Via per Arnesano, I-73100 Lecce, Italy and*
*INFN, Sezione di Lecce, Via per Arnesano, I-73100 Lecce, Italy*
*Email: depaolis@le.infn.it*

Achille A. Nucita

*Department of Mathematics and Physics "E. De Giorgi", University of Salento, Via per Arnesano, I-73100 Lecce, Italy and*
*INFN, Sezione di Lecce, Via per Arnesano, I-73100 Lecce, Italy*
*Email: nucita@le.infn.it*

The rotation of the galactic objects has been seen by asymmetric Doppler shift in the CMB data. Molecular hydrogen clouds at virial temperature may contribute to the galactic halo dark matter and they might be the reason for the observed rotational asymmetry in the galactic halos. We present a method to constrain the parameters of these virial clouds given that they are composed of a single fluid. The method is such that it should be possible to extend it to more than one fluid.

*Keywords*: Cosmic microwave background; virial theorem; dark matter; molecular hydrogen clouds.

## 1. Introduction

Half of the baryonic matter in the Universe is dark,[1–4] some of it may consist of molecular hydrogen clouds left over from the process of star formation, which would continue to collapse till their virial temperature reaches the CMB. To observe them one could look for $\gamma$-ray scintillation due to strikes by cosmic rays on $H_2$ molecules. A $\gamma$-ray corona around the Milky Way *was* discovered[5] with the expected angular distribution,[6] but it was not not clear how much, if any, of the carpet might be due to the proposed clouds. An alternative suggestion was to look for an asymmetric Doppler shift of these clouds in M31.[7]

The predicted effect was seen at a sufficient but low level in the WMAP data in 2011[8] and confirmed in the Planck data in 2014.[9] Objections that the blue shift may be due to "hot spots" were countered by the argument that the red shift could not be "cold spots" and were soon dispelled. The precision was sufficient to be

able to use for studying the rotation of the M31. This was followed by the same effect being seen in NGC 5128, M33, M82, M81.[10–13] However, the verification of the effect that it is due to the molecular clouds which led to its prediction, does not prove that it is due to the cause that led to the prediction. It could be partially or totally due to other interstellar matter. To be able to determine the nature of the cause, we need to model the molecular hydrogen clouds or dust clouds produced by the mechanism mentioned, or some mixture of the two, and compare with the data. One could then try to look at whether contamination of the spectrum by radiation from higher temperature sources could be distinguished.

## 2. The Virial Cloud Model

As these clouds are immersed in a heat bath of the CMB, they are isothermal. Since they have formed due to the Jeans instability, they must have a precisely defined Jeans radius and consequent Jeans mass. The density profile would then be given by the solution of the Lane-Emden equation subject to the boundary conditions that the density be literally *zero* (not approximately zero) at the boundary and the density profile be flat at the core. The latter condition says that the derivative of the density be zero at the core. We call them "virial clouds". Since James Jeans had applied the analysis for the formation of a normal star, he had taken the boundary condition that the density of the cloud merge with that of the interstellar medium (ISM) (which has been called "the Jeans fiddle").[14] In that case the central density had to be put in on an ad hoc basis. We do not need to do this. Further, we will want to extend the analysis (in later work) to a two-fluid model. As such, we obtain the equations from first principles in any case.

For an isothermal gas sphere, the temperature, $T$, is directly related to the mass of the molecule, $m_H$ in our case, and the average speed. The speed of the perturbation would have to be taken to be the speed of sound in the cloud, $c_s$. This gives

$$c_s = \sqrt{\frac{\gamma k T_{CMB}}{m_H}} \ , \tag{1}$$

where $\gamma$ is the adiabatic factor given by the ratio of the heat capacity at constant pressure to that at constant volume. At the densities and temperatures that we are considering no higher degrees of freedom would be excited and we would only have the translational degree, making the ideal gas approximation extremely good and so $\gamma = 5/3$. For our purpose, $m_H \approx 2.016$ g/mol and $T \approx 2.726^o K$, so

$$c_s \simeq 1.110 \times 10^4 \text{ km s}^{-1} \ . \tag{2}$$

Now the virial theorem $2K + U = 0$ gives the Jeans mass as[15]

$$M_J^2 \simeq \left(\frac{81}{32\pi\rho_c}\right)\left(\frac{3c_s^2}{5G}\right)^3 \ , \tag{3}$$

and the corresponding Jeans radius as

$$R_J^2 = \frac{27c_s^2}{20\pi\rho_c G} \ , \tag{4}$$

where, $\rho_c$ is the central density of the virial cloud, and $G$ is Newton's gravitational constant. We need to use the canonical ensemble distribution in order to calculate the density profile of the clouds which is given by the relation[16]

$$f(r,p) = \frac{1}{h^{3N}N!}\frac{1}{Z}exp\left(-\frac{H(r,p)}{kT_{CMB}}\right) \ , \tag{5}$$

where, $H(r,p)$ is the Hamiltonian of the system. The partition function $Z$ is the normalization factor, thus being the integral over phase space of the rest of this distribution function. The Hamiltonian is position dependent

$$H(r,p) = \frac{p^2}{2m_H} + \frac{GM(r)m_H}{r} \ , \tag{6}$$

and so the density distribution, is given by

$$\rho(r) = 8m_H^{5/2}\left(\frac{G\ \rho_c}{3k\ T_{CMB}}\right)^{3/2} exp\left(-\frac{GM(r)m_H}{rkT_{CMB}}\right), \tag{7}$$

where $M(r)$ is the mass of the cloud interior to $r$ given by

$$GM(r) = 4\pi\int_0^r \rho(q)q^2dq \ . \tag{8}$$

Our problem, then, is that the density is defined in terms of the integral of the density, i.e. we have an integral equation. Taking the natural logarithm of Eq. (7) we obtain

$$\int_0^r q^2\rho(q)dq = -\frac{rkT_{CMB}}{4\pi Gm_H}\ln\left(\frac{\rho_(r)}{\zeta}\right) \ , \tag{9}$$

where $\zeta = 8m_H^{5/2}(G\rho_c/3kT_{CMB})^{3/2}$. Taking the derivative of the above equation with respect to $r$, and substituting Eq. (7) we obtain the ordinary differential equation

$$r\frac{d\rho(r)}{dr} - r^2\left(\frac{4\pi Gm_H}{kT_{CMB}}\right)\rho^2(r) - \rho(r)\ln\left(\frac{\rho(r)}{\zeta}\right) = 0 \ . \tag{10}$$

It will be noticed that this is a nonlinear differential equation that cannot be solved by separation of variables. However, we do have the initial condition $\rho'(r)|_{r=0} = 0$ and the boundary condition that $\rho(R_J) = 0$. As such, we can solve it numerically by assuming a central density and the initial condition. If the density at the boundary is too high, we lower the central value and try again, till it becomes too low. At that stage we raise the central density a bit till it again becomes too high. Iteratively we bring the boundary density close enough to zero to be acceptable. The resulting density profile is shown in Fig. 1. One sees that the central density of the cloud is $\rho_c \simeq 1.60 \times 10^{-18}\ kgm^{-3}$, the radius is $R_J \simeq 0.030$pc, and the mass is $M_J \simeq 0.798M_\odot$.

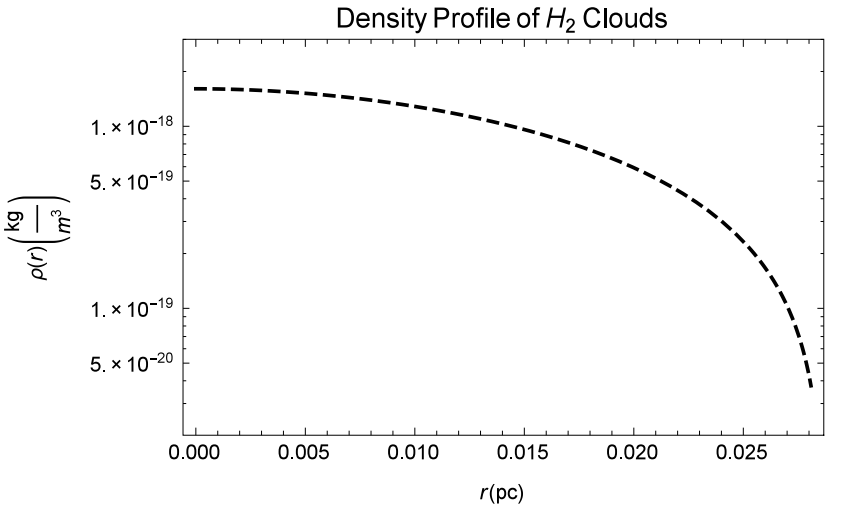



Fig. 1. The density of the virial molecular hydrogen cloud. It decreases monotonically from the central density, $\rho_c \simeq 1.60 \times 10^{-18}\ kgm^{-3}$, to zero at $r = R_J \simeq 0.030$ pc.

## 3. Luminosity of the Clouds

Since the virial clouds are at the CMB temperature, their luminosity is the CMB luminosity, except for their Doppler shift. Thus, we only need to look at the differential frequency shift, $\Delta\nu/\nu = -(v_{rot}/c)\cos\theta$, where $v_{rot}$ is the rotational velocity of the galactic halo, $c$ the speed of light and $\theta$ the angle between the direction of the velocity and the line of sight. Now $\Delta\nu/\nu = \Delta E/E = \Delta T/T$, so the differential energy is given by $\Delta E = (\Delta T/T)E$, where $E$ is the energy in the radiation for a patch of the sky corresponding to the size of the cloud. Now, as the luminosity, $L$, is proportional to the energy, we get the luminosity of the CMB for that patch. Then the luminosity of the cloud will be[17]

$$L_c = (\Delta\nu/\nu)L \simeq 1.486 \times 10^{-4} L\cos\theta\ . \tag{11}$$

Thus, as we scan across the halo, the luminosity will not change linearly but with $\cos\theta$. We can use this check for the Doppler shift explanation and obtain the energy associated with each cloud. This would directly give an estimate of the total number of clouds being seen.

## 4. Results and Discussions

The virial clouds are completely constrained by physical requirements, and so precisely determinable. They provide a satisfactory explanation of the observed asymmetry of the CMB Doppler shift in galactic halos. We have only considered pure molecular hydrogen clouds, but we also need to model these molecular hydrogen clouds contaminated with more or less interstellar dust which will be dealt with

elsewhere. We have not discussed the possible contamination of radiation by matter at high temperature. That needs to be done. The effect is probably negligible, but cannot be neglected without checking. One also needs to investigate how the clouds would evolve over billions of years, from the time when the CMB temperature was higher than it is now to the present value. Precisely how did the clouds lose energy over this period?

**Acknowledgments**

We are grateful to the Salento University of Lecce for hospitality at Lecce, where part of this work was done, and AQ is grateful to IUPAP, ESA, ICRANet and AS-ICTP for support to participate in MG15.

**References**

1. G. Bertone, *Particle dark matter: observations, models and searches* (Cambridge University Press, 2010).
2. R. Cen and J. P. Ostriker, Where are the baryons? *The Astrophysical Journal* **514**, p. 1 (1999).
3. R. Cen and J. P. Ostriker, Where are the baryons? ii. Feedback effects, *The Astrophysical Journal* **650**, p. 560 (2006).
4. A. Fraser-McKelvie, K. A. Pimbblet and J. S. Lazendic, An estimate of the electron density in filaments of galaxies at $z \sim 0.1$, *Monthly Notices of the Royal Astronomical Society* **415**, 1961 (2011).
5. D. Dixon, D. Hartmann, E. Kolaczyk, J. Samimi, R. Diehl, G. Kanbach, H. Mayer-Hasselwander and A. Strong, Evidence for a Galactic gamma-ray halo, *New Astronomy* **3**, 539 (1998).
6. F. De Paolis, G. Ingrosso, P. Jetzer and M. Roncadelli, Gamma-ray astronomy and baryonic dark matter, *The Astrophysical Journal Letters* **510**, p. L103 (1998).
7. F. De Paolis, G. Ingrosso, P. Jetzer, A. Quadir and M. Roncadelli, Observing molecular hydrogen clouds and dark massive objects in galactic halos, *Astronomy & Astrophysics* **299**, p. 647 (1995).
8. F. De Paolis, V. Gurzadyan, G. Ingrosso, P. Jetzer, A. Nucita, A. Qadir, D. Vetrugno, A. Kashin, H. Khachatryan and S. Mirzoyan, Possible detection of the M31 rotation in WMAP data, *Astronomy & Astrophysics* **534**, p. L8 (2011).
9. F. De Paolis, V. Gurzadyan, A. Nucita, G. Ingrosso, A. Kashin, H. Khachatryan, S. Mirzoyan, E. Poghosian, P. Jetzer, A. Qadir et al., Planck confirmation of the disk and halo rotation of M31, *Astronomy & Astrophysics* **565**, p. L3 (2014).
10. F. De Paolis, V. Gurzadyan, A. Nucita, G. Ingrosso, A. Kashin, H. Khachatryan, S. Mirzoyan, G. Yegorian, P. Jetzer, A. Qadir et al., Planck revealed bulk motion of Centaurus a lobes, *Astronomy & Astrophysics* **580**, p. L8 (2015).

11. V. Gurzadyan, F. De Paolis, A. Nucita, G. Ingrosso, A. Kashin, H. Khachatryan, S. Sargsyan, G. Yegorian, P. Jetzer, A. Qadir et al., Planck view of the M82 galaxy, *Astronomy & Astrophysics* **582**, p. A77 (2015).
12. F. De Paolis, V. Gurzadyan, A. Nucita, L. Chemin, A. Qadir, A. Kashin, H. Khachatryan, S. Sargsyan, G. Yegorian, G. Ingrosso et al., Triangulum galaxy viewed by Planck, *Astronomy & Astrophysics* **593**, p. A57 (2016).
13. V. Gurzadyan, F. De Paolis, A. Nucita, A. Kashin, A. Amekhyan, S. Sargsyan, G. Yegorian, A. Qadir, G. Ingrosso, P. Jetzer et al., Messier 81's Planck view versus its halo mapping, *Astronomy & Astrophysics* **609**, p. A131 (2018).
14. J. H. Jeans, The stability of a spherical nebula, *Proceedings of the Royal Society of London* **68**, 454 (1901).
15. S. Chandrasekhar, *An introduction to the study of stellar structure* (Courier Corporation, 1957).
16. P. M. Morse and B.W. Shore, Thermal physics, *Physics Today* **17**, p. 60 (1964).
17. N. Tahir, F. De Paolis, A. Qadir and A. Nucita, Constraining baryons in the M31 halo using Planck data, *International Journal of Modern Physics D*, accepted in (2019).

# Diffuse cosmic dipoles

Tiziana Trombetti*

*INAF, Istituto di Radioastronomia*
*Via Piero Gobetti, 101, 40129 Bologna, Italy*
**E-mail: trombetti@ira.inaf.it*

The boosting effects induced by the peculiar motion of an observer with respect to the Cosmic Microwave Background (CMB) rest frame can be explored to analyze the frequency dependence of the dipole. The improvements achievable with future CMB missions on our knowledge of CMB spectral distortions and Cosmic Infrared Background spectrum are discussed considering realistic uncertainties in relative calibration and foreground subtraction.

*Keywords*: CMB experiments, spectral distortions, reionization.

## 1. Introduction

The cosmic dipoles have a crucial relevance in cosmology since they are linked with the isotropy and homogeneity of the Universe at the largest scales. The observed dipole is a mixture of different contributions, including the observer motion with respect to the Cosmic Microwave Background (CMB) rest frame as well as dipoles from astrophysical (extragalactic and Galactic) sources. The CMB and Cosmic Infrared Background (CIB) spectra can be investigated by future anisotropy missions, as CORE[1] or LiteBIRD,[2] by exploring the frequency spectral behaviour of the dipole amplitude, thus without requiring an independent absolute calibration. Tiny CMB spectral distortions are predicted to be generated at different cosmic times in the evolving plasma. In the present work, I will focus on two well-defined types of signal, namely Bose-Einstein (BE) and Comptonization distortions. The precise interpretation of these signals could contribute to constraint primordial cosmological processes otherwise unexplored. The CIB spectrum amplitude and shape, still not well known, can provide a better understanding of the dust-obscured star-formation phase of galaxy evolution.

## 2. The CMB spectrum

One of the main features of the CMB is its frequency spectrum. To first approximation, it emerges from the thermalization epoch ($z \sim 10^6 - 10^7$) with a thermal black body (BB) shape owing to highly efficient interaction processes in the cosmic plasma, able to re-establish matter-radiation thermal equilibrium. Since COBE/FIRAS, no remarkable improvements have been achieved in the knowledge of CMB spectrum at $\nu \gtrsim 30$ GHz. Although gaining an absolute calibration precision of 0.57 mK, the FIRAS characterization of CIB amplitude and shape still presents a substantial uncertainty. To overcome the absolute calibration problem, it is possible to investigate the CMB and CIB spectra by looking at the frequency spectral behaviours of their dipole amplitude thanks to future CMB missions, as e.g.

CORE. The CMB spectrum can set constraints on various types of non-standard processes, including non evaporating BH spin, small scale magnetic fields power spectra, vacuum energy density and particle decay.

### 2.1. *Spectral distortions*

Different energy injections in the radiation field due to fundamental unavoidable processes occurring at different cosmic times cause the departure of CMB spectrum from a perfect BB.

At early times, a BE-like distorted spectrum with a positive (dimensionless) chemical potential is produced by the dissipation of primordial perturbations at small scales,[3] damped by photon diffusion and non-linear processes and then invisible in CMB anisotropies.[4] Moreover, because of the matter temperature faster decrease with respect to that of radiation temperature in an expanding Universe, colder electrons induce Bose condensation of CMB photons, mainly resulting into a BE-like distortion with negative chemical potential. The photon occupation number of the BE spectrum is $\eta_{BE} = 1/(e^{x_e+\mu} - 1)$, where $\mu$ is the chemical potential that quantifies the fractional energy, $\Delta\varepsilon/\varepsilon_i$, exchanged in the plasma during the interaction, $x_e = x/\phi(z)$, $\phi(z) = T_e(z)/T_{CMB}(z)$, being $T_e(z)$ the electron temperature, $T_{CMB} = T_0(1+z)$, $T_0$ the present CMB temperature in the BB approximation and $x = h\nu/kT_{CMB}$ a redshift independent, dimensionless frequency. For small distortions, $\mu \simeq 1.4\Delta\varepsilon/\varepsilon_i$ and $\phi_{BE} \simeq (1 - 1.11\mu)^{-1/4}$.[a]

Cosmological reionization associated with the early stages of structure and star formation induces electron heating that is responsible for late type distortions, described, at high frequency,[b] by a Comptonization spectrum.[7] It is quantified by the Comptonization parameter $u(t) = \int_{t_i}^{t}[(\phi - \phi_i)/\phi](k_B T_e/m_e c^2) n_e \sigma_T c dt$. For small energy injections and integrating over the relevant epochs $u \simeq (1/4)\Delta\varepsilon/\varepsilon_i$.

## 3. Future CMB missions

Among the many different classes of CMB missions, here I will focus on three recent proposals with concepts and designs aimed at improving the measure of polarization anisotropies and at minimizing the impact of foregrounds on CMB maps using a large number of frequency channels, with relevant outcomes for e.g. our knowledge of the reionization process and the primordial background of gravitational waves.

The Primordial Inflation Explorer[8] (PIXIE) is an Explorer-class mission submitted to NASA in 2011 aimed at mapping the CMB and the diffuse astrophysical foregrounds over the full sky, from 30 GHz to 6 THz (1 cm to 50 $\mu$m) with degree

[a] At decreasing frequencies, photon emission processes produce a frequency dependent, decreasing chemical potential and the convergency to a blackbody at matter temperature at extremely low frequencies.[5] For positive (negative) $\mu$ they imply a minimum (maximum) in the brightness temperature, as well as the avoiding of the divergency of the BE spectrum for negative $\mu$.

[b] Associated free-free distortions are instead relevant at low frequencies.[6]

resolution and high signal to noise ratio in absolute intensity and linear polarization. It is designed to perform absolute spectroscopy measurements and to simultaneously observe the large-scale CMB B-modes. It will provide crucial constraints on Universe ionization history, on CIB spectrum and anisotropies. Indeed, PIXIE will trace CIB monopole, dipole and higher order power spectrum up to $z \sim 3$ for matter distribution analyses. The cross correlation of temperature and polarization anisotropies will produce more stringent constraints on the optical depth parameter, the gas temperature and the Universe reionization at $z \sim 10$. Great hopes on CMB spectral distortion parameters are expected from PIXIE, thanks to its absolute calibration of the CMB temperature measure with a precision about $10^3$ times better than FIRAS.

LiteBIRD is a Lite (Light) Satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection at the extremely early Universe proposed to JAXA in February 2015. It is a highly-targeted, low-cost Japanese B-mode mission concept, aimed at the B-modes detection at the level of $r \sim 10^{-3}$, its main scientific goal being one order of magnitude better than what can be done from the ground. Its primary strategy is the focus on $r$ measurement in synergy with ground based super telescopes which are necessary for an accurate subtraction of the lensing contribution to B-mode.

Cosmic ORigins Explorer (CORE), submitted to the European Space Agency (ESA) in October 2016 in response to a call for future medium-sized space mission proposals for the M5 launch opportunity of ESAs Cosmic Vision programme, is a satellite dedicated to microwave polarization. CORE is targeted to provide and exploit definitive maps of CMB polarization anisotropies at large and medium angular scales, with a suppression of all systematic effects at an extreme accuracy level. The instrument will host about 19 frequency channels over a range spanning the 60–600 GHz interval, in order to control astrophysical foreground emissions.

## 4. CMB and CIB dipoles

A relative velocity between an observer and the CMB rest frame induces a dipole in the observed CMB sky temperature through the Doppler effect. This dipole, the so called $\ell = 1$ anisotropy, is likely dominated by the velocity of the Solar System, $\vec{\beta_S} = \vec{v_S}/\vec{c}$, with respect to the CMB (Solar dipole), with a seasonal modulation due to the velocity of the Earth or the satellite, $\vec{\beta_o}$, with respect to the Sun (orbital dipole).

In addition, an intrinsic dipole is induced by the Sachs-Wolfe effect at the last-scattering surface and by a large-scale dipolar Newtonian potential.[9] In a ΛCDM cosmology, this dipole should be of order of the Sachs-Wolfe plateau amplitude ($10^{-5}$), but it could be larger in the case of more exotic models. Neglecting the orbital dipole (useful for calibration aims), and denoting with $\vec{\beta}$ the relative velocity of the Solar dipole, it is possible to forecast the CORE improvements in the dipole recovery including also potential foreground and calibration residuals.[10] The

dipole amplitude is directly proportional to the first (logarithmic) derivative with respect to the frequency of the photon occupation number, $\eta(\nu)$, which is related to the thermodynamic temperature, $T_{therm}(\nu)$, defined as the temperature of the blackbody having the same $\eta(\nu)$ at the frequency $\nu$, by:

$$T_{therm} = \frac{h\nu}{k_B \ln(1 + 1/\eta(\nu))}. \tag{1}$$

Differencing the measure of $T_{therm}$ in the direction of motion and in its perpendicular direction one gets:[11]

$$\Delta T_{therm} = \frac{h\nu}{k_B} \left\{ \frac{1}{\ln[1 + 1/\eta(\nu)]} - \frac{1}{\ln[1 + 1/\eta(\nu(1+\beta))]} \right\}. \tag{2}$$

To first order in $\beta$, $\Delta T_{therm}$ is approximated by:

$$\Delta T_{therm} \simeq -\frac{x\beta T_0}{(1+\eta)\ln^2(1+1/\eta)} \frac{d\ln\eta}{d\ln x}. \tag{3}$$

A direct determination of the CIB spectral shape is not trivial since it requires absolute intensity measurements and it is also limited by foreground signal, that, in the case of Galactic emission, has a similar shape. Even though the dipole amplitude is about $10^{-3}$ of the monopole, its spatial distribution is known, hence, an indirect approach may provide in the future a robust measurement of the CIB. The analytic form of the CIB spectrum, observed at present time, is:[12]

$$\eta_{CIB} = \frac{c^2}{2h\nu^3} I_{CIB}(\nu) = I_0 \left( \frac{k_B T_{CIB}}{h\nu_0} \right)^{k_F} \frac{x_{CIB}^{k_F}}{e^{x_{CIB}} - 1}, \tag{4}$$

with $T_{CIB} = (18.5 \pm 1.2)$ K, $x_{CIB} = h\nu/k_B T_{CIB} = 7.78(\nu/\nu_0)$, $\nu_0 = 3 \times 10^{12}$ Hz and $k_F = 0.64 \pm 0.12$. Here $I_0$ sets the CIB spectrum amplitude, its best-fit value being $(1.3 \pm 0.4) \times 10^{-5}$. The CIB dipole should be detectable by CORE in its highest frequency bands. Defining the motion vector of the observer, from the dipole direction of the *Planck* 2015 release, I produced the maps at a given observational frequency, $\nu_{obs}$, from the photon distribution function, $\eta_{BB,dist}$, for the assumed type of spectra (BB, CIB, BE or Comptonization (C)) at the frequency $\nu_{obs}$ multiplied by $(1 - \hat{n} \cdot \vec{\beta})/(1 - \beta^2)^{1/2}$ to account for all the possible sky directions with respect to the observer peculiar velocity. The observed signal map in thermodynamic temperature is a generalization of Eq. (1):

$$T_{therm}^{BB/dist}(\nu, \hat{n}, \vec{\beta}) = \frac{xT_0}{\ln(1/(\eta(\nu, \hat{n}, \vec{\beta})^{BB/dist} + 1))}, \tag{5}$$

where $\eta(\nu, \hat{n}, \vec{\beta}) = \eta(\nu')$ with $\nu' = \nu((1 - \hat{n} \cdot \vec{\beta})/(1 - \beta^2)^{1/2}$. Decomposing the maps into spherical harmonics and reproducing them from the $a_{\ell m}$ up to a desired multipole $\ell_{max}$, I derived the expected signal (see Fig. 1) finding that it is important for the dipole, can be considerable for the quadrupole and, depending on the distortion parameters, still not negligible for the octupole (although depending on the amplitude relative to experimental noise levels). For higher-order multipoles, the signal is essentially negligible.

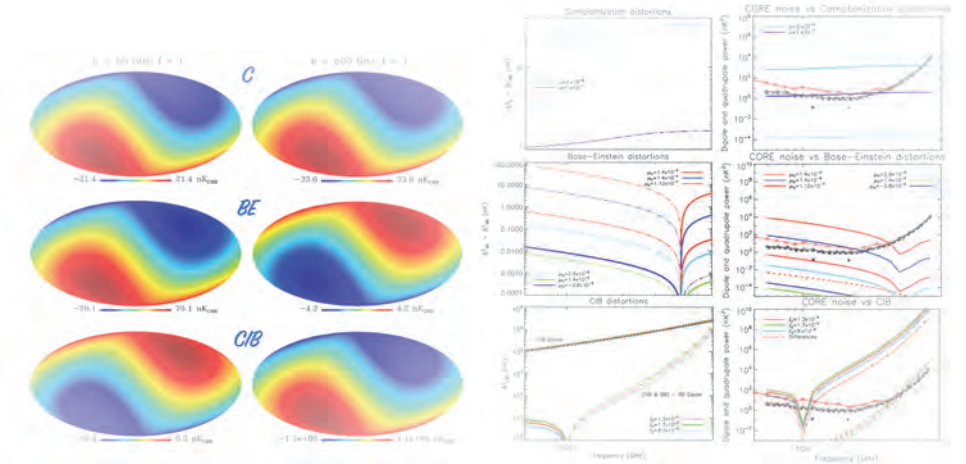



Fig. 1. Left: typical maps of dipole pattern (after the subtraction of the CMB blackbody dipole) for the three considered types of signal. Middle: dipole frequency behaviour. Right: angular power spectrum of the maps of dipole pattern compared with the sensitivity (diamonds) of CORE (black) and LiteBIRD (red). The quadrupole signal (dashes) is also shown in one case of BE distortion and in the case of CIB. See also the legend and the text. A dapted from Ref. 10 [©SISSA Medialab Srl. Reproduced by permission of IOP Publishing. All rights reserved].

## 5. Results and Conclusions

To recover the dipole parameters (amplitude and direction) it is essential to perform Markov Chain Monte Carlo (MCMC) analyses. I assumed the *Planck* common mask 76 (in temperature), publicly available from the *Planck* Legacy Archive (PLA),[13] exploiting its extension that excludes all the pixels at $|b| \leq 30°$. I first explored the dipole reconstruction ability for different frequency channels and then the implications for CMB spectral distortions and CIB starting from the ideal case (i.e., without calibration errors or sky residuals). The dipole can be parametrized as $d(\hat{n}) = A\hat{n} \cdot \hat{n}_0 + T_0$, where $\hat{n}$ and $\hat{n}_0$ are the unit vectors defined respectively by the Galactic longitudes and latitudes $(l, b)$ and $(l_0, b_0)$. The dipole map used in the simulations was generated assuming the measured dipole amplitude best-fit values, $A = (3.3645 \pm 0.002)$ mK, and direction, $l_0 = 264.00 \pm 0.03$ and $b_0 = 48.24 \pm 0.02$, as found in *Planck* 2015 release.[14]

To quantify the ideal CORE sensitivity to spectral distortion parameters and CIB amplitude, I carried out detailed simulations assuming a certain model and quantifying the accuracy level at which (in the presence of noise and of potential residuals) the key parameters can be recovered. I consider twelve, physically or observationally motivated, reference cases, from a (reference) blackbody spectrum, a CIB spectrum at the FIRAS best-fit amplitude, BE and C spectra with different $\mu_0$ and $u$ values. For each model I generate an ideal sky (the prediction) and a sky with white Gaussian noise realizations (simulated data). I calculated $\Delta\chi^2$ to quantify the significance level at which each model can be potentially detected or

ruled out. I consider three approaches: (a) analyzing each of the 19 frequency channels (assumed independent from each other); (b) using the 171 (19 ·18/2) combinations of the maps differences from pairs of frequency bands; (c) combining (a) and (b).

Approach (a) compares the dipole amplitude of a distorted spectrum with that of the blackbody, being sensitive to the overall difference between the two cases, while approach (b) compares the dipole signal at different frequencies for each type of spectrum, being so sensitive to its slope. The significance of the CIB amplitude recovery is increased in case b), because of the very steep frequency shape of its dipole spectrum. In general, this does not occur for CMB distortion parameters, and, in particular, approach (b) can make the recovery of the Comptonization distortion more difficult. Method c) typically results in an overall advantage. Including potential residuals from imperfect foreground subtraction and calibration (characterized by effective parameters $E_{\rm cal}$ and $E_{\rm for}$) may affect these results, but still the improvement for $I_0^{CIB}$ ranges from a factor of 4 (a) to a factor of about 15 or 20 for (b) and (c), respectively. The recovery of CMB spectral distortion parameters is also very promising. Including the extended mask in case c), I find a significant improvement with respect to the full sky: the significance of the CIB amplitude recovery improves by about 50% and that on the BE distortion by about 20%, indicating the relevance of optimising the selection of the sky region. Tab. 1 summarizes the main results expressed in terms of a factor characterizing the improvement with respect to FIRAS.

Table 1. Predicted improvement in the recovery of the distortion parameters with respect to FIRAS for different calibration and foreground residual assumptions. "P06" stands for the *P*lanck common mask, while "P06ext" is the extended P06 mask. When not stated, all values refer to $E_{\rm cal}$ and $E_{\rm for}$ at $N_{\rm side} = 64$. From Ref. 10 [].

| | $E_{\rm cal}$ (%) | $E_{\rm for}$ (%) | CIB amplitude | BE | C |
|---|---|---|---|---|---|
| Ideal case, all sky | - | - | $\simeq 4.4 \times 10^3$ | $\simeq 10^3$ | $\simeq 6.0 \times 10^2$ |
| All sky | $10^{-4}$ | $10^{-2}$ | $\simeq 15$ | $\simeq 42$ | $\simeq 18$ |
| P76 | $10^{-4}$ | $10^{-2}$ | $\simeq 19$ | $\simeq 42$ | $\simeq 18$ |
| P76ext | $10^{-2}$ | $10^{-2}$ | $\simeq 17$ | $\sim 4$ | $\sim 2$ |
| P76ext | $10^{-4}$ | $10^{-2}$ | $\simeq 22$ | $\simeq 47$ | $\simeq 21$ |
| P76ext | $10^{-4}$ | $10^{-3}$ | $\simeq 2.1 \times 10^2$ | $\simeq 2.4 \times 10^2$ | $\simeq 1.1 \times 10^2$ |
| P76ext | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-2}$ | $\simeq 19$ | $\simeq 26$ | $\simeq 11$ |
| P76ext | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-3}$ | $\simeq 48$ | $\simeq 35$ | $\simeq 15$ |
| P76ext, $N_{\rm s} = 128$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-2}$ | $\simeq 38$ | $\simeq 51$ | $\simeq 23$ |
| P76ext, $N_{\rm s} = 128$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-3}$ | $\simeq 43$ | $\simeq 87$ | $\simeq 39$ |
| P76ext, $N_{\rm s} = 256$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-2}$ | $\simeq 76$ | $\simeq 98$ | $\simeq 44$ |
| P76ext, $N_{\rm s} = 256$ | $10^{-3}_{(\leq 295)}$–$10^{-2}_{(\geq 340)}$ | $10^{-3}$ | $\simeq 85$ | $\simeq 1.6 \times 10^2$ | $\simeq 73$ |

**Acknowledgments**

I acknowledge partial supports from the INAF PRIN SKA/CTA project FORE-CaST, from ASI/INAF agreement n. 2014-024-R.1 for the *Planck* LFI Activity of Phase E2 and from the ASI/Physics Department of the University of Roma–Tor Vergata agreement n. 2016-24-H.0 for study activities of the Italian cosmology community. It is a pleasure to thank the many colleagues of the *Planck* and CORE collaborations for constructive discussions. Some of the results in this paper have been derived using the HEALPix[15] package.

**References**

1. Delabrouille, J., de Bernardis, P., Bouchet, F. R., et al. 2018, JCAP, 4, 014
2. Ishino, H., Akiba, Y., Arnold, K., et al. 2016, Proc. SPIE, 9904, 99040X
3. Hu, W., Scott, D., & Silk, J. 1994, ApJL, 430, L5
4. Chluba, J., & Sunyaev, R. A. 2012, MNRAS, 419, 1294
5. Danese, L., & de Zotti, G. 1982, A&A, 107, 39
6. Trombetti, T., & Burigana, C. 2014, MNRAS, 437, 2507
7. Ya. B. Zel'dovich, A.F. Illarionov, R.A. Syunyaev, JETP, 1972, Vol. 35, No. 4, p. 643
8. Kogut, A., Fixsen, D. J., Chuss, D. T., et al. 2011, JCAP, 7, 025
9. Roldan, O., Notari, A., & Quartin, M. 2016, *JCAP*, 6, 026
10. Burigana, C., Carvalho, C. S., Trombetti, T., et al. 2018, JCAP, 4, 021
11. Danese, L., & de Zotti, G. 1981, A&A, 94, L33
12. Fixsen, D. J., Dwek, E., Mather, J. C., Bennett, C. L., & Shafer, R. A. 1998, ApJ, 508, 123
13. Planck 2015 Release Explanatory Supplement, `https://www.cosmos.esa.int/documents/387566/387639/Planck_2015_results_explanatory_supplement.pdf/8b18ecea-49b2-45fa-89e4-9b74e9fb051b`
14. Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016, *A&A*, 594, A1
15. Górski, K. M., Hivon, E., Banday, A. J., et al. 2005, ApJ, 622, 759

# New extragalactic research paths opened by *Planck*

Gianfranco De Zotti*

*INAF-Osservatorio Astronomico di Padova,*
*Vicolo dell'Osservatorio 5, I-35122 Padova, Italy*
** E-mail: gianfranco.dezotti@inaf.it*

Matteo Bonato

*INAF-Istituto di Radioastronomia, and Italian ALMA Regional Centre,*
*Via Gobetti 101, I-40129, Bologna, Italy*

Mattia Negrello

*School of Physics and Astronomy, Cardiff University,*
*The Parade, Cardiff CF24 3AA, UK*

Tiziana Trombetti

*INAF-Istituto di Radioastronomia, Via Piero Gobetti 101, I-40129 Bologna, Italy*
*and Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Giuseppe Saragat 1, I-44122 Ferrara, Italy*
*and INFN, Sezione di Ferrara, Via Giuseppe Saragat 1, I-44122 Ferrara*

Carlo Burigana

*INAF-Istituto di Radioastronomia, Via Piero Gobetti 101, I-40129 Bologna, Italy*
*and Dipartimento di Fisica e Scienze della Terra, Università di Ferrara, Via Giuseppe Saragat 1, I-44122 Ferrara, Italy*
*and INFN, Sezione di Bologna, Via Irnerio 46, I-40126 Bologna*

The *Planck* mission has clearly demonstrated that, although not specifically designed for the observation of extragalactic sources, the space-borne experiments aimed at investigating the Cosmic Microwave Background (CMB) have the potential to bring breakthrough science also in this field. One example is the detection of high-$z$ galaxies with extreme gravitational amplifications. The combination of flux boosting and of stretching of the images has allowed the investigation of the structure of galaxies at $z \simeq 3$ with the astounding spatial resolution of $\simeq 60$ pc. Another example is the detection of proto-clusters of dusty galaxies at high $z$, when they may not yet possess the hot intergalactic medium allowing their detection in X-rays or via the Sunyaev-Zeldovich effect.

*Keywords*: Galaxies: high-redshift – galaxies: evolution – submillimetre: galaxies – galaxies: clusters: general.

## 1. Introduction

Extragalactic sources are one of the foregrounds that contaminate the maps produced by experiments aimed at mapping the Cosmic Microwave Background (CMB). On the other hand, space borne CMB experiments, like *Planck*, have properties that allow them to produce unique surveys of extragalactic sources:

- they cover a broad spectral range hardly, if at all, accessible from the ground and only lightly covered by other space missions;

- they offer an all-sky coverage, ideal to look for rare phenomena;
- they have a poor angular resolution that is, in general, a strong limitation for extragalactic surveys but, as we will see, becomes a resource to detect high-$z$ proto-clusters of dusty galaxies.

Two main populations of extragalactic sources show up in *Planck*'s frequency range, extending from 30 to 857 GHz: blazars, that dominate up to about 150–200 GHz, and active star-forming galaxies dominating at higher frequencies (Fig. 1).

*Planck* has provided unique information on these source populations. In this paper we will confine ourselves to only two examples: strongly gravitationally lensed high-$z$ galaxies (Sect. 2) and proto-clusters of dusty galaxies (Sect. 3). In Sect. 4 we summarize our main conclusions.

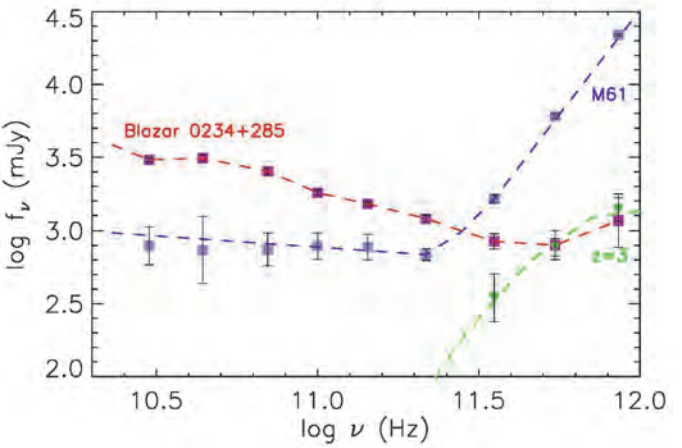



Fig. 1. Examples of spectral energy distributions (SEDs) of the extragalactic sources detected by *Planck*: a blazar, a nearby star-forming galaxy (M61) and a strongly lensed galaxy at $z \simeq 3$. Data points are *Planck* measurements.

## 2. Strongly lensed galaxies

*Herschel* surveys have demonstrated that bright sub-mm galaxies, such as those detected by *Planck*, have a strikingly bimodal redshift distribution (Fig. 2). Unlensed galaxies are local (only a few are at $z \geq 0.06$), while the gravitationally lensed ones are at $z > 1$. This implies that the two sub-populations can be easily distinguished[2,3].

Another nice property of sub-mm selected strongly lensed galaxies is that they show up in a waveband different from that of the foreground lens: the former are bright at far-IR/sub-mm wavelengths but very faint in the optical because of dust obscuration; on the contrary, the latter are mostly passively evolving early-type galaxies, essentially invisible at far-IR/sub-mm wavelengths, but bright in the

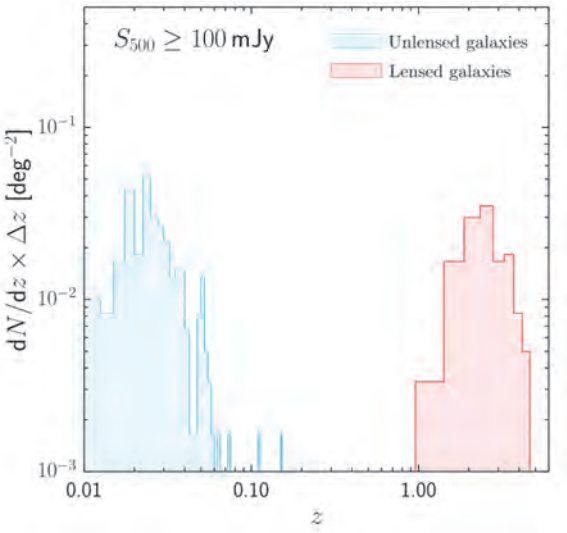



Fig. 2. Redshift distribution of dusty galaxies detected by the *Herschel*–ATLAS survey with $S_{500\mu m} > 100$ mJy (Ref. 1).

optical and in the near-IR. Hence there is no mutual contamination of the two images.

The 11 strongly lensed galaxies discovered on *Planck* maps[4] have redshifts in the range 2.2–3.6. For comparison, those discovered by optical searches have mostly $z < 1$ (cf. Fig. 7 of Ref. 5). Thus optical/near-IR and (sub-)mm surveys are nicely complementary: the former see evolved galaxies while the latter catch them in the most active star-formation phase.

This brings us to the main reason why the (sub-)mm selection of strongly lensed galaxies is important: high spatial and spectral resolution follow-up of these objects will have a crucial role in providing answers to major, still open issues on galaxy formation and evolution: which are the main physical mechanisms shaping the galaxy properties: in situ processes? interactions? mergers? cold flows from the intergalactic medium? How do feedback processes work? To settle these issues we need direct information on the structure and the dynamics of high-$z$ galaxies. But these are compact, with typical sizes of 1–2 kpc (e.g., Ref. 6), corresponding to angular sizes of 0.1–0.2 arcsec at $z \simeq 2$–3. Thus they are hardly resolved even by ALMA and by the HST. If they are resolved, high enough S/N ratios per resolution element are achieved only for the brightest galaxies, probably not representative of the general population.

Strong gravitational lensing provides a solution to these problems. The *Planck* dusty GEMS (Gravitationally Enhanced subMillimetre Sources)[4] have estimated magnifications, $\mu$, in the range 10–50, making them the brightest sub-mm sources in the sky. The gravitational stretching of the images has allowed ALMA observations

of one of them, PLCK_G244.8+54.9 at $z \simeq 3.0$ with $\mu \simeq 30$, to reach the astounding spatial resolution of $\simeq 60$ pc, substantially smaller than the size of Galactic giant molecular clouds[7].

Cañameras et al.[7] have also obtained CO spectroscopy, measuring the kinematics of the molecular gas with an uncertainty of 40–50 km/s. This spectral resolution makes possible a direct investigation of massive outflows driven by AGN feedback at high $z$. In this way Spilker et al.[8] were able to detect a fast (800 km/s) molecular outflow due to feedback in a strongly lensed galaxy at $z = 5.3$. The outflow carries mass at a rate close to the SFR and can thus remove a large fraction of the gas available for star-formation.

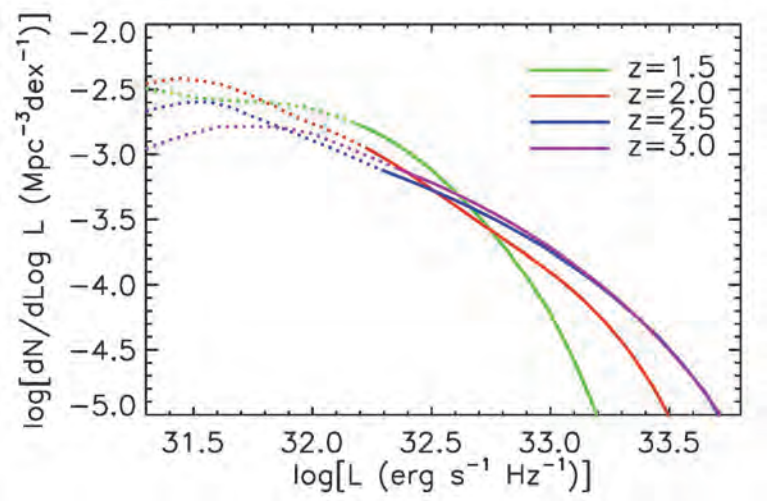



Fig. 3. Portions of the 500 $\mu$m galaxy luminosity function at various $z$ covered by *Herschel* surveys to a flux limit of 50mJy. The corresponding fraction of resolved luminosity density ranges from 29% to 66%.

## 3. Protoclusters

Classical techniques for detecting galaxy clusters (optical/near-IR "red sequence", X-ray emission, Sunyaev-Zeldovich (SZ) effect) preferentially or exclusively select evolved objects, with mature galaxy populations and a hot intra-cluster medium. As a result, most known clusters are at $z < 1.5$, i.e. below the peak of global star-formation activity.

This reflects the anti-correlation between density and specific star-formation rate (sSFR), well established in the low-$z$ universe. However the sSFR in clusters increases with increasing redshift faster than in the field matching field values at $z \gtrsim 1.2$[11]. Correspondingly the "red sequence" disappears and the hot intergalactic gas, making clusters visible in X-rays or via the SZ effect, is no longer necessarily in place.

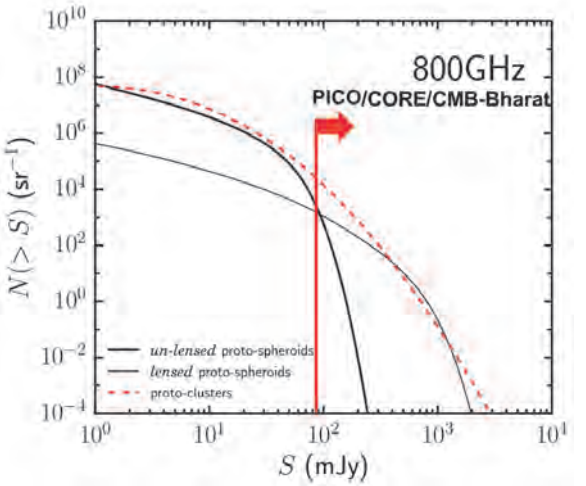



Fig. 4. Integral number ounts at 800 GHz of lensed and unlensed dusty galaxies predicted by the model by Ref. 9 and of proto-clusters[10]. The vertical red line shows the estimated detection limit of planned next generation CMB experiments.

On the other hand, member galaxies become increasingly bright at far-IR/sub-mm wavelengths. As mentioned above, individual high-$z$ galaxies are not detectable by space-borne CMB experiments unless their flux densities are boosted by gravitational lensing. However, the summed emission of galaxies within a strong over-density (proto-cluster) shows up as an intensity peak in the maps.

As illustrated by Fig. 3, proto-clusters stand out more clearly in low-resolution sub-mm maps, such as those that have been provided by *Planck* or will hopefully be provided by planned next-generation experiments like PICO, CORE and CMB Bharat, than in much higher resolution point source surveys at the same wavelengths because the available point source surveys miss a substantial fraction of the cluster. This is in keeping with the findings[12,13] that the sub-mm flux densities of proto-cluster candidates measured by *Planck* are about 2 to 3 times larger than the summed luminosities of member galaxies detected with *Herschel* within the *Planck* beam, although part of the difference is to be attributed to the 'flux boosting' affecting the low signal-to-noise *Planck* measurements.

Although *Planck* has demonstrated the power of low-resolution surveys for the study of large-scale structure, its resolution was too poor to detect individual proto-clusters[10]. Studies of the high-$z$ 2-point correlation function and *Herschel* images of the few sub-mm bright protoclusters detected so far, at $z$ of up to 4, indicate sizes of $\simeq 1'$ for the cluster cores, nicely matching the FWHM of the highest frequency channels of planned next generation experiments.

As illustrated by Fig. 4, such experiments will detect many tens of thousands of these objects as peaks in sub-mm maps, in addition to the evolved ones, detected by the SZ effect. This will constitute a real breakthrough in the observational

validation of the formation history of the most massive dark matter halos, traced by clusters, a crucial test of models for structure formation. Follow-up observations will characterize the properties of member galaxies, probing the galaxy evolution in dense environments and shedding light on the complex physical processes driving it.

## 4. Conclusions

Planck has demonstrated the unique capability of all-sky CMB experiments to explore astrophysical phenomena otherwise unaccessible to the present day instrumentation.

Those presented are only examples. Other examples are blazar astrophysics and high frequency polarization of extragalactic sources, including dusty galaxies.

Proposed next generation CMB experiments, like PICO, CORE and CMB Bharat, have the capability to make a giant leap forward in these fields by detecting several thousands high-$z$ strongly lensed galaxies at least up to $z \simeq 5$ and proto-clusters at least up to $z \simeq 4$.

## Acknowledgments

GDZ acknowledges support from ASI/INAF agreement n. 2014-024-R.1 for the *Planck* LFI Activity of Phase E2 and from the ASI/Physics Department of the university of Roma–Tor Vergata agreement n. 2016-24-H.0 for study activities of the Italian cosmology community.

## References

1. M. Negrello, S. Amber, A. Amvrosiadis, et al., *MNRAS* **465**, 3558 (2017).
2. M. Negrello, F. Perrotta, J. González-Nuevo, et al., *MNRAS* **377**, 1557 (2007).
3. M. Negrello, R. Hopwood, G. De Zotti, et al., *Science* **330**, 800 (2010).
4. R. Cañameras, N. P. H. Nesvadba, D. Guery, et al., *A&A* **581**, A105 (2015).
5. T. Treu, *Ann. Rev. Astr. Ap.* **48**, 87 (2010).
6. S. Fujimoto, M. Ouchi, K. Kohno, et al., *ApJ* **861**, 7 (2018).
7. R. Cañameras, N. Nesvadba, R. Kneissl, et al., *A&A* **604**, A117 (2017).
8. J. S. Spilker, M. Aravena, M. Béthermin, et al., *Science* **361**, 1016 (2018).
9. Z.-Y. Cai, A. Lapi, J.-Q. Xia, et al., *ApJ* **768**, 21 (2013).
10. M. Negrello, J. Gonzalez-Nuevo, G. De Zotti, et al., *MNRAS* **470**, 2253 (2017).
11. C. R. Wagner, S. Courteau, M. Brodwin, et al., *ApJ* **834**, 53 (2017).
12. D. L. Clements, F. G. Braglia, A. K. Hyde, et al., *MNRAS* **439**, 1193 (2014).
13. Planck Collaboration Int. XXXIX, *A&A* **596**, A100 (2016).

# Cosmological perturbations and gravitational instability of the Bianchi I model with a magnetic field

Federico Di Gioia

*Physics Department, "La Sapienza" University of Rome*
*P.le A. Moro 5, 00185 Roma, Italy*
*and*
*INFN, Roma 1 Section*
*P.le A. Moro 5, 00185 Roma, Italy*
*E-mail: federico.digioia@uniroma1.it*

Giovanni Montani

*ENEA*
*C.R. Frascati (Rome), Italy*
*Via E. Fermi 45, 00044 Frascati (Roma), Italy*
*and*
*Physics Department, "La Sapienza" University of Rome*
*P.le A. Moro 5, 00185 Roma, Italy*
*E-mail: giovanni.montani@enea.it*

We study the effects of a uniform magnetic field on the evolution of cosmological perturbations, in a full relativistic framework, considering the effects of anisotropies. We take advantage of the synchronous gauge to simplify the equations. We study both super-horizon scales in radiation dominated universe and sub-horizon scales in both a full relativistic framework and the Newtonian limit in case of adiabatic sound speed.

*Keywords*: Cosmology; Cosmological perturbations; Magnetic fields.

## 1. Introduction

There is strong evidence of the presence and importance of magnetic fields in cosmology[1] and it is possible that such fields, although small,[2–4] are related to the anisotropy of large scale structures (for a detailed introduction, see[5]). There are a lot of studies in literature about their effects in both Newtonian limit, considering the full anisotropy[5], and in full relativistic treatment, but inside a FRW model and discarding anisotropic effects.[6,7] Our aim is to incorporate anisotropic effects inside the full relativistic treatment, while using the synchronous gauge to manage the increased mathematical complexity.

In Sec. 2 we develop some corrections to a former analysis of the background model,[3] then in Sec. 3 we obtain the perturbed equations. We discuss the problem of gauge freedom in Sec. 4 and we analyse the solutions in physical interesting regimes in Sec. 5.

## 2. Background model

We study a uniform Bianchi I model with a perfect barotropic fluid and a magnetic field aligned in the $z$ direction. Our metric is

$$g_{\mu\nu} = \mathrm{diag}\left(-1, a^2(t), a^2(t), c^2(t)\right). \tag{1}$$

The background fluid has velocity $u^\mu = (1, 0, 0, 0)$, energy density $\rho$ and pressure $p = w\rho$, with $w = \text{const}$. The magnetic field is $B^\mu = (0, 0, 0, B^3)$, with energy density $B^2 = B_\mu B^\mu = c^2 B^3 B^3$ in the Lorentz–Heaviside units.

We follow exactly[3] to solve this system, however we find slightly different solutions due to different assumptions and we add isotropic corrections, while that paper only cares about anisotropic ones.

Defining the isotropic expansion $H$ and the anisotropy parameter $S$ as

$$3H = 2\frac{\dot{a}}{a} + \frac{\dot{c}}{c}, \qquad HS = \frac{\dot{a}}{a} - \frac{\dot{c}}{c} \tag{2}$$

and using the Alfvn speed

$$v_A^2 = \frac{B^2/2}{\rho} \tag{3}$$

to describe the magnetic field we write the evolution equations for the system. The Einstein equation with components 00 reads

$$3\dot{H} + H^2\left(3 + \frac{2}{3}S^2\right) = -\left[\frac{1}{2}(1 + 3w) + v_A^2\right]\rho, \tag{4}$$

subtracting the Einstein 33 equation from the 11 equation we get

$$H\dot{S} + \dot{H}S + 3H^2 S = 2v_a^2\rho \tag{5}$$

and contracting the Einstein equations over the spatial indices

$$3\dot{H} + 9H^2 = \left[\frac{3}{2}(1 - w) + v_A^2\right]\rho. \tag{6}$$

The fluid energy conservation and the magnetic field energy conservation read

$$\dot{\rho} + 3H(1 + w)\rho = 0, \qquad (\dot{B^2}) + 4\frac{\dot{a}}{a}B^2 = 0; \tag{7}$$

the last one gives

$$(\dot{v_A^2}) = v_A^2 H\left(3w - 1 - \frac{4}{3}S\right). \tag{8}$$

Considering the magnetic field small and solving the equations perturbatively, at 0-order one simply recovers FRW

$$H_{(0)} = \frac{2}{3(1 + w)t}, \qquad \rho_{(0)} = 3H_{(0)}^2, \qquad S_{(0)} = 0. \tag{9}$$

Substituting this solution inside the equations we solve in order Eq. (8) that gives the temporal law of $v_A^2$, Eq. (5) that describes the evolution of $S$, Eq. (4) and Eq. (6) that describe together the isotropic corrections to $H$ and $\rho$. We find the following solutions.

### 2.1. *Radiation dominated universe*

The background fluid is ultra-relativistic with $w = 1/3$ and the solution is

$$v_A^2 = v_{A0}^2 = \text{const}, \qquad t_0 = \text{const}, \qquad S = 6v_A^2 = 6v_{A0}^2 \tag{10}$$

$$a \propto \left(\frac{t}{t_0}\right)^{1/2} \left(1 + v_{A0}^2 \ln\left(\frac{t}{t_0}\right)\right), \qquad c \propto \left(\frac{t}{t_0}\right)^{1/2} \left(1 - 2v_{A0}^2 \ln\left(\frac{t}{t_0}\right)\right) \tag{11}$$

$$H = \frac{1}{2t}, \qquad \rho = \frac{3}{4t^2}(1 - v_{A0}^2). \tag{12}$$

### 2.2. *Matter dominated universe*

The background fluid is non-relativistic with $w = 0$ and the solution is

$$v_A^2 = v_{A0}^2 \left(\frac{t}{t_0}\right)^{-2/3}, \qquad v_{A0}^2, t_0 = \text{const}, \qquad S(t) = 12v_A^2(t) \tag{13}$$

$$a \propto \left(\frac{t}{t_0}\right)^{2/3} - 3v_{A0}^2, \qquad c \propto \left(\frac{t}{t_0}\right)^{2/3} + 9v_{A0}^2 \tag{14}$$

$$H = \frac{2}{3t}(1 - v_A^2(t)), \qquad \rho = \frac{4}{3t^2}\left(1 - 3v_A^2(t)\right). \tag{15}$$

## 3. Perturbed model

We perturb the background model allowing small inhomogeneities. We take the starting equations from[6], but instead of solving them in full covariant formalism we focus on the synchronous gauge; this choice will make their solution a lot easier. Enforcing the synchronous gauge and the spatiality of the magnetic field, represented by $B_\mu u^\mu = 0$, the perturbed variables are

$$g_{\mu\nu} = g_{\mu\nu}^{\text{back}} + \delta g_{\mu\nu}, \qquad \gamma_{\mu\nu} = \delta g_{\mu\nu}, \qquad \gamma_{0\nu} = \gamma_{\mu 0} = 0 \tag{16a}$$

$$g_{\mu\rho} g^{\rho\nu} = \delta_\mu^\nu \implies \delta g^{\mu\nu} = -\gamma^{\mu\nu} \tag{16b}$$

$$u^\mu = u^\mu_{\text{back}} + \delta u^\mu, \qquad u_\mu u^\mu = -1 \implies \delta u^0 = 0 \tag{16c}$$

$$\rho = \rho^{\text{back}} + \delta\rho, \qquad p = p^{\text{back}} + \delta p, \qquad \delta p = v_S^2 \delta\rho \tag{16d}$$

$$B^2 = B^2_{\text{back}} + \delta(B^2), \qquad \delta\left(B_\mu B^\mu\right) = \delta\left(B^2\right) = \gamma_{33} B^3 B^3 + 2c^2 B^3 \delta B^3 \tag{16e}$$

$$B_\mu u^\mu = 0 \implies \delta B^0 = c^2 B^3 \delta u^3, \tag{16f}$$

where "back" marks background variables and the indices of $\gamma_{\mu\nu}$ are raised and lowered with the unperturbed metric. For simplicity, we define the following perturbation variables:

$$\Delta = \frac{\delta\rho}{(1+w)\rho}, \qquad G = \frac{1}{2}\gamma_i^i, \qquad T = \frac{1}{2}\gamma_3^3, \qquad M = \frac{\delta\left(B^2\right)}{B^2}. \tag{17}$$

We will now write the relevant equations in synchronous gauge. The full covariant equations can be found in[6]; they form a very complicated system, so we will write them and solve them in the limit of small magnetic fields: we will make a

Fourier expansion for the spatial components $e^{ik_j x^j}$ and keep only terms up to first order in $v_A^2$. The sound speed is related to the barotropic index $w$ through

$$\dot{w} = -3H(1+w)(v_S^2 - w) \implies v_S^2 = w \text{ if } w = \text{const}, \tag{18}$$

but we will distinguish them if $v_S^2$ is multiplied by $k_i k^i$ or $k_3 k^3$ and we will simplify them in all other circumstances in which they cancel each other; we will explain this choice in the following.

Through some algebraic manipulations, the perturbed equations take the following forms (we refer to[8] for the full calculations): the fluid energy conservation is

$$\dot{G} = -\dot{\Delta} - \partial_i \delta u^i; \tag{19}$$

the magnetic energy conservation is

$$\dot{M} = -2\left(\partial_i \delta u^i - \partial_3 \delta u^3 + \dot{G} - \dot{T}\right) = 2\left(\dot{\Delta} + \dot{T} + \partial_3 \delta u^3\right); \tag{20}$$

the Einstein 00 equation, through the use of Eq. (19), reads

$$\begin{aligned} &\ddot{\Delta} + 2H\left(1 + \frac{1}{3}S\right)\dot{\Delta} - \frac{1}{2}(1+3v_S^2)(1+w)\rho\Delta + \partial_0\partial_i\delta u^i \\ &\quad + 2H\left(1+\frac{1}{3}S\right)\partial_i\delta u^i + \frac{4}{3(1+w)}S\frac{\dot{T}}{t} - \frac{4}{3(1+w)^2}v_A^2\frac{M}{t^2} = 0. \end{aligned} \tag{21}$$

from the Einstein 33 and 03 equations we get

$$\begin{aligned} &3(1+w)(k_i k^i \dot{T} - k_3 k^3 \dot{G}) + 3(1+w)\dddot{T} + 10\frac{\ddot{T}}{t} + 2\frac{1-3w}{1+w}\frac{\dot{T}}{t^2} + \frac{2}{3}\frac{1-3w}{1+w}\frac{\dot{G}}{t^2} \\ &\quad + 2\frac{\ddot{G}}{t} - 2\left[1 - v_S^2 - (\dot{v_S^2})\right]\frac{\dot{\Delta}}{t^2} + \frac{4}{3}(1-v_S^2)\frac{1+3w}{1+w}\frac{\Delta}{t^3} - 8\frac{\partial_3\delta u^3}{t^2} = 0; \end{aligned} \tag{22}$$

the divergence of the momentum conservation is

$$\begin{aligned} &\left(1 + \frac{2}{1+w}v_A^2\right)\partial_0\partial_i\delta u^i + \left[(2-3w)H + \left(\frac{v_A^2}{1+w} + \frac{1}{3}S\right)\frac{4}{3(1+w)}\frac{1}{t}\right]\partial_i\delta u^i \\ &\quad = -v_S^2\partial_i\partial^i\Delta - \frac{v_A^2}{1+w}\partial_i\partial^i M \\ &\qquad + \frac{2}{1+w}v_A^2\partial_0\partial_3\delta u^3 - 2\left(\frac{v_A^2}{1+w} + S\right)\frac{2}{3(1+w)}\frac{\partial_3\delta u^3}{t} \\ &\qquad - \frac{2}{1+w}v_A^2\left[\ddot{T} + \frac{2}{1+w}\frac{\dot{T}}{t} + \frac{2}{3(1+w)}\frac{\dot{G}}{t}\right] + \frac{4}{3(1+w)}(1-v_S^2)v_A^2\frac{\Delta}{t^2}; \end{aligned} \tag{23}$$

the momentum conservation along the background magnetic field at 0-order in $v_A^2$, $v_S^2$, but keeping the Laplacian gives

$$\partial_0\partial_3\delta u^3 + (2-3w)H\partial_3\delta u^3 + \partial_3\partial^3(v_S^2\Delta) = 0. \tag{24}$$

## 4. Gauge modes

The synchronous gauge does not end the freedom of choice of the reference frame: the residual sub-gauge freedom is itself a solution of the perturbed equations, so we must find and discard it. Following a standard procedure, shown for example in[9], one can calculate the exact residual gauge freedom and show that the resulting gauge mode for the energy density is[8]

$$\delta\rho_{\rm gauge} \propto \dot{\rho}^{\rm back} \propto H(\rho^{\rm back} + p^{\rm back}) \propto H\rho^{\rm back}. \tag{25}$$

We must find this solution solving the perturbed equations, and we must discard it because it is a pure mathematical effect and it is unphysical.

## 5. Solutions

### 5.1. *Large scale, radiation dominated universe*

We study two kind of solutions. At large scales and during radiation dominated universe we have $k_i \approx 0$, $w = v_S^2 = 1/3$ and the density perturbations follow a power law solution:

$$\Delta = \frac{\Delta_{\rm gauge}}{t} + \Delta_{\rm phys} t^{1-2v_{A0}^2}; \tag{26}$$

the decaying solution is the gauge one, while the growing one is the physical solution. This is slightly different from the solution of[6].

### 5.2. *Small scales, matter dominated universe*

We will now ignore Eq. (18) and use an adiabatic sound speed, with adiabatic index $\gamma$, as in[5,10]:

$$v_S^2 \propto t^{2(1-\gamma)}; \tag{27}$$

after recombination and before baryon decoupling (for perturbations, the decoupling happens effectively at $z \approx 100$, see[5]) we have $\gamma = 4/3$, while after baryon decoupling $\gamma = 5/3$.[5,10] This is valid in the Newtonian approximation,[5,10] where we do not have Eq. (18), but assuming $v_S^2 = 0$ we recover the full relativistic case of[6]. We define the constants

$$\Lambda_S^2 = v_S^2 k_i k^i t^{2\gamma-2/3}, \qquad \Lambda_A^2 = v_A^2 k_i k^i t^2 \tag{28}$$

ad we solve the system at 0-order in $v_S^2$, $v_A^2$, considering $k_i k^i \gg H^2$ and so keeping $k_i k^i v_S^2$, $k_i k^i v_A^2 \gg v_S^2$, $v_A^2$. Defining an anisotropy parameter $\mu$ through $k_3 k^3 = \mu^2 k_i k^i$, with some algebra we find the same final equation of[5]: this solutions is different from[6] and shows that anisotropies induced by the magnetic field are important and must be considered with more care. Considering the correct relativistic treatment, with $v_S^2 = 0$, this solves the open problem of the former study of Bianchi I model of[11]. It can be shown[5] that for $\mu \neq 0$ it holds the usual Jeans

threshold $k_J$, while for $\mu = 0$ there is an additional critical value determined by the magnetic field:

$$k_J = \sqrt{\frac{24\pi G(\gamma - 4/3)^2 \rho^{\rm back}}{v_S^2}}, \qquad k_A = \sqrt{\frac{12\pi^2 G\rho^{\rm back}}{B^2_{\rm back}}}, \tag{29}$$

where $G$ is Newton's constant. For a more accurate analysis of the results see[8].

## 6. Conclusions

We highlighted the importance of the anisotropy in studies regarding cosmological perturbations with magnetic fields and we showed that the equations describing cosmological perturbations take a simpler form when expressed in the synchronous gauge, once we are able to recognize and discard the unphysical gauge modes. We incorporated anisotropic effects in the full relativistic analysis and we found corrections to the usual solutions in isotropic FRW models due to the anisotropy, controlled by the background magnetic field.

## References


1. M. Giovannini, The magnetized universe, *International Journal of Modern Physics D* **13**, 391 (2004).
2. J. D. Barrow, P. G. Ferreira and J. Silk, Constraints on a primordial magnetic field, *Phys. Rev. Lett.* **78**, 3610 (1997).
3. J. D. Barrow, Cosmological limits on slightly skew stresses, *Phys. Rev. D* **55**, 7451 (1997).
4. Planck Collaboration, P. A. R. Ade, N. Aghanim, M. Arnaud, F. Arroja, M. Ashdown, J. Aumont, C. Baccigalupi, M. Ballardini, A. J. Banday and et al., Planck 2015 results. xix. constraints on primordial magnetic fields, *ArXiv e-prints* (2015).
5. M. Lattanzi, N. Carlevaro and G. Montani, Gravitational instability of the primordial plasma: Anisotropic evolution of structure seeds, *Phys. Lett. B* **718**, 255 (2012).
6. J. D. Barrow, R. Maartens and C. G. Tsagas, Cosmology with inhomogeneous magnetic fields, *Physics Reports* **449**, 131 (2007).
7. C. G. Tsagas, A. Challinor and R. Maartens, Relativistic cosmology and large-scale structure, *Physics Reports* **465**, 61 (2008).
8. F. Di Gioia and G. Montani, Linear perturbations of an anisotropic bianchi i model with a uniform magnetic field, In preparation, arXiv 1807.00434.
9. G. Montani, M. V. Battisti, R. Benini and G. Imponente, *Primordial Cosmology* (World Scientific, 2011).
10. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley & Sons, 1972).
11. C. G. Tsagas and R. Maartens, Cosmological perturbations on a magnetized bianchi i background, *Classical and Quantum Gravity* **17**, 2215 (2000).

# Neutrino properties and the cosmological tensions in the ΛCDM model

Stefano Gariazzo

*Instituto de Física Corpuscular (CSIC-Universitat de València), Paterna (Valencia), Spain*
*E-mail: gariazzo@ific.uv.es*

In this talk I will review the current status of the constraints on the neutrino properties from cosmological measurements, with a particular focus on their mass and effective number. I will also discuss the existing tensions within the context of the ΛCDM model, including the discrepancies on the Hubble parameter and on the matter fluctuations at small scales, and how neutrinos could help to alleviate the aforementioned problems.

*Keywords*: Neutrinos; Neutrinos in Cosmology.

One of the most amazing experimental results of the recent years has been achieved by the Planck collaboration, which measured the Cosmic Microwave Background (CMB) radiation with an unprecedented precision.[1–4] Such result allowed us to improve our knowledge of the Universe history, which can be very well described by the six-parameters model named ΛCDM, after the cosmological constant (Λ) and the Cold Dark Matter (CDM) fluids, which represent most of the current energy density in the Universe.

Within the context of the ΛCDM model and using CMB data, we can derive constraints on a number of cosmological quantities, as for example the Hubble parameter $H_0$ or the matter perturbations at small scales, usually quantified through the parameter $\sigma_8$, which describes the mean matter fluctuations in a sphere with a radious of $8h^{-1}$ Mpc. The values of $H_0$ and $\sigma_8$ obtained from a fit of the CMB data within the context of the ΛCDM model are not in agreement with the values that are measured in the local universe, at small redshifts. For instance, while Planck data[2] point towards $H_0 = 67.27 \pm 0.66$ km s$^{-1}$Mpc$^{-1}$, estimates from the local Universe[5] indicate $H_0 = 73.24 \pm 1.74$ km s$^{-1}$Mpc$^{-1}$: a $\sim 3.4\sigma$ tension exists. In the same way, cosmological and local determinations of $\sigma_8$ typically exhibit a $\sim 2\sigma$ tension with local determinations, as performed for example by the KiDS[6] or DES[7] experiments. These tensions may be the result of an incomplete understanding of systematics in the experimental determinations or of the presence of new physics which is not considered in the ΛCDM model.

Under a different point of view, cosmology is nowadays one of the most powerful tools to study particular aspects of particle physics, such as some neutrino properties, and in particular their absolute masses or the way the interact and thermalize in the early universe. In the standard cosmological analyses, neutrinos are usually treated in the minimal way, assuming minimal masses ($\Sigma m_\nu = 0.06$ eV, corresponding to a massless lightest neutrino and two massive neutrinos, for which the masses come from the mass splittings measured in neutrino oscillation experiments and assuming normal ordering, see e.g. Ref. 8) and a contribution to the effective number of relativistic species in the early Universe $N_{\rm eff} = 3.05$.[9] Non-standard

neutrino properties may alter the cosmological evolution and have an effect on the CMB spectrum, as well as the formation of structures at later times, so that both parameters ($\Sigma m_\nu$ and $N_{\rm eff}$) can be well constrained.

First of all, neutrinos decoupled in the early universe when they were still relativistic and contribute to $N_{\rm eff}$. Variations of $N_{\rm eff}$ influence the universe evolution before the photon decoupling, and as a consequence the CMB spectrum through a variation of the amplitude of the peaks and their angular scale. In order to compensate the effects of having $N_{\rm eff}$ different from 3.05, as it would be in non-standard cases (modified neutrino properties or presence of new neutrino-like particles), one can alter the energy densities of CDM and of dark energy. In this case, the angular scale of the CMB peaks can be left unvaried, at the expense of increasing $H_0$ and having a damping of the CMB spectrum at small scales.

After the non-relativistic transition, which occurs at different times for each neutrino eigenstate depending on its mass, neutrinos contribute as regular massive particles and must be considered when computing the energy density of matter. Their influence on the CMB spectrum is related to the modification of the late-time expansion, both through the variations of the amplitude of the first peak, related to the early integrated Sachs-Wolf (ISW) effect, and of the angular position of the CMB peaks, which comes from the modifications of the angular diameter distance in presence of massive neutrinos. This latter effect can be partially compensated with a rescaling of the Hubble parameter $H_0$, so that an anti-correlation between $H_0$ and $\Sigma m_\nu$ appears.

Massive neutrinos, however, have a proportionally more important impact on the evolution of large scale structures than on the CMB spectrum. Neutrinos, indeed, possess a large thermal velocity which allows them to escape the gravitational attraction at very small scales. For this reason, no neutrino overdensity can grow at scales smaller than a characteristic scale, named the free-streaming scale, which depends on the neutrino mass. When one compares two universes which share the same total amount of matter, one with massless and the other with massive neutrinos, the universe with massive neutrinos will therefore have a suppressed matter power spectrum at the smallest scales, because the free-streaming impedes the growth of neutrino perturbations and meanwhile reduces the possibility for the gravitational attraction to accrete the total matter overdensities. The fact that for massive neutrinos there is a reduction of the matter fluctuations at small scales means that an increase of the neutrino masses correlates with a decrease of $\sigma_8$.

Given these main effects that neutrinos produce on the cosmological quantities and the resulting degeneracies between the neutrino properties ($\Sigma m_\nu$ and $N_{\rm eff}$), $H_0$ and $\sigma_8$, one may be tempted to say that neutrinos can reduce or even solve the tensions that appear within the context of the ΛCDM model between the CMB preferred cosmology and local universe probes. Unfortunately, this does not happen. The analyses of CMB data, indeed, show that it is difficult to alleviate both the tensions at the same time. If one considers CMB measurements from Planck 2015,[2]

the best fit values for $N_{\rm eff}$ and $\Sigma m_\nu$ lie very close to the expected values: $N_{\rm eff} = 3.0 \pm 0.2$ (68% CL) and $\Sigma m_\nu < 0.17$ eV (95% CL). Naively, in order to allieviate the $H_0$ tension, an higher $N_{\rm eff}$ would be useful. In the same way, heavier neutrinos would allow to reduce the matter fluctuations at small scales, $\sigma_8$. The problem is that an increase in $N_{\rm eff}$ requires to be compensated by a larger matter density at all times, so that at the practical level also $\sigma_8$ increases. In the same way, heavier neutrinos require a smaller $H_0$: it is therefore easy to see that the two tensions cannot be solved at the same time solely varying neutrino properties.

One last word on Planck measurements in connection with neutrino masses. The Planck 2015 estimates of CMB lensing from the distortions of the TT spectrum[2] are in slight tension with the reconstruction of the lensing profile from the 4-point correlation function,[10] as quantified by the phenomenological lensing amplitude parameter, $A_L = 1.22 \pm 0.10$ (68% CL), which is expected to be 1 in general relativity. Massive neutrinos, again thanks to their free-streaming, can allieviate the tension. This is the reason for which the 95% CL limits change from $\Sigma m_\nu < 0.17$ eV (Planck TT,TE,EE + lowP + BAO) to $\Sigma m_\nu < 0.22$ eV (Planck TT,TE,EE + lowP + BAO + lensing) when the lensing information is considered in the analyses.

Let us now discuss non-standard neutrino degrees of freedom. In the very recent years, for the first time we have had model-independent indications that neutrino oscillations at a baseline that is not compatible with the standard three-neutrino oscillations may exist: the NEOS[11] and DANSS[12] reactor experiments observed electron antineutrino disappearance at distances of $\sim 24$ m and $\sim 10 - 12$ m, respectively. Combined model-independent analyses[13,14] show that the combination of the two experiments give a $\sim 3.5\sigma$ preference for the three active plus one light sterile (3+1) neutrino scenario over the standard three neutrino case, implying the existence of a fourth neutrino, with a mass around 1.1 eV. When considering the global picture, one must also take into account data from muon neutrino disappearance (e.g. from MINOS+[15] or IceCube[16,17]) or electron neutrino appearance in a muon neutrino beam (e.g. from LSND[18] or MiniBooNE[19]). The first problem is that there is a tension between appearance (lead by LSND and MiniBooNE) and disappearance data (lead by NEOS, DANSS and MINOS+), see e.g. Ref. 20, so that the situation is still unclear. The second problem is that if the 3+1 neutrino oscillation parameters as obtained from the global fit are considered to compute the neutrino decoupling in the early universe, one would have a fully thermalized fourth neutrino, with $N_{\rm eff} \simeq 4$,[21,22] which is in clear tension with the constraints on $N_{\rm eff}$. This means that if the light sterile neutrino will be confirmed, some additional mechanism will be needed to suppress its thermalization in the early universe. As in the case of active neutrinos, the additional sterile neutrino and its properties can only marginally reduce the tension between the local universe probes of $H_0$ and $\sigma_8$ and the corresponding values inferred by CMB observations in the context of an extended $\Lambda$CDM model.

As a summary, our current knowledge of the universe and its history shows a number of mild tensions, in particular regarding the Hubble parameter $H_0$ and the matter fluctuations at small scales, encoded by $\sigma_8$. A mild tension is also present among the two different determinations of the CMB lensing by Planck.[2] Neutrino properties (considering both active or sterile states) can alleviate each of these tensions if considered separately, but in the global picture the situation cannot be easily solved. If these cosmological tensions or the existence of a light sterile neutrino will be confirmed in the future, some new physical mechanism will be needed in order to obtain a consistent model that can explain the entire evolution of the universe.

## Acknowledgments

Work supported by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie individual grant agreement No. 796941, and by the Spanish grants SEV-2014-0398 and FPA2017-85216-P (AEI/FEDER, UE), and PROMETEOII/2018/165 (Generalitat Valenciana).

## References

1. R. Adam *et al.*, Planck 2015 results. I. Overview of products and scientific results, *Astron. Astrophys.* **594**, p. A1 (2016).
2. P. A. R. Ade *et al.*, Planck 2015 results. XIII. Cosmological parameters, *Astron. Astrophys.* **594**, p. A13 (2016).
3. Y. Akrami *et al.*, Planck 2018 results. I. Overview and the cosmological legacy of Planck (2018).
4. N. Aghanim *et al.*, Planck 2018 results. VI. Cosmological parameters (2018).
5. A. G. Riess *et al.*, A 2.4% Determination of the Local Value of the Hubble Constant, *Astrophys. J.* **826**, p. 56 (2016).
6. S. Joudaki *et al.*, KiDS-450: Testing extensions to the standard cosmological model, *Mon. Not. Roy. Astron. Soc.* **471**, p. 1259 (2017).
7. T. M. C. Abbott *et al.*, Dark Energy Survey Year 1 Results: Cosmological Constraints from Galaxy Clustering and Weak Lensing (2017).
8. P. de Salas, D. Forero, C. Ternes, M. Tórtola and J. Valle, Status of neutrino oscillations 2018: 3 $\sigma$ hint for normal mass ordering and improved CP sensitivity, *Phys. Lett.* **B782**, 633 (Jul 2018).
9. P. F. de Salas and S. Pastor, Relic neutrino decoupling with flavour oscillations revisited, *JCAP* **07**, p. 051 (2016).
10. P. A. R. Ade *et al.*, Planck 2015 results. XV. Gravitational lensing, *Astron. Astrophys.* **594**, p. A15 (2016).
11. Y. J. Ko *et al.*, Sterile Neutrino Search at the NEOS Experiment, *Phys. Rev. Lett.* **118**, p. 121802 (2017).
12. I. Alekseev *et al.*, Search for sterile neutrinos at the DANSS experiment (2018).

13. S. Gariazzo, C. Giunti, M. Laveder and Y. Li, Model-Independent $\bar{\nu}_e$ Short-Baseline Oscillations from Reactor Spectral Ratios, *PLB* **782**, 13 (2018).
14. M. Dentler, A. Hernández-Cabezudo, J. Kopp, M. Maltoni and T. Schwetz, Sterile Neutrinos or Flux Uncertainties? - Status of the Reactor Anti-Neutrino Anomaly, *JHEP* **11**, p. 099 (2017).
15. P. Adamson *et al.*, Search for sterile neutrinos in MINOS and MINOS+ using a two-detector fit, *Phys. Rev. Lett.* (2017).
16. M. G. Aartsen *et al.*, Search for sterile neutrino mixing using three years of IceCube DeepCore data, *Phys. Rev. D* **95**, p. 112002 (2017).
17. M. G. Aartsen *et al.*, Searches for Sterile Neutrinos with the IceCube Detector, *Phys. Rev. Lett.* **117**, p. 071801 (2016).
18. A. Aguilar-Arevalo *et al.*, Evidence for neutrino oscillations from the observation of anti-neutrino(electron) appearance in a anti-neutrino(muon) beam, *Phys. Rev. D* **64**, p. 112007 (2001).
19. A. Aguilar-Arevalo *et al.*, Observation of a Significant Excess of Electron-Like Events in the MiniBooNE Short-Baseline Neutrino Experiment.
20. M. Dentler, Á. Hernández-Cabezudo, J. Kopp, P. A. N. Machado, M. Maltoni, I. Martinez-Soler and T. Schwetz, Updated global analysis of neutrino oscillations in the presence of eV-scale sterile neutrinos (2018).
21. S. Hannestad, I. Tamborra and T. Tram, Thermalisation of light sterile neutrinos in the early universe, *JCAP* **07**, p. 025 (2012).
22. A. Mirizzi, N. Saviano, G. Miele and P. D. Serpico, Light sterile neutrino production in the early universe with dynamical neutrino asymmetries, *Phys. Rev. D* **86**, p. 053009 (2012).

# $H_0$ from cosmic chronometers and Type Ia supernovae, with Gaussian processes and the weighted polynomial regression method

Adrià Gómez-Valent*

*Departament de Física Quàntica i Astrofísica, and Institute of Cosmos Sciences, Universitat de Barcelona, Av. Diagonal 647, E-08028 Barcelona, Catalonia, Spain*
*adriagova@fqa.ub.edu*

Luca Amendola

*Institut für Theoretische Physik, Ruprecht-Karls-Universität Heidelberg Philosophenweg 16, 69120 Heidelberg, Germany*
*l.amendola@thphys.uni-heidelberg.de*

We derive new constraints on the Hubble parameter $H_0$ using the available data on $H(z)$ from cosmic chronometers (CCH), and the Hubble rate data points from the supernovae of Type Ia (SnIa) of the Pantheon compilation and the Hubble Space Telescope (HST) CANDELS and CLASH Multy-Cycle Treasury (MCT) programs. We employ two alternative techniques, Gaussian Processes (GPs) and the Weighted Polynomial Regression (WPR) method, to reconstruct the Hubble function, determine the derived values of $H_0$, and compare them with the local HST measurement provided by Riess *et al.* (2018), $H_0^{\rm HST} = (73.48 \pm 1.66)$ km/s/Mpc, and with the Planck+ΛCDM value, $H_0^{\rm P18} = (66.88 \pm 0.92)$ km/s/Mpc. With GPs we obtain $H_0 = (67.99 \pm 1.94)$ km/s/Mpc, and with the WPR method $H_0 = (68.90 \pm 1.96)$ km/s/Mpc. Both are fully compatible at $< 1\sigma$ c.l., and also with $H_0^{\rm P18}$. In contrast, they are in $\sim 2\sigma$ tension with $H_0^{\rm HST}$.

*Keywords*: Dark energy experiments; supernova type Ia - standard candles; cosmological parameters.

## 1. Introduction

The Hubble-Lemaître constant, $H_0$, is a crucial parameter in Cosmology. It does not only tell us about the current velocity of recession of distant astronomical objects due to the cosmological dynamics, but it is also essential to compute cosmic distances and spans of time. Now, ninety years after the discovery of the Universe's expansion and first measurement of $H_0$ by Edwin Hubble[1], and after the many works in the literature since then aiming to improve the estimations of this parameter, we can state with high statistical confidence level that $H_0$ lies in the range of $65 - 75$ km/s/Mpc. Nevertheless, this parameter is still a matter of many discussions between the members of the cosmological community, mainly focused on the origin of the huge tension between the local (distance ladder) determination of $H_0$ obtained with the Hubble Space Telescope (HST)[2] and the value inferred from the temperature and polarization anisotropies of the cosmic microwave background (CMB) radiation measured by the Planck satellite[3], by assuming the ΛCDM theoretical framework. The former reads $H_0^{\rm HST} = (73.48 \pm 1.66)$ km/s/Mpc, whereas the TT+lowE CMB Planck+ΛCDM analysis leads to $H_0^{\rm P18} = (66.88 \pm 0.92)$ km/s/Mpc.

*Speaker

These values are in $\sim 3.5\sigma$ tension. Finding the ultimate reason that generates this mismatch has become one of the most important issues in Cosmology. In principle, unaccounted systematics affecting the HST and/or Planck measurements could play a role here. Regarding this point, it is important to mention that many other works making use of the cosmic distance ladder find values of $H_0$ also lying in the high region preferred by Ref. 2, see e.g. Refs. 4, 5, 6. There are also discrepant voices in the literature, though. For instance, the author of Ref. 7 argues that the use by Riess et al. of the 2M++ density field map (which covers redshifts $z \leq 0.06$) to compute peculiar velocity flows could be biasing their results, since there is evidence of the existence of local radial inhomogeneities extending in different directions up to a redshift of about 0.07[8], and according to Ref. 7 the 40% of the Cepheids used in Ref. 2 would be affected. Moreover, the authors of Ref. 9 claim that the GAIA parallax distances of Milky Way Cepheids employed in Ref. 2 in the first step of the cosmic distance ladder may be underestimated a $\sim 7 - 18\%$. This would produce an important decrease of their measured value of $H_0$. The discussion on the validity of these arguments is, though, still open and intense.[10]

In regards to the potential systematics affecting Planck's data, it is worth to highlight the study of Ref. 11, in which the authors showed that the use of CMB data independent from Planck together with data from baryon acoustic oscillations (BAO) and the big bang nucleosynthesis (BBN) gives rise to values of $H_0$ fully compatible with the Planck preferred range and, hence, it seems quite improvable that systematics affect Planck's analysis in a decisive way.

Assuming the $\Lambda$CDM, the authors of Refs. 12, 13 analyzed the gravitational time delay of the light rays coming from multiply imaged quasar systems and found values of $H_0$ unable to discriminate between the two values in dispute due to their large error bars. It has also been possible to measure the Hubble parameter using the gravitational wave signal of the neutron star merger GW17081716 and its electromagnetic counterpart[14–16], but still with very big uncertainties too. This method is interesting because it is cosmology-independent and represents an alternative to the cosmic distance ladder measurement[2]. According to Ref. 17, a sample of $\sim 50$ binary neutron star standard sirens (detectable within the next decade) will be able to arbitrate between the local and CMB estimates. The impact of the cosmic variance on the local determination of $H_0$ has been also studied in Refs. 18, 19, 20. The authors conclude that its effect cannot explain the whole discrepancy between the HST and Planck's values.

In view of all these facts and the unsuccessful efforts of the community of finding a theoretical model able to relieve in an efficient way the $H_0$-tension (see e.g. Refs. 21, 22), it is important to go on with the investigations on the value of the Hubble parameter. Particularly relevant are those studies that try to extract values from observations at intermediate redshifts, as model-independent as possible, and also utilizing approaches different from the one applied in Ref. 2, where the authors made use of the cosmic distance ladder and low-redshift SnIa ($z < 0.15$). We present

here the main results of our dedicated work[23], in which we employed data on cosmic chronometers (CCH) and supernovae of Type Ia (SnIa), and two different methods – Gaussian Processes (GPs) and the novel Weighted Polynomial Regression (WPR) technique – to reconstruct the Hubble function and derive an extrapolated value of $H_0$.

## 2. Data

In this section we limit ourselves to cite the references from which we have collected the data on CCH and SnIa, and confer the reader to Ref. 23 and references therein for a more detailed account and description of these data sets. On the one hand, we employ the first 5 (correlated) effective points on the normalized Hubble rate, i.e. $E(z) = H(z)/H_0$, from the Pantheon+MCT sample[24,25], which includes 1063 SnIa. As it is shown in Ref. 25, the compression of the information contained in such SnIa sample is carried out in a very efficient way. It is well-known that SnIa cannot be used alone to determine $H_0$, since this parameter is fully degenerated with the SnIa absolute magnitude. To break such degeneracy, we use the 31 (uncorrelated) Hubble function data points from cosmic chronometers provided in Refs. 26, 27, 28, 29, 30, 31, 32, 33. These data cover a redshift range up to $z \sim 2$ and are obtained without assuming any particular cosmological model, by applying spectroscopic dating techniques of passively-evolving galaxies. They do not rely neither on the Cepheid distance scale nor parallaxes. There are, though, other sources of systematic uncertainties, as the ones associated to the modeling of stellar ages, which is carried out making use of the so-called stellar population synthesis (SPS) models, e.g. the BC03[34] and MaStro[35] ones. Let us detail which SPS models have been used in obtaining the CCH data points of Refs. 26, 27, 28, 29, 30, 31, 32, 33. In Refs. 30, 28, 33 the authors only provide the values of $H(z_i)$ obtained with the BC03 model. This constitutes a $\sim 25\%$ of the whole CCH data set. In Ref. 31 only the combined MaStro/BC03 values are available, whereas in Refs. 26, 27 an alternative SPS model is used, different from the MaStro and BC03 ones. These points constitute the $\sim 32\%$ of the CCH data set. In contrast, in Refs. 29, 32 the authors provide both, the BC03 and MaStro values. This group includes the remaining $\sim 43\%$ of the data. We have opted to use the BC03 values of these two references in our main analyses so as to incorporate consistently the data from Refs. 30, 28, 33, namely to avoid the use of a mixture of MaStro and BC03 values while maximizing the number of data points entering the calculations. The impact of considering alternative data sets has been discussed in detail in Sect. 3.4 of Ref. 23. We will also comment later on on (*sic*) the effect of using a more conservative data set which takes into account the combination of BC03 and MaStro values of Refs. 29, 32 instead of only the BC03 ones. This data set coincides with the one provided in Table 2 of Ref. 36. We refer the reader to this paper for details.

## 3. Reconstructing $H(z)$ from observations

### 3.1. *With Gaussian Processes*

Gaussian distributions are defined for a finite set of quantities, and are characterized by a vector of mean values and the corresponding covariance matrix, which is in charge of controlling the uncertainties' size and the correlations. Gaussian Processes are their direct generalization to the continuum, and are analogously characterized by a mean function $\mu(z)$ and a two-point covariance function $\mathcal{C}(z,z')$, see e.g. Ref. 37,

$$\xi(z) \sim \mathcal{GP}\left(\mu(z), \mathcal{C}(z,z')\right) , \tag{1}$$

where the curve $\xi(z)$ is a realization of the Gaussian process. The covariance $\mathcal{C}(z,z')$ is defined as follows: (i) when $z$ and/or $z'$ do *not* coincide with any point contained in the data set that we want to use to reconstruct the function, then $\mathcal{C}(z,z') = \mathcal{K}(z,z')$, where $\mathcal{K}(z,z')$ is the so-called kernel function, which must be of course symmetric and at this stage is unknown; (ii) when both, $z$ and $z'$, *do* coincide with the $z$-values of one or more points contained in our data set, then $\mathcal{C}(z,z') = \mathcal{D}(z,z') + \mathcal{K}(z,z')$. Here we incorporate the information of the known covariance matrix $\mathcal{D}(z,z')$ of our data points. The kernel function plays a crucial role in the GPs, and must be selected beforehand. Three of the most famous ones with only two degrees of freedom are the following: (a) The Gaussian kernel, $\mathcal{K}(z,z') = \sigma_f^2 \mathrm{Exp}\left[-\frac{1}{2}\left(\frac{z-z'}{l_f}\right)^2\right]$; (b) The Cauchy kernel, $\mathcal{K}(z,z') = \sigma_f^2 l_f/[(z-z')^2 + l_f^2]$; and (c) The Matérn kernel, $\mathcal{K}(z,z') = \sigma_f^2\left[1+\frac{\sqrt{3}}{l_f}|z-z'|\right]\mathrm{Exp}\left[-\frac{\sqrt{3}}{l_f}|z-z'|\right]$. They all depend on $\sigma_f$ and $l_f$, the so-called hyperparameters of the kernel function. The first one controls the uncertainties' size and the strength of the correlations, whereas the second somehow limits the scope of these correlations in $z$. Although it is possible to construct more elaborated kernels, e.g. kernels which do not only depend on the distance $|z-z'|$, but also on the locations $z$ and $z'$ themselves, in such a way that the symmetry under the interchange $z \leftrightarrow z'$ is kept intact, we will stick to the more simple ones mentioned above and, more concretely, only to the Gaussian one, since as we showed in Ref. 23 the reconstructed $H(z)$ obtained with these three kernels are compatible at $< 1\sigma$ c.l. for all the redshift range under study. Of course, the results also depend on the values of the hyperparameters. Once we choose the kernel, how can we properly select $\sigma_f$ and $l_f$? To do so we must make use of our data. In the GPs philosophy, our data set is conceived as part of a subset of realizations of the Gaussian process. The hyperparameters are usually chosen so as to maximize the probability of the GP to produce our data set. If we marginalize the GP (1) over the points at $z^*$, i.e. those at which we do not have data, we get the following multivariate normal distribution,

$$\vec{\xi} \sim \mathcal{N}\left(\{\mu_i(\tilde{z}_i)\}, \mathcal{C}\right) , \tag{2}$$

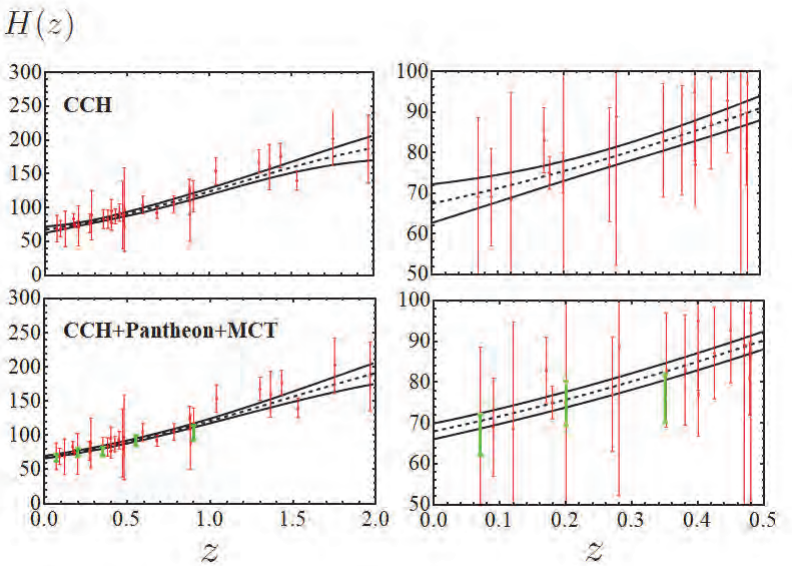



Fig. 1. Reconstructed $H(z)$ in [km/s/Mpc] with the corresponding $1\sigma$ bands obtained with GPs when only CCH are used (figures on the top), and when we add the Pantheon+MCT data (the ones at the bottom). In both cases we zoom in the redshift range $z \in [0, 0.5]$ (plots on the right) so as to better appreciate our determination for $H_0$ and how the uncertainty bands narrow down when the SnIa data are included. The processed ($E(z_i) \to H(z_i)$) Pantheon+MCT points are plotted in green, whereas the CCH data are in red. See related comments in the main text.

where $i = 1, ..., N$, with $N$ being the dimension of the vector of data points $\vec{y} \equiv \{\tilde{z}_i, y_i\}$ at our disposal, and $\mu_i(\tilde{z}_i)$ can be set e.g. to $\vec{0}$ $\forall i$, since the result is almost insensitive to this. Thus, the hyperparameters will be obtained upon the minimization of

$$-2\ln\mathcal{L}(\sigma_f, l_f) = N\ln(2\pi) + \ln|\mathcal{C}(\sigma_f, l_f)| + \vec{y}^T\mathcal{C}^{-1}(\sigma_f, l_f)\vec{y}, \tag{3}$$

with $\mathcal{L}$ being the marginal likelihood and $|\mathcal{C}|$ the determinant of $\mathcal{C}$. Having done this, and using (1), we can compute the conditional probability of finding a given realization of the Gaussian process in the case in which $\xi(\tilde{z}_i) = y_i(\tilde{z}_i)$. The resulting mean and variance functions extracted from such conditioned GP read, respectively,

$$\bar{\xi}(z^*) = \sum_{i,j=1}^{N} \mathcal{C}^{-1}(\tilde{z}_i, \tilde{z}_j) y(\tilde{z}_j) \mathcal{K}(\tilde{z}_i, z^*)\,, \tag{4}$$

$$\sigma^2(z^*) = \mathcal{K}(z^*, z^*) - \sum_{i,j=1}^{N} \mathcal{C}^{-1}(\tilde{z}_i, \tilde{z}_j) \mathcal{K}(\tilde{z}_i, z^*) \mathcal{K}(\tilde{z}_j, z^*)\,. \tag{5}$$

The results derived using this formalism are shown in Fig. 1. When only CCH are included in the analysis we obtain an extrapolated value of $H_0 = (67.42 \pm 4.75)$ km/s/Mpc. This result coincides with the one provided in Ref. 38. We have

incorporated the SnIa information with the following procedure. We have firstly used the value of $H_0$ extracted from the GP-reconstruction of $H(z)$ with CCH to promote (via a Monte Carlo routine) the Hubble rates of Ref. 25 to values on $H(z_i)$. They can be used now in combination with the CCH data in a new GP-reconstruction to obtain the corresponding new value of $H_0$. The result reads, $H_0 = 67.99 \pm 1.94$ km/s/Mpc. The error bars has decreased a factor $\sim 2.5$ with respect to when only CCH are employed. The central value is fully compatible with the Planck+ΛCDM result, and it is in $2.2\sigma$ tension with $H_0^{\rm HST}$. Using the more conservative CCH data set of Ref. 36 in combination with the Pantheon+MCT data one obtains $H_0 = 69.00 \pm 2.35$ km/s/Mpc. In this case the tension with the HST value loosens, being now of $1.7\sigma$. For a discussion on the effect of the propagation of the hyperparameters' uncertainties we refer the reader to Sect. 3.3 of Ref. 23.

### 3.2. *With the Weighted Polynomial Regression method*

Cosmographical analyses are based on truncated expansions (in $z$ or some other variable) of cosmological quantities, e.g. cosmic distances or the Hubble function, around $z = 0$, see e.g. Ref. 39. People usually truncate these expansions (which do not rely on any particular cosmological model) at a concrete order and subsequently fit the resulting expression to the data in order to extract kinematic information of our Universe. The truncation order should not be determined in an *ad hoc* way, but applying some well-motivated criterion, e.g. looking for the order that makes the value of the reduced chi-squared statistic to be as close to one as possible, the one that maximizes the Bayesian evidence (see e.g. Ref. 40), or just the one that minimizes some approximate information criterion statistic, as the Akaike[41] or Bayesian[42] ones. But even these approaches can prove insufficient to obtain objective (non-biased) constraints on the fitting parameters, as it has been highlighted in Refs. 23, 36. However, it is possible to improve the usual cosmographical methodology by making use of the so-called Weighted Function Regression technique, which does not consist on choosing only one particular order of the truncated series, but on weighting in a consistent way the contribution of the various truncated expansions.

Let us explain how it works. Firstly, imagine that we have the following function, which is linear in the parameters,

$$f(z) = \sum_{i=0}^{n} a_i b_i(z)\,, \tag{6}$$

with the $b_i$'s being the so-called basis functions and $\vec{a}$ the vector of coefficients. If we have the vector of mean values $\vec{\bar{a}}$ and the associated covariance matrix $D$, then we can compute the mean function $\bar{f}_M(z)$ and covariance matrix $\mathrm{cov}[f_M(z), f_M(z')]$ as follows:

$$\bar{f}_M(z) = \sum_{i=0}^{n} \bar{a}_i b_i(z)\,, \qquad \mathrm{cov}[f_M(z), f_M(z')] = \sum_{i,j=0}^{n} D_{ij} b_i(z) b_j(z')\,. \tag{7}$$

$\vec{a}$ and $D$ can be computed exactly in the case of Gaussian-distributed data with flat priors (cf. Ref. 23) or, in a more general case, computationally making use of a Monte Carlo routine. The variance of the reconstructed function is just $\sigma_M^2(z) = \mathrm{cov}[f_M(z), f_M(z)]$. These tools can be used, e.g. to reconstruct $H(z)$ in the cosmographic scenario. We just have to Taylor-expand $H(z)$ around e.g. $z = 0$,

$$H(z) = H_0 + \frac{dH}{dz}\Big|_{z=0} z + \frac{1}{2}\frac{d^2H}{dz^2}\Big|_{z=0} z^2 + ... \tag{8}$$

In this case $b_i(z) = z^i$, $a_0 = H_0$, $a_1 = H_0(q_0 + 1)$, etc. Let us call $M_0$, $M_1$,..., $M_{N-1}$ the cosmographic polynomials of order $n = 0$, 1,..., $N-1$, respectively, with $N$ being again the number of data points entering the analysis. That is, let us conceive each polynomial as a different model, and compute the probability density associated to the fact of having a certain shape of the function $f(z)$ as follows,

$$P[f(z)] = k \cdot [P(f(z)|M_0)P(M_0) + ... + P(f(z)|M_{N-1})P(M_{N-1})]\,, \tag{9}$$

where $k$ is just a normalization constant that must be fixed by imposing $\int [\mathcal{D}f]\, P[f(z)] = 1$. Taking into account that $\int [\mathcal{D}f]\, P(f(z)|M_J) = 1 \quad \forall J \in [0, N-1]$ and $\sum_{J=0}^{N-1} P(M_J) = 1$, we find $k = 1$ and therefore:

$$P[f(z)] = \sum_{J=0}^{N-1} P(f(z)|M_J)P(M_J)\,. \tag{10}$$

We now denote $M_*$ as the most probable model and identify $\frac{P(M_J)}{P(M_*)}$ with the Bayes ratio $B_{J*}$[40]. The last expression can be finally written as

$$P[f(z)] = \frac{\sum\limits_{J=0}^{N-1} P(f(z)|M_J)B_{J*}}{\sum\limits_{J=0}^{N-1} B_{J*}}\,. \tag{11}$$

This is the central expression of the WPR method, where the weights are directly given by the Bayes factors. Notice that using (11) we can compute the (weighted) moments and related quantities too, e.g. the weighted mean and variance read,

$$\bar{f}(z) = \frac{\sum\limits_{J=0}^{N-1} \bar{f}_J(z)B_{J*}}{\sum\limits_{J=0}^{N-1} B_{J*}}\,, \qquad \sigma^2(z) = \frac{\sum\limits_{J=0}^{N-1} [\sigma_J^2(z) + (\bar{f}_J(z))^2]B_{J*}}{\sum\limits_{J=0}^{N-1} B_{J*}} - (\bar{f}(z))^2\,, \tag{12}$$

where $\bar{f}_J(z)$ and $\sigma_J^2(z)$ can be computed by means of (7). For the obtention of the Bayes factors we invoke the time-honored Bayesian and Akaike information criteria, BIC and AIC, defined as[41,42]:

$$\mathrm{AIC}_J = \chi^2_{\mathrm{min},J} + \frac{2n_J N}{N - n_J - 1}\,, \qquad \mathrm{BIC}_J = \chi^2_{\mathrm{min},J} + n_J \ln N\,, \tag{13}$$

Table 1. Fitting results obtained for the models with largest weight entering the WPR-reconstruction of $H(z)$ with the CCH data

| Polyn. degree (d) | $\chi^2_{\rm min}$ | BIC | AIC | $H_0$ | $B_{\rm d1}$ (BIC) | $B_{\rm d1}$ (AIC) |
|---|---|---|---|---|---|---|
| 1 | 16.62 | 23.48 | 21.04 | $62.3 \pm 3.1$ | 1 | 1 |
| 2 | 14.74 | 25.04 | 21.63 | $67.8 \pm 5.1$ | 0.47 | 0.74 |
| 3 | 14.78 | 28.51 | 24.31 | $68.9 \pm 4.5$ | 0.08 | 0.19 |
| 4 | 14.80 | 31.97 | 27.20 | $68.2 \pm 4.2$ | 0.014 | 0.05 |
| 5 | 14.95 | 35.56 | 30.45 | $68.4 \pm 4.0$ | 0.002 | 0.009 |
| 6 | 15.01 | 39.05 | 33.88 | $68.6 \pm 4.4$ | $4 \times 10^{-4}$ | 0.002 |
| 7 | 15.27 | 42.74 | 37.82 | $68.3 \pm 4.2$ | $6 \times 10^{-5}$ | $2 \times 10^{-4}$ |
| 8 | 15.72 | 46.63 | 42.29 | $68.8 \pm 4.4$ | $9 \times 10^{-6}$ | $2 \times 10^{-5}$ |
| 9 | 16.38 | 50.72 | 47.38 | $68.9 \pm 3.9$ | $1 \times 10^{-6}$ | $2 \times 10^{-6}$ |

*Note*: Some quantities obtained from the fitting analysis of the first nine (cosmographic) polynomials involved in the WPR-reconstruction, by only using the CCH data. In the second column we show the minimum value of the $\chi^2$ function. The third and fourth columns contain the values of the Bayesian and Akaike information criteria, as defined in (13). In the fifth column we show the values of the Hubble parameter in [km/s/Mpc] together with its $1\sigma$ uncertainty. The last two columns contain the values of the Bayes ratio that are obtained by using the BIC and AIC, respectively, as defined in (14). See comments in the text.

with $\chi^2_{{\rm min},J}$ being the minimum of the $\chi^2$ function in the model $M_J$, and $n_J = J+1$ the dimension of $\vec{a}_J$. In this way the Bayes factor between the model $M_J$ and the most probable model $M_*$ can be approximated by

$$B_{J*} = e^{\frac{BIC_* - BIC_J}{2}} \qquad \text{or} \qquad B_{J*} = e^{\frac{AIC_* - AIC_J}{2}}\,, \tag{14}$$

depending on the criterion used, the BIC or AIC, respectively. The most probable model $M_*$ is defined as the one with lowest BIC or AIC, depending again on the chosen criterion. It is crystal-clear from these expressions that the competing models with more parameters used to analyze the same data receive a suitable penalty and, therefore, our weighted method implements in practice Occam's razor principle.

We apply the WPR formalism to the family of models for $H(z)$ that result from truncating the Taylor series (8) at different orders. We force the coefficients to be positive so as to fulfill: (i) $H_0 \neq 0$; (ii) $H'(z=0) \neq 0$; and (ii) $dH/dz(z) > 0\,\forall z > 0$, cf. Ref. 23 for further details. We present in Table 1 the most relevant quantities obtained from the individual fitting analyses of the truncated expansions entering the WPR-reconstruction with CCH data. Therein we show the values of $H_0$ and the associated uncertainties derived from the individual fits and also the relative weight with respect to the most probable result, which in this case is obtained for the linear polynomial. Applying (12) we are led to $H_0 = (64.44 \pm 4.72)$ km/s/Mpc [$H_0 = (65.27 \pm 4.98)$ km/s/Mpc] when BIC [AIC] is used. When we add the Pantheon+MCT data we obtain $H_0 = (68.90 \pm 1.96)$ km/s/Mpc regardless of the information criteria. If we use the conservative CCH data set from Ref. 36 we obtain $H_0 = (70.50 \pm 2.60)$ km/s/Mpc. These results are completely compatible with the ones obtained using GPs.

## 4. Conclusions

In this work we have presented some estimations of $H_0$ obtained with two alternative reconstruction techniques and data on cosmic chronometers and supernovae of Type Ia. The results are independent from those obtained in Ref. 2, in which the authors made use of low-redshift SnIa and the cosmic distance ladder. Our results do not rely on any particular cosmological model, and seem to favor the low estimates of $H_0$ provided by the Planck Collaboration[3]. Our results are also resonant with the conclusions of Ref. 43, where the authors provided some indication that the local measurement[2] of the Hubble-Lemaître constant is an outlier, by using the so-called index of inconsistency to test the consistency between the preferred values of $H_0$ derived from alternative data sources: CMB from Planck, SnIa+BAO+BBN, data on large scale structure formation, gravitational time delay, and local determination of $H_0$. Other authors[17,38,44,45] have also obtained values of $H_0$ lying in the lower range preferred by Planck, using model-independent techniques and different combinations of intermediate-redshift data from BAO, SnIa, and CCH. Moreover, several analyses have shown that it is possible to loosen or even solve the $\sigma_8$-tension at the expense of keeping the one concerning $H_0$, see e.g. Refs. 22, 46, 47. A low value of $H_0$ would automatically allow to remove the existing tension in the $\sigma_8$-parameter.

Our methodology has proved to be a viable route for estimating $H_0$. Revisiting it in the future, when new and more precise data are available, would be interesting, especially when the theoretical uncertainties associated to the SPS models employed to extract the CCH data from observations decreases in a significant way.

## Acknowledgements

AGV would like to thank Prof. Joan Solà for his invitation to give this talk as part of the CM4 parallel sessions on "Tensions on ΛCDM cosmological model and model-independent constraints" at the 15th Marcel Grossmann Meeting, in the beautiful city of Rome. He is also grateful to the organizers and the *Department of Quantum Physics and Astrophysics* of the University of Barcelona for their economical support. LA is supported by the DFG through TRR33 "The Dark Universe".

## References

1. E. Hubble, Proc. Nat. Acad. Sci. **15** (1929) 168.
2. A.G. Riess *et al.*, ApJ **855** (2018) 136.
3. N. Aghanim *et al.* (Planck 2018 results), arXiv:1807.06209.
4. W. Cardona, M. Kunz and V. Pettorino, JCAP **1703** (2017) 056.
5. B.R. Zhang *et al.*, MNRAS **471** (2017) 2254.
6. S. Jang and M.G. Lee, ApJ **836** (2017) 74.
7. A.E. Romano, IJMPD **27** (2018) 1850102.
8. R.C. Keenan, A.J. Barger and L.L. Cowie, ApJ **775** (2013) 62.

9. T. Shanks, L.M. Hogarth and N. Metcalfe, MNRAS **484** (2019) L64.
10. A.G. Riess, S. Casertano, D. Kenworthy, D. Scolnic and L. Macri, arXiv:1810.03526; T. Shanks, L.M. Hogarth and N. Metcalfe, arXiv:1810.07628.
11. G.E. Addison *et al.*, ApJ **853** (2018) 119.
12. V. Bonvin *et al.*, MNRAS **465** (2017) 4914.
13. S. Birrer *et al.*, MNRAS **484** (2019) 4726.
14. B.P. Abbott *et al.*, Nature **551** (2017) 85.
15. C. Guidorzi *et al.*, ApJ **851** (2017) L36.
16. K. Hotokezaka *et al.*, arXiv:1806.10596.
17. S.M. Feeney *et al.*, PRL **122** (2019) 061105.
18. V. Marra, L. Amendola, I. Sawicki and W. Valkenburg, PRL **110** (2013) 241305.
19. H-Y. Wu and D. Huterer, MNRAS **471** (2017) 4946.
20. D. Camarena and V. Marra, PRD **98** (2018) 023537.
21. J.L. Bernal, L. Verde and A.G. Riess, JCAP **1610** (2016) 019.
22. J. Solà, A. Gómez-Valent and J. de Cruz Pérez, PLB **774** (2017) 317.
23. A. Gómez-Valent and L. Amendola, JCAP **1804** (2018) 051.
24. D.M. Scolnic *et al.,* ApJ **859** (2018) 101.
25. A.G. Riess *et al.,* ApJ **853** (2018) 126.
26. R. Jiménez, L. Verde, T. Treu & D. Stern, ApJ **593** (2003) 622.
27. J. Simon, L. Verde & R. Jiménez, Phys. Rev. D **71** (2005) 123001.
28. D. Stern *et al.,* JCAP **1002** (2010) 008.
29. M. Moresco *et al.,* JCAP **1208** (2012) 006.
30. C. Zhang *et al.,* Res. Astron. Astrophys. **14** (2014) 1221.
31. M. Moresco, MNRAS **450** (2015) L16.
32. M. Moresco *et al.,* JCAP **1605** (2016) 014.
33. A.L. Ratsimbazafy *et al.,* MNRAS **467** (2017) 3239.
34. G. Bruzual and S. Charlot, MNRAS **344** (2003) 1000.
35. C. Maraston and G. Strömbäck, MNRAS **418** (2011) 2785.
36. A. Gómez-Valent, arXiv:1810.02278, accepted for publication in JCAP (2019).
37. C.E. Rasmussen and C.K.I. Williams, *Gaussian processes for machine learning*, MIT Press, U.S.A. (2006).
38. H. Yu, B. Ratra and F-Y. Wang, ApJ **856** (2018) 3.
39. M. Visser, CQG **21** (2004) 2603; GRG **37** (2005) 1541.
40. L. Amendola and S. Tsujikawa, *Dark energy. Theory and observations*, Cambridge Univ. Press, Cambridge (2010).
41. H. Akaike, IEEE Trans. Autom. Control. **19** (1974) 716.
42. G. Schwarz, Ann. Statist. **6** (1978) 461.
43. W. Lin and M. Ishak, PRD **96** (2017) 083532.
44. B.S. Haridasu *et al.*, JCAP **1810** (2018) 015.
45. E. Macaulay *et al.* (DES Collab.), MNRAS **486** (2019) 2184.
46. A. Gómez-Valent and J. Solà, MNRAS **478** (2018) 126; EPL **120** (2017) 39001.
47. S. Anand, P. Chaubal, A. Mazumdar and S. Mohanty, JCAP **1711** (2017) 005.

# Lambda-CDM model and small-scale-cosmology "crisis": from astrophysical explanations to new fundamental physics models

Nick E. Mavromatos

*Theoretical Particle Physics and Cosmology Group, Department of Physics, King's College London, Strand, London WC2R 2LS, UK*
*Nikolaos.Mavromatos@kcl.ac.uk*

I review briefly three problems where tensions between predictions based on numerical simulations of the Lambda-CDM prototype and observations at small (galactic) scales occur. These include (i) the core-cusp problem on the galactic profiles, (ii) the missing satellite problem, and the (ii) too-big-to-fail problem. I explain what these problems are and present potential resolutions, first through some astrophysical mechanisms, which however, as I argue, fail to alleviate completely the problems, at least currently. Then, I discuss fundamental modifications of the Lambda-CDM model, through the inclusion of self-interacting dark matter (SIDM). I argue that a simple model of SIDM, with (warm) self-interacting right-handed neutrinos (RHN), that exist in minimal extensions of the Standard model of particle physics, appears promising in providing a resolution of the aforementioned "small-scale-Cosmology crisis", in particular the core-cusp problem, and an observationally consistent description of the core-halo structure in galaxies.

*Keywords*: Cosmological observations; self-interacting interacting dark matter; right-handed-neutrinos.

## 1. Introduction

A plethora of precision cosmological observations[1], based on type Ia supernovae, cosmic microwave background (CMB) radiation, large scale structure, baryon acoustic oscillations and strong and weak lensing data, point towards the fact that only $\sim 4\%$ of the current energy budget of the Universe consists of the known form of matter (mainly baryons), while the rest seems to be of uknown nature, specifically $\sim 69\%$ dark energy (DE) and $\sim 27\%$ dark matter (DM). The current universe appears to enter an acceleration phase (for a second time after inflation), and the data point towards an equation of state for the dark energy to be close to that of a de Sitter space-time, $w \simeq -1$, corresponding to a positive cosmological-constant, $\Lambda > 0$. The best fit to the large-scale data seems to be provided by the so-called $\Lambda$-Cold-Dark-Matter ($\Lambda$CDM) model for the Universe, according to which the DE is provided by a positive cosmological constant $\Lambda > 0$, while the DM content of the Universe is thermal, of "Cold" type, corresponding to a mass range for the (assumed) dominant DM particle species $m_\chi > 100$ keV. Several particle physics models, in particular Supersymmetry, provide candidates for this type of DM, although currently there is no experimental evidence from collider experiments that such extensions of the Standard Model (SM) are realised in nature[2], thus leaving us only with Cosmological evidence for the existence of DM.

Despite the success of the $\Lambda$CDM model at large (cosmological) scales, however, there are tensions at galactic scales which, if not due to observational fluctuations, or admitting astrophysical resolution, may hint new physics beyond the $\Lambda$CDM

paradigm. In this talk we mention briefly some of these challenges, and discuss their potential resolution by modifying the ΛCDM model via the inclusion of self interactions in the Dark Matter sector.[a]

## 2. Small Scale Challenges of the ΛCDM Model

There are three challenges to the ΛCDM model at galactic scales, collectively called some times as "*small-scale cosmology crisis*, which we focus our attention upon in this section:

**(i)** *The Core-Cusp problem*, that is, a discrepancy between the observed dark matter density profiles of low-mass galaxies and the density profiles predicted by cosmological N-body simulations. Characteristically, all the ΛCDM-based (DM only) simulations form dark matter halos which have "cuspy" dark matter distributions, with the density increasing steeply, i.e. as $\rho \propto r^{-1}$, at small radii. This is, e.g., evidenced in the standard Navarro-Frenk-White (NFW) DM profile.[5] On the contrary, the rotation curves of most of the observed dwarf galaxies indicate flat central density profiles ("cores").[6]

**(ii)** *The "missing satellite or dwarf galaxy problem"*, which pertains to a discrepancy between observations and ΛCDM-based numerical cosmological simulations, which predict the evolution of the distribution of matter in the universe. Such simulations point towards a hierarchical clustering of DM, where smaller halos merge to form larger halos. Although there seem to be enough observed normal-sized galaxies to account for such a numerical distribution, the number of (observed) dwarf galaxies is orders of magnitude lower than that expected from the simulations.[7]

**(iii)** *The too big to fail problem*, which is a discrepancy arising between the most massive subhaloes predicted in (dissipationless) ΛCDM simulations and the observed dynamics of the brightest dwarf spheroidal (dSph) galaxies in the Milky way. In other words, the ΛCDM simulations predict that the most massive subhaloes of the Milky way are too dense to host any of its bright satellites, with luminosity higher than $10^5$ the luminosity of the Sun.[8]

The problems may admit astrophysical explanations, and some of them may be alleviated by improved measurements. For instance, many studies have shown that including baryonic feedback (particularly feedback from supernovae and active galactic nuclei) can "flatten out" the core of a galaxy's dark matter profile, since feedback-driven gas outflows produce a time-varying gravitational potential that transfers energy to the orbits of the collisionless dark matter particles.[9] The missing satellite problem also admits standard Astrophysical solutions.[10] One is that the smaller halos do exist but only a few of them end up becoming visible because they have not been able to attract enough baryonic matter to create a visible dwarf

[a] We do not discuss here other recent challenges of ΛCDM, e.g. those associated with the different value of the Hubble parameter $H_0$, obtained between CMB/Planck-Coll. measurements[1] and Cepheid-galaxies observations[3], which could be alleviated via modified DE models, such as the "running vacuum".[4]

galaxy. In support of this, Keck observations in 2007 of eight newly discovered ultra-faint Milky Way dwarf satellites showed that six were almost exclusively composed of DM, around 99.9% (with a mass-to-light ratio of about 1000).[11] Such ultra-faint dwarfs substantially alleviate the discrepancy between the predicted and observed numbers of satellites around the Milky Way, but there are still discrepancies by a factor of about four too few dwarf galaxies over a significant range of masses. See also the recent work Ref. 12 in this direction. Other solutions may be that dwarf galaxies tend to be merged into or tidally stripped apart by larger galaxies due to complex interactions. This tidal stripping has been part of the problem in identifying dwarf galaxies in the first place, which is an extremely difficult task since these objects have low surface brightness and are highly diffused, so much that they are virtually unnoticeable. Finally, the Too-Big-to-Fail Problem may also be tackled by taking into account baryonic feedback.

From the point of view of the ΛCDM simulations, all three problems have their root in the fact that the cold DM particles, which the ΛCDM simulations rely upon, have too short free streaming length during the epochs of galaxy formation, and therefore they form too clumped and too many structures compared to those observed. The interesting question therefore is whether these problems, especially the core-casp problem, provide hint towards a paradigm shift of the underlying fundamental physics, implying deviations from ΛCDM framework. One such deviation is to include self interactions among the DM particles.

## 3. Resolution through Self Interacting Dark Matter Models

Self interactions have been argued to play an important rôle in galactic structure already in Ref. 13. The original idea of Self-Interacting DM (SIDM) was implemented for CDM particles with rest masses above 1 MeV/$c^2$ (up to 10 GeV/$c^2$), consistent with the nature of the effective interactions and the mean free paths considered in that work. This way of thinking regarding self-interactions was applied uniquely on DM halo scales with typical densities of $10^{-2} M_\odot/\mathrm{pc}^3$, suggesting that normalized total cross-sections of order $\sigma/m \sim 0.1 - 100$ cm$^2$/g, would imply observational effects in the inner regions of the DM halos. It was also shown that a SIDM regime with these values of $\sigma/m$ would generate shallower inner DM profiles, with a necessary reduction in the amount of sub-structures, thereby alleviating the core-cusp and the missing satellite problems of collisionless ΛCDM[14]. However, contemporaneously, some tension with upper limits in the DM cross sections as obtained from lensing studies on galaxy cluster scales emerged. In a subsequent work,[15] motivated by updated analysis of the Bullet Cluster,[16] new cosmological simulations within CDM were performed, with the aim of further scrutinizing the effects of SIDM on inner halo cores of galaxies and galaxy clusters. The authors of Ref. 15 concluded that $\sigma/m \sim 0.2$ barn GeV$^{-1} = 0.1$ cm$^2$ g$^{-1}$ is consistent with all the observational constraints. In general, SIDM would make no difference from ΛCDM at large scales,

but individual galaxies would appear more cored and spherical, with higher velocity dispersion.

Stringent constraints of the interaction cross section of SIDM have been inferred in Ref. 17 by means of new DM observables in merging galaxies, taking into account *the DM drag*. In particular, the study of 72 mergers using this new technique imposed a more stringent upper bound on the SIDM cross section per unit DM mass, thus defining a new range for the total SIDM cross section[b]

$$0.1 \leq \frac{\sigma_{\rm SIDM}/m}{{\rm cm}^2\, g^{-1}} \leq 0.47 \tag{1}$$

This might lead to a potential resolution of the three "small scale cosmology problems" of DM, discussed previously. For instance, self interactions may tackle the core-cusp and too big to fail problems by reducing the central density. In this latter respect we mention the works of Ref. 20, according to which self interactions can lead to both deceleration and evaporation of a DM halo when the latter moves through a background of DM particles. This results in a shift of the halo's centroid relative to the collisionless stars and galaxies.

## 4. Self-Interacting Right-Handed Neutrinos as (Warm) DM and Galactic Structure

A concrete model of SIDM, which may have significant effects on the galactic structure, and which plays an important rôle in particle physics, has been studied in Ref. 21. It is based on a class of self-interacting models of massive right-handed neutrinos that exist in minimal extensions of the standard model[22]. The model constitutes a microscopic (particle physics) realisation of previous ideas that populations of electrically neutral fermionic particles might play an important rôle for the galactic core-halo structure[23]. In such a scenario, the DM fermion is played by the right-handed neutrino (RHN), which has a mass of a order a few tens of keV, and implies a universal and novel DM density profile (*compact core - dilute halo*), with important implications for the very central and halo regions of galaxies. The model is in agreement with observations, provided sufficiently strong self-interactions among the RHNs are considered[21]. The Lagrangian of the right-handed neutrino sector, including gravity, reads (in units $\hbar = c = 1$, which we use throughout here):

$$\mathcal{L} = \mathcal{L}_{GR} + \mathcal{L}_{N_{R\,1}} + \mathcal{L}_V + \mathcal{L}_I \tag{2}$$

[b]However, some tension with the upper bound of (1) still remains, since the analysis and assumptions of Ref. 18 on their study of the Abell Cluster 3827 have been questioned in Ref. 19, which presented corrected estimates for the SIDM cross section: $\sigma_{\rm SIDM}/m \simeq 3\ {\rm cm}^2\, g^{-1}$, when self interactions result in a drag force, and $\sigma_{\rm SIDM}/m \simeq 1.5\ {\rm cm}^2\, g^{-1}$ in the case of contact interactions.

where

$$\mathcal{L}_{GR} = -\frac{R}{16\pi G},\ \mathcal{L}_{N_{R\,1}} = i\,\overline{N}_{R\,1}\gamma^{\mu}\,\nabla_{\mu}\,N_{R\,1} - \frac{1}{2}m\,\overline{N^{c}}_{R\,1}N_{R\,1},$$
$$\mathcal{L}_{V} = -\frac{1}{4}V_{\mu\nu}V^{\mu\nu} + \frac{1}{2}m_{V}^{2}V_{\mu}V^{\mu},\ \mathcal{L}_{I} = -g_{V}V_{\mu}J_{V}^{\mu} = -g_{V}V_{\mu}\overline{N}_{R\,1}\gamma^{\mu}N_{R\,1}\,, \quad (3)$$

with $R$ the Ricci scalar for the static spherically symmetric metric background considered in Ref. 21, $\nabla_{\mu}$ the gravitational covariant derivative acting on a Majorana spinor, $m$ the mass of the sterile RHN neutrino, and $m_V$ the mass of the dark vector boson providing the carrier of the self-interactions among RHN. When the Lagrangian (3) is expressed in terms of the Majorana field $N^c = N$, the massive-vector-mesons $V_\mu$ couple to the axial current $\overline{N}\gamma^{\mu}\gamma^{5}N$.

In general one may add to (2) a Yukawa term, coupling RHNs to the active neutrino sector (see, e.g., the case of $\nu$MSM in Ref. 22)

$$\mathcal{L}_{\rm Yuk} = \sum_{\alpha,\,I} F_{\alpha I}\,\overline{\ell}_{\alpha}\,N_{R\,I}\phi^{c} + \text{h.c.}\,, \quad I = 1,2,3, \quad \alpha = e,\mu,\tau, \quad (4)$$

where $\ell_\alpha$ are the lepton doublets of the SM, $\alpha = e,\mu,\tau$, $F_{\alpha I}$ are appropriate Yukawa couplings, and $\phi^c$ is the SM conjugate Higgs field, *i.e.* $\phi^{c} = i\tau_{2}\phi^{\star}$, with $\tau_2$ the $2\times 2$ Pauli matrix. If one ignores self interactions among the RHN, then, upon considering such a coupling, one obtains stringent X-ray and Big Bang Nucleosynthesis (BBN) constraints[24] of the mixing angle and mass of $N_{R\,1}$, given that (4) implies decays of the heavy neutrinos $N_I \to \nu H$, where $H$ is the Higgs excitation field, $\phi = v + H$, with $v$ the vacuum expectation value. In the context of $\nu$MSM[22], the decay time of the lightest of the RHN neutrinos can be longer than the age of the universe, hence the latter can be considered as stable for all practical purposes, thus playing the rôle of the entirety of DM. This assumption requires sufficiently low mixing with the standard model sector, with the corresponding mixing angle given by $\sin^2\theta = \sum_{\alpha=e,\mu.\tau}\frac{\langle\phi\rangle^2 F_{\alpha\,I}}{m^2}$. On the other hand, the requirement that the sterile neutrino constitutes the entirety of DM, requires it to be produced in sufficient amounts in the early Universe, and hence, if no self-interactions are assumed, there are lower bounds for $\theta$, depending on the RHN mass $m$. The best fit to date[24] points to a narrow allowed region in the RHN mass $m \sim 10$ keV and mixing angle $\sin^2(2\theta) \sim 10^{-12}$ (slightly wider regions for $m$ are allowed[25] if one uses the galactic density profile of Ref. 23).

In the RHN SIDM model of Ref. 21 the mixing with the SM sector was ignored, $F_{\alpha 1} = 0$, in which case the lightest neutrino is absolutely stable. The important feature are the self-interactions of the right-handed neutrino, which ensured phenomenologically correct values for the radius and mass of the galactic core within certain mass density profiles[23]. Such interactions among RHNs also lead to production mechanisms of the RHN DM that no longer require lower bounds on the mixing angle $\theta$ with the standard model sector, hence the stringent X-ray constraints of Ref. 24 can be evaded. The analysis of Ref. 21 showed that, under the assumption of the absence of a central black hole in the galaxy[23], the allowed range for RHN

DM mass that fit the Milky Way observables is: 47 keV/$c^2 < m <$ 350 keV/$c^2$, the upper bound being set by reaching the critical core mass for gravitational collapse, $M_c^{cr} \propto M_{\rm Pl}^3/m^2 \approx 4.4 \times 10^6 M_\odot$. One also concludes that the inclusion of sufficiently strong interactions in the dark sector among the sterile neutrinos can lead to significantly more compact cores and higher central degeneracies than in the free case[23c].

The total RHN-RHN scattering cross-section in the quantum core of the Galaxy has been calculated in Ref. 21 in a perturbative regime $g < 1$ for the dimensionless interaction coupling, $\sigma_{core}^{tot} \approx 29\,\frac{(g_V/m_V)^4}{4^3\pi}\,m^2$ $(p^2/m^2 \ll 1)$. This allows for a comparison of these results with the generic ones existing in the literature and mentioned in the previous section, on the required range of the total cross section per unit DM mass, $\sigma/m$, in order to resolve the small-scale Cosmology "crisis" (Eq. (1)). To put things in perspective, one can normalize the interaction field strength in terms of the visible sector (SM) weak interaction dimensionfull coupling, the Fermi "constant" $G_F$. Thus, if, e.g. one constrains the total cross-section to the N-body simulation value $\sigma^{tot}/m = 0.1$ cm$^2$/g (or, in general, to lie in the region (1) discussed above), the coupling constant $\overline{C}_V$ would be constrained to the value $\overline{C}_V \in (2.6 \times 10^8, 7 \times 10^8)$, for RHN-DM masses in the range $m \in (47, 350)$ keV. It worths noticing that for $C_V \sim 10^8\, G_F$, the mass of the massive-vector meson would be constrained to values $m_V \lesssim 3 \times 10^4$ keV, in order to satisfy $g_V \lesssim 1$ as requested by the self-consistency of the perturbation scheme applied to the computation of the cross-section[21].

As stressed in Ref. 21, the inclusion of baryonic matter is not expected to change the basic conclusions that the introduction of WDM fermion self interactions affects the core/halo structure and in particular induces higher central degeneracies and higher compactness of the inner quantum core of galaxies. Moreover, the model of self-interacting sterile neutrinos provides a natural resolution to the core-cusp problem of DM at small scales, because the density profiles based on fermionic phase-space distributions develop always an extended plateau on halo scales, in a way that resemble Burkert or cored Einasto profiles. Another important feature is the fact that the right-handed neutrino DM mass is 'colder' by a few keV as compared to most of the Warm Dark Matter (WDM) models available in the literature, which implies that the model does not suffer from such standard WDM problems.

It goes without saying, that the presence of self-interacting right-handed neutrinos with a mass range of a few tens of keV may co-exist harmonically with other types of DM, given that the latter may consist (like ordinary matter) of more than one dominant species. In this way, at large scales, the predictions of the Lambda-CDM model might still be valid, while at galactic scales the rôle of self-interacting

[c]If the assumption that the RHN DM co-exists with a central black hole is made, then one may accommodate the narrow window of Ref. 24 in the analysis of Ref. 21, although, as already mentioned, the existence of self interactions of RHN evades the constraints of Ref. 24.

RHN WDM may play a crucial rôle in providing agreement of the predicted galactic structure and the pertinent DM distribution with observations.

**Acknowledgments**

NEM wishes to thank the organisers of the CM4 session of MG15 for the invitation to give this short review. The work is supported in part by the UK Science and Technology Facilities research Council (STFC) under the research grant ST/P000258/1.

**References**

1. P. A. R. Ade *et al.* [Planck Collaboration], *Planck 2015 results. XIII. Cosmological parameters*, arXiv:1502.01589 [astro-ph.CO].
2. V. A. Mitsou, arXiv:1903.11589 [hep-ph], talk in DM1 session of MG15, these proceedings, and references therein.
3. A. G. Riess *et al.*, Astrophys. J. **730**, 119 (2011) Erratum: [Astrophys. J. **732**, 129 (2011)] doi:10.1088/0004-637X/732/2/129, 10.1088/0004-637X/730/2/119 [arXiv:1103.2976 [astro-ph.CO]].
4. J. Solà, A. Gómez-Valent and J. de Cruz Pérez, Phys. Lett. B **774**, 317 (2017) doi:10.1016/j.physletb.2017.09.073 [arXiv:1705.06723 [astro-ph.CO]]. Astrophys. J. **836**, no. 1, 43 (2017) doi:10.3847/1538-4357/836/1/43 [arXiv:1602.02103 [astro-ph.CO]]; see also talk by J. Sola in DM1 session of MG15, these proceedings.
5. J. F. Navarro, C. S. Frenk, and S. D. M. White, Astrophys. J. **490**, 493 (1997) http://adsabs.harvard.edu/abs/1997ApJ...490..493N [astro-ph/9611107].
6. B. Moore, *Nature* 370, 629 - 631 (25 August 1994); doi:10.1038/370629a0; W. J. G. de Blok, *The Core-Cusp Problem*, Advances in Astronomy 2010 789293 (2010) [arXiv:0910.3538]; Se-Heon Oh *et al.*, Astrophys. J. **149** (6), 96 (2015).
7. B. Moore, S. Ghigna, F. Governato, G. Lake, T. R. Quinn, J. Stadel and P. Tozzi, Astrophys. J. **524**, L19 (1999) doi:10.1086/312287 [astro-ph/9907411]; A. A. Klypin, A. V. Kravtsov, O. Valenzuela and F. Prada, Astrophys. J. **522**, 82 (1999) doi:10.1086/307643 [astro-ph/9901240]; E. Polisensky and M. Ricotti, Phys. Rev. D **83**, 043506 (2011) doi:10.1103/PhysRevD.83.043506 [arXiv:1004.1459 [astro-ph.CO]].
8. M. Boylan-Kolchin, J. S. Bullock and M. Kaplinghat, Mon. Not. Roy. Astron. Soc. **415**, L40 (2011) [arXiv:1103.0007]; *ibid.* **422**, 1203 (2012) [arXiv: 1111.2048].
9. J. Navarro *et al.* Mon. Not. Roy. Astron. Soc. **283** L72 (1996) doi:10.1093/mnras/283.3.72L; A. Pontzen *et al.*, *Nature* **42**, 3464 (2012) doi:10.1111/j.1365-2966.2012.20571.
10. J.S. Bullock, *Notes on the Missing Satellites Problem* [arXiv:1009.4505], Bibliographic Code: 2010arXiv1009.4505B, and references therein.
11. J. D. Simon and M. Geha, Astrophys. J. **670**, 313 (2007) doi:10.1086/521816 [arXiv:0706.0516 [astro-ph]].

12. S. Y. Kim, A. H. G. Peter and J. R. Hargis, Phys. Rev. Lett. **121**, no. 21, 211302 (2018) doi:10.1103/PhysRevLett.121.211302 [arXiv:1711.06267 [astro-ph.CO]].
13. D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. **84**, 3760 (2000) doi:10.1103/PhysRevLett.84.3760 [astro-ph/9909386].
14. M. Lovell *et al.,* Mon. Not. Roy. Astron. Soc. **420**, 2318 (2012). doi:10.1111/j.1365-2966.2011.20200.x; O. Elbert *et al.* Mon. Not. Roy. Astron. Soc. **453**, 29 (2015) doi:10.1093/mnras/stv1470
15. M. Rocha, A. H. G. Peter, J. S. Bullock, M. Kaplinghat, S. Garrison-Kimmel, J. Onorbe and L. A. Moustakas, Mon. Not. Roy. Astron. Soc. **430**, 81 (2013) doi:10.1093/mnras/sts514 [arXiv:1208.3025 [astro-ph.CO]].
16. S. W. Randall, M. Markevitch, D. Clowe, A. H. Gonzalez and M. Bradac, Astrophys. J. **679**, 1173 (2008) doi:10.1086/587859 [arXiv:0704.0261 [astro-ph]].
17. D. Harvey, R. Massey, T. Kitching, A. Taylor, E. Tittley *Science* **347**, 6229, 1462 (2015) and references therein.
18. R. Massey, L. L. R. Williams, R. Smit, M. Swinbank, T. Kitching *et al.*, Mon. Not. Roy. Astron. Soc. **449**, 3393 (2015).
19. F. Kahlhoefer, K. Schmidt-Hoberg, J. Kummer and S. Sarkar, Mon. Not. Roy. Astron. Soc. **452**, no. 1, L54 (2015) doi:10.1093/mnrasl/slv088 [arXiv:1504.06576 [astro-ph.CO]].
20. F. Kahlhoefer, K. Schmidt-Hoberg, M. T. Frandsen and S. Sarkar, Mon. Not. Roy. Astron. Soc. **437**, no. 3, 2865 (2014) doi:10.1093/mnras/stt2097 [arXiv:1308.3419 [astro-ph.CO]].
21. C. R. Argüelles, N. E. Mavromatos, J. A. Rueda and R. Ruffini, **JCAP** 04 (2016) 038 http://dx.doi.org/10.1088/1475-7516/2016/04/038 arXiv:1502.00136 [astro-ph.GA].
22. T. Asaka, S. Blanchet and M. Shaposhnikov, Phys. Lett. B **631**, 151 (2005) doi:10.1016/j.physletb.2005.09.070 [hep-ph/0503065]; A. Boyarsky, O. Ruchayskiy and M. Shaposhnikov, Ann. Rev. Nucl. Part. Sci. **59**, 191 (2009) doi:10.1146/annurev.nucl.010909.083654 [arXiv:0901.0011 [hep-ph]] and references therein.
23. R. Ruffini, C. R. Argüelles and J. A. Rueda, Mon. Not. Roy. Astron. Soc. **451**, no. 1, 622 (2015) doi:10.1093/mnras/stv1016 [arXiv:1409.7365 [astro-ph.GA]].
24. K. Perez, K. C. Y. Ng, J. F. Beacom, C. Hersh, S. Horiuchi and R. Krivonos, Phys. Rev. D **95**, no. 12, 123002 (2017) doi:10.1103/PhysRevD.95.123002 [arXiv:1609.00667 [astro-ph.HE]].
25. R. Yunis, C. R. Argüelles, N. E. Mavromatos, A. Moliné, A. Krut, J. A. Rueda and R. Ruffini, arXiv:1810.05756 [astro-ph.GA].

# Modified gravity, gravitational waves and the large-scale structure of the Universe: A brief report

Ippocratis D. Saltas

*CEICO, Institute of Physics of the Czech Academy of Sciences, Na Slovance 2, 182 21 Praha 8, Czechia*

Luca Amendola

*Institut fr Theoretische Physik, Ruprecht-Karls-Universitt Heidelberg, Philosophenweg 16, 69120 Heidelberg, Germany*

Martin Kunz

*Dpartment de Physique Thorique and Center for Astroparticle Physics, Universit de Genve, Quai E. Ansermet 24, CH-1211 Genve 4, Switzerland*

Ignacy Sawicki

*CEICO, Institute of Physics of the Czech Academy of Sciences, Na Slovance 2, 182 21 Praha 8, Czechia*

The goal of this short report is to summarise some key results based on our previous works on model independent tests of gravity at large scales in the Universe, their connection with the properties of gravitational waves, and the implications of the recent measurement of the speed of tensors for the phenomenology of general families of gravity models for dark energy.

*Keywords*: Dark energy, gravitational waves, large-scale structure of the Universe.

## 1. Introduction

The physics behind the late-time acceleration of the Universe remains elusive. Although the ΛCDM model provides a good and simple phenomenological paradigm for its description, its fundamental understanding suffers due to the current paradigm to explain the nature of the cosmological constant.[1–3]

A description of the accelerating Universe might be embedded within an infrared modification of the standard gravitational paradigm of General Relativity (GR). This idea has sparked the construction of theories which modify the effective gravitational degrees of freedom at large scales, and are usually labelled as dark energy and/or modified gravity models (for detailed expositions, see e.g., Refs. 4, 5). Typical examples are the general scalar-tensor (Horndeski and beyond as well as the so–called DHOST theories), vector-tensor (Einstein-Aether, generalised Proca) and massive/bi-metric theories of gravity. These are the kind of models that will concern us in the discussion to follow.

The current major task of research in cosmology is the investigation of possible such departures from the standard framework of GR. From an observational standpoint, the currently unknown detailed physics of the dark sector make this particularly challenging, due to the unobservability of the *individual* dark components

under pure gravity probes. Here we will explain how *the gravitational slip* parameter is a powerful model independent discriminator between large classes of gravity models in cosmology, and discuss its intimate relation to the properties of gravitational waves (GWs),[6,7] that provides a phenomenological bridge between the physics of GWs and the large scale structure of the Universe. In light of the recent measurement of the speed of GWs, a reverse engineering allows to predict the consequences for large classes of gravity models at large scales.

This summary paper is based on our previous original works which appeared in Refs. 8–12, which we refer to for a full list of references.

## 2. The gravitational slip and model independence

[a]Large scale inhomogeneities are sourced by scalar fluctuations of the metric around the homogeneous and isotropic FLRW spacetime. We assume that scalar metric fluctuations are described according to $ds^2 = -(1+2\Psi(\mathbf{x},t))dt^2 + a(t)^2(1-2\Phi(\mathbf{x},t))d\mathbf{x}^2$. At the linear level, the anisotropy constraint reads as

$$\Phi - \Psi = \sigma(t)\Pi(t,k). \tag{1}$$

Here $\Pi$ is a functional of background quantities and linear perturbation variables (time and scale dependent), while $\sigma(t)$ depends on background quantities only[b]. The gravitational slip parameter $\eta$ is a direct measure of scalar (linear) anisotropic stress and is defined as

$$\eta \equiv \frac{\Phi}{\Psi}. \tag{2}$$

Anisotropic stress therefore sources a gravitational slip, i.e $\eta \neq 1$. At late times, the free streaming of neutrinos makes a negligible contribution to the anisotropic pressure part of the total energy-momentum in the Universe. Instead, its only source in the presence of perfect fluid matter is a possible modification of gravity. For minimally coupled scalar field theories (quintessence, k–essence) or $\Lambda CDM$, it is $\eta = 1$, while $\eta \neq 1$ in non-minimally coupled models such as $f(R)$/Brans–Dicke. Obviously this discussion needs to be made on firm grounds – the distinction between different models will be made precise in section 3 when we will introduce our working definition of modified gravity.

We will now show how $\eta$ can be reconstructed in a model-independent manner, closely following the original work of Refs 8 and 9. Let us start by assuming that galaxies move on geodesics, a valid assumption as long as non-linearities associated with e.g spacetime effects or galaxy-galaxy interactions can be neglected. In addition, it uniquely fixes the conformal frame, an issue implicit in theories with

[a]This section closely follows the original work presented in Refs. 8 and 9.
[b]It is not hard to see that, for any model the anisotropy constraint can be always brought into this convenient form.

different conformal representations (the so–called Jordan/Einstein frame). The resulting geodesic equation reads as

$$\left(a^2\theta_{\rm gal}\right)' = a^2 H k^2 \Psi\,. \tag{3}$$

Therefore, the observable galactic velocity field $\theta_{\rm gal}$ can be simply seen as a measurement of the gravitational potential $\Psi$ at different redshifts and scales.

Weak lensing observations provide a measurement of the lensing potential defined through

$$k^2\Phi_{\rm lens} \equiv k^2(\Psi+\Phi) \equiv -\frac{3(1+z)^3}{2E^2}L\,, \tag{4}$$

where $L \equiv \Omega_{\rm m0}GY(1+\eta)\delta_{\rm m,0}$. Therefore, tomographic weak lensing observations provide a measurement of the linear combination of $\Phi$ and $\Psi$. Now, combining equations (3) and (4) we can derive a model-independent relation for the gravitational slip solely in terms of observable quantities as[8]

$$\eta = \Phi/\Psi = \frac{3(1+z)^3}{2E^2\left(\frac{R'}{R}+2+\frac{E'}{E}\right)}\frac{L}{R} - 1\,, \tag{5}$$

where we introduced the velocity-field observable $R \equiv -\theta_{\rm gal}/H(z)$, and the background evolution $E(z) \equiv H(z)/H_0$, while $L$ was defined after (4).

## 3. Modified gravity and gravitational waves

The distinction between dark energy and modified gravity is sometimes blurry. For a meaningful and unambiguous discussion on how to discriminate between $\Lambda$CDM/dark energy and modified gravity models, we need to have a consistent definition of what we mean by dark energy and modified gravity in a given context. Here, following Ref. 10 we will define a modification of gravity as any theory that changes the propagation of the dynamical degrees of freedom of GR, i.e the gravitational waves (GWs). Assuming that GWs are far away from their source, we write

$$h_{ij}'' + \big(2+\nu(t)\big)Hh_{ij}' + c_{\rm T}^2(t)k^2h_{ij} + a^2\mu^2h_{ij} = a^2\Gamma(t)\gamma_{ij}(t)\,, \tag{6}$$

where the field $h_{ij}$ is the spin-two fluctuation of the metric. FLRW symmetry ensures that, general extensions of GR can modify the propagation of GWs only through one of the following parts: the friction ($\nu(t)$), the propagation speed ($c_{\rm T}^2$), mass ($\mu^2$) and source term ($\Gamma\gamma_{ij}$). Let us now introduce the models that we will be considering, which we conveniently divide into models introducing a new dynamical field of spin-0 (scalar-tensor), spin-1 (vector-tensor) and a spin-2 (massive/bigravity) respectively:

*Scalar-tensor theories:* These start with Brans-Dicke gravity and its generalisations, i.e Horndeski,[13] Beyond Horndeski and DHOST theories.[14–16] The presence of a non-minimal coupling between curvature and the scalar field will induce a running of the effective Planck mass which in turn modifies the friction term in (6) with

$\nu \equiv H^{-1} d \ln M(t)^2/dt$. The existence of a non-trivial kinetic mixing between the scalar and the metric alters the light-cone structure of GWs with respect to that of photons, leading to $c_T^2 \neq 1$. All other modifications in (6) equal zero for this class of models.

*Vector-tensor theories:* This class of theories includes Einstein-Aether (EA),[17] which is related to the low-energy limit of Horava-Lifsitz gravity.[18] The extension of the old (massive) Proca theory is described by the generalised Proca theories,[19,20] corresponding to the general covariant theory of a massive vector on a curved background. All of them lead to $c_T^2 \neq 1$, related to the non-trivial derivative interactions between the dynamical vector field and the metric.

*(Massive)Bi-gravity theories:* This family of theories describes the dynamics of two, non-minimally interacting coupled spin-two fields.[21] The non-minimal character of the interactions inevitably gives a mass to the standard (GR's) graviton, $\mu^2 \neq 0$. What is more, the second graviton sources the evolution of the standard one through the term on the right-hand side of (6), while all other terms are standard.

## 4. The link between gravitational slip and gravitational waves

### 4.1. *Non-standard propagation of gravitational waves implies a gravitational slip*

Let us assume that gravity is modified, according to its definition in Section 3. This means that the propagation of GWs will be different to GR in one way or the other according to (6). In Ref. 10 it was shown that a direct link exists between the propagation of GWs at any scale in the Universe and the existence of gravitational slip in the presence of perfect fluid matter at large scales: *For general extensions of GR introducing an extra spin-0, spin-1 and spin-2 two field respectively, the theory-space parameters that enter the coefficient of anisotropic stress, are exactly those that modify the propagation of GWs at any scale.*

In view of the anisotropy and the GW propagation equation, (1) and (6) respectively, this translates to the theory-space parameters entering $\sigma$ to be the same modifying the propagation of GWs. This link establishes a phenomenological bridge between two apparent disconnected regimes, connecting the physics of large-scale structures in the Universe with the propagation of GWs at any scale.

## 5. Dark energy phenomenology after GW170817

[c]Recently, the speed of GWs was measured from an observation of a merging neutron star and an electromagnetic counterpart, yielding the extremely strong constraint $|c_T/c - 1| \leq 10^{-15}$.[6] The observation has essentially ruled out all gravitational interactions which change the light cone structure of GWs – these are typically

[c]This section closely follows the results presented in Ref. 12.

terms which mix kinetically the scalar with the metric, beyond the simple conformal coupling (Brans–Dicke/$f(R)$). In light of the previously discussed link between GWs and the large-scale structure of the Universe, this new constraint can be translated to phenomenological predictions at large scales. Below we discuss the implications for the theory space of different families of models following the results and notation of the original work.[12]

*Scalar-tensor theories:* For Horndeski theories, the surviving theory space is described by

$$\mathcal{L} = \frac{f(\phi)}{2}R + K(X,\phi) - G(X,\phi)\Box\phi. \tag{7}$$

Notably, the only non-minimal coupling between scalar and curvature allowed is of the conformal type. The second term in (7) encompasses the k–essence/quintessence theory, while the third one the so–called kinetic gravity braiding.[22] At the linear level of scalar perturbations, the theory (7) is described by three functions of time, namely the so–called kineticity ($\alpha_K(t)$), braiding ($\alpha_B(t)$) and Planck mass rate ($\alpha_M(t)$). The first one measures the amount of fluctuations of the scalar field, the second the braiding between scalar and metric fluctuations, while the third the rate of change of the evolving Planck mass. In the quasi-static limit, it turns out that for the effective Newton's constant for the theory (7) reads as

$$Y = 1 + \frac{(\kappa+\alpha_M)^2}{2N}\,, \tag{8}$$

with $Y(t,k) \equiv -2k^2M_*^2\Psi/(a^2\rho_m\delta_m)$. Here, $\kappa \equiv \alpha_M + \alpha_B$ is sourced only through the third operator in (7) and is zero otherwise, since $\alpha_M = -\alpha_B$ for Brans-Dicke/$f(R)$ and $\alpha_M = 0 = \alpha_B$ in minimally-coupled models. $N$ is associated with the sound speed of the scalar and is required to be positive. For the gravitational slip and lensing parameter one finds that,

$$\eta - 1 = -\frac{2\alpha_M(\kappa+\alpha_M)}{2N+(\kappa+\alpha_M)^2}, \qquad \Sigma \equiv \frac{Y(\eta+1)}{2}. \tag{9}$$

From these relations it is straightforward to conclude for the *updated phenomenology* of scalar-tensor theories at large scales: The strength of clustering can be only made stronger than or equal to that in GR, that is, $Y \geq 1$. The lensing is the same as in GR ($\Sigma = 1$) as soon as the third operator in (7) (braiding term) is absent ($\alpha_B = -\alpha_M$ or $\alpha_B = 0 = \alpha_M$) – this is to be expected, as conformally- or minimally-coupled models do not affect null geodesics. However, as soon as the braiding term is included it is $\Sigma \neq 1$. As regards the slip parameter $\eta$, that can be on either side of its GR value, i.e $\eta < 1$ or $\eta \geq 1$.

*Vector theories:* Let us start with Einstein-Aether theories. The theory introduces three constant parameters $\beta_i$ associated with each operator in the action. The requirement of $c_T^2 = 1$ imposes that $\beta_1 = -\beta_3$, and at the same time eliminates the vector's anisotropic stress contribution, fixing $\eta = 1$. For the effective Newton's

coupling, it is $Y = (1+3\beta_2)/(1-\beta_1)$. Given that stability requires $\beta_2, \beta_1 > 0$, then $Y \geq 1$, as was the case in Horndeski models.

On the other hand, generalised Proca theories are characterised by six free functions at the level of the action, and similar to Horndeski theory, these combine to form a set of time-dependent functions that control scalar linear perturbations. The terms in the action that affect the GW propagation are those that kinetically mix the vector with the metric – It turns out that, for the associated free functions of these operators, $G_4 = \text{const}$ and $G_5 = 0$.[23,24] For this case the gravitational slip parameter becomes always equal to unity $\eta = 1$, while for the effective Newton's coupling (within the quasi–static approximation) we find that,

$$Y = 1 + \frac{-w_3 w_2^2}{N} \geq 1, \tag{10}$$

where we used the notation of Ref. 25 for the background functions $w_i(t)$. We further defined the quantity $N \equiv 2\mu_2/\phi + w_3 w_2^2$, which is related to the sound speed of the new degree of freedom, and the functions $w_2, w_3, \mu_2$ are defined in Ref. 25. Investigation of the explicit expression for $w_3$, along with the requirement of absence of ghost instabilities implies that $w_3 < 0$ and on similar grounds, the stability of scalar fluctuations requires that $N > 0$. From these, it follows for this case too that, $Y \geq 1$, that is, the clustering of matter can be made either equal or stronger compared to GR for these models ($Y \geq 1$), similar to the previous models.

As we discussed in Section 3, *massive/bi-gravity theories* predict that GWs propagate with $c_{\rm T}^2 = 1$, hence the recent measurement has no implications for their theory space.

*Consistency conditions:* Based on the above, and working within the small-scale limit of the quasi–static approximation (i.e $k \to \infty$), we can extract two new consistency conditions for *scalar–tensor theories* at large scales[d]:

1. $\eta < 1$ and $\Sigma = 1$: An imprint of a conformally coupled scalar field (Branks-Dicke/$f(R)$).

2. $\eta > 1$ and/or $\Sigma \neq 1$: An imprint of the presence of a kinetic braiding term in the action, in addition to the conformal coupling to curvature.

For *vector models*, an important corollary is the characteristic combination $\eta = 1$ and $Y \geq 1$.

It is interesting to remark that future direct observations of GWs can probe the running rate of the Planck mass in the surviving scalar-tensor sector as was argued in Refs. 27, 28, 29.

[d]For minimally-coupled scalar-tensor models it is easy to see that $\eta = 1$ and $Y > 1$. Notice that here, we do not consider the possibility of special configurations such as the scenario of No-Slip gravity of Ref. 26.

Table 1. Phenomenological implications for the gravitational slip and lensing parameters $\eta$ and $\Sigma$ respectively, after the recent measurement of the speed of GWs (see Section 5).[12] The first two rows describe the surviving theory space of Horndeski models, while the second and third one the Einstein–Aether and generalised Proca vector theories respectively. In all models, it now turns out that for the effective Newton's coupling at large scales it is $Y \geq 1$, i.e, clustering of matter is equal to or stronger than in GR.

| **Model** | $\eta \equiv \Phi/\Psi$ | $\Sigma \equiv (1/2)(\eta+1)Y$ | $Y$ |
|---|---|---|---|
| $M_p^2 R + K(\phi, X)$ | 1 | 1 | $Y \geq 1$ |
| $f(\phi)R + K(\phi, X)$ | $\leq 1$ | 1 | $Y \geq 1$ |
| $f(\phi)R + K(\phi, X) + G(X, \phi)\Box\phi$ | $< 1$ or $\geq 1$ | $\neq 1$ | $Y \geq 1$ |
| Einstein-Aether | $= 1$ | $\neq 1$ | $Y \geq 1$ |
| Generalised Proca | $= 1$ | $\neq 1$ | $Y \geq 1$ |

## 6. Summary

In this short paper we summarised key results from our previous works on the significance and the construction of model independent tests of gravity at cosmological scales, their intimate relation with the properties of GWs and the most recent phenomenological implications of GW observations for dark energy models. The recent measurement of the speed of GWs placed surprisingly stringent constraints on the theory space of gravity models, and allowed for robust cosmological predictions that can be accurately tested with the future Euclid satellite mission.[30] The new predictions and consistency conditions for the gravitational slip, weak lensing and effective gravity strength parameter summarised in Section 5, provide a challenge for the surviving models that will be tested with future surveys.

## Acknowledgments

I.D.S and I.S are supported by ESIF and MEYS (Project CoGraDS – CZ.02.1.01/0.0/0.0/15_003/0000437). This report is based on a talk given by I.D.S at the 15th Marcel Grossmann meeting. I.D.S would like to thank Luca Amendola and Joan Solá for the invitation.

## References

1. S. Weinberg, "*The Cosmological Constant Problem*," *Rev. Mod. Phys.*, vol. 61, pp. 1–23, 1989.
2. P. Bull *et al.*, "*Beyond ΛCDM: Problems, solutions, and the road ahead*," *Phys. Dark Univ.*, vol. 12, pp. 56–99, 2016.
3. A. Padilla, "*Lectures on the Cosmological Constant Problem*," 2015.
4. T. Clifton, P. G. Ferreira, A. Padilla, and C. Skordis, "*Modified Gravity and Cosmology*," *Phys. Rept.*, vol. 513, pp. 1–189, 2012.

5. L. Amendola and S. Tsujikawa, *Dark Energy: Theory and Observations*. Cambridge University Press, 2010.
6. B. P. Abbott *et al.*, "*GW170817: Observation of Gravitational Waves from a Binary Neutron Star Inspiral*," *Phys. Rev. Lett.*, vol. 119, no. 16, p. 161101, 2017.
7. B. P. Abbott *et al.*, "*Gravitational Waves and Gamma-Rays from a Binary Neutron Star Merger: GW170817 and GRB 170817A*," *ApJ*, vol. 848, no. 2, p. L13, 2017.
8. L. Amendola, M. Kunz, M. Motta, I. D. Saltas, and I. Sawicki, "*Observables and unobservables in dark energy cosmologies*," *Phys.Rev.*, vol. D87, p. 023501, 2013.
9. M. Motta, I. Sawicki, I. D. Saltas, L. Amendola, and M. Kunz, "*Probing Dark Energy through Scale Dependence*," *Phys.Rev.*, vol. D88, p. 124035, 2013.
10. I. D. Saltas, I. Sawicki, L. Amendola, and M. Kunz, "*Anisotropic Stress as a Signature of Nonstandard Propagation of Gravitational Waves*," *Phys. Rev. Lett.*, vol. 113, no. 19, p. 191101, 2014.
11. I. Sawicki, I. D. Saltas, M. Motta, L. Amendola, and M. Kunz, "*Nonstandard gravitational waves imply gravitational slip: On the difficulty of partially hiding new gravitational degrees of freedom*," *Phys. Rev.*, vol. D95, no. 8, p. 083520, 2017.
12. L. Amendola, M. Kunz, I. D. Saltas, and I. Sawicki, "*The fate of large-scale structure in modified gravity after GW170817 and GRB170817A*," 2017.
13. G. W. Horndeski, "*Second-order scalar-tensor field equations in a four-dimensional space*," *Int. J. Theor. Phys.*, vol. 10, pp. 363–384, 1974.
14. J. Gleyzes, D. Langlois, F. Piazza, and F. Vernizzi, "*Healthy theories beyond Horndeski*," *Phys. Rev. Lett.*, vol. 114, no. 21, p. 211101, 2015.
15. J. Ben Achour, M. Crisostomi, K. Koyama, D. Langlois, K. Noui, and G. Tasinato, "*Degenerate higher order scalar-tensor theories beyond Horndeski up to cubic order*," *JHEP*, vol. 12, p. 100, 2016.
16. D. Langlois, M. Mancarella, K. Noui, and F. Vernizzi, "*Effective Description of Higher-Order Scalar-Tensor Theories*," *JCAP*, vol. 1705, no. 05, p. 033, 2017.
17. T. Jacobson and D. Mattingly, "*Gravity with a dynamical preferred frame*," *Phys.Rev.*, vol. D64, p. 024028, 2001.
18. D. Blas, O. Pujolas, and S. Sibiryakov, "*Models of non-relativistic quantum gravity: the good, the bad and the healthy*," 2010.
19. G. Tasinato, "*Cosmic Acceleration from Abelian Symmetry Breaking*," *JHEP*, vol. 04, p. 067, 2014.
20. L. Heisenberg, "*Generalization of the Proca Action*," *JCAP*, vol. 1405, p. 015, 2014.
21. S. Hassan and R. A. Rosen, "*Bimetric Gravity from Ghost-free Massive Gravity*," *JHEP*, vol. 1202, p. 126, 2012.
22. C. Deffayet, O. Pujolas, I. Sawicki, and A. Vikman, "*Imperfect Dark Energy from Kinetic Gravity Braiding*," *JCAP*, vol. 1010, p. 026, 2010.

23. T. Baker, E. Bellini, P. G. Ferreira, M. Lagos, J. Noller, and I. Sawicki, "*Strong constraints on cosmological gravity from GW170817 and GRB 170817A*," *Phys. Rev. Lett.*, vol. 119, no. 25, p. 251301, 2017.
24. J. M. Ezquiaga and M. Zumalacárregui, "Dark Energy After GW170817: Dead Ends and the Road Ahead," *Phys. Rev. Lett.*, vol. 119, no. 25, p. 251304, 2017.
25. A. De Felice, L. Heisenberg, R. Kase, S. Mukohyama, S. Tsujikawa, and Y.-l. Zhang, "*Effective gravitational couplings for cosmological perturbations in generalized Proca theories*," *Phys. Rev.*, vol. D94, no. 4, p. 044024, 2016.
26. E. V. Linder, "*No Slip Gravity*," *JCAP*, vol. 1803, no. 03, p. 005, 2018.
27. E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "Gravitational-wave luminosity distance in modified gravity theories," *Phys. Rev.*, vol. D97, no. 10, p. 104066, 2018.
28. L. Amendola, I. Sawicki, M. Kunz, and I. D. Saltas, "*Direct detection of gravitational waves can measure the time variation of the Planck mass*," *JCAP*, vol. 1808, no. 08, p. 030, 2018.
29. E. Belgacem, Y. Dirian, S. Foffa, and M. Maggiore, "Modified gravitational-wave propagation and standard sirens," *Phys. Rev.*, vol. D98, no. 2, p. 023510, 2018.
30. L. e. Amendola, "*Cosmology and fundamental physics with the Euclid satellite*," *Living Rev.Rel.*, vol. 16, p. 6, 2013.

# Linking multipole vectors and pseudo entropies for CMB analysis

Marvin Pinkwart[1*], Peter Schupp[1] and Dominik J. Schwarz[2]

[1] *Department of Physics and Earth Sciences, Jacobs University, 28759 Bremen, Germany*
[2] *Fakultät für Physik, Universität Bielefeld, 33501 Bielefeld, Germany*
[*] *E-mail: m.pinkwart@jacobs-university.de*

Multipole vectors and pseudo entropies provide powerful tools for a numerically fast and vivid investigation of possible statistically anisotropic, resp. non-Gaussian signs in CMB temperature fluctuations. After reviewing and linking these two conceptions we compare their application to data analysis using the Planck 2015 NILC full sky map.

*Keywords*: Cosmic Microwave background; data analysis; entropy.

## 1. Motivation

Over the past 50 years the cosmic microwave background (CMB) has become the main source of information about the early universe. It displays a nearly isotropic black body with an order $10^{-3}$ dipole modulation due to peculiar motion and order $10^{-5}$ anisotropies which depict energy density fluctuations on the last scattering surface (LSS). Motivated by the simplest inflationary scenarios and linear perturbation theory these anisotropies are commonly assumed to be statistically isotropic and Gaussian, which means that if one decomposes the relative temperature fluctuations on the celestial sphere into spherical harmonics

$$\delta T(\theta,\phi) = \sum_{l=1}^{\infty} \sum_{m=-l}^{l} a_{lm} Y_{lm}(\theta,\phi), \tag{1}$$

the joint distribution of the spherical harmonic coefficients $a_{lm}$ for a given multipole number $l$ should be isotropic

$$p(\hat{R}\{a_{lm}\}) = p(\{a_{lm}\}), \tag{2}$$

where the rotation acts on the $a_{lm}$ via Wigner symbols, and Gaussian

$$p(\{a_{lm}\}) = \mathcal{N} \exp\left(-\frac{1}{2} \sum_{mm'} a^*_{lm'} \mathcal{D}_{l,mm'} a_{lm}\right). \tag{3}$$

Since the one-point function of the $a_{lm}$ vanishes, Gaussianity implies that all information about $\delta T$ is encoded in the two-point functions $\langle a_{lm} a_{lm'} \rangle = \mathcal{D}^{-1}_{l,mm'}$ and statistical isotropy further reduces the degrees of freedom to one per multipole number by diagonalizing the covariance $\mathcal{D}^{-1}_{l,mm'} = C_l \delta_{mm'}$. This yields the power spectrum which is commonly used as the main source of information about the CMB. Hence, the two basic assumptions reduce the real degrees of freedom per multipole from $2l + 1$ to one. If one or both of the assumptions are relaxed, one needs to find additional data compressing measures which are easier and can be interpreted more directly and vividly than the $a_{lm}$. In this contribution we describe and link two approaches to find such measures which are furthermore complementary to

the power spectrum in information content. The first approach concerns multipole vectors whose use is motivated by large scale anisotropy anomalies that have been found recently and the second approach introduces pseudo entropy measures on the space of spin states associated to temperature fluctuation fields which provide a non brute-force method to investigate non-Gaussianities and are motivated by the analogy to statistical mechanics where the large number of microscopic degrees of freedom is essentially captured by a few macroscopic quantities such as entropy.

## 2. Multipole vectors and pseudo entropies

Multipole vectors (MPVs) have been introduced to CMB data analysis in Ref. 1. and were used to identify one of the three basic "atoms" of large scale isotropy anomalies. It was found that the quadrupole and octupole are correlated, that the octupole is unusually planar and that both quadrupole and octupole are unusually orthogonal to the Ecliptic Plane and aligned with the Solar Dipole, see the review in Ref. 2. Contrary to the behavior of spherical harmonic coefficients, MPVs rotate rigidly with the celestial sphere and therefore put themselves forward as the correct basic constituents of measures of statistical anisotropy. Let $\vec{e} = (e^1, e^2, e^3)^T = \vec{e}(\theta, \phi)$ be a unit vector, then the basic idea is to observe that the $l$-multipole-part of $\delta T$ is a homogeneous polynomial of degree $l$ and to decompose it according to

$$(\delta T)_l\,(\vec{e}(\theta,\phi)) = A^{(l)}\left[v_{i_1}^{(l,1)}\cdots v_{i_l}^{(l,l)}\right]\left[e^{i_1}\cdots e^{i_l}\right] = B^{(l)}\prod_{i=1}^{l}\vec{e}\cdot\vec{v}^{(l,i)} + F^{(l)}(\theta,\phi), \quad (4)$$

where $A^{(l)}$ and $B^{(l)}$ are real constants, $F^{(l)}$ is a polynomial of defree $l-2$ and Einstein's summation convention was used.[3] The $\vec{v}^{(i,l)}$ are the MPVs and contain all information content of the temperature map up to a constant amplitude. Since the power spectrum is contained within this amplitude, the information content of the MPVs lies in the complement of the space of $C_l$-information. Hence, MPVs provide an interesting tool to test if this subspace is non-empty, i.e. if the temperature map is statistically anisotropic. In fact, as shown in Ref. 4 the statistical behavior of MPVs is the same for the larger class of all completely random sets of $a_{lm}$, i.e. $p(\{a_{lm}\}) = f(\sum_m |a_{lm}|^2)$, which contains the standard ΛCDM case but also some isotropic non-Gaussian maps. Note, that the $l$-multipole part of $\delta T$ can also be viewed as a normalized spin-$l$ state

$$|\delta T\rangle_l = \sum_m \widetilde{a}_{lm}|lm\rangle \quad (5)$$

up to an amplitude which contains the information about $C_l$, where the $a_{lm}$ have been normalized to $\widetilde{a}_{lm}$. Then the contraction of Bloch coherent states with such a temperature state and stereographic projection yield a complex polynomial

$${}_l\langle\Omega|\delta T\rangle_l \sim \sum_m \sqrt{\binom{2l}{l+m}}\widetilde{a}_{lm}z^{l+m} \sim \prod_{i=1}^{l}(z-z^{(l,i)})\left(z+\frac{1}{(z^{(l,i)})^*}\right), \quad (6)$$

whose $2l$ zeros are the stereographic projections $z^{(l,i)}$ of the MPVs $\vec{v}^{(l,i)}$ and their antipodes.[5] Bloch coherent states are given by the rotation of the highest spin state $|\Omega\rangle_l = \hat{R}(\Omega)|l,l\rangle$ by angles $\Omega = (\theta, \phi)$. For a detailed review on MPV constructions see Ref. 6. Now one can define the most classical quantum entropy which is named Wehrl entropy

$$S_l^{\mathrm{W}} = S^{\mathrm{W}}(|\delta T\rangle_l) = -\frac{(2l+1)}{4\pi}\int \mathrm{d}\Omega\, |{}_l\langle\Omega|\delta T\rangle_l|^2 \log\left(|{}_l\langle\Omega|\delta T\rangle_l|^2\right), \tag{7}$$

and which is used for CMB data analysis as a measure of randomness of temperature fluctuations.[7] The disadvantage of the Wehrl entropy is its numerically expensive form and hence one seeks measures that approximate the Wehrl entropy reasonably well but are much easier to compute. Since the temperature map describes a pure state, the usual von Neumann entropy is trivial. The basic idea to overcome this problem is to apply a quantum channel $\Phi$, which respects isotropy, to the pure temperature state $\rho_l = |\delta T\rangle_l\langle\delta T|$ to obtain a mixed density $\rho_l^{\Phi}$ and afterwards apply the von Neumann entropy to get a mixed pseudo entropy

$$S_l^{\Phi}(\delta T) = -\mathrm{Tr}\left(\rho_l^{\Phi}\log(\rho_l^{\Phi})\right). \tag{8}$$

The term "pseudo" shall indicate that the resulting quantity loses some properties entropies usually possess. From now on we drop this adjunct. In Ref. 8 we introduced two convenient choices of rotationally symmetric quantum channels:

- The $j$-projection entropy uses the mixed density

$$\rho_l^{\mathrm{proj},(j)} = \Phi^{\mathrm{proj},(j)}(\rho_l) = \frac{2l+1}{2(l+j)+1}\hat{P}_{l+j}\left(\rho_l \otimes 1_j\right)\hat{P}_{l+j}, \tag{9}$$

  where $\hat{P}_{l+j}$ denotes the projection to the $[l+j]$-subspace of $[l]\otimes[j]$. It converges to $S_l^{\mathrm{W}} + \log\left(\frac{2(l+j)+1}{2l+1}\right)$ for $j\to\infty$, but not uniformly. The additional term does not depend on the data, but just on the numbers $l$ and $j$ and henceforth in data analysis this term does not need to be taken into account. Therefore one may say that in the realm of data analysis the projection entropy converges to the Wehrl entropy.
- The angular entropy uses the mixed density

$$\rho_l^{\mathrm{ang}} = \Phi^{\mathrm{ang}}(\rho_l) = \frac{1}{l(l+1)}\sum_{i=1}^{3}\hat{L}_i\rho_l\hat{L}_i^{\dagger}, \tag{10}$$

  where the $L_i$ are the angular momentum operators acting on spin space. For $l=2$ it measures the repulsion of MPVs and shows the exact same behavior as the Wehrl entropy once an overall shift and dilation in absolute value has been accounted for. Let $\epsilon$ be the squared chordal distance between the two MPVs on

a sphere of radius 1/2, then the angular and Wehrl entropy can be expressed as

$$S_2^{\mathrm{ang}}(\epsilon) = -\frac{c(\epsilon)}{2}\left((1-\epsilon)^2\log(1-\epsilon)^2 + \epsilon^2\log\left(\epsilon^2\right)\right) - \log\left(\frac{c(\epsilon)}{2}\right) \tag{11}$$

$$S_2^{\mathrm{W}}(\epsilon) = c(\epsilon) - \log(c(\epsilon)) + \frac{32}{15} - \log(6), \tag{12}$$

where $c(\epsilon) := (1-(1-\epsilon))^{-1}$. Motivated by this result, the convergence of the projection to the Wehrl entropy and the numerically similar behavior, we propose that the similarity of Wehrl and angular entropy as well as the interpretation as a measure of repulsion between MPVs should be valid on all scales, at least approximately. Since repulsive behavior of zeros of random polynomials is connected to the degree of Gaussianity of the polynomial, we consider the angular entropy as a simple measure for deviations from Gaussianity, at least as long the deviation causes an exit from the space of completely random distributions. Furthermore, we showed that the probability distribution of the angular entropy is highly sensitive to "small" deviations from Gaussianity, in the sense that if one replaces $i$ MPVs from a Gaussian map by uniformly distributed unit vectors — which constitutes a deviation from Gaussianity without violating isotropy — the shift in distribution is sizable and the largest portion of shift is obtained from the first replaced vector.

Since all entropy measures use the normalized temperature state, their information content is complementary to the $C_l$, which is also the case for MPVs. Using the latter to build vivid scalar quantities yields additional measures, which we interpret as measures of statistical anisotropy. In Ref. 6 we used the following statistics:

- The internal/inner statistics

$$S^{||}(l) = \frac{2}{l(l-1)}\sum_{i<j}|\vec{v}^{(l,i)}\cdot\vec{v}^{(l,j)}| \tag{13}$$

$$S^{v}(l) = \frac{6}{l(l-1)(l-2)}\sum_{i<j<k}\left|\left(\vec{v}^{(l,i)}\times\vec{v}^{(l,j)}\right)\cdot\vec{v}^{(l,k)}\right| \tag{14}$$

  measure internal alignment and non-planarity of multipoles. In structure and interpretation these statistics share some similarities with the entropy measures.
- The external/outer statistics

$$S^{||}_{\vec{D}}(l) = \frac{2}{l(l-1)}\sum_{i}|\vec{v}^{(l,i)}\cdot\vec{D}| \tag{15}$$

$$S^{v}_{\vec{D}}(l) = \frac{6}{l(l-1)(l-2)}\sum_{i<j<k}\left|\left(\vec{v}^{(l,i)}\times\vec{v}^{(l,j)}\right)\cdot\vec{D}\right|, \tag{16}$$

  where $\vec{D}$ denotes some given outer physically motivated direction, measure outer directional or planary influences on the data.

## 3. Application to data analysis

Both methods have been applied to Planck 2015 and Planck 2018 full sky maps in order to compare them to statistically isotropic and Gaussian random maps and investigate if any signs of possible violations appear on multipole numbers up to 50 (MPVs), resp. 1000 (entropies).[6,8] In Fig. 1 we compare $S^{\text{ang}}$ (8)(10), -$S^{||}$(13), and $S^{||}_{\vec{D}}$ (15) with the Solar Dipole as a given physical direction $\vec{D}$ using the NILC 2015 full sky map. We use the negative of $S^{||}$ because it allows for an interpretation as a measure of non-alignment which is related to the interpretation of $S^{\text{ang}}$ as a measure of repulsion. One observes that the general behavior of both $S^{\text{ang}}$ and $S^{||}$ as a function of $l$ approximately agrees while unlikely multipoles are much more pronounced in $S^{\text{ang}}$. $S^{||}_{\vec{D}}$ clearly deviates from the others and the clustering of low-likelihood multipoles at $l \in [2, 5]$ reproduces the corresponding atom of CMB anomalies. Note that for the sake of clarity $1/3\sigma$-boundaries are not shown due to their different relative distances among the statistics, but note that $1\sigma \neq \frac{1}{2}(2\sigma)$.

Eventually we summarize further results that were obtained in Refs. 6, 8. With the MPV statistics the behavior of three full sky maps shows up to be similar even without applying masks, and no global anisotropies could be identified on $l \in [2, 50]$. After masking, the behavior of SEVEM approaches that of the other maps. The direction that is visible in the data most is the Solar Dipole, which appears in the large scale anomalies and via a clustering of $1\sigma$ (anti-)alignments around $l = 20$. Using the entropies the findings on masking, non-masked SEVEM, similarity between the other maps and overall normal behavior agree with those obtained with MPV statistics, but the entropy shows above $2/3\sigma$ unusual values at $l = 5/28$, which are curious because their values are higher than expected. It is unclear which kind of physical temperature map can result in too high entropies. The large-scale anomalies are visible when applying the range angular entropy, which takes into account ranges and possible correlations of multipoles. A comparison between 2015

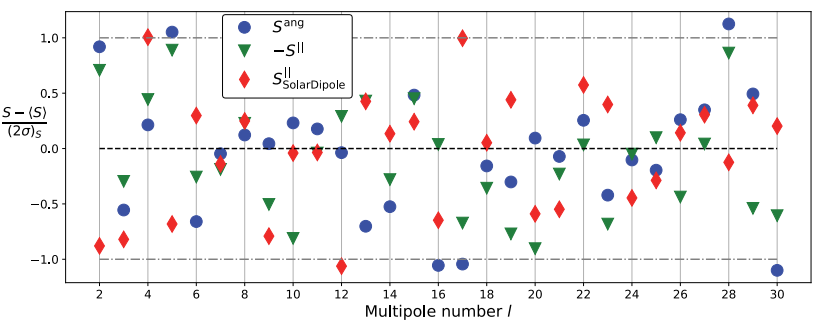



Fig. 1. Comparison of MPV statistics and angular entropy via the deviation of NILC 2015 full sky data from isotropic, Gaussian expectation in units of $2\sigma$-deviation calculated with $10^4$ (entropy) and $10^3$ (MPVs) ensembles on the range $l \in [2, 30]$.

and 2018 data revealed that the change in NILC and SMICA is negligible while unmasked SEVEM has been enhanced. The angular entropy can also be used on very small angular scales with moderate computational expense.

## Conclusions

MPV statistics yield vivid measures for identifying possible statistical anisotropies, possibly induced by physical influences measured via effects of outer directions. On the other hand our entropy methods yield fast measures for non-Gaussianity which allow for investigation down to small angular scales $l \geq 1000$. Both methods are mathematically related, see Eqs. (6), (11) and (12), and yield information content that is complementary to that of $C_l$. Note that in the future effects of inhomogeneous noise and masking should be taken into account more thoroughly. Our methods need full sky data but using e.g. spherical cap harmonics or other local methods one could make more sense of the statistics when applying a mask. An interesting future task would also be to redo the analysis using full focal plane simulations. Finally an efficient measure for unlikeliness when picking out ranges of random size needs to be worked out.

## Acknowledgments

We acknowledge financial support by the DFG RTG 1620 Models of Gravity.

## References

1. C. J. Copi, D. Huterer and G. D. Starkman, Multipole vectors - A New representation of the CMB sky and evidence for statistical anisotropy or non-Gaussianity at $2 \leq l \leq 8$, *Phys. Rev.* **D70**, p. 043515 (2004).
2. D. J. Schwarz, C. J. Copi, D. Huterer and G. D. Starkman, CMB Anomalies after Planck, *Class. Quant. Grav.* **33**, p. 184001 (2016).
3. G. Katz and J. Weeks, Polynomial interpretation of multipole vectors, *Phys. Rev.* **D70**, p. 063527 (2004).
4. M. R. Dennis and K. Land, Probability Density of the Multipole Vectors for a Gaussian Cosmic Microwave Background, *Mon. Not. Roy. Astron. Soc.* **383**, p. 424 (2008).
5. P. Schupp, On Lieb's Conjecture for the Wehrl Entropy of Bloch Coherent States, *Communications in Mathematical Physics* **207**, 481 (1999).
6. M. Pinkwart and D. J. Schwarz, Multipole vectors of completely random microwave skies for $l \leq 50$, *Phys. Rev.* **D98**, p. 083536 (2018).
7. R. C. Helling, P. Schupp and T. Tesileanu, CMB statistical anisotropy, multipole vectors, and the influence of the dipole, *Phys. Rev. D* **74**, p. 063004 (Sep 2006).
8. M. Minkov, M. Pinkwart and P. Schupp, Entropy methods for CMB analysis of anisotropy and non-Gaussianity, *astro-ph.CO/1809.08779*.

# Preliminary analysis of the optical system of the LSPE-STRIP instrument

Sabrina Realini* and Cristian Franceschet

*Physics Department, Universitá degli Studi di Milano,*
*Via Celoria 16, Milano, Italy*
**E-mail: sabrina.realini@unimi.it*

Maura Sandri and Fabrizio Villa

*INAF OAS Bologna,*
*Via Gobetti 101, Bologna, Italy*

We present a preliminary analysis of the optical system of the STRIP instrument of the Large Scale Polarization (LSPE) experiment, which aims at polarization measurements of the Cosmic Microwave Background on large angular scales. STRIP will observe approximately 25% of the Northern sky from the Observatorio del Teide in Tenerife, using an array of forty-nine coherent polarimeters at 43 GHz (Q-band), coupled to a 1.5 m fully rotating crossed-Dragone telescope. An additional frequency channel with six-elements at 95 GHz (W-band) will be exploited as an atmospheric monitor.

Non-idealities in the optical system may introduce limitations in achieving high precision measurements, if not well understood and controlled. For this reason, we studied the optical design of STRIP, its characteristics in terms of performance on angular resolution, sidelobes, main beam symmetry, polarization purity and feedhorns orientation, by means of electromagnetic simulations.

*Keywords*: CMB polarization, ground-based telescope, microwave optical simulations.

## 1. Introduction

The characterization of the polarized signal of the Cosmic Microwave Background (CMB) is, nowadays, one of the most important branches of observational cosmology. A direct measurement of the slight trace left by primordial gravitational waves on the CMB polarization pattern (the so called B-modes) would shed light on the first seconds of our universe.

The amplitude of the B-mode component is parametrized by the tensor-to-scalar ratio $r$ and, to date, only an upper bound at $r < 0.07$ at 95% C.L. has been set by the joint BICEP2/Keck Array and Planck analysis.[1]

Besides a strict control of the instrumental noise and systematic effects, a deep knowledge of astrophysical foregrounds is a key element for polarization analysis. In fact, both synchrotron and thermal dust emissions are partially linearly polarized. Since there is no region of the sky, at any frequency, where Galactic foreground contamination lies below a B-mode signal, it is important for all future experiments to monitor the foreground emissions with several frequency channels.

In this context, we find the STRIP instrument of the "Large Scale Polarization Explorer" (LSPE), which is the Italian Space Agency's upcoming experiment for the observation of the CMB polarization on large angular scales.

LSPE[2] combines ground-based and balloon-borne observations and aims at improving the limit on the tensor-to-scalar ratio down to $r = 0.03$ at 99.7% confidence level. A second target is to produce wide maps of foreground polarization generated in our Galaxy by synchrotron and interstellar dust emissions in a range of frequencies between 40 and 250 GHz, observing a 25% fraction of the sky in the Northern hemisphere.

## 2. The STRIP Instrument

STRIP is the ground-based instrument of the LSPE experiment that will observe the sky for two years from the "Observatorio del Teide" in Tenerife, starting in early 2020.[3] STRIP will scan the sky with a continuous azimuthal spin at 1 r.p.m., while maintaining the telescope boresight at a fixed elevation angle (approximately 20° from the Zenith at Tenerife latitude).

STRIP consists of an array of forty-nine coherent polarimeters, arranged in seven modules, operating in a 18% frequency band centered at 43 GHz (Q-band) and six polarimeters operating in a frequency band centered at 95 GHz (W-band) for atmosphere monitoring. Receivers are cooled down to 20 K in a cryostat placed in the focal surface of a dual reflector telescope, with a 1.5 m primary mirror.

The STRIP goal sensitivity on Stokes parameters is $\delta Q(U)_{deg} \leq 1.2\,\mu$K for the Q-band and $\delta Q(U)_{deg} \leq 4.5\,\mu$K for the W-band. Each Q-band receiver includes a circular corrugated feedhorn, a polarizer, an orthomode transducer (OMT) and a polarimeter module, which enables the detection of the Q and U Stokes parameters.

The W-band channel consist of a corrugated feedhorn and a septum polarizer, matched to a correlation polarimeter module. This channel will be used especially to track the amount of water vapor, which is the primary cause of atmospheric opacity. Characterizing the Tenerife observing site at W-band, i.e., near the foregrounds minimum, is of high interest also for potential future CMB experiments covering the Northern hemisphere.

## 3. The STRIP Optics

### 3.1. *The telescope*

The STRIP dual-reflector telescope is based on a Dragonian cross-fed design, with a projected aperture of 1.5 m. The crossed Dragone design is characterized by a parabolic primary mirror and hyperbolic secondary arranged to provide a large field of view and to aberrations at the center of the field. Moreover, the Dragonian configuration is the best in terms of polarization purity and symmetry over a wide focal region.[4]

The primary mirror of the STRIP telescope has an offset parabolic shape, with a 1500 mm aperture. The secondary mirror is an offset concave hyperboloid, with a 1719.76 mm × 1658.7 mm wide elliptical rim. The telescope provides an angular resolution of $\sim 20'$ in the Q-band and $\sim 10'$ in the W-band. The entire dual-reflector system has an equivalent focal length of 2700 mm, resulting in a F#$\sim$ 1.8.

The whole array of feedhorns is placed in the focal region of the telescope, ensuring no obstruction of the field of view. All feedhorns modules have been placed according to the telescope focal surface (see Fig. 1) and oriented towards the main reflector center, so that an optimum spillover is obtained while guaranteeing low level of cross–polarization contamination.

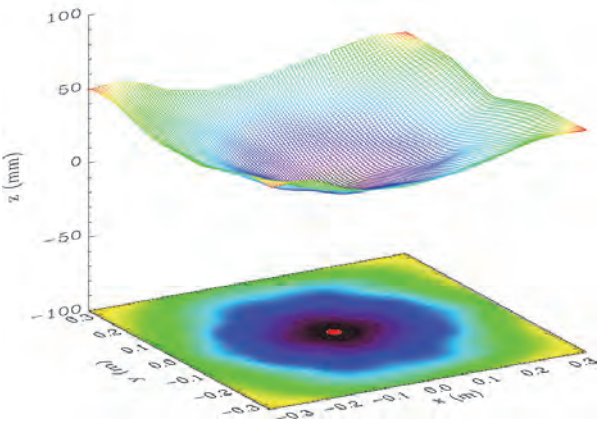



Fig. 1. Focal surface of the STRIP dual-reflector telescope.

### 3.2. *Mount and shielding structures*

The telescope is surrounded by a comoving baffle made of seven aluminum plates in order to reduce the contamination due to the sidelobes. This optical assembly is hold by a mount, which allows the STRIP optics and receivers to point at any direction in the sky. The telescope mount, already built for the CLOVER experiment[5], is a three-axes system that can fully rotate in azimuth and reach elevations from 0° to 89.5°. The telescope elevation angle will be fixed at approximately 20°.

### 3.3. *Electromagnetic model*

The STRIP optics have been modeled with the software GRASP[a]. The model includes the two nominal reflectors, forty-nine Q-band feedhorns and six W-band feedhorns, and the shielding structures, which redistribute the power that is radiated by the feedhorns and is not reflected by the telescope (see Fig. 2).

Main beams and sidelobes have been simulated and analyzed separately, because of the different sampling in the $\theta$ and $\phi$ coordinates and the different simulation

[a]The GRASP software is developed by TICRA (Copenhagen, DK) for analysing general reflector antennas.

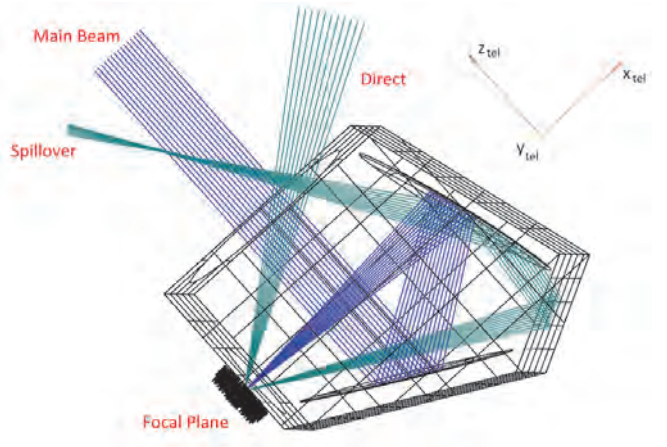


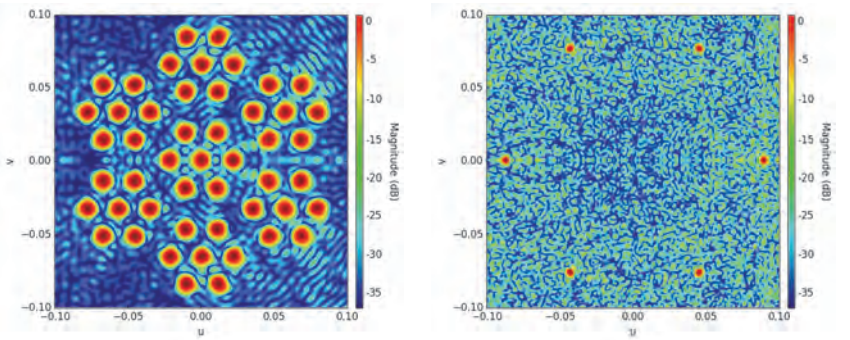

Fig. 2. Telescope and shielding structure model with the ray-tracing of some MrGTD contributions in the symmetry plane.

technique. Main beams have been simulated using the Physical Optics (PO) method, which computes the induced currents on the reflecting surfaces in order to obtain the total field. This method is the most accurate for predicting beams; however, in the sidelobes region, the PO integrand oscillates and a finer integration grid is required, leading to an increasing computation time. For this reason, the sidelobes have been computed using the Multi-Reflector Geometrical Theory of Diffraction (MrGTD), which represents a suitable method for predicting the full-sky radiation pattern of complex mm-wavelength optical systems, because it computes the scattered field from the reflectors performing a backward ray-tracing.

## 4. Performance Analysis

The radiation pattern of the feedhorns in the STRIP optics have been simulated in their main beam coordinate system at 43 and 95 GHz, exciting each feedhorn individually with a linearly polarized signal. Fig. 3(a) shows the footprint of the forty-nine beams at 43 GHz in the sky, while Fig. 3(b) shows the footprint of the six beams at 95 GHz.

Given the telescope configuration and the feedhorn off-axis location on the focal surface, the main beams are not perfectly Gaussian, so that they cannot be described by a single parameter. For a complete characterization of the main beams, several descriptive parameters have been evaluated: the angular resolution (FWHM), the ellipticity (e), the main beam directivity (D), the cross-polar discrimination factor (XPD) and the spillover. For the Q-band, the angular resolution is around $21'$ and the directivity is around 54.7 dBi for all the Q-band channels, while the XPD ranges




(a) (b)

Fig. 3. Q-band channel (a) and W-band channel (b) footprints on the sky as seen by an observer looking towards the telescope along its optical axis. The origin of the uv-coordinate system is at the center of the focal plane. The z-axis is along the line-of-sight and points towards the observer.

from 40.8 to 44.5 dB. For the W-band, the angular resolution is around $9.5'$ and the directivity is around 61.4 dBi.

We computed also the sidelobes of the STRIP telescope up to the $3^{rd}$ order of interaction (i.e. reflection or diffraction), for some representative channels. They are unevenly distributed and concentrated mainly in two areas, namely the "direct contribution" and the "spillover" (see Fig. 4). The direct contribution is generated by the rays entering the feedhorns without any interaction with the reflectors; its shape and power level are given by the feedhorn radiation pattern pointing at about 60° from the line of sight of the telescope. The spillover is primarily due to rays reflected by one of the shields and then reflected by the sub-reflector. The radiation patterns of both the Q- and W-band corrugated feedhorns have been measured in an anechoic chamber to check the compliance with the requirements and verify the reliability of the GRASP model.

## 5. Conclusions

In this paper, we presented a preliminary analysis of the optical system of the STRIP instrument of the LSPE experiment. All the main beams have been computed and we retrieved the main descriptive parameters. Concerning the sidelobes, we analyzed only four representative channels, which revealed that particular attention should be paid to multiple reflections inside the shielding structure.

The analysis will be repeated evaluating the response in the bandwidth, because accurate predictions of the telescope response are essential for an in-depth knowledge of the whole instrument response and for the development of the data reduction pipeline, where also instrumental systematic effects are taken into account.

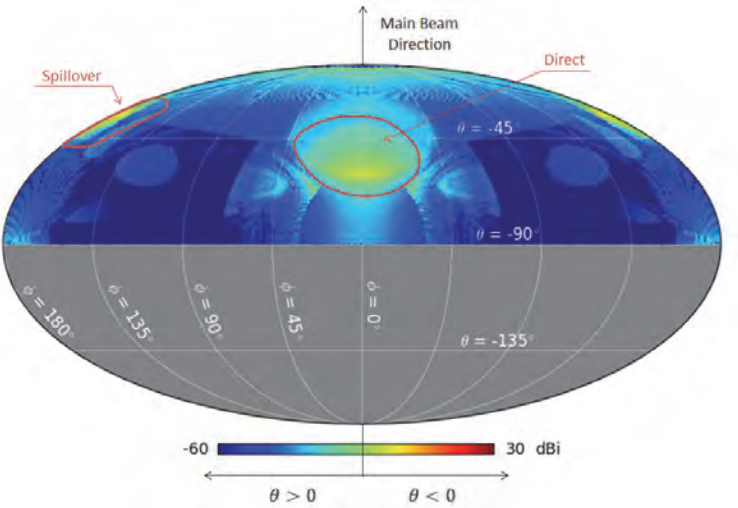



Fig. 4. Far sidelobes at 43 GHz for the feedhorn placed in the telescope focus. The main beam points to the top of the map ($\theta = 0°$) and peaks at 54.76 dBi. The direct contribution is due to the feedhorn sidelobe and peaks at about -5 dBi, while the spillover is due to a double reflection inside the shielding structure and peaks at about -0.8 dBi.

## Acknowledgments

We gratefully acknowledge support from ASI through contract I-022-11-0 “LSPE” and from the Istituto Nazionale di Fisica Nucleare (INFN).

## References

1. BICEP2/Keck Collaboration, Planck Collaboration, P. A. R. Ade, N. Aghanim, Z. Ahmed, R. W. Aikin and et al., Joint Analysis of BICEP2/Keck Array and Planck Data, *Physical Review Letters* **114**, p. 101301 (March 2015).
2. The LSPE collaboration, S. Aiola, G. Amico, P. Battaglia, E. Battistelli, A. Baù and et al., The Large-Scale Polarization Explorer (LSPE), *ArXiv e-prints*, (August 2012).
3. C. Franceschet, S. Realini, A. Mennella and et al., The STRIP instrument of the Large Scale Polarization Explorer: microwave eyes to map the Galactic polarized foregrounds, *Proc. SPIE,* **10708,** (2018).
4. H. Tran, A. Lee, S. Hanany, M. Milligan and T. Renbarger, Comparison of the crossed and the Gregorian Mizuguchi-Dragone for wide-field millimeter-wave astronomy, *Applied Optics* **47**, 103 (January 2008).
5. P. K. Grimes, P. A. R. Ade, M. D. Audley, C. Baines, R. A. Battye, M. L. Brown and et al., Clover - Measuring the Cosmic Microwave Background B-mode Polarization, in *Twentieth International Symposium on Space Terahertz Technology*, April 2009.

# Reconstructing the CMB *B*-mode power spectrum using the ABS method

L. Santos[1] and W. Zhao[2]

[1] *Center for Gravitation and Cosmology, College of Physical Science and Technology, Yangzhou University, Yangzhou 225009, China*

[2] *CAS Key Laboratory for Researches in Galaxies and Cosmology/School of Astronomy and Space Science, Department of Astronomy, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230026, China*

[*] *E-mail: larissa@ustc.edu.cn*

One of the main goals of modern cosmology is to probe inflationary theories by looking on the imprint of primordial gravitational waves in the cosmic microwave background (CMB) polarization field. Future CMB experiments face the great challenge to search for this primordial B-mode signal. However, the CMB sky is also filled with secondary B-modes, including CMB lensing and astrophysical foregrounds. Extracting the CMB B-mode polarization from astrophysical contaminations is a primordial task towards detection of the primordial signal. We use the analytical method of blind separation (ABS) proposed by Zhang, P., et al. (2019) to reconstruct the CMB B-mode power spectrum in the presence of foregrounds and white noise considering a full sky analysis for $r = 0$.

*Keywords*: CMB, foregrounds, statistics.

## 1. Introduction

The temperature anisotropies of the cosmic microwave background (CMB) radiation have been extensively addressed in the literature during the past three decades, being well characterized by various experiments such as WMAP [1, 2] and Planck satellites [3]). It is well known that these surveys detect not only the CMB signal itself, but a mixture of different emissions, including astrophysical foregrounds. One of the challenges in analyzing the CMB data has been the component separation, a fundamental step towards the efficient recovery of the CMB power spectra, thus the estimation of the cosmological parameters. Ongoing experiments targeting the polarized signal already face the difficulties of extracting the $E$-mode [4, 5] and specially the $B$-mode [6] polarization field from foregrounds.

Focusing on the CMB $B$-mode field that can probe inflation on large angular scales, we use the analytical method of Blind Separation (ABS) developed recently by [7] to extract the CMB simulated signal from foregrounds considering a multi-frequency future experiment. Contrary to other "blind" methods, such as the internal linear combination (ILC) [8–15] or the the Spectral Matching Independent Component Analysis (SMICA) [16–19], the ABS do not involve heavy computation. Moreover, it consists of the measured cross band power between different frequency bands, which is derived analytically in order to avoid multiple parameter fitting. Moreover, the ABS method was already successfully tested against

simulated temperature Planck maps [20]. Here, we start the study of the ABS performance on the CMB polarized signal.

## 2. The CMB polarization field

The linearly polarized CMB field is described by $Q$ and $U$ (Stokes parameters) which can be combined into a spin-(2) and spin-(-2) fields, such as $P_{\pm}(\hat{n}) = Q(\hat{n}) \pm iU(\hat{n})$. For full sky, these spin fields can be expanded over spin-weighted harmonic functions basis [21]:

$$P_{\pm}(\hat{n}) = \sum_{\ell m} a_{\pm 2,\ell m}\ {}_{\pm 2}Y_{\ell m}(\hat{n}). \tag{1}$$

The polarization field can also be written as the $E$ and $B$ components, which are defined in terms of the coefficients $a_{\pm 2,\ell m}$ in the harmonic space as

$$E_{\ell m} \equiv -\frac{1}{2}[a_{2,\ell m} + a_{-2,\ell m}], \qquad B_{\ell m} \equiv -\frac{1}{2i}[a_{2,\ell m} - a_{-2,\ell m}]. \tag{2}$$

The above definition allows $E(\hat{n})$ and $B(\hat{n})$ polarization sky maps to be written in terms of spherical harmonics,

$$E(\hat{n}) \equiv \sum_{\ell m} E_{\ell m} Y_{\ell m}(\hat{n}), \qquad B(\hat{n}) \equiv \sum_{\ell m} B_{\ell m} Y_{\ell m}(\hat{n}). \tag{3}$$

The power spectra being

$$C_{\ell}^{EE} \equiv \frac{1}{2\ell+1} \sum_{m} \langle E_{\ell m} E_{\ell m}^{*} \rangle, \qquad C_{\ell}^{BB} \equiv \frac{1}{2\ell+1} \sum_{m} \langle B_{\ell m} B_{\ell m}^{*} \rangle. \tag{4}$$

The brackets indicate the average over all realizations. In this work, we will consider only the $B$-mode field and our ability to recover the $B$-mode power spectrum from a multi-frequency future CMB experiment.

## 3. The ABS method

The ABS method analytically and blindly recovers the CMB power spectra from a contaminated map, relying on the measured cross band power between different frequency bands. The data in the multipole bin $\ell$ can be written as

$$\mathcal{D}_{ij}^{\rm obs}(\ell) = f_i f_j \mathcal{D}^{\rm cmb}(\ell) + \mathcal{D}_{ij}^{\rm fore}(\ell) + \delta\mathcal{D}_{ij}^{\rm noise}(\ell). \tag{5}$$

Here, $\mathcal{D}_{ij}^{\rm obs}(\ell)$ represents the cross band power spectrum of the B-mode power spectra at the $i$- and $j$-th frequency channels, where $i, j = 1, 2 \cdots N_f$ and $N_f$ is the total number of frequency channels.The three main contributions to the signal are $\mathcal{D}_{ij}^{\rm fore}$, which expresses the cross band power matrix of the foregrounds, $\mathcal{D}^{\rm cmb}$ being

the CMB signal and $\mathcal{D}_{ij}^{\rm noise}(\ell)$, which stands for the noise contribution to the total B-mode power spectrum. $f_i = 1$ for all channels in the units of thermodynamic temperature.

The CMB power spectrum can then be analytically derived such as [7]

$$\mathcal{D}^{\rm cmb} = \left( \sum_{\mu=1}^{M+1} G_\mu^2 \lambda_\mu^{-1} \right)^{-1} , \tag{6}$$

for $M < N_f$, where $M \equiv rank(\mathcal{D}_{ij}^{\rm fore})$, which depends on the number of independent foreground components. The order of $\mathcal{D}_{ij}^{\rm obs}(\ell)$ is $N_f$. In the above equation, $G_\mu = \mathbf{f} \cdot \mathbf{E}^\mu$ and $\mathbf{f} = (f_1, \ldots, f_{N_f})^T$. The $\mu$-th eigenvector and the associated eigenvalue of $\mathcal{D}_{ij}^{\rm obs}(\ell)$ are represented by $\mathbf{E}^\mu$ and $\lambda_\mu$, respectively. The eigenvectors are normalized as $\mathbf{E}^\mu \cdot \mathbf{E}^\nu = \delta_{\mu\nu}$.

Taking the instrumental noise into account, we modify Eq. 6 such as

$$\hat{\mathcal{D}}^{\rm cmb} = \left( \sum^{\tilde{\lambda}_\mu \geq \lambda_{\rm cut}} \tilde{G}_\mu^2 \tilde{\lambda}_\mu^{-1} \right)^{-1} - \mathcal{S} \, . \qquad \tilde{\mathcal{D}}_{ij}^{\rm obs} \equiv \frac{\mathcal{D}_{ij}^{\rm obs}}{\sqrt{\sigma_{\mathcal{D},i}^{\rm noise} \sigma_{\mathcal{D},j}^{\rm noise}}} + \tilde{f}_i \tilde{f}_j \mathcal{S} \, ,$$

$$\tilde{f}_i \equiv \frac{f_i}{\sqrt{\sigma_{\mathcal{D},i}^{\rm noise}}} \, , \quad \tilde{G}_\mu \equiv \tilde{\mathbf{f}} \cdot \tilde{\mathbf{E}}^\mu \, . \tag{7}$$

The matrix $\tilde{\mathcal{D}}_{ij}^{\rm obs}$ has $\tilde{\mathbf{E}}^\mu$ and $\tilde{\lambda}_\mu$ as the $\mu$-th eigenvector and corresponding eigenvalue, respectively. To avoid unphysical eigenmodes a threshold corresponding to $\tilde{\lambda}_{\rm cut}$ in Eq. 7 is taken into account, since the instrumental noise can lead to noise dominated eigenmodes with eigenvalues of $|\tilde{\lambda}_\mu| \lesssim 1/2$ in $\tilde{\mathcal{D}}_{ij}^{\rm obs}$ (see the details at [7, 20]). Moreover, the free shift parameter $\mathcal{S}$ is responsible for stabilizing the numerical calculations, specially in the case of the CMB B-mode power spectrum due to its small amplitude together with a high instrumental noise. In the next Section, we apply the ABS method to our polarized simulations.

## 4. Our simulations and results

We consider a future-like experiment with 10 frequency bands described in Table 1. The polarized maps are generated for these bands using the LensPix software [22, 23] with resolution of $N_{\rm side} = 1024$. The predicted CMB $B$-mode power spectrum was obtained by running the public available code CAMB [24], considering the best fit parameters from the standard cosmological model [25] for $r = 0$.

We added to the CMB $Q$ and $U$ maps two polarized foregrounds adopting the nominal `PySM` model [26] as the fiducial foreground model: thermal dust and synchrotron. Finally, a Gaussian instrumental noise, uncorrelated from pixel to pixel and from channel to channel, is also considered. In order to properly test the ABS method in the presence of noise, we generated 50 independent noise maps for each frequency band, being the CMB $B$-mode power spectrum recovered by averaging

Table 1. The considered experimental setup

| Band center (GHz) | Beam FWHM (arcmin) | noise level ($\mu K_{CMB}$- arcmin) |
|---|---|---|
| 030 | 28.3 | 12.4 |
| 043 | 22.2 | 7.9 |
| 075 | 10.7 | 4.2 |
| 090 | 9.5 | 2.8 |
| 108 | 7.9 | 2.3 |
| 129 | 7.4 | 2.1 |
| 155 | 6.2 | 1.8 |
| 223 | 3.6 | 4.5 |
| 268 | 3.2 | 3.1 |
| 321 | 2.6 | 4.2 |

over these sky maps. The associated statistical errors are obtained from its dispersion. It is important to point out that the CMB signal and the foreground components are fixed for all realizations. The results are shown in Figure 1 for $l_{max} = 1050$.

Finally, in order to visualize the difference between the recovered and the "true" $B$-mode power spectrum, we also show in Fig. 1 the relative error, $\hat{\mathcal{D}}_\ell^{\rm cmb}/\mathcal{D}_\ell^{\rm real} - 1$, in the percentage level. For full sky, the recovery of the $B$-mode power spectrum using the ABS method is in agreement with the "true" one for all multipole range within 20%, even though the 1-$\sigma$ confidence level for the reconstructed $B$-mode power spectrum is quite large at low-$\ell$ region.

## 5. Conclusions

We used the ABS methodology to estimate the CMB $B$-mode power spectrum considering a full sky approach. The ABS estimator was applied to simulated mutifrequency maps at 10 different frequency bands centered at 030, 043, 075, 090, 108, 129, 155, 223, 268 and 321 GHz mimicking a possible future CMB experiment. As for foregrounds, we added thermal dust and synchrotron, in addition to the Gaussian uncorrelated noise.

The results for the ABS method for recovering the CMB $B$-mode power spectrum in the presence of foregrounds are promising. We found that for the full range of considered multipoles the ABS is able to perform robustly, being the $B$-mode power spectrum recovered within 20% compared with the input spectrum. Moreover, considering ground based experiments, which are the near future surveys designed to measure the $B$-mode signal, a partial sky analysis must be taken into account. Such analysis is in preparation [27].

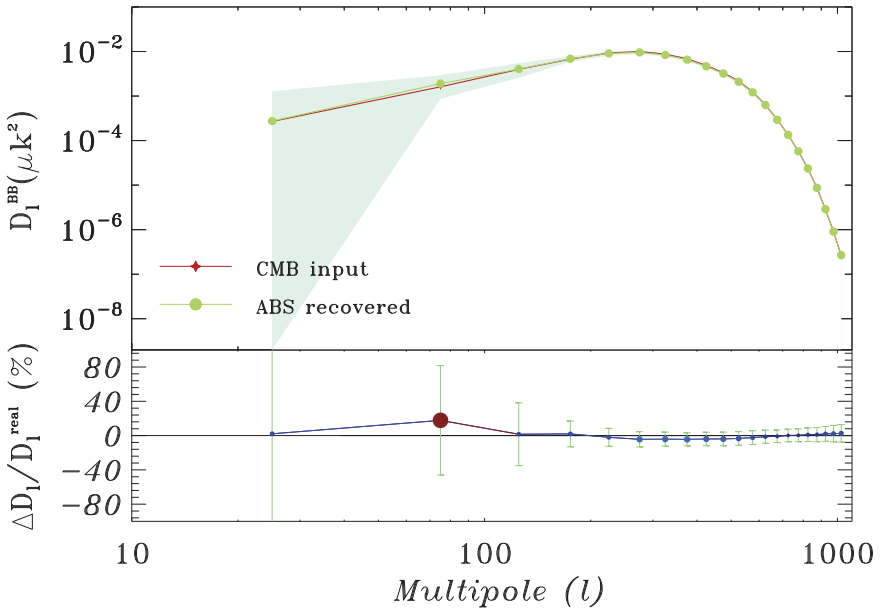



Fig. 1. Upper pane: The CMB binned $B$-mode power spectrum ($\Delta l = 50$) estimated from the ABS approach from a future experiment specified in Table 1 (green curve), considering $r = 0$ and smoothed to a Gaussian beam with FWHM=28.3′. The red curve corresponds to the CMB $B$-mode power spectrum from the sky realization without noise or foregrounds ("true" spectrum). The associated 1-$\sigma$ statistical error are also shown in the shadow region, based on 50 independent realizations of the instrumental noise. Lower panels: In comparison with the true power spectrum, the relative error, $\hat{\mathcal{D}}_\ell^{\rm cmb}/\mathcal{D}_\ell^{\rm real} - 1$ in the percentage level. The symbol colors and sizes illustrate deviations from 0% in respect to the "true" spectrum.

## Acknowledgment

L. Santos and W. Zhao are supported by NSFC No. 11603020, 11633001, 11173021, 11322324, 11653002, project of Knowledge Innovation Program of Chinese Academy of Science and the Fundamental Research Funds for the Central Universities.

## References

[1] Bennett, C. L., Halpern, M., Hinshaw, G., et al. 2003, APJS, 148, 1.
[2] Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, APJS, 208, 19.
[3] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2014, A&A, 571, A16.
[4] Planck Collaboration, Adam, R. et al. 2016, A&A , 594, A1.
[5] Planck Collaboration, Aghanim, N. et al. 2016, A&A, 594, A11.
[6] BICEP2 Collaboration, Ade, P. A. R., Aikin, R. W., et al. 2014, PRL, 112, 241101.
[7] Zhang, P., Zhang, J., & Zhang, L. 2019, MNRAS, 484, 1616.
[8] Tegmark, M., & Efstathiou, G. 1996, MNRAS, 281, 1297.
[9] Tegmark, M., de Oliveira-Costa, A., & Hamilton, A. J. 2003, PRD, 68, 123523.

[10] Bennett, C. L., Hill, R. S., Hinshaw, G., et al. 2003, APJS, 148, 97.
[11] Saha, R., Jain, P., & Souradeep, T. 2006, APJL, 645, L89.
[12] Delabrouille, J., Cardoso, J.-F., Le Jeune, M., et al. 2009, A&A, 493, 835.
[13] Remazeilles, M., Delabrouille, J., & Cardoso, J.-F. 2011, MNRAS, 418, 467.
[14] Leach, S. M., Cardoso, J.-F., Baccigalupi, C., et al. 2008, A&A, 491, 597.
[15] Fernández-Cobos, R., Vielva, P., Barreiro, R. B., & Martínez-González, E. 2012, MNRAS, 420, 2162.
[16] Delabrouille, J., Cardoso, J.-F., & Patanchon, G. 2003, MNRAS, 346, 1089.
[17] Martínez-González, E., Diego, J. M., Vielva, P., & Silk, J. 2003, MNRAS, 345, 1101.
[18] Aumont, J., & Macías-Pérez, J. F. 2007, MNRAS, 376, 739.
[19] Cardoso, J.-F., Martin, M., Delabrouille, J., Betoule, M., & Patanchon, G. 2008, arXiv:0803.1814.
[20] Yao, J., Zhang, L., Zhao, Y., Zhang, P., Santos, L. & Zhang, J. 2018, APJS, 239, 36.
[21] Seljak, U. and Zaldarriaga, M. 1996, APJ, 469, 437.
[22] Lewis, A., 2005, PRD, 71, 083008.
[23] Lewis, A., Challinor, A. & Hanson, D. 2011, JCAP, 03, 018.
[24] Lewis, A., Challinor, A., & Lasenby, A. 2000, APJ, 538, 473.
[25] Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016, A&A, 594, A13.
[26] Thorne, B., Dunkley, J., Alonso, D., & Næss, S. 2017, MNRAS, 469, 2821.
[27] Santos, L. et al. 2019, arXiv:1908.07862.

# Tackling cosmic strings by knot polynomials

Xinfei LI[*,**], X. LIU[**,†,‡] and Yong-chang Huang[**]

[*] *School of Science, Guangxi University of Science and Technology, Guangxi, 545026, China*

[**] *Institue of Theoretical Physics, Beijing University of Technology, Beijing, 100124, China*

[†] *Beijing-Bublin International College, Beijing University of Technology, Beijing, 100124, China*

[‡] *xin_liu@bjut.edu.cn*

Cosmic strings in the early universe have received revived interest in recent years. In this work we derive these structures as topological defects from singular distributions of the quintessence field of dark energy. Emphasis is placed on the topological charge of tangled cosmic strings, which originates from the Hopf mapping degree and is a Chern-Simons type action possessing strong inherent tie to knot topology. It is shown that the HOMFLYPT (Jones) polynomial can be constructed in terms of this charge, serving as a topological invariant much stronger than the traditional Gauss linking numbers in characterizing string topology. This method induces applications in two aspects. One is to search a tool to measure the cascade of decreasing topological complexity and that of decreasing energy/entropy. The other is arousing a promising mathematical approach for the study of physical breaking-reconnection processes of tangled cosmic strings.

## 1. Complexity and cosmic strings

Cosmic strings as one dimensional relics might have been produced in the early universe, which lead to galaxies and primordial supermassive black holes. Interestingly, tangled cosmic strings compacitificated in a region (contexture) could collapse, causing cold/hot spots in CMB.[1] Dynamical simulations show that tangled string structures whose components tend to move to each other and then reconnect at cross sites quickly become simpler ones to carry less energy/entropy. It is hard to analyze this topology-changing process from a pure dynamical point of view; fortunately, similar process has been investigated via topological complexity in classical fluids. By means of analogy with the knotted fluid structures, we intend to study the evolution of tangled cosmic strings via topological complexity.

Let us start with a superconducting cosmic string (SCS) model with two massive complex scalars, $\phi$ and $\sigma$,

$$S=\int d^4x\left[\left|\widetilde{D}_\mu\phi\right|^2+\left|D_\mu\sigma\right|^2-\frac{1}{4}\widetilde{F}^{\mu\nu}\widetilde{F}_{\mu\nu}-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}+V\left(\phi,\sigma\right)\right], \tag{1}$$

where $V\left(\phi,\sigma\right)=\frac{1}{4}\lambda_\phi\left(\left|\phi\right|^2-\eta_\phi^2\right)^2+\frac{1}{4}\lambda_\sigma\left(\left|\sigma\right|^2-\eta_\sigma^2\right)^2+\beta\left|\phi\right|^2\left|\sigma\right|^2$. This model breaks the $\phi$-related $U(1)$ symmetry in the ground state and keeps the $\sigma$-related one yielding current-carrying SCSs with orientations. With fine-tuning of the Higgs sector, currents are held stable in the cores of the SCSs which can be treated as line defects when observed from far away.

The equation of motion derived from (1) gives a distribution of the fundamental field $\phi$ on the base manifold. Introducing a covariant derivative $D_\mu^{\mathrm{ind}}\phi=\partial_\mu\phi-iA_\mu$,

which indeed corresponds to the parallel field condition in general relativity,[2,3] we obtain a $U(1)$ gauge potential induced by the $\phi$-distribution on the manifold. This $A_\mu$, through the $\phi$-distribution, is able to reflect the topology of the tangled SCSs. Then, based on $A_\mu$ a low-energy classical Yang-Mills model can be proposed to present stable knotted soliton structures in three dimensions, where a topological charge $Q$ is introduced to characterize the tangledness of the system,[4]

$$Q = \int_\Omega A \wedge F, \tag{2}$$

where the field strength is defined as $F = dA$. This $Q$ is recognized to be an integral of an abelian Chern-Simons (CS) type 3-form; traditionally, it can be delivered by topological numbers as $Q(\mathcal{L}) = \sum_{k=1}^{N} W_k^2 SL(\gamma_k) + \sum_{k,l=1(k<l)}^{N} W_k W_l Lk(\gamma_k, \gamma_l)$, where $\mathcal{L}$ denotes the link, $\mathcal{L} = \bigoplus_k \gamma_k$. $W_k$ is the topological charge (i.e., winding number) carried by the closed string $\gamma_k$. The $SL(\gamma_k)$ is the self-linking number of a knot $\gamma_k$, and $Lk(\gamma_k, \gamma_l)$ the Gauss linking number between two link components $\gamma_k$ and $\gamma_l$.[5] This $Q$ was used to study knotted cosmic strings in references.[6]

(a) (b)

Fig. 1. Different topological configurations have the same self-linking number 0.

(a) (b) (c)

Fig. 2. Different topological configurations have the same linking number 0.

However, the (self-)linking numbers are unfortunately weak invariants which fail to distinguish some typical knots/links such as in Figs. 1 and 2. In the following we will introduce more powerful tools of knot theory to characterize topology of SCS structures. Our starting point will still be the CS 3-form $Q$, which is able to generate various kinds of knot topological invariants thanks to the quantum CS topological field theory, such as the Kauffman bracket, Jones and HOMFLYPT polynomials, the Vassiliev finite type invariants via the Kontsevich integral, Khovanov homology, and so on.

## 2. HOMFLYPT polynomial for SCSs

SCSs have a chance to obtain released twist during evolutions, although the latter is not easy to be demonstrated mathematically. The Călugăreanu-White theorem states that writhing and twisting both have independent contributions to string topology, hence we intend to construct the two-variable HOMFLYPT polynomial to catch the topological information of writhes and twists in SCSs.

*Writhe contribution*: Integrating off two dimensions from the triple integral (2), the CS charge $Q$ becomes $Q(\mathcal{L}) = \sum_k W_k \oint_{\gamma_k} A_i dx^i$, where for simplicity one usually takes $W_k = 1$. According to Ref. 7, we propose to consider the exponential form $\exp(Q(\mathcal{L})) = \exp\left(\oint_{\mathcal{L}} A_i dx^i\right)$; in the following a bracket symbol is used, $\langle \mathcal{L} \rangle \equiv e^{Q(\mathcal{L})}$. Obviously, $\langle \bigcirc \rangle = 1$.


Fig. 3. The different states of four almost-the-same knots/links which differ only at one particular crossing site: (a) $L_+$, over-crossing; (b) $L_-$, under-crossing; (c) non-crossings: (*left*) $L_0$, left-right open; (*right*) $L_\infty$, up-down open.

First, let us consider un-oriented strings. Preparing a path-dependent quantity $\alpha$, defined by $\alpha \equiv \langle \gamma_+ \rangle$ and $\alpha^{-1} \equiv \langle \gamma_- \rangle$ with $\gamma_+ =$  and $\gamma_- =$  , we can follow the routine of knot theory to examine the basic states of Fig. 3.

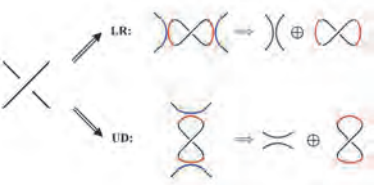



Fig. 4. Two channels for the over-crossing $L_+$: left-right (LR) and up-down (UD) splittings.

Second, based on an ergodic statistical hypothesis, a crossing site decomposes into two equivalent channels with the aid of two pairs of added imaginary strands which cancel each other. For instance, the over-crossing state $L_+$ in Fig. 4 has the splitting $\langle L_+ \rangle = \langle L_+ - \text{LR} \rangle + \langle L_+ - \text{UD} \rangle$, where $\langle L_+ - \text{LR} \rangle = e^{\oint_{L_0 \oplus \gamma_+}} = \alpha e^{\oint_{L_0}}$ and $\langle L_+ - \text{UD} \rangle = e^{\oint_{L_\infty \oplus \gamma_-}} = \alpha^{-1} e^{\oint_{L_\infty}}$. Performing a similar procedure for $L_-$ one obtains $\left\langle \times \right\rangle = \alpha \left\langle )( \right\rangle + \alpha^{-1} \left\langle \asymp \right\rangle$ and $\left\langle \times \right\rangle = \alpha^{-1} \left\langle )( \right\rangle + \alpha \left\langle \asymp \right\rangle$. Then, endowing the components of $\mathcal{L}$ with orientations, and introducing a R-polynomial $R[\mathcal{L}] = \alpha^w \langle \mathcal{L} \rangle$, with $w$ denoting the writhe of $\mathcal{L}$, we achieve

$$R\left(\times\right) - R\left(\times\right) = zR\left()(\right), \tag{3}$$

where $z = \alpha^2 - \alpha^{-2}$ stands for the writhe contribution. This presents one of the skein relations (i.e., the definition formulae) of the R-polynomial.

Fig. 5. Dehn surgery. (a) A cosmic string with intrinsic twist can be regarded as a twisted ribbon. (b) A $2\pi$-twist is added on a torus.

Fig. 6. A twisted ribbon can be decomposed into the sum of an untwisted ribbon, a trivial annulus and a positive twisting operation.

*Twist contribution*: In the context of physical knots, the Dehn surgery is considered as an idealized technique to treat insertion/deletion of intrinsic twists of knots, as showed in Figs. 5 and 6. The twisted string is mathematically identical to the sum of an untwisted ribbon, a trivial annulus and a positive twisting operation,[3,8] *i.e.*, $\exp\left(\oint \right) = a\exp\left(\oint \right)\exp\left(\oint_{\bigcirc}\right)$, where $a = \exp\left(\oint_{\text{positive twisting operation}}\right)$. Introducing the so-called HOMFLYPT polynomial $P(a,z) = a^{-w}R(a,z)$, with $a^{-w}$ being a compensation factor,

**Comment.** : $w\left[\ \right] = +1,\quad w\left[\ \right] = -1,\quad w\left[\ \right] = w\left[\ \right] = w\left[\ \right] = 0.$

we arrive at the HOMFLYPT skein relations of $P(\mathcal{L})$:

$$(1)\ P\left(\bigcirc\right) = 1, \qquad (2)\ aP\left(\times\right) - a^{-1}P\left(\times\right) = zP\left(\ \right). \tag{4}$$

Furthermore, if setting $z = \tau^{1/2} - \tau^{-1/2}$ and $a = \tau^{-1}$, (4) degenerates to the skein relations of the Jones polynomial:[8]

$$(1)\ V\left(\bigcirc\right) = 1,\quad (2)\ \tau^{-1}V\left(\times\right) - \tau V\left(\times\right) = (\tau^{-\frac{1}{2}} - \tau^{\frac{1}{2}})V\left(\ \right). \tag{5}$$

## 3. Example and conclusion

*Example*: The recurrence relation of torus knots/links, $T(2,n)$, is computed:[9] $P_{T(2,n)} = \left[\frac{k^{n-2}-(-k)^{-(n-2)}}{a^{n-3}(k+k^{-1})^{-1}}\right]P_{T(2,3)} + \left[\frac{k^{n-3}-(-k)^{-(n-3)}}{a^{n-3}(k+k^{-1})^{-1}}\right]P_{T(2,2)}$, where $P_{T(2,3)} = 2a^{-2} + a^{-2}z^2 - a^{-4}$ and $P_{T(2,2)} = a^{-1}z + (a^{-1} - a^{-3})z^{-1}$ with $z = k - k^{-1}$. Direct computation of the HOMFLYPT values of $T(2,n)$ allows us to associate a sequence of decreasing numerical values to a cascade of states with decreasing topological complexity, as shown in Figs. 7 and 8.

*Conclusion*: In this paper a new knot theoretical tool, the HOMFLYPT polynomial, has been introduced into the study of cosmic string topology. It can be used to characterize the complexity of SCSs. Typical examples of application are also provided. We hope this method could be useful for investigating topological non-conservative processes such as reconnections.

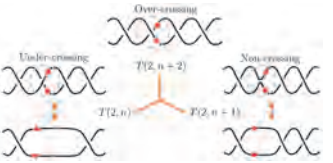



Fig. 7. Analytical implementation of the polynomial skein relations to the cascade of a family of torus knots/links examined. (From Ref. 9)

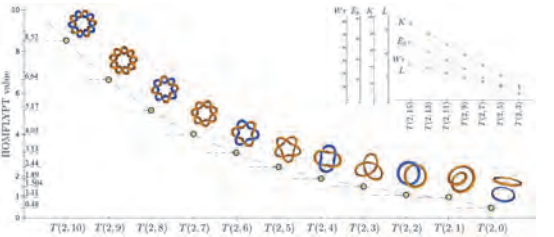



Fig. 8. The HOMFLYPT values of torus knots/links $T(2, n)$ monotonically decrease with $n$. (From Ref. 9)

## Acknowledgements

In this work, Xinfei Li is supported by the Doctor Start-up Foundation of Guangxi University of Science and Technology with Grant No. 19Z21 and by International program for graduate students of Beijing University of Technology. Yong-Chang Huang is supported by National Science Foundation of China (No.11875081). Xin Liu is supported by National Science Foundation of China (No. 11572005) and the Youth Fellowship of the Haiju Project of Beijing.

## References

1. N. Turok and D. Spergel, *Phys. Rev. Lett.* **64**, 2736 (1990).
2. X. Li, X. Liu and Yong-Chang Huang, *Int. J. Mod. Phys.* **A 32**, 1750164 (2017).
3. X. Li and X. Liu, *Int. J. Mod. Phys.* **A 33**, 1850156 (2018).
4. L.D. Faddeev and A.J. Niemi, *Nature* **387**, 58 (1997).
5. H.K. Moffat, *J. Fluid Mech.* **35**, 117 (1969).
6. J.D. Bekenstein, *Phys. Lett.* **B 282**, 44 (1992).
7. X. Liu and R.L. Ricca, *J. Phys. A: Math. Theor.* **45**, 205501 (2012).
8. X. Liu and R.L. Ricca, *J. Fluid Mech.* **773**, 34 (2015).
9. X. Liu and R.L. Ricca, *Scientific Report* **7**, 24118 (2017).

# Cosmic cable

Colin MacLaurin*

*School of Mathematics and Physics,*
*University of Queensland, Brisbane, Australia*
**E-mail: colin.maclaurin@uqconnect.edu.au*

I investigate the relativistic mechanics of an extended "cable" in an arbitrary static, spherically symmetric spacetime. Such hypothetical bodies have been proposed as tests of energy and thermodynamics: by lowering objects toward a black hole, scooping up Hawking radiation, or mining energy from the expansion of the universe. I review existing work on stationary cables, which demonstrates an interesting "redshift" of tension, and extend to a case of rigid motion. By using a partly restrained cable to turn a turbine, the energy harvested is up to the equivalent of the cable's rest mass, concurring with the quasistatic case. Still, the total Killing energy of the system is conserved.

## 1. Introduction

Hypothetical ropes/strings/cables/tethers in curved spacetime have been applied as test scenarios for energy, thermodynamics, and mechanics: hence are related to major open research questions. The cable considered here obeys ordinary energy conditions for macroscopic matter, and is conceived as engineered rather than naturally occurring. These properties are distinct from other strings in theoretical physics including Nambu-Goto relativistic strings, cosmic strings, and quantum strings such as in superstring theory.

Penrose proposed lowering a mass towards a black hole to extract energy.[1] Bekenstein considered lowering a hot object towards a black hole, to convert heat into work and violate the second law of thermodynamics, inspired by a talk of Geroch.[2] However Gibbons was the first to analyse the cable itself, including its tension.[3] Unruh & Wald suggested a resolution to Geroch's thermodynamic paradox, now in the context of Hawking radiation, and showed a "redshift" of tension along the cable.[4] Redmount used Weyl's axisymmetric solution to analyse distortions of Schwarzschild spacetime by two masses suspended on ropes.[5]

Others considered cables in Friedmann-Lemaître-Robertson-Walker spacetime, to test the dynamics of an expanding universe. Davies showed the futility of scooping up Hawking radiation from the de Sitter horizon.[6] Harrison proposed latching a rope onto distant receding bodies, to mine mechanical energy from the expansion.[7] Others also considered a "tethered galaxy",[8] or a network of strings reminiscent of Szekeres' gravitational compass. There are plenty of other sources, even omitting the cosmic string and string theory literature.

## 2. Newtonian cables

To show the concept is perfectly reasonable, consider a spool of cable tied to a tree branch at height $L$ above ground. Set up a steady state where a segment always stretches from winch to ground, and moves downwards at constant speed $\beta$. Then

under Galilean gravitational acceleration $g$ and linear mass density $\mu$, a constant force $\mu g L$ is exerted downwards. This gives a power $\mu\beta g L$ which can be harnessed at the winch, at the expense of gravitational potential energy.

Next consider a Newtonian "de Sitter" cosmology: an empty universe with cosmological constant $\Lambda$. Suppose the winch is anchored at the coordinate origin. Poisson's law $\vec{\nabla}^2\Phi + \Lambda = 4\pi G\rho$ has solution $\Phi = -\Lambda(x^2+y^2+z^2)/6$ in this case, where $\rho = 0$ since we ignore the cable's own gravitational field. The force on a static cable particle of mass $m$ at location $(x, y, z)$ is $\vec{f} = -m\,\vec{\nabla}\Phi = (x, y, z)m\Lambda/3$. The total force at the winch is hence $-\Lambda\mu L^2/6$ for a cable of length $L$. By allowing the cable to recede, work is gained, but the cable is gradually lost to space.

## 3. Stationary cable

This section reviews results of Gibbons and others, with some additions.[3,5,9] Consider an arbitrary spacetime containing a cable with linear mass density $\mu(\mathbf{x})$ and tension $T(\mathbf{x})$. With no sideways rigidity, the cable's stress-energy tensor is $\mathbf{T} = \tilde{\mu}\mathbf{u}\otimes\mathbf{u} + \tilde{T}\mathbf{q}\otimes\mathbf{q}$. Here $\mathbf{u}$ is the 4-velocity field of cable particles, $\mathbf{q}$ is a unit spatial vector field pointing along the cable, and $\tilde{\mu} := \mu/A$ and $\tilde{T} := T/A$ are scaled by the cross-sectional area.

Conservation of stress-energy implies the vector $\operatorname{div}\mathbf{T}$ vanishes, which when contracted with $\mathbf{u}$ and $\mathbf{q}$ leads to $d\tilde{\mu}/d\tau = -\tilde{\mu}\operatorname{div}\mathbf{u} + \tilde{T}\dot{\mathbf{q}}\cdot\mathbf{u}$ and $d\tilde{T}/dL = -\tilde{T}\operatorname{div}\mathbf{q} - \tilde{\mu}\mathbf{q}\cdot\dot{\mathbf{u}}$ respectively. Here $L$ is the proper-frame length along the cable, $\dot{\mathbf{u}} := \nabla_{\mathbf{u}}\mathbf{u}$ is 4-acceleration, $\dot{\mathbf{q}} := \nabla_{\mathbf{q}}\mathbf{q}$, and a dot between vectors implies the metric inner product. However stress-energy need not be conserved, for instance if the cable's gravitational field is ignored so $\mathbf{T}$ is not the source term in the Einstein field equations. Still, $\mathbf{u}\cdot\operatorname{div}\mathbf{T} = 0$ is guaranteed.[10]

Now suppose the spacetime contains a timelike Killing vector field $\boldsymbol{\xi}$, and that the cable particles are stationary. Hence $\mathbf{u} = \boldsymbol{\xi}/V$, where $V := \sqrt{-\boldsymbol{\xi}\cdot\boldsymbol{\xi}}$ is the redshift factor. In this case stress-energy conservation is guaranteed. It follows $d\tilde{\mu}/d\tau = 0$, and if $\mu$ is also constant over space then the tension varies as

$$\boxed{T = W\frac{V_{\text{end}}}{V} + \mu\Big(1 - \frac{V_{\text{end}}}{V}\Big)} \tag{1}$$

where $W$ is a weight (mass times magnitude of proper acceleration, if a point mass) hanging at the end. This exhibits a curious effect: the tension is "redshifted" by transmission along the cable, due to gravitational time-dilation. This is most evident for a massless cable $\mu \equiv 0$,[4] a limiting case of which is familiar in surface gravity as the "force at infinity" to support a unit mass particle at a horizon.[11] The cases $W = \mu$, or a cable ending at the horizon with $W = 0$, both yield a constant tension $T \equiv \mu$.[5,12] Alternatively, one can extend conservation of energy arguments used by Einstein and Bondi for redshift of photons, to justify redshift of force.

## 4. Moving cable

This section describes the kinematics and dynamics of a moving cable. For further explanation see §7, §9, and §11 of preliminary work.[13]

### 4.1. *Apparatus and background spacetime*

Consider a static, spherically symmetric spacetime with general metric

$$ds^2 = -e^{2\alpha(r)}dt^2 + e^{2\zeta(r)}dr^2 + r^2(d\theta^2 \sin^2\theta\, d\phi^2). \tag{2}$$

Static observers form a useful reference or comparison, these have 4-velocity

$$u^\mu_{\text{static}} = \big(e^{-\alpha}, 0, 0, 0\big), \tag{3}$$

where the $r$-dependence is omitted for brevity. The 4-acceleration is $a^\mu = (0, \alpha' e^{-2\zeta}, 0, 0)$ with magnitude $|\alpha'|e^{-2\zeta}$, and is directed towards increasing $r$ if $\alpha' > 0$ and vice versa. A dash will always indicate derivative with respect to $r$. The gravitational redshift factor, determined as usual by static observers, is $\sqrt{-\partial_t \cdot \partial_t} = e^\alpha$.

Assume the cable is hanging "downwards" in an equilibrium state, so as a whole the 4-velocity field is unchanged by translation in $\partial_t$. As in Section 2, set a winch or spool at the "top" $r = r_0$ say, attached to a turbine which allows only a fixed angular velocity. At the winch location, we assume $e^{\alpha(r_0)} = 1$, which interprets the winch as free of gravitational redshift (otherwise one could introduce an extra constant, or rescale the $t$-coordinate via $\bar{t} := e^{\alpha(r_0)}t$). Suppose the cable is cut repeatedly, so its end is fixed at $\approx r_{\text{end}}$, which simplifies the calculations. (One could imagine a cutting robot tethered to a static "service cable".)

### 4.2. *Kinematics*

Parametrise radial 4-velocities using the speed $\beta(r)$ relative to a local static observer:

$$u^\mu = \gamma\big(e^{-\alpha}, \beta e^{-\zeta}, 0, 0\big), \tag{4}$$

where $\gamma(r) := (1-\beta^2)^{-1/2}$ is the corresponding Lorentz factor, and we allow $\beta < 0$ via $\operatorname{sgn}(\beta) := \operatorname{sgn}(dr/d\tau)$. The cable's stress-energy is not conserved in general, as mentioned previously, but particle number

$$^{(1)}\nabla_i(nu^i) = 0 \tag{5}$$

is conserved. Here $n(r)$ is the number density, and the "(1)" indicates divergence in the radial direction only, which is a sum over $i = 0, 1$, in our coordinates. This evaluates to the first-order ODE $(n\beta\gamma)' + (n\beta\gamma)\alpha' = 0$, with solution

$$n\beta\gamma = Ce^{-\alpha} \tag{6}$$

for some constant $C$, ignoring trivial cases. In a local static frame, the number flux vector $n\mathbf{u}$ has components $n\gamma(1, \beta, 0, 0)$, hence $n\beta\gamma$ is interpreted as the local

number flux density in the radial direction, so $C \equiv n\beta\gamma e^{\alpha}$ is the redshift-corrected value — that is, the number flux density determined at the winch.

An equation of state gives an additional constraint. Instead of relativistic elasticity, for simplicity we assume Born-rigidity, meaning the expansion tensor is zero (constant proper-frame distance between neighbouring particles) in the radial direction. While Born-rigidity is not possible as an *implicit* property of a physical material, we treat it as a toy model, with discretion.[a]

Zero divergence $\nabla_i u^i = 0$, $i = 0, 1$, simplifies to $(\beta\gamma)' + (\beta\gamma)\alpha' = 0$ with solution

$$\beta\gamma = Ke^{-\alpha} \tag{7}$$

for some constant $K$. $\beta\gamma$ is the "proper speed" relative to the local static observer, hence $K$ is interpreted as the redshift-corrected proper speed; we treat $K \in (-\infty, \infty)$ as provided. From Equation 6, $n = C/K$, and it follows the density $\mu$ is constant. Also,

$$\gamma = \sqrt{1 + K^2 e^{-2\alpha}} \qquad\qquad \beta = K/\sqrt{K^2 + e^{2\alpha}}, \tag{8}$$

which follows from Equation 7. The cable 4-velocity is hence

$$u^{\mu} = e^{-\alpha}(\sqrt{1 + K^2 e^{-2\alpha}}, Ke^{-\zeta}, 0, 0). \tag{9}$$

Brotas gives similar results for Schwarzschild spacetime.[14] The 3-velocity $\vec{u}$ relative to the local static frame is $(K/\sqrt{K^2 + e^{2\alpha}}, 0, 0)$, from the decomposition $\mathbf{u} = \gamma(1, \vec{u})$ in a static frame, with $\beta = \|\vec{u}\|$ as expected. While it may seem *a priori* that rigid kinematics are trivial, various authors including Harrison do not correctly treat the frame dependence of length, i.e. "length-contraction". Now given the above motion, a cable particle has 4-acceleration $\mathbf{a} := \nabla_{\mathbf{u}}\mathbf{u}$ with magnitude

$$a = \sqrt{\mathbf{a} \cdot \mathbf{a}} = \frac{|\alpha'|e^{-\zeta}}{\sqrt{1 + K^2 e^{-2\alpha}}} \tag{10}$$

directed in the $\mathrm{sgn}(\alpha')$ direction of the $r$-axis. Physically, this acceleration is due to the support of the cable above.

### 4.3. *Dynamics, energy, and power*

The surplus of energy entering and exiting the system due to the cable's motion gives the power harvested. Consider firstly various energy fluxes measured locally in a static frame at given $r$:

- mass flux: $\mu\beta_r\gamma_r \equiv \mu K e^{-\alpha_r}$, locally at $r$
- kinetic energy flux: $\mu\beta_r\gamma_r(\gamma_r - 1) \equiv \mu K(\gamma_r - 1)e^{-\alpha_r}$, locally
- mass plus kinetic flux: $\mu\beta_r\gamma_r^2 \equiv \mu K\gamma_r e^{-\alpha_r}$, locally

[a] Physical intuition is applied here. For instance we derive tension and power, but not the speed of travelling waves. Similarly, our usage of "tension" is not necessarily compatible with Lorentz transformation of the stress-energy tensor between frames, but should be self-consistent.

In particular at the winch, the passing cable has mass flux $\mu K$ and kinetic energy flux $\mu K(\sqrt{1+K^2}-1)$, with sum $\mu K\sqrt{1+K^2}$. These are the energy rates entering the system. Now for a given power in a static frame at some $r$, transmitted to a static frame at $r_1$, the received power is redshifted by $e^{\alpha_r-\alpha_1}$ twice. (Imagine transmission via photons, then both the number rate and individual wavelengths are affected.)

Subtract the total flux at the start and end of the cable, to determine the overall power as transmitted to the winch frame. This calculation invokes time symmetry. However the incoming kinetic energy must also be provided,[b] hence at the winch:

$$\boxed{\mu K\big(1-\sqrt{K^2+e^{2\alpha_{\text{end}}}}\big).} \tag{11}$$

The winch end, with redshift 1, is included implicitly. The term in parentheses is the energy profit per mass, which for a slow cable ending near a Killing horizon ($e^\alpha \approx 0$) approaches 1, or 100% conversion of $E=mc^2$! If the technology exists to recover the outgoing kinetic energy, the power profit increases to $\mu K(1-e^{\alpha_{\text{end}}})$.

We check using an alternate derivation from forces. The 4-force on a particle is best defined as $\mathbf{f} := \nabla_{\mathbf{u}}\mathbf{p}$, where $\mathbf{p} := m\mathbf{u}$ is the 4-momentum. It follows $\mathbf{f} = dm/d\tau\,\mathbf{u} + m\mathbf{a}$, but $dm/d\tau = 0$ in our case, a "pure force".[15] In an arbitrary orthonormal frame, $\mathbf{f}$ has components $\gamma(\dot{E},\vec{f})$ where $\vec{f}$ is 3-force and $\dot{E}$ is the rate of local energy change in this frame. In a local static frame, $\mathbf{f}$ has components

$$f^\mu = \gamma\Big(\frac{m\beta^2\alpha' e^{\alpha-\zeta}}{K}, \frac{m\alpha' e^{-\zeta}}{\gamma}, 0, 0\Big), \tag{12}$$

which satisfy $\dot{E} = \vec{f}\cdot\vec{u} + \gamma^{-2}dm/d\tau$ as expected.[15] The 3-force is unchanged under boosts in the radial direction. Now a coordinate interval $dr$ contains a mass $\mu\gamma e^\zeta dr$ of cable, according to a *static* frame. This is because a static observer measures a proper length $e^\zeta dr$ with its ruler, and so a greater interval $\gamma e^\zeta dr$ of cable fits, according to the cable's length-contracted rulers in this frame.[16] Replacing $m$ in $\dot{E}$ with this mass density, the power contributed from this cable element is $\mu\beta\alpha' dr$, which is redshifted twice to the winch. This integrates to $\mu K\sqrt{K^2+e^{2\alpha}}$, which concurs with Equation 11 after evaluating it at the cable ends and subtracting the initial kinetic energy rate. This is the total power gained. The reason for focusing on static frames is that they respect time symmetry, hence measure *sustainable* rates.

The 3-force contributed locally from an interval $dr$ is $\mu\alpha' dr$. After redshifting once to a given $r_1$, then integrating over all cable below $r_1$, the tension at $r_1$ is

$$\boxed{\mu\big(1-e^{\alpha_{\text{end}}-\alpha_1}\big)} \tag{13}$$

[b] In practice the cable tension would cause this motion. However in our accounting system, all work done by the cable is harvested, hence the incoming kinetic energy must be considered separately. Also the incoming rest mass is treated as *gratis* or expendable, in order to achieve any profit.

in the static frame or cable frame, and is independent of $K$. At the winch the tension is $\mu(1 - e^{\alpha_{\text{end}}})$. For a cable ending near a Killing horizon, the winch tension is $\approx \mu$, which saturates various energy conditions, and $T/\mu \approx 1$ is the Planck force.

## 5. Conclusions and future work

A cable in curved spacetime illustrates interesting relativistic effects including "redshift" of tension. In an idealised scenario, usable energy equivalent to 100% of a cable's rest mass can be extracted; I have extended calculations to a moving cable. The overall Killing energy is conserved. There are many ways this research could be extended. Instead of ignoring backaction, string-like exact solutions could be analysed. Other possible avenues are quantum effects, elasticity, more general motions, application to quasilocal energies, and thermodynamics.

## References


1. R. Penrose, Gravitational collapse: The role of general relativity, *Nuovo Cimento Rivista Serie* (1969).
2. J. Bekenstein, Black holes and the second law, *Nuovo Cimento Lettere* (1972).
3. G. Gibbons, On lowering a rope into a black hole, *Nature Physical Science* (1972).
4. W. Unruh and R. Wald, Acceleration radiation and the generalized second law of thermodynamics, *Physical Review D* (1982).
5. I. Redmount, Topics in black-hole physics, PhD thesis, California Institute of Technology (1984).
6. P. Davies, Mining the universe, *Physical Review D* (1984).
7. E. Harrison, Mining energy in an expanding universe, *Astrophysical Journal* (1995).
8. T. Davis, C. Lineweaver and J. Webb, Solutions to the tethered galaxy problem in an expanding universe and the observation of receding blueshifted objects, *American Journal of Physics* (2003).
9. I. Fouxon, G. Betschart and J. Bekenstein, Bound on viscosity and the generalized second law of thermodynamics, *Physical Review D* (2008).
10. M. Wernig-Pichler, Relativistic elastodynamics, PhD thesis, Universität Wien (2006).
11. R. Wald, *General relativity* (University of Chicago, 1984).
12. A. Brown, Tensile strength and the mining of black holes, *Physical Review Letters* (2013).
13. C. MacLaurin, Expanding space, redshifts, and rigidity: Conceptual issues in cosmology, Master's thesis, University of Queensland (2015).
14. A. Brotas, Fishing in black holes, *arXiv e-prints* (2006).
15. M. Tsamparlis, *Special relativity* (Springer, 2010).
16. C. MacLaurin, Clarifying spatial distance measurement, in *Proceedings of the 2018 Marcel Grossmann conference*, submitted. arXiv:1911.07500.

# Multi-scale search machine to detect cosmic strings

Alireza Vafaei Sadr[1,2,3], Seyed Mohammad Sadegh Movahed[*1], Marzieh Farhang[1], Christophe Ringeval[4], François R. Bouchet[5], Bruce Bassett[3,6,7,8] and Martin Kunz[2]

[1] *Department of Physics, Shahid Beheshti University, Velenjak, Tehran 19839, Iran*
[*] *E-mail: m.s.movahed@ipm.ir*
[2] *Département de Physique Théorique and Center for Astroparticle Physics, Université de Genéve, 24 Quai Ernest Ansermet, 1211 Genéve 4, Switzerland*
[3] *African Institute for Mathematical Sciences, 6 Melrose Road, Muizenberg, 7945, South Africa*
[4] *Centre for Cosmology, Particle Physics and Phenomenology, Université Catholique de Louvain, Louvain-la-Neuve B-1348, Belgium*
[5] *Institut d' Astrophysique de Paris (UMR7095: CNRS & UPMC-Sorbonne Universities), F-75014, Paris, France*
[6] *South African Astronomical Observatory, Observatory, Cape Town, 7925, South Africa*
[7] *SKA South Africa, The Park, Park Road, Pinelands, Cape Town 7405, South Africa*
[8] *Department of Maths and Applied Maths, University of Cape Town, Cape Town, South Africa*

Cosmic topological defects possibly formed during phase transitions in the very early universe and are theoretically well-motivated. Cosmic strings (CSs) would leave distinct imprints on CMB fluctuations. We apply certain topological and geometrical measures in a multi-scale edge-detection algorithm to CMB maps with contributions from CS network to assess the capability of the CS footprints. On the noiseless sky maps with an angular resolution of $0.9'$, we show that our pipeline is capable of detecting CSs with $G\mu$ as low as $G\mu > 4.3 \times 10^{-10}$. We also explore two powerful tree-based machine learning algorithms to perform feature importance analysis. This analysis would improve future CS searches to focus on the most significant and informative features. Our machine-learning algorithm can detect CSs with $G\mu > 2.1 \times 10^{-10}$ at $3\sigma$ level, for a noise-less experiment and angular resolution of $0.9'$.

*Keywords*: Cosmic background radiation; early Universe; large-scale structure; data analysis.

## 1. Introduction

The line-like version of topological defects called cosmic strings (CSs) are theoretically expected to be produced in the early Universe (see [23, 22] and references therein). A lot of effort has been put into developing powerful statistical tools for cosmic string network detection and putting tight upper and lower bounds on the CS tension, parameterized by $G\mu$, where $G$ and $\mu$ represent Newton's constant and the string's tension, respectively. Mathematical description of the string tension is intimately related to the energy of the phase transition epoch as $\frac{G\mu}{c^2} = \mathcal{O}(\varpi^2/M^2_{\rm Planck})$, with the symmetry breaking energy scale, $\varpi$. The CS network can leave imprints on the CMB temperature and polarization anisotropies through different mechanisms such as ordinary and integrated Sachs-Wolfe effect [8] and lensing. The Gott-Kaiser-Stebbins effect produces line-like discontinuities on the CMB temperature anisotropies of the form: $\frac{\delta T}{T} \sim 8\pi G\mu v_s$ [8, 21]. Here $v_s$ is the transverse velocity of the string. CMB-based approaches to search for CS have been conducted to wide range of constraint on the $G\mu$. This range covers the lower value form $G\mu \gtrsim 6.3{\times}10^{-11}$ for noise-less maps [5] to upper limits of $G\mu < 8.8{\times}10^{-7}$ [1, 23, 22].

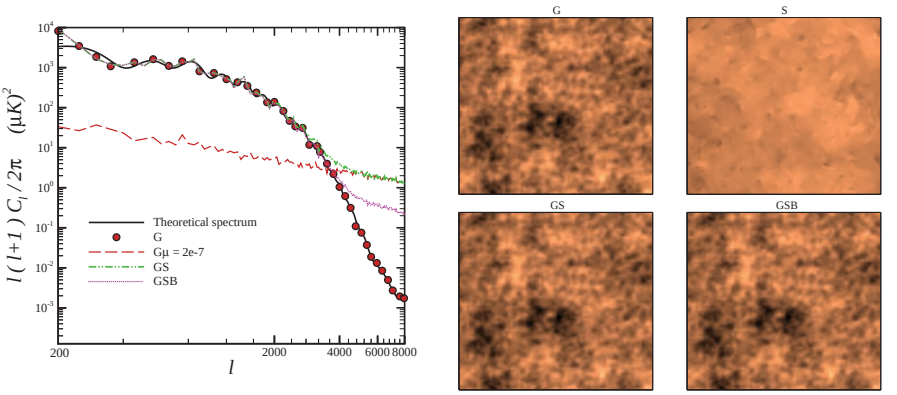



Fig. 1. Left panel: The CMB power spectrum for different components. Right panel includes the Gaussian CMB, CS-induced anisotropies, the combination of the Gaussian and string, smeared by the beam named by G, S, GS and GSB maps, respectively [23].

Here, we propose a multi-scale edge-detection algorithm to search for the CS network on the CMB anisotropies. We also develop machine-learning (ML)-based algorithms for a systematic search of CSs. ML algorithms are ideally capable to exploit all available information in the data and extract the most relevant features for CS detection.

## 2. Simulation the components

Our pipeline for simulation of mock CMB map consists of three components: (1) the Gaussian inflation-induced contribution denoted by $G$, as well as the secondary lensing signal, (2) the CS contribution, $G\mu \times S$, where $S$ represents the normalized simulated template for the string signal and $G\mu$ sets its amplitude, and (3) the experimental noise indicated by $N$. The full simulated map $T(x, y)$, with $x$ and $y$ representing pixel coordinates, would then be

$$T(x, y) = B\,[G(x, y) + G\mu \times S(x, y)] + N(x, y) \tag{1}$$

where $B$ characterizes the beam function [23]. For the CS-induced CMB anisotropies, we use high-resolution flat-sky CMB maps obtained from numerical simulations of Nambu-Goto string networks [3, 18]. They are also produced by stacking maps from various redshifts [3, 18, 17]. Our simulated 100 square maps of side $\Theta = 7.2^\circ$ are well in the flat sky limit, with $1024 \times 1024$ pixels with the resolution of $R = 0.42'$ before convolution with an experimental beam. Figure 1 shows the power spectrum of various components together the corresponding real space fluctuations.

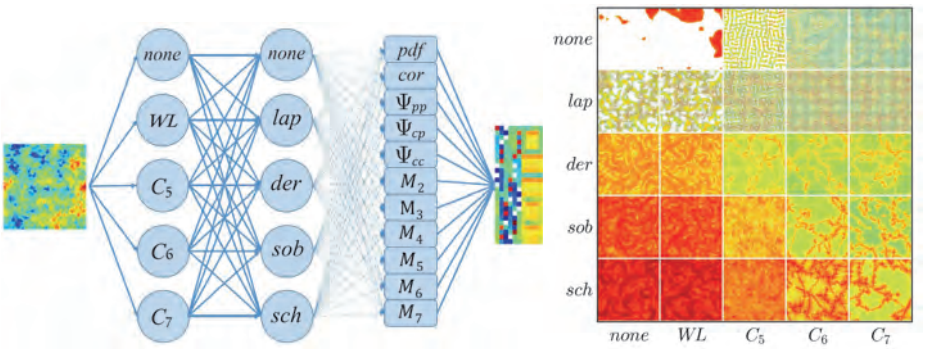



Fig. 2. Left panel: A schematic view of the feature vector generation. For a CMB map including a 275-dimensional feature vector, here presented as a $25 \times 11$ array (right side). Right panel: All of the 25 outputs of the image processing layers of the algorithm applied to a map with $G\mu = 1.0 \times 10^{-7}$. The color scale is logarithmic. These are then passed to the 11 statistical measures, yielding the full set of 275 features[22].

## 3. Image processing based methods

Our goal in this work is to evaluate the performance of various sequences of image-processing and statistical tools in the detection of the trace of possible CS networks on CMB maps. Our proposed pipeline comprises two major steps: (1) Processing CMB maps using multi-scaling method accompanying applying filters and (2) Using various statistical analysis of CMB map to quantify the detectability of CSs signature on the filtered maps from the first step.

### 3.1. *Multi-scale image processing*

We apply several image-processors with the aim to isolate or/and enhance the CSs imprint on CMB maps. The two pillars of this step are a multi-scaling analysis through curvelet-decomposition of the input maps and the generation of filtered maps through extended Canny algorithm. Curvelet decomposition and extended Canny algorithm are used to enhance the string detectability as indicated in Figure 2.

### 3.2. *Statistical measures*

In order to quantify the statistical measures, we start with probabilistic framework. The 2D CMB stochastic field is represented by a map $T \in \mathrm{L}^2(\mathbb{R}^2)$ (Eq. (1)). We also denote a vector variable, $\mathcal{A}$, at each spatial location on the CMB map as:

$$\{\mathcal{A}\} \equiv \{\delta_T, \eta_x, \eta_y, \xi_{xx}, \xi_{yy}, \xi_{xy}\}$$

where $\delta_T \equiv T$ (temperature fluctuation) for the CMB map. We define $\eta_x \equiv \partial\delta_T/\partial x$, $\eta_y \equiv \partial\delta_T/\partial y$ and $\xi_{xy} \equiv \partial^2\delta_T/\partial x\partial y$. The characteristic function of $\mathcal{A}$, is defined

by:

$$\mathcal{Z}(\lambda) = \int_{-\infty}^{+\infty} \mathrm{d}^6 \mathcal{A} \mathcal{P}(\mathcal{A}) e^{\boldsymbol{i}\lambda.\mathcal{A}}, \tag{2}$$

where $\lambda$ is an array with the same dimension as $\mathcal{A}$. Expansion form of $\mathcal{Z}$ becomes [11]:

$$\mathcal{Z}(\lambda) = \exp\left( -\frac{1}{2}\lambda^T.\mathcal{C}.\lambda + \left[ \sum_{j=3}^{\infty} \frac{\boldsymbol{i}^j}{j!} \left( \sum_{\mu_1}^{N} \sum_{\mu_2}^{N} ... \sum_{\mu_j}^{N} \mathcal{K}^{(j)}_{\mu_1,\mu_2,...,\mu_j} \lambda_{\mu_1}\lambda_{\mu_2}...\lambda_{\mu_j} \right) \right] \right). \tag{3}$$

where $\mathcal{K}^{(n)}_{\mu_1,\mu_2,...,\mu_n} \equiv \langle \mathcal{A}_{\mu_1}\mathcal{A}_{\mu_2}...\mathcal{A}_{\mu_n} \rangle_{\mathrm{c}}$ are the *connected* cumulants. Here $\mathcal{C} \equiv \langle \mathcal{A} \otimes \mathcal{A} \rangle$ represents the covariance matrix of $\mathcal{A}$ at each spatial point. Using the temperature power spectrum of CMB, the various spectral parameters read as:

$$\sigma_m^2 = \sum_{\ell} \frac{(2\ell+1)}{4\pi} \left[\ell(\ell+1)\right]^m C_\ell^{TT} W_\ell^2. \tag{4}$$

The smoothing kernel associated with beam transfer function is $W_\ell = \exp\left(-\theta^2_{\mathrm{beam}}\ell(\ell+1)/2\right)$ and $\theta_{\mathrm{beam}} = \theta_{\mathrm{FWHM}}/\sqrt{8\ln(2)}$ [4, 6, 7]. Also $C_\ell^{TT}$ is CMB temperature power spectrum.

Now, we have robust approaches in hand to deduce statistical measures applied on CMB map. Here we only notice to geometrical measures and for the rest part refer to [23, 22].

### 3.2.1. *The n-point joint PDF*

The joint probability density function (JPDF) can be calculated from the inverse Fourier Transform of the corresponding characteristic function according to:

$$\mathcal{P}(\mathcal{A}) = \frac{1}{(2\pi)^6} \int_{-\infty}^{+\infty} \mathrm{d}^6\lambda \mathcal{Z}(\lambda) e^{-\boldsymbol{i}\lambda.\mathcal{A}}. \tag{5}$$

Inserting Eq. (3) in Eq. (5) gives

$$\mathcal{P}(\mathcal{A}) = \exp\left[ \sum_{j=3}^{\infty} \frac{(-1)^j}{j!} \left( \sum_{\mu_1=1}^{6} ... \sum_{\mu_j=1}^{6} \mathcal{K}^{(j)}_{\mu_1,\mu_2,...,\mu_j} \frac{\partial^j}{\partial\mathcal{A}_{\mu_1}...\partial\mathcal{A}_{\mu_j}} \right) \right] \mathcal{P}_{\mathrm{G}}(\mathcal{A}), \tag{6}$$

where

$$\mathcal{P}_{\mathrm{G}}(\mathcal{A}) = \frac{1}{\sqrt{(2\pi)^6|\mathcal{C}|}}\, e^{-\frac{1}{2}(\mathcal{A}^T.\mathcal{C}^{-1}.\mathcal{A})}. \tag{7}$$

The probability density function for the CMB temperature induced by CS network deviates from Gaussian distribution depending on the value of $G\mu$ [16].

3.2.2. *The weighted TPCF*

The weighted two-point correlation function (TPCF) which is defined as $\mathcal{C}_{TT} = \langle \delta_T(\mathbf{r}_1)\delta_T(\mathbf{r}_2)\rangle$ is another statistical measure used for searching the footprint of CS. Here $\mathbf{r}_1$ and $\mathbf{r}_2$ represent the coordinates of the points.

3.2.3. *The unweighted TPCF*

The unweighted TPCF of a certain feature is referred to the its excess probability. Here, we compare the clustering of the local maxima and up-crossing of CMB maps for Gaussian-only fluctuations with those affected by the CS network as well. The excess probability of finding two different features represented by symbols, $\oplus$ and $\otimes$ is $\Psi_{\oplus-\otimes}$ separated by distance $r = |\mathbf{r}_1 - \mathbf{r}_2|$, at thresholds $\vartheta_1 \equiv \alpha_1/\sigma_0$ and $\vartheta_2 \equiv \alpha_2/\sigma_0$ is defined as:

$$\Psi_{\oplus-\otimes}(r;\vartheta_1,\vartheta_2) = \frac{\langle n_{\oplus}(\mathbf{r}_1,\vartheta_1) n_{\otimes}(\mathbf{r}_2,\vartheta_2)\rangle}{\bar{n}_{\oplus}(\vartheta_1)\bar{n}_{\otimes}(\vartheta_2)} - 1. \tag{8}$$

The mathematical description of local maxima number density, $\bar{n}_{\rm pk}(\vartheta)$, reads:

$$\bar{n}_{\rm pk}(\vartheta) = \langle \delta_{\rm d}(\delta_T - \vartheta\sigma_0)\delta_{\rm d}(\eta)\,|\det(\xi)|\rangle\,. \tag{9}$$

The second derivative tensor of the CMB field ($\xi_{ij}$) should be *negative definite* at peak position. Its analytical expression for a 2D homogenous Gaussian field was calculated in [2, 4]. While the mean number of up-crossings of a 2D CMB field $\delta_T$ at a given threshold $\vartheta$ for an arbitrary 1D slice denoted by $\diamond$ is given by [15, 20, 19, 10, 12, 11, 14, 13]:

$$\bar{n}^{\diamond}_{\rm up}(\vartheta) = \langle \delta_{\rm d}(\delta_T - \vartheta\sigma_0)\Theta(\eta^{\diamond})\,|\eta^{\diamond}|\rangle\,, \tag{10}$$

where $\Theta(\eta)$ is the unit step function and $|\eta^{\diamond}|$ is the absolute value of the first derivative of temperature fluctuations in direction $\diamond$ [11]. One can choose any direction $\diamond$ on the statistically isotropic CMB map, without loss of generality. To reduce the boundary effect as much as possible in computing excess probability, $\tilde{\Psi}_{\oplus-\otimes}(r;\vartheta)$, following estimator has been proposed [9]:

$$\tilde{\Psi}_{\oplus-\otimes}(r;\vartheta) = \left[\frac{D_{\oplus}D_{\otimes}(r,\vartheta)}{R_{\oplus}R_{\otimes}(r,\vartheta)}\right]\frac{N_{R_{\oplus}}N_{R_{\otimes}}}{N_{D_{\oplus}}N_{D_{\otimes}}} - 1. \tag{11}$$

Here, $RR(r,\vartheta)$ and $DD(r,\vartheta)$ are the number of feature pairs in random and data catalogs, respectively, separated by distance $r$ from each other. Similarly, $N_D$ and $N_R$ are the total number of feature in data and random catalogs, respectively.

We utilize above statistical measures for the third layer to make feature vector (see Figure 2). Therefore, we compare the results with null hypothesis. On noiseless sky maps with an angular resolution of $0.9'$, we show that our pipeline detects CSs with $G\mu$ as low as $G\mu \gtrsim 4.3\times10^{-10}$. At the same resolution, but with a noise level typical to a CMB-S4 phase II experiment, the detection threshold would increase to $G\mu \gtrsim 1.2\times10^{-7}$ [23].

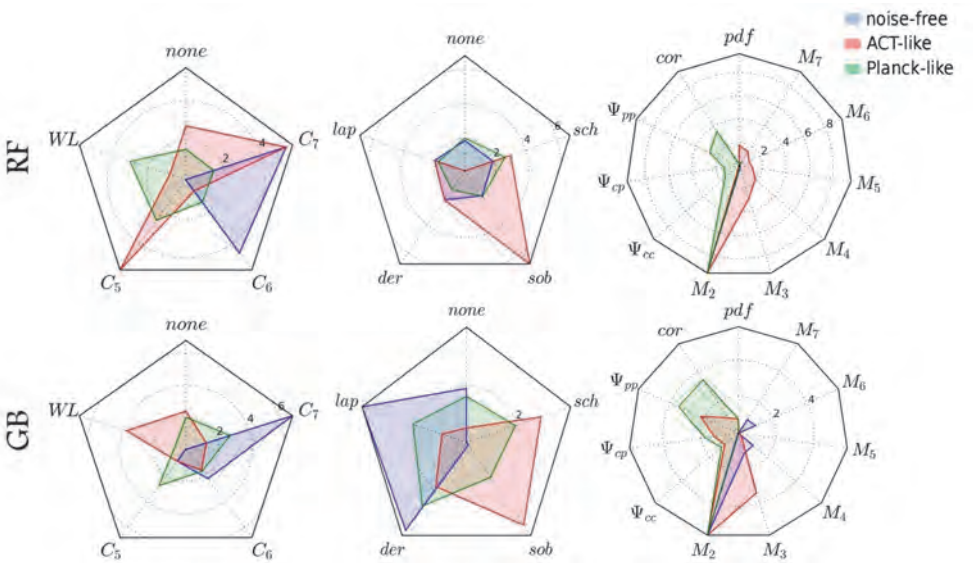



Fig. 3. Feature importance report: the average number of times each pre-processing tool appeared among the top ten features, for each layer of the pre-processor, for the RF (top) and GB (bottom) learner [22].

## 4. Machine-learning approach

The purpose is to develop a detection strategy through optimally exploiting the available information accessible to the multi-scale pipeline presented by Vafaei et al. [23]. We use two tree-based supervised classifiers: random forest (RF) and gradient boosting (GB) [22]. The information in the maps is compressed into feature vectors before being passed to the learning units. The feature vectors contain various statistical measures of the processed CMB maps that boost cosmic string detectability (see Figure 2). Our proposed classifiers, after training, give results similar to or better than claimed detectability levels from other methods for string tension. Our results confirm that the minimum detectable $G\mu$ for a noise-free experiment is $2.1 \times 10^{-10}$. This bound is, to the best of our knowledge, below the claimed detectability levels by other methods on noise-less maps. The minimum detectable tension in this work for a CMB-S4-like (II) experiment, $G\mu \gtrsim 3.0 \times 10^{-8}$, is a major improvement over the claimed detectability level by the above multi-scale pipeline, $G\mu \gtrsim 1.2 \times 10^{-7}$. For a Planck-like case, the minimum detectable $G\mu$ is $5 \times 10^{-7}$, comparable to the current upper bounds from Planck data [1]. Figure 3 compares the feature importance of the most significant features in each of the three pre-processing layers for noise-free, ACT-like and Planck-like cases, and for the two tree-based algorithms considered in this work, namely RF and GB. Among the 275 features, we choose the ten most significant ones by comparing the number of their occurrences as the splitting feature in all trees of a MLM. The plot shows the

number of occurrences, averaged over all MLMs, of the pre-processing tools among these top ten features, separated by their parent layer, i.e., decomposers, filters and statistical measures. Final remark is that, the scale of curvelet components should be matched to the effective resolution of experiments in the presence of experimental noise, larger-scale curvelet components are the more important decomposers. For filters it is difficult to make a definite recommendation, while the second moment is the most important statistical measure in the classification process.

## 5. Conclusion

In this review, we introduced multiscaling approach followed by some excursion and critical sets in order to explore the footprint of CSs network superimposed on the CMB simulation maps. Particularly, we focused on this question: What is the minimum value of CS's tension which could be recognized in the CMB map? In order to answer this question, at first we proposed a pipeline and accordingly, the minimum detectable value in noiseless CMB map is $G\mu > 4.3\times10^{-10}$ while applying more advanced searching method, namely machine learning approach, enables us to find CS with $G\mu > 2.1 \times 10^{-10}$. In this algorithm, we are also able to evaluate the significance of each pre-processors in detecting algorithm for different CMB observational strategies.

## References


1. Ade, P. A. R., et al. Planck 2013 results. XXV. Searches for cosmic strings and other topological defects. *Astron. Astrophys. 571* (2014), A25.
2. Bardeen, J. M., Bond, J. R., Kaiser, N., and Szalay, A. S. The Statistics of Peaks of Gaussian Random Fields. *Astrophys. J. 304* (1986), 15–61.
3. Bennett, D. P., and Bouchet, F. R. High-resolution simulations of cosmic-string evolution. I. Network evolution. *Phys. Rev. D41* (Apr. 1990), 2408–2433.
4. Bond, J. R., and Efstathiou, G. The statistics of cosmic background radiation fluctuations. *Mon. Not. Roy. Astron. Soc. 226* (1987), 655–687.
5. Hammond, D. K., Wiaux, Y., and Vandergheynst, P. Wavelet domain Bayesian denoising of string signal in the cosmic microwave background. *Mon. Not. Roy. Astron. Soc. 398* (2009), 1317.
6. Heavens, A. F., and Sheth, R. K. The correlation of peaks in the microwave background. *Monthly Notices of the Royal Astronomical Society 310*, 4 (1999), 1062–1070.
7. Hikage, C., Komatsu, E., and Matsubara, T. Primordial non-gaussianity and analytical formula for minkowski functionals of the cosmic microwave background and large-scale structure. *The Astrophysical Journal 653*, 1 (2006), 11.
8. Kaiser, N., and Stebbins, A. Microwave Anisotropy Due to Cosmic Strings. *Nature 310* (1984), 391–393.
9. Landy, S. D., and Szalay, A. S. Bias and variance of angular correlation functions. *Astrophys. J. 412* (1993), 64.

10. Matsubara, T. Statistics of isodensity contours in redshift space. *Astrophys. J. 457* (1996), 13–17.
11. Matsubara, T. Statistics of Smoothed Cosmic Fields in Perturbation Theory. 1. Formulation and Useful Formulae in Second Order Perturbation Theory. *Astrophys. J. 584* (2003), 1–33.
12. Movahed, S., and Khosravi, S. Level Crossing Analysis of Cosmic Microwave Background Radiation: A method for detecting cosmic strings. *JCAP 1103* (2011), 012.
13. Musso, M., and Sheth, R. K. On the Markovian assumption in the excursion set approach: the approximation of Markovian Velocities. *Mon. Not. Roy. Astron. Soc. 443*, 2 (2014), 1601–1613.
14. Musso, M., and Sheth, R. K. The importance of stepping up in the excursion set approach. *Mon. Not. Roy. Astron. Soc. 438*, 3 (2014), 2683–2693.
15. Rice, S. O. Mathematical analysis of random noise. *Bell Labs Technical Journal 23*, 3 (1944), 282–332.
16. Ringeval, C. Cosmic strings and their induced non-Gaussianities in the cosmic microwave background. *Adv. Astron. 2010* (2010), 380507.
17. Ringeval, C., and Bouchet, F. R. All Sky CMB Map from Cosmic Strings Integrated Sachs-Wolfe Effect. *Phys. Rev. D86* (2012), 023513.
18. Ringeval, C., Sakellariadou, M., and Bouchet, F. Cosmological evolution of cosmic string loops. *JCAP 0702* (2007), 023.
19. Ryden, B. S. The collapse of density peaks in a gaussian random field. *The Astrophysical Journal 333* (1988), 78–89.
20. Ryden, B. S., Melott, A. L., Craig, D. A., Gott, III, J. R., Weinberg, D. H., Scherrer, R. J., Bhavsar, S. P., and Miller, J. M. The Area of Isodensity Contours in Cosmological Models and Galaxy Surveys. *Astrophys. J. 340* (1989), 647–660.
21. Stebbins, A. Cosmic strings and the microwave sky. I - Anisotropy from moving strings. *Astrophys. J. 327* (Apr. 1988), 584–614.
22. Vafaei Sadr, A., Farhang, M., Movahed, S., Bassett, B., and Kunz, M. Cosmic string detection with tree-based machine learning. *Monthly Notices of the Royal Astronomical Society 478*, 1 (2018), 1132–1140.
23. Vafaei Sadr, A., Movahed, S., Farhang, M., Ringeval, C., and Bouchet, F. A multiscale pipeline for the search of string-induced cmb anisotropies. *Monthly Notices of the Royal Astronomical Society 475*, 1 (2017), 1010–1022.

# Evidence of cosmic strings by the observation of the alignment of quasar polarization axes on Mpc scale

Reinoud J. Slagter*

*ASFYON, Astronomisch Fysisch Onderzoek Nederland, Bussum, The Netherlands*
*and*
*Department of Physics, University of Amsterdam, The Netherlands*
**E-mail: info@asfyon.com*
*www.asfyon.com*

Topological defects formed in the early stages of our universe can play a crucial role in understanding anisotropic deviations of the Friedmann Lemaître Robertson Walker model we observe today. These defects are the result of phase transitions associated with spontaneous symmetry breaking in gauge theories at the grand unification energy scale. The most interesting defects are cosmic strings, vortex-like structures in the famous gauged U(1) abelian Higgs model with a "Mexican-hat" potential. Other defects, such as domain walls and monopoles are probably ruled out, because they should dominate otherwise the energy density of our universe. This local gauge model is the basis of the standard model of particle physics, where the Higgs-mechanism provides elementary particles with mass. It cannot be a coincidence that this model also explains the theory of superconductivity. The decay of the high multiplicity ($n$) super-conducting vortex into a lattice of $n$ vortices of unit magnetic flux is energetically favourable and is experimentally confirmed. It explains the famous Meissner effect. This process could play an essential role by the entanglement of cosmic strings just after the symmetry breaking. The stability of the lattice depends critically on the parameters of the model, especially when gravity comes into play. The questions is how the imprint of the cosmic strings could be observed at present time. Up to now, no evidence is found. The recently found alignment of the spinning axes of quasars in large quasar groups on Mpc scales, could be a first indication of the existence of these cosmic strings. The temporarily broken axial symmetry will leave an imprint of a preferred azimuthal-angle on the lattice. This effect is only viable when a scaling factor is introduced. This can be realized in a warped five dimensional model. The warp factor plays the role of a dilaton field on an equal footing with the Higgs field. The resulting field equations can be obtained from a conformal invariant model. Conformal invariance, the missing symmetry in general relativity, will then spontaneously be broken, just as the Higgs field. The dilaton field, or equivalently, the warp factor, could also contribute to the expansion of the universe as it can act as a dark energy term coming from the bulk spacetime. It makes the cosmic string temporarily super-massive. This process could solve the cosmological constant and hierarchy problem. It is conjectured that the dilaton field has a dual meaning. At very early times, when the dilaton field approaches zero, it describes the small-distance limit of the model, while at later times it is a warp (or scale) factor that determines the dynamical evolution of the universe. When more data of quasars of high redshift will become available, one could prove that the alignment emerged after the symmetry breaking scale and must have a cosmological origin. The effect of the warp factor on the second-order perturbations could also be an indication of the existence of large extra dimensions.

*Keywords*: Quasar alignment, Brane world models, U(1) scalar-gauge field, Cosmic strings, Dilaton field, Multiple-scale analysis.

## 1. Introduction

The standard model (SM) of the electroweak and strong interactions is a successful framework in which one studies elementary particles and includes the principles of quantum mechanics (QM). On the other hand, general relativity (GR) is also a very impressive theory constructed by theoretical physicists. It describes large scale structures in our universe and one can construct solutions which are related to real physical objects, for example the Kerr solution, the end stage of a collapsing spinning star. A legitimate question is if there are other axially symmetric solutions in GR. It came as a big surprise that there exist vortex-like solutions in Einstein's theory. These vortex solutions occur as topological defects at the symmetry breaking scale in the Einstein-abelian U(1) scalar-gauge model, where the gauge field is coupled to a complex charged scalar field [1,2]. The solution shows a surprising resemblance with type II superconductivity of the Ginzburg-Landau (GL) theory, where the electro-magnetic(EM) gauge invariance is broken and the well-known Meissner effect occurs [3]. One says that the phase symmetry is spontaneously broken and the EM field acquires a length scale, which introduces a penetration depth of the gauge field $A_\mu$ in the superconductor and a coherence length of $\Phi$. In the relativistic case one says that the photon acquires mass. Because we have three space dimensions, these solutions of the GL theory behave like magnetic flux vortices (or Nielsen-Olesen (NO) strings [2]) extended to tubes and carry a quantized magnetic flux $2\pi n$, with $n$ an integer, the topological charge or winding number of the field. It was discovered by Abrikosov [3] that these vortices can form a lattice. These localized vortices (or solitons) in the GL-theory are observed in experiments. The phenomenon of magnetic flux quantization in the theory of superconductivity is characteristic for so-called ordered media. The vortex solution possesses mass, so it will couple to gravity. The resulting self-gravitating cosmic strings (CS) still show all the features of superconductivity, but the stability conditions complicate considerably. The stability of the formed lattices depends critically on the parameters of the model, certainly when gravity comes into play. The force between the gauged vortices depends on the the strength of the self-interaction potential of the Higgs field, the gauge-coupling constant, the energy scale at which the phase transition takes place and the spacetime structure. When the mass of the Higgs field is greater than the mass of the gauge field, vortices will repel each other. So gravity could balance the vortices. The energy of the vortex grows by increasing multiplicity $n$, so configurations with $n > 1$ can be seen as multi-soliton states and it is energetically favourable for these to decay into $n$ well separated $n = 1$ solitons. Vortices with high multiplicity can be formed during the symmetry breaking. The total vortex number $n$ is the sum of multiplicities $n_1, n_2, ..$ of isolated points (zero's of $\Phi$). Our universe, described by a spatially homogeneous and isotropic Friedmann Lemaître Robertson Walker (FLRW) spacetime, shows significant large-scale inhomogeneous structures, for example, the cosmic web of voids with galaxies and clusters in sheets, filaments and knots, the angular distribution in the cosmic microwave background (CMB)

radiation and the recently found alignment of polarization axes of quasars in large quasar groups (LQG's) on Mpc-scales[4,5]. The question is if these complex nonlinear structures of deviation from isotropy and homogeneity have a cosmological origin at a moment in the early stage of the universe. One possibility of this origin could be a CS-network formed by the self-gravitating Einstein-scalar-gauge model. A pleasant fact is that this model has very few parameters and hence more appealing than other models such as inflationary models. It is believed that the mass per unit length of the CS is of the order of the GUT scale, $G\mu \approx 10^{-7}$. Observational bounds, however, predict a negligible contribution of CS's to initial density perturbation from which galaxies and clusters grew. Besides the inconsistencies with the power spectrum of the CMB, radiative effects of the CS embedded in a FLRW spacetime are rapidly damped in any physical regime[6]. Further, the lensing effect of these CS's are not found yet. There is, however, another possibility to detect the presence of CS's. On a warped spacetime, the fields can become temporarily super-massive by the warp factor in the framework of string theory (or M-theory). Naively one expect that gravity will play a subordinate role compared with the other fields. In 4D counterpart models this is true, but not in warped spacetimes. The super-massive CS's can be formed at a symmetry breaking scale much higher than the GUT scale, i.e., $G\mu >> 1$. So their gravitational impact increases considerably, because the CS builds up a huge mass in the bulk space. Here we consider the warped brane world model of Randall-Sundrum (RS)[7,8], with one large extra dimension. The result is that effective 4D Kaluza-Klein (KK) modes are obtained from the perturbative 5D graviton. These KK modes will be massive from the brane viewpoint. The modified Einstein equations on the brane and scalar gauge field equations will now contain contributions from the 5D Weyl tensor[9]. In order to explore these effective field equations, we apply an approximation scheme, i.e., a multiple scale method(MSM). In this method one can handle the decay of the $n$-vortex in a perturbative way. The MSM or high-frequency method is an approved tool to handle nonlinearities and secular terms arising in the partial differential equations (PDE) in GR. When there is a high curvature situation, a linear approximation of the Einstein equations is not suitable[10,11]. Other issue related to our 5D warped spacetime is the behavior at small scales, i.e., when the warp factor or scale factor of the spacetime becomes very small. We conjecture that our warp factor becomes the dilaton field which is needed to make the Lagrangian conformal invariant. Breaking of the conformal symmetry (which will occur when other fields come into play; after all we experience today a huge discrepancy in scales), can be compared with the Brout-Englert-Higgs Mechanism (BEH) in the standard model of particle physics.

## 2. The superconducting string model in warped spacetime

### 2.1. *Outline of the model*

Our model will be based on a warped five-dimensional FLRW spacetime

$$ds^2 = \mathcal{W}(t,r,y)^2\Big[e^{2(\gamma(t,r)-\psi(t,r))}(-dt^2+dr^2)+e^{2\psi(t,r)}dz^2+r^2e^{-2\psi(t,r)}d\varphi^2\Big]+dy^2, \tag{1}$$

with $\mathcal{W}=W_1(t,r)W_2(y)$ is the warp factor. All standard model fields reside on the brane, while gravity can propagate into the bulk. An extended treatment of the warped five-dimensional spacetime and the effective gravitational equations can be found in Shiromizu et al.[9] and Slagter[12]. The 5D Einstein equations are

$${}^{(5)}G_{\mu\nu} = -\Lambda_5\,{}^{(5)}g_{\mu\nu}+\kappa_5^2\delta(y)\Big(-\Lambda_4\,{}^{(4)}g_{\mu\nu}+{}^{(4)}T_{\mu\nu}\Big), \tag{2}$$

with $\kappa_5 = 8\pi^{(5)}G = 8\pi/{}^{(5)}M_{pl}^3$, $\Lambda_4$ the brane tension, ${}^{(4)}g_{\mu\nu} = {}^{(5)}g_{\mu\nu}-n_\mu n_\nu$, and $n^\mu$ the unit normal to the brane. The effective 4D Einstein-Higgs-gauge field equations are

$${}^{(4)}G_{\mu\nu} = -\Lambda_{eff}\,{}^{(4)}g_{\mu\nu}+\kappa_4^2\,{}^{(4)}T_{\mu\nu}+\kappa_5^4\mathcal{S}_{\mu\nu}-\mathcal{E}_{\mu\nu}, \tag{3}$$

$$D^\mu D_\mu\Phi = 2\frac{dV}{d\Phi^*}, \qquad {}^{(4)}\nabla^\mu F_{\nu\mu} = \frac{1}{2}ie\Big(\Phi(D_\nu\Phi)^*-\Phi^*D_\nu\Phi\Big), \tag{4}$$

with $D_\mu\Phi \equiv {}^{(4)}\nabla_\mu\Phi+ieA_\mu\Phi$, ${}^{(4)}\nabla_\mu$ the covariant derivative with respect to ${}^{(4)}g_{\mu\nu}$, $V(\Phi)=\frac{1}{8}\beta(\Phi\Phi^*-\eta^2)^2$ the potential of the Abelian Higgs model and $\eta$ the symmetry breaking scale. $F_{\mu\nu}$ is the Maxwell tensor. The righthand side of the Einstein equations contains a contribution $\mathcal{E}_{\mu\nu}$ from the 5D Weyl tensor and carries information of the gravitational field outside the brane. The quadratic term in the energy-momentum tensor, $\mathcal{S}_{\mu\nu}$, arising from the extrinsic curvature terms in the projected Einstein tensor. ${}^{(4)}T_{\mu\nu}$ represents the matter content on the brane, in our case the scalar and gauge fields[9,12]. These terms are responsible for the effective 4D modes, i.e., KK-modes, of the perturbative 5D graviton. We parameterize the self-gravitating scalar gauge field as

$$\Phi = \eta X(t,r)e^{in\varphi}, \qquad A_\mu = \frac{n}{e}\big[P(t,r)-1\big]\nabla_\mu\varphi, \tag{5}$$

with n the topological charge or winding number of the scalar field. The warp factor can be solved from the 5D Einstein equations:

$$\mathcal{W} = W_2(y)W_1(t,r) = \frac{\pm e^{\sqrt{-\frac{1}{6}\Lambda_5}(y-y_0)}}{\sqrt{\tau r}}\sqrt{\Big(d_1e^{(\sqrt{2\tau})t}-d_2e^{-(\sqrt{2\tau})t}\Big)\Big(d_3e^{(\sqrt{2\tau})r}-d_4e^{-(\sqrt{2\tau})r}\Big)}, \tag{6}$$

On the one hand, the constants $(d_i,\tau)$ will be fixed at late time by the features of the FLRW-model. They will appear in the modified second Friedmann equation. On the other hand, on the very small scale, the warp factor equals the dilaton field in conformal invariant models and these constant will be determined by constraint equations.

### 2.2. *The approximation*

In order to study perturbations in the model, one can apply the multiple-scale approximation, very suited at high curvature situations [10,11]. The method is called a "two-timing" method, because one considers the relevant fields $V_i$ in point $\mathbf{x}$ on a manifold M dependent on different scales $(\mathbf{x}, \xi, \chi, ...)$:

$$V_i = \sum_{n=0}^{\infty} \frac{1}{\omega^n} F_i^{(n)}(\mathbf{x}, \xi, \chi, ...). \tag{7}$$

Here $\omega$ represents a dimensionless parameter, which will be large (the "frequency", $\omega >> 1$). So $\frac{1}{\omega}$ is a small expansion parameter. Further, $\xi = \omega\Theta(\mathbf{x})$, $\chi = \omega\Pi(\mathbf{x}), ...$ and $\Theta, \Pi, ...$ scalar (phase) functions on M. The parameter $\frac{1}{\omega}$ can be the ratio of the characteristic wavelength of the perturbation to the characteristic dimension of the background. On warped spacetimes it could also be the ratio of the extra dimension y to the background dimension or even both. When one substitutes the expansions of the field variables

$$\begin{aligned} g_{\mu\nu} &= \bar{g}_{\mu\nu}(\mathbf{x}) + \frac{1}{\omega} h_{\mu\nu}(\mathbf{x}, \xi) + \frac{1}{\omega^2} k_{\mu\nu}(\mathbf{x}, \xi) + ..., \\ A_\mu &= \bar{A}_\mu(\mathbf{x}) + \frac{1}{\omega} B_\mu(\mathbf{x}, \xi) + \frac{1}{\omega^2} C_\mu(\mathbf{x}, \xi) + ..., \\ \Phi &= \bar{\Phi}(\mathbf{x}) + \frac{1}{\omega} \Psi(\mathbf{x}, \xi) + \frac{1}{\omega^2} \Xi(\mathbf{x}, \xi) + ..., \end{aligned} \tag{8}$$

into the equations, one obtains first order equations in $u = t - r$ for the first and second order perturbations [13,14]. They are of the form

$$\partial_u \dot{\mathbf{U}}_1 = \bar{\mathbf{A}}, \qquad \partial_u \dot{\mathbf{U}}_2 = \bar{D}_1 \dot{\mathbf{U}}_2 + D_2 \dot{\mathbf{U}}_1 + D_3 \tag{9}$$

where $\bar{A}$ and $\bar{D}_1$ depends solely on the background fields, while $D_2, D_3$ depend on the first order perturbations and background fields. The dot represents differentiation with respect to the rapid scale $\xi$. In principle one could push the approximation to higher orders. In this way one obtains a wavelike approximation which is asymptotically finite. To highest order in $\omega$, one obtains the equations $l^\alpha(\ddot{h}_{\alpha\nu} - \frac{1}{2}\bar{g}_{\alpha\nu}\ddot{h}) = 0, l_\alpha l^\alpha \ddot{\Psi} = 0$ and $l^\alpha \ddot{B}_\alpha = 0$, where $l_\alpha \equiv \frac{\partial\Theta}{\partial x^\mu}$ is the wave vector perpendicular to the hyper-surface. These equations deliver constraints on $h, B$ and $C$. In other approximations, they are a priori used as gauge conditions. Moreover, the original symmetry on the gauge field will be broken by the appearance of $B_t$ besides $B_\varphi$.

In our approximation scheme we can gain a lot of insight in the behavior of the clustering of vortices when gravity is present. The equations (Eq. (9)) are hard to solve. However, the energy-momentum tensor components can tell us a lot about the behavior of the model.

### 2.3. *Excitation of the vortices and the quasar link*

In the expansion of Eq. (8) we parameterized the scalar field in subsequent orders as[13,14]

$$\bar{\Phi} = \eta \bar{X}(t,r)e^{in_1\varphi}, \qquad \Psi = Y(t,r,\xi)e^{in_2\varphi}, \qquad \Xi = Z(t,r,\xi)e^{in_3\varphi}. \tag{10}$$

It turns out that the solution to second order is no longer axially symmetric. There appear terms like $\sin(n_2 - n_1)\varphi$ in the field equations. The most interesting information can be found in the energy-momentum tensor components

$${}^4T^{(0)}_{t\varphi} = \bar{X}\bar{P}\dot{Y}n_1 \mathbf{sin}[(n_2 - n_1)\varphi], \tag{11}$$

$${}^4T^{(0)}_{tt} = \dot{Y}^2 + \dot{Y}(\partial_t \bar{X} + \partial_r \bar{X})\mathbf{cos}[(n_2 - n_1)\varphi] + \frac{e^{2\bar{\psi}}}{\bar{W}_1^2 r^2 e}\Big(e\dot{B}^2 + n_1 \dot{B}(\partial_r \bar{P} + \partial_t \bar{P})\Big), \tag{12}$$

While ${}^4\bar{T}_{t\varphi} = 0$, we conclude from Eq. (12) that the axial symmetry is broken already to first order. The energy ${}^4T^{(0)}_{tt}$ contribution to first order contains the warp factor in the denominator. So the energy depends crucially on the age of the universe. Terms in the energy can dominate at early times and are negligible at late times in the evolution of the universe. In the second order contributions there appear terms like $\cos(n_3 - n_1)\varphi$. The azimuthal-angle dependency are expressed in trigonometrical functions with extrema which differ $mod(\frac{\pi}{k})$. After the excitation of the vortex with multiplicity n, it will decay into n vortices of unit flux in a regular lattice (figure 1).

There is another characterization of the winding number. It is the total vortex number, i.e., the number of points in the plane with multiplicity taken into count where $\Phi = 0$. The zero's of $\Phi$ are then a set of n isolated points $z_i, i = 0..n$ in $\mathbb{C}$ such that $\Phi(z, z^*) \sim c_j(z - z_j)^{n_j}$ with $n_j$ the multiplicity of $z_j$ and $n = \sum_{z_j} n_j$.

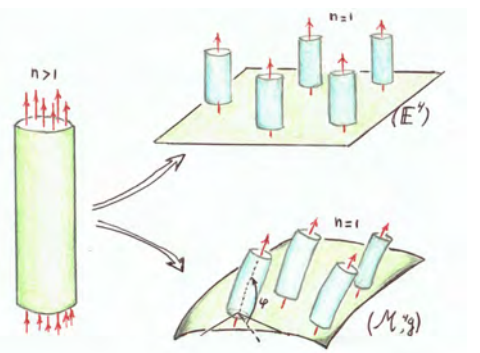



Fig. 1. Excitation and decay of a high multiplicity vortex into correlated vortices of unit flux $n = 1$. Top: the Abrikosov lattice in Euclidean space. Bottom: correlated vortices with preferred azimuthal angle $\varphi$ in curved spacetime after the symmetry breaking.

This n-vortex solution represents a finite energy configuration with n flux quanta, provided $\Phi$ and $A$ satisfy the boundary conditions

$$\lim_{r\to 0} X = 0, \quad \lim_{r\to 0} P = 1, \quad \lim_{r\to\infty} X = 1, \quad \lim_{r\to\infty} P = 0. \tag{13}$$

The collection of n vortices of unit flux is energetically more appealing than a n-flux vortex. The energy density is peaked around the zero's of $\Phi$. Hence they can be identified by the location of the vortices. In general, one must solve the time-dependent GL equations in order to get insight in the stability issues. In this case there is a gradient flow, which makes the analysis very complicated. The PDE's are badly nonlinear and one relies often on numerical simulation. The temporarily broken axial symmetry will be the onset of emission of electro-magnetic and gravitational waves. From Eq. (11) we see that the angular momentum will fade away when $n_2$ approaches $n_1$ and the axial symmetry is restored. The first and second order perturbations of the scalar and gauge fields in higher winding number decay into NO strings of n=1. In order to understand the azimuthal-angle ($\varphi$) preference, one must consider the terms of ${}^4T_{\varphi\varphi}$, for example

$${}^4T^{(0)}_{\varphi\varphi} = e^{-2\gamma} r^2 \dot{Y}(\partial_t \bar{X} - \partial_r \bar{X})\mathbf{cos}[(n_2 - n_1)\varphi] + \frac{n_1 e^{2\bar{\psi}-2\bar{\gamma}}}{\bar{W}_1^2 e}\dot{B}(\partial_r \bar{P} - \partial_t \bar{P}), \tag{14}$$

$T_{\varphi\varphi}$ plays an important role in the interaction of the strings. Positive terms in the expression indicate "pressure" in the direction of the Killing vector field $(\frac{\partial}{\partial\varphi})^i$ (and negative "tension"). The result is that the interaction contribution can change sign dynamically (dependent of the warp factor). This Killing vector must be normalized such that, along a closed integral curve, the parameter $\varphi$ varies van 0 to $2\pi$ with $\varphi = 0$ and $\varphi = 2\pi$ identified. This will provide boundary conditions for the metric fields close to the axis of the string, such as $\partial_r(r^2 e^{-2\psi})(0) = 1$. We observe in the expressions of ${}^4T^{(0)}_{\varphi\varphi}$ and ${}^4T^{(0)}_{t\varphi}$ that when $sin(n_2 - n_1)\varphi$ becomes zero, $cos(n_2 - n_1)\varphi$ has its maximum. So there is an emergent imprint of a preferred azimuthal angle $\varphi$ on the lattice of vortices when the ground state is reached ($n = 1$). This effect can also be seen in the ${}^4T^{(0)}_{zz}$ component which is not equal to $-{}^4T^{(0)}_{tt}$ as is the case in static models. The second order contribution ${}^4T^{(1)}_{\varphi\varphi}$ contains terms like $cos(n_3 - n_1)\varphi$ and produces a complicated extremum[14]. The recently observed alignment of the spinning axes of quasars in LQG's on Mpc-scales can be explained by our model. The observations were carried out at the European Southern Observatory, Paranal with the Very Large Telescope equipped with the FORS2 instrument. There was a confirmation of the alignment for radio galaxies by the Giant Metrewave Radio Telescope in the ELAIS-N1 field. This curious effect cannot be the result of statistical fluctuations[16]. The origin must be found in the early universe just after the symmetry breaking, as described in our model. Specially, the two preferred orientations perpendicular to each other in quasar groups of less richness could be the second order effect in our model by the appearance of the trigonometrical terms with periodicity difference of $\frac{\pi}{2}$. The correlated $n = 1$ vortices with preferred

azimuthal angle, emerged on a correlation length smaller than the horizon on that moment and took place at the Ginzburg temperature $\sim \frac{1}{\beta\eta}$. These correlated regions will survive to later times, because at this moment the gravity contribution from the 5D bulk comes into play. The warp factor will have different contributions to the field equations for different times. The mass per unit length will contain the warp factor. Just after the symmetry breaking, the vortex will acquire a huge mass $G\mu > 1$ and will initiate the perturbations of high-frequency and justifies our high-frequency approximation. This is the reason that the regions with $(n = 1, \varphi = \varphi_0)$ will stick together and are observed in LQG's with aligned polarization axes [4,5]. The most striking observation is the alignment of the spin axes of the quasars parallel to the major axes of their host LQG, while the spin axes can also become perpendicular to the the LQG major axes when the richness decreases. This can be explained in our model as second order effect: the higher multiplicity terms like $\cos(n_3 - n_1)\varphi$ enters the energy-momentum tensor component ${}^4T^{(1)}_{\varphi\varphi}$ and is out phase with the term $\cos(n_2 - n_1)\varphi$. In figure 2 we sketched a global distribution of the polarization axes as found by the observations. Some specific features of this alignment which must be confirmed by more observations on quasars and radio sources at high redshift. Alignment at high redshifts would confirm that the mechanism took place indeed in the early universe.

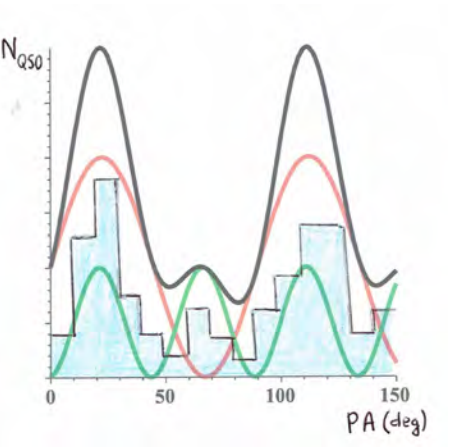



Fig. 2. Possible fit of the theoretical predicted curve (black) on the distribution of a number of observed polarization angles. This curve can be the sum of the first (red) and second (green) contribution.

### 2.4. *Breaking of the axial symmetry from a different viewpoint*

Self-gravitating objects in equilibrium exhibit a striking analogue with the mathematical model of the Maclaurin-Jacobi sequences and its bifurcation points [17]. Bifurcation points that are of particular interest to us here are those marking the

onset secular instability, i.e., the dynamical breaking of axially symmetry (or better formulated: the spacetime possesses 2 in stead of 3 Killing vectors). This means the appearance of an off-diagonal metric function. In our model it is the transition from the Weyl metric (after the substitution $t \to iz, z \to it$) to the Papapetrou metric, expressed by the appearance of the $T_{t\varphi}$ components in first and second order. It is remarkable that this symmetry breaking can be compared with the second order phase transition in type II superconductivity[18–20], which is the basis of our model (see also Slagter[14] for more details). An initial axially symmetric configuration, as is the case in our perturbative model, can dynamically spontaneously be broken, where equatorial eccentricity plays the role of order-parameter. The equatorial eccentricity $\varepsilon \equiv \frac{b}{a}$, with b and a the two equatorial axes, can be expressed through the azimuthal-angle $\varphi(t)$. The particular orientation of the ellipsoid in the frame $(r, \varphi, z)$ expressed through $\varphi_0 \equiv \varphi(t_0)$, will be at $t > t_0$ determined by the transformation $\varphi \to \varphi_0 - Jt$, where $J$ is the rotation frequency (circulation or "angular momentum") of the coordinate system. The angle $\varphi_0$ is fixed arbitrarily at the onset of symmetry breaking. This arbitrariness of $\varphi_0$, i.e., the orientation of the ellipsoid at $t = t_0$ can be compared with the massless Goldstone-boson modes of the spontaneously broken symmetry of continuous groups. The phase transition take place on the same time scale that the vorticity is destroyed by dissipative mechanism and $J$ is lost. The end point is a lower energy state that belongs to the Jacobi or Dedekind sequence of equilibrium ellipsoids[21]. In the original paper of Chandrasekhar and Lebovitz[18], in the Newtonian case, the deformations of the axisymmetric configuration by an infinitesimal non-axisymmetric deformation is described in terms of a Lagrangian displacement $\varsigma^a(r, z, \varphi) = \bar{\varsigma}^a(r, z)e^{in\varphi}$, with $n$ an integer. However, the real part of the $e^{in\varphi}$ must be put in by hand, in contrast to our result: it appears in a perturbative way as a first and second order effect. The temporarily broken axial symmetry will be the onset of emission of electro-magnetic and gravitational waves, while the string relaxes to the NO configuration. It is a consequence of the coupled system of PDE's that a high-frequency scalar field can create through an electro-magnetic field, a high frequency gravitational field and conversely. It is the appearance of the term $sin(n_2 - n_1)\varphi$ in the first order term ${}^4T^{(0)}_{t\varphi}$ (Eq.(11)) and explained in section 2c, which triggers this angular momentum and the axially symmetry will be restored when $n_2$ becomes equal to $n_1$ again. The second order contribution ${}^4T^{(1)}_{tt}$ shows terms like $\dot{B}^2 h_{44}$, indicating the interaction between the high-frequency EM and gravitational waves. It contains the warp factor in the denominator. In the early stages of the universe $W_1$ is still small and the term is significant. As time increases, it will fade away.

## 3. Conclusion

By considering an axially symmetric warped five-dimensional warped spacetime, were the standard model fields are confined to the brane, we find in a nonlinear approximation, an emergent azimuthal-angle dependency of Nielsen-Olesen vortices

just after the symmetry breaking at GUT scale. Using a approximation scheme, the azimuthal-angle dependency appears in the first and second order field equations as trigonometrical functions $\sin(n_i - n_j)\varphi$ and $\sin(n_i - n_j)\varphi(i > j)$, with $n_i$ the multiplicities of subsequent perturbation terms of the scalar field. Vortices with high multiplicity decay into a lattice with entangled Abrikosov vortices. The stability of this lattice of correlated flux $n = 1$ vortices with preferred azimuthal-angle is guaranteed by the contribution from the bulk spacetime by means of the warp factor: the cosmic string becomes super-massive for some time during the evolution. These so-called super-massive cosmic strings arise in a natural way in string theory or M-theory., i.e., brane-world models and are produced when the universe underwent phase transitions at energies much higher than the GUT scale, $G\mu > 1$. The imprint on the effective 4D brane spacetime can be caused by wavelike disturbances triggered by the huge mass of the cosmic strings in the bulk. One conjectures that these disturbances could act as an effective dark energy field. Another possible effect can be the formation of massive KK-modes as the imprint of the 5D gravitational field on the 4D brane.

We used the azimuthal-angle correlation for the explanation of the recently observed alignment of polarization axes of quasars in large quasar groups. The detailed behavior of this alignment can be explained with our model. The two different orientations perpendicular to each other in quasars groups of less richness could be a second order effect in our model.

When gravity is coupled to standard model fields and one demands the validity on all distance scales, one runs into problems. These are: the dark energy problem, the cosmological constant problem, the hierarchy problem and the problem how probe the small distance structure of our spacetime. Conformal invariance could be the solution for at least some of these problems. It could be the missing symmetry of nature. In our model, by identifying the warp factor as a dilaton field, one will not encounter singular behavior when the dilaton field becomes very small. At present time, the warp-like manifestation of the dilaton field describes the exponential expansion of our universe. Moreover, the exceptional smallness of the cosmological constant, $\Lambda \sim 10^{-120}$ compared to the calculated vacuum energy could be explained in the warped 5D spacetime by the warp factor. More data of high-redshift quasars will be needed in order to test the second order effect predicted in our model.

## References


1. Vilenkin, A. and Shellard, E. P. S. (1994) Cosmic Strings and Other Topological Defects. Cambridge University press, Cambrigde, UK.
2. Nielsen, H. B. and Olesen, P. (1973) *Nucl. Phys.* **B61**, 45.
3. Abrikosov, A. A. (1957) *Soviet Physics JETP* **5**, 1774.
4. Hutsemekers, D., Braibant, L., Pelgrims, V. and Sluse, D. (2014) *Astron. Astrophys.* **572**, A18.

5. Taylor, A. R. and Jagannathan, P. (2016) *Mon. Not. Roy. Astr. Soc.* **459**, *459*, L36.
6. Gregory, R. (1989) *Phys. Rev. D* **39**, 2108.
7. Randall, L. and Sundrum, R. (1999) *Phys. Rev. Lett.* **83**, 3370.
8. Randall, L. and Sundrum, R. (1999) *Phys. Rev. Lett.* **83**, 4690.
9. Shiromizu, T., Maeda, K. and Sasaki, M. (2000) *Phys. Rev. D* **62**, 024012.
10. Choquet-Bruhat, Y. (1977) *Gen. Rel. Grav.* **8**, 561.
11. Slagter, R. J. (1986) *Astroph. J.* **307**, 20.
12. Slagter, R. J. and Pan, S. (2016) *Found. of Phys.* **46**, 1075.
13. Slagter, R. J. (2016) *J. of Mod. Phys.* **7**, 501.
14. Slagter, R. J. (2017) *J. of Mod. Phys.* **8**, 163.
15. Bogomol'nyi, E. (1976) *Sov. J. Nucl. Phys.* **24**, 449.
16. Pelgrims, V. Thesis: ArXiv: astr-ph/160405141v1
17. Lebovitz, N. R. (1967) Bifurcation and stability problems in astrophysicsm, in Applications of bifurcation theory. Academic Press, New York, U.
18. Chandrasekhar, S. and Lebovitz, N. R. (1973) *Astroph. J* **185**, 19.
19. Chandrasekhar, S. and Friedman, J. L. (1973) *Astrophys. J* **185**, 1.
20. Bertin, G. and Radicati, L. A. (1976) *Astrophys. J* **206**, 815.
21. Christodoulou, D. M., Kazanas, D. Shlosman, I and Tohline, J. E. (1995) *Astrophys. J* **446**, 472.

# Linking dark energy and a Kaluza–Klein treatment of a graviton via the deceleration parameter $Q(z)$ by quintessence?

Andrew Walcott Beckwith

*Physics Department, Chongqing University, Chongqing 401331, People's Republic of China*
*E-mail: Rwill9955b@gmail.com; abeckwith@uh.edu*
*english.cqu.edu.cn*

We examine a dark-energy equation of state. [1] From there, we link early-universe graviton production by varying the DE equation from $-1$ to 1 with a Kaluza–Klein treatment of reacceleration and comment on how modification of gravitational $\frac{1}{r}$ potentials and the reacceleration of the universe also bridge between the first and second parts of this document, in terms DE physics and gravitons, as well as changing $Q(z)$ reacceleration behavior. We close with a suggestion on redoing initial contributions to DE.

*Keywords*: Massive Gravitons, Entropy, Kaluza Klein, Dark Matter, Dark Energy PACS: 98.80.-k, 98.80.Cq, 89.70 CF.

## 1. DE Equation of State, as a Starting Point

*If we quantize the gravitational field as an effective field, we find that it too comes in set quanta, called gravitons.* In short, we argue that to come up with a graviton-based DE model with reacceleration of the universe, and to have it coincide with the modification of the $1/r$ potential, we are really developing a program to find if gravity can be quantized. What we have done is also to specify a relic black hole genesis for early-universe gravitons, which complements the picture of turbulence in the electroweak era. This, in turn, is relatable to the question of if micro black holes could contribute to cosmology. [2] We are trying to reconcile $-1 \leq \omega$ from the DE equation of state with the idea of graviton production from the early states of the universe to the present era, where we have acceleration of the universe. What motivates this presentation as we will try reconcile relic conditions for GW / graviton production. [3,4] The models seem irreconcilable and we do have a huge problem. A possible reconciliation [1 p. 114] suggests an examination of the DE equation of state for reacceleration of the universe, $\widetilde{m} > 1$:

$$\omega = \frac{\dot{\phi}^2 - 2V(\phi)}{\dot{\phi}^2 + 2V(\phi)} \approx \frac{\dot{\phi}^2 - 2V_0 \exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right)}{\dot{\phi}^2 + 2V_0 \exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right)}. \tag{1}$$

Again, this looks like a very hard problem, but in a nutshell, we need make the following identification, which may allow for us to make a concrete bridge between formalisms which otherwise look unlinkable. There is there a way to link $\phi$ of Eq. (1) with $\varphi_k$ of Eq. (3) below so as to come up with an acceptable form of $\phi$ that would satisfy the following equation. Here we need for acceleration of universe $\widetilde{m} > 1$.

$$V(\phi) = V_0 \exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right). \tag{2}$$

Then, we should try to link $\phi$ of Eq. (1) with $\varphi_k$ of Eq. (3).[1,4]

$$\left(\frac{3}{8\pi G}\frac{\kappa}{a^2}+\frac{\kappa^2}{a^2}\varphi_k^2-\frac{16}{3}c_1B^4\right)\approx\frac{-\alpha c_2}{r^\alpha}\equiv\frac{-\alpha c_2}{(Planck\text{–}length)^\alpha}$$
$$\Leftrightarrow\left[\frac{\alpha}{(Planck\text{–}length)^\alpha}\right]\approx\frac{16}{3}\frac{c_1}{c_2}B^4-\frac{\frac{3}{8\pi G}\cdot\frac{\kappa}{a^2}+\frac{k^2}{a^2}\cdot\varphi_k^2}{c_2}. \tag{3}$$

To make this equivalence, we would, after considerable work, be able to link to the scale factor $a$ which shows up in Eq. (1) above. That is, having results in terms of the scale factor, and $\phi$ while making full identification with Eq. (3) above, may be a way to reformulate the linkage with early DE. In addition, such restraints upon scale factor $a$ may also begin to explain some of the scale-factor behavior. Also, the final equation-of-state constraint we have to keep in mind is[1,3,4]

$$-2m_{\text{Planck}}^2\dot{H}=\dot{\phi}^2. \tag{4}$$

We can use this to connect the final and initial parts of expansion of the universe. We will start with, initially going to a Kaluza–Klein (KK) treatment of gravitons.[2,5]

## 2. Reacceleration of the Universe and KK Gravitons with a Nonzero Rest Mass. The KK Treatment

Consider if there is, then, also a small graviton mass:[6,7]

$$m_n=\sqrt{\frac{n^2}{L^2}+(m_{\text{graviton rest mass}}=10^{-65}\text{ g})^2}=\frac{n}{L}+10^{-65}\text{ g}. \tag{5}$$

Note that Rubakov works with KK gravitons, without the tiny mass term for a four-dimensional rest mass included in Eq. (5). To obtain the KK-graviton/DM candidate representation along $RS\,\mathrm{d}S$ Brane world, Rubakov obtains his values for graviton mass and graviton physical states in space–time after using the normalization $\int\frac{\mathrm{d}z}{a(z)}h_m(z)h_{\widetilde{m}}(z)\equiv\delta(m-\widetilde{m})$ and $J_1$, $J_2$, $N_1$, and $N_2$, which are different forms of Bessel functions.[8] His representation of a graviton state is given by Eq. (4). If so, then the wave function for a graviton with a tiny four-dimensional space–time rest mass can be written as[8]

$$h_m(z)=\sqrt{\frac{m}{k}}\frac{J_1\left(\frac{m}{k}\right)N_2\left[\frac{m}{k}\exp(kz)\right]-N_1\left(\frac{m}{k}\right)J_2\left[\frac{m}{k}\exp(kz)\right]}{\sqrt{\left[J_1\left(\frac{m}{k}\right)\right]^2+\left[N_1\left(\frac{m}{k}\right)\right]^2}}. \tag{6}$$

Equation (6) is for KK gravitons having 1 TeV magnitude mass $M_Z\sim k$ (i.e., for mass values at 0.5 TeV to above 1 TeV) on a negative tension RS Brane. It would be useful to relate this KK graviton, which is moving with a speed proportional to $H^{-1}$ with regards to the negative tension Brane with $h\equiv h_m(z\to 0)=const\sqrt{\frac{m}{k}}$ as an initial starting value for the KK graviton mass. If Eq. (5) is for a "massive" graviton with a small four-dimensional graviton rest mass and if $h\equiv h_m(z\to 0)=const\sqrt{\frac{m}{k}}$ represents an initial state, then one may relate the mass of the KK graviton moving at high speed with its initial rest mass. This rest mass of a graviton

is $m_{\text{graviton}}(\text{4D GR}) \sim 10^{-48}$ eV, opposed to $M_X \sim M_{\text{KK}} \sim 0.5 \times 10^9$ eV. Whatever the range of the graviton mass, it may be a way to make sense of what was presented for a graviton mass using CMBR measurements of $M_{\text{KK}} \sim 10^{-20}$ eV.[9] Also, Eq. (6) will be the starting point used for a KK tower version. So[10]

$$\dot{a}^2 = \left\{ \left[ \frac{\widetilde{\kappa}^2}{3} \left( \rho + \frac{\rho^2}{2\lambda} \right) \right] a^2 + \frac{\Lambda a^2}{3} + \frac{m}{a^2} - K \right\}. \tag{7}$$

Maartens[10] also gives a second Friedman equation.

$$\dot{H}^2 = \left\{ - \left[ \frac{\widetilde{\kappa}^2}{2} \cdot (p + \rho) \cdot \left( 1 + \frac{\rho^2}{\lambda} \right) \right] + \frac{\Lambda \cdot a^2}{3} - 2\frac{m}{a^4} + \frac{K}{a^2} \right\}. \tag{8}$$

Also, an observer is in the low-redshift regime for cosmology, for which $\rho \approx -P$ and one obtains exact equality, $\rho = -P$ for redshift, $z$, between zero and five. The net effect will be to obtain, based on Eq. (7), assuming $\Lambda = 0 = K$ and using $a \equiv \frac{a_0 = 1}{1+z}$ to get a deceleration parameter,

$$q = -\frac{\ddot{a}a}{\dot{a}^2} = -1 + \frac{2}{1 + \widetilde{\kappa}^2 \left( \frac{\rho}{m} \right) \cdot (1+z)^4 \cdot \left( 1 + \frac{\rho}{2\lambda} \right)} \approx -1 + \frac{2}{2 + \delta(z)}. \tag{9}$$

These set values allow a graviton-based substitute for DE. $\Lambda = 0 = K$ plus a small rest mass for a graviton in four dimensions allows for "massive gravitons" that behave the same as DE. Setting $\Lambda = 0 = K$, with a modified behavior for the density expression, for a Friedman equation with small four-dimensional graviton mass replaces dark energy by a small four-dimensional rest mass for a graviton.

## 3. How a Small Graviton Mass Reaccelerates the Universe

Revising Alves et al.,[11] Beckwith[6,7] used a higher-dimensional Brane-world model with KK graviton towers, per Maartens,[10] and a Brane-world energy density $\rho$ in the Friedman equation similar to Alves et al.[11] for a nonzero graviton:[6,7]

$$\rho = \rho_0 \cdot (1+z)^3 - \left[ \frac{m_{\text{g}} \cdot (c=1)^6}{8\pi G(\hbar = 1)^2} \right] \cdot \left[ \frac{1}{14 \cdot (1+z)^3} + \frac{2}{5 \cdot (1+z)^2} - \frac{1}{2} \right]. \tag{10}$$

At $z \approx 4$, a billion years ago, acceleration of the universe increased, as shown in Fig. 1.[6,7] Figure 1 is, if confirmed, a good verification of the Ng hypothesis,[12] and would be a starting point to investigate the role of gravitons in cosmology. Buonnano[13] assumes a much lower range of initial frequencies for relic GW. Beckwith[6,7] obtained a reacceleration of the universe as given in Fig. 1. The contribution of a low rest mass for four-dimensional gravitons, as given in Eq. (10) leads to reacceleration of the expansion of the universe a billion years ago. That is, for a redshift slightly smaller than .5. We will state what is needed to explain, in part, Fig. 1.

## 4. Reconciling Graviton Impact and Graviton Generation

The key is to consider the following, namely Eq. (4), Eq. (9), and if $\rho_{\text{DE}} = \rho$.[1,4]

$$3m_{\text{Planck}}^2 H^2 = \rho_{\text{DE}} = \rho = 3m_{\text{Planck}}^2 \left\{ \left[ \frac{\widetilde{\kappa}^2}{3} \left( \rho + \frac{\rho^2}{2\lambda} \right) \right] + \frac{\Lambda}{3a^2} + \frac{m}{a^4} - \frac{K}{a^2} \right\}. \tag{11}$$

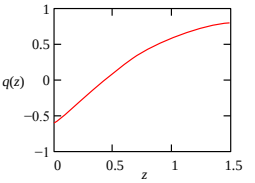



Fig. 1. Reacceleration of the universe.[6] Note that $q(z) < 0$ if $z < .423$.

Leading to

$$\left[\frac{\alpha}{(Planck\text{–}length)^{\alpha}}\right] \approx \frac{16}{3}\frac{c_1}{c_2}B^4 - \frac{\frac{3}{8\pi G}\cdot\frac{\kappa}{a^2} + \frac{k^2}{a^2}\cdot\varphi_k^2}{c_2}, \tag{12}$$

and also satisfying $V \cong \frac{c_2}{r^{\alpha}}$ for deignated $r$ values, $-1 \leq \omega \leq 1$ for the DE equation of state,

$$\omega = \frac{\dot{\phi}^2 - 2V(\phi)}{\dot{\phi}^2 + 2V(\phi)} \approx \frac{\dot{\phi}^2 - 2V_0\exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right)}{\dot{\phi}^2 + 2V_0\exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right)}, \tag{13}$$

and

$$V_0\exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right) \leftrightarrow V \cong \frac{c_2}{r^{\alpha}}. \tag{14}$$

We would have restricted versions of $\alpha$ with this identification. The idea would be if say early gravitons are released by this mechanism,[3] we could eventually come up with a polynomial reconstruction for an inflaton, $\phi$ which would be commensurate with both Eq. (12) and a description of say early universe black holes (micro sized) shedding relic gravitons.[3]

## 5. Conclusions

*To put it mildly, we are bridging between the initial and second physics realms as to graviton physics and cosmology.* This in turn means keeping in mind the following.[2] The diagram[14] gives the relation between GW frequency and GW energy density for different cosmological models. And link Beckwith with Li et al.[1,3,4] and see how they compare with LIGO.[14] That is, we refer you to the figure[14] showing the relation between GW frequency and GW energy density for different cosmological models. Then, we should try to reconcile if is there a way to link $\phi$ of Eq. (1) with $\varphi_k$ of Eq. (3) and Eq. (15)

$$V_0\exp\left(-\sqrt{\frac{2}{\widetilde{m}}}\frac{\phi}{m_{\text{Planck}}}\right) \leftrightarrow V \cong \frac{c_2}{r^{\alpha}}. \tag{15}$$

By using $-1 \leq \omega \leq 1$ of Eq. (1) with the physics of the situation outlined in Eq. (15) and finding $\phi$ of Eq. (15) with $\varphi_k$ of Eq. (3), we hope to eventually reach o final

theory of the evolution of DE and, in doing so, explain the state of gravitons released by initial micro black holes,[3,4] with the physics given in Fig. 1, which represents gravitons as a form of DE, reaccelerating the universe from one billion years ago.

## 6. Now for a Very Important Final Insight from Ng, Using the Idea of Holographic Foam Cosmology

We examine two formulas[15 p. 77] for a cosmic density function, and entropy defined in Eq. (16), with $R_{\rm H}$ being a Hubble radius, $l_{\rm Planck}$ being a Planck radius, and entropy $\approx (R_{\rm H} l_{\rm Planck})^2 \stackrel{?}{\to} n$ (graviton count)

$$\rho_{\rm Cosmic} = \frac{3}{8\pi}(R_{\rm H} l_{\rm Planck})^{-2} \approx \left(\frac{H}{l_{\rm Planck}}\right)^2 . \tag{16}$$

If $\rho_{\rm Cosmic} \propto \rho_{\rm DE} \approx \rho$, can $R_{\rm H}$ be defined rigourously?

As of today, we have the Hubble volume approximately equal to $10^{31}$ cubic light years, approximately the cube of the radius, of Eq. (16), but how this Hubble radius evolves may have much to do with the evolution of dark energy. This is an important detail to work out. Furthermore, in line with Eq. (16), we have another line of inquiry to pursue:[15] If $\rho_{\rm DE} \approx \rho$ and

$$\rho \approx \frac{\hbar k_{\rm Planck}^4}{16\pi^2} \approx \frac{\hbar\left[k_{\rm Planck}^4 = \left(\frac{2\pi}{\Lambda}\right)^4\right]}{16\pi^2} \approx 10^{123}\rho_{\rm observed}, \tag{17}$$

reset $\Lambda$ to $\Lambda_{\rm initial} = 10^{\frac{123}{4}}\Lambda$, so that

$$\rho \approx \frac{\hbar k_{\rm Planck}^4}{16\pi^2} \approx \frac{\hbar\left[k_{\rm Planck}^4 = \left(\frac{2\pi}{\Lambda_{\rm initial}}\right)^4\right]}{16\pi^2} \approx \rho_{\rm observed}. \tag{18}$$

Key to this would be, perhaps, putting Eq. (17) initially into Eq. (16), If this is feasible or not needs to be explored. It could though be part of a new gravitons–DE tapestry. A possible issue is, do we have, after formation of DE density, some sort of invariance of this dark-energy density value, say after the electroweak era?

### Acknowledgements

This work is supported in part by National Natural Science Foundation of China grant No. 11375279

### References

1. M. Li, X.-D. Li, S. Wang and Y. Wang, *Dark Energy*, Peking University Advance Physics Series, Vol. 1 (World Scientific, Hackensack, NJ, 2015).
2. A. Beckwith, Structure formation in the early universe as a result of non linear electrodynamics influencing scale factor size with attendant changes in gravitational potential and its relationship to the three-body problem, *J. High Energy Phys. Gravitation Cosmol.* **forthcoming** (2018).

3. A. Beckwith, Using "enhanced quantization" to bound the cosmological constant, and computing quantum number $n$ for production of 100 relic mini black holes in a spherical region of emergent space–time, *J. High Energy Phys. Gravitation Cosmol.* **4**, 549 (2018), https://doi.org/10.4236/jhepgc.2018.43033.
4. A. Beckwith, Identifying a Kaluza Klein treatment of a graviton permitting a deceleration parameter $q(z)$ as an alternative to standard DE, *J. Cosmol.* **13**, 1 (2011), http://journalofcosmology.com/BeckwithGraviton.pdf.
5. K. Inomata, M. Kawasaki, K. Mukaida and T. T. Yanagida, Double inflation as a single origin of primordial black holes for all dark matter and LIGO observations, *Phys. Rev. D* **97**, p. Article 043514 (2018), https://doi.org/10.1103/PhysRevD.97.043514.
6. A. Beckwith, Applications of Euclidian Snyder geometry to the foundations of space time physics, *Electron. J. Theor. Phys.* **7**, 241 (2010), http://www.ejtp.com/articles/ejtpv7i24p241.pdf.
7. A. Beckwith, Energy content of gravitation as a way to quantify both entropy and information generation in the early universe, *J. Mod. Phys.* **2**, 58 (2011), https://doi.org/10.4236/jmp.2011.22010.
8. V. A. Rubakov, *Classical Theory of Gauge Fields* (Princeton University Press, Princeton, NJ, 2002).
9. S. Dubovsky, R. Flauger, A. Starobinsky and I. Tkachev, Signatures of a graviton mass in the cosmic microwave background, *Phys. Rev. D* **2014**, p. Article 023523 (2010), https://doi.org/10.1103/PhysRevD.81.023523.
10. R. Maartens, Brane-world cosmology, in *The Physics of the Early Universe*, ed. E. Papantonopoulos, Lecture Notes in Physics, Vol. 653 (Springer Verlag, Amsterdam, 2005) pp. 213–247.
11. M. E. S. Alves, O. D. Miranda and J. C. N. de Araujo, Can massive gravitons be an alternative to dark energy? (2010), http://arxiv.org/pdf/0907.5190.
12. Y. J. Ng, Spacetime foam: From entropy and holography to infinite statistics and nonlocality, *Entropy* **10**, 441 (2008), https://doi.org/10.3390/e10040441.
13. A. Buonanno, Gravitational waves, in *Particle Physics and Cosmology: The Fabric of Spacetime*, eds. F. Bernardeau, C. Grojean and J. Dalibard (Elsevier, Oxford, UK, 2007) pp. 10–52.
14. LIGO Scientific Collaboration and Virgo Collaboration, An upper limit on the stochastic gravitational-wave background of cosmological origin, *Nature* **460**, 991 (2009), https://doi.org/10.1038/nature08278.
15. Y. J. Ng, Vacuum fluctuations in domains with moving boundaries and the dark energy issue, in *4th International Workshop on the Dark Side of the Universe*, ed. S. Khalil (American Institute of Physics, Washington, DC, 2008) pp. 74–79.

# A five-dimensional approach to dark matter and dark energy

Michael B. Gerrard* and Timothy J. Sumner

*Physics Department, Imperial College London, London, UK*

**E-mail: gerrardmb@icloud.com*

*www.imperial.ac.uk*

A naive five-dimensional model simultaneously explains dark energy and the flat rotation curves of galaxies and enhanced gravitational lensing usually attributed to dark matter. An additional (fifth) dimension is invoked which corresponds to the radius of curvature of four-dimensional space-time and does not represent a degree of freedom of motion. The universe so modeled has two modes of expansion: that of three-dimensional space and that of the fifth dimension itself, where the latter has the characteristics partly of space and partly of time. The boundary between four-dimensional space-time and the fifth dimension is modeled by contours along which energy is conserved. The model is able to reproduce: (i) the observed relationship between distance modulus and red shift for Type 1a supernovae; (ii) the Tully-Fisher relationship; and (iii) gravitational lensing effects beyond General Relativity for galaxies and clusters of galaxies.

*Keywords*: Dark Energy; Dark Matter; 5 dimensions.

## 1. Introduction

A discrepancy exists between the observed extent of gravitating matter in the universe and that inferred from galaxy rotation curves, galaxy cluster dynamics, gravitational lensing and simulations of evolution of the universe. The existence of abundant unobserved (or dark) matter has been proposed by many cosmologists[1] to account for this discrepancy, although to date dark matter has not been directly detected.[2–4] Observations concerning the luminosity distance of Type 1a supernovae, their red shift and the rate of expected Hubble expansion of the universe have led to the conclusion that the expansion has been accelerating since about six billion years ago. This has led many to propose that dark energy permeates the universe, counteracting the decelerating effect of gravity.[5,6] However, to date, the source of dark energy has not been identified.[7]

It is the purpose of this paper to show that a universe comprised of five rather than four large-scale dimensions and without the assumed presence of dark matter and dark energy, could account for these same observations concerning galaxies, clusters of galaxies and the expansion rate of the universe. The possibility of additional dimensions of space-time has, of course, long been under discussion within theoretical physics.[8–11] However, the distinctive features of the dimension considered in this paper include its large scale and its hybrid character, being partly like space and partly like time.

## 2. A Fifth Large Scale Dimension

We consider a naive model universe in which three-dimensional space, $s(x, y, z)$, is isotropic, closed and expanding. The radius of curvature in this universe is given the symbol, $r$, and dimensions of length and is defined to be orthogonal to the three space dimensions, $s$, and time, $t$. If Newton's Second Law can be extended to five dimensions, then energy propagating (at the speed of light, $c$, through three-dimensional space that is closed and forming the "surface volume" of a four-dimensional hyper-sphere of radius $r$, will experience a virtual acceleration $a_r$ (1) acting in the direction of $r$ increasing. It is assumed to be fundamental that energy is conserved within four-dimensional space-time, notwithstanding that it is embedded within five dimensions.

$$a_r = \frac{c^2}{r}. \tag{1}$$

## 3. Angular Velocity Vectors in 5 Dimensions

An axial vector, $\Omega$, representing angular velocity is aligned normal to the plane of rotation $xy$ of the rotating body it describes. However, the direction of $\Omega$ can lie anywhere in the $zr$ plane, as both dimension $z$ and the additional dimension, $r$, are orthogonal to the $xy$ plane. Consequently, more information is required to define uniquely the orientation of $\Omega$ within a five-dimensional framework.

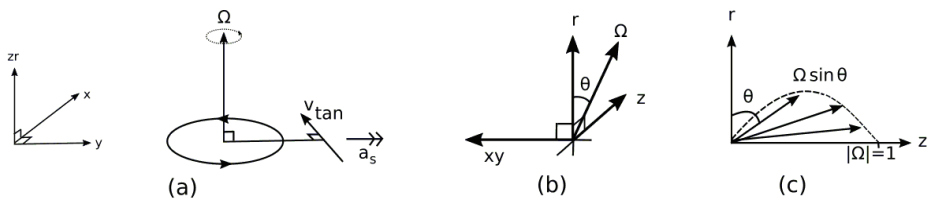



Fig. 1. Axial vector ($\Omega$), its orientation in 5 dimensions and resolved magnitude in 4 dimensions.

This additional information comes from including the influence of the additional acceleration $a_r$ (defined above) on the rotating body, which defines the orientation $\theta$ of the angular velocity vector $\Omega$ in the $zr$ plane (Fig. 1). The centrifugal acceleration, $a'_s$, experienced by the rotating body within a five-dimensional framework is given by (2), where: $\tan\theta = a_s/a_r$; $a_s$ is the four-dimensional centrifugal acceleration of the body; and $v_{tan}$ is the tangential velocity of the body about its centre of rotation. The possibility of forces arising from an additional dimension and perturbing dynamics observed within four dimensions has been previously identified by Wesson.[12]

$$a'_s = \Omega\,(\sin\theta + i\cos\theta)\,v_{tan}. \tag{2}$$

This five-dimensional analysis of rotational motion ultimately leads to the derivation of (3) and so to the Tully-Fisher relationship[13] and predictions of flat rotation curves for certain types of galaxies[14], without the need for dark matter to be assumed, where $M$ is the baryonic mass of the galaxy and $G$ is the gravitational constant. However, consistent with observations, this analysis does not predict any departure from the expected four-dimensional Newtonian behaviour for stars at the fringe of galaxies, as regards their motion normal to their planes of rotation, nor their motion radially from the centre of the galaxy.[15,16]

$$|a_r GM| = v_{tan}^4. \tag{3}$$

Comparing the parameters predicted by this analysis with those reported by Stark[17] and McGaugh[18] places values for $a_r$ in the range $1.2 - 1.9 \times 10^{-10}\,\mathrm{ms}^{-2}$ which is also consistent with values for the MOND constant, $a_0$, reported by Milgrom.[19,20] From (1), the implied range of values for $r$ is $4.7 - 7.5 \times 10^{26}\,\mathrm{m}$. A steady increase in $r$ with cosmological time, described in section 5. implies a steadily reducing value for $a_r$.

## 4. Gravitational Lensing

The principle of energy conversation along contours that define the boundary between four-dimensional space-time, $s, t$ and the fifth dimension, $r$, can be used to generate the Lorentz contraction formula of Special Relativity and the dilation formula of General Relativity, as they apply to the dimension $r$. The latter is shown in (4) where $r_u$ is the value for $r$ remote from gravitational fields.

$$r = r_u \left(1 - \frac{2GM}{sc^2}\right)^{1/2}. \tag{4}$$

In five dimensions, the acceleration, $a$, experienced by a beam of light propagating through a gravitational field (generated by mass $M$) is given by (5), where $g$ is the gravitational acceleration.

$$a = g + \left(\frac{c}{r} \wedge \dot{r}\right). \tag{5}$$

Considering only the real term in (2), the induced angular rate of perturbation of the beam of light, $\dot{\epsilon}$, can be written as in (6), where the relevant gradient $\tan\theta = g/a_r$ and the tangential velocity is the speed of light, $c$.

$$\dot{\epsilon} = \frac{a}{c \sin\theta}. \tag{6}$$

Integration along the path of the beam of light in five dimensions yields the expected expression (7) of General Relativity for the angular deflection, $\Delta\epsilon$, of the beam of light, where $s$ is the impact parameter; however, with an additional term that accounts for enhanced lensing effects at the fringes of galaxies and galaxy clusters.

$$\Delta\epsilon = \frac{4GM}{sc^2 f\,(\sin\theta)}. \tag{7}$$

Whereas for the Sun, $f(\sin\theta) = 1$ as $g >> a_r$ at the impact parameter, for galaxies and clusters of galaxies $g < a_r$ and, correspondingly $f(\sin\theta) < 1$ leading to lensing effects that are beyond GR expectation (7). The distribution of matter within the lensing system determines the local profile of $\sin\theta$ and could, in principle, give rise to a centre of lensing which did not coincide with the centre of mass of the system. This would be relevant to consideration of lensing observed for the bullet cluster and similar systems.[21]

## 5. The Expansion of Space

If matter is assumed on a large scale to be evenly spread across the universe, it is possible to derive from (4) an expression for the background value for $r$, given the symbol $r_b$, of the form $r_b = f(r_u, G, \rho, H_H, c)$ where $\rho$ is the baryonic mass density of the universe (a value of $4.2 \times 10^{-28}\,\mathrm{kgm}^{-3}$ has been used for the current era[22]) and $H_H$ is the Hubble horizon. The derivatives of $r_b$ show how a universe defined in five dimensions (rather than four) has two modes of expansion: the first concerns the expansion of three-dimensional space at a constant value of the fifth-dimensional parameter $r_b$; and the second concerns the expansion of space due to increases in the parameter $r_b$ itself. Accordingly, in five dimensions Hubble's constant can be decomposed into two terms, referred to respectively as $H3$ and $H5$, whose values can be shown to contribute to a Hubble constant of value $68.1\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$,[23,24] as follows: $H3 = 66.5\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ and $H5 = 1.6\,\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. The competing effects of these two modes of expansion can give rise to either an accelerating or decelerating expansion of the universe, within a five-dimensional framework.

Moreover, the conservation of energy at the boundary between four-dimensional space-time and the fifth dimension must necessarily impose an important constraint on the second derivative of $r_b$, shown in (8), and values for these two terms can be independently derived[25] and shown to satisfy

$$\frac{c^2}{r_b} + \ddot{r}_b = 0. \tag{8}$$

The balance between accelerations directed in three-dimensional space and those directed in the fifth dimension determines the gradient of contours $r(s)$ which define the boundary between four-dimensional space-time and the fifth dimension. This condition of balance (or equilibrium) between different accelerations can, alternatively, be thought of as defining a "surface tension" within the contour at the boundary between four-dimensional space-time the fifth dimension, which in $(x, y, z, r)$ dimensions become branes. This approach to the determination of the local curvature of space-time suggests a more than coincidental basis for the rubber sheet analogy of General Relativity.

Finally, we use the five-dimensional model described in this paper to calculate the relationship between distance modulus, $\mu$ and red shift, $z$ for Type 1a SN and compare this with currently available observational data from

(http://supernova.lbl.gov/Union/Union2.html).[26] The results of this fit are shown in Fig 2. The residuals trend indicates that the model slightly over-estimates $\mu$ at lower red shift and, correspondingly, under-estimates it for larger $z$. This kind of residuals trend may be an artefact of the way the data are calibrated for the absorption of light by inter-stellar gas. The relative evolution of the terms $H3$ and $H5$ in the model (see above) predict a point of inflection in the rate of expansion of the universe circa 6 billion years ago, but from an accelerating to a decelerating expansion. That is, the reverse of what has been deduced from fitting the SN data to four-dimensional models.[5,6] The five-dimensional model does not indicate the presence of an unaccounted accelerating expansion of the universe, which has been taken to indicate the presence of dark energy.

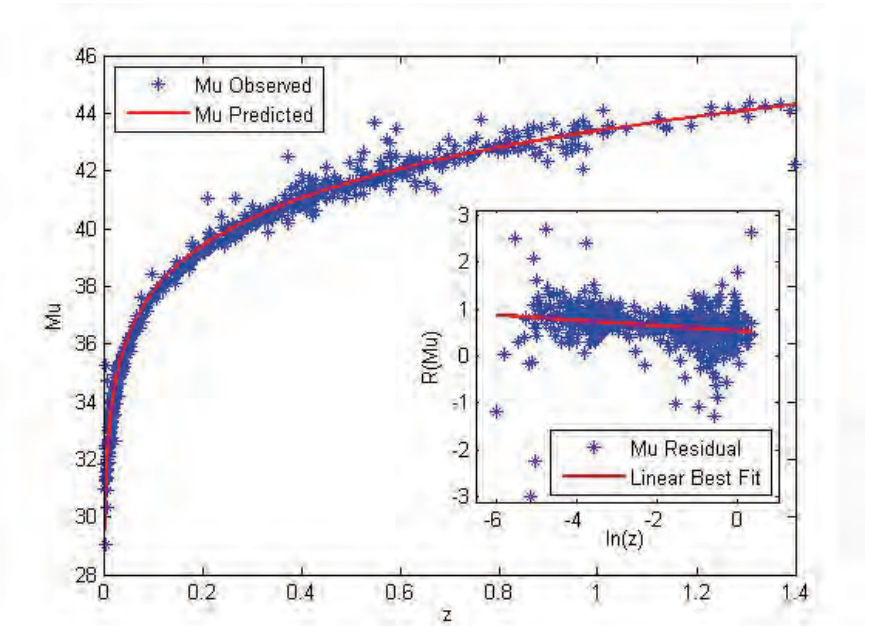



Fig. 2. Type 1a SN luminosity distance modulus ($\mu$) vs red shift, $z$, comparison between observation and five-dimensional model predictions.

## 6. Conclusion

If five-dimensional effects are taken into account, then predictions can be made for orbital dynamics within galaxies, the gravitational lensing of galaxies and galaxy clusters, and the relationship between distance modulus and red shift for Type 1a SN, which are consistent with observations, without the need to introduce dark matter or dark energy. The additional dimension introduced in this analysis has characteristics both of space dimensions (e.g. being a radius of curvature and a

distance parameter within vector algebra) and a time dimension (e.g. being neither a degree of freedom of motion nor being directly observable).

**Acknowledgments**

We would like to acknowledge computational work done by James H Tull.

**References**

1. L. Bergstrom, *Ann. Physik* **524**, 479 (2012).
2. D. S. Akerib et al., *Phys. Rev. Letts.* **118**, 021303 (2017).
3. X. Cui et al., *Phys. Rev. Letts.* **119**, 181302 (2017).
4. E. Aprile et al., *Phys. Rev. Letts.* **121** 111302 (2018).
5. A. G. Riess et al., *AJ* **116**, 1009 (1998).
6. S. Perlmutter et al., *Ap. J.* **517**, 565 (1999).
7. P. A. R Ade et al., *A&A* **594**, A14 (2016).
8. T. Kalusa *Sitz. Preuss. Akad. Wiss. (Berlin)* 966 (1921).
9. O. Klein, *Zeits. fur Physik* **37**, 895 (1926).
10. J. Overduin and P. Wesson *Phys. Reps.*, **402**, 467 (2004).
11. P. S. Wesson, *Int. J. Mod. Phys. D* **24**, 15300001 (2015).
12. P. S Wesson and B. Mashhoon, and H. Liu and W. N. Sajko *Phys. Letts. B* **456**, 34 (1999).
13. R. B. Tully and J. R. Fisher, *A&A* **54**, 34 (1977).
14. V. Rubin and N. Thonnard and W. K. J. Ford, *Ap. J.* **225**, L107 (1978).
15. G. W. Angus et al., *MNRAS*, **451**, 3551 (2015).
16. S. S. McGaugh, *ApJ* **816**, 42 (2016).
17. D. V. Stark, S. S. McGaugh and R. Swaters, *AJ* **138**, 392 (2009).
18. S. S. McGaugh and J. M. Schombert, *ApJ* **802**, 18 (2015).
19. M. Milgrom, *ApJ* **270**, 365 (1983).
20. M. Milgrom, *Phys. Rev. D* **91**, 044009 (2015).
21. F. Gastaldello et al., *MNRAS*, **442**, L76 (2014).
22. K. Olive et al., *Chin. Phys. C.* **38**, 090001 (2014).
23. P. A. R. Ade et al., *A&A* **571**, A1 (2014).
24. N. Jarosik et al., *ApJ. Suppl.* **192**, 14 (2011).
25. M. B. Gerrard and T. J. Sumner, *arXiv:0704.0332* (2007).
26. R. Amanullah et al., *ApJ* **716**, 712 (2010).

# Non-metric fields in quantum gravity

Kaushik Ghosh*

*Vivekananda College,*
*Kolkata, West Bengal 700091, India*
** E-mail: kaushikhit10@gmail.com*
*www.vckolkata63.org*

In this manuscript, we will discuss the construction of a theory of gravity with nonmetricity and general affine connections. We will consider a simple potential formalism with symmetric affine connections and symmetric Ricci tensor. Corresponding affine connections introduce two massless scalar fields. One of these fields contributes a stress-tensor with opposite sign to the sources of Einstein's equation when we state the equation using the Levi-Civita connections. This means we have a massless scalar field with negative stress-tensor in the familiar Einstein equation. These scalar fields can be useful to explain dark energy and inflation. These fields bring us beyond strict local Minkowski geometries.

*Keywords*: Connections, nonmetricity, scalar fields, negative stress tensor.

## 1. Introduction

In this manuscript, we will try to find if quantum gravity can introduce additional fields besides metric. We will discuss this with reference to the construction of covariant derivative operator in quantum gravity. We have earlier found it is a perceptive alternative to use affine connections more general than metric compatible affine connections in quantum gravity.[1] General affine connections introduce additional fields in the theory that can be useful to explain cosmological observations and can introduce new effects.[2] These fields are non-localized similar to dark energy and inflation. In this article, we have considered a simple case with symmetric Ricci tensor. Corresponding affine connections are symmetric and introduce only two massless scalar fields. One of these gives a stress tensor with opposite sign to the sources of Einstein's equation in the classical theory when we state the equation using the Levi-Civita connections. This means we have a source with negative stress tensor in the familiar Einstein equation. These scalar fields can be possible candidates for dark energy and inflation. These two fields break strict local Minkowski structure of spacetime. Global splitting of spacetime into space and time no longer remains exact in such spaces. This is an important issue in quantum gravity.[3] Finding the particle interpretations and other possible interactions of the above fields are non-trivial problems without local Minkowski structure. However, experiments suggest that such effects are very small in the present universe and the corresponding scalars mentioned above are also small. Their effects can be observed in large scale phenomena like the present cosmic acceleration. This is also consistent with the smallness of the parameters like the cosmological constant required to explain dark energy in some models mentioned at the beginning. The effects of these fields

will be large in the quantum gravity domain. We note that vacuum fluctuations of these fields can also be useful in cosmology.

## 2. General Affine Connections and Scalar Fields

In this section, we will consider applications of the potential formalism of the Palatini theory discussed in to introduce finite on shell non-metricity.[1] General affine connections are given by: $\Gamma^{\alpha}_{\ \mu\nu} + \tilde{C}^{\alpha}_{\ \mu\nu}$, where $\Gamma^{\alpha}_{\ \mu\nu}$ is the Levi-Civita connections and $\tilde{C}^{\alpha}_{\ \mu\nu}$ is symmetric in the lower indices and the sources do not couple with these fields. We consider symmetric Ricci tensors only. This is the most relevant case when we consider semiclassical and classical limits of quantum gravity. This example will introduce two massless scalar fields. One of them will contribute a negative stress tensor to Einstein's equation in the semiclassical or classical theories. We will use the geometrized units in the following where, $G = c = 1$. We have discussed earlier that we can introduce non-metricity by defining the affine connections to be compatible with a symmetric covariant field: $b_{\mu\nu} = g_{\mu\nu} + a_{\mu\nu}$; $a_{\mu\nu} \neq k g_{\mu\nu}$, where $k$ is a constant including zero.[1] Such a set of affine connections give a symmetric Ricci tensor. We now break $a_{\mu\nu}$ into a trace and a traceless part:

$$a_{\mu\nu}(x) = \Phi(x) g_{\mu\nu} + \bar{a}_{\mu\nu}(x); \quad \Phi(x) = \frac{a(x)}{4} \tag{1}$$

Where, $\Phi$ is a scalar field, $a(x)$ is the trace of $a_{\mu\nu}$ and $\bar{a}_{\mu\nu}$ is trace-free. We can express corresponding $\tilde{C}^{\alpha}_{\ \mu\nu}$ in the following way:

$$\tilde{C}^{\alpha}_{\ \mu\nu} = \delta^{\alpha}_{\ (\mu} \nabla'_{\nu)} [ln(1+\Phi)] - \frac{1}{2} g_{\mu\nu} \nabla'^{\alpha} [ln(1+\Phi)] + E^{\alpha}_{\ \mu\nu} \tag{2}$$

The first two terms in the *r.h.s* gives the contribution of the trace part of $b_{\mu\nu}$ given by: $(g_{\mu\nu} + \Phi(x) g_{\mu\nu})$.[4] A further contraction of the third term in the covaraiant indices leads to the following expression for a general $a_{\mu\nu}$:

$$\tilde{C}^{\alpha}_{\ \mu\nu} = \delta^{\alpha}_{\ (\mu} \nabla'_{\nu)} [ln(1+\Phi)] - \frac{1}{2} g_{\mu\nu} \nabla'^{\alpha} [ln(1+\Phi)] + g_{\mu\nu} \nabla'^{\alpha} \Psi + g_{\mu\nu} E^{\alpha} + \bar{E}^{\alpha}_{\ \mu\nu} \tag{3}$$

Where, $E^{\alpha}$ is a vector field which is not a gradient and $\bar{E}^{\alpha}_{\ \mu\nu}$ is traceless in the lower indices. $\Phi$ gives the trace-scalar of $a_{\mu\nu}$. The scalar $\Psi$ is also expected. In the Minkowski space, we have a spin zero boson associated with $\bar{a}_{\mu\nu}$ in addition to a spin one boson and a spin two boson. We will discuss this aspect further below Eq. (14). We now consider the simplest cases of Eq. (3). We first consider the case where the traceless part of $a_{\mu\nu}$ vanish and which introduces a scalar field only:

$$b_{\mu\nu} = g_{\mu\nu} + \Phi(x) g_{\mu\nu} = \chi(x) g_{\mu\nu} \tag{4}$$

We use affine connections that are compatible with $b_{\mu\nu}$ which is conformal to metric. In this case, $\tilde{C}^{\alpha}_{\ \mu\nu}$ and the modified curvature scalar are given by the following expressions:[4]

$$\tilde{C}^{\alpha}_{\ \mu\nu} = \delta^{\alpha}_{\ (\mu}\nabla'_{\nu)}[ln(1+\Phi)] - \frac{1}{2}g_{\mu\nu}\nabla'^{\alpha}[ln(1+\Phi)] \tag{5}$$
$$\mathcal{R} = R' - \frac{3}{2}\frac{1}{(1+\Phi)^2}(\nabla'\Phi)^2$$

Here $(\nabla'\Phi)^2 = (\nabla'_{\ \mu}\Phi)(\nabla'^{\mu}\Phi)$, is the norm of the gradient of $\Phi$ and the primed quantities are evaluated using the Levi-Civita connections. For small $\Phi$ ($<< 1$), we have the following modification of the *r.h.s* of Einstein's equation where covariant derivatives are evaluated using the Levi-Civita connections:[4]

$$\nabla'_{\ \kappa}\nabla'^{\kappa}\chi + \frac{1}{\chi}(\nabla'\chi)^2 = 0 \ \approx \ \nabla'_{\ \kappa}\nabla'^{\kappa}\Phi + (\nabla'\Phi)^2 \tag{6}$$
$$G'_{\ \mu\alpha} = 8\pi P'_{\ \mu\alpha} + \frac{3}{2\chi^2}[(\nabla'_{\ \mu}\chi)(\nabla'_{\ \alpha}\chi) - \frac{1}{2}g_{\mu\alpha}(\nabla'\chi)^2]$$
$$\approx 8\pi P'_{\ \mu\alpha} + \frac{3}{2}[(\nabla'_{\ \mu}\Phi)(\nabla'_{\ \alpha}\Phi) - \frac{1}{2}g_{\mu\alpha}(\nabla'\Phi)^2]$$
$$= 8\pi[P'_{\ \mu\alpha} + \frac{3}{16\pi}P'_{\ \mu\alpha}(\Phi)]$$

Where, we have considered terms upto second order in $\Phi$. $P'_{\ \mu\alpha}$ is the stress tensor of ordinary matter. $P'_{\ \mu\alpha}(\Phi)$ is the stress tensor of an ordinary masless scalar field. Both $\chi$ and $\Phi$ are massless. We find that the scalar field $\Phi$ behaves like a massless scalar field with its stress tensor coming as a source term in Einsrein's equation. This field can be useful to explain inflation. We can modify the metric-affine action by adding potential term for $\chi$ or $\Phi$. To ensure non-metricity, we have the following condition:

$$\nabla_{\mu}g_{\alpha\beta} = -g_{\mu(\alpha}\nabla'_{\beta)}[ln(1+\Phi)] \neq 0 \tag{7}$$

This is a mild condition. In the present case, matter fields do not couple with $\Phi$. The only observable effect of $\Phi$ is to produce a massless scalar field stress tensor in Einstein's equation. We now consider the other case where only $\Psi$ is finite:

$$\tilde{C}^{\alpha}_{\ \mu\nu} = g_{\mu\nu}\nabla'^{\alpha}\Psi \tag{8}$$

We have the following expressions for different quantities:

$$\mathcal{R}_{\mu\alpha} = R'_{\ \mu\alpha} + g_{\mu\alpha}(\nabla'\Psi)^2 - (\nabla'_{\ \mu}\Psi)(\nabla'_{\ \alpha}\Psi) \tag{9}$$
$$\mathcal{R} = R' + 3(\nabla'\Psi)^2$$

We now solve the corresponding extremization problem and obtain the following set of equations:

$$\begin{aligned} \nabla'_{\kappa}\nabla'^{\kappa}\Psi &= 0 \\ G'_{\mu\alpha} &= 8\pi P'_{\mu\alpha} - 3[(\nabla'_{\mu}\Psi)(\nabla'_{\alpha}\Psi) - \frac{1}{2}g_{\mu\alpha}(\nabla'\Psi)^2] \\ &= 8\pi[P'_{\mu\alpha} - \frac{3}{8\pi}P'_{\mu\alpha}(\Psi)] \end{aligned} \tag{10}$$

Where, $P'_{\mu\alpha}$ is the stress-tensor of ordinary sources and $P'_{\mu\alpha}(\Psi)$ is the stress-tensor of the massless scalar field $\Psi$. This stress tensor satisfy all the energy and pressure conditions that the stress-tensor of a massless scalar field satisfies but it comes with an opposite sign in the *r.h.s* of Einstein's equation, when we state the equation using the Levi-Civita connections. Thus, the effect of $\Psi$ introduced to generalize the Levi-Civita connections, is to contribute a negative massless scalar field stress tensor to the sources of the Einstein equation obtained from the Einstein-Hilbert action formalism. This gives us an alternate way to explain effects that dark energy is proposed for. Negative stress-tensor is also important in Hoyle-Narlikar theory of gravity and also in wormhole and warp drive. In this case, we can not generate the cosmological constant with ordinary massless scalar field $\Psi$ for which the energy density is positive definite. We have the following expression for non-metricity:

$$\nabla_{\mu}g_{\alpha\beta} = -2g_{\mu(\alpha}\nabla'_{\beta)}\Psi \tag{11}$$

This expression is similar to the non-metricity obtained for $\Phi$ apart from a factor of 2. We now consider the case when both $(\Psi, \Phi)$ are present. $\tilde{C}^{\alpha}_{\mu\nu}$ is given by the following expression:

$$\tilde{C}^{\alpha}_{\mu\nu} = \delta^{\alpha}_{(\mu}\nabla'_{\nu)}[ln(1+\Phi)] - \frac{1}{2}g_{\mu\nu}\nabla'^{\alpha}[ln(1+\Phi)] + g_{\mu\nu}\nabla'^{\alpha}\Psi \tag{12}$$

We obtain the following expression for the modified curvature scalar:

$$\mathcal{R} = R' - \frac{3}{2}\frac{1}{(1+\Phi)^2}(\nabla'\Phi)^2 + 3(\nabla'\Psi)^2 + \frac{3}{(1+\Phi)}[(\nabla'_{\kappa}\Phi)(\nabla'^{\kappa}\Psi)] \tag{13}$$

We have the following generalization of Einstein's equation:

$$\begin{aligned} \nabla'_{\kappa}\nabla'^{\kappa}\Psi + \frac{1}{2}\frac{1}{(1+\Phi)}\nabla'_{\kappa}\nabla'^{\kappa}\Phi &= 0 \\ \frac{3}{2}\nabla'_{\kappa}\nabla'^{\kappa}\Phi + \frac{1}{(1+\Phi)}(\nabla'\Phi)^2 - (\nabla'_{\kappa}\Psi)(\nabla'^{\kappa}\Phi) &= 0 \\ G'_{\mu\alpha} &= 8\pi[P'_{\mu\alpha} + \frac{3}{16\pi}P'_{\mu\alpha}(\Phi) - \frac{3}{8\pi}P'_{\mu\alpha}(\Psi) - \frac{3}{8\pi}P'_{\mu\alpha}(\Psi,\Phi)] \\ P'_{\mu\alpha}(\Psi,\Phi) &= \frac{1}{(1+\Phi)}[(\nabla'_{(\mu}\Psi)(\nabla'_{\alpha)}\Phi) - \frac{1}{2}g_{\mu\alpha}(\nabla'_{\kappa}\Psi)(\nabla'^{\kappa}\Phi)] \end{aligned} \tag{14}$$

The first two equations are the equations for $\Psi$ and $\Phi$ respectively. We find that coupling of $\Psi$ with $\Phi$ gives another contribution to source stress tensor which can be positive or negative. This is again important for dark energy research.

We now briefly discuss the geometrical significance of $(\Psi, \Phi)$ and corresponding non-metricities given by Eqs. (7,11). We define the the non-metricity tensor by:

$$Q_{\mu\alpha\beta} = -\nabla_\mu g_{\alpha\beta} = g_{\mu(\alpha}\nabla'_{\beta)}[ln(1+\Phi)]; \quad 2g_{\mu(\alpha}\nabla'_{\beta)}\Psi \tag{15}$$

Both $\Psi$ and $\Phi$ can be present in the manifold when non-metricity is required. We can split $Q_{\mu\alpha\beta}$ into a trace $Q_\mu$ and traceless part $\bar{Q}$ in the last two indices:[3]

$$Q_{\mu\alpha\beta} = Q_\mu g_{\alpha\beta} + \bar{Q}_{\mu\alpha\beta} \tag{16}$$

Both trace and traceless parts of $Q_{\mu\alpha\beta}$ are finite for the above cases. Corresponding connections do not preserve the light cone under parallel transport and we no longer have the local Minkowski structure of spacetime.[3] Thus, we can not have exact $(3+1)$ -splitting of the underlying manifold into space and time. We find that departure from local Minkowski geometry can give new fields like dark matter, dark energy and inflation not found ordinarily. A few examples of non-metric affine connections with vanishing $\bar{Q}$, that are local gauge theories for the Weyl group (Poincare group plus dilatation), are discussed in the references given at the footnote 24 of Ref. 3. We can have finite or vanishing torsion in such theories. These theories preserve the light cone under parallel transport and are locally Minkowski due to the reparameterization invariance of geodesics.[4]

The scalar fields $(\Psi, \Phi)$ are purely quantum gravitational in origin. They are kinematically required in quantum gravity. In this domain, $Q_{\mu\alpha\beta}$ can not vanish due to the arguments given in Sect. III. $(\Psi, \Phi)$ are finite every where and hence, they are non-localized similar to dark energy and inflation. They break the local Minkowski structure of spacetime, and are not present in the matter-gravity coupling part of the semiclassical theory of quantum gravity which presently is locally Lorentz invariant quantum field theory in curved spaces. It is unexpected that any significant coupling between $(\Psi, \Phi)$ and ordinary matter like fermions, present in the full quantum theory, will be lost in the semiclassical limit. Thus, we assume that $(\Psi, \Phi)$ can be present without corresponding sources from ordinary matter. They contribute non-trivially to the source stress tensor of Einstein's equation and are possible candidates for dark energy and inflation. Dark energy is non-localized, have negative pressure and is primarily observed by their gravitational effects. These make $\Psi$ a possible candidate for dark energy. $\Phi$ can be a possible candidate for inflation. The amount of dark energy is much higher than ordinary matter and $(\Psi, \Phi)$ need not to have ordinary matter with finite hypermomentum as their sources. A complete theory of quantum gravity will illuminate this issue further. $(\Psi, \Phi)$ are in line with other scalars introduced similarly, like the dilaton.[5] Observed local Lorentz invariance in many experiments indicate that $(\Psi, \Phi)$ are presently very small. This can explain the smallness of parameters like the cosmological constant required to explain present cosmological acceleration in some models. With $(\Psi, \Phi)$ being small

and non-localized, their effects are usually observed in large scale phenomena and are not much significant in small scale astrophysics like that of the Solar system. This is another characteristic aspect of dark energy. Quantum fluctuations in $(\Psi, \Phi)$, including vaccuum fluctuations, can also be useful to explain different cosmological eras. $(\Psi, \Phi)$ will be coupled with each other when we include both of them. This is given by Eq. (14). When required, we can include $E^{\alpha}, \bar{E}^{\alpha}{}_{\mu\nu}, q_{\mu\nu}$ and torsion potentials to explain different cosmological observations. A related problem is to find out possible interactions and particle interpretations of fields that correspond to different representations of the Lorentz group when the local Minkowski structure of spacetime is broken strongly. This will happen in the quantum gravitational domain. This will be important to understand the relations between the complete set of fields including ordinary matter and gauge fields. We can continue to describe gauge theories by potentials. This is consistent with the Palatini formalism, since the potentials are analogous to connections in the geometric theory of gauge fields.

## References

1. K. Ghosh, *https://hal.archives-ouvertes.fr/hal-02105422* (accepted in *Physics of the Dark Universe*).
2. L. Amendola and S. Tsujikawa, *Dark Energy*, (Cambridge University Press, Cambridge, 2010).
3. F. W. Hehl *et al*, *Rev. Mod. Phys*, **48** (1976) 3641.
4. R. M. Wald, *General Relativity* (The University of Chicago Press, Chicago and London, 1984).
5. M. Gasperini and G. Veneziano, *Phys. Rept.* **373**, 1 (2003).

# *p*-form quintessence: exploring dark energy of *p*-forms coupled to a scalar field

Alejandro Guarnizo[1,2 *], Juan P. Beltrán Almeida[1] and César A. Valenzuela-Toledo[2]

[1] *Departamento de Física, Universidad Antonio Nariño, Cra 3 Este # 47A-15, 110231, Bogotá DC, Colombia*
[2] *Departamento de Física, Universidad del Valle, Ciudad Universitaria Meléndez, 760032, Santiago de Cali, Colombia*
[*] *E-mail: alejandro.guarnizo@correounivalle.edu.co*

We consider a model based on $p$-form kinetic Lagrangians in the context of dark energy. The Lagrangian of the model is built with kinetic terms of the field strength for each $p$-form coupled to a scalar field $\phi$ through a kinetic function. We assume that this scalar field is responsible for the present accelerated expansion of the Universe. Since we are interested in cosmological applications, we specialize the analysis to a 4-dimensional case, using an anisotropic space-time. By studying the dynamical equations, we investigate the evolution of the dark energy density parameter, the effective equation of state and the shear induced by the anisotropic configuration.

*Keywords*: Dark energy; $p$-forms; dynamical systems.

## 1. Motivation

The predictions coming from the inflationary paradigm 1, 2 had been successfully confirmed with measures of the fluctuations in temperature of the Cosmic Microwave Background (CMB), and probes in the Large Scale Structure (LSS) of the Universe, with a significantly increase of precision during the last decades. In its simple form, based on a single scalar field (the *inflaton*) with a slow-roll potential, inflation predicts a statistical Gaussian distribution function and an isotropic power spectrum. However, some anomalies present in current data, need models beyond the standard slow-roll description in order to be fully addressed. These anomalies are related with statistical anisotropies and signals of parity violation 3.

One simple attempt relies in the inclusion of vector fields (or 1-forms), due to the intrinsic preferred directions they dictate. Models which couples a Maxwell kinetic term and a scalar field as $f(\phi)F^{\mu\nu}F_{\mu\nu}$ with $F_{\mu\nu}$ the field strenght of a vector field $A_\mu$, had been studied in the context of inflation 4, as well as modifications like $\phi F_{\mu\nu}\tilde{F}^{\mu\nu}$ with $\tilde{F}$ the dual of $F$ 5. With the same spirit, the anisotropic spectrum of models including terms as $H_{\mu\nu\sigma}H^{\mu\nu\sigma}$ being $H_{\mu\nu\sigma}$ the corresponding field strength of a 2-form field $B_{\mu\nu}$, had been considered 6.

Besides the applications to cosmic inflation, general $p$-forms had potential interest to explain the current acceleration of the Universe. In particular, anisotropic dark energy coming from a quintessence field $\phi$ coupled to a vector field were studied in Ref. 7. A similar analysis was carried out in Ref. 8, but this time using the field strength of a 2-form coupled to the scalar $\phi$. In both references, possible scenarios

where dark energy domination era is plausible after radiation and matter epochs, were found. We can go further the standard approach of Maxwell-like terms of the $p$-forms and allow for couplings between them, as in Ref. 9, where this construction was made. The aim of this short paper is to begin studying the cosmological consequences of coupled $p$-forms. We will focus in the case of a 4-dimensional space-time and will consider the effect of a combination of a 1-form and a 2-form fields coupled to a kinetic function of the quintessence field.[a]

## 2. *p*-form-scalar model

We will consider the standard Lagrangian for a scalar field composed by its kinetic term and a potential $V(\phi)$ as:

$$\mathcal{L}_\phi = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + V(\phi). \tag{1}$$

For the $p$-form sector we start with basic definitions. Given a $p$-form $A_{p\,\mu_1,\mu_2\cdots\mu_p}$, its dynamics is introduced by the field strength $F_{p\,\mu_1\mu_2\cdots\mu_{p+1}} \equiv \partial_{[\mu_1}A_{p\,\mu_2\mu_3\cdots\mu_{p+1}]}$. In this simple case, the Lagrangian that we are going to construct will be built out of the appropriate combinations of the field strengths of the $p$-forms, coupled to the scalar field $\phi$ through arbitrary functions $f(\phi)$. In four dimensions, only two terms remain 9, thus the Lagrangian simply reads[b]

$$\mathcal{L}_p(\phi, A_p) = f^2(\phi)\left(\frac{1}{4}F_{1\,\mu_1\mu_2}F_1^{\mu_1\mu_2} + \frac{1}{12}F_{2\,\mu_1\mu_2\mu_3}F_2^{\mu_1\mu_2\mu_3}\right). \tag{2}$$

Assuming standard gravity, the action of our model can be written as

$$S_p = \int \mathrm{d}^4x\sqrt{-g}\left[\frac{M_\mathrm{p}^2}{2}R - \mathcal{L}_\phi - \mathcal{L}_p(\phi, A_p)\right], \tag{3}$$

where $M_\mathrm{p}$ the Planck mass and $R$ the Ricci scalar.

## 3. Background equations

The Einstein equations could be written as

$$R_{\mu\nu} - \frac{1}{2}Rg_{\mu\nu} = 8\pi G(T^m_{\mu\nu} + T^\phi_{\mu\nu} + T^p_{\mu\nu}), \tag{4}$$

where we split the energy momentum tensor, $T_{\mu\nu}$, in three parts: $T^\phi_{\mu\nu}$, $T^p_{\mu\nu}$ and $T^m_{\mu\nu}$, representing the contributions of the scalar field, the $p$-forms and the standard

[a] In 4 dimensions there is a non-vanishing coupling term between a 1-form and a 2-form, which can support anisotropic inflation 10. We leave the study of these term in the context of dark energy for a forthcoming publication.

[b] The coupling functions $f_i(\phi)$ for each $p$-form are in general different. Here we assume them equal just for simplicity.

matter, respectively.

$$T^{\phi}_{\mu\nu} = \partial_\mu\phi\partial_\nu\phi - \frac{1}{2}g_{\mu\nu}\partial_\sigma\phi\partial^\sigma\phi - g_{\mu\nu}V(\phi), \tag{5}$$

$$T^{p}_{\mu\nu} = f^2\left[F_{1\,\nu\gamma}F_{1\,\mu}^{\ \gamma} + F_2^{\alpha\beta}{}_{\nu}F_{2\,\alpha\beta\mu} - g_{\mu\nu}\left(\frac{1}{4}F_1^2 + \frac{1}{12}F_2^2\right)\right], \tag{6}$$

with the shorthand notation $F_1^2 \equiv F_{1,\mu_1\mu_2}F^{1,\mu_1\mu_2}$ and $F_2^2 \equiv F_{2,\mu_1\mu_2\mu_3}F^{2,\mu_1\mu_2\mu_3}$. For the matter sector we assume a perfect fluid contribution $T^m_{\mu\nu} = \text{diag}\,(\rho_f, p_f, p_f, p_f)$ with $\rho_f$ the energy density and $p_f$ the preassure. Taking into account the Lagrangian of the scalar field $\mathcal{L}_\phi$ given in eq. (1), variation w.r.t. $\phi$ gives

$$\Box\phi - V_{,\phi} + 2ff_{,\phi}\left(F_1^2 + F_2^2\right) = 0, \tag{7}$$

with $f_{,\phi} \equiv \frac{\mathrm{d}f(\phi)}{\mathrm{d}\phi}$. In which follows, we will use the gauge freedom $A_0 = \partial^i A_i = 0$, to choose the vector field along the $x$ direction $A_1 = A_1(t)\mathrm{d}x$, and the 2-form along the plane $y-z$, this is $A_2 = A_2(t)\mathrm{d}y\wedge\mathrm{d}z$ 6. Thus, due to the rotational symmetry of $A_{1\,\mu_1}$ and $A_{2\,\mu_1\mu_2}$ we use a Bianchi I metric:

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + e^{2\alpha(t)}\left[e^{-4\sigma(t)}\mathrm{d}x^2 + e^{2\sigma(t)}(\mathrm{d}y^2 + \mathrm{d}z^2)\right], \tag{8}$$

being $e^\alpha \equiv a$, with $a$ the scale factor, and $\sigma$ the spatial shear. The equation of motion (e.o.m.) for the fields $A_1$ and $A_2$ are

$$\ddot{A}_1 + \left[2\frac{f'}{f}\dot{\phi} + \dot{\alpha} + 4\dot{\sigma}\right]\dot{A}_1 = 0, \qquad \ddot{A}_2 + \left[2\frac{f'}{f}\dot{\phi} - (\dot{\alpha} + 4\dot{\sigma})\right]\dot{A}_2 = 0, \tag{9}$$

the solutions are simply

$$\dot{A}_1 = \tilde{p}_1 f(\phi)^{-2}e^{-\alpha-4\sigma}, \qquad \dot{A}_2 = \tilde{p}_2 f(\phi)^{-2}e^{\alpha+4\sigma}, \tag{10}$$

with $\tilde{p}_1$ and $\tilde{p}_2$ integration constants. If we define the energy densities of the $p$-forms as

$$\rho_1 = \frac{f^2}{2}e^{-2\alpha+4\sigma}\dot{A}_1^2, \qquad \rho_2 = \frac{f^2}{2}e^{-4\alpha-4\sigma}\dot{A}_2^2, \tag{11}$$

the Friedmann equations, coming from eq. (4), and the e.o.m for the scalar field could be written as

$$\dot{\alpha}^2 = \dot{\sigma}^2 + \frac{1}{3M_{\rm pl}^2}\left[\rho_m + \rho_r + \frac{1}{2}\dot{\phi}^2 + V(\phi) + \rho_1 + \rho_2\right], \tag{12}$$

$$\ddot{\alpha} = -3\dot{\alpha}^2 + \frac{1}{M_{\rm pl}^2}\left[\frac{\rho_m}{2} + \frac{\rho_r}{3} + V(\phi) + \frac{1}{3}\rho_1 + \frac{2}{3}\rho_2\right], \tag{13}$$

$$\ddot{\sigma} = -3\dot{\alpha}\dot{\sigma} + \frac{1}{M_{\rm pl}^2}\left[\frac{2}{3}\rho_1 - \frac{2}{3}\rho_2\right], \tag{14}$$

$$\ddot{\phi} = -3\dot{\alpha}\dot{\phi} - V_{,\phi} + 2\frac{f_{,\phi}}{f}(\rho_1 + \rho_2), \tag{15}$$

where we take into account contributions of non-relativistic matter and radiation.

## 4. Cosmological dynamics

Let us introduce the following dimensionless quantities

$$\Sigma \equiv \frac{\dot{\sigma}}{\dot{\alpha}}, \quad X \equiv \frac{\dot{\phi}}{\sqrt{6}M_{\rm pl}H}, \quad Y \equiv \frac{\sqrt{V}}{\sqrt{3}M_{\rm pl}H}, \quad \Omega_m \equiv \frac{\rho_m}{3M_{\rm pl}^2H^2}, \tag{16}$$

$$\Omega_r \equiv \frac{\rho_r}{3M_{\rm pl}^2H^2}, \quad \Omega_1 \equiv \frac{\rho_1}{3M_{\rm pl}^2H^2}, \quad \Omega_2 \equiv \frac{\rho_2}{3M_{\rm pl}^2H^2}, \tag{17}$$

where $H \equiv \dot{\alpha}$. Thus, eq. (12) can be written as $\Omega_m = 1 - \Sigma^2 - \Omega_{DE} - \Omega_r$, where $\Omega_{DE} \equiv X^2 + Y^2 + \Omega_1 + \Omega_2$, is the dark energy density parameter . The effective equation of state (e.o.s.) is defined as $w_{\rm eff} \equiv -1 - \frac{2\dot{H}}{3H^2}$, where the ratio $\frac{\dot{H}}{H^2}$ can be computed from eq. (13) as

$$\frac{\dot{H}}{H^2} = -\frac{1}{2}\left(3 + 3X^2 - 3Y^2 + 3\Sigma^2 + \Omega_1 - \Omega_2 + \Omega_r\right). \tag{18}$$

In addition, we define the dark energy density and pressure as

$$\rho_{\rm DE} = \frac{\dot{\phi}}{2} + V(\phi) + \rho_1 + \rho_2 + 3M_{\rm pl}^2H^2\Sigma^2, \tag{19}$$

$$p_{\rm DE} = \frac{\dot{\phi}}{2} - V(\phi) + \frac{\rho_1}{3} - \frac{\rho_2}{3} + 3M_{\rm pl}^2H^2\Sigma^2. \tag{20}$$

The e.o.s for dark energy becomes

$$w_{\rm DE} \equiv \frac{p_{\rm DE}}{\rho_{\rm DE}} = \frac{3(X^2 - Y^2 + \Sigma^2) + \Omega_1 - \Omega_2}{3(X^2 + Y^2 + \Sigma^2 + \Omega_1 + \Omega_2)}. \tag{21}$$

In order to get a closed system of equations, is necessary to define explicitly the form of the potential $V(\phi)$ and the coupling function $f(\phi)$. We choose them to be of exponential type $V(\phi) \propto e^{-\frac{\lambda\phi}{M_{\rm pl}}}$, $f(\phi) \propto e^{-\frac{\mu\phi}{M_{\rm pl}}}$, where $\lambda$ and $\mu$ are dimensionless constants 7, 8. Thus, by differentiating w.r.t the number of $e$-folds $N \equiv \ln a$ each one of the variables given in eq. (16) and eq. (17) we find

$$\Sigma' = \frac{\Sigma}{2}\left(3X^2 - 3Y^2 + 3\Sigma^2 - 3 + \Omega_1 - \Omega_2 + \Omega_r\right) + 2\Omega_1 - 2\Omega_2, \tag{22}$$

$$X' = \frac{3X}{2}\left(X^2 - Y^2 + \Sigma^2 - 1 + \frac{\Omega_1}{3} - \frac{\Omega_2}{3} + \frac{\Omega_r}{3}\right) - \sqrt{6}\left(\mu(\Omega_1 + \Omega_2) - \frac{\lambda Y^2}{2}\right), \tag{23}$$

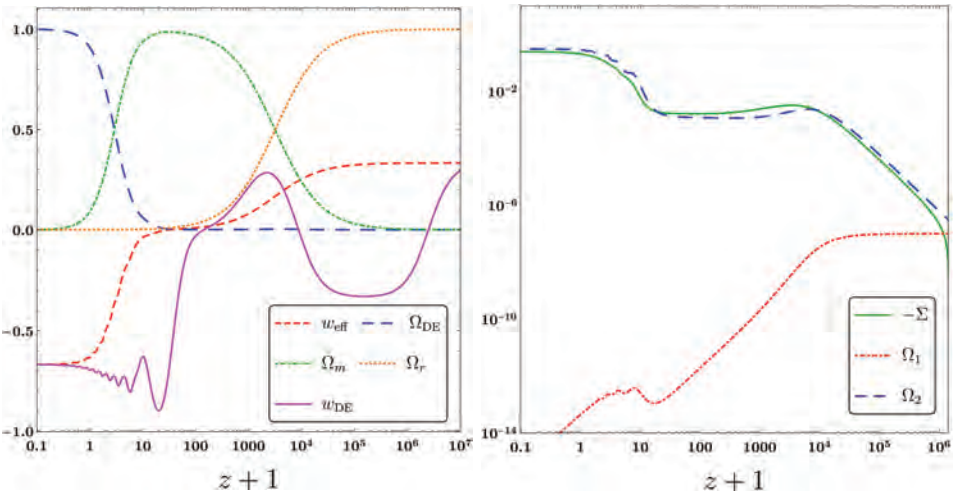
$$Y' = \frac{Y}{2}\left(3X^2 - 3Y^2 + 3\Sigma^2 + 3 + \Omega_1 - \Omega_2 + \Omega_r - \sqrt{6}\lambda X\right), \tag{24}$$

$$\Omega_1' = \Omega_1\left(3X^2 - 3Y^2 + 3\Sigma^2 + 4\Sigma - 1 + 2\sqrt{6}\mu X + \Omega_1 - \Omega_2 + \Omega_r\right), \tag{25}$$

$$\Omega_2' = \Omega_2\left(3X^2 - 3Y^2 + 3\Sigma^2 + 4\Sigma + 1 + 2\sqrt{6}\mu X + \Omega_1 - \Omega_2 + \Omega_r\right), \tag{26}$$

$$\Omega_r' = \Omega_r\left(3X^2 - 3Y^2 + 3\Sigma^2 - 1 + \Omega_1 - \Omega_2 + \Omega_r\right). \tag{27}$$

Instead of the standard analysis of critical points and stability for the previous autonomous system, we decide the make numerical integrations of the equations to obtain a general behavior of the solutions. A complete analysis of this system goes




(a) (b)

Fig. 1. (a) Evolution of $\Omega_{\rm DE}$, $\Omega_r$, $\Omega_m$, $w_{\rm DE}$, and $w_{\rm eff}$ versus $z+1$ for $\lambda = \mu = 10$, with the initial conditions $X = 10^{-13}$, $Y = 10^{-14}$, $\Sigma = 0$, $\Omega_1 = 10^{-7}$, $\Omega_2 = 10^{-10}$, and $\Omega_r = 0.99996$ at the redshift $z = 7.9 \times 10^7$. (b) Evolution of $-\Sigma$, $\Omega_1$ and $\Omega_2$ versus $z+1$ for $\lambda = \mu = 10$ with the same set of initial conditions as figure (a).

beyond this short paper and is left for a forthcoming work. In Fig. 1 we shown the numerical integration of the set of eqs. (22) to (27) where the couplings constants were fixed to be $\lambda = \mu = 10$. Typically, we search for a transition from radiation dominance, to matter dominance, and finllay reach an epoch dominated by dark energy today. In terms of effective equation of state $w_{\rm eff}$ those transitions are of the type $w_{\rm eff} \sim \frac{1}{2} \longrightarrow w_{\rm eff} \sim 0 \longrightarrow w_{\rm eff} \leq -\frac{1}{3}$, as we can see in Fig. 1a, where the sequence is observed in terms of the density parameters. For large redshifts the radiation dominates and start to decrease at an approximate redshift of $z \sim 3000$ to a matter dominated epoch ($\Omega_m \sim 1$), as expected. The evolution of the shear $\Sigma$ and the density parameters $\Omega_1$ and $\Omega_2$ are shown in Fig. 1b. The contribution of the 2-form always dominate over the 1-form, except for the initial condition at high redshift ($\sim 10^7$); the shear increases until a constant value in the present time. In contrast with the case presented in Ref. 8, where only the 2-form is considered, $w_{\rm eff}$ and $w_{\rm DE}$ we will not reach the asymptotic value $\sim -1$. Nevertheless, we check numerically by evolving $w_{\rm DE}$ with different (larger) values of the coupling $\mu$, that this value is realized.

As we anticipated, a complete analysis including the stability of fixed points of the autonomous system, will be presented in a future work, where we also want to elucidate the effect of a non-vanishing coupling between the 1- and 2-form fields.

**Acknowledgments**

This work was supported by COLCIENCIAS Grant No. 110656933958 RC 0384-2013 and by COLCIENCIAS grant 110278258747 RC-774-2017 (DAAD-Procol program). AG also acknowledges financial support from the MG15 organizing committee and from the ESA travel funds.

**References**

1. A. H. Guth and S.-Y. Pi, Fluctuations in the new inflationary universe, *Phys. Rev. Lett.* **49**, 1110 (1982).
2. A. Starobinsky, Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations, *Phys Lett. B* **117**, 175 (1982).
3. P. A. R. Ade et al., Planck 2015 results. XVI. Isotropy and statistics of the CMB, *Astron. Astrophys.* **594** A16 (2016).
4. M.-A. Watanabe, S. Kanno and J. Soda, Inflationary Universe with Anisotropic Hair, *Phys. Rev. Lett.* **102**, 191302 (2009).
5. K. Dimopoulos and M. Karciauskas, Parity Violating Statistical Anisotropy, *JHEP* **06**, 040 (2012).
6. J. Ohashi, J. Soda and S. Tsujikawa, Observational signatures of anisotropic inflationary models, *JCAP* **1312**, 009 (2013).
7. M. Thorsrud, D. F. Mota and S. Hervik, Cosmology of a Scalar Field Coupled to Matter and an Isotropy-Violating Maxwell Field, *JHEP* **10**, 066 (2012).
8. J. P. Beltrán Almeida, A. Guarnizo, R. Kase, S. Tsujikawa and C. A. Valenzuela-Toledo, Anisotropic 2-form dark energy, *Phys. Lett. B* **793**, 396 (2019).
9. J. P. Beltrán Almeida, A. Guarnizo and C. A. Valenzuela-Toledo, Arbitrarily coupled $p$-forms in cosmological backgrounds, `[arXiv:1810.05301 [astro-ph.CO]]` (2018).
10. J. P. Beltrán Almeida, A. Guarnizo, R. Kase, S. Tsujikawa and C. A. Valenzuela- Toledo, Anisotropic inflation with coupled $p$-forms, *JCAP* **1903**, 025 (2019).

# Cosmological constant: the possible nature and relationship with quantum gravity and fundamental particle masses

Ho-Ming Mok

*RHU, Hong Kong SAR, China*
*E-mail: a8617104@graduate.hku.hk*

In order to reconcile the quantum mechanics and general relativity, the equivalence principle of quantum gravity is introduced for extending the equivalence principle of general relativity to the observer frames of reference which are in quantum mechanical motions. The equivalence principle of quantum gravity is that the laws of physics must be of such a nature that they apply to systems of reference in any kind of motions, both classical and quantum mechanical. Under such principle, the quantum gravity should be formulated in the quantum space-time-matter space with local conformal symmetry and the mathematical expressions for the cosmological constant as well as the masses of fundamental particles can be found by the theory.

*Keywords*: Cosmological constant, equivalence principle of quantum gravity, Higgs condensate, quantum space-time-matter geometry, local conformal symmetry, Weyl geometry.

## 1. Introduction

It is well known that the two main pillars of modern physics, the general relativity (GR) and quantum mechanics (QM), are fundamentally inconsistent with each other. In order to reconcile the QM and GR, the equivalence principle of quantum gravity is introduced for extending the equivalence principle of GR to the observer frames of reference which are in quantum mechanical motions[1]. The equivalence principle of quantum gravity is that the laws of physics must be of such a nature that they apply to systems of reference in any kind of motions, both classical and quantum mechanical. That means the quantum state of a particle is relative to the observer frame of reference which can also be itself in a quantum mechanical state. Under the principle, the quantum gravity should be formulated in the quantum space-time-matter (QSTM) space with local conformal symmetry. It was also expected that the ultimate theory of quantum gravity would explain the cosmological constant problem[2,3].

## 2. Quantum space-time-matter geometry

In relativity, the space and time are defined by the coordinate system established by measuring-rods and synchronised clocks, which are in classical motions and their measuring results correspond to the points in the 4-dimensional spacetime coordinate system. However, the measuring-rods and clocks themselves are subject to the quantum uncertainty as well as quantum mechanical motions. Therefore, the space-time cannot simply be a 4-dimensional coordinate system with spacetime points but should be associated with the quantum states of measuring-rods and clocks specified

by their state parameters. The quantum state of the measuring-rod and clock, that is the quantum state of the spacetime by definition, can be defined[4] as $|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle$, where $\overrightarrow{x}$ is the classical spacetime vector and $\Delta\overrightarrow{x}$ is the associated uncertainty. As the quantum uncertainties depends upon the energy scales, the modified definition of spacetime should therefore be observation energy scale dependent. The quantum state $|\alpha\rangle$ of a particle can be projected to such quantum spacetime state to become a wavefunction $\phi(\overrightarrow{x}, \Delta\overrightarrow{x})$ as $\phi(\overrightarrow{x}, \Delta\overrightarrow{x}) = \langle \overrightarrow{x}, \Delta\overrightarrow{x}|\alpha\rangle$. Furthermore, the particle quantum state can be observed in another quantum spacetime state $|\overrightarrow{y}, \Delta\overrightarrow{y}\rangle$ which can be expressed in general as linear combination of the eigenstate of $|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle$. The transformation can be expressed as

$$\langle \overrightarrow{y}, \Delta\overrightarrow{y}|\alpha\rangle = \int \langle \overrightarrow{y}, \Delta\overrightarrow{y}|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle \langle \overrightarrow{x}, \Delta\overrightarrow{x}|\alpha\rangle \tag{1}$$

The integration sign denotes the summation or integration of all the possible eignestates of $|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle$. Actually, such notation can be omitted by identifying the operator $|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle\langle\overrightarrow{x}, \Delta\overrightarrow{x}|$ as the internal index of the summation or integration as analogous to the Einstein's convention.

For introducing the geometrical concept to QM and the equivalence priniple of quantum gravity in quantum gravity, the spacetime should be merged with matter together to become the QSTM space with local conformal symmetry under the Weyl like geometry[1,4]. As a consequence, all the physical laws apply not just to the frames of reference in any kind of classical motion as in GR but also the same in the reference frames in quantum mechanical motions as well. It implies that the general physical laws of nature are required to be expressed by equations which hold good for all systems of quantum spacetime states which are covariant under the transformation in Eq (1) and the quantum mechanical motion is relative in nature. The QSTM state $|\Psi\rangle$ is the combination of the quantum state of particle and spacetime as $|\Psi\rangle = |\phi\rangle|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle$. The length of a general small line element $dL$ can be defined in a small region of QSTM space as the inner product of $\delta|\Psi\rangle$ as

$$dL^2 = \delta|\Psi\rangle \cdot \delta|\Psi\rangle = \delta(|\phi\rangle|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle) \cdot \delta(|\phi\rangle|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle) \tag{2}$$

where the inner product can be defined as the usual inner product used for quantum mechanical states and thus $dL^2$ has a real value. Let us also introduce the operator $\widehat{\delta\overrightarrow{x}_{\Delta x}}$, which can extract the spacetime length from the quantum spacetime state, for small changes on spacetime state with the property as

$$\widehat{\delta\overrightarrow{x}_{\Delta x}}\delta|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle = \delta\overrightarrow{x}_{\Delta x}\delta|\overrightarrow{x}, \Delta\overrightarrow{x}\rangle \tag{3}$$

When there is no local mixing between the quantum matter particle state and the quantum spacetime state, another line element $dl$ can be defined in the QSTM

space and when expressing it in generalised coordinates $\xi_m^\mu$ in the QSTM space gives

$$dl^2 = \delta\left|\phi\right\rangle \widehat{\delta\overrightarrow{x}_{\Delta x}}\delta\left|\overrightarrow{x},\Delta\overrightarrow{x}\right\rangle \cdot \delta\left|\phi\right\rangle \widehat{\delta\overrightarrow{x}_{\Delta x}}\delta\left|\overrightarrow{x},\Delta\overrightarrow{x}\right\rangle = G_{\mu\nu}^{mn} d\xi_m^\mu d\xi_n^\nu \tag{4}$$

where $G_{\mu\nu}^{mn}$ is the combined metric of the QSTM space and the extracted spacetime length and the index $\mu$ and $\nu$ are the usual spacetime indexes in GR while $m$ and $n$ are the indexes associated with the quantum spacetime and matter states. Such formulation works well on discrete eigenstates as well as continuous eignestates. Let us define the local conformal operator $\Omega$ on the QSTM state as $\delta\left|\Psi\right\rangle \rightarrow \Omega\delta\left|\Psi\right\rangle$. By using the closest admissible action in Weyl geometry analogous to the Einstein's action for local conformal symmetry[5],

$$S = \int \Re^2 \sqrt{-G} d^p\xi \tag{5}$$

where $\Re$ is the curvature scalar defined for $G_{\mu\nu}^{mn}$ with the QSTM and $p$ is the total number of dimensions of it. It is a generalized action of quantum gravity in the formulation of the QSTM space. The variation of the action under a suitable local conformal transformation with imposing the gauge $\Re \sim 4\Lambda$ for fixing the scale of the QSTM space, where $\Lambda$ is a constant, gives

$$S = \int \left(\widetilde{\Re} - 2\Lambda\right) \sqrt{-G} d^p\xi \tag{6}$$

In the special case that the QSTM space is a flat space and, for simplicity, assuming that the matter field is a scalar field, the Eq (4) can be written as

$$dl^2 = \left\langle \phi^2 \overline{g}\left(\Delta x\right)_{\mu\nu} \right\rangle dx^\mu dx^\nu \tag{7}$$

where $\left\langle \phi^2 \overline{g}\left(\Delta x\right)_{\mu\nu} \right\rangle$ is the quantum probability weighted spacetime metric, which is specified with measurement uncertainty, combined with the matter field. Also, for a flat space, the metric $G_{\mu\nu}^{mn}$ does not vary with $\xi_m$ and the action can then be expressed in terms of $\overline{g}_{\mu\nu}$ and $\phi$ with variation on spacetime as

$$S = \int \left(\phi^2 \overline{R} - 2\Lambda\phi^4 + 6\left(\partial\phi\right)^2\right) \sqrt{-\overline{g}} d^4x \tag{8}$$

Aligning the term $\left(\partial\phi\right)^2$ as the kinetic term of the $\phi$ field gives

$$S' = \int \left(\frac{1}{12}\phi^2 \overline{R} - \frac{1}{6}\Lambda\phi^4 + \frac{1}{2}\left(\partial\phi\right)^2\right) \sqrt{-\overline{g}} d^4x \tag{9}$$

If $\overline{R}$ is a constant, the spontaneous symmetry breaking (SSB) on the $\phi$ field gives the relation $\overline{R}/4\Lambda = \phi_0^2$. Defining the conformal operator $\omega$, which is associated with the change of scale of observation, in the spacetime on the metric and scalar field as $\overline{g}_{\mu\nu} \rightarrow \omega^2 \overline{g}_{\mu\nu}$ and $\phi \rightarrow \omega^{-1}\phi$. The action in Eq (9) is invariant under

such conformal transformation. The quantum gravity therefore possesses a double conformal structures. This point is very cruical as it allows the change scale in one degree of freedom and then compensate by the other to make the physical laws invariant.

## 3. Particle masses and cosmological constant

Considering a shift field $h$ around the minimum of the potential as $\phi = \phi_0$ with an energy scale $\phi_0^2 = 6m'^2$ in the SSB of the action in Eq (9), the equation becomes

$$S = \int \left( \frac{1}{2} m'^2 \overline{R} - 3\Lambda m'^4 - \frac{1}{2} (\partial h)^2 + \frac{1}{2} \overline{R} h^2 + ... \right) \sqrt{-\overline{g}} d^4x \tag{10}$$

Under the change the scale of observation by applying conformal transformation on the spacetime as $\overline{g}_{\mu\nu} \to (m'/M_p)^2 \overline{g}_{\mu\nu}$ and $\phi \to (m'/M_p)^{-1} \phi$, where the mass scale $M_p$ is the reduced Planck mass, the equation can be expressed as

$$S = \int \left( \frac{1}{2} M_p^2 \overline{R'} - 3\Lambda M_p^4 - \frac{1}{2} (\partial h')^2 + \frac{1}{2} \overline{R'} h'^2 + ... \right) \sqrt{-\overline{g'}} d^4x \tag{11}$$

The term $\overline{R'} h'^2$ acts as the mass term of the excitation field $h'$. Since $\overline{R'}$ is the conformal transformation of the term related to the vacuum energy of spacetime in empty space and since $\overline{R}\sqrt{-\overline{g}}$ should be transformed as $\overline{g}$ and therefore $\overline{R'}$ should be equal to $(m'/M_p)^2 \overline{R}$ which is the particle squared mass term. That imply the existence of a particle with mass $(m'/M_p) \overline{R}^{\frac{1}{2}}$. We can change the observer frame of reference for observing such particle due to the double conformal symmetry and this will lead to the change on the broken scale. Recalling that the scale of the QSTM space is fixed by the gauge $\Re \sim 4\Lambda$ and let the factor for the associated scale change is $\Omega^4$, the scale of $\Lambda$ is therefore changed as $\Lambda \to \Omega^4 \Lambda$. The scale change of $\Lambda$ implies that the SSB condition $\overline{R}/4\Lambda = \phi_0^2$ requires to be varied so that the mass value $(m'/M_p) \overline{R}^{\frac{1}{2}}$ in Eq (11) of the said particle will be changed also. As the mass value is determined by the ratio $\overline{R}\sqrt{-\overline{g}}/\phi^2$ and under the change of scale for Eq (11) following the conformal transformation $\omega$ on the spacetime and matter field induce the transformation as

$$\frac{\Omega^4 \overline{R} \sqrt{-\overline{g}}}{\phi^2} \to \frac{\omega^4 \Omega^4 \overline{R} \sqrt{-\overline{g}}}{\phi^2} \tag{12}$$

Therefore, for keeping the mass unchanged through the change of observer frame of reference requires that $\omega = 1/\Omega$. The structure of the QSTM space under broken symmetry can be explored by repeatedly performing the above processes of changing observer frame such that in every changed frame the mass of the particle are still equal to $m'/M_p$, although the $M_p$ varies with the transformed frame of reference. A set of scale factors can be found for the QSTM space scale

$$(\Omega)^{2n} = \left( \frac{m'}{M_p} \right)^2 \Rightarrow \Omega = \left( \frac{m'}{M_p} \right)^{\frac{1}{n}} \tag{13}$$

where $n$ is the number of the said transformation to achieve the mass factor $m'/M_p$ in a frame. Suppose $m'/M_p \sim 10^{-15}$, that is $m' = 2.435TeV$, and since the particle mass $m_i$ is equal to $\Omega_{mi}\overline{R}^{\frac{1}{2}}$, if $\overline{R}^{\frac{1}{2}} = 10TeV$, a set of mass values can be found and compared with the lepton masses data of the Particle Data Group[7] as $m_1 = 10^{-15}\overline{R}^{\frac{1}{2}} = 0.01eV \sim m_\nu$?; $m_2 = 3.16 \times 10^{-8}\overline{R}^{\frac{1}{2}} = 316.23keV \sim m_e$; $m_3 = 10^{-5}\overline{R}^{\frac{1}{2}} = 100MeV \sim m_\mu$; $m_4 = 1.778 \times 10^{-4}\overline{R}^{\frac{1}{2}} = 1.778GeV \sim m_\tau$, where $m_\nu$, $m_e$, $m_\mu$ and $m_\tau$ are respectively the mass of neutrinos, electron, muon and tau lepton. The actual neutrino masses might not necessary be degenerate as the radiative corrections is not considered in the above calculation. The $m_2$ state is of the order of the electron mass and the difference could be due to the radiative correction to the QED vacuum. The $m_3$ state is equal to the muon mass up to about 5 percent and the $m_4$ state is even equal to the tau lepton mass up to only about 0.06 percent. For the mass states with $n > 4$, for instance, the mass value of $m_5$ state is consistent with the dark matter candidate with a mass of 10 GeV as proposed by some researchers[8]. In extending the argument to half-integer value of $n$, $m_{4.5} = 4.64GeV$ and $m_{8.5} = 171.9GeV$. The values are very close to the mass values of bottom and top quarks. That means the half integer values of $n$ may correspond to the quark masses. However, due to the QCD vacuum, the calculated values for the lighter quark masses would have greater discrepancy with the actual current quark masses.

As the Eq. (9) and Eq. (11) is related by the conformal transformation of spacetime and the matter field and the SSB which is determined by the observation scale. In fact, the Eq (9) resembles the following action of the Higgs potential when $\overline{R}$ is a constant in one hand

$$S' = \int \left(\frac{1}{2}\mu^2\phi^2 - \frac{1}{4}\lambda\phi^4 + \frac{1}{2}(\partial\phi)^2\right)\sqrt{-\overline{g}}d^4x \tag{14}$$

On the other hand, it can become the action of gravitational field coupled to matter field after the local conformal transformation and SSB as discussed above. Comparing the coefficients between the Eq (9) and Eq (14) gives $\overline{R}/12 = \mu^2/2$ and $\Lambda/6 = \lambda/4$. Furthermore, by comparing Equation (11) with the gravitational action and eliminating $\Lambda$ in the above equations, we get $3\Lambda M_p^2 = \Lambda_{cc}$ and $\Lambda_{cc} = 4.5\lambda M_p^2$, where $\Lambda_{cc}$ is the cosmological constant. If then applying a conformal transformation on the QSTM space metric to change the observer frame of reference, under the same factor as the one in obtaining the fundamental particle masses with $n = 2$, and since the cosmological constant transforms as the square of the metric, that is conformal factor $\Omega^4$, with respect to the conformal factor of $\Omega^2$ on the metric, we have

$$\Omega = \left(\frac{m'}{M_p}\right)^2 \Rightarrow \Lambda_{cc} = 4.5\lambda\frac{m'^4}{M_p^2}\left(\frac{m'}{M_p}\right)^4 \tag{15}$$

By putting $\lambda = 0.258$ which is based on the 125 GeV Higgs particle mass and 246 GeV electroweak VEV value, we get $\Lambda_{cc} = 6.49 \times 10^{-66}eV^2$. Such calculated

cosmological constant value is in very good agreement with the observation value of $4.33 \times 10^{-66} eV^2$ of Planck CMB probe when connecting to the fundamental particle masses.

## 4. Discussions and conclusions

The theory of quantum gravity is formulated in the QSTM space with local conformal symmetry under the equivalence principle of quantum gravity. The QSTM space possesses a double conformal symmetry geometrical structure. Different exponent $n$ values of the associated conformal factor $\Omega = (m'/M_p)$ gives the fundamental particle mass values and the observed cosmological constant value with very good agreement. It indicates that the fundamental particle masses and the cosmological constant are related in the theory. Furthermore, the above mass formula allows the existence of some mass states that are so far not experimentally observed that might provide new opportunities for discussing the nature of such mass states.

## References


1. H. M. Mok, "Cosmological Constant Problem and Equivalence Principle of Quantum Gravity" in *Proceedings of the 13th Marcel Grossmann Meeting on General Relativity, Stockholm, 2012* (World Scientific, Singapore, 2015).
2. H. M. Mok, "A Possible Solution to the Cosmological Constant Problem by Discrete Space-time Hypothesis" in *On the Nature of Dark Energy - Proceedings of the 18th Institut d'Astropysique de Paris Colloquium* (Frontier Group, 2002) (arXiv:astro-ph/0105513).
3. H. M. Mok, "The Solution of Cosmological Constant Problem by Discrete Space-time at Electroweak Scale" in *Proceedings of the 12th Marcel Grossmann Meeting on General Relativity, Paris, 2009* (World Scientific, Singapore, 2012).
4. H. M. Mok, "Cosmological Constant and Particle Masses in Conformal Quantum Gravity" in *Cosmology* (IntechOpen, London, 2018) (DOI:10.5772/intechopen.81875).
5. R. Adler *et al.*, *Introduction to General Relativity*, (McGraw-Hill Kogakusha, Tokyo, 1975).
6. C. Rovelli, *Quantum Gravity*, (Cambridge University Press, Cambridge, 2004).
7. Tanabashi *et al.* (Particle Data Group), *Phys. Rev. D* **98**, 030001 (2018).
8. D. Hooper, *Dark Universe* **1**, 1-23 (2012).

# Some cosmological solutions of $F(R)$ gravity with $f$-essence

K. R. Myrzakulov*, S. S. Bekov and R. Myrzakulov

*Eurasian International Center for Theoretical Physics,*
*L. N. Gumilyov Eurasion National University*
*Nur-Sultan, 010008, Kazakhstan*
**E-mail: krmyrzakulov@gmail.com*
*www.enu.kz*

In this work, we study the $F(R)$ gravity with $f$ -essence for the flat and homogeneous Friedman-Robertson-Walker universe. For this model, we have presented the point-like Lagrangian and the corresponding field equations. To describe the dynamics of the universe, we have investigated some cosmological solutions for $K$, $F$ and $h$ functions. It is shown that these solutions describe the late time accelerated expansion of the Universe.

*Keywords*: Cosmology, FRW metric, $F(R)$ gravity, fermion field, scale factor.

## 1. Introduction

To theoretically investigate the accelerated expansion of the Universe, many alternative gravity theories have been proposed. One of these theories is $F(R)$ gravity, where $R$ is a Ricci scalar [1–3]. In this theory, the Einstein-Hilbert action is generalized to action for $F(R)$, i.e. as arbitrary function of $R$. On the other hand, it is known that higher degrees of $R$ in the Einstein-Hilbert action can lead to an inflation effect accelerated expansion in the early stages of the Universe [4]. Each form of $F(R)$ generates a new cosmological model. So it is important that this theory be tested for compliance not only with cosmological, but also with other requirements, such as stability of solutions.

As one of the other modifications in modern cosmological theory, is additional fermions fields (Dirac or spinor fields) are considered, such candidates for inflation or dark energy have been studied in the works [5,6]. Some cosmological model with fermion field in teleparallel gravity studied in works [7,8]. And in the work [9] were study the cosmic expansion from boson and fermion fields. Generalized fermion fields were considered in [10] and known as $f$-essence. In work [11] was considered some cosmological model in $F(T)$ gravity with $f$-essence. Starobinsky model with $f$-essence was considered in work [12].

This paper organizing as following: In Sect. 2 we review $F(R)$ gravity. In Sect. 3 we consider the generalization of $F(R)$ gravity with $f$-essence and derived field equations. In Sect. 4 we apply Noether symmetry to find the form of $F(R)$ gravity, coupling function and potential. The cosmological solutions are investigated in Sect. 5.

The signature of space-time metric used is $(-,+,+,+)$ and units have been chosen so that $8\pi G = c = \hbar = k = 1$.

## 2. $F(R)$ gravity

For four dimensions space-time, the action in $F(R)$ gravity with matter is written as

$$S = \int d^4x\sqrt{-g}\,[F(R) + L_m] \tag{1}$$

were $g = det(g_{ik})$ is the determinant of the metric tensor matrix, $F(R)$ is a differentiable function of the Ricci scalar $R$, $L_m$ is the matter Lagrangian and for our case $8\pi G = 1$. Variation of the action (1) with respect to metric tensor $g_{ik}$ yields following field equations

$$F'(R)R_{ik} - \frac{1}{2}g_{ik}F(R) - \nabla^i\nabla_k F'(R) + g_{ik}\Box F'(R) = kT_{ik} \tag{2}$$

where the prime a symbol means differentiation with respect to $R$ and $R_{ik}$ is a Riemann curvature tensor and $k$ is a coupling constant in gravitational units. Also, here $\Box = \nabla^i\nabla_k$ with $\nabla_k$ is a covariant derivative. Here $T_{\mu\nu}$ is the energy-momentum tensor and write this as

$$T_{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta\left(\sqrt{-g}L_m\right)}{\delta g^{\mu\nu}} = g_{\mu\nu}L_m - 2\frac{\partial L_m}{\partial g^{\mu\nu}}. \tag{3}$$

## 3. $F(R)$ gravity with $f$-essence

Action for model with a generalized fermionic field that is non-minimally coupled to gravity in the framework of $F(R)$ gravity written as

$$S = \int d^4x\sqrt{-g}\,\{h(u)F(R) + 2K(Y,u)\}\,, \tag{4}$$

where $F$ is some function of the scalar Ricci $R$, $K$ is the Lagrangian density of the fermionic field, $u = \bar{\psi}\psi$ where $\psi$ and $\bar{\psi} = \psi^\dagger\gamma^0$ denote the fermion field and its adjoint, respectively, the dagger represents complex conjugation, $h(u)$ is generic function, representing the coupling with gravity and fermionic field.

The canonical kinetic term has the form

$$Y = \frac{1}{2}i\left[\bar{\psi}\Gamma^\mu D_\mu\psi - (D_\mu\bar{\psi})\Gamma^\mu\psi\right], \tag{5}$$

where $\Gamma^\mu = e^\mu_a\gamma^a$ are the generalized Dirac-Pauli matrices satisfying the Clifford algebra $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}$, where the braces denote the anti-commutation relation, the covariant derivatives $e^a_\mu$ are given by

$$D_\mu\psi = \partial_\mu\psi - \Omega_\mu\psi,\ D_\mu\bar{\psi} = \partial_\mu\bar{\psi} + \bar{\psi}\Omega_\mu.$$

Above, the fermionic connection $\Omega_\mu$ is defined by

$$\Omega_\mu = -\frac{1}{4}g_{\rho\sigma}\left[\Gamma^\rho_{\mu\delta} - e^\rho_b\partial_\mu e^b_\delta\right]\Gamma^\sigma\Gamma^\delta,$$

with $\Gamma^\rho_{\mu\delta}$ denoting the Christoffel symbols.

We will consider the flat, homogeneous and isotropic Fridmann-Robertson-Walker (FRW) metric in the folowing form:

$$ds^2 = -dt^2 + a^2(t)\left(dx^2 + dy^2 + dz^2\right), \tag{6}$$

where $a(t)$ is the scale factor of the Universe. Then for this metric we can define the following expressions

$$\sqrt{-g} = a^3,\ R = 6\left(\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2}\right),\ Y = 0.5i\left(\bar{\psi}\gamma^0\dot{\psi} - \dot{\bar{\psi}}\gamma^0\psi\right).$$

where the dot represents differentiation with respect to time $t$. For the FRW metric (6) the point-like Lagrangian we can write the following form

$$L = a^3hF - a^3hRF_R - 6a\dot{a}^2hF_R - 6a^2\dot{a}F_Rh^{'}\dot{u} - 6a^2\dot{a}\dot{R}hF_{RR} + 2a^3K. \tag{7}$$

here the spinor field depends only on time and the prime denotes the derivative with respect to the bilinear function $u$. Then for the action (4) and of the metric (6) the equations of motion are defined as

$$3\frac{\dot{a}}{a}\dot{R}F_{RR} + \left(3\frac{\dot{a}^2}{a^2} + 3\frac{\dot{a}}{a}\frac{\dot{h}}{h} - \frac{1}{2}R\right)F_R + \frac{1}{2}F = \frac{1}{h}\rho_f, \tag{8}$$

$$F_{RRR}\dot{R}^2 + \left[\ddot{R} + 2\left(\frac{\dot{a}}{a} + \frac{\dot{h}}{h}\right)\dot{R}\right]F_{RR} + \left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{\ddot{h}}{h} + 2\frac{\dot{a}}{a}\frac{\dot{h}}{h} - \frac{1}{2}R\right)F_R + \frac{1}{2}F = -\frac{1}{h}p_f, \tag{9}$$

$$K_Y\dot{\psi} + \frac{1}{2}\left(3\frac{\dot{a}}{a}K_Y + \dot{K_Y}\right)\psi - iK_u\psi\gamma^0 + \\ + \frac{i}{2}\left[\left(R - 6\frac{\ddot{a}}{a} - 6\frac{\dot{a}^2}{a^2}\right)h^{'}\psi\gamma^0 - 6\frac{\dot{a}}{a}\left(h^{'}\right)_t\psi\gamma^0 + 6\frac{\dot{a}}{a}\left(h^{'}\right)_{\bar{\psi}}\dot{u}\gamma^0\right]F_R - \frac{i}{2}Fh^{'}\psi\gamma^0 = 0, \tag{10}$$

$$K_Y\dot{\bar{\psi}} + \frac{1}{2}\left(3\frac{\dot{a}}{a}K_Y + \dot{K_Y}\right)\bar{\psi} + iK_u\bar{\psi}\gamma^0 - \\ - \frac{i}{2}\left[\left(R - 6\frac{\ddot{a}}{a} - \frac{\dot{a}^2}{a^2}\right)h^{'}\bar{\psi}\gamma^0 - 6\frac{\dot{a}}{a}\left(h^{'}\right)_t\bar{\psi}\gamma^0 + 6\frac{\dot{a}}{a}\left(h^{'}\right)_{\psi}\dot{u}\gamma^0\right]F_R + \frac{i}{2}Fh^{'}\bar{\psi}\gamma^0 = 0, \tag{11}$$

where $\rho_f = V$ is the energy density and $p_f = YK_Y - K$ is the pressure of the fermion field. In the following section we will use the Noether symmetry approach for find form function $F(R)$, the coupling function $h(u)$ and function $K(Y, u)$.

## 4. The Noether symmetries approach

One can express the Lagrangian for metric (6) in terms of the components of the spinor field $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)^T$ and its adjoint $\bar{\psi} = ({\psi_1}^\dagger, {\psi_2}^\dagger, -{\psi_3}^\dagger, -{\psi_4}^\dagger)$, the point-like Lagrangian (7), we can rewritten as follows

$$L = a^3 hF - a^3 hRF_R - 6a\dot{a}^2 hF_R - 6a^2 \dot{a}F_R h^{'} \dot{\psi}_i^\dagger \psi_i - 6a^2 \dot{a}F_R h^{'} \psi_i^\dagger \dot{\psi}_i - 6a^2 h\dot{a}\dot{R}F_{RR} + 2a^3 K. \tag{12}$$

The symmetry an important role in physics because it is directly associated with the consevation laws of a dynamical systen. First used the Noether symmetry approach in cosmology models and gravitational theories in the works[13,14]. This approach tells us that Lie derivative of the Lagrangian function $L$ with respect to a given vector field $\mathbf{X}$ vanishes, i.e.

$$\mathbf{X}L = 0. \tag{13}$$

A vector fied $\mathbf{X}$ for the Lagrangian (12) is given by

$$\mathbf{X} = \alpha \frac{\partial}{\partial a} + \beta \frac{\partial}{\partial R} + \dot{\alpha} \frac{\partial}{\partial \dot{a}} + \dot{\beta} \frac{\partial}{\partial \dot{R}} + \sum_{j=0}^{3} \left( \eta_j \frac{\partial}{\partial \psi_j} + \dot{\eta}_j \frac{\partial}{\partial \dot{\psi}_j} + \chi_j \frac{\partial}{\partial \psi_j^\dagger} + \dot{\chi}_j \frac{\partial}{\partial \dot{\psi}_j^\dagger} \right), \tag{14}$$

Here

$$\dot{\alpha} = \frac{\partial \alpha}{\partial a}\dot{a} + \frac{\partial \alpha}{\partial R}\dot{R} + \sum_{j=0}^{3} \left( \frac{\partial \alpha}{\partial \psi_j}\dot{\psi}_j + \frac{\partial \alpha}{\partial \psi_j^\dagger}\dot{\psi}_j^\dagger \right), \tag{15}$$

$$\dot{\beta} = \frac{\partial \beta}{\partial a}\dot{a} + \frac{\partial \beta}{\partial R}\dot{R} + \sum_{j=0}^{3} \left( \frac{\partial \beta}{\partial \psi_j}\dot{\psi}_j + \frac{\partial \beta}{\partial \psi_j^\dagger}\dot{\psi}_j^\dagger \right), \tag{16}$$

$$\dot{\eta}_j = \frac{\partial \eta_j}{\partial a}\dot{a} + \frac{\partial \eta_j}{\partial R}\dot{R} + \sum_{j=0}^{3} \left( \frac{\partial \eta_j}{\partial \psi_j}\dot{\psi}_j + \frac{\partial \eta_j}{\partial \psi_j^\dagger}\dot{\psi}_j^\dagger \right), \tag{17}$$

$$\dot{\chi}_j = \frac{\partial \chi_j}{\partial a}\dot{a} + \frac{\partial \chi_j}{\partial R}\dot{R} + \sum_{j=0}^{3} \left( \frac{\partial \chi_j}{\partial \psi_j}\dot{\psi}_j + \frac{\partial \chi_j}{\partial \psi_j^\dagger}\dot{\psi}_j^\dagger \right), \tag{18}$$

where $\alpha, \beta, \eta_j$ and $\chi_j$ are unknown functions of the variables $a, R, \psi_j$ and $\psi_j^\dagger$. Then we can written at the following system of differential equations as

$$\alpha F_R + \beta a F_{RR} + 2aF_R \frac{\partial \alpha}{\partial a} + a^2 F_{RR} \frac{\partial \beta}{\partial a} + + aF_R \frac{h_u}{h} \sum_{i=0}^{3} \left( \epsilon_i \eta_j \psi_i^\dagger + \epsilon_i \chi_j \psi_i \right) + a^2 F_R \frac{h_u}{h} \sum_{i=0}^{3} \left( \frac{\partial \eta_j}{\partial a}\psi_i^\dagger + \frac{\partial \chi_j}{\partial a}\psi_i \right) = 0, \tag{19}$$

$$-6a^2 hF_{RR}\frac{\partial \alpha}{\partial R} = 0, \ -6a^2 F_R h_u \psi_i^\dagger \frac{\partial \alpha}{\partial \psi_j} = 0, \ -6a^2 F_R h_u \psi_i \frac{\partial \alpha}{\partial \psi_j^\dagger} = 0, \tag{20}$$

$$2\alpha F_{RR}+\beta aF_{RRR}+aF_{RR}\frac{\partial\alpha}{\partial a}+2F_R\frac{\partial\alpha}{\partial R}+aF_{RR}\frac{\partial\beta}{\partial R}+aF_{RR}\frac{h_u}{h}\sum_{i=0}^{3}\Big(\epsilon_i\eta_i\psi_i^\dagger+\epsilon_i\chi_i\psi_i\Big)+$$
$$+aF_R\frac{h_u}{h}\sum_{i=0}^{3}\left(\frac{\partial\eta_j}{\partial R}\psi_i^\dagger+\frac{\partial\chi_j}{\partial R}\psi_i\right)=0, \quad (21)$$

$$\left(2\alpha F_R+\beta aF_{RR}+aF_R\frac{\partial\alpha}{\partial a}\right)h_u\psi_i^\dagger+aF_R\sum_{i=0}^{3}\left(\frac{\partial h_u}{\partial\psi_i}\eta_i+\frac{\partial h_u}{\partial\psi_i^\dagger}\chi_i\right)\psi_i^\dagger+aF_R\chi_ih_u+$$
$$+2hF_R\frac{\partial\alpha}{\partial\psi_j}+ahF_{RR}\frac{\partial\beta}{\partial\psi_j}+aF_Rh_u\sum_{i=0}^{3}\left(\frac{\partial\eta_j}{\partial\psi_j}\psi_i^\dagger+\frac{\partial\chi_j}{\partial\psi_j}\psi_i\right)=0, \quad (22)$$

$$\left(2\alpha F_R+\beta aF_{RR}+aF_R\frac{\partial\alpha}{\partial a}\right)h_u\psi_i+aF_R\sum_{i=0}^{3}\left(\eta_i\frac{\partial h_u}{\partial\psi_i}+\chi_i\frac{\partial h_u}{\partial\psi_i^\dagger}\right)\psi_i+aF_R\eta_ih_u+$$
$$+2hF_R\frac{\partial\alpha}{\partial\psi_i^\dagger}+ahF_{RR}\frac{\partial\beta}{\partial\psi_i^\dagger}+aF_Rh_u\sum_{i=0}^{3}\left(\frac{\partial\eta_j}{\partial\psi_j^\dagger}\psi_i^\dagger+\frac{\partial\chi_j}{\partial\psi_j^\dagger}\psi_i\right)=0, \quad (23)$$

$$\sum_{i=0}^{3}\left(F_Rh_u\psi_i^\dagger\frac{\partial\alpha}{\partial R}+hF_{RR}\frac{\partial\alpha}{\partial\psi_i}\right)=0,\ \sum_{i=0}^{3}\left(F_Rh_u\psi_i\frac{\partial\alpha}{\partial R}+hF_{RR}\frac{\partial\alpha}{\partial\psi_i^\dagger}\right)=0, \quad (24)$$

$$\sum_{i=0}^{3}\left(\frac{\partial\alpha}{\partial\psi_j}\psi_i+\frac{\partial\alpha}{\partial\psi_j^\dagger}\psi_i^\dagger\right)=0,\ \sum_{i=0}^{3}\left(\frac{\partial\eta_j}{\partial a}\psi_i^\dagger-\frac{\partial\chi_j}{\partial a}\psi_i\right)=0,\ \sum_{i=0}^{3}\left(\frac{\partial\eta_j}{\partial R}\psi_i^\dagger-\frac{\partial\chi_j}{\partial R}\psi_i\right)=0, \quad (25)$$

$$3\alpha\psi_j^\dagger+a\chi_j+a\sum_{i=0}^{3}\left(\frac{\partial\eta_i}{\partial\psi_j}\psi_i^\dagger-\frac{\partial\chi_i}{\partial\psi_j}\psi_i\right)=0, \quad (26)$$

$$3\alpha\psi_j+a\eta_j-a\sum_{i=0}^{3}\left(\frac{\partial\eta_i}{\partial\psi_j^\dagger}\psi_i^\dagger-\frac{\partial\chi_i}{\partial\psi_j^\dagger}\psi_i\right)=0, \quad (27)$$

$$3\alpha\left(F-RF_R\right)-\beta aRF_{RR}+a\frac{h_u}{h}\left(F-RF_R\right)\sum_{i=0}^{3}\Big(\epsilon_i\eta_i\psi_i^\dagger+\epsilon_i\chi_i\psi_i\Big)=0, \quad (28)$$

$$3\alpha\left(K-YK_Y\right)+aK_u\sum_{i=0}^{3}\Big(\epsilon_i\eta_i\psi_i^\dagger+\epsilon_i\chi_i\psi_i\Big)=0. \quad (29)$$

After some mathematical calculations, we find the following particular solutions for the generators as

$$\alpha=\alpha_0a^{\frac{1}{3}}, \quad (30)$$

$$\beta = 2\alpha_0 a^{-\frac{2}{3}} R, \tag{31}$$

$$\eta_j = -\left(\frac{3}{2}\alpha_0 a^{-\frac{2}{3}} + \epsilon_j \eta_0\right)\psi_j, \tag{32}$$

$$\chi_j = -\left(\frac{3}{2}\alpha_0 a^{-\frac{2}{3}} - \epsilon_j \eta_0\right)\psi_j^{\dagger}, \tag{33}$$

where $\alpha_0$ and $\eta_0$ are some constant of integration. And find the functions $h$, $K$ and $F$ the following form

$$h = h_0 u^{\frac{5}{3}} \tag{34}$$

$$K = K_0 (Y - \nu u) \tag{35}$$

$$F = C_1 R, \tag{36}$$

where $h_0$, $K_0$, $(\nu)$ and $C_1$ are a integrable constant.

## 5. Cosmological solutions

In previous section we reconstructed the forms of $h$, $K$ and $F$ by using Noether symmetry. Now we're ready to solve equations of motion and find exact cosmological solutions. As a first step, we need to substitute the solutions (34), (35) and (36) in the field equations (10), (11). Then, we have

$$\dot{u} + 3\frac{\dot{a}}{a}u = 0, \tag{37}$$

so that

$$u = \frac{u_0}{a^3} \tag{38}$$

where $u_0$ is an integration constant.

If we put the equations (34)-(36) in equation (8), we find

$$\dot{a} = a_0 a. \tag{39}$$

Solution of differential equation

$$a(t) = C_2 e^{a_0 t}, \tag{40}$$

where $C_2$ is a constant of integration and $a_0$ is

$$a_0 = \frac{\nu}{2h_0 u^{\frac{2}{3}_0} C_1}. \tag{41}$$

The Hubble parameter is

$$H = \frac{\dot{a}}{a} = a_0 C_2, \tag{42}$$

Also, we find the energy density and pressure in these form

$$\rho = 3H^2 = 3(a_0 C_2)^2, \tag{43}$$

$$p = -3H^2 - 2\dot{H} = -3(a_0 C_2)^2. \tag{44}$$

For our model the equation of state parameter can be define as

$$\omega = \frac{p}{\rho} = -1. \tag{45}$$

We conclude that the fermionic field behaves as dark energy.

## 6. Conclusions

In this paper we investigated $F(R)$ gravity with $f$ -essence for Friedman-Robertson-Walker space-time metric. For this model, we have derived the point-like Lagrangian and the corresponding field equations. By using Noether symmetry approach we obtained solutions for $K$, $F$ and $h$ functions. By substitude these expressions we have some cosmological solutions. It is shown that this solutions describe the late time accelerated expansion of the Universe.

## Acknowledgments

The authors would like to thank the MG15 Meeting organisers for the kind invitation to present this talk. This work was supported in part by the Ministry of Edication and Science of Kazakhstan under grant 0118RK00935 as well as by grant 0118RK00693.

## References

1. H. A. Buchdahl, Non-Linear Lagrangians and Cosmological Theory, *Mon. Not. Roy. Astron. Soc*, **150**, 1 (1970).
2. T. P. Sotiriou, V. Faraoni, $f(R)$ theories of gravity, *Rev. Mod. Phys*, **82**, 451 (2010).
3. S. Capozziello, C. A. Mantica, L. G. Molinari, Cosmological perfect-fluids in $f(R)$ gravity, *Int. J. Geom. Meth. Mod. Phys.*, **16**, 1950008 (2019).
4. A. A. Starobinsky, A new type of isotropic cosmological models without singularity, *Phys. Lett.* B, **91**, 99 (1980).
5. R. C. de Souza and G. M. Kremer, Noether symmetry for non-minimally coupled fermion fields, *Class. Quant. Grav.*, **25**, 225006 (2008).
6. G. Gecim, Y. Kucukakca, Y. Sucu, Noether Gauge Symmetry of Dirac Field in (2+1)-Dimensional Gravity, *Adv. in High En. Phys.*, **2015**, 567395 (2015).
7. Y. Kucukakca, Teleparallel dark energy model with a fermionic field via Noether symmetry, *Eur. Phys. J.* C, **74**, 3086 (2014).
8. G. Gecim, Y. Sucu, Dirac Field as a Source of the Inflation in Dimensional Teleparallel Gravity, *Adv. in High En. Phys.*, **2017**, 2056131 (2017).

9. R. C. de Souza and G. M. Kremer, Cosmic expansion from boson and fermion fields, *Class. Quant. Grav.*, **28**, 125006 (2011).
10. M. Jamil, D. Momeni, N. S. Serikbayev, R. Myrzakulov, FRW and Bianchi type I cosmology of f-essence, *Astrophys Space Sci.*, **339**, 37 (2012).
11. K. Myrzakulov, P. Tsyba, R. Myrzakulov, Noether symmetry in $F(T)$ gravity with f-essence, arXiv:1601.07357.
12. S. Myrzakul, K. Myrzakulov, S. Bekov, T. Myrzakul, R. Myrzakulov, Starobinsky model with f-essence, arXiv:1710.08413.
13. R. de Ritis, G. Marmo, G. Platania, C. Rubano, P. Scudellaro, C. Stornaiolo, New approach to find exact solutions for cosmological models with a scalar field, *Phys. Rev.* D, **42**, 1091 (1990).
14. S. Capozziello, R. de Ritis, Relation between the potential and nonminimal coupling in inflationary cosmology, *Phys. Lett.* A, **177**, 1 (1993).

# Cosmological model of $f(T)$ gravity with fermion fields in (2+1) dimensions

N. A. Myrzakulov*, S. S. Bekov and K. R. Myrzakulov

*Department of General & Theoretical Physics, Eurasian National University, Nur-Sultan 010008, Kazakhstan*
** E-mail: nmyrzakulov@gmail.com*
*www.enu.kz*

In this paper, we explore $f(T)$ gravity which non-minimally coupled to fermionic fields in (2+1) spatially flat Friedmann-Robertson-Walker (FRW) dimensions. Friedmann equations, equations for fermionic fields are derived. Forms of gravitational coupling, self-interaction potential and $f(T)$ gravity are reconstructed by Noether symmetry approach for the point-like Lagrangian. Cosmological solution of model which corresponding to the late-time accelerating expansion are obtained.

*Keywords*: $f(T)$ gravity; fermion fields; Friedman-Robertson-Walker metric; Noether symmetry.

## 1. Introduction

Cosmological observations indicate that our Universe has two acceleration phases: early-time inflation and late-time cosmic acceleration so called dark energy.[1,2] To describe these phenomena has been proposed several alternative models of gravity as $f(R)$ gravity, $f(T)$ gravity, Gauss-Bonnet gravity, massive gravity, scalar-tensor gravity and so on. All these theories assume as theoretical candidates to explain all evolutons of Universe.

To modify usual curvature formulation of general relativity (GR), one starts by the replaces the Ricci scalar $R$ in the Einstein Hilbert action by arbitrary functions of it. However, alternatively one can use torsional general relativity, namely the so called teleparallel equivalent of general relativity (TEGR). In teleparallel gravity the gravitational field is described not by the curvature tensor but by the torsion, namely the torsion scalar $T$. Hence, similarly to the $f(R)$ extension of GR, one can construct the $f(T)$ extension of TEGR.[3,4]

In cosmological framework *symmetry* technique like Noether symmetry,[5] Lie symmetry,[6] Hojman symmetry[7] has revealed a useful tool in order to examine exact solutions. Also exact cosmological solutions of the Dirac (fermion) equation in curved space-time are useful tools in examine the physical behavior of particles in the particular space-time. These solutions have considerable importance in cosmology. In the 3D theory the exact solutions of the Dirac equations have been recently found for various potentials in the flat and curved space-time.

One of the first works on gravity in (2+1) dimensions appeared in the literature.[8] Recently, the three dimensional gravity have also gained considerable importance. The (2+1) dimensional gravity has similar properties as the (3+1) dimensional theories of gravity. In 3D gravity Riemann tensor is reduced to the Ricci tensor in the (2+1) dimensions. It is more simple case than the (3+1) dimensional gravity.

Therefore, the (2+1) dimensional gravity is assume as toy model for a regular Einstein theory in (3+1) dimensions. Additionally, the (2+1) Dirac equations is less complicated than the (3+1) Dirac equation. The (3+1) Dirac equation includes the Dirac matrices and the Dirac spinors with the four components. However, in (2+1) dimensions, the Dirac matrices are reduced to Pauli matrices.[9] In[10,11] has been studied Dirac equations in framework of (2+1) dimensional Einstein-Hilbert and teleparallel gravity by Noether gauge symmetry approach, respectively. The aim of this study is to find exact solutions of the Dirac equation in a (2+1) dimensional $f(T)$ contracting and expanding curved space-time via Noether symmetry.

The structure of this paper organizing as following. In Sect. 2 we find field equations for $f(T)$ gravity which non-minimally coupled to fermionic fields in (2+1) dimensions. In Sect. 3 we apply Noether symmetry approach to find the form of $f(T)$ gravity, coupling function and the self-interaction potential. The cosmological solution is considered in Sect. 4.

The signature of metric used is $(+,-,-)$ and units have been chosen so that $8\pi G = c = \hbar = k = 1$.

## 2. Action and field equations

We start by writing the action in the framework of the (2+1) dimensional $f(T)$ gravity non-minimally coupled with Dirac field,

$$\mathcal{S} = \int d^3x|e|\Bigg\{h(u)f(T) + \frac{\imath}{2}\Big[\bar{\psi}\sigma^{\mu}(x)(D_{\mu}\psi) - (D_{\mu}\bar{\psi})\sigma^{\mu}(x)\psi\Big] - V(u)\Bigg\}, \quad (1)$$

where the $h(u)$ and $V(u)$ are the coupling functions and the self-interaction potential, respectively. These scalar functions depend on only function of a bilinear Dirac field $u = \bar{\psi}\psi$. Bilinear field $u$ is a quantity, where $\bar{\psi}$ is the adjoint of the Dirac field $\psi$, which $\bar{\psi}{=}\psi^{\dagger}\sigma^3$.

The (2+1) dimensional FRW spacetime background that is spatially flat, homogeneous and isotropic universe,

$$ds^2 = dt^2 - a^2(t)[dx^2 + dy^2], \quad (2)$$

where $a(t)$ is the scale factor of the universe. Using the metric (2), the torsion scalar becomes

$$T = -2\frac{\dot{a}^2}{a^2}. \quad (3)$$

Based on metric (2) we rewrite action (1) as

$$S = \int d^3xe\Big[h(u)f(T) - \frac{\imath a^2}{2}\left(\bar{\psi}\sigma^3\dot{\psi} - \dot{\bar{\psi}}\sigma^3\psi\right) - V(u)\Big]. \quad (4)$$

By using Lagrange multiplier

$$S = \int d^3xe\Big[h(u)f(T) + \lambda\left(T + 2(\frac{\dot{a}^2}{a^2})\right) - \frac{\imath a^2}{2}\left(\bar{\psi}\sigma^3\dot{\psi} - \dot{\bar{\psi}}\sigma^3\psi\right) - V(u)\Big]. \quad (5)$$

The variation of this action with respect to $T$ gives us

$$\lambda = hf_T. \tag{6}$$

By substitute (6) to above action we have

$$S = \int d^3xe\Big[h(u)f(T) + hf_T\left(T + 2(\frac{\dot{a}^2}{a^2})\right) - \frac{\imath a^2}{2}\left(\bar{\psi}\sigma^3\dot{\psi} - \dot{\bar{\psi}}\sigma^3\psi\right) - V(u)\Big]. \tag{7}$$

Considering the background (2), it is possible to obtain the point-like Lagrangian

$$L = a^2hf_TT - a^2hf + 2\dot{a}^2hf_T - \frac{\imath a^2}{2}\left(\bar{\psi}\sigma^3\dot{\psi} - \dot{\bar{\psi}}\sigma^3\psi\right) + a^2V. \tag{8}$$

Equation of motion for fermion fields and its adjoint $\psi$ and $\bar{\psi}$, respectively as

$$\dot{\psi} + H\psi + \imath V'\sigma^3\psi + 2\imath H^2f_Th'\sigma^3\psi + \imath f_TTh'\sigma^3\psi - \imath fh'\sigma^3\psi = 0 \tag{9}$$

$$\dot{\bar{\psi}} + H\bar{\psi} - \imath V'\sigma^3\bar{\psi} - 2\imath H^2f_Th'\sigma^3\bar{\psi} - \imath f_TTh'\sigma^3\bar{\psi} + \imath fh'\sigma^3\bar{\psi} = 0. \tag{10}$$

On the other hand, from the point-like Lagrangian (8) and by considering the Dirac's equations, we find the acceleration equation for $a$ from the Euler-Lagrange equation as

$$\frac{\ddot{a}}{a} = -\frac{p_{_f}}{2hf_T} \tag{11}$$

and Hamiltonian constraint (energy condition) corresponding to Lagrangian $\mathcal{L}$ defined by

$$E_L = \frac{\partial L}{\partial\dot{a}}\dot{a} + \frac{\partial L}{\partial\dot{T}}\dot{T} + \frac{\partial L}{\partial\dot{\psi}}\dot{\psi} + \dot{\bar{\psi}}\frac{\partial L}{\partial\dot{\bar{\psi}}} - L. \tag{12}$$

By substituting (8) into (12) with the energy condition, we obtain

$$H^2 = \frac{\rho_{_f}}{2hf_T}. \tag{13}$$

In the acceleration and Friedmann equations, $\rho_{_f}$ and $p_{_f}$ are the effective energy density and pressure of the fermion (spinor) field, respectively, so that they have the following forms

$$\rho_{_f} = hf_TT - hf + V, \tag{14}$$

$$p_{_f} = 2h'\ddot{u}Hf_T - 8hf_{TT}H^2\dot{H} - (fh' - f_TTh' + 2f_Th'H^2 - V')u + hf - hf_TT - V. \tag{15}$$

In order to solve the field equations we have to determine a form for the coupling function and the potential density of the theory. To do this, in the following section we will use the Noether symmetry approach.

## 3. Noether symmetry

The existence of Noether symmetries lead to a specific form of coupling function and the self-interaction potential, which they are important for obtaining the exact solutions of the fields equations. Noether symmetry approach tells us that Lie derivative of the Lagrangian with respect to a given vector field $\mathbf{X}$ vanishes, i.e.

$$\pounds_{\mathbf{X}} L = 0. \tag{16}$$

Letting the spinor field $\psi = (\psi_1, \psi_2)^T$ and its adjoint, $\bar{\psi} = \psi^\dagger \sigma^3$, the point-like Lagrangian (14) is reduced to the following

$$L = a^2 h f_T T - a^2 h f + 2\dot{a}^2 h f_T - \frac{\imath a^2}{2}\left[\sum_{i=1}^{2}(\psi_i^\dagger \dot{\psi}_i - \dot{\psi}_i^\dagger \psi_i)\right] + a^2 V. \tag{17}$$

Hence the Noether condition (16) and requiring the coefficients of $\dot{a}^2, \dot{a}\dot{T}, \dot{a}\dot{\psi}_j, \dot{a}\dot{\psi}_j^\dagger, \dot{a}, \dot{T}, \dot{\psi}_j$ and $\dot{\psi}_j^\dagger$ to be zero, leads to the system of equations. Further, by solve this system of equations analytically, we can get solution for $\alpha$ as

$$\alpha(a) = \alpha_0 a^k \tag{18}$$

and solution for $f(T)$ as

$$f(T) = f_0 T^{\frac{3s-2}{2k+2}}. \tag{19}$$

Generators $\nu$ and $\delta$ defined as

$$\nu_j = -(\alpha_0 a^{k-1} + \epsilon_j \nu_0)\psi_j, \tag{20}$$

$$\delta_j = -(\alpha_0 a^{k-1} - \epsilon_j \nu_0)\psi_j^\dagger. \tag{21}$$

Coupling function and potential written as

$$h = h_0 u^s, \tag{22}$$

$$V = V_0 u, \tag{23}$$

where $h_0$ and $V_0$ are an integration constants.

For Noether symmetry the first integral or conserved quantity (Noether charge) defined as

$$Q = \xi^j \frac{\partial \mathcal{L}}{\partial \dot{q}^j} = 4\alpha h f_T \dot{a} = const. \tag{24}$$

Here $\xi^j$ is some symmetry generators, $q^j$ is variables.

In the next section we will search cosmological solutions of the field equations using the obtained coupling functions $h(u)$ and potential $V(u)$.

## 4. Cosmological solutions

To find the time evolution of the scale factor, $a(t)$, firstly, we need to know the form of bilinear function, $u$. Thus, we insert the coupling function $h$ given by (22) in the fermionic equations (9) and in its adjoint (10) we can get

$$\dot{u} + 2\frac{\dot{a}}{a}u = 0. \tag{25}$$

Then, the form of the bilinear function is calculated as

$$u = \frac{u_0}{a^2}, \tag{26}$$

where $u_0$ is a constant of integration.

From the above solution, the acceleration and Friedmann equations become only a function of the cosmic scale factor. Using the potential (23) in the Friedmann equation together with the equation (14), the time evolution of the scale factor can be easily calculated as

$$\dot{a} = a_0 a^{\frac{2sk-s-2k}{3s-2}}, \quad \text{where} \quad a_0 = \frac{v^{\frac{k+1}{3s-2}} u_0^{\frac{(1-s)(k+1)}{3s-2}}}{\sqrt{2}}, \quad v = \frac{V_0}{h_0 F_0}\left(\frac{k+1}{3s-k-3}\right) \tag{27}$$

The general solution of the equation is

$$a(t) = (tv - ntv + c_1)^{-\frac{1}{n-1}}, \quad \text{where} \quad n = \frac{2sk-s-2k}{3s-2} \tag{28}$$

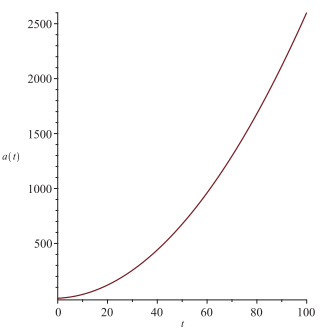



Fig. 1. Scale factor $a$ depends on cosmic time $t$ at $v = 1$, $c_1 = 1$, $n = \frac{1}{2}$.

In Fig. 1. shows the dependence of scale factor $a$ on cosmic time $t$. For our model in order to describe the accelerated expansion of the Universe, it is necessary that $0 < n < 1$.

## 5. Conclusion

In this paper, we have investigated a model with a fermionic field that is non-minimally coupled to gravity in the framework of $f(T)$ gravity in (2+1) dimensions. By using the Noether symmetry approach we determined form of the potential and coupled function as function of bilinear $u$. We also find scale factor for our model. Our result shown that fermion field may considered as source of dark energy in framework $f(T)$ gravity.

## Acknowledgments

The authors would like to thank the MG15 Meeting organisers for the kind invitation to present this talk. This work was supported in part by the Ministry of Edication and Science of Kazakhstan under grant 0118RK00935 as well as by grant 0118RK00693.

## References

1. S. Perlmutter et al., Measurements of Ω and Λ from 42 High-Redshift Supernovae, *ApJ*, **517**, 565 (1999).
2. A. Riess et al., Observational Evidence from Supernovae for an Accelerating Universe and a Cosmological Constant, *ApJ*, **116** 3, 1009 (1998).
3. Y-F. Cai et al., $f(T)$ teleparallel gravity and cosmology, *Rep. Prog. Phys*, **79** 106901 (2016).
4. E. Güdekli, N. Myrzakulov, K. Yerzhanov, R. Myrzakulov, Trace-anomaly driven inflation in $f(T)$ gravity with a cosmological constant, *Astrophys Space Sci*, **357**, 45 (2015).
5. S. Capozziello, N. Frusciante, D. Vernieri, New spherically symmetric solutions in $f(R)$-gravity by Noether symmetries, *Gen Relativ Gravit*, **44**, 1881 (2012).
6. A. Paliathanasis, L. Karpathopoulos, A. Wojnar, S. Capozziello. Wheeler–DeWitt equation and Lie symmetries in Bianchi scalar-field cosmology, *Eur. Phys. J.* C, **76**, 225 (2016).
7. S. Capozziello, M. Roshan. Exact cosmological solutions from Hojman conservation quantities, *Phys. Lett.* B, **726**, 471-480 (2013).
8. A. Staruszkiewicz, Gravitation Theory in Three-Dimensional Space, *Acta Phys. Polon*, **24**, 734 (1963).
9. Y. Sucu, N. Unal, Exact solution of Dirac equation in 2+1 dimensional gravity, *Jour. of Math. Phys.*, **726**, 471-480 (2007).
10. G. Gecim, Y. Kucukakca, Y. Sucu, Noether Gauge Symmetry of Dirac Field in (2+1)-Dimensional Gravity, *Adv. in High En. Phys.*, **2015**, 567395 (2015).
11. G. Gecim, Y. Sucu, Dirac Field as a Source of the Inflation in 2+1 Dimensional Teleparallel Gravity, *Adv. in High En. Phys.*, **2017**, 2056131 (2017).

# Forecasting interacting dark energy models with future experiments

L. Santos[1,2] and W. Zhao

[1] *Center for Gravitation and Cosmology, College of Physical Science and Technology, Yangzhou University, Yangzhou 225009, China.* [2] *CAS Key Laboratory for Researches in Galaxies and Cosmology/School of Astronomy and Space Science, Department of Astronomy, University of Science and Technology of China, Chinese Academy of Sciences, Hefei, Anhui 230026, China*

[*] *E-mail: larissa@ustc.edu.cn*

We perform a forecast analysis on the ability of future baryonic acoustic oscillation (BAO) and cosmic microwave background (CMB) experiments in constraining 3 specific interacting dark energy models using the well known Fisher-matrix formalism. In addition to a future ground-based CMB experiment, we consider a Euclid-like experiment, which is supposed to put tight constraints on the dark sector parameters. In the interacting dark energy scenario, a coupling between dark matter and dark energy modifies the conservation equations such that the fluid equations for both constituents are conserved as the total energy density of the dark sector. In this context, we consider three phenomenological models which have been deeply investigated in literature in the past years. We find that the combination of both CMB and BAO can break degeneracies among the parameters for every studied model. We found powerful constraints on, for example, the coupling constant when comparing it with present limits for two of the models, and their future statistical 3-$\sigma$ bounds could potentially exclude the null interaction for the combination of probes that is considered.

*Keywords*: Dark energy, dark matter, BAO, CMB.

## 1. Introduction

Considering the standard cosmological model, the dark energy assumes its simplest form as the cosmological constant $\Lambda$, leading to the so-called $\Lambda$CDM model. Despite of successfully explaining the observations [1], the the standard model faces some difficulties, especially in the dark sector. The dark matter particles are far from detection, and its physics is still unknown. The cosmological constant theoretical predictions estimate the value of a vacuum energy density orders of magnitude larger than its actual observed value. In addition, the present values of the dark matter and dark energy densities are at the same other of magnitude even though they do not share the same cosmological evolutionary behavior, problem known as cosmic coincidence [2]. To overcome some of these problems, models where dark matter and dark energy interact started to be considered, becoming very useful to alleviate this coincidence problem. An interacting dark matter and dark energy scenario would affect the overall evolution of the universe and its expansion history, thus observationally distinguishable from the $\Lambda$CDM model. The interaction can then be constrained by the data, becoming a testable theory for the universe.The present observations, however, are not able to confidently distinguish between these alternative interacting dark energy models and the $\Lambda$CDM. Our goal is to test the ability of future experiments to constraint the interacting dark energy models described in the next section.

## 2. The interacting dark energy models

In the standard cosmological model the energy density of radiation, baryons, cold dark matter and dark energy is conserved separately, for each component. Conversely, in an interacting dark energy model, the fluid equations for the dark energy and dark matter are not conserved individually, but together as the total energy density of the dark sector such that $\dot{\rho}_{DM} + 3H\rho_{DM} = +Q$ and $\dot{\rho}_{DE} + 3H(1 + \omega_{DE})\rho_{DE} = -Q$, where $H$ is the Hubble parameter, $\rho_{DM}$ and $\rho_{DE}$ are the energy densities for dark matter and dark energy, respectively, and $\omega_{DE} = P_{DE}/\rho_{DE}$ is the dark energy EoS. $Q$ represents the interaction kernel that can be written phenomenologically as $Q = 3H(\xi_1\rho_{DM} + \xi_2\rho_{DE})$, being the coupling coefficients ($\xi_1$ and $\xi_2$) constants to be determined by observations [4, 5]. The energy flow from dark energy to dark matter is defined by $Q > 0$, and the opposite for $Q < 0$. Considering the stability of the model (see, for instance, [6]), two choices are made: The first, $\xi_1 = 0$ and $\xi_2 \neq 0$, satisfying a constant dark energy EoS within the range $-1 < \omega_{DE} < 1$ (described as model 1), or $\omega_{DE} < -1$ (model 2). The second, $\xi_2 = 0$ and $\xi_1 \neq 0$ for $\omega_{DE} < -1$ (model 3). For all three models, the other components follow the standard conservation equations. For a review on the topic refer to [7].

## 3. The fisher formalism

The BAO is an important observable currently used to constrain the cosmological parameters, more efficiently in combination with other probes, such as the CMB. The information stored in the BAO can precisely determine the Hubble parameter $H(z)$ and the angular diameter distance $Da(z)$ as a function of the redshift, which subsequently enables the calculation of the dark energy parameters constraints (for details on this methodology see [8]). The energy densities for dark matter and dark energy for the models considered here can be found in [9]. For the matter power spectrum obtained from galaxy surveys the Fisher matrix is given by [11][a]

$$F_{ij} = \int_{-1}^{1}\int_{k_{\min}}^{k_{\max}} \frac{\partial \ln P(k,\mu)}{\partial p_i}\frac{\partial \ln P(k,\mu)}{\partial p_j} V_{\text{eff}}(k,\mu)\frac{2\pi k^2 dk d\mu}{2(2\pi)^3}, \tag{1}$$

being $V_{\text{eff}}$ the effective volume of the survey. We present the expected cosmological implications of the BAO measurements for an Euclid-like survey (for specifications on Euclid, see, for example, [12]).

Furthermore, we use the CMB information as a second probe to forecast the parameters of the interacting dark energy models described here. We do not consider

[a]The matter power spectrum was generated using a modified version of CAMB software package [10].

the primordial $B-$mode contribution. We then construct the Fisher matrix for the CMB temperature anisotropy and polarization [13].

$$F_{ij} = \sum_l \sum_{XY} \frac{\partial C_l^X}{\partial p_i} (Cov_l^{-1})_{XY} \frac{\partial C_l^Y}{\partial p_j}, \tag{2}$$

being $C_l^X$ the power in the $l$-th multipole, $X$ stands for $TT$ (temperature), $EE$ (E-mode polarization), $TE$ (temperature and E-mode polarization cross-correlation) and $(Cov_l^{-1})_{XY}$ the covariance matrix. For the definition and the elements of the covariance matrix refer to [13]. We consider the Advanced Atacama Cosmology telescope (ACTadv) instrumental setup as described in [14]. The ACTadv is supposed to obtain precise measurements of the CMB small-scale polarization, which can lead on probing alternative cosmological models.

## 4. Results and conclusions

We combine the Fisher matrices for the BAO and CMB future measurements from Euclid and ACTadv surveys, respectively. The marginalized error for the dark energy EoS in model 1 improves drastically for the combined analysis, being $\sigma(\omega) = 0.026$ for Euclid, $\sigma(\omega) = 0.028$ for the ACTadv and $\sigma(\omega) = 0.0044$ for their combination: an improvement by a factor of $\sim 6$ when compared with each individual probe. The constraint on the dark matter density improves by a factor of $\sim 3$ for the combined analysis ($\sigma(h^2\Omega_c) = 0.00053$), compared with Euclid alone ($\sigma(h^2\Omega_c) = 0.0017$). A similar improvement occurs for the coupling constant, where we find $\sigma(\xi_2) = 0.0037$ for Euclid alone and $\sigma(\xi_2) = 0.0019$ for ACTadv + Euclid. Such stringent constraint would exclude the null interaction correspondent to the $\Lambda$CDM model with high confidence. Present constraints on $h^2\Omega_c$, $\omega$ and $\xi_2$ for a combination of probes (Planck+BAO+SNIa+H0) show $h^2\Omega_c = 0.0792^{+0.0348}_{-0.0166}$, $\omega = -0.9191^{+0.0222}_{-0.0839}$ and $\xi_2 = -0.1107^{+0.085}_{-0.0506}$ [3]. The constraints on model 1 from the present datasets are affected by the degeneracies among the parameters, more evident between $h^2\Omega_c$ and $\xi_2$. It is clear by our analysis that Euclid information can help break the degeneracies between these parameters, therefore providing tight constraints on $h^2\Omega_c$, $\omega$ and $\xi_2$. As an example, we show the Fisher contours on the dark sector parameters for model 1 in Figure 1.

The same occurs for model 2. The combined result leads to stringent constraints on $\sigma(h^2\Omega_c)$, $\sigma(\omega)$ and $\sigma(\xi_2)$, the latter being $\xi_2 = 0.03798 \pm 0.00310$ at $1\sigma$. A zero positive interaction is excluded by [3] with $\xi_2 = 0.02047^{+0.00565}_{-0.00667}$ at $1\sigma$. The future combination of ACTadv and Euclide-like surveys would be able to improve this constraint by a factor of $\sim 2$.

As for model 3, degeneracies between the dark sector parameters do not play such an important role compared with models 1 and 2. However, the significance on the $\xi_1$ constraint is low at the current observational stage, being $\xi_1 = 0.0007127^{+0.000256}_{-0.000633}$ at $1\sigma$, considering a combination of probes: Planck, SNIa, BAO, $H_0$ [3]. We found $\xi_2 = 0.0007273 \pm 0.00034$ (ACTadv + Euclid), which does not improve the actual

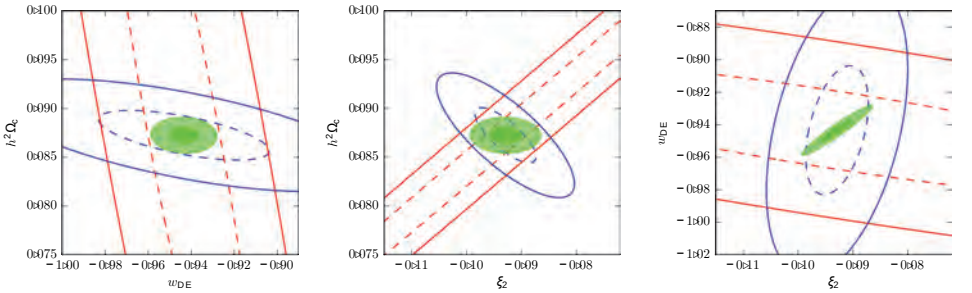



Fig. 1. Fisher forecast contours for model 1 with CMB and BAO information using AdvACT (red curves) and Euclid (blue curves) experimental setups, respectively. The dashed curves represent 68% C. L. and the solid curves represent 99.9% C. L. The combined contours are shown by the green filled ellipses. Similarly, the darker ellipses represent 68% C. L. and the fainter ones represent 99.9% C. L.

best constraint. For this model, a combination of other probes is still needed in order to tighten the present limits.

It is well known the advantages of combining different observational probes in constraining cosmological parameters, and its implication to interacting dark energy models has been widely addressed. In our context, for models 1 and 2, stringent constraints were found in the dark sector parameters for the combined probes, especially for the coupling constant, being the $1-\sigma$ bound of $\xi_2 = -0.0929 \pm 0.0019$ and $\xi_2 = 0.03798 \pm 0.00310$, respectively. Future CMB and BAO experiments combined, such as presented here, would be able to exclude the null interaction with more than $3\sigma$ C.L. (for the full reference on these results refer to [16]).The present dataset and the future CMB information alone are affected by degeneracies that can be broken by the addition of Euclid BAO measurements, thus tighten the constraints on the dark sector cosmological parameters, and enabling a deeper discussion on these interacting dark energy scenarios. Conversely, the constraint on the coupling constant for model 3 is not improved by the combination of future CMB and BAO information compared with its constraint derived by present dataset. Extra information is still necessary for probing this model.

## Acknowledgment

L. Santos and W.Zhao are supported by NSFC No. 11603020, 11633001, 11173021, 11322324, 11653002, project of Knowledge Innovation Program of Chinese Academy of Science and the Fundamental Research Funds for the Central Universities.

## References

[1] Planck Collaboration and P. A. R. Ade, N. Aghanim, M. Arnaud *et al.*, A&A, **594**, A13 (2016)
[2] I. Zlatev, L. Wang, P. Steinhardt, Phys. Rev. Lett, **82**, 896 (1999).
[3] A. A. Costa, X. D. Xu, B. Wang, E. Abdalla , JCAP, **1**, 028 (2017).

[4] C. Feng, B. Wang, E. Abdalla and R. K. Su, Phys. Lett. B, **665**, 111 (2008).
[5] J. H. He, B. Wang, E. Abdalla, PRD, **83**, 063515 (2011).
[6] J. H. He, B. Wang, E. Abdalla, Phys. Lett. B, **671**, 139 (2009).
[7] B. Wang, E. Abdalla, F. Atrio-Barandela and D. Pavón, Reports on Progress in Physics, **79**, 096901 (2016).
[8] H.-J. Seo and D. J. Eisenstein, APJ, **598**, 720-740 (2003).
[9] J. H. H. and B. Wang J. Cosmol. Astropart. Phys., **0806**,010 (2008).
[10] A. Lewis, A. Challinor and A. Lasenby, APJ, **538**, 473-476 (2000).
[11] M. Tegmark, PRL, **79**, 3806-3809 (1997).
[12] L. Amendola *et al.*, Living Reviews in Relativity, **16**, 6 (2013).
[13] M. Zaldarriaga and U. Seljak, PRD, **55**, 1830-1840 (1997).
[14] E. Calabrese, R. Hložek, N. Battaglia *et al.*, J. Cosmol. Astropart. Phys., **8**,010 (2014).
[15] D. Eisenstein, W. Hu, and M. Tegmark, APJ, **518**, 2-23 (1999).
[16] L. Santos, W. Zhao, E. Ferreira and E. Quintin, PRD, **96**, 103529 (2017).

# Cosmographic analysis as framework to evaluate cosmological models

Hugo Solís-Sánchez*[12], Manuel Ortega-Rodríguez[1], Luis A. Álvarez[1], Esteban Dodero[1], E. Gabriela Barrantes[2], and José M. Gamboa[1]

[1] *Escuela de Física and* [2] *Centro de Investigaciones en Tecnologías de la Información y Comunicación, Universidad de Costa Rica, San José, 11501-2060, Costa Rica*

* *E-mail: hugo.solis@ucr.ac.cr*

*www.ucr.ac.cr*

By using a cosmographic analysis of the redshift data of type Ia supernovae, we are able to get the expansion of the scale factor, obtaining the current values of the Hubble, deceleration, jerk and snap parameters. Our data is then used to compare the fitness of various proposed alternative cosmological models. Since our method assumes only the validity of general relativity at the cosmic scale, along with the isotropy and homogeneity of the universe, they are very useful for comparison between different cosmological models, including the fitness of ΛCDM model. Our method is based on the order expansion of the scale factor present in the FRW metric and using a Monte Carlo integration to find the best fit order parameters of the scale factor to reproduce the observed data, we make use of parallel paradigm to improve the computational time behind the model. We find the known result of an accelerated expansion of the universe. With access to better measurements of type Ia supernovae redshifts and more data, the cosmographic results will be significantly improved.

*Keywords*: Cosmography; Type Ia Supernovae; ΛCDM; deceleration parameter; Monte Carlo Integration.

## 1. Introduction

Although there is almost complete agreement on the accelerated expansion of the universe[1], ΛCDM being the most favored cosmological model by the current data, there is no consensus to the current specific value of this decelerated expansion and to whether or not this model is still prevalent in the future[2]. The decelerated expansion $q_0$ is the parameter indicating the current rate of acceleration of the expansion of the universe. Therefore, it is the first indicator of the fitness of any cosmological model, so a precise determination of its value is necessary for the diagnostic of any cosmological model. There recently have been plenty of research but there is no consensus for its current value[3].

We obtain the expansion terms of the scale factor in the FRW line element to test different cosmological models. We compare the expected value of the apparent magnitude in terms of the measured redshift to each supernova with the measured value of its apparent magnitude. This comparison is made by means of a likelihood ratio test to find the current values of the coefficients in the expansion of the factor scale that best fit the measured data.

Our interest is in comparing the results from our analysis with proposed cosmological models (as in ref. 4 and ref. 5). Even though cosmographic analysis has been previously studied (as in ref. 6 and ref. 7), we are performing the analysis using parallel computing what permits to make the expansion to higher orders than the

usuals. We found that some results have significant changes when calculations are improved and how the cosmography works as a framework to assess cosmological models.

## 2. Friedmann-Robertson-Walker and cosmography

### 2.1. *Friedmann-Robertson-Walker*

The line element for all the homogeneous and isotropic models of the universe is the Friedmann-Robertson-Walker metric, shown in 1. Where $a(t)$ is the expansion factor of the universe, which gives us the rate at which the universe is expanding. This value depends on the content and matter-energy densities for the universe and it is theoretically found using the Friedmann equation, which is obtained from Einstein's field equations. The constant $k$ is only determined by its sign, if $k < 0$ the universe is said to be open and the spatial hypersurfaces have negative constant curvature, if $k > 0$ the universe is said to be closed and spatial hypersurfaces have positive constant curvature, for $k = 0$ the universe is said to be flat with the spatial hypersurfaces being Euclidean with curved spacetimes.

$$ds^2 = c^2dt^2 - a(t)^2\left[\frac{dr^2}{(1-kr^2)} + r^2(d\theta^2 + \sin^2\theta d\phi^2)\right] \tag{1}$$

### 2.2. *Cosmographic analysis*

We proceed by using a cosmographic approach to determine the value of cosmological parameters. We seek to make a comparison of the measured values of the apparent magnitude ($m$) of the supernovae with the expected values given its measured redshift ($z$). The apparent magnitude is given by (2) in terms of the luminosity distance ($d_L$) and the absolute magnitude ($M$) which is known for supernovae to constant. Since the luminosity distance (3) is given in terms of the physical distance ($r_0$) between the source signal and the observer and the measured $z$ we need to express $r_0$ in terms of the measured redshift. We do this by using the null geodesic in FRW and the cosmographic redshift. By the null geodesic we have:

$$m = 5\log\frac{d_L}{10} + M \tag{2}$$

$$d_L = (1+z)r_0a_0 \tag{3}$$

$$-c\int_{t_*}^{t_0}\frac{dt}{R(t)} = f(r_0) = \int_{r_0}^{0}\frac{dr}{\sqrt{(1-kr^2)}} \tag{4}$$

$$\text{Where: } f(r_0) = \begin{cases} -\sin^{-1}(r_0) & (k=+1) \\ -r_0 & (k=0) \\ -\sinh^{-1}(r_0) & (k=-1) \end{cases}$$

We expand the scale factor in FRW:

$$R(t) = R(t_0)\left[1 + H_0(t-t_0) - \frac{1}{2!}q_0 H_0^2 (t-t_0)^2 + \frac{1}{3!}j_0 H_0^3 (t-t_0)^3 + \frac{1}{4!}s_0 H_0^4 (t-t_0)^4 + \cdots\right]$$

and using the cosmological redshift relation with the scale factor expansion[8], we obtain the flight time from the source to us ($T \equiv t_0 - t_*$ where $t_*$ is the time at which the signal was emitted) as a function of the measured redshift.

$$z + 1 = \frac{R(t_0)}{R(t_*)}$$

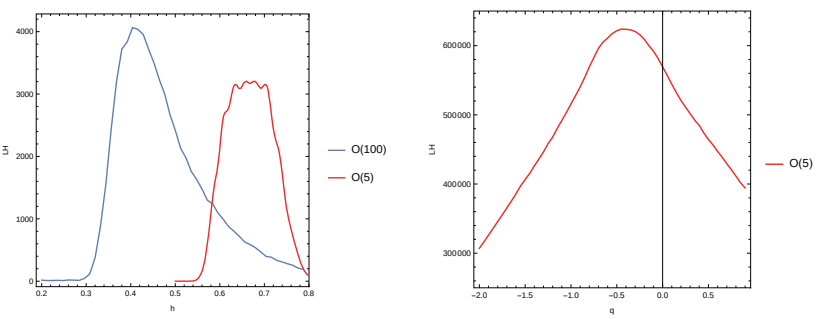
$$\frac{R(t_0)}{R(t_*)} = 1 + H_0 T + \frac{2+q_0}{2} H_0^2 T^2 + \frac{6(1+q_0)+j_0}{6} H_0^3 T^3 + \frac{24 - s_0 + 8j_0 + 36q_0 + 6q_0^2}{24} H_0^4 T^4 + \cdots$$

Numerically inverting:

$$T\left(\frac{z}{H_0}\right)^{-1} = 1 - \left[1 + \frac{q_0}{2}\right] z + \left[1 + q_0 + \frac{q_0^2}{2} - \frac{j_0}{6}\right] z^2$$

$$- \left[1 + \frac{3}{2}q_0(1+q_0) + \frac{5}{8}q_0^3 - \frac{1}{2}j_0 - \frac{5}{12}q_0 j_0 - \frac{s_0}{24}\right] z^3 + \ldots$$

We solve the left side integral in eq. 4 with the expansion of the scale factor and substitute $T$ in terms of $z$ as found above. Therefore we are able to use the luminosity distance in terms of the redshift, to which we then employ a marginal likelihood ratio analysis in order to find the best fit values for the cosmological parameters. We are then able to use this data (from ref. 6) to compare with any cosmological model.




(a) (b)

Fig. 1. Here are the results of our cosmographic analysis. (a) The Hubble parameter (b) The deceleration parameter.

In figure 1 (a), we are showing the results for the Hubble parameter for two orders of expansion. It is visible how when the order is higher the average is lower than the usual value, for 5 order we found $\langle H\rangle = 100\,\langle h_0\rangle = 67$ and for 100 order is $\langle H\rangle = 47$. The figure 1 (b) shows the deceleration parameter which in agreement with recent results is negative[3] $\langle q_0\rangle = -0.48$.

## 3. Comparison of different models

We compare our obtained factor scale with two different theoretically proposed cosmological models. The first one[4] proposes the gravitational constant ($G$) and the cosmological constant ($\Lambda$) are not constants but instead functions of time. With the standard Friedmann equations derivation and proposing they relate to each other by $G\rho = \frac{\eta\Lambda}{8\pi}$ with $\rho$ the density of the perfect fluid and $\eta$ a constant, there are two possible scale factors and deceleration parameters:

- Case A: $n \neq 0$

$$a(t) = (nlt + C_1)^{\frac{1}{n}} \tag{5}$$

$$q = n - 1 \tag{6}$$

- Case B: $n = 0$

$$a(t) = C_2 e^{lt} \tag{7}$$

$$q = -1 \tag{8}$$

The second one[5] proposes measuring the average expansion rate in a universe in which a set of spherically symmetric sub-regions expand in an accelerated way, *Average Expansion Rate Approximation (AvERA)*. It has the appeal that it conclusively resolves the tension between the measured values of the *Hubble constant* but the great drawback is that it is difficult to match with the homogeneity observed in the CMB.

We plot them in figure 2 with our obtained results to see how they compare to each other. We see that although one has several parameters to adjust to get greater similarity with our results can be more closely approximated by Case A of the given model, Case B is an exponential, so it starts at the value one and never approximates the rest of the curves.

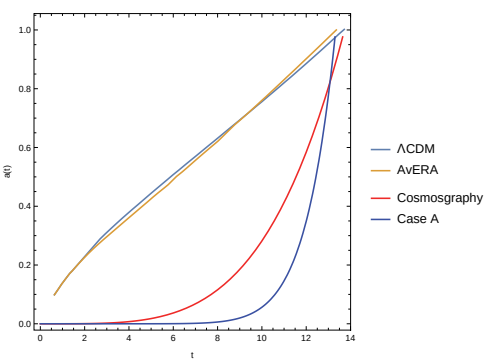



Fig. 2. Comparison of different models with the cosmographic result.

**Acknowledgments**

This work is partially supported by MICITT and CONICIT grant FI-0204-2012 and by grant 112-B8-107 of the Universidad de Costa Rica's Vicerrectoría de Investigación.

**References**

1. A. G. Riess, S. A. Rodney, D. M. Scolnic et al. *ApJ* **853**, 126 (2018).
2. R. G. Cai and Z. L. Tuo. *Phys. Lett. B* **706**, 116 (2011).
3. I. Tutusaus, B. Lamine, A. Dupays and A. Blanchard. *A&A* **602**, A73 (2017).
4. S. K. Tripathi and R. K. Dubey. *Indian J. Sci. Res.* **2**, 95 (2011).
5. G. Rácz, L. Dobos, R. Beck, I. Szapudi and I. Csabai. *MNRAS: Letters* **469**, L1-L5 (2017).
6. M. V. John. *ApJ* **614**, 1 (2004).
7. S. Capozziello, V. F. Cardone, and V. Salzano. *Phys. Rev. D* **78**, 063504 (2008).
8. M. Visser. *Class. Quantum Grav* **21**, 2603 (2004).

# Accelerated expansion of the Universe and the Higgs true vacuum

Muhammad Usman* and Asghar Qadir†

*Department of Physics,*
*School of Natural Sciences (SNS),*
*National University of Sciences and Technology (NUST),*
*Sector H-12 Islamabad 44000 Pakistan.*
**E-mail: muhammad_usman_sharif@yahoo.com*
†*E-mail: asgharqadir46@gmail.com*

Scalar fields which are favorite among the possible candidates for the dark energy usually have degenerate minima at $\pm\phi_{min}$. In the presented work, we discuss a two Higgs doublet model with the non-degenerate vacuum named inert uplifted double well type two-Higgs doublet model (UDW-2HDM) for the dark energy. It is shown that when the both Higgs doublets lie in their respective true minima then one Higgs doublet can cause the current accelerated expansion of the Universe.

*Keywords*: Dark energy; Uplifted double well two Higgs doublet model.

## 1. Introduction

Modeling the accelerated expansion of the Universe which is also in agreement with at least the standard model of the Particle Physics is a very hot question. We in this context earlier showed that the Higgs field(s) of the inert doublet model (IDM) and the general two Higgs doublet model (2HDM) can provide the current accelerated expansion of the Universe. In this short work we show that a new kind of two Higgs doublet model where both the Higgs vacuums are non-degenerate can also be used for dark energy modeling. We call this model to be uplifted double well two Higgs doublet model (UDW-2HDM). We in this short review give the very brief review and the results with necessary model information. The complete full length paper will be submitted to an ISI indexed journal later.

## 2. Uplifted double well two-Higgs double model

The Lagrangian which describes any model in Particle Physics is

$$\mathcal{L} = \mathcal{L}_{gf}^{SM} + \mathcal{L}_Y + T_H - V_H \ , \tag{1}$$

where $\mathcal{L}_{gf}^{SM}$ is the $SU_C(3)\otimes SU_L(2)\otimes U_Y(1)$ SM interaction of the fermions and gauge bosons (force carriers), $\mathcal{L}_Y$ is the Yukawa interaction of fermions with the Higgs field(s), $T_H$ is the kinetic term of the Higgs field and $V_H$ is the potential of the Higgs field. The last two terms form the Higgs Lagrangian $\mathcal{L}_{Higgs}$. In UDW-2HDM

$$\begin{aligned} T_H = (D_{1\mu}\phi_1)^\dagger (D_1{}^\mu \phi_1) + (D_{2\mu}\phi_2)^\dagger (D_2{}^\mu \phi_2) \\ + \left[\chi (D_{1\mu}\phi_1)^\dagger (D_2{}^\mu \phi_2) + \chi^* (D_{2\mu}\phi_2)^\dagger (D_1{}^\mu \phi_1)\right] , \end{aligned} \tag{2}$$

and

$$
\begin{aligned}
V_H = {} & \rho_1 \exp(\Lambda_1 m_{11}^2 (\phi_1 - \phi_{1_0})^\dagger (\phi_1 - \phi_{1_0})) \\
& + \rho_3 \exp\left(\frac{1}{2}\Lambda_3 \lambda_1 ((\phi_1 - \phi_{1_0})^\dagger (\phi_1 - \phi_{1_0}))^2\right) \\
& + \rho_2 \exp(\Lambda_2 m_{22}^2 (\phi_2 - \phi_{2_0})^\dagger (\phi_2 - \phi_{2_0})) \\
& + \rho_4 \exp\left(\frac{1}{2}\Lambda_4 \lambda_2 ((\phi_2 - \phi_{2_0})^\dagger (\phi_2 - \phi_{2_0}))^2\right) \\
& + \lambda_3 (\phi_1^\dagger \phi_1)(\phi_2^\dagger \phi_2) + \lambda_4 (\phi_1^\dagger \phi_2)(\phi_2^\dagger \phi_1) + \left[m_{12}^2 (\phi_1^\dagger \phi_2) + \frac{\lambda_5}{2}(\phi_1^\dagger \phi_2)^2\right. \\
& \left. + \lambda_6 (\phi_1^\dagger \phi_1)(\phi_1^\dagger \phi_2) + \lambda_7 (\phi_2^\dagger \phi_2)(\phi_1^\dagger \phi_2) + \text{h.c.}\right], \qquad (3)
\end{aligned}
$$

where

$$
D_{p_\mu} = \partial_\mu + i\frac{g_p}{2}\sigma_i W^i{}_\mu + i\frac{g_p'}{2} B_\mu, \qquad p = 1 \text{ or } 2 \,,
$$

$$
\phi_i = \begin{bmatrix} \phi_i^+ \\ \eta_i + i\chi_i + \nu_i \end{bmatrix} \,, \qquad \phi_{i_0} = \begin{bmatrix} 0 \\ \tau_i \end{bmatrix} \qquad \text{and} \qquad \phi_i^\dagger = \begin{bmatrix} \phi_i^- & \eta_i - i\chi_i + \nu_i \end{bmatrix} .
$$

The Higgs fields $\phi_i^+$, $\phi_i^-$, $\eta_i$ and $\chi_i$ are hermitian ($\phi_i^\pm$ are charged whereas the others are neutral), $\nu_i$ is the VeV of the doublet $\phi_i$, $\phi_{i_0}$ in the potential is the true minimum of the field $\phi_i$. The shape of the UDW-2HDM Higgs potential is shown in fig. 1.

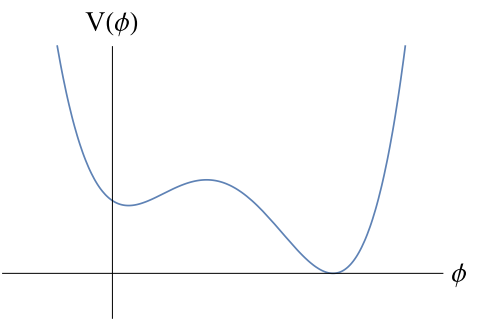



Fig. 1. The uplifted double well potential

The $Z_2$ symmetry ($\phi \to -\phi$) is very important in determining the stability of the Higgs fields. There are two types of $Z_2$ symmetry breaking: 1) soft; and 2) hard. The term containing $m_{12}^2$ describes the soft $Z_2$ symmetry breaking, whereas the terms containing $\lambda_6$ and $\lambda_7$ describe the hard $Z_2$ symmetry breaking. In the absence of these terms along with no cross kinetic term i.e. $\chi = 0$, the UDW-2HDM's Higgs Lagrangian has a perfect $Z_2$ symmetry.

The extrema of the potential are found by taking

$$\frac{\partial V_H}{\partial \phi_1}\Big|_{\phi_1=\langle\phi_1\rangle} = \frac{\partial V_H}{\partial \phi_1^\dagger}\Big|_{\phi_1^\dagger=\langle\phi_1^\dagger\rangle} = 0 \quad \text{and} \quad \frac{\partial V_H}{\partial \phi_2}\Big|_{\phi_2=\langle\phi_2\rangle} = \frac{\partial V_H}{\partial \phi_2^\dagger}\Big|_{\phi_2^\dagger=\langle\phi_2^\dagger\rangle} = 0. \quad (4)$$

The most general solution of the conditions (4) is

$$\langle\phi_1\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \nu_1\end{pmatrix} \qquad \text{and} \qquad \langle\phi_2\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix}u\\ \nu_2\end{pmatrix}.$$

We must remember that now $\nu^2 = \nu_1^2 + \left|\nu_2^2\right| + u^2$, where $\nu = 1/\sqrt[4]{2{G_F}^2} \approx 246$ GeV is the VeV of the Higgs field in the SM. For a nonzero $u$, the "charged" type dark energy is obtained, which obviously is not the case. For this reason, choose $u = 0$. Solving eq. (4) for the potential truncated to the forth power of field gives

$$\begin{aligned}&\rho_1\Lambda_1 m_{11}^2(\nu_1-\tau_1) + \rho_1\Lambda_1^2 m_{11}^4(\nu_1-\tau_1)^3 + \rho_3\Lambda_3\lambda_1\nu_1(\nu_1^2-\tau_1^2) + {m_{12}^2}^*\nu_2\\ &+(\lambda_3+\lambda_4+\lambda_5^*)\nu_1\nu_2^2 + (\lambda_6+2\lambda_6^*)\nu_1^2\nu_2 + \lambda_7^*\nu_2^3 = 0,\end{aligned} \quad (5)$$

$$\begin{aligned}&\rho_2\Lambda_2 m_{22}^2(\nu_2-\tau_2) + \rho_2\Lambda_2^2 m_{22}^4(\nu_2-\tau_2)^3 + \rho_4\Lambda_4\lambda_2\nu_2(\nu_2^2-\tau_2^2) + m_{12}^2\nu_1\\ &+(\lambda_3+\lambda_4+\lambda_5)\nu_1^2\nu_2 + \lambda_6\nu_1^3 + (2\lambda_7+\lambda_7^*)\nu_1\nu_2^2 = 0.\end{aligned} \quad (6)$$

Imposing the $Z_2$ symmetry ($\chi = 0, \quad m_{12}^2 = 0, \quad \lambda_6 = 0, \quad \lambda_7 = 0$), which makes the lightest Higgs field is stable, on eqs. (5, 6) with $\lambda_3+\lambda_4+\lambda_5 = 0$ [1a] we get four solutions for $\nu_1$ and $\nu_2$, which are

$$\nu_1 = \tau_1\ , \qquad \nu_2 = \tau_2\ ; \quad (7)$$

$$\nu_1 = \tau_1\ , \qquad \nu_2 = \frac{2m_{22}^4\Lambda_2^2\rho_2\tau_2 - \lambda_2\Lambda_4\rho_4\tau_2 - \sqrt{\Xi_2}}{2\left(m_{22}^4\Lambda_2^2\rho_2 + \lambda_2\Lambda_4\rho_4\right)}\ ; \quad (8)$$

$$\nu_1 = \frac{2m_{11}^4\Lambda_1^2\rho_1\tau_1 - \lambda_1\Lambda_3\rho_3\tau_1 - \sqrt{\Xi_1}}{2\left(m_{11}^4\Lambda_1^2\rho_1 + \lambda_1\Lambda_3\rho_3\right)}\ , \qquad \nu_2 = \tau_2\ ; \quad (9)$$

$$\nu_1 = \frac{2m_{11}^4\Lambda_1^2\rho_1\tau_1 - \lambda_1\Lambda_3\rho_3\tau_1 - \sqrt{\Xi_1}}{2\left(m_{11}^4\Lambda_1^2\rho_1 + \lambda_1\Lambda_3\rho_3\right)}\ , \qquad \nu_2 = \frac{2m_{22}^4\Lambda_2^2\rho_2\tau_2 - \lambda_2\Lambda_4\rho_4\tau_2 - \sqrt{\Xi_2}}{2\left(m_{22}^4\Lambda_2^2\rho_2 + \lambda_2\Lambda_4\rho_4\right)}\ ; \quad (10)$$

where

$$\Xi_1 = -4m_{11}^6\Lambda_1^3\rho_1^2 - 4m_{11}^2\lambda_1\Lambda_1\Lambda_3\rho_1\rho_3 - 8m_{11}^4\lambda_1\Lambda_1^2\Lambda_3\rho_1\rho_3\tau_1^2 + \lambda_1^2\Lambda_3^2\rho_3^2\tau_1^2,$$

$$\Xi_2 = -4m_{22}^6\Lambda_2^3\rho_2^2 - 4m_{22}^2\lambda_2\Lambda_2\Lambda_4\rho_2\rho_4 - 8m_{22}^4\lambda_2\Lambda_2^2\Lambda_4\rho_2\rho_4\tau_2^2 + \lambda_2^2\Lambda_4^2\rho_4^2\tau_2^2.$$

If we choose $\tau_2 = 0$ then the fields $\phi_1^\pm$ and $\chi_1$ become Goldstone bosons and the other fields become physical. With $\tau_2 = 0$, the Yukawa interactions are described by the interaction of $\phi_1$ with fermions (as $\phi_2$ does not couple to fermions but appears in loops). This scenario makes the lightest field of $\phi_2$ stable.

[a] With $\lambda_3+\lambda_4+\lambda_5 \neq 0$ the solution to the VeV contain complex part.

## 3. Higgs fields as dark energy

For the field $\phi_2$ to be the dark energy field, we need to solve the Euler-Lagrange equations in FRW flat, ($\kappa$), Universe, ($\sqrt{-g} = a(t)^3$), which for the fields $\phi_2^{\pm}$, $\eta_2$ and $\chi_2$ are

$$\ddot{\eta}_2 + 3\frac{\dot{a}}{a}\dot{\eta}_2 + \frac{1}{2}\eta_2\left(\nu^2\left(\lambda_3+\lambda_4+\lambda_5\right) + 2m_{22}^2\Lambda_2\rho_2 e^{m_{22}^2\Lambda_2\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)/2}\right.$$
$$\left. + \lambda_2\Lambda_4\rho_4 e^{\lambda_2\Lambda_4\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)^2/8}\left({\chi_2}^2+2{\phi_2^c}^2\right)\right)$$
$$+ \frac{1}{2}\lambda_2\Lambda_4\rho_4 e^{\lambda_2\Lambda_4\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)^2/8}{\eta_2}^3 = 0, \tag{11}$$

$$\ddot{\chi}_2 + 3\frac{\dot{a}}{a}\dot{\chi}_2 + \frac{1}{2}\chi_2\left(\nu^2\left(\lambda_3+\lambda_4-\lambda_5\right) + 2m_{22}^2\Lambda_2\rho_2 e^{m_{22}^2\Lambda_2\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)/2}\right.$$
$$\left. + \lambda_2\Lambda_4\rho_4 e^{\lambda_2\Lambda_4\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)^2/8}\left({\eta_2}^2+2{\phi_2^c}^2\right)\right)$$
$$+ \frac{1}{2}\lambda_2\Lambda_4\rho_4 e^{\lambda_2\Lambda_4\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)^2/8}{\chi_2}^3 = 0, \tag{12}$$

$$\ddot{\phi}_2^c + 3\frac{\dot{a}}{a}\dot{\phi}_2^c + \phi_2^c\left(\nu^2\lambda_3 + 2m_{22}^2\Lambda_2\rho_2 e^{m_{22}^2\Lambda_2\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)/2}\right.$$
$$\left. + \lambda_2\Lambda_4\rho_4 e^{\lambda_2\Lambda_4\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)^2/8}\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)\right) = 0, \tag{13}$$

where $c$ is $+$ or $-$. The energy density and pressure after expansion of UDW-2HDM Higgs Lagrangian for physical fields become

$$\rho_{\rm DE}/{\rm P}_{\rm DE} = \frac{1}{2}{\dot{\chi}_2}^2 + \frac{1}{2}{\dot{\eta}_2}^2 + \frac{1}{2}\dot{\phi}_2^{c2} \pm \left(\rho_1+\rho_3 + \frac{1}{2}\nu^2\lambda_3\phi_2^{c2} + \frac{1}{4}\nu^2\left(\lambda_3+\lambda_4+\lambda_5\right){\eta_2}^2\right.$$
$$+ \frac{1}{4}\nu^2\left(\lambda_3+\lambda_4-\lambda_5\right){\chi_2}^2 + \rho_2 e^{m_{22}^2\Lambda_2\left({\chi_2}^2+{\eta_2}^2+2{\phi_2^c}^2\right)/2}$$
$$\left. + \rho_4 e^{\lambda_2\Lambda_4\left({\chi_2}^4+2{\chi_2}^2{\eta_2}^2+{\eta_2}^4+4{\chi_2}^2\phi_2^{c2}+4{\eta_2}^2\phi_2^{c2}+4\phi_2^{c4}\right)/8}\right). \tag{14}$$

The initial conditions used are $\eta_{2ini} = M_P$, $\chi_{2ini} = M_P$, $\phi_{2ini}^c = 0$, $\dot{\chi}_{2ini} = 0$, $\dot{\chi}_{2ini} = 0$, $\dot{\phi}_{2ini}^c = 0$. The masses of the Higgs bosons in the analysis are taken to be $m_{\eta_2} = m_{\chi_2} = 1.0247 \times 10^{-59}$ GeV, the charged Higgs mass is arbitrary. The solution of the eqs. (11, 12, 13) along with Friedmann equations is shown below in the graphs.

During the initial stages $Z \gg 1$, the evolution of the Higgs fields, $\eta_2$, $\chi_2$ and $\phi_2^{\pm}$, is frozen, and acts as a negligibly small vacuum energy component with $\omega = -1$. As time proceeds the Higgs fields begin to evolve towards the minimum of the potential, the energy density in the Higgs fields starts to dominate cosmologically. During the evolution, $\omega_{Higgs}$ starts to increase and becomes $> -1$ as shown in fig. 2. In the very late (future) Universe ($Z \ll 0$), the fields come to rest at the minimum of the potential and a period with $\omega = -1$ is re-achieved to give an accelerating Universe similar to a pure cosmological constant.

As discussed before, the Higgs field stability is provided by imposing $Z_2$ symmetry. The lightest Higgs fields, $\eta_2$ and $\chi_2$, do not decay into any other Higgs field

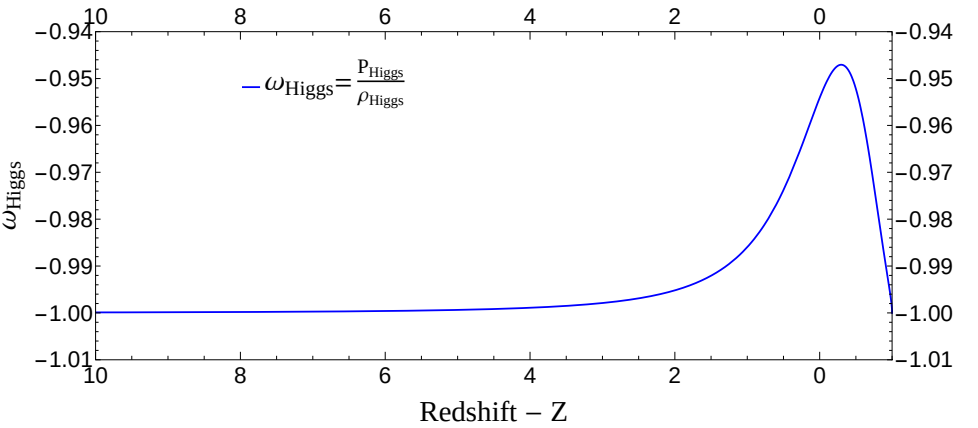



Fig. 2. Effective equation of state parameter for Higgs fields $\omega_{Higgs}$, as seen it starts with $-1$ then evolves towards quintessence regime after large enough time it comes back at $-1$.

(since these fields are lighter than the SM-like and charged Higgs) or into fermions (since they do not couple to them at tree level).

Note that the initial conditions for the charged field took the dark energy (vacuum) not to be charged. Since it does not contribute to relic density, this does not mean that it can not exist, it can appear in the loop process.

## 4. Conclusion

In the work presented, we assumed that dark energy is actually some scalar field which is present as the Higgs in a model where the potential has the non-degenerate vacua, we called this model uplifted double well two-Higgs doublet model (UDW-2HDM).

We found that if the present Universe is described by the true vacua of UDW-2HDM then the component fields of the second doublet $\phi_2$ (which acts as the inert doublet) can be one possible candidate for the dark energy. As the present contribution of the dark energy to the critical energy density is about 0.7, this value is obtained by taking the mass of the CP-even field's mass small ($O(10^{-59})$GeV). The most important thing is that with the initial conditions set, the mass of the charged ($\phi_2^{\pm}$) field becomes arbitrary. Hence this model will fit for any value of mass of $\phi_2^{\pm}$. One also needs to keep in mind that the values of masses were chosen arbitrarily so as to get dark energy relic density $\approx 0.7$. Changing the values of the masses, the relic density does not change much.

It should also be mentioned that if we remove the $Z_2$ symmetry, the second Higgs doublet does not remain inert. Thus in the case of $Z_2$ violation (soft or hard), the CP-even Higgs fields will mix by an angle $\beta$. In that case a new parameter ($\beta$) will arise in the theory. Obtaining a dark energy candidate in that model will require fine tuning in the Yukawa interactions in such a way that either the dark energy field does not couple or couple very weakly with the fermions.

## References


1. Muhammad Usman, Dark energy via multi-Higgs doublet models: accelerated expansion of the Universe in inert doublet model scenario, *ArXiv e-prints* (June 2015).
2. M. Usman and A. Qadir, The extra scalar degrees of freedom from the two-higgs doublet model for dark energy, *International Journal of Modern Physics D* **26**, p. 1741003 (2017).
3. M. Usman and A. Qadir, *Higgs Dark Energy in Inert Doublet Model*, in *Mathematical Physics*, pp. 180–185.
4. Muhammad Usman and Asghar Qadir, Dark energy via multi-Higgs doublet models: accelerated expansion of the Universe in the two Higgs doublet model scenario (2016), submitted.
5. E. Greenwood, E. Halstead, R. Poltis and D. Stojkovic, Electroweak vacua, collider phenomenology, and possible connection with dark energy, *Phys. Rev. D* **79**, p. 103003 (May 2009).
6. S. V. Ketov and N. Watanabe, On the Higgs-like quintessence for dark energy, *Modern Physics Letters A* **29**, p. 1450117 (2014).
7. S. M. Carroll, M. Hoffman and M. Trodden, Can the dark energy equation-of-state parameter $w$ be less than $-1$?, *Phys. Rev. D* **68**, p. 023509 (Jul 2003).
8. A. W. E. Kaffas, W. Khater, O. M. Ogreid and P. Osland, Consistency of the two-higgs-doublet model and {CP} violation in top production at the {LHC}, *Nuclear Physics B* **775**, 45 (2007).
9. I. F. Ginzburg, K. A. Kanishev, M. Krawczyk and D. Sokołowska, Evolution of the universe to the present inert phase, *Phys. Rev. D* **82**, p. 123533 (Dec 2010).
10. G. Branco, P. Ferreira, L. Lavoura, M. Rebelo, M. Sher and J. P. Silva, Theory and phenomenology of two-Higgs-doublet models, *Physics Reports* **516**, 1 (2012).

# Structure formation in the early universe, the three-body problem, and gravity modification w.r.t. possible replacement of DE

Andrew Walcott Beckwith

*Physics Department, Chongqing University, Chongqing 401331, People's Republic of China*
*E-mail: Rwill9955b@gmail.com; abeckwith@uh.edu*
*english.cqu.edu.cn*

We find that having the scale factor close to zero due to a given magnetic field value in an early-universe magnetic field affects how we would interpret Mukhanov's chapter on self-reproduction of the universe. A stronger early-universe magnetic field suggests a greater likelihood of production of about 20 new domains of size $\frac{1}{H}$, with $H$ the early-universe Hubble constant per Planck time interval in evolution. Fluctuations in the Hubble expansion parameter, $H$, may affect structure. Finally, $\alpha$ in a gravitational potential is proportional to $r^{-\alpha}$. This adjustment affects the three-body problem.

## 1. Introduction

Part 1: We examine how a scale factor is affected by the NLED paradigm, also linked to *self-reproduction,*[1] another way to outline how this affects the evolution of density in the early universe. Next, after having done this, we give a description of an equation for setting the value of $\alpha$ in a gravitational potential proportional to $r^{-\alpha}$. This $\alpha$ has real and complex values, unlike the Newtonian real value.

Part 2: We summarize what this has to do with possible revisions of the three-body problem[2] with a particular mention as to how this affects astrophysics.[2] (p. 141)

## 2. How a Scale Factor Is Affected by the NLED Paradigm

First we look at structure formation, mainly self-reproduction of inhomogeneity in terms of early-universe conditions.[1] In this, the starting point is the meme of chaotic inflation. That is, inflation generated by a potential:[1,3]

$$V_{\text{potential}} \sim \phi^2. \tag{1}$$

In this, one can look at a scalar field at the end of (chaotic) inflation,[1] with an amplitude given by $\phi_i$ for the initial value of the inflaton such that

$$\delta_{\phi_{\text{Max}}} \sim m \cdot \phi_i^2, \tag{2}$$

where $m$ will be determined by NLED inputs to be brought up later. In terms of the initial inflaton, inhomogeneities do not form if the initial inflaton is bounded:[1]

$$m^{-1} > \phi_i > m^{-\frac{1}{2}}. \tag{3}$$

This leads to (low?) inflation-generated inhomogeneity in space–time. Inflation is eternal[1,3] if there is only the inequality

$$\phi_i > m^{-\frac{1}{2}}. \tag{4}$$

### 2.1. *NLED Applied to Eq. (4) Plus Details of Structure Formation*

We examine the following treatment of mass with this as our starting point. If $l_{\text{Planck}}$ is the Planck length and $\alpha > 0$, then

$$m \sim 10^{\alpha} \cdot l_{\text{Planck}}^{3} \cdot \rho_{\text{density}}. \tag{5}$$

Then we can consider the following formulation of density. If we do not wish to consider a rotating universe, then Camara et al.[4] has an expression as to density, with a $B$-field contribution to density, and we also can used the Weinberg result[5] of scaling density with one over the fourth power of a scale factor, which we will remark upon in the general section, as well as the Corda and Mosquera Cuesta result[6] for density is for a star, whereas the Camarade et al. result[4] is for a universe). In addition, Corda and Mosquera Cuesta[6] use quintessential density to falsify the null energy condition of a Penrose theorem.[7] With further details of what Penrose was trying to do as to this issue of GR,[8] and to answer how to violate the null-energy condition, one should go to[6] for quintessential density defined, with the constant in Eq. (4) greater than zero. Then in both the massive star and the early universe, this density result is applicable.

$$\rho_{\gamma} = \frac{16}{3} \cdot c_1 \cdot B^4. \tag{6}$$

Keeping in mind what was said as to choices of what to do about density, and its relationship to Eq. (5), we then can reference structure formation as follows.[1] Look at how a Hubble parameter changes with respect to cosmic evolution. It changes with respect to $H_{\text{today}}$ being the Hubble parameter in the recent era, and the scale factor $\alpha$, with this scale factor being directly responsive to changes in density.[5] That is, $\rho \sim a^{-4}$. In the next section, we examine how to vary the scale factor cited in Eq. (7).[4] In this section, we note of what the scale factor does[4] to the Hubble parameter given in Eq. (7), and then in the section afterwards review a possible reconciliation of how Eq. (6) and Eq. (7) define early-universe parameters. To understand, we consider what happens to the Hubble parameter:

$$H \sim \frac{H_{\text{today}}}{a^{\frac{3}{2}}}. \tag{7}$$

If Eq. (4) holds,[1] then inhomogeneous patches of space–time appear in a causal region of space–time for which

$$\textit{Causal–domain} \sim H^{-1} \sim \frac{a^{\frac{3}{2}}}{H_{\text{today}}}. \tag{8}$$

Furthermore, about 20 such domains are created in a Hubble time interval $\Delta t_H \propto H^{-1}$.[1] As a function of say $10^{\alpha}$ times Planck time, for a domain size given by Eq. (8) and this requires then a clear statement as to how the scale factor changes[4] and reconciling the density expressions given in Eq. (6) and Eq. (7).

### 2.2. *Nonzero Initial Universe Radius from Nonlinear Space–Time*

We assert a scaled parameter $\lambda$,[1] and $a_0$, paired with $\alpha_0$. For the sake of argument, we will set $a_0 \propto \sqrt{t_{\text{Planck}}}$, with $t_{\text{Planck}} \sim 10^{-44}$ s. Also, $\lambda$ is a cosmological "constant" parameter described later as quintessence:[4,9]

$$\alpha_0 = \sqrt{\frac{4\pi G}{3\mu_0 c}} B_0, \text{ and } \lambda = \frac{\Lambda c^2}{3}. \tag{9}$$

Then if, initially, Eq. (9) is large, due to a very large $\Lambda$ the time[4] (Eq. 53) is such that, even though there is an expanding and contracting universe, the key time parameter may be set to $t_{\min} \approx t_0 \equiv t_{\text{Planck}} \sim 10^{-44}$ s. Whenever one sees the coefficient like the magnetic field, with the small 0 subscript, for large values of $\Lambda$, this should be the initial coefficient at the beginning of space–time which helps us make sense of the nonzero but tiny minimum scale factor.[1]

$$a_{\min} = a_0 \cdot \left[ \frac{\alpha_0}{2\lambda} \left( \sqrt{a_0^2 + 32\lambda\mu_0\omega B_0^2} - \alpha_0 \right) \right]^{\frac{1}{4}}. \tag{10}$$

The minimum time most likely means, due to large $\Lambda$ that Eq. (10) is of the order of about $10^{-55}$, 33 orders of magnitude smaller than the square root of Planck time. We next will be justifying the relative size of the $\Lambda$.

### 2.3. *An NLED Large, Varying $\Lambda$ and Early-Universe Scale Factor*

Dispite the temperature variation[9] for the cosmological Hubble parameter, we can set $\Lambda(t) \sim H^2_{\text{inflation}}$.[4] In short, we obtain its equivalence with $\Lambda_{\text{Max}} \sim c_2 T \widetilde{\beta}$.[9] Then, comparing Eq. (6) and Eq. (7) leads to the following constraints.

$$(\rho \sim a^{-4})^{-1} \sim a^4 \sim \frac{16}{3} \cdot c_1 \cdot B^4 \sim \frac{a_0^{-4}}{\frac{\alpha_0}{2\lambda}\left( \sqrt{a_0^2 + 32\lambda\mu_0\omega B_0^2} - \alpha_0 \right)}. \tag{11}$$

Equation (11) suggests a large $B$ field in Eq. (11). Note that[10]

$$\frac{\Delta H}{H} \sim \Omega_m^2 h^2 \Delta_{\Re} \sim 10^{-5}. \tag{12}$$

Here, if there is a flat universe[11] and taking $H^2 = \frac{8\pi}{3} \cdot \rho$ and using the magnetic field's relation to density,[12,6] then the number of operations, $n$,

$$n \sim \rho_{\text{crit}} \times t^4 \sim \left( \frac{t}{t_p} \right)^2 \Leftrightarrow \rho_{\text{crit}} \sim \frac{1}{t^2} \propto \rho_\gamma = \frac{16}{3} \cdot c_1 \cdot B^4$$
$$\Leftrightarrow n \sim \left( \frac{t}{t_p} \right)^2 \cdot \frac{1}{B^8} \sim \frac{1}{t_p B^4} \sim \frac{1}{B^4}. \tag{13}$$

Adding Lloyd's treatment of information,[12] we estimate an early-universe fluctuation at, if we have a baseline number for initial (expansion) value of the Hubble

parameter, $H_{\text{baseline}}$. This is a starting point for an expanding universe with $n$[12] as an initial function of entropy. So then, we may see in the CMBR

$$\Delta H_{\text{thermal}} \sim h_{\text{baseline}} \cdot n^{\frac{1}{4}} \cdot 10^{-5} \cdot \sqrt{\frac{t}{t_{\text{Planck}}}}. \tag{14}$$

In Eq. (15)–(16), with $H$ given by Eq. (16) and, as in Adamek et al.,[10] (Fig. 2) we have, perhaps, the beginning of how NLED may impact fluctuations in $H$. That, in turn, may lead to the issue that opened our discussion. Equation (14) may give some insight as to the fluctuations in Adamek et al.[10] (Fig. 2)

## 3. The Problem of $\alpha \propto r^{-\alpha}$ in a Gravitational Potential

To review this, we use the following treatment of the Klein–Gordon equation:[13]

$$\ddot{\varphi}_k + 3H\dot{\varphi}_k + \tfrac{k^2}{a^2}\varphi_k = 0, \quad \varphi_k \approx \tfrac{H\tau}{\sqrt{2k}}\left(1 + \tfrac{1}{ik\tau}\right)\mathrm{e}^{-ik\tau}, \quad \text{and } \tau = -\tfrac{\mathrm{e}^{-Ht}}{H}. \tag{15}$$

Here, $k$ is the wave number and $H$ is constant in the early universe. Thus $k = \frac{2\pi}{\lambda}$, with $\lambda$ proportional to the "width" of a would-be preuniverse "bubble"[14] in place of a singularity. Also, one would have for a constant $H$, during this time,[15]

$$H = \sqrt{\frac{8\pi G}{3} \cdot \rho - \frac{\kappa}{a^2}} = \sqrt{\frac{8\pi G}{3}} \cdot \sqrt{\rho - \frac{3}{8\pi G} \cdot \frac{\kappa}{a^2}} \approx \sqrt{\frac{8\pi G}{3}} \cdot \sqrt{V(\varphi) - \frac{\dot{\varphi}^2}{2}}$$
$$\Leftrightarrow \frac{\dot{\varphi}^2}{2} = \left(\frac{3}{8\pi G} \cdot \frac{\kappa}{a^2} + V(\varphi)\right) - \rho, \tag{16}$$

where $\rho$ is energy density and $\kappa$ is curvature. Chaotic inflation uses $V(\varphi) \approx \frac{k^2}{a^2} \cdot \varphi^2$, the time derivative is $\frac{\mathrm{d}}{\mathrm{d}\tau}$, and $\varphi = \varphi_k$:

$$\frac{\dot{\varphi}^2}{2} = \frac{3}{8\pi G} \cdot \frac{\kappa}{a^2} + \frac{k^2}{a^2} \cdot \varphi_k^2 - \rho = \frac{3}{8\pi G} \cdot \frac{\kappa}{a^2} + \frac{k^2}{a^2} \cdot \varphi_k^2 - \frac{16}{3} \cdot c_1 \cdot B^4 \approx \Delta E. \tag{17}$$

The last term of Eq. (17) states that, if we apply it to the Pre-Planckian-to-Planckian regime, there will be a change in the energy, We call this shift in energy equivalent to a change in *kinetic* energy, and then reference the Virial theorem:[16]

$$\left\langle \psi \middle| \Delta E \approx \frac{3}{8\pi G}\frac{\kappa}{a^2} + \frac{k^2}{a^2}\varphi_k^2 - \frac{16}{3}c_1 B^4 \middle| \psi \right\rangle \approx \left\langle \psi \middle| r \cdot \Delta\left(V_{\text{PE}} \approx \frac{c_2}{r^\alpha}\right) \middle| \psi \right\rangle, \tag{18}$$

where $V_{\text{PE}}$ is potential energy. In Pre-Planckian–Planckian space–time, we approximate, in the instant before time is initialized, the mean-value theorem for the computed values of both Eq. (17) and (18) with our results:

$$\frac{3}{8\pi G}\frac{\kappa}{a^2} + \frac{k^2}{a^2}\varphi_k^2 - \frac{16}{3}c_1 B^4 \approx -\frac{\alpha}{r^\alpha} \equiv -\frac{\alpha}{l^\alpha} \Leftrightarrow \frac{\alpha}{l^\alpha} \approx \frac{16}{3}c_1 B^4 - \frac{3}{8\pi G}\frac{\kappa}{a^2} - \frac{k^2}{a^2}\varphi_k^2. \tag{19}$$

Here, the magnetic field is determined by $B$, the scale factor $a$ (Eq. (12)), and $\varphi_k$ (Eq. (19)). $\alpha$ is no longer strictly real and is influenced by the complex $\varphi_k$. We next conclude with how the three-body problem, and in fact other experimental science could be impacted by Eq. (19). This merits scrutiny of the Lagrangian equilateral

triangle[2] (p. 134) and possible orbits from the three-body problem.[17] (p. 15) Here, we consider both in terms of KAM theory,[2,15] a highly restricted set of equations:

> When the system is slightly perturbed most of the invariant tori are not destroyed but only slightly shifted in the phase space. This has important implications on stability of orbits in the general and restricted three body problem. The proof of the KAM theorem by Moser [1962] and Arnold [1963] also demonstrated that convergent power series solutions exist for the three-body (as well as for the $n$-body) problem. The KAM theorem seems to be very useful for studying the global stability in the three-body problem [Robutel, 1993a, Montgomery, 2001, Simó, 2002]; however, some of its applications are limited only to small masses of the third body.[17] (p. 15)

The limitation of KAM theory is that, most of the time, we have that the KAM results require a third body to be low mass, a classic dynamic-system result.

## 4. Conclusion

Via careful use of Eq. (19), we must remove dependence upon third-mass size, and examine if this can still recover much of the stability analysis. Later, this question in terms of a serious application of the value of Eq. (19) will be pursued, Second, orbital changes in encounters with planets is a restricted three-body problem,[2] (p. 154) frequently used as to the interaction of say comets (comparatively small mass) with planets circulating the Sun, where we have two "massive" masses, and the third body. This could be used to show how the hyperbolic trajectory of a minor mass object (comet) is impacted in a simple solar-system model. We hope that, by judicious investigation of the arguments given for Eq. (19), the restrictions as to the smallness of the third mass may be partly ameliorated. If this is done, and it will require through investigations, then a template as to how to reliably simulate $N$ bodies interacting may be a doable problem. To do this, one must investigate if the classical KAM problem may be generalized beyond its present restrictions.

## Acknowledgements

This work is supported in part by National Natural Science Foundation of China grant No. 11375279.

## References

1. V. Mukhanov, *Physical Foundations of Cosmology* (Cambridge University Press, New York, 2005).
2. M. Valtonen and H. Karttunen, *The Three-Body Problem* (Cambridge University Press, New York, 2005).
3. A. H. Guth, Eternal inflation and its implications, *J. Phys. A* **40**, 6811 (2007), https://doi.org/10.1088/1751-8113/40/25/S25.

4. C. S. Camara, M. R. de Garcia Maia, J. C. de Carvalho and J. A. de Lima, Nonsingular FRW cosmology and non linear dynamics, *Phys. Rev. D* **69**, p. Article 063501 (2004).
5. S. Weinberg, *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity* (John Wiley and Sons, New York, 1972).
6. C. Corda and H. J. M. Cuesta, Removing black hole singularities with non linear electrodynamics, *Mod. Phys. Lett. A* **25**, 2423 (2010).
7. R. Penrose, Gravitational collapse and space–time singularities, *Phys. Rev. Lett.* **14**, 57 (1965), https://doi.org/10.1103/PhysRevLett.14.57.
8. R. Penrose, Gravitational collapse: The role of general relativity, *Riv. Nuovo Cim.* **1**, 252 (1969), [Gen. Rel. Grav. 34,1141 (2002)].
9. D. Park, H. Kim and S. N. Tamarayan, Nonvanishing cosmological constant of flat universe in Brane world scenario, *Phys. Lett. B* **535**, 5 (2002).
10. J. Adamek, C. Clarkson, R. Durrer and M. Kunz, Does small scale structure significantly affect cosmological dynamics? *Phys. Rev. Lett.* **114**, p. Article 051302 (2015), https://doi.org/10.1103/PhysRevLett.114.051302.
11. A. H. Guth, Phase transitions in the early universe, in *The Very Early Universe, Proceedings of the Nuffield Workshop*, eds. G. W. Gibbons, S. W. Hawking and S. T. C. Siklos (Cambridge University Press, Cambridge, UK, 1982) pp. 171–204.
12. S. Lloyd, A theory of quantum gravity based on quantum computation (2018), http://arxiv.org/abs/quant-ph/0501135.
13. C. Stahl, On early and late phases of acceleration of the expansion of the universe (2017), Sapienza University of Rome. https://arxiv.org/pdf/1702.05630.pdf.
14. A. Beckwith, How a minimum time step and formation of initial causal structure in space–time may void the Penrose singularity theorem, as in Hawking's and Ellis's 1973 write ups (2017), http://vixra.org/abs/1705.0213.
15. S. Carroll, *Spacetime and Geometry* (Pearson India, New Deli, 2016).
16. S. Gasiorowitz, *Quantum Physics*, 1st edn. (John Wiley and Sons, New York, 1974).
17. Z. E. Musielak and B. Quarles, The three-body problem, *Rep. Prog. Phys.* **77**, p. Article 065901 (2014), https://doi.org/10.1088/0034-4885/77/6/065901.

# Constraining time-dependent dark energy with the flux power spectrum of the Lyman $\alpha$ forest

G. J. Mathews, J. W. Coughlin, L. A. Phillips, A. P. Snedden and I.-S. Suh

*Department of Physics, Center for Astrophysics, University of Notre Dame, Notre Dame, IN 46556, USA*
** E-mail: gmathews@nd.edu*
*https://physics.nd.edu/people/faculty/grant-j-mathews/*

Calculations of the flux power spectrum of the Lyman $\alpha$ forest are performed as a means to quantify the possible effects of time-dependent dark energy. We use a parameterized version of the time-dependent dark energy equation of state consistent with the *Planck* analysis. We have run high-resolution, large-scale cosmological simulation with a modified version of the publicly available SPH code `GADGET-2`. These simulations were used to extract synthetic Lyman $\alpha$ forest spectra. These were then used to simulate the flux power spectrum at various observed quasar redshifts. We conclude that the effect of time-dependent dark energy on the flux power spectrum is of marginal statistical significance compared to the intrinsic cosmic variance.

*Keywords*: Cosmology: dark energy — cosmology: theory — methods: numerical.

## 1. Introduction

The origin of the present cosmic acceleration remains a mystery. Within the dark energy paradigm several time-dependent models have been proposed as the physical mechanism for cosmic acceleration.[1]. The simplest model is that of a cosmological constant. Another possibility, however, is a self-coupled, slowly evolving scalar field that manifests as either a cosmic quintessence[2] or k-essence[3]. The empirical difference between the cosmological constant and all other dark energy models is that the energy density of the later can vary in time. Hence, discriminating between various dark energy models can be made by observing how the dark energy changes in time.

Current observational constraints on time-dependent dark energy, however, are quite weak[4]. In the work described here[5] we consider using the Lyman-$\alpha$ forest as a means to constrain the time evolution of dark energy. The motivation for using the Lyman-$\alpha$ forest is that the effects of dark energy, should be most apparent on the morphology of voids in the cosmic web.[6–12] The absorbers responsible for the Lyman-$\alpha$ forest should reside primarily in the clusters and filaments.[13,14] Along the line of sight to distant quasars these absorbers will be separated by the voids. As such, the separation of these absorbers in redshift space should act as a tracer of the evolution of the voids.[15] Hence, the flux power spectrum acts as a proxy for the matter power spectrum and should contain complementary information about how each dark energy model affects the expansion of the universe. Thus, studying the Lyman-$\alpha$ forest could be an independent approach to searches based on the SNIa redshift-distance relation, the CMB, BAO, ISW, and gravitational lensing, and one that has received comparatively little attention in the literature.

## 2. Simulations

Previously Ref. [15] considered the effects on the Lyman-$\alpha$ forest from various values of cosmological constant in the context of a semi-analytic treatment of the Lyman-$\alpha$ forest. In Ref. [5] we expanded upon that work by making use of high-resolution N-body simulations, from which we extracted synthetic Lyman-$\alpha$ spectra. Moreover, we have explored dynamical models of dark energy.

Our simulations were run with the publicly available smoothed-particle-hydrodynamics (SPH) code GADGET-2[16], modified[5] to work with dynamical dark energy.[17] Simulating the Lyman-$\alpha$ forest requires a very high resolution. It has been suggested[18] that a resolution of $\approx 40\ h^{-1}$ kpc in a box of size $L \approx 40\ h^{-1}$ Mpc is needed to adequately resolve the structure of the Lyman-$\alpha$ forest and achieve convergence in the power spectrum. With these requirements in mind, we simulated $1024^3$ dark matter particles in a box of length $40\ h^{-1}$ comoving Mpc.

Due to the high resolution requirements of our simulations, we evolve a distribution of only dark matter particles out of consideration for the total run-time. This is justified because in the low-density, mildly non-linear environments typically responsible for the Lyman-$\alpha$ forest, the baryon distribution largely follows that of the dark matter on large scales.[19] Additionally, the effects of non-linear baryonic physics, such as galactic winds, have been shown[20] to be small at large scales where the effects of dark energy should be most prominent.

Our cosmological parameters are those given in Ref. [21]. Other simulation parameters are given in Ref. [5]. We ran five simulations both with a cosmological constant and with various dynamical dark energy where $w(a)$ is the linear slope of the dark energy equation of state with redshift. The dynamical models were chosen such that their parameters were as close to the edges of the allowed 95% confidence range for the $(w_0, w_a)$ parameter space given in the *Planck* analysis.[4] We chose points at the fringes of the allowed parameter space since these models should produce flux power spectra with the largest differences between them. The only exception to this is model DE2-40-1024, $(w_0, w_a) = (-1.1, 1.3)$ which was deliberately chosen from the region of the $(w_0, w_a)$ parameter space that is outside the bounds permitted by *Planck*'s marginalized posteriors. This was done for two reasons: first, we wanted to determine if very extreme values the dark energy parameters were capable of producing a distinct signature in the flux power spectrum of the Lyman-$\alpha$ forest, and second, we were interested in whether or not the flux power spectrum provided constraints on the dark energy parameter space that were in accord with the results determined from other observational probes. The values of $w_0$ and $w_a$ that we used are $(w_0, w_a) = (-1.0, 0.0), (0.0, -3.0), (-1.1, 1.3), (-2.0, 0.0), (-2.0, 2.0)$.

Each of our simulations was started from the same initial conditions and evolved from $z = 49$ to $z = 2.2$. Our initial conditions were generated using the publicly available second-order Lagrangian perturbation theory code `2LPTIC`[22]. We generated snapshots of each simulation for quasars located at observed redshifts of $z = 4.2, 3.8, 3.0, 2.7$, and $z = 2.2$ as described in Ref. [5]. Our simulations

required on the order of $\approx 10$ days on 72 processors to run. Following this, the post-processing took on the order of one week per snapshot, with the majority of the time (about four or five days) being devoted to halo-finding.

## 3. Implementation of Dynamical Dark Energy

The publicly available version of GADGET-2 assumes that dark energy arises from the cosmological constant. However, we make use of the parameterization given in 5. Nevertheless, modifying GADGET-2 to include the effects of dynamical dark energy was relatively straight forward. Calculating a synthetic spectrum requires knowledge of the densities, temperatures, and H*I* neutral fractions for each of the simulation particles. Since these are not properties of dark matter particles in GADGET-2, and we ran dark matter only simulations, we calculated these quantities in post-processing. Details of the GADGET calculations and the extraction of the Lyman-$\alpha$ spectra are given in Ref. [5].

The absorption in the Lyman-$\alpha$ forest serves as a way of mapping out the large-scale structure between the observer and the distance source quasar. Thus, the flux power spectrum serves as a proxy for the power spectrum of the underlying matter field that gives rise to the absorption in the Lyman-$\alpha$ forest. Since the matter power spectrum is a measure of the density amplitudes as a function of scale, and these amplitudes depend upon the expansion history of the universe, we can, in principle, use the flux power spectrum to discriminate between dark energy models.

Following Ref. [23], we do not analyze the flux directly but instead consider the quantity:

$$F_p = \frac{e^{-\tau}}{\langle e^{-\tau} \rangle} - 1. \tag{1}$$

$F_p$ is used in place of $F$ because $F$ is sensitive[23] to changes in the mean flux $\langle e^{-\tau} \rangle$. We then take the Fourier transform of $F_p$ using the publicly available package `FFTW3`[a] to calculate the power spectrum. That is, the power spectrum is constructed from the Fourier transform of $F_p$, which we denote $F_{p,k}$,

$$P_{F_p}(k) = |F_{p,k}|^2. \tag{2}$$

We normalize the spectrum by dividing out the total counts within each bin of frequency $k$. The frequencies are found via:

$$k_i = \frac{2\pi i}{T}, \tag{3}$$

where $i$ indicates the bin index and $T$ is the period. For discrete Fourier Transforms (DFTs), the signal is assumed to be periodic over the range in which there is data, so the period is simply the length of the spectrum in velocity space.

[a] `http://www.fftw.org/`

## 4. Results

Using the procedure described in Ref. [5], we extract one-thousand synthetic spectra for each model at each redshift. This pool of spectra was then used to calculate an average power spectrum for each dark energy model at each redshift, along with errors. Figure 1 shows the spectrum through the center of the simulation volume for each of our dark energy models. These are offset from one another for visual clarity. All of the spectra, save for the DE2-40-1024 model, are nearly identical to one another.

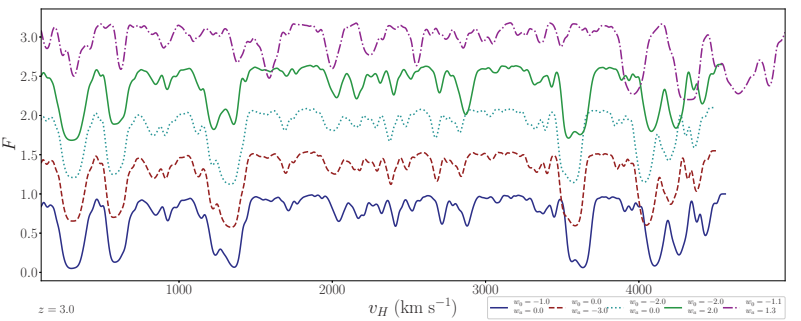



Fig. 1. The synthetic spectrum through the center of the simulation volume for each of our dark energy models at $z = 3.00$. This figure highlights the fact that the spectrum for each of our dark energy models, save for DE2-40-1024 (magenta), are all nearly identical to one another. Each spectrum is offset from the others for reasons of visual clarity.

In order to calculate an average power spectrum at each redshift, we bootstrap a sample of eight-hundred synthetic Lyman-$\alpha$ forest spectra from our pool of one-thousand. The process described in Ref. [5] is then applied to this bootstrapped sample in order to calculate one instance of the power spectrum. This process is then repeated one thousand times to generate one thousand instances of the power spectrum. These instances are then averaged on a $k$-by-$k$ basis in order to obtain a mean value of $P$ for each $k$ along with a standard deviation, which we take to be the errors on our simulated power spectra.

We applied[5] a k-sample Anderson-Darling (AD) test[24] to analyze the statistical significance of the differences among power spectra. This test places more emphasis on the tails of the distribution. Since dark energy is a large-scale phenomenon, we expect the differences between power spectra to be greatest on the largest scales (smallest $k$). The key feature of the AD test is the fact that it is distribution-free. This is because the underlying distribution function giving rise to the power spectrum is unknown. With no *a priori* reason to prefer one distribution function over any other, a reasonable place to start is to assume a Gaussian distribution, though we would not necessarily expect this to be the case given that cosmic structure is not normally distributed.

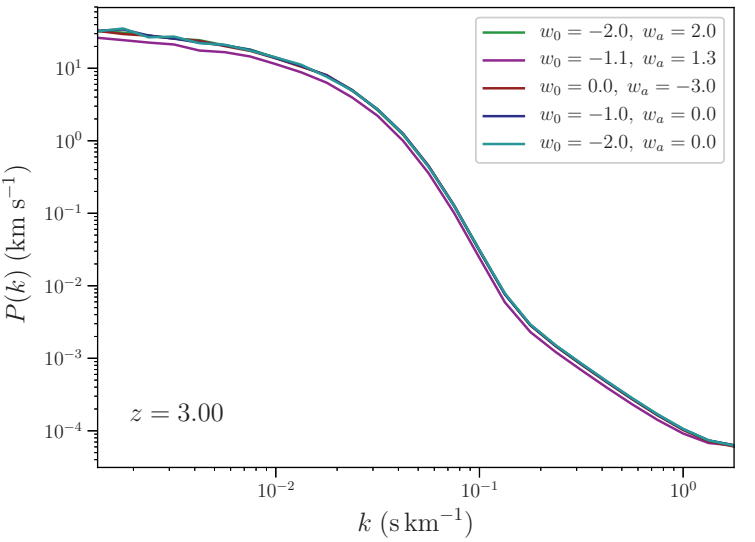



Fig. 2. Power spectrum from each of our dark energy models at $z = 3$. This figure shows that effect of various time-dependent dark energy models on the flux power spectrum of the Lyman-$\alpha$ forest is small.

In summary, his work has explored whether the flux power spectrum of the Lyman-$\alpha$ forest can be used as a probe of time-dependent dark energy. We extracted synthetic Lyman-$\alpha$ forest spectra from high-resolution N-body simulations and used these to calculate the flux power spectrum. We used five different dark energy models, including the cosmological constant and four dynamical, parameterized dark energy models. These models were chosen from the $(w_0, w_a)$ posterior distributions as determined by *Planck*. In particular, of the four dynamical dark energy models we employed, three of them were chosen to lie at the fringes of this posterior, while the fourth was deliberately chosen to lie outside of the 95% C.L. bounds determined by *Planck* to serve as an extreme example from which we might discern an effect on the power spectrum.

Based our pool of synthetic Lyman-$\alpha$ forest spectra, we calculated an average power spectrum for each dark energy model at each redshift. We used k-sample Anderson-Darling test, which shows that there are only marginal differences between the power spectra. Thus, the effects of dark energy on the Lyman-$\alpha$ forest are less significant than the probes based upon the baryonic physics in the IGM. This implies that it is challenging to discriminate among dark-energy models based upon the power spectrum of the Lyman-$\alpha$ forest.

**Acknowledgments**

This work was supported by the U.S. Department of Energy under grant DE-FG02-95-ER40934.

**References**

1. Amendola, L., & Tsujikawa, S. 2010, Dark Energy: Theory and Observations (Cambridge: Cambridge Univ. Press).
2. Caldwell, R. R., Dave, R., & Steinhardt, P. J. 1998, PhRvL, 80, 1582.
3. Armendriz-Picn, C., Mukhanov, V., & Steinhardt, P. J. 2000, PhRvL, 85, 4438.
4. Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, A&A, 594, A14.
5. J. W. Coughlin, G. J. Mathews, L. A. Phillips, A. P. Snedden & I.-S. Suh, Astrophys. J., 874, 11 (1029).
6. Park, D., & Lee, J. 2007, PhRvL, 98, 081301.
7. Bos, E. G. P., van de Weygaert, R., Dolag, K., & Pettorino, V. 2012, MNRAS, 426, 440.
8. Lee, J., & Park, D. 2009, ApJL, 696, L10.
9. Biswas, R., Alizadeh, E., & Wandelt, B. D. 2010, PhRvD, 82, 023002.
10. Lavaux, G., & Wandelt, B. D. 2010, MNRAS, 403, 1392.
11. Lavaux, G., & Wandelt, B. D. 2012, ApJ, 754, 109.
12. Shoji, M., & Lee, J. 2012, arXiv:1203.0869.
13. Cen, R., Miralda-Escud, J., Ostriker, J. P., & Rauch, M. 1994, ApJL, 437, L9.
14. Bi, H., Ge, J., & Fang, L.-Z. 1995, ApJ, 452, 90.
15. Viel, M., Matarrese, S., Theuns, T., Munshi, D., & Wang, Y. 2003, MNRAS, 340, L47.
16. Springel, V. 2005, MNRAS, 364, 1105.
17. Dolag, K., Bartelmann, M., Perrotta, F., et al. 2004, A&A, 416, 853.
18. McDonald, P. 2003, ApJ, 585, 34.
19. Meiksin, A., & White, M. 2001, MNRAS, 324, 141.
20. Bertone, S., & White, S. D. M. 2006, MNRAS, 367, 247.
21. Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016a, A&A, 594, A13.
22. Scoccimarro, R., Hui, L., Manera, M., & Chan, K. C. 2012, PhRvD, 85, 083002.
23. Hui, L., Burles, S., Seljak, U., et al. 2001, ApJ, 552, 15.
24. Anderson, T. W., & Darling, D. A. 1952, Ann. Math. Statist., 23, 193.

# Hubble trouble or Hubble bubble?

Antonio Enea Romano

*Theoretical Physics Department, CERN, CH-1211 Geneva 23, Switzerland*
*Instituto de Fisica, Universidad de Antioquia, A.A.1226, Medellin, Colombia*

The recent analysis of low-redshift supernovae (SN) has increased the apparent tension between the value of $H_0$ estimated from low and high redshift observations such as the cosmic microwave background (CMB) radiation. At the same time other observations have provided evidence of the existence of local radial inhomogeneities extending in different directions up to a redshift of about 0.07. About 40% of the Cepheids used for SN calibration are directly affected because are located along the directions of these inhomogeneities. We compute with different methods the effects of these inhomogeneities on the low-redshift luminosity and angular diameter distance using an exact solution of the Einstein's equations, linear perturbation theory and a low-redshift expansion. We confirm that at low redshift the dominant effect is the non relativist Doppler redshift correction, which is proportional to the volume averaged density contrast and to the comoving distance from the center. We derive a new simple formula relating directly the luminosity distance to the monopole of the density contrast, which does not involve any metric perturbation. We then use it to develop a new inversion method to reconstruct the monopole of the density field from the deviations of the redshift uncorrected observed luminosity distance respect to the $\Lambda CDM$ prediction based on cosmological parameters obtained from large scale observations.

The inversion method confirms the existence of inhomogeneities whose effects were not previously taken into account because the $2M++$ [1] density field maps used to obtain the peculiar velocity [2] for redshift correction were for $z \leq 0.06$, which is not a sufficiently large scale to detect the presence of inhomogeneities extending up to $z = 0.07$. The inhomogeneity does not affect the high redshift luminosity distance because the volume averaged density contrast tends to zero asymptotically, making the value of $H_0^{CMB}$ obtained from CMB observations insensitive to any local structure. The inversion method can provide a unique tool to reconstruct the density field at high redshift where only SN data is available, and in particular to normalize correctly the density field respect to the average large scale density of the Universe.

*Keywords*: Luminosity distance; inversion method; Hubble parameter.

## 1. Introduction

The recent analysis of low-redshift supernovae (SN) luminosity distance measurements [3] have given an estimate of the the Hubble parameter $73.24 \pm 1.74$ km s$^{-1}$ Mpc$^{-1}$ which is more than 9% larger than the one obtained from CMB data [4], $66.93 \pm 0.62$ km s$^{-1}$ Mpc$^{-1}$. The difference is significant at about $3.4\sigma$ confidence level, making it a discrepancy which is definitely worth investigating and cannot be easily attributed to a statistical fluke. On the other hand there is evidence [5] from luminosity density observations that the radial density profile is not homogeneous in certain directions. In one direction, corresponding to subregion 3 in [5], the radial profile is underdense up to about 300 Mpc/h.

In order to check the effects of local inhomogeneities on the luminosity distance and that redshift correction is sufficiently accurate we compute the full relativistic effects on the luminosity distance using an exact solution of the Einstein's equations and compare to results obtained using linear perturbations theory and a low-redshift

expansion. In the perturbative limit we obtain a formula relating directly the density field to the luminosity distance, showing that the effects are proportional to the volume average of the density field, not to the local value of the density contrast. This explains naturally why large scale observations are not affected by low redshift inhomogeneities, since their volume average is negligible on large scales.

We use this formula to derive a new inversion method to obtain the density contrast from redshift uncorrected observed luminosity distance observation.The method gives a value of the density contrast in subregion 3 in agreement with[5] observations.Note that no assumption about a large void extending in all directions is made, and only the radial inhomogeneity in subregion 3 is considered, but since about 40% of the Cepheids are affected by this inhomogeneity, the entire analysis can be affected even if the local Universe is homogeneous in other directions.

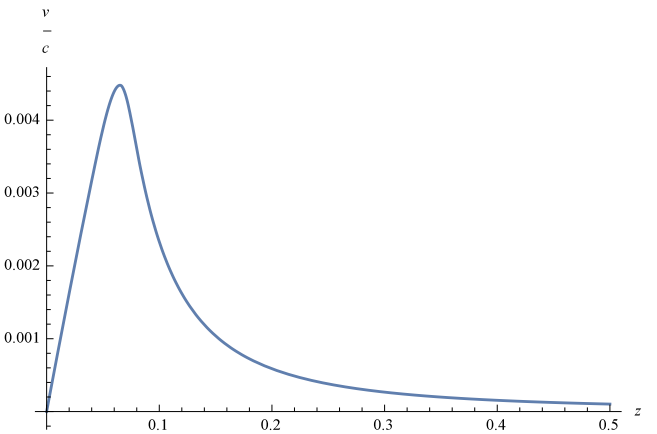



Fig. 1. The peculiar velocity associated to an inhomogeneity profile as the one shown in the upper panel of fig. 2 is plotted in units of the speed of light $c$ as a function of redshift. As can be seen the effect reaches its peak around the hedge of the inhomogeneity and is then asymptotically suppressed due to volume averaging.

Due to the insensitivity of high redshift luminosity distance to local structure, we argue that any local deviation from the theoretical prediction of a FRW model based on cosmological parameters estimated using large scale observations such as the CMB radiation, should be considered an evidence of local structures with a size larger than the depth of the density maps used to apply redshift correction to observational data.

The inversion method we developed could be hence particularly useful to reconstruct the density field from SN luminosity distance observations on scales where galaxy surveys data is not available, not just to resolve the $H_0$ tension.

## 2. Effects of inhomogeneities on the luminosity distance

Unless differently specified, we use a system of units in which $8\pi G = c = 1$. The effects of inhomogeneities on the angular diameter distance were calculated to first

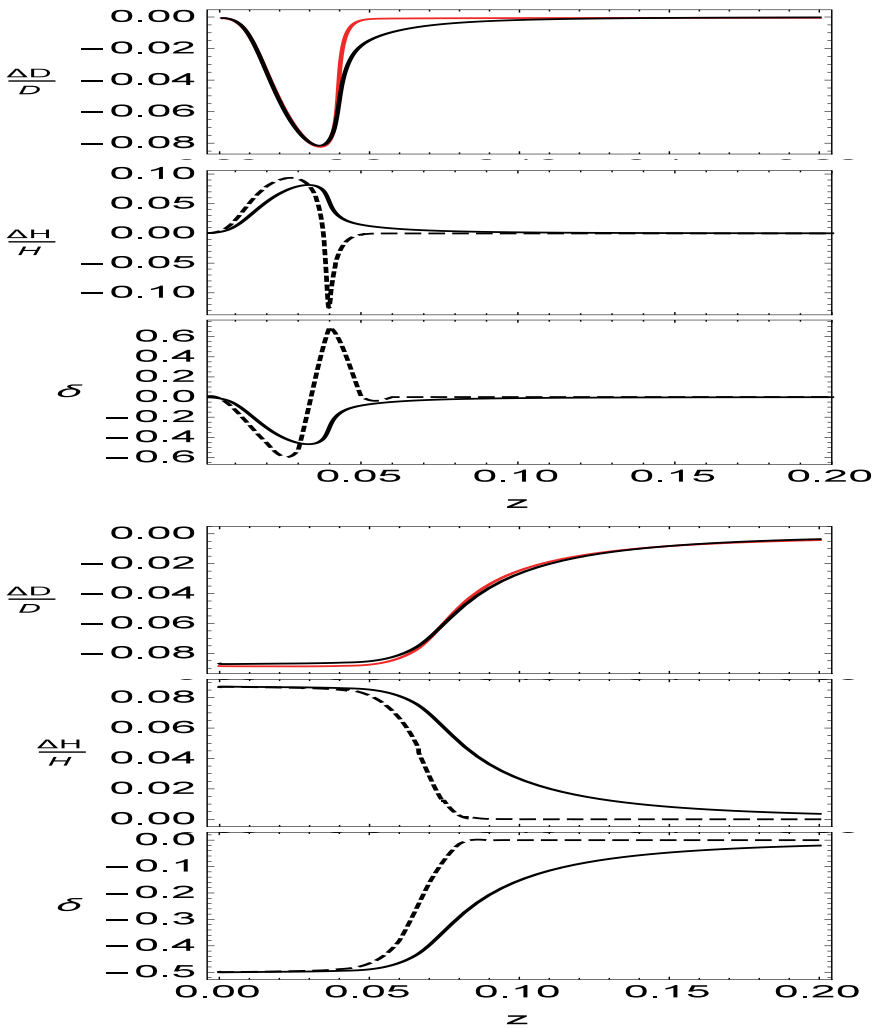



Fig. 2. The fractional difference of the luminosity distance $\Delta D_L/\overline{D_L}$ and the local Hubble parameter $\Delta H/\overline{H_0}$ are plotted as a function of the redshift for a compensated (top) inhomogeneity such as the one studied in[6] and an uncompensated (bottom) void. For $\Delta D/D$ the effects of the inhomogeneity are computed with a non pertubative approach using a LTB metric (red line) and are in good agreement with the approximation (black line) given in eq. (11), confirming that $k_v$ in eq. (9) produces the dominant effect at low redshift. The relative fractional difference $\Delta H/H$ is computed with eq. (14) (black line) and with eq. (12) (dashed line). For the compensated case the difference is particularly important since eq. (12) would predict a negative variation, while eq. (14) gives the correct sign, in agreement with the results of a local fitting procedure.The volume averaged fractional density contrast $\overline{\delta}$ defined in eq. (7) is plotted with a black line and the local density contrast $\delta$ with a dashed line.

order in perturbation theory in[7] and to second order in[8]. The results in the Newton gauge can be written in this form[9,10]

$$D_A(z) = \overline{D}_A(z)\,[1 - k(z)]\ , \tag{1}$$

where $\overline{D}_A(z)$ is the diameter distance for a homogeneous Universe. The convergence $k(z) = \sum_i k_i(z)$ is the sum of different terms among which the most important ones are

$$k(z) \approx k_\delta(z) + k_v(z)\,, \tag{2}$$

$$k_v = \left[1 - \frac{1}{a_e \chi_e H_e}\right] v_e \cdot n + \frac{1}{a_e \chi_e H_e} v_o \cdot n\,, \tag{3}$$

$$k_\delta = \frac{3}{2} H_0^2 \Omega_m \int_0^{\chi_e} \mathrm{d}\chi \frac{(\chi_e - \chi)}{a_e \chi_e} \chi \delta(\chi)\,, \tag{4}$$

where $\chi$ is the comoving distance, we are denoting respectively with a lower-script $e$ or $o$ any quantity evaluated at the point of emission of the photons or at the observer, and the unit vector $n$ is in the direction of propagation from the emitter to the observer. The term $k_\delta$ is known as the gravitational lensing magnification [11] while $k_v$ is due to the peculiar velocities. The other $k_i$ terms are related to the line of sight integral of the gravitational potential and its time derivatives, such as the integrated Sachs-Wolf effect for example, and are sub-dominant at low-redshift [9]. We will show later using both an exact solution of the Einstein's equations and a low-redshift Taylor expansion that the most important term is $k_v$.

## 3. Peculiar velocity and density maps

In the Newtonian limit the peculiar velocity field can be related to the density field by integrating the Euler's equation [12]

$$v(\chi) = \frac{afH}{4\pi} \int^{R_{Max}} \delta(\chi') \frac{\chi' - \chi}{|\chi' - \chi|^3} d^3\chi'\,. \tag{5}$$

In [3,13,14] $R_{Max}$ corresponds to $z < 0.06$, which is less than the size of the local inhomogeneity, $z \approx 0.07$. While the redshift correction method is quite precise at low redshift as we will show later, the velocity field they obtain is missing the peculiar velocity component due to the local void which can be obtained only by integrating eq. (5) over scales larger than the size of the void. The correct background density value $\overline{\rho}$ entering the definition of the fractional density contrast $\delta = \delta\rho/\overline{\rho}$ should be the volume average of the density $\rho$ on a scale larger than the void size, otherwise the underdensity of the void respect to the rest of the Universe will not be taken into account.

Density maps for $z < 0.06$ cover regions inside the void and cannot be used to find the relative density difference respect to the outside region, located at $z > 0.07$. In other words if we only consider density maps for $z < 0.06$ we cannot detect the presence of the local void, while extending the analysis to a higher redshift range [5] gives $\delta \approx -0.5$ inside the void. It turns out that the component of the peculiar velocity due to the void is crucial to explain the apparent tension in the $H_0$ estimation.

## 4. Perturbative monopole correction

The peculiar velocity associated to a spherically symmetric inhomogeneity can be obtained in the Newtonian limit as[12]

$$v(\chi) = -\frac{1}{3} a f H \overline{\delta}(\chi) \chi \,, \tag{6}$$

$$\overline{\delta}(\chi) = \frac{3}{4\pi\chi^3} \int^{\chi} 4\pi \chi'^2 \delta(\chi')\, d\chi' \tag{7}$$

where $\overline{\delta}$ is the density contrast averaged over the sphere of comoving radius $\chi$, it has been assumed that the density contrast can be factorized as the product of a space and a time dependent function as $\delta = A(x)D(t)$, and $f = \frac{1}{H}\frac{\dot{D}}{D}$ is the growth factor. The negative sign in front of eq. (6) implies that for an underdensity the velocity is directed outward from the center, which is what we intuitively would expect since the region outside the void is denser. From the above equation we can see that the monopole component of the peculiar velocity is zero for a central observer, implying that the observer velocity is not affected by the monopole component of the local structure. Since the effects of the inhomogeneity depends on $\overline{\delta}(z)$, they extend slightly beyond the edge of the void, because the volume averaged density contrast does not go to zero immediately after it.

Eq. (6) is used to plot in fig. 1 the peculiar velocity as function of the red-shift inside an inhomogeneity with a density profile given in fig. 2. As expected the velocity is zero at the center, reaches its pick at the edge of the inhomogeneity and tends asymptotically to zero due to the volume average.

Using eq. (6), and only considering the effects due to the emitter velocity $v_e$, since, as shown in eq. (6), a spherical symmetric inhomogeneity does not affect the observer velocity, we can re-write $k_v$ in terms of the averaged density density and get

$$k_v = \frac{1}{3} f \overline{\delta}\, (aH\chi - 1) \,. \tag{8}$$

Note that we have used that $v_e \cdot n = \frac{1}{3} f \overline{\delta} a H \chi$ because $n$ is directed from the emitter to the observer, while in eq. (6) the unit vector is in the opposite direction, i.e the positive radial direction, since we are assuming a coordinate system centered at the observer.

## 5. Low redshift effects of a local inhomogeneity

At low redshift we have $aH\chi \approx z$ and consequently only the second term in eq. (8) is important at leading order in $z$, giving

$$k_v(z) = -\frac{1}{3} f \overline{\delta}(z) \,. \tag{9}$$

In the case of $k_\delta$ we can also perform a low-redshift expansion and assuming $\overline{\delta} = \delta_c + \delta_1 z + ..$, the leading order term is

$$k_\delta = \frac{3}{4} \Omega_m \delta_c z^2 \,, \tag{10}$$

which is a second order term which can be safely neglected at low redshift.

Now that we know that the dominant contribution at low redshift is $k_v$, we can compute the leading order correction to the angular diameter distance

$$D_A(z) = D_L(z) = \overline{D_L}(z)\left[1 + \frac{1}{3}f\overline{\delta}(z)\right] . \tag{11}$$

In the above equation we have used that at leading order in redshift the reciprocity relation $D_L(z) = (1+z)^2 D_A(z)$ implies that $D_L(z) = D_A(z)$.

As shown in fig. 2 eq. (11) is in very good agreement with the results obtained from the calculation of the full relativistic effects of the inhomogeneity using the LTB solution, confirming that the approximations used to derive eq. (11) are well justified.

$$\frac{\Delta H}{\overline{H}} = -\frac{1}{3}f\delta\,. \tag{12}$$

$$H_0^{loc} = \overline{H_0} + \Delta H(z)\,, \quad \overline{H_0} = \lim_{z\to 0}\frac{z}{\overline{D}_L(z)} \tag{13}$$

$$\frac{\Delta H(z)}{\overline{H_0}} = -\frac{\Delta D_L(z)}{\overline{D}_L(z)} = -\frac{1}{3}f\overline{\delta}(z) \tag{14}$$

## Acknowledgments

I thank Misao Sasaki, George P. Efstathiou, Adam Riess and Malcolm Fairbairn for useful discussions and comments.

## References

1. G. Lavaux and M. J. Hudson, *Mon. Not. Roy. Astron. Soc.* **416** (October 2011) 2840, arXiv:1105.6107.
2. J. Carrick, S. J. Turnbull, G. Lavaux and M. J. Hudson, *Mon. Not. Roy. Astron. Soc.* **450** (2015) 317, arXiv:1504.04627 [astro-ph.CO].
3. A. G. Riess *et al.*, *Astrophys. J.* **826** (2016) 56, arXiv:1604.01424 [astro-ph.CO].
4. Planck Collaboration (N. Aghanim *et al.*) (2016) arXiv:1605.02985 [astro-ph.CO].
5. R. C. Keenan, A. J. Barger and L. L. Cowie, *Astrophys. J.* **775** (2013) 62, arXiv:1304.2884 [astro-ph.CO].
6. A. E. Romano and S. A. Vallejo, *Europhys. Lett.* **109** (2015) 39002, arXiv:1403.2034 [astro-ph.CO].
7. M. Sasaki, *Mon. Not. Roy. Astron. Soc.* **228** (October 1987) 653.
8. E. Barausse, S. Matarrese and A. Riotto, *Phys. Rev.* **D71** (2005) 063537, arXiv:astro-ph/0501152 [astro-ph].
9. C. Bonvin, R. Durrer and M. A. Gasparini, *Phys. Rev.* **D73** (2006) 023523, arXiv:astro-ph/0511183 [astro-ph], [Erratum: Phys. Rev.D85,029901(2012)].
10. L. Hui and P. B. Greene, *Phys. Rev.* **D73** (2006) 123526, arXiv:astro-ph/0512159 [astro-ph].
11. N. Kaiser, *APJ* **388** (April 1992) 272.
12. P. J. E. Peebles, *Principles of physical cosmology* (Princeton University Press, 1993).
13. D. Scolnic *et al.*, *Astrophys. J.* **795** (2014) 45, arXiv:1310.3824 [astro-ph.CO].
14. SNLS Collaboration (J. D. Neill, M. J. Hudson and A. J. Conley), *Astrophys. J.* **661** (2007) L123, arXiv:0704.1654 [astro-ph].

# Relativistic corrections to large scale structures

Lina Castiblanco, Radouane Gannouji, Jorge Noreña and Clément Stahl

*Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4950, Valparaíso, Chile*
*clement.stahl@pucv.cl*

We investigate the relativistic corrections to the standard model of formation of large scale structures. In matter domination and in the Poisson gauge, we use the weak-field approximation which allows to keep compact expressions even for the one-loop bispectrum. Whereas in the Newtonian limit, the choice of gauge is marginally important as all gauge coincides, when relativistic corrections are taken into account, it matters as a change of gauge may induce a change of gravitational potential and introduce fictitious modes in the finial result for the power spectrum. It is precisely what happens in the example presented in this talk as the equivalence principle is not fulfilled in the Poisson gauge and the cancellation of the IR divergence at one-loop does not occur. We will discuss how other choices of gauge may solve this issue.

*Keywords*: Cosmology, Large Scale Structures, Relativistic corrections, Bispectrum, Non-gaussianities.

## 1. Introduction

Most of the large scale structures observables can be computed within a Newtonian approximation (see Ref. 1 for a review). On the large scales, at early enough time, the universe is in the linear regime and all gauges coincide with the Newtonian one. On the small scales, local processes for the formation of structures happen and gravity is well described by Newton's theory. However, reporting on Ref. 2, we argue that a proper calculation of the one-loop bispectrum requires to take into account relativistic corrections. Unlike the power spectrum which depends only on one Fourier mode by translation and rotation invariances, the bispectrum couples the scales[a] and therefore could get non-linear relativistic corrections that requires to go beyond the Newtonian gauge. We will present some steps toward calculating the one-loop bispectrum. In section 2, we present a weak field scheme which allows to keep the leading relevant relativistic corrections for cosmological spacetimes. In section 3, we describe the equations of motions for the dark matter perturbations and we solve them using perturbation theory. We present our results for the bispectrum in section 4 and finally present some further subtleties of this calculation and conclude in section 5.

## 2. Weak field approximation

The task of computing relativistic corrections in full GR is quite challenging even at second order, see *eg.* Refs. 3–5. However, on cosmological scales, gravity is

[a]It is possible to consider triangle configurations that have one large scale and two small scales: the so-called squeezed limit.

weak and to obtain reliable results, one does not need to consider the full GR. Mathematically, the weak field approximation[6–8] relies on the fact that the root mean square of the velocity of dark matter perturbations is small: $v \sim 10^{-3}$ and furthermore the gravitational potential is also small $\phi \sim 10^{-5}$. This allows to tread those two quantities perturbatively in the equation of motion. We note however that to derive the equations of motion of section 3, nothing is assumed regarding the density contrast of the dark matter field: $\delta \sim 1$ can happen in the small scales, at shell-crossing. A more systematic description of the weak field approximation can be found in table 1 of Ref. 2 and in the text around it.

## 3. Equations of motion

The metric we consider reads:

$$ds^2 = -(1+2\phi)dt^2 + 2\omega_i dx^i dt + a(t)^2\left[(1-2\psi)\delta_{ij} + h_{ij}\right]dx^i dx^j\,, \tag{1}$$

where we considered two gravitational potentials $\phi$ and $\psi$, frame dragging effects $\omega_i$ and gravitational waves $h_{ij}$. Applying the weak field scheme of section 2 and calculating the first two moments of the Boltzmann equation, supplemented by the Einstein equations, one finds[2]:

$$\dot{\delta}_R + \theta_R = -\int_{\boldsymbol{k}_1,\boldsymbol{k}_2} \alpha(\boldsymbol{k}_1,\boldsymbol{k}_2)(\theta_R(\boldsymbol{k}_1)\delta_N(\boldsymbol{k}_2) + \theta_N(\boldsymbol{k}_1)\delta_R(\boldsymbol{k}_2)) + \mathcal{S}_\delta[\delta_N]\,, \tag{2}$$

$$\dot{\theta}_R + 2H\theta_R + \frac{3}{2}H^2\delta_R = -2\int_{\boldsymbol{k}_1,\boldsymbol{k}_2} \beta(\boldsymbol{k}_1,\boldsymbol{k}_2)\theta_N(\boldsymbol{k}_1)\theta_R(\boldsymbol{k}_2) + \mathcal{S}_\theta[\delta_N]\,. \tag{3}$$

We have split the dynamical quantities ($\delta$, $\theta \equiv \partial_i v^i$) between a Newtonian part and a relativistic part, denoted by superscripts $N$ and $R$. The left hand sides of these equations describe the linear terms, an expression for the relativistic sources is given in appendix C of Ref. 2. The standard kernels $\alpha$ and $\beta$ describes usual non-linearities terms.

To solve Eqs. (2)-(3), we perform perturbation theory in the usual sense (see *eg.* Ref. 9) and expand the dark matter fields in powers of the linear density contrast $\delta_l$. This allows to define Newtonian kernels $F_n$ together with their relativistic counterparts $F_n^R$:

$$\delta(\boldsymbol{k},t) = \sum_{n=1}^{\infty} a^n(t) \int_{\boldsymbol{k}_1\ldots\boldsymbol{k}_n} \Big[F_n(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n) + a^2(t)H^2(t)F_n^R(\boldsymbol{k}_1,\ldots,\boldsymbol{k}_n)\Big]\delta_l(\boldsymbol{k}_1)\ldots\delta_l(\boldsymbol{k}_n)\,. \tag{4}$$

In appendix D of Ref. 2, expressions for the kernel up to $n = 4$ are presented in two gauges of interest: Poisson and synchronous-comoving.

## 4. Results

After solving the equations of motion, we are in position to present our main results: a plot of the bispectrum defined as

$$\langle\delta(\boldsymbol{k}_1,t)\delta(\boldsymbol{k}_2,t)\delta(\boldsymbol{k}_3,t)\rangle = (2\pi)^3\delta_D(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B(k_1,k_2,k_3,t)\,. \tag{5}$$

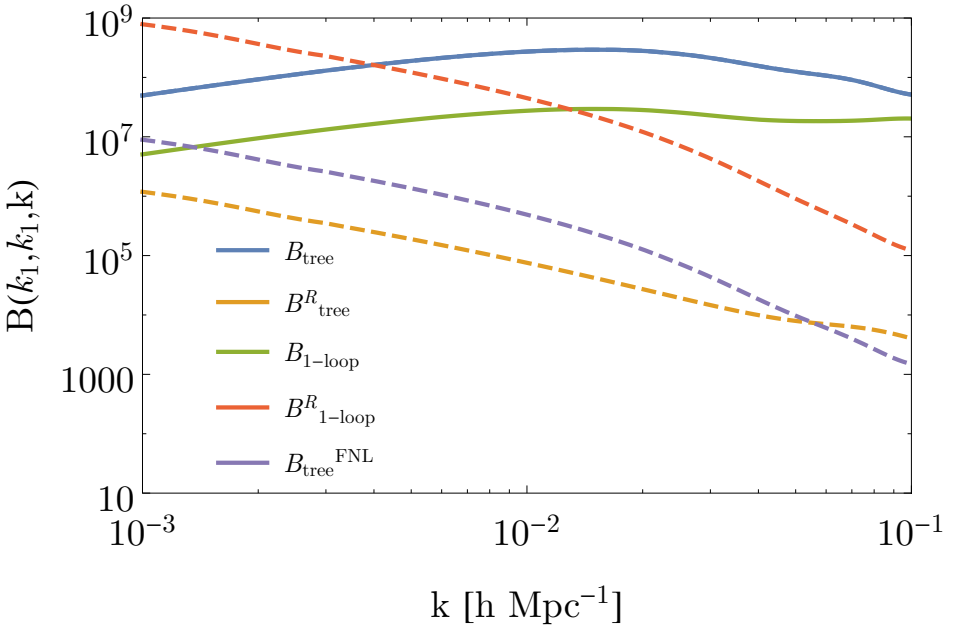



Fig. 1. In the squeezed configuration: $B(k_1, k_1, k)$, one-loop corrections to the linear bispectrum in Poisson gauge, with $k_1 = 0.1$ h Mpc$^{-1}$. Doted lines denotes a negative contribution. The qualitative reasoning discussed in the introduction holds and indeed relativistic corrections are important at large scales (small $k$) in squeezed configurations. Observe even that the relativistic one-loop results become larger than the tree-level Newtonian results and the perturbation theory breaks down in the IR. In purple, we ploted a signal that could come from primordial non-gaussianities with $f_{NL} = 1$.

Those results are contrasted with a synchronous gauge in Ref. 2.

## 5. Conclusions and perpectives

Due to lack of space, several issues were eluded in this proceeding, in particular, in section 5 of Ref. 2, some details are given about the renormalization of the background which is required in order to avoid infinities, the UV and IR behavior of the loop-integrals is also discussed there.

We have reported in this proceeding our new results on the one-loop bispectrum. We argued that taking into account GR is necessary as the bispectrum couples scales. We indeed found that the relativistic corrections becomes of the same order of magnitude than the Newtonian (linear) results.

These encouraging results invite to perform several steps in order to observe the bispectrum:

- Take into account, not only dark matter but galaxies which is done by introducing a set of bias parameters [10,11].

- Take into account the fact that what is observed in our telescopes is a redshift and a direction in the sky and not a position and a time. To do so, one goes to redshift space and takes into account the propagation of photons in our cuspy universe: those effects are sometimes dubbed *redshift space distortions* [12,13].

Other interesting roads includes to more precisely model th short scales, either with effective field theory [14], resummation [15] or renormalization group [16] techniques. The inclusion of a cosmological constant [17] $\Lambda$ is also in order opening also the door to further generalize to modified gravity scenario for cosmology.

## Acknowledgments

It is a pleasure to thank all the participants and speakers of this session for stimulating scientific discussions ans interaction during the meeting and during the nights. I am also very grateful to all the people who organized the Marcel Grossmann meeting: professors, staffs, secretaries, students....

## References

1. L. Castiblanco, R. Gannouji and C. Stahl, arXiv:1910.03931 [gr-qc].
2. L. Castiblanco, R. Gannouji, J. Noreña and C. Stahl, JCAP **1907** (2019) no.07, 030 [arXiv:1811.05452 [astro-ph.CO]].
3. S. Matarrese, S. Mollerach and M. Bruni, Phys. Rev. D **58** (1998) 043504 [astro-ph/9707278].
4. L. Boubekeur, P. Creminelli, J. Noreña and F. Vernizzi, JCAP **0808** (2008) 028 [arXiv:0806.1016 [astro-ph]].
5. A. L. Fitzpatrick, L. Senatore and M. Zaldarriaga, [arXiv:0902.2814 [astro-ph.CO]].
6. S. R. Green and R. M. Wald, Phys. Rev. D **83** (2011) 084020 [arXiv:1011.4920 [gr-qc]].
7. J. Adamek, D. Daverio, R. Durrer and M. Kunz, Nature Phys. **12** (2016) 346 [arXiv:1509.01699 [astro-ph.CO]].
8. J. Adamek, D. Daverio, R. Durrer and M. Kunz, JCAP **1607** (2016) no.07, 053 [arXiv:1604.06065 [astro-ph.CO]].
9. M. Simonović, T. Baldauf, M. Zaldarriaga, J. J. Carrasco and J. A. Kollmeier, JCAP **1804** (2018) no.04, 030 [arXiv:1708.08130 [astro-ph.CO]].
10. V. Desjacques, D. Jeong and F. Schmidt, Phys. Rept. **733** (2018) 1 [arXiv:1611.09787 [astro-ph.CO]].
11. O. Umeh, K. Koyama, R. Maartens, F. Schmidt and C. Clarkson, JCAP **1905** (2019) 020 [arXiv:1901.07460 [astro-ph.CO]].
12. C. Bonvin and R. Durrer, Phys. Rev. D **84** (2011) 063505 [arXiv:1105.5280 [astro-ph.CO]].

13. E. Di Dio, R. Durrer, R. Maartens, F. Montanari and O. Umeh, JCAP **1904** (2019) 053 [arXiv:1812.09297 [astro-ph.CO]].
14. J. J. M. Carrasco, M. P. Hertzberg and L. Senatore, JHEP **1209** (2012) 082 [arXiv:1206.2926 [astro-ph.CO]].
15. L. Senatore and M. Zaldarriaga, JCAP **1502** (2015) no.02, 013 [arXiv:1404.5954 [astro-ph.CO]].
16. S. Floerchinger, M. Garny, N. Tetradis and U. A. Wiedemann, JCAP **1701** (2017) no.01, 048 [arXiv:1607.03453 [astro-ph.CO]]; A. Erschfeld and S. Floerchinger, JCAP **1906** (2019) no.06, 039 [arXiv:1812.06891 [astro-ph.CO]].
17. E. Villa and C. Rampf, JCAP **1601** (2016) no.01, 030 Erratum: [JCAP **1805** (2018) no.05, E01] [arXiv:1505.04782 [gr-qc]].

# Self interacting dark matter and small scale structure

N. Q. Lan, G. J. Mathews, J. A. Coughlin and I.-S. Suh

*Center for Astrophysics, Department of Physics, University of Notre Dame*
*Notre Dame, IN 46556, USA*
**E-mail: gmathews@nd.edu*
*https://physics.nd.edu/people/faculty/grant-j-mathews*

The core-cusp problem remains as one of the unresolved challenges between observation and simulations in the standard ΛCDM model for the formation of galaxies. Basically, the problem is that ΛCDM simulations predict that the center of galactic dark matter halos contain a steep power-law mass density profile. However, observations of dwarf galaxies in the Local Group reveal a density profile consistent with a nearly flat distribution of dark matter near the center. A number of solutions to this dilemma have been proposed. We summarize investigations the possibility that the dark matter particles themselves self interact and scatter. The scattering of dark matter particles then can smooth out their profile in high-density regions. We also summarize theoretical theoretical models as to how self- interacting dark matter may arise. We summarize our own implementation this form in simulations of self-interacting dark matter in models for galaxy formation and evolution. Constraints on self-interacting dark matter are then summarized.

*Keywords*: Interacting Dark Matter; Dwarf Galaxies; Galactic Structure.

## 1. Introduction

The nature of most of the matter in the universe remains as one of the most challenging questions in modern physics.[1] Nevertheless, observational evidence from the Cosmic Microwave Background, galaxies cluster, weakly lensing, and the Lyman-$\alpha$ forest agree with the predictions of the ΛCDM model. Such models contain a mixture of roughly 25% collisionless cold dark matter such as WIMPs, axions, etc. interacting through the weak and gravitational forces only, and 70% vacuum energy or quintessence.[2,3] Only a fraction of the present total matter can be made of ordinary baryons has an unknown, nonbaryonic origin[4].

However, it is now appreciated (e.g. Ref. [5] and refs. therein) that conventional models of collisionless cold dark matter lead to problems with regard to galactic structure. They are only able to fit the observations on large scales ($\gg 1Mpc$). In particular, high-resolution $N$-body simulations in these models result in a central singularity of the galactic halos[6] and a large number of sub-halos[7] than observed. This is called the *Core-Cusp Problem*. A number of other inconsistencies are discussed in Refs. [5, 8, 9, 10]. In particular, the mass density profile for CDM halos increases toward the center, scaling approximately as $\rho_{\rm dm} \sim r^{-1}$. However, many observed rotation curves of disk and dwarf spheroidal galaxies prefer[11–13] a constant density profile $\rho_{\rm dm} \sim r^0$, as evidenced by a linearly rising circular velocity in the inner regions. This is most evident in dwarf and low surface brightness (LSB) galaxies since they are highly DM-dominated.

A possible way to avoid these problems is to hypothesize *self-interacting dark matter*[14] (SIDM). [Although self-interacting models lead to spherical halo centers in clusters. This is not in agreement with the inferred ellipsoidal centers indicated

by gravitational lensing[15] and by Chandra observations[16].]

Self-interacting dark matter models are well motivated as a model for particle dark matter. The key property of dark-matter particles is that they are non-relativistic and have have a weak scattering cross-section. The Spergel-Steinhard model[14] has motivated many follow-up studies[4,17,18] to identify a dark matter particle with self interactions. Here we constrain one particular model[19] with right-handed (RH) neutrinos for self-interacting dark matter based upon simulations of galactic structure. We deduce that the mass of the SIDM in this model in about an MeV. We compare numerical simulations to the observations.

The range of mass for SIDM is from ∼MeV with a mean free path and total cross section over mass from 0.1 cm$^2$ g$^-$1 to 100 cm$^2$ g$^{-1}$ can solve the core-cusp and the missing satellite problems of the ΛCDM model. In our model the cross section over mass $\sigma/m$ is in the range of 4 to 5 cm$^2$ g$^{-1}$. In large scale structure there is no difference between that of normal CDM and SIDM. However, on small scales galactic cores are consistent with all of the observational constraints. We analyze the small scale structure of the dark matter based upon the hydrodynamic simulations described in this work.

## 2. Model for self-interaction dark matter

The SM offers no options for dark matter. The first gauge model[20] for SIDM was found in the 3-3-1 model. To keep the Higgs sector with three triplets in that model the existence of *exotic leptons* was proposed. The 3-3-1 models were proposed with an independent motivation[21–23]. These models have the intriguing features that they are anomaly free only if the number of families $N$ is a multiple of three. If one adds the condition of QCD asymptotic freedom, which is valid only if the number of families of quarks is to be less than five, it follows that $N$ is equal to 3.

I this work we summarize the argument that the 3-3-1 model with right-handed (RH) neutrinos[24] contains such self-interacting dark matter.

The main properties that a good dark matter candidate must satisfy are stability and neutrality. Therefore, one should consider the scalar sector of the model, more specifically, the neutral scalars.[14]. In addition, such dark matter particles must not overpopulate the Universe. On the other hand, since our dark matter particle is not imposed arbitrarily to solve this specific problem, one must check that the necessary values of the parameters do not spoil the other bounds of the model.

Under the assumption of discrete symmetry $\chi \to -\chi$, the most general potential can then be written in the following form,[25]

$$\begin{aligned} V(\eta,\rho,\chi) = {} & \mu_1^2\eta^+\eta + \mu_2^2\rho^+\rho + \mu_3^2\chi^+\chi + \lambda_1(\eta^+\eta)^2 + \lambda_2(\rho^+\rho)^2 + \lambda_3(\chi^+\chi)^2 \\ & +(\eta^+\eta)[\lambda_4(\rho^+\rho) + \lambda_5(\chi^+\chi)] + \lambda_6(\rho^+\rho)(\chi^+\chi) + \lambda_7(\rho^+\eta)(\eta^+\rho) \\ & +\lambda_8(\chi^+\eta)(\eta^+\chi) + \lambda_9(\rho^+\chi)(\chi^+\rho) + \lambda_{10}(\chi^+\eta + \eta^+\chi)^2. \end{aligned} \quad (1)$$

Next one can write the expansion of the scalar fields which acquire a VEV:

$$\eta^o = \frac{1}{\sqrt{2}}\,(v + \xi_\eta + i\zeta_\eta)\,;\ \rho^o = \frac{1}{\sqrt{2}}\,(u + \xi_\rho + i\zeta_\rho)\,;\ \chi^o = \frac{1}{\sqrt{2}}\,(w + \xi_\chi + i\zeta_\chi)\,. \quad (2)$$

For the prime neutral fields which do not have a VEV, one finds analogously:

$$\eta'^{o} = \frac{1}{\sqrt{2}}\left(\xi'_{\eta} + i\zeta'_{\eta}\right);\ \chi'^{o} = \frac{1}{\sqrt{2}}\left(\xi'_{\chi} + i\zeta'_{\chi}\right). \tag{3}$$

To satisfy the requirements of stability and neutrality, one can go to the scalar sector of the model, more specifically to the neutral scalars. One must then whether any of them can be stable and are self-interacting. [14] In addition, such a dark matter particle must freeze out at the correct cosmic abundance. One can check through a direct calculation by employing the relevant Lagrangians that the Higgs scalar $h_0$ and $H_3^0$ can, satisfy these criteriona. Remarkably, they do not interact directly with any SM field except for the standard Higgs $H_1^0$. However, $h_0$ must be favored, since it is easier to obtain a large scattering cross section.

In contrast to the singlet models of Refs. 4, 18, 28, 29, 26, 27, where an extra symmetry must be imposed to account the stability of the dark matter, the decay of the $h_0$ scalar is automatically forbidden in all orders of perturbative expansion. This is because this scalar comes from the triplet $\chi$, that induces the spontaneous symmetry breaking of the 3-3-1 model to the standard model. Therefore, the SM fermions and the standard gauge bosons cannot couple with $h^0$. Also, the $h^0$ scalar comes from the imaginary part of the Higgs triplet $\chi$.The imaginary parts of $\eta$ and $\rho$ are pure massless Goldstone bosons. Thus, there are no physical scalar fields which can mix with $h^0$ and the only interactions of $h^0$ come from the scalar potential. They are $H_3^0 h^0 h^0$ and $H_1^0 h^0 h^0$. If $v \sim u \sim (100 \sim 200)$ GeV and $-1 \leqslant a_5 \sim a_6 \leqslant 1$, the $h^0$ can interact only weakly with ordinary matter through the Higgs boson of the standard model $H_1^0$. The relevant quartic interaction for scattering is thus $h^0 h^0 h^0 h^0$.

## 3. Structure formation in a SIDM Universe

In the work of Lan and Long, [19] it was deduced that the mass range for the dark matter is from 4.7MeV to 29 MeV. Our dark matter is non-relativistic with decoupling temperature $\sim$ 1eV. Dark matter does not interact with any particles in the standard model except Higgs Boson, so one need not deal with any collision terms. The self-interacting dark matter in the 3-3-1 model of Ref. [19] is non-relativistic driving the decoupling era with a decoupling temperature about 1eV. So we have $g_{\mu\nu}P^{\mu}P^{\nu} = m^2$, where m is the mass of the dark matter particles, $g_{\mu\nu}$ is the full metric tensor, and $P^0 = \sqrt{q^2 + m^2 a^2}(1 - \phi)$. We can define the energy $q = \sqrt{(p^2 + m^2)}$, where $p$ the is magnitude of momentum.

If one defines $\beta = n_h/T^3$ in the radiation dominated era one write the evolution Boltzmann Equation as

$$\frac{d\beta}{dT} = -\frac{\Gamma\beta}{KT^3} = -\frac{\alpha}{8\pi^3 K e^{m_1/T}}\left(\frac{\Theta}{T^2}\right)^2 \tag{4}$$

where $K^2 = 4\pi^3 g(T)/45m_{Pl}^2$ and $\beta = \frac{n_h}{T^3}$ are parameters of thermal equilibrium and $m_{PL}$ is the Planck mass. We take $T = m_1$ so that the cosmic density of the $h_0$

scalar is

$$\Omega_h = 2g(T_\Gamma)T_\gamma^3 \frac{m_h \beta}{\rho_c g(T)} \tag{5}$$

where $T_\gamma = 2.4 \times 10^{-4}$ eV is the present photon temperature, $g(T_\gamma) = 2$ is the photon degree of freedom and $\rho_c = 7.5 \times 10^{-47} h^2$, with h = 0.71, being the critical density of the Universe. Let us take $m_h = 7.75$ MeV, v = 174 GeV, $a_5 = 0.65$, $-a_6 = 0.38$ (actually in our calculations, we have used a better precision for $a_5$ and $a_6$) and $m_1 = 150$ GeV. Thus, from we obtain $\Omega_h = 0.3$. Therefore, without imposing any new fields or symmetries, the 3-3-1 model possesses a scalar field that can satisfy all the properties required for the self-interacting dark matter and does not overpopulate the Universe.

## 4. Constraint with CMB and galaxy simulation

We have simulated of the formation of a dwarf galaxy of $10^{10} M_\odot$ halos of self interacting dark matter (SIDM) with cross section over mass in the range $\sigma/m = 3.7$ to 5.2 cm$^2$ g$^{-1}$. We used a cosmological model with parameters $\Omega_\Lambda = 0.734$, $\Omega_m = 0.266, \Omega_b = 0.0449, n_s = 0.963, h = 0.71, \sigma_8 = 0.801$ (Komatsu et al. 2011). We start with isolated halo galaxies with a stellar mass of $M_{star} = 1.4 \times 10^{11} M_\odot$ and temperature $T = 10^4$K with in a box size of $50 Mpch^{-1}$.

Figure 1 shows an example of our inferred SIDM radial profile for a best fit of $\sigma/m = 4$ cm$^2$ g$^{-1}$. It shows a flattening toward the center as required by observation.

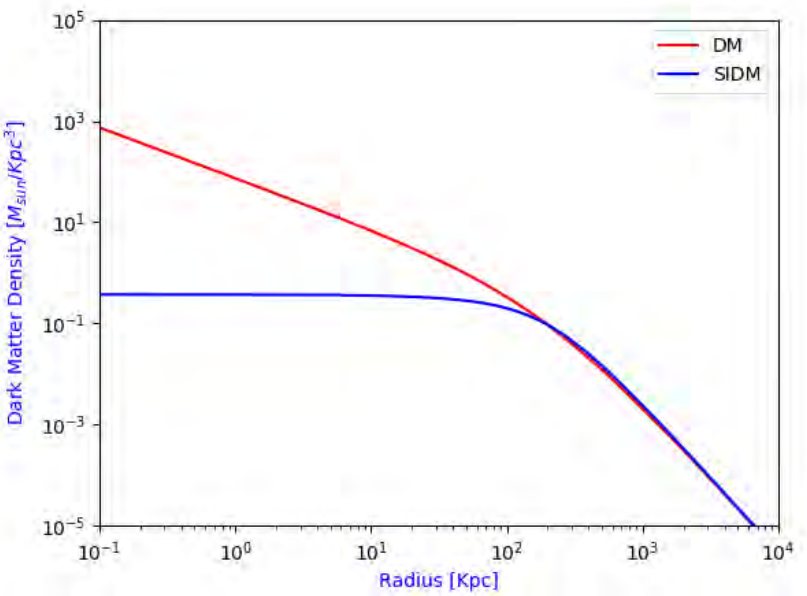



Fig. 1. Galaxy profile calculated with (blue line) and without (red line) SIDM from a fit to the numerical simulations.

**Acknowledgment**

This work was supported by the U.S. Department of Energy under grant DE-FG02-95-ER40934.

**References**

1. M. S. Turner, Phys. Rep. 333-334 (2000) 619.
2. N. A. Bahcall *et al.*, Science 284 (1999) 1481.
3. L. Wang *et al.*, Astrophys. J. 530 (2000) 17.
4. C. P. Burgess *et al.*, Nucl. Phys. B 619 (2001) 709
5. S. Tulin and H.-B. Yu, Phys. Rep., 730, 18 (2017).
6. S. Ghigna *et al.*, Astrophys. J. 544 (2000) 616.
7. B. Moore *et al.*, Astrophys. J. 524 (1999) L19.
8. R. Davé *et al.*, Astrophys. J., 547 (2001) 574.
9. J. S. Bullock *et al.*, Astrophys. J. 539 (2000) 517.
10. G. J. Mathews, A. Snedden, L. A. Phillips, I.-S. Suh, J. Coughlin, A. Bhattacharya, X. Zhao, and N. Q. Lan, Mod. Phys. Lett., A29, 1430012 (2014).
11. B. Moore, T. R. Quinn, F. Governato, J. Stadel, and G. Lake, Mon. Not. R. Astron. Soc. 310 (1999) 1147.
12. R. A. Flores and J. R. Primack, Astrophys. J. 427, L1 (1994), astro-ph/9402004.
13. B. Moore, Nature 370, 629 (1994).
14. D. N. Spergel and P. J. Steinhardt, Phys. Rev. Lett. 84 (2000) 3760.
15. N. Yoshida *et al.*, Astrophys. J. 544 (2000) L87.
16. D. A. Buote *et al.*, Astrophys. J. 577 (2002) 183.
17. J. P. Ostriker, Phys. Rev. Lett. 84 (2000) 5258.
18. J. McDonald, Phys. Rev. Lett. 88 (2002) 091304.
19. Lan, N. Q and Long, H. N, Astrophys. Space Sci. 305 (2006) 225-233.
20. D. Fregolente and M. D. Tonasse, Phys. Lett. B555 (2003) 7.
21. F. Pisano and V. Pleitez, Phys. Rev. D 46, (1992) 410.
22. P. H. Frampton, Phys. Rev. Lett. 69, (1992) 2889.
23. R. Foot *et al.,* Phys. Rev. D 47, (1993) 4158.
24. R. Foot, H. N. Long, and Tuan A. Tran, Phys. Rev. D 50, (1994) R34.
25. H. N. Long, Mod. Phys. Lett. A 13 (1998) 1865.
26. D. E. Holz and A. Zee, Phys. Lett. B 517 (2001) 239.
27. V. Silveira and A. Zee, Phys. Lett. B 161 (1985) 136.
28. M. C. Bento *et al.*, Phys. Rev. D 62 (2000) 041302.
29. M. C. Bento *et al.*, Phys. Lett. B 518 (2000) 276.

# Dark matter: experimental and observational status

Vasiliki A. Mitsou

*Instituto de Física Corpuscular (IFIC), CSIC – Universitat de València,*
*C/ Catedrático José Beltrán 2, E-46980 Paterna (Valencia), Spain*
*E-mail: vasiliki.mitsou@ific.uv.es*
*webific.ific.uv.es/web*

This brief review covers recent results on searches for dark matter in collider experiments, as well as from direct and indirect detection observatories. It focuses on generic searches for dark matter signatures at the LHC, e.g. mono-$X$, dijets, etc. Recently observed astrophysical signals that may provide hints of dark matter are also discussed.

*Keywords*: Dark matter; Direct detection; Indirect detection; LHC.

## 1. Introduction

Dark matter (DM) still remains one of the most puzzling and fascinating mysteries in Fundamental Physics nowadays. The quest for unveiling its nature encompasses Cosmology, Astroparticle and Particle Physics. Observations over the past decades — obtained by combining a variety of astrophysical data, such as type-Ia supernovae, cosmic microwave background (CMB), baryon oscillations and weak lensing data — indicate that most of our Universe energy budget consists of unknown entities: $\sim 27\%$ is dark matter and $\sim 68\%$ is dark energy,[1] a form of ground-state energy.

## 2. DM Candidates

Weakly interacting massive particles (WIMPs) are the leading class of candidates for cold DM (CDM). They are thermal relics from Big Bang and the measured relic density implies that the DM annihilation cross section is of the same order as the one characterising the weak interaction scale, constituting the so-called "WIMP miracle". Specific theoretical models may provide naturally a DM particle, such as supersymmetry (SUSY), extra dimensions and little Higgs models.

Other non-WIMP possibilities to explain the DM observations are superWIMPs (gravitinos, axinos), axions, (sterile) neutrinos, fuzzy CDM, Q-balls, WIMPzillas and macroscopic objects, such as primordial black holes.[2] New paradigms are continuously being proposed, partly as a result of interplay with observations. For instance, self-interacting DM has been proposed to ameliorate observed tensions between $N$-body simulations of collisionless CDM and astrophysical observations on galactic scales: cusp-vs-core problem, too-big-to-fail problem, missing-satellite problem, diversity problem.[3–5]

## 3. Direct Detection

Direct detection (DD) of DM involves the observation of elastic scattering of WIMP off nuclei. It is sensitive to the WIMP mass $m_\chi$ and the cross section $\sigma_{\chi\text{–nucleon}}$.

A low-energy threshold in the WIMP-induced recoils is required for an efficient DD experiment, as well as reliable shielding of the detector from radioactive sources and cosmic backgrounds. The background suppression relies on the ability to distinguish nuclear recoils against other possible processes, such as electrons and $\alpha$-particles. Lastly, stable detector operation during annual and diurnal modulation is required.

The use of different targets safeguards against nuclei-related systematic uncertainties. Liquid noble gases, such as Xenon and Argon, offer sensitivity over the widest WIMP-mass range from 5 GeV to 1 TeV (Darkside, DARWIN, DEAP3600, LUX, LZ, Panda-X, XENON). The (oldest) technology of cryogenic crystals presents new innovations and covers $m_\chi \sim 1 - 10$ GeV (CRESST, EDELWEISS, SuperCDMS). Alternative targets with unique properties include NaI crystals and bubble chambers (ANAIS, COSINE, DAMA/LIBRA, SABRE, PICO). A recent review of DD concepts and status is given in Ref. 6.

No signal of dark matter in direct detection has been observed so far. The only persisting "anomaly" over several years is the annually modulating signal observed by DAMA/LIBRA, a massive array of low-background NaI(Tl) crystals installed in the Gran Sasso underground laboratory. This modulation has been seen in various phases of the detector and its period and phase are consistent with the expectation from the standard DM halo model.[7]

As the sensitivity in cross section lowers, it reaches the so-called "neutrino floor", i.e. where the (irreducible) neutrino-flux background becomes dominant. Some DM models, e.g. SUSY or Composite Higgs[8] models, are still viable in regions of the $\sigma_{\chi\text{–nucleon}}$-vs.-$m_\chi$ plane. Some caveats apply to the DD results. The astrophysical uncertainties on the local DM density and DM velocity distribution are typically large. On the particle physics side, the way DM interacts with the detector is only partially known, while other nuclear-physics uncertainties may be considerable.

## 4. Indirect Detection

In the indirect detection (ID) of DM, the focus is on the DM-particle decay or annihilation products in the galactic centre, dwarf galaxies, etc. Several probes are used as messengers, such as neutrinos, photons, antiprotons, and positrons. This class of observations may distinguish among different WIMP candidates: neutralinos, KK states, etc. Detection and analysis techniques used in ID and results are detailed in Ref. 9. Several hints of DM annihilation have been observed in data over the years frequently being attributed to astrophysical origin. Here, we outline two recent observations: excess of GeV gamma rays and the 21-cm signal line.

A spatially extended excess of $\sim 1 - 3$ GeV $\gamma$ rays from the region surrounding the Galactic Center has been identified, consistent with the emission expected from annihilating dark matter. High resolution $\gamma$-ray maps, such as the one shown in Fig. 1, render the excess robust and highly statistically significant, with a spectrum, angular distribution, and overall normalisation that is in good agreement with that predicted by simple annihilating $36 - 51$ GeV DM particle annihilating to $b\bar{b}$.[10]

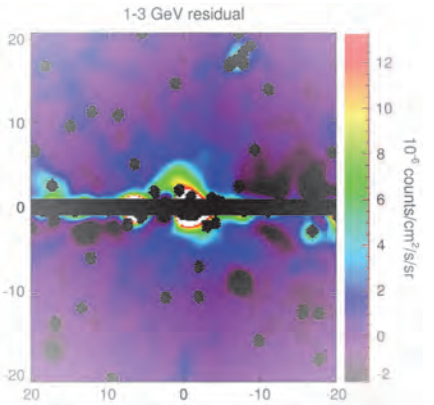



Fig. 1. Gamma-ray sky map after subtracting the point source model and the best-fit Galactic diffuse model, Fermi Bubbles, and isotropic templates. From Ref. 10.

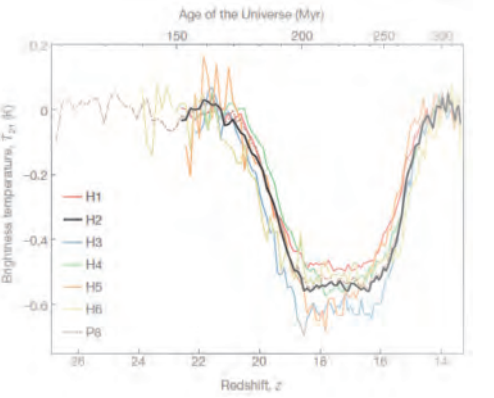



Fig. 2. Best-fitting 21-cm absorption profiles for different hardware cases of the EDGES experiment. The thick black line is the model fit with the highest signal-to-noise ratio. From Ref. 11.

The Experiment to Detect the Global EoR Signature (EDGES) is a radio-telescope for detection of hydrogen signatures from the Epoch of Reionisation (EoR), soon after the formation of the first stars and galaxies. The Collaboration reported the observation of an unexpectedly deep absorption in the radio background at 78 MHz, shown in Fig. 2, based on their low-band instruments and interpreted it as a redshifted 21-cm line.[11] It is unlikely that radiation from stars and stellar remnants to account for this discrepancy. Cooling of gas as a result of interactions between dark matter and baryons seems to explain the observed amplitude.

## 5. Searches at the LHC

The ATLAS and CMS experiments have embarked upon searches for signals of DM produced at the LHC early on during data taking.[12] Theoretical models, such as supersymmetry[13] or theories with extra-dimensions[14] provide a natural DM candidate, hence searches targeting these models implicitly cover dark matter, too. Since the candidate is only weakly interacting with matter, the common feature among these analyses is the requirement for large missing momentum $E_{\mathrm{T}}^{\mathrm{miss}}$. These models offer definite predictions, however they are characterised by a large number of parameters and this approach is profoundly model dependent.

The first generic approach towards DM searches involved the deployment of effective field theories (EFTs), extensively used in Run 1 (2010–2012). In this method, the interaction between DM and Standard Model (SM) particles is described by effective operators. It is clearly less model independent, yet it is only valid for low-momentum transfers $Q^2 \ll M_{\mathrm{med}}^2$, where $M_{\mathrm{med}}$ is the mediator mass.

A third approach extensively followed in Run 2 (2015–2018) assumes simplified topologies where the DM and SM particles interact via a mediator $(Z', H)$.[15] Its

advantage is that it covers features of a whole class of models and remains valid at high energies. In addition, it is described by a relatively small number of free parameters, namely the mediator and DM masses $M_{\text{med}}$ and $m_\chi$ and the mediator couplings to SM and DM particles $g_{\text{SM}}$ and $g_{\text{DM}}$, respectively.

Concerning the final states, DM cannot be directly observed at the LHC, however DM-pair production can be detected via the presence of an imbalance in measured transverse momenta of visible particles. The DM pair may only give rise to large $E_{\text{T}}^{\text{miss}}$, if it recoils to an energetic particle $X$, hence the so-called mono-$X$ searches, where $X$ is a jet,[16] a photon,[17] a top, a (hadronically[16] or leptonically[18] decaying) $W/Z$ boson. When $X = H(\to b\bar{b}, \gamma\gamma, \tau\tau)$ coupled to DM through a BSM effective vertex, it provides a direct probe of the DM–SM coupling. Additionally, the associated production of DM with $t\bar{t}$[19,20] and $b\bar{b}$[20] pairs is considered.

In addition to the dedicated analyses for DM, constraints can be extracted from recasting searches for heavy resonances. For an assumed $Z'$ mediator, the final states can be two quarks ("dijet"),[21,22] two Higgs bosons ($\to b\bar{b}, \tau\tau$), $H\gamma \to b\bar{b}\gamma$. Such searches are only sensitive to the mediator–SM particle coupling, e.g. the mediator–quark coupling in the case of dijet resonances. A wide range of mediator masses are covered with these analyses with/without jet/$\gamma/W$ initial state radiation.

The complementarity between the results from DM direct detection and related LHC searches is shown in Fig. 3. CMS limits do not include relic-density constraints and their relative importance strongly depend on the chosen coupling and model

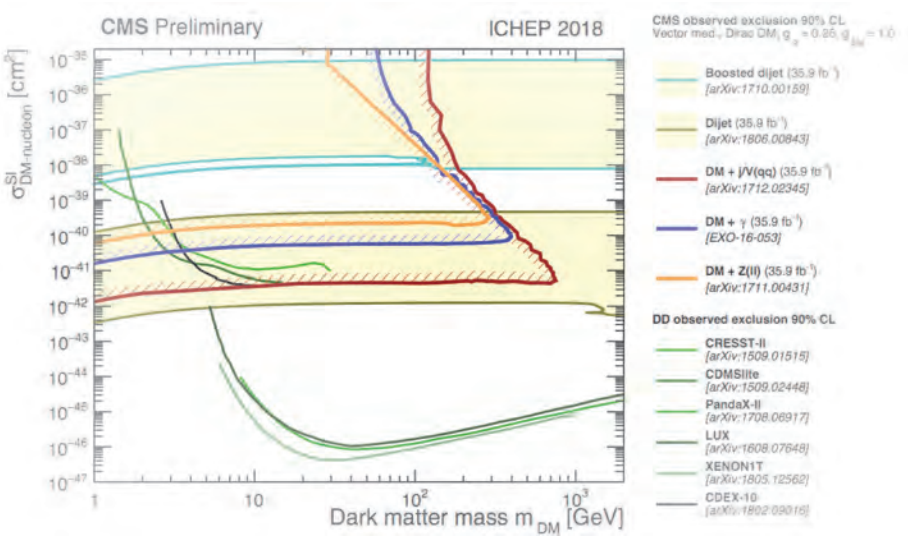



Fig. 3. CMS DM exclusion limits at 90% CL in the $m_{\text{DM}} - \sigma_{\text{SI}}$ plane for a vector mediator, Dirac DM and couplings $g_q = 0.25$ and $g_{\text{DM}} = 1.0$ compared with the XENON1T,[23] LUX,[24] PandaX-II,[25] CDMSLite[26], CDEX[27] and CRESST-II[28] limits, which constitute the strongest constraints in the shown mass range. From Ref. 29.

scenario. Therefore, the shown CMS exclusion regions in this plot are not applicable to other choices of coupling values or models.

## 6. Summary

Signals of dark matter are sought after in direct and indirect detection and in production in colliders. Additional constraints are obtained from cosmological observations on its nature (thermal/non-thermal, super/sub/relativistic, etc.) and on its relic density. Some still unexplained evidence in DD and ID are under investigation taking into account additional systematic uncertainties that matter–DM interactions may hinder and the strong dependence of the ID & DD results interpretation on astrophysical assumptions. In parallel, the LHC experiments search for DM in a variety of channels following different approaches and strategies characterised by high dependency on theoretical models and/or assumptions. Some approaches provide access to the mediator nature than to DM itself, rendering a possible signal difficult to be assigned to DM. The quest for dark matter continues in all fronts.

## Acknowledgements

The author would like to thank the MG15 Meeting organisers for the kind invitation to present this talk. This work was supported by the Generalitat Valenciana via the Project PROMETEO-II/2017/033, by the Spanish MICIU via the grant FPA2015-65652-C4-1-R, by the Severo Ochoa Excellence Centre Project SEV-2014-0398, and by a 2017 Leonardo Grant for Researchers & Cultural Creators, BBVA Foundation.

## References

1. Planck Collaboration, N. Aghanim *et al.* (2018), arXiv:1807.06209 [astro-ph.CO].
2. A. Kusenko and L. J. Rosenberg, *Working Group Report: Non-WIMP Dark Matter*, in *Proceedings, 2013 Community Summer Study on the Future of U.S. Particle Physics: Snowmass on the Mississippi (CSS2013): Minneapolis, MN, USA, July 29-August 6, 2013*, (2013). arXiv:1310.8642 [hep-ph].
3. D. N. Spergel and P. J. Steinhardt, *Phys. Rev. Lett.* **84**, 3760 (2000), arXiv:astro-ph/9909386 [astro-ph].
4. N. E. Mavromatos, C. R. Arguelles, R. Ruffini and J. A. Rueda, *Int. J. Mod. Phys.* **D26**, 1730007 (2016).
5. N. E. Mavromatos, *ΛCDM Model and Small-Scale-Cosmology "Crisis": from astrophysical explanations to new fundamental physics models*, in *15th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories (MG15) Rome, Italy, July 1-7, 2018*, (2019). p. to appear.
6. M. Schumann (2019), arXiv:1903.03026 [astro-ph.CO].
7. R. Bernabei *et al.*, *Eur. Phys. J.* **C73**, 2648 (2013), arXiv:1308.5109 [astro-ph.GA].

8. J. Yepes (2018), arXiv:1811.06059 [hep-ph].
9. T. R. Slatyer, *Indirect Detection of Dark Matter*, in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: Anticipating the Next Discoveries in Particle Physics (TASI 2016): Boulder, CO, USA, June 6-July 1, 2016*, (2018). pp. 297–353. arXiv:1710.05137 [hep-ph].
10. T. Daylan, D. P. Finkbeiner, D. Hooper, T. Linden, S. K. N. Portillo, N. L. Rodd and T. R. Slatyer, *Phys. Dark Univ.* **12**, 1 (2016), arXiv:1402.6703 [astro-ph.HE].
11. J. D. Bowman, A. E. E. Rogers, R. A. Monsalve, T. J. Mozdzen and N. Mahesh, *Nature* **555**, 67 (2018), arXiv:1810.05912 [astro-ph.CO].
12. V. A. Mitsou, *Int. J. Mod. Phys.* **A28**, 1330052 (2013), arXiv:1310.1072 [hep-ex].
13. V. A. Mitsou, *PoS* **CORFU2016**, 133 (2017).
14. ATLAS Collaboration, V. A. Mitsou, *EPJ Web Conf.* **182**, 02089 (2018).
15. G. Busoni *et al.* (2016), arXiv:1603.04156 [hep-ex], arXiv:1603.04156.
16. CMS Collaboration, A. M. Sirunyan *et al.*, *Phys. Rev.* **D97**, 092005 (2018), arXiv:1712.02345 [hep-ex].
17. ATLAS Collaboration, M. Aaboud *et al.*, *Eur. Phys. J.* **C77**, 393 (2017), arXiv:1704.03848 [hep-ex].
18. CMS Collaboration, A. M. Sirunyan *et al.*, *Eur. Phys. J.* **C78**, 291 (2018), arXiv:1711.00431 [hep-ex].
19. ATLAS Collaboration, M. Aaboud *et al.*, *JHEP* **06**, 108 (2018), arXiv:1711.11520 [hep-ex].
20. ATLAS Collaboration, M. Aaboud *et al.*, *Eur. Phys. J.* **C78**, 18 (2018), arXiv:1710.11412 [hep-ex].
21. CMS Collaboration, A. M. Sirunyan *et al.*, *JHEP* **01**, 097 (2018), arXiv:1710.00159 [hep-ex].
22. CMS Collaboration, A. M. Sirunyan *et al.*, *JHEP* **08**, 130 (2018), arXiv:1806.00843 [hep-ex].
23. XENON Collaboration, E. Aprile *et al.*, *Phys. Rev. Lett.* **121**, 111302 (2018), arXiv:1805.12562 [astro-ph.CO].
24. LUX Collaboration, D. S. Akerib *et al.*, *Phys. Rev. Lett.* **118**, 021303 (2017), arXiv:1608.07648 [astro-ph.CO].
25. PandaX-II Collaboration, X. Cui *et al.*, *Phys. Rev. Lett.* **119**, 181302 (2017), arXiv:1708.06917 [astro-ph.CO].
26. SuperCDMS Collaboration, R. Agnese *et al.*, *Phys. Rev. Lett.* **116**, 071301 (2016), arXiv:1509.02448 [astro-ph.CO].
27. CDEX Collaboration, H. Jiang *et al.*, *Phys. Rev. Lett.* **120**, 241301 (2018), arXiv:1802.09016 [hep-ex].
28. CRESST Collaboration, G. Angloher *et al.*, *Eur. Phys. J.* **C76**, 25 (2016), arXiv:1509.01515 [astro-ph.CO].
29. CMS Collaboration, *CMS Exotica Summary plots for 13 TeV data* https://twiki.cern.ch/twiki/bin/view/CMSPublic/SummaryPlotsEXO13TeV, (2018).

# Signs of interacting vacuum and dark matter in the universe

Joan Solà Peracaula*, Adrià Gómez-Valent and Javier de Cruz Pérez

*Departament de Física Quàntica i Astrofísica (FQA),*
*and Institute of Cosmos Sciences (ICCUB)*
*Universitat de Barcelona, Av. Diagonal 647, E-08028 Barcelona, Catalonia, Spain*
*sola@fqa.ub.edu, adriagova@fqa.ub.edu, decruz@fqa.ub.edu*

We consider the impact of dynamical dark energy (DDE) in the possible solution of the existing tensions in the ΛCDM. We test both interacting and non-interacting DE models with dark matter (DM). Among the former, the running vacuum model (RVM) interacting with DM appears as a favored option. The non-interacting scalar field model based on the potential $V \sim \phi^{-\alpha}$, and the generic XCDM parametrization, also provide consistent signs of DDE. The important novelty of our analysis with respect to the existing ones in the literature is that we use the matter bispectrum, together with the power spectrum. Using a complete and updated set of cosmological observations on SNIa+BAO+$H(z)$+LSS+CMB, we find that the crucial triad BAO+LSS+CMB (i.e. baryonic acoustic oscillations, large scale structure formation data and the cosmic microwave background) provide the bulk of the signal. The bispectrum data is instrumental to get hold of the DDE signal, as our analysis shows. If the bispectrum is not included, the DDE signal could not be currently perceived at a significant confidence level.

*Keywords*: Cosmological constant; dark energy theory; cosmological parameters; structure formation; cosmic microwave background.

## 1. Introduction

The cosmological constant problem in the context of quantum field theory (QFT) is well known. Already in the standard model of particle physics the value of the vacuum energy density $\rho_\Lambda$ is predicted to be some 55 orders of magnitude higher than the measured value, $\rho_\Lambda \sim 10^{-47}$ GeV$^4$ see Refs. 1–5. Rather than attempting to solve this problem here, we adopt a pragmatic point of view and consider the possibility to test whether the vacuum energy, and in general the dark energy (DE), is dynamical or not in the light of the current observations. This might help to shed some light not only on how to cure the CC-problem eventually, but also on how to solve some of the existing tensions between the ΛCDM model and observations, like the $H_0$-tension (the mismatch between the Planck measurements of $H_0$ and the local measurements based on the distance ladder approach)[6–8] and the $\sigma_8$-tension, which is described in terms of the combined observable $f(z)\sigma_8(z)$, where $f(z)$ is the linear growth rate and $\sigma_8(z)$ is the RMS matter fluctuation on scales of $R_8 = 8h^{-1}$ Mpc at redshift $z$. The corresponding prediction of the ΛCDM is known to be exceedingly large and hence the description of the LSS data is poorly accounted for by the concordance model.[9,10] In the following we briefly review how dynamical dark energy (DDE) can lend a helping hand on trying to solve some of these problems, although the evidence of DDE (or lack of it) varies significantly if taking different

*Speaker

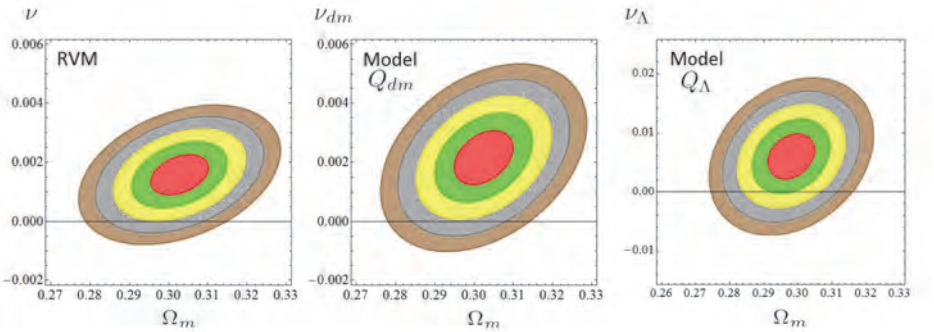



Fig. 1. Likelihood contours for the three DVMs (3)-(5) in the $(\Omega_m, \nu_i)$ plane up to $5\sigma$ c.l. after marginalizing over the rest of the fitting parameters. Use is made of SNIa+BAO+$H(z)$+LSS data and compressed CMB Planck 2015 data. The $\Lambda$CDM ($\nu_i = 0$) appears disfavored in front of the DVMs. In particular, the RVM and the $Q_{dm}$ are favored at $\sim 3\sigma$ c.l. See Ref. 14 for details on the SNIa+BAO+$H(z)$+LSS+CMB data used in our fit. We note, in particular, that bispectrum data is also included.[20]

studies in the literature, see Refs. 11–17 and the more recent analyses 18, 19 and references therein.

## 2. Cosmology beyond $\Lambda$ = const.

The Friedmann and acceleration equations when the vacuum evolves with the cosmic expansion, $\rho_\Lambda = \rho_\Lambda(t)$ can be written formally similar to the standard case when $\rho_\Lambda$ =const. If we focus on flat FLRW metric, the field equations that incorporate vacuum dynamics read

$$3H^2 = 8\pi G(\rho_m + \rho_\Lambda(t))\,, \qquad 2\dot{H} + 3H^2 = -8\pi G(w_m\rho_m + w_\Lambda\rho_\Lambda(t))\,, \tag{1}$$

where $w_m = 1/3, 0$ for relativistic and nonrelativistic matter, and $w_\Lambda = -1$ for the vacuum (irrespective of its time evolution). Hereafter we assume $w_\Lambda = -1$ – see 12 for the analysis of the general case with $w_\Lambda = -1+\epsilon$ ($|\epsilon| \ll 1$), i.e. for quasivacuum models. The total matter density $\rho_m$ involves the contributions from baryons and cold dark matter (DM): $\rho_m = \rho_b + \rho_{dm}$. We assume that the DM component is the only one that interacts with vacuum, whilst radiation and baryons are taken to be self-conserved, hence $\rho_r(a) = \rho_r^0\, a^{-4}$ and $\rho_b(a) = \rho_b^0\, a^{-3}$, where the superscript zero denotes the current values. The DM component, instead, exchanges energy with the vacuum, and the local conservation law reads

$$\dot{\rho}_{dm} + 3H\rho_{dm} = Q\,, \qquad \dot{\rho}_\Lambda = -Q\,. \tag{2}$$

The form of the source function $Q$ depends on the assumptions made on the dynamical vacuum model (DVM). Let us consider three model possibilities, one is the "running vacuum model" (RVM) — see Refs. 3–5. and references therein — and the other two are called the $Q_{dm}$ and $Q_\Lambda$ models, whose names bear relation to the

structure of the interaction source. For convenience they can also be labeled I, II and III:

$$\text{Model I (RVM)}: \; Q = \nu\, H(3\rho_m + 4\rho_r) \tag{3}$$

$$\text{Model II } (Q_{dm}): \; Q_{dm} = 3\nu_{dm} H \rho_{dm} \tag{4}$$

$$\text{Model III } (Q_\Lambda): \; Q_\Lambda = 3\nu_\Lambda H \rho_\Lambda\,. \tag{5}$$

Each model is characterized by a (dimensionless) coupling parameter $\nu_i = \nu, \nu_{dm}, \nu_\Lambda$ in the interaction source, to be fitted to the observational data. In the RVM case, the source function $Q$ in (3) can be derived from the following vacuum energy density:

$$\rho_\Lambda(H) = \frac{3}{8\pi G}\left(c_0 + \nu H^2\right)\,, \tag{6}$$

which can be motivated in QFT, see Ref. 3.[a] The other two models do not possess an analogous property and therefore are more *ad hoc*. The additive constant $c_0$ in (6) is fixed by the boundary condition $\rho_\Lambda(H = H_0) = \rho_\Lambda^0$, where $H_0$ is the current Hubble rate and $\rho_\Lambda^0$ is the present measured value for $\rho_\Lambda$. The dimensionless parameter $\nu$ encodes the dynamics of the vacuum and can be related with the $\beta$-function of the $\rho_\Lambda$-running.[3] Therefore we naturally expect $|\nu| << 1$ in QFT. Here we will treat $\nu$ as a free parameter of the RVM and, as indicated, it will be fitted to the data.

In Fig. 1 we display the fitting results on the $(\Omega_m, \nu_i)$-planes for the three models I, II and III (corresponding to $\nu_i = \nu, \nu_{dm}, \nu_\Lambda$), where $\Omega_m$ is the usual matter parameter at present. In Fig. 2 we can better appraise the impact of the various data sources on these contour plots in the specific case of the RVM, see Ref. 14 for details. The DDE option ($\nu \neq 0$, actually $\nu > 0$) is clearly favored in the last plot when we combine the crucial triad of BAO+LSS+CMB data.

## 3. Analytical solution of the dynamical vacuum models

One can solve Models I, II and III analytically in terms of the scale factor.[14] The DM density in the case of the RVM (Model I) as a function of the scale factor is

$$\rho_{dm}^{(I)} = \rho_{dm}^0\, a^{-3(1-\nu)} + \rho_b^0\left(a^{-3(1-\nu)} - a^{-3}\right) - \frac{4\nu\rho_r^0}{1+3\nu}\left(a^{-4} - a^{-3(1-\nu)}\right) \tag{7}$$

and that of the vacuum energy density reads

$$\rho_\Lambda^{(I)} = \rho_\Lambda^0 + \frac{\nu}{1-\nu}\left[\rho_m^0\left(a^{-3(1-\nu)} - 1\right) + \rho_r^0\left(\frac{(1-\nu)\,a^{-4} + 4\nu a^{-3(1-\nu)}}{1+3\nu} - 1\right)\right]. \tag{8}$$

The total matter density is just the sum of the conserved baryonic and radiation densities and the above DM density. As can be easily checked, for $\nu \to 0$ we recover the corresponding results for the ΛCDM, as it should.

[a] The structure of the vacuum energy density in the running vacuum model can also be motivated from the effective behavior of the power-law solutions of Brans-Dicke gravity, see Ref. 21.

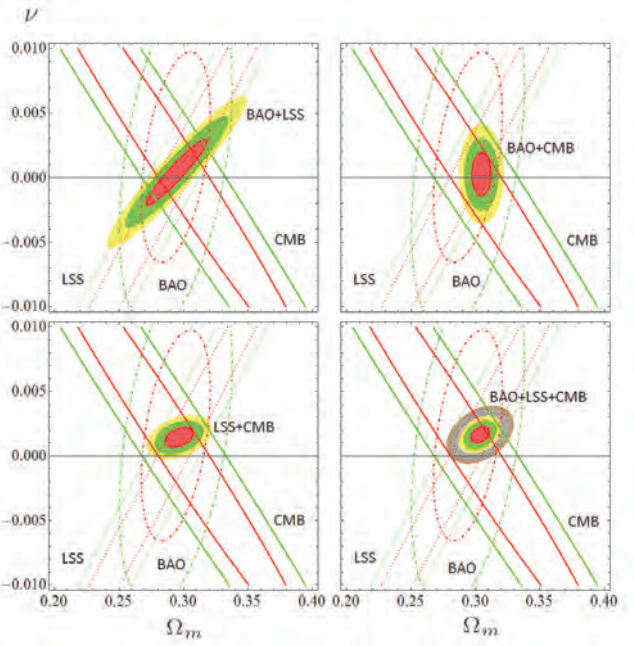



Fig. 2. Contour lines for the RVM by considering the effect of only BAO, LSS and CMB in all the possible pairings. Same data as in Fig. 1. Remarkably the simultaneous combination of these three observables is able to capture signs of running vacuum at a confidence level of $\gtrsim 3\sigma$ .

For Model II:

$$\rho_{dm}^{(II)}(a) = \rho_{dm}^{0}\, a^{-3(1-\nu_{dm})} \tag{9}$$

and

$$\rho_{\Lambda}^{(II)}(a) = \rho_{\Lambda}^{0} + \frac{\nu_{dm}\,\rho_{dm}^{0}}{1-\nu_{dm}}\left(a^{-3(1-\nu_{dm})} - 1\right). \tag{10}$$

And for Model III:

$$\rho_{dm}^{(III)}(a) = \rho_{dm}^{0}\, a^{-3} + \frac{\nu_{\Lambda}}{1-\nu_{\Lambda}}\rho_{\Lambda}^{0}\left(a^{-3\nu_{\Lambda}} - a^{-3}\right) \tag{11}$$

and

$$\rho_{\Lambda}^{(III)}(a) = \rho_{\Lambda}^{0}\, a^{-3\nu_{\Lambda}}\,. \tag{12}$$

Let us also briefly consider the simplest parametrization for DDE, namely the XCDM,[22] in which the DE density as a function of the scale factor is simply given by $\rho_X(a) = \rho_{X0}\, a^{-3(1+w_0)}$, with $\rho_{X0} = \rho_{\Lambda 0}$. Here $w_0$ is the (constant) EoS parameter. Last, but not least, we also study a scalar field model, denoted as $\phi$CDM. In this type of models the DE has a well-defined local Lagrangian description in terms

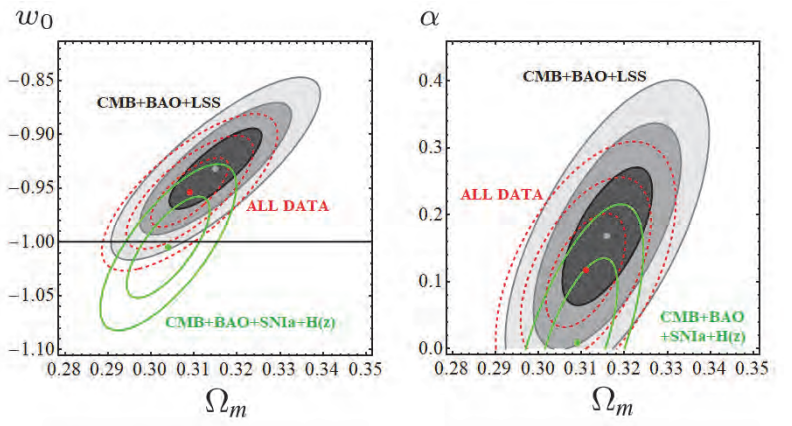



Fig. 3. Contour plots for the XCDM parametrization (left) and $\phi$CDM (right), using all sources of data (including LSS with bispectrum) as compared to using all data except LSS. Here we use Planck 2018 data with compressed CMB likelihood. When LSS (with bispectrum) is included, DDE appears once more favored at $\sim 3\sigma$ c.l. See Ref. 18 for details.

of a scalar field. Taking $\phi$ dimensionless the energy density and pressure read

$$\rho_\phi = \frac{M_P^2}{16\pi}\left(\frac{\dot\phi^2}{2} + V(\phi)\right) \qquad p_\phi = \frac{M_P^2}{16\pi}\left(\frac{\dot\phi^2}{2} - V(\phi)\right), \tag{13}$$

where $M_P$ is the Planck mass, which in natural units takes the value $M_P = 1/\sqrt{G}$ $= 1.2211\times10^{19}$ GeV. Note that in our conventions $V(\phi)$ has dimension 2 in natural units. The scalar field satisfies the Klein-Gordon equation in the context of the FLRW metric $\ddot\phi + 3H\dot\phi + \partial V(\phi)/\partial\phi = 0$. Here, as a particular case, we consider the original quintessence potential by Peebles and Ratra:[23]

$$V(\phi) = \frac{1}{2}\kappa M_P^2 \phi^{-\alpha}, \tag{14}$$

where $\kappa$ and $\alpha$ are dimensionless. The parameter $\alpha$ should be positive and sufficiently small so that $V(\phi)$ can mimic approximately a CC slowly decreasing with time, such that it can eventually surface (near our time) over the rapidly decaying matter density. In Fig. 3 we can see the corresponding contour lines for both the XCDM and $\phi$CDM, using SNIa+BAO+$H(z)$+LSS data (including the bispectrum component for the LSS+BAO data from Ref. 20) and compressed CMB Planck 2018 likelihood. The DDE option is consistently favored in both cases since $w_0 > -1$ (for the XCDM) and $\alpha > 0$ (for the $\phi$CDM) at $\lesssim 3\sigma$ c.l. after marginalizing over $\Omega_m$ [b]. Failure of including the bispectrum component (which seems to be particularly sensitive to the dynamics of the DE) may explain why other recent analyses on interacting vacuum and $\phi$CDM could not trace signs of DDE, see e.g., Refs. 16, 17.

[b]We should point out that a similar level of evidence is obtained using Planck 2015 data with full (i.e. non-compressed) CMB likelihood, see Ref. 18 for details.

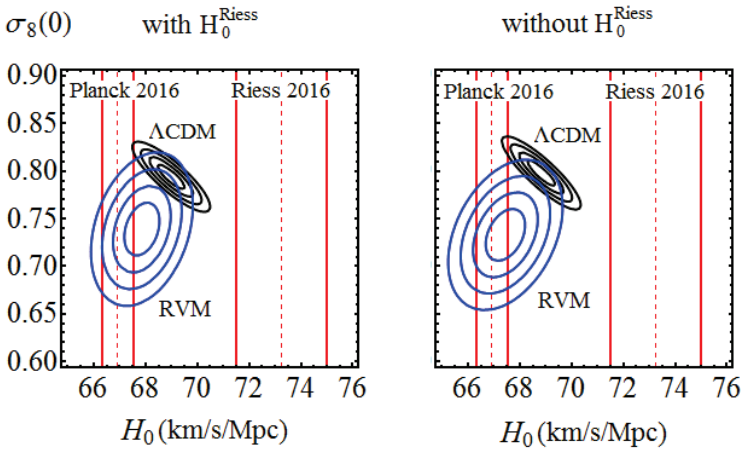



Fig. 4. Contour lines for the ΛCDM (black) and RVM (blue) up to $4\sigma$ in the $(H_0, \sigma_8(0))$-plane[12]. On the left plot we have included the R16 point[8] on $H_0$ in our fit, whereas on the right plot it was not included. In both cases the favored range for $H_0$ is the Planck range. At the same time the optimal (lower) $\sigma_8(0)$ values are attained by the RVM, see Eq. (3), whereas the ΛCDM yields too large values.

## 4. Dealing with the tensions through vacuum dynamics

As indicated in the introduction, one of the persisting tensions with the ΛCDM is the discrepancy between the Planck value of the Hubble parameter $H_0$ obtained from the CMB anisotropies, and the local HST measurement (based on distance ladder estimates from Cepheids). The latter reads $H_0 = 73.24 \pm 1.74$ km/s/Mpc (referred to as R16) and $H_0 = 73.48 \pm 1.66$ km/s/Mpc (R18).[8] In contrast, the CMB value is $H_0 = 67.51 \pm 0.64$ km/s/Mpc, as extracted from Planck 2015 TT,TE,EE+lowP+lensing data, or $H_0 = 66.93 \pm 0.62$ km/s/Mpc, based on Planck 2015 TT,TE,EE+SIMlow data.[6] In both cases there is a tension above $3\sigma$ c.l. with respect to the local measurement. This controversial situation has stimulated a lot of discussion in the literature. For instance, in Ref. 24 the phantom DE option is exploited as a means to solve the tension. In contrast, in Ref. 12 it is proposed that vacuum and quasivacuum dynamics support the Planck measurement of $H_0$ against the local measurement with no need of phantom behavior. A similar conclusion is reached with the XCDM and $\phi$CDM models studied here.[18] In addition, a variety of model-independent approaches such as the Inverse Distance Ladder method or the Multi-Task Gaussian Process[25] as well as other techniques (e.g., the Weighted Polynomial Regression[26]) do also support a lower value of $H_0$ in the ballpark of Planck CMB measurements and hence in full consistency with DDE.

The $H_0$-tension is closely connected with another type of problem, which is the role played by the LSS data, namely the data on $f(z)\sigma_8(z)$. There seems to be no way at present to correctly account for both the CMB and the LSS data within

the ΛCDM. This new pitfall is at the root of the so-called $\sigma_8$-tension, one of the most intriguing phenomenological problems of the ΛCDM.[9,10] The problem has been dealt with in the literature from a variety of points of view, see e.g., Refs. 27, 28. Here we will focus once more on the possibility of DDE.[12–14,18]

In the presence of interacting vacuum the perturbations equations of the ΛCDM must be modified appropriately.[13,14] The detailed analysis of these matters performed in the mentioned references provides the basis for a possible solution to the $\sigma_8$-tension within the RVM. In Fig. 4, we present the likelihood contours in the $(H_0, \sigma_8(0))$-plane. In the left plot we show the situation when the local $H_0$ value R16 is included in the fit, whereas in the right plot we show the case when that local value is *not* included. The plots for the more updated R18 value would be almost indistinguishable from the R16 ones. From Fig. 4 we can see that in order to reach the R16 neighborhood the contours should be extended beyond the $5\sigma$ c.l., which would lead to a too large value of $\sigma_8(0)$ for the RVM. One can also show that it would lead to a too low value of $\Omega_m$.[12]

Notice that $H_0$ and $\sigma_8(0)$ are positively correlated in the RVM (i.e. $H_0$ increases/decreases when $\sigma_8(0)$ increases/decreases, respectively), whilst they are anticorrelated in the ΛCDM ($H_0$ increases when $\sigma_8(0)$ decreases, and vice versa). The simultaneous $H_0$ and $\sigma_8(0)$ tensions could only be resolved simultaneously if the contours are extended well beyond $5\sigma$ at the expense of values of $\Omega_m$ below 0.27. However, this does not seem to be the most likely solution. Vacuum dynamics offers a better scenario. The opposite correlation feature with respect to the ΛCDM allows e.g., the RVM to improve the LSS fit in the region where both $H_0$ and $\sigma_8(0)$ are smaller than the respective ΛCDM values (cf. Fig. 4). This explains why the Planck range for $H_0$ is clearly preferred by the RVM, as it allows a much better description of the LSS data. One can also derive similar conclusions using the XCDM and $\phi$CDM, see Ref. 18, where again a positive correlation between $H_0$ and $\sigma_8(0)$ is found and the LSS tension can then be relaxed in the range of Planck values of $H_0$.

## 5. Conclusions

We have analyzed the ability of dynamical dark energy (DDE) models to fit the overall cosmological observations as compared to the ΛCDM. We find that there is non-negligible evidence that these models are more favored. Furthermore, DDE is capable of smoothing out the LSS tensions of the concordance model with the data, and they favor the (lower) Planck value of $H_0$ over the local one.

## Acknowledgments

We are funded by projects FPA2016-76005-C2-1-P (MINECO), 2017-SGR-929 (Generalitat de Catalunya) and MDM-2014-0369 (ICCUB). JdCP also by a FPI fellowship associated to FPA2016.

## References


1. S. Weinberg, Rev. of Mod. Phys. **61** (1989) 1.
2. P.J.E. Peebles & B. Ratra, Rev. Mod. Phys. **75** (2003) 559.
3. J. Solà, J. Phys. Conf. Ser. **453** (2013) 012015; AIP Conf.Proc. **1606** (2014) 19.
4. J. Solà & A. Gómez-Valent, Int. J. Mod. Phys. D**24** (2015) 1541003.
5. A. Gómez-Valent, PhD Thesis, arXiv:1710.01978.
6. Planck 2015 results: P.A.R Ade *et al.*, A&A **594** (2016) A13: Planck 2016 intermediate results: N. Aghanim *et al.*, A&A **596** (2016) A107.
7. Planck 2018 results: N. Aghanim *et al.*, arXiv:1807.06209.
8. A.G. Riess *et al.*, ApJ. **826** (2016) 56; A.G. Riess *et al.*, ApJ. **855** (2018) 136.
9. E. Macaulay, I.K. Wehus I.K. & H.C. Eriksen, Phys. Rev. Lett. **111** (2013) 161301.
10. S. Basilakos & S. Nesseris, Phys. Rev. D**96** (2017) 063517.
11. J. Solà, A. Gómez-Valent & J. de Cruz Pérez, ApJ **811** (2015) L14; ApJ. **836** (2017) 43; Mod. Phys. Lett. A**32** (2017) 1750054.
12. J. Solà, A. Gómez-Valent & J. de Cruz Pérez, Phys. Lett. B**774** (2017) 317; Int. J. Mod. Phys. A**32** (2017) 1730014.
13. A. Gómez-Valent & J. Solà, MNRAS **478** (2018) 126; EPL 120 (2017) 39001.
14. J. Solà, J. de Cruz Pérez & A. Gómez-Valent, MNRAS **478** (2018) 4357; EPL **121** (2018) 39001.
15. G.B. Zhao *et al.*, Nat. Astron. **1** (2017) 627.
16. A.A. Costa *et al.*, JCAP **1701** (2017) 028; Y.H. Li *et al.* Phys. Rev. D**93** (2016) 023002; Phys. Rev. D**90** (2014) 123007; Phys. Rev. D90 (2014) 063005.
17. C-G. Park & B. Ratra, ApJ **868** (2018) 83; arXiv:1803.05522; J. Ooba, B. Ratra & N. Sugiyama, arXiv:1802.05571.
18. J. Solà, A. Gómez-Valent & J. de Cruz Pérez, "Signs of Dynamical Dark Energy in Current Observations", arXiv:1811.03505 (to appear in Phys. Dark Univ., 2019).
19. M. Martinelli, N.B. Hogg, S. Peirone, M. Bruni & D. Wands, arXiv:1902.10694.
20. H. Gil-Marín *et al.*, MNRAS **465** (2017) 1757.
21. J. Solà, Int. J. of Mod. Phys. D**27** (2018) 1847029; J. de Cruz Pérez & J. Solà, Mod. Phys. Lett. A33 (2018) 1850228.
22. S.M. Turner & M. White, Phys. Rev. D**56** (1997) R4439.
23. P.J.E. Peebles & B. Ratra, ApJ **325** (2018) L17; Phys. Rev. D**37** (1988) 3406.
24. E.D. Valentino, A. Melchiorri & O. Mena, Phys. Rev. D**96** (2017) 043503.
25. S.M. Feeney *et al.,* Phys. Rev. Lett. **122** (2019) 061105; B.S. Haridasu, V.V. Luković, M. Moresco & N. Vittorio, JCAP **1810** (2018) 015; E. Macaulay *et al.* (DES Collab.), arXiv:1811.02376.
26. A. Gómez-Valent & L. Amendola, JCAP **1804** (2018) 051; A. Gómez-Valent, arXiv:1810.02278.
27. S. Anand, P. Chaubal, A. Mazumdar & S. Mohanty, JCAP **1711** (2017) 005.
28. R. An, C. Feng & B. Wang, JCAP **1802** (2018) 038.

# Tests of Pauli Exclusion Principle as a new frontier of quantum gravity phenomenology

Andrea Addazi*

*Department of Physics & Center for Field Theory and Particle Physics, Fudan University, 200433 Shanghai, China*
** E-mail: andrea.addazi@lngs.infn.it*

Pauli Exclusion Principle violating transitions in atoms and nuclei are predicted from many models of non-commutative space-time, such as $\theta$-Poincaré theory. Exclusion sensitivity curves from DAMA/LIBRA and BOREXINO are so powerful to exclude many non-commutative space-time models until the Planck energy scale. A new frontier of quantum gravity phenomenology is opened, inspiring future underground projects with a higher sensitivity then current experiments.

*Keywords*: Pauli Exclusion Principle violations, quantum gravity, underground experiments.

The Pauli Exclusion Principle (PEP) states that no any electrons, or other identical fermions, can be in the same quantum mechanical state [1]. On the basis of it, electrons organize themself in the well known tower of atomic levels, as similarly nucleons in nuclei. A state with two electrons with the same aligned spins on the same orbital level is prohibited. Therefore, PEP is directly connected to the same anthropic existence since related to atomic and nuclei matter stability. However, the anthropic argument is not the only reason of PEP in Nature: such a principle emerges out from the Spin Statistic Theorem (SST) of relativistic quantum field theory – defining the quantum statistical properties of bosons (integer spins) and fermions (1/2 spin).

Many examples, in which the SST is effectively violated, can be found in many exotic new condensed matter materials, where, of course, pseudo-particle fields propagate in a non-Lorentz invariant background. Typical pseudo-particle excitations found in superconductors, dubbed anyons, have a mix fermionic/bosonic statistics not described by standard Bose-Einstein and Fermi-Dirac distributions. Inspired by such a rich phenomenology in materials, it is highly not obvious that electron creation/annihilation operators would obey to the same quantum mechanical anti-commutation rules at energies close to the Planck quantum gravity scale. Such an issue may be intimately connected with the Nature of Space and Time as well as Quantum Gravity UV completion. Indeed, the SST is formulated on the basis of Lorentz invariance, locality and microcausality. If one of these principles is violated or deformed, any particle fields may exhibit energy dependent anyonic statistics. Possible effects of high energy Lorentz symmetry deformations, such as in many cases the PEPV, can be probed with high accuracy, competitive with panckian

suppressions! Non-commutative space-time models, such as $\theta$-Poincaré, provide examples of theories which substitute or deform the Lorentz invariance. Quantum field theories can be reformulated in many non-commutative frameworks through the redefinition of product rules such as the Moyal plane model for $\theta$-Poincaré [2, 3, 4]. Non-commutative space-time theories like $\theta$-Poincaré typically predict a energy scale dependent deformation of the co-algebra which percolates on a deformation of commutator and anticommutator relations for fields creation/annihilation operators. The energy scales entering in quantum gravity corrections are depending from the specific models and they can be related to the particle masses, kinetic energy of PEP violating energy transitions of the specific process considered. The predicted energy dependence can be parametrized as $(\bar{E}/\Lambda)^k$, where $\bar{E}$ is typically $<< \Lambda$ in every realistic non-commutative quantum gravity models and in every processes in underground experiments. Such a factor corrects the creation/annihilation commutators as (for boson) $[a_i, a_j^\dagger]_{\bar{E},\Lambda,k} = a_i a_j^\dagger - [1+\eta(\bar{E}/\Lambda)^k/2+O(\bar{E}/\Lambda)^{k+1})]a_j^\dagger a_i = \delta_{ij}$, where $\eta = \pm 1$ just determines the sign of the quantum gravity correction. It is worth to note that for $\Lambda \to \infty$, the (anti)commutator just reduces to the standard model one. In non-commutative space-time models, the scale $\Lambda$ is related to the non-commutative critical energy scale. Around (lenght) $\simeq \Lambda^{-1}$, space-time coordinates start to fully exhibit a non-commutative nature as $[x_\mu, x_\nu] = i\Lambda^{-2}\theta_{\mu\nu}$, where $\theta_{\mu\nu}$ is a fully antisymmetric tensor. Once again, one can check that for $\Lambda \to \infty$, standard commutative space-time coordinates are re-obtained. The deformation of the spin statistics can be directly related to transition rate processes. Indeed the decay width of a single nucleon transition to an ith already filled state, $\Gamma_i$, can be expressed as $\Gamma_i = (\bar{E}/\Lambda)^k \Gamma_i^{(0)}$ where $\Gamma_i^{(0)}$ is the width of the corresponding PEP preserving transition in all possible final states. More generically, there can also be an angular dependence of the processes that here we ignore but are fully considered in Refs. [3, 4].

Considering the comparison between the best constrains from DAMA/LIBRA and BOREXINO, we infer very strong bounds on the $(k, \Lambda)$ parameter space (see Fig. 1 and Ref. [2]). As we can see from Fig.1, both DAMA/LIBRA [8] and BOREXINO [9] are able to exclude corrections up to the order $(\bar{E}/\Lambda)^2$ until the Planck energy scale, i.e. the case of $\Lambda \simeq M_{Pl}$ and $k = 2$ is excluded. Such high precision measures can be achieved by virtue of an extremely low backgrounds in the 10 MeV-range, of both the experiments, as well as a large statistical sample for increasing the limits on PEPV transition time scale. It is also worth to note that exotic scenarios where $\Lambda \simeq 10 \div 100\,\mathrm{TeV}$ are ruled out at practically every order of k-power series. From the theoretical side, non-commutative space-time models are highly motivated in both string theory and loop quantum gravity frameworks and this may be not a coincidence but a matter of a hidden duality between the two theories [5, 6, 7]. On the other hand, alternatives for quantum gravity phenomenology from gravitational waves seem to be much less realistic, as discussed in Ref. [10]. In a certain class of models, the non-commutativity of space-time may also be related

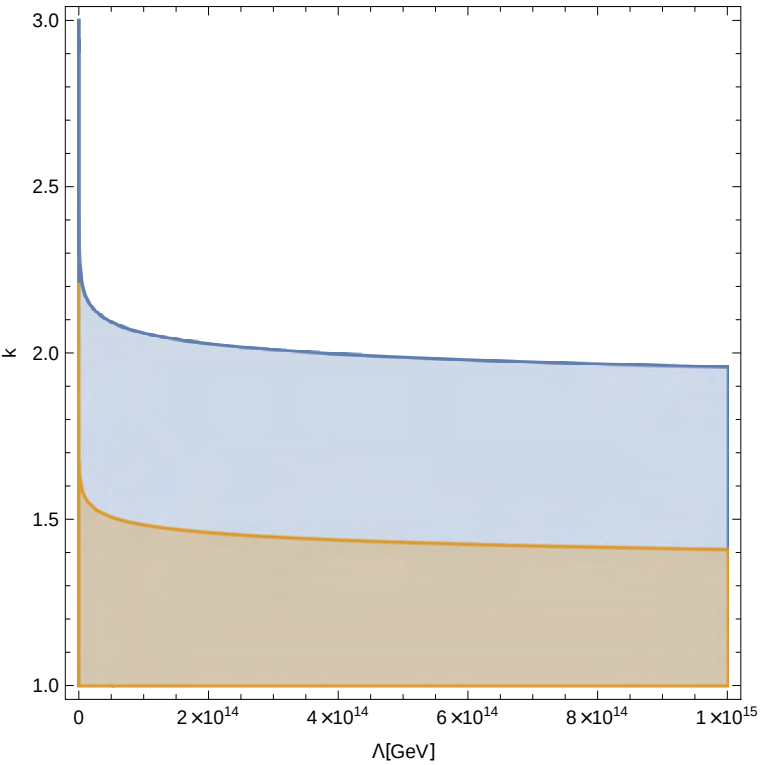



Fig. 1. The best $(k, \Lambda)$ contour parameter limits from DAMA and BOREXINO are displayed.

to non-associativity of quantum mechanics, which can be highly constrained from atomic and space physics [11]. It is worth to remark that, from the phenomenological side, VIP-2 experiment, using atomic techniques and not nuclear process measures such as in the case of BOREXINO, provides important alternative tests of other PEPV processes which do not destabilize nuclei [12].

## Acknowledgments

I would like to thank P. Belli, R. Bernabei and A. Marciano for our many exciting discussions on these fascinating aspects.

## References

1. W. Pauli, Phys. Rev. **58** (1940) 716. doi:10.1103/PhysRev.58.716.
2. A. Addazi, P. Belli, R. Bernabei and A. Marciano, Chin. Phys. C **42** (2018) no.9, 094001 doi:10.1088/1674-1137/42/9/094001 [arXiv:1712.08082 [hep-th]].
3. A. Addazi and A. Marciano, to appear in IJMP. arXiv:1811.06425 [hep-ph].
4. A. Addazi and R. Bernabei, to appear in MPLA, arXiv:1901.00390 [hep-ph].

5. A. Addazi and A. Marciano, arXiv:1802.09940 [hep-th].
6. A. Addazi and A. Marciano, Sci. China Phys. Mech. Astron. **61** (2018) no.12, 120421 doi:10.1007/s11433-018-9263-0 [arXiv:1707.05347 [hep-th]].
7. A. Addazi and A. Marciano, arXiv:1905.06673 [hep-th].
8. R. Bernabei *et al.*, Eur. Phys. J. C **62** (2009) 327. doi:10.1140/epjc/s10052-009-1068-1
9. G. Bellini *et al.* (Borexino Collaboration), Phys. Rev. C **81** (2010) 034317.
10. A. Addazi, A. Marciano and N. Yunes, Phys. Rev. Lett. **122** (2019) no.8, 081301 doi:10.1103/PhysRevLett.122.081301 [arXiv:1810.10417 [gr-qc]].
11. A. Addazi, A. Marciano, arXiv:1905.06807.
12. A. Pichler *et al.*, J. Phys. Conf. Ser. **718** (2016) no.5, 052030 doi:10.1088/1742-6596/718/5/052030 [arXiv:1602.00898 [physics.ins-det]].

# First results from DAMA/LIBRA phase2

R. Bernabei, P. Belli, A. Bussolotti, R. Cerulli, A. Di Marco, V. Merlo and F. Montecchia*

*Dip. di Fisica, Università di Roma "Tor Vergata" and INFN, sez. Roma "Tor Vergata", Rome, Italy*

F. Cappella, A. d'Angelo, A. Incicchitti and A. Mattei

*Dip. di Fisica, Università di Roma "La Sapienza" and INFN, sez. Roma, Rome, Italy*

V. Caracciolo

*INFN Laboratori Nazionali del Gran Sasso, Assergi, Italy*

C. J. Dai, H. L. He, X. H. Ma, X. D. Sheng and Z. P. Ye†

*Key Laboratory of Particle Astrophysics, Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, China*

The DAMA/LIBRA experiment has been collecting data deep underground in the Gran Sasso National Laboratory (LNGS) in its phase2 to further investigate the Dark Matter (DM) particles of the Galactic halo. The first data release of DAMA/LIBRA–phase2 corresponds to 6 full annual cycles for an exposure of 1.13 ton × yr. The data confirm the evidence of a signal that meets all the requirements of the model independent DM annual modulation signature at 9.5 $\sigma$ C.L. in the energy region (1-6) keV. Considering the previous data collected by DAMA/NaI and by DAMA/LIBRA–phase1 in 2-6 keV, the achieved C.L. for the full exposure (2.46 ton × yr) is 12.9 $\sigma$.

*Keywords*: Dark matter; low background NaI(Tl) scintillators; model-independent annual modulation signature.

## 1. The DAMA/LIBRA–phase2

The DAMA/LIBRA experiment has been carried out with the aim to investigate the DM particles in the Galaxy by exploiting the DM annual modulation signature. The detector consists of 25 ultra low background NaI(Tl) detectors for an exposed mass of about 250 kg. It is the successor of the pioneer DAMA/NaI set-up[1–13] and it has been taking data in two phases: the first one, called DAMA/LIBRA-phase1[1,14–30] and the present DAMA/LIBRA-phase2[30–32]. This last phase is the result of an upgrade of the experiment in which all the photomultipliers (PMTs) were replaced by second generation Hamamatsu R6233MOD PMTs, having higher quantum efficiency and lower background with respect to those used in phase1[19,33]. The commissioning of the experiment was successfully performed in 2011, allowing the achievement of the software energy threshold at 1 keV.

The DM annual modulation signature is due to the Earth's revolution around the Sun which is moving in the Galaxy. Because of the Earth and Sun velocities

*Also: Dip. Ingegneria Civile e Ingegneria Informatica, Università "Tor Vergata", Rome, Italy.
†Also: University of Jinggangshan, Ji'an, Jiangxi, China.

composition in the Galactic frame, a larger flux of DM particles crossing the Earth is expected around $\simeq$ 2 June, while a smaller one is expected around $\simeq$ 2 December. The DM signal must satisfy simultaneously many requirements: the rate must contain a component modulated as a cosine function (1), with period of one year (2), and phase peaked around $\simeq$ 2 June (3); this modulation should be present only in a well-defined low energy range, where DM particle induced events can be present (4); it applies only to those events in which just one detector of many actually "fires" (*single-hit* events), since the DM particle multi-interaction probability is negligible (5); in the region of maximal sensitivity the modulation amplitude must be $\simeq 7\%$ for usually adopted halo distributions, but it can be larger (even up to $\simeq 30\%$) for some cases (6). Many DM candidates, interaction types and scenarios can be explored. Only systematic effects or side reactions accounting for the measured modulation amplitude and satisfying all the peculiarities of the signature could mimic it, but none is available[1,3–5,14–17,20,24,25,30].

## 2. The results

The commissioning of DAMA/LIBRA–phase2 occurred in 2011 and that year was dedicated to preliminary studies[19]; the full annual cycles of DAMA/LIBRA–phase2 considered for the DM annual modulation analysis are six (exposure of 1.13 ton×yr). The cumulative exposure, including the former DAMA/NaI and DAMA/LIBRA–phase1, is 2.46 ton×yr. The same procedures already adopted for the DAMA/LIBRA–phase1[1,14–17] have also been exploited in the analysis of DAMA/LIBRA–phase2[30–32].

Many analyses have been performed on the data to study the peculiarities expected from the DM annual modulation signature. Here just some results and examples will be reported. More details have been addressed in Ref. 30–32. As example, in Figure 1 the residual rates of the *single-hit* scintillation events measured in the former DAMA/LIBRA–phase1 and the new DAMA/LIBRA–phase2 are reported; the data are from 2 keV, the software energy threshold of DAMA/LIBRA–phase1, up to 6 keV. The residual rates are calculated from the measured rate of the *single-hit* events after subtracting the unmodulated part (see e.g. Refs. 1,4,5,15–17). The null modulation hypothesis is rejected at very high C.L. by $\chi^2$ test: $\chi^2/d.o.f. = 199.3/102$, corresponding to a P-value $= 2.9 \times 10^{-8}$. The residual rates versus time for 1 keV energy threshold are reported in Ref. 30–32. The *single-hit* residual rates shown in Figure 1 and those of DAMA/NaI have been fitted with the function: $A\cos\omega(t - t_0)$, considering a period $T = \frac{2\pi}{\omega} = 1$ yr and a phase $t_0 = 152.5$ day (June $2^{nd}$) as expected by the DM annual modulation signature. The $\chi^2/d.o.f.$ is equal to 113.8/138 and a modulation amplitude $A = (0.0102 \pm 0.0008)$ cpd/kg/keV is obtained[30]. If the period and the phase of the signal are kept free in the fit, the achieved C.L. for the full exposure (2.46 ton × yr) is 12.9 $\sigma$; the modulation amplitude of the *single-hit* scintillation events is: $(0.0103 \pm 0.0008)$ cpd/kg/keV, the measured phase is $(145 \pm 5)$ days and the measured period is $(0.999 \pm 0.001)$ yr, all these values are well in agreement with those expected for DM particles.

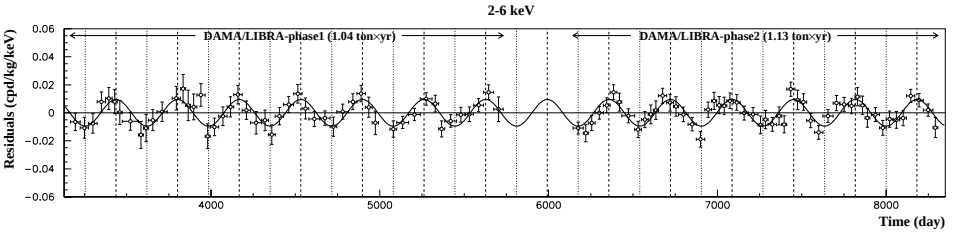



Fig. 1. Residual rate for *single-hit* scintillations measured by DAMA/LIBRA–phase1 and DAMA/LIBRA–phase2 in the (2–6) keV energy interval. The superimposed curve is the cosinusoidal functional forms $A\cos\omega(t-t_0)$ with a period $T=\frac{2\pi}{\omega}=1$ yr, a phase $t_0=152.5$ day (June $2^{nd}$) and modulation amplitude, $A$, equal to the central value obtained by best fit on the data points of DAMA/LIBRA–phase1 and DAMA/LIBRA–phase2. The residual rates versus time for 1 keV energy threshold are reported in Ref. 30–32.

The DAMA/LIBRA–phase2 data confirm the evidence of a signal satisfying all the requirements of the model independent DM annual modulation signature. No systematics or side reactions have been found able to mimic the exploited DM signature (i.e. to account for the whole measured modulation amplitude and to simultaneously satisfy all the requirements of the signature).

Other analyses, not shown here, exclude the presence of any background modulation in the energy regions above 6 keV. For example, the measured rate integrated above 90 keV, $R_{90}$, as a function of the time has been analysed [1,17,30]; similar conclusion are also obtained considering data in other energy regions.

Among other performed analyses, a further investigation on DAMA/LIBRA–phase2 data was done – as in the previous DAMA/NaI and DAMA/LIBRA–phase1 experiments [1,5,15–17] – by applying to the *multiple-hit* residual rate[a] the same hardware and software procedures used to acquire and to analyse the *single-hit* ones. Since the probability that a DM particle interacts in more than one detector is negligible, a DM signal is expected only in the *single-hit* residual rate. Therefore, by comparing *single-hit* and *multiple-hit* results is equivalent to compare the cases of DM particles beam-on and beam-off. In this way a further test is possible on the background behaviour in the same energy range in which the positive effect is observed. Figure 2 shows the residual rates of the *single-hit* scintillation events of DAMA/LIBRA–phase2, as collected in a single cycle, and the residual rates of the *multiple-hit* events, in the (1–6) keV energy range. A clear modulation, satisfying all the peculiarities of the DM annual modulation signature, is present in the *single-hit* events, while the fitted modulation amplitude for the *multiple-hit* residual rate is well compatible with zero: $(0.0004\pm0.0004)$ cpd/kg/keV. Thus, again evidence of annual modulation with the proper peculiarities of the exploited signature is present in the *single-hit* residuals (events class to which the DM particle induced events belong), while it is absent in the *multiple-hit* residual rate (event class to which only background events belong).

[a]Multiple-hit is an event with a deposition of energy in coincidence in more than one detector.

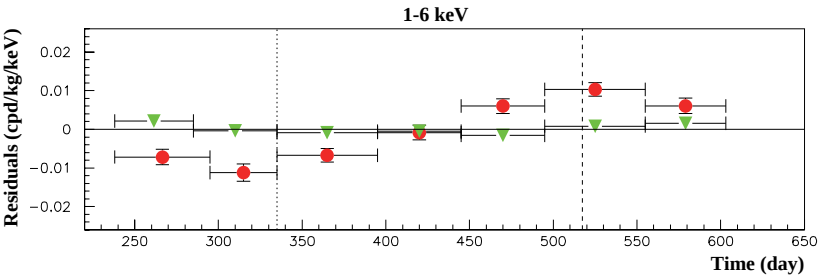



Fig. 2. Experimental residual rates of DAMA/LIBRA–phase2 *single-hit* events (circles), class of events to which DM events belong, and for *multiple-hit* events (filled triangles), class of events to which DM events do not belong. Data were considered as collected in a single annual cycle; in both cases the same hardware and software procedures have been applied. The time scale is the same as the previous DAMA papers for consistency [1,5,15–17]. The experimental points present the errors as vertical bars and the widths of the associated time bins as horizontal bars.

In conclusion, no background that can mimic the exploited DM signature, simultaneously accounting for the measured modulation amplitude and satisfying all the peculiarities of the signature, has been found or suggested by anyone throughout some decades thus far (see e.g. in Ref. 1,14–17,20,25,30)[b].

The annual modulation measured at low energy can also be highlighted by studying as a function of energy the modulation amplitude, $S_m$, obtained by analyzing the events with the maximum likelihood method and fixing $T = 1$ yr and $t_0 = 152.5$ day. The modulation amplitudes as a function of the energy for the whole data sets: DAMA/NaI, DAMA/LIBRA–phase1 and DAMA/LIBRA–phase2 (total exposure 2.46 ton×yr) are plotted in Figure 3; the data below 2 keV refer only to the DAMA/LIBRA–phase2 exposure (1.13 ton×yr). It can be deduced that positive signal is present in the (1–6) keV energy range, while $S_m$ values compatible with zero are present just above. All this confirms the previous analyses. The test of the hypothesis that the $S_m$ values in the (6–14) keV energy range have random fluctuations around zero yields $\chi^2$ equal to 19.0 for 16 $d.o.f.$ (upper tail probability of 27%).

[b]Considering some remarks recently appeared in literature, any hypothetical effect due to environmental He diffusion inside the PMTs can be excluded, e.g., by considering the following arguments: i) the PMTs are kept in HP nitrogen 5.5 atmosphere the gas being continuously fluxed through the apparatus ($\simeq$ 250 liters / hour); thus no helium accumulation process can take place (typical characteristic time for accumulation of helium in PMTs through glass, considering the permeability of materials is $\simeq$ 1 year). An estimation of helium concentration within the DAMA shield is less than $5 \times 10^{-11}$ ppm; thus any hypothetical effect due to He-correlated events is negligible; ii) any migration of helium into PMT would cause their irreversible continuous degradation which would be clearly seen over time and this is not observed in the DAMA data. e.g. the Dark Noise of the PMTs goes from 40 to 500 Hz and tends to improve over time on the contrary of what it is expected by He migration inside PMTs; iii) the PMTs used in DAMA/LIBRA-phase2 and DAMA/LIBRA-phase1 are different with different voltage values on the first dynode; this should have produced different modulation amplitudes in the two phases; this has not been observed.

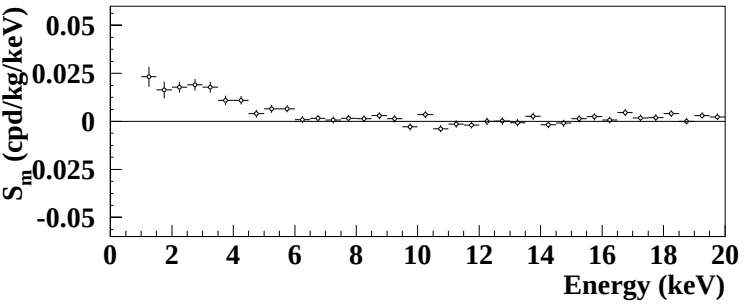



Fig. 3. Modulation amplitudes, $S_m$, for the whole data sets: DAMA/NaI, DAMA/LIBRA–phase1 and DAMA/LIBRA–phase2 (total exposure 2.46 ton×yr) above 2 keV; below 2 keV only the DAMA/LIBRA-phase2 exposure (1.13 ton × yr) is available and used. The energy bin $\Delta E$ is 0.5 keV. A modulation is evident in the lowest energy interval, and only $S_m$'s compatible with zero are present just above. In particular, the $S_m$ have random fluctuations around zero in (6–20) keV energy region with $\chi^2$ equal to 42.6 for 28 *d.o.f.* (upper tail probability of 4%); see text for comments.

For the case of (6–20) keV energy interval $\chi^2/d.o.f. = 42.6/28$ (upper tail probability of 4%). The obtained $\chi^2$ value is rather large due mainly to two data points, whose centroids are at 16.75 and 18.25 keV, far away from the (1–6) keV energy interval. The P-values obtained by excluding only the first and either the two points are 11% and 25%.

As already done for the other data releases[1,15–17,28], the observed annual modulation effect has been verified to be well distributed in all the 25 detectors.

No modulation has been found in any possible source of systematics or side reactions; therefore, cautious upper limits on possible contributions to the DAMA/LIBRA measured modulation amplitude have been obtained (see Ref. 3–5,15–17,20,25). They do not account for the measured modulation amplitudes, and also are not able to simultaneously satisfy all the many requirements of the signature. Similar analyses were also performed for the DAMA/NaI data[4,5].

## 3. Conclusions

The long-standing annual-modulation evidence observed by the DAMA is model-independent and it can be compatible with a wide set of DM scenarios.

Preliminary results recently reported by other experiments using NaI(Tl) detectors have not enough sensitivity to be comparable with the annual modulation evidence reported by DAMA. They require longer data taking and further developments. Both the negative results and all the possible positive hints, achieved so-far in the field, can be compatible with the DAMA result in many scenarios.

To efficiently disentangle among at least some of the many possible candidates and scenarios the lowering of the energy threshold below 1 keV is important. In this context related R&D's have been founded and are in progress.

## References


1. Bernabei, R. et al., *Int. J. of Mod. Phys. A* **2013**, *28*, 1330022.
2. Bernabei, R. et al., *Il Nuovo Cim. A* **1999**, *112*, 545-575.
3. Bernabei, R. et al., *Eur. Phys. J. C* **2000**, *18*, 283-292.
4. Bernabei, R. et al., *La Rivista del Nuovo Cimento* **2003**, *26*, 1-73.
5. Bernabei, R. et al., *Int. J. Mod. Phys. D* **2004**, *13*, 2127-2159.
6. Bernabei, R. et al., *Phys. Lett. B* **1996**, *389*, 757-766. Bernabei, R. et al., *Phys. Lett. B* **1998**, *424*, 195-201. Bernabei, R. et al., *Phys. Lett. B* **1999**, *450*, 448-455. Belli, P. et al., *Phys. Rev. D* **2000**, *61*, 023512. Bernabei, R. et al., *Phys. Lett. B* **2000**, *480*, 23-31. Bernabei, R. et al., *Phys. Lett. B* **2001**, *509*, 197-203. Bernabei, R. et al., *Eur. Phys. J. C* **2002**, *23*, 61-64.
7. Belli, P. et al., *Phys. Rev. D* **2002**, *66*, 043503.
8. Bernabei, R. et al., *Int. J. Mod. Phys. A* **2006**, *21*, 1445-1469.
9. Bernabei, R. et al., *Eur. Phys. J. C* **2006**, *47*, 263-271.
10. Bernabei, R. et al., *Int. J. Mod. Phys. A* **2007**, *22*, 3155-3168.
11. Bernabei, R. et al., *Eur. Phys. J. C* **2008**, *53*, 205-213.
12. Bernabei, R. et al., *Phys. Rev. D* **2008**, *77*, 023506.
13. Bernabei, R. et al., *Mod. Phys. Lett. A* **2008**, *23*, 2125-2140.
14. Bernabei, R. et al., *Nucl. Instr. and Meth. A* **2008**, *592*, 297-315.
15. Bernabei, R. et al., *Eur. Phys. J. C* **2008**, *56*, 333-355.
16. Bernabei, R. et al., *Eur. Phys. J. C* **2010**, *67*, 39-49.
17. Bernabei, R. et al., *Eur. Phys. J. C* **2013**, *73*, 2648.
18. Belli, P. et al. , *Phys. Rev. D* **2011**, *84*, 055014.
19. Bernabei, R. et al., *J. of Instr.* **2012**, *7*, P03009.
20. Bernabei, R. et al., *Eur. Phys. J. C* **2012**, *72*, 2064.
21. Bernabei, R. et al., *Eur. Phys. J. C* **2009**, *62*, 327-332.
22. Bernabei, R. et al., *Eur. Phys. J. C* **2012**, *72*, 1920.
23. Bernabei, R. et al., *Eur. Phys. J. A* **2013**, *49*, 64.
24. Bernabei, R. et al., *Eur. Phys. J. C* **2014**, *74*, 2827.
25. Bernabei, R. et al., *Eur. Phys. J. C* **2014**, *74*, 3196.
26. Bernabei, R. et al., *Eur. Phys. J. C* **2015**, *75*, 239.
27. Addazi, A. et al., *Eur. Phys. J. C* **2015**, *75*, 400.
28. DAMA coll. Issue dedicated to DAMA. *Int. J. of Mod. Phys. A* **2016**, *31*.
29. Cerulli, R. et al., *Eur. Phys. J. C* **2017**, *77*, 83.
30. Bernabei, R. et al., *Nucl. Phys and Atomic Energy* **2018**, *19*, 307.
31. Bernabei, R. et al., *Universe* **2018**, *4*, 116.
32. Bernabei, R. et al., *Bled Workshops in Physics* **2018**, *19* n. 2, 27.
33. Bernabei, R. and Incicchitti, A. Low background techniques in NaI(Tl) setups. *Int. J. Mod. Phys. A* **2017**, *32*, 1743007, doi:10.1142/S0217751X17430072.

# Investigation of rare processes with high purity germanium detectors

Matthias Laubenstein*

*Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali del Gran Sasso*
*Via G. Acitelli 22, I-67100 Assergi (AQ), Italy*
**E-mail: matthias.laubenstein@lngs.infn.it*

The exceptional sensitivity and high resolution of high purity germanium detectors in gamma-ray spectrometry and their use in underground laboratories has increasing application because of the important science and technology that they allow to be studied. In fundamental physics their application is focussed on rare phenomena, e.g. double beta decay, rare nuclear decays and Dark Matter search. In the past years there has been a growing number of underground measurements also in other fields such as environmental monitoring, surveillance of nuclear activities, benchmarking of other physical techniques, Life Science and material selection for equipment which require materials with extremely low levels of radioactivity. This report describes the state of the art in underground gamma-ray spectrometry with high purity germanium detectors. Various applications in underground laboratories are presented.

*Keywords*: Germanium detectors; ultra low background; material screening; underground physics.

## 1. Introduction

In the underground laboratories all over the world many experiments are probing different aspects of the standard model of particle physics and beyond to very high precision.[1] Especially two fields of phenomena in the low energy region are explored, the investigation of the nature of dark matter search on the one hand, and of the neutrino on the other. The exceptional sensitivity and high resolution of high purity germanium detectors makes them a natural candidate for being used as detector. Over the last decade there have been at least seven collaborations in the field of Dark Matter search (CDEX-10[2], CDMS[3], COGENT[4], EDELWEISS[5], MALBEK[6], SuperCDMS[7], and TEXONO[8]) and five in the field of neutrino studies, especially checking whether it is a Majorana particle or not and its mass by searching for the neutrinoless double beta decay ($0\nu\beta\beta$ decay) (GERDA[9], LEGEND[10], and MAJORANA[11]) and measuring the coherent elastic neutrino-nucleus scattering (COHERENT[12], and CONUS[13]). The dark matter experiments are searching directly for interactions of Dark Matter particle candidates (nuclear recoils) and for possible oscillations of this signal if present. The experiments searching for the $0\nu\beta\beta$ decay in the isotope $^{76}$Ge have achieved exceptionally low background levels and the next generation of experiments will be able to explore completely the region of inverted mass hierarchy for neutrinos.

Nevertheless, there is a growing number of underground measurements also in other fields such as environmental monitoring, surveillance of nuclear activities, benchmarking of other physical techniques, Life Science and material selection

for equipment which require materials with extremely low levels of radioactivity. Various applications in underground laboratories will be presented in the following.

## 2. Ultra low-level gamma-ray spectrometry

In this paper the acronym ULGS stands for ultra low-level gamma-ray spectrometry. This means that additional measures (such as using a muon shield or an underground laboratory) have been taken compared to low-level gamma-ray spectrometry, where a detector and shield made from selected radio pure materials are used.[14] When performing ULGS with an active shield (or "veto counter" or "muon shield"), one gains a lot above ground and at shallow depth laboratories.[15] It is, however, not possible to avoid the cosmogenic activation of detector and shield so a muon shield can not completely substitute an underground location.[16]

The Italian Institute for Nuclear Physics (INFN) performs ULGS in the laboratory STELLA (STELLA is the acronym for SubTerranean Low-Level Assay)[17], which is located in the underground laboratories of the Gran Sasso National Laboratory (LNGS) at a depth of 1400 m (3800 m of water equivalent). Since from the beginning in the late 1970s material selection was done for the LNGS in its construction phase in order to select the most radio pure construction materials available at that time (1979), and later on for the experiments that were installed at the LNGS searching for very rare events under conditions of cosmic silence, i.e. in the absence of cosmic rays. The definite ULGS work started in 1990 in small scale with three high purity germanium (HPGe) detectors as an exploratory research project.[18] The usefulness of the technique led to an increase of the capacity. For the moment there are 15 HPGe-detectors in STELLA. A wide range of measurements are necessary, which led to the decision of equipping the laboratory with general purpose detectors rather than detectors specialised for a specific type of measurements. The non routine character of the work makes calculation of the detection efficiency for many types of samples a requirement. Reference samples are used whenever possible, but the general rule is that the efficiency is calculated using the Monte Carlo code Geant4[19,20].

## 3. Applications

Several interesting developments in the use of underground gamma-ray spectrometry have arisen in recent years. As described before the major achievements were in the field of fundamental physics and double beta decay in particular. As the benefits of the technique became clearer to a wider community the number of applications has grown. The examples that follow illustrate the versatility of the technique.

### 3.1. *Small samples*

For special investigations the amount of available material can be very small (only few grams). A lot can be gained using ULGS where the background is negligible

in order to extract valuable information. At the LNGS meteorite samples of few grams were measured to give important information about the exposure history of meteorite falls[21] or to verify that the investigated specimen was really originating from a fresh meteorite fall and belonged to a group of other samples already measured[22]. Furthermore also for material screening it has become ever more important to measure extremely small samples (surface mounted electronics, Silicon photomultipliers etc). The use of ULGS gives the possibility to screen even either single components[17] or larger sets of very small samples to improve the sensitivity on the single component reaching levels below $\mu$Bq per piece.

### 3.2. *Physics measurements*

The extremely ow background in ULGS combined with the use of ultra pure or extremely purified materials[23], is a very good opportunity to look for rare decays in several radionuclides of interest. At the LNGS in the past years new decays to excited levels of radioactive isotopes were discovered. It started with the discovery[24] of the $^{115}$In decay to the first excited level of $^{115}$Sn, which was later again investigated with better accuracy[25]. Further experiments on this decay are in progress to investigate possible effects due to its very low Q value of the matrix in which the $^{115}$In is implemented. Later on in a similar way the $\alpha$ decay of $^{190}$Pt to the first excited level of $^{186}$Os was detected[26]. Higher precision measurements are currently in progress. Finally very recently the decay of $^{50}$V was remeasured, aiming at the detection of the $\beta$ decay branch. Unfortunately, only the electron capture decay branch was confirmed and remeasured with high precision, whereas for the $\beta$ decay only a lower limit could be determined[27], now reaching almost the theoretical predictions[28].

### 3.3. *Method validation*

For LNGS as well as for other physics institutes it is essential to have alternative analytical techniques in order to either resolve issues or disputes relating to measurements or to do more comprehensive measurements on the whole natural decay chains of $^{238}$U and $^{232}$Th, considering parent nuclides and later on the daughter nuclides, in order to look for secular equilibrium through out the decay chains. LNGS could confirm that ULGS and ICP-MS gave the same results in many different matrices down to very low concentrations.[29]

### 3.4. *Radiopurity studies*

In order to benefit from the reduction of the background induced by cosmic rays in underground laboratories, it is essential that all materials used for constructing a detector and its shield are radio pure. STELLA has contributed to radio purity studies for all large and small scale detectors for rare events located at the LNGS. (e.g. BOREXINO neutrino detector[30], GERDA[31] and CUORE[32] detectors for double beta decay, XENON1T[33,34] detector for dark matter search).

### 3.5. *Activation*

Activated samples are most suitable samples for measurement using ULGS. The reason for this is that very often there is no interference from other radionuclides that increase the detection limit for the radionuclide of interest. This fortunate situation is either due to the fact that the activated material was selectively chosen beforehand, like in the case of flux monitors in nuclear reactors[35], or like in the case of environmental sampling after an incident involving neutron irradiation outside a controlled area[36]. Another example of ULGS measurements of activated samples involves solving discrepancies between model calculations and measurements of activation induced by the A-bomb in 1945. Komura et al. could solve the problem with $^{152}$Eu in granite by using the Ogoya underground laboratory[37] and Gasparro et al. showed that older measurements of $^{60}$Co in steel could be significantly too high[38].

## 4. Conclusions

The LNGS of the INFN host the underground ultra-low background facility STELLA. Its primary mission is to perform the material screening for the fundamental physics experiments installed in the underground laboratories. The ultra-low background of the detector systems allow for detecting extremely low radioactivity levels in materials down to the $\mu$Bq kg$^{-1}$ level. The exceptional sensitivity and high resolution of high purity germanium detectors in gamma-ray spectrometry and their use in underground laboratories has increasing application in other fields of science. A growing number of underground measurements is done in fields such as environmental monitoring, surveillance of nuclear activities, benchmarking, Life Science and material selection for experiments, which require materials with extremely low levels of radioactivity. In the future it is envisaged to further increase the performance of the ULGS systems carefully selecting the materials. Moreover, there might be further developments including using underground production and storage of Ge-crystals, transfer the technology developed for fundamental physics experiments to screening detector systems, and for special applications to use special coincidence techniques. Nevertheless for the HPGe-detectors acting as all-purpose work horses it is important to keep the system simple in order to minimise the labour cost for analyses. In the future new interesting applications making use of ULGS could turn up.

## Acknowledgments

The author expresses his gratitude towards the organisers of the XV Marcel Grossman Meeting 2018 in Rome for having been given the opportunity to present an invited talk on his work at this important and renowned conference.

## References


1. A. Bettini, New underground laboratories: Europe, Asia and the Americas, *Phys. Dark Univ.* **4**, 36 (2014).
2. CDEX-10 Collab. (H. Jiang *et al.*), Limits on Light Weakly Interacting Massive Particles from the First 102.8 kg day Data of the CDEX-10 Experiment, *Phys. Rev. Lett.* **120**, 241301 (2018).
3. CDMS Collab. (Z. Ahmed *et al.*), Dark Matter Search Results from the CDMS II Experiment, *Science* **327**, Issue 5973, p. 1619 (2010).
4. CoGeNT Collab. (C.E. Aalseth *et al.*), CoGeNT: A Search for Low-Mass Dark Matter using p-type Point Contact Germanium Detectors, *Phys. Rev. D* **88**, 012002 (2012).
5. EDELWEISS-III Collab. (E. Armengaud *et al.*), Constraints on low-mass WIMPs from the EDELWEISS-III dark matter search, *J. Cosmo. Astropart. Phys.* **05**, p. 19 (2016).
6. G.K. Giovanetti for the MALBEK collaboration, A Dark Matter Search with MALBEK, in *Phys. Procedia* **61**, p. 19 (2015).
7. SuperCDMS Collab. (R. Agnese *et al.*), Results from the Super Cryogenic Dark Matter Search (SuperCDMS) experiment at Soudan, *Phys. Rev. Lett.* **120**, 061802 (2018).
8. TEXONO Collab. (M. Deniz *et al.*), Constraints on scalar-pseudoscalar and tensorial nonstandard interactions and tensorial unparticle couplings from neutrino-electron scattering, *Phys. Rev. D*, **95**, p. 033008 (2017).
9. GERDA Collab. (M. Agostini *et al.*), Background free search for neutrinoless double-$\beta$ decay with GERDA, *Nature* **544**, 47 (2017).
10. LEGEND Collab. (N. Abgrall *et al.*), The large enriched germanium experiment for neutrinoless double beta decay (LEGEND), in *AIP Conf. Proc.*, **1894**, 020027 (2017).
11. S. Elliott for the MAJORANA collaboration, Initial Results from the Majorana Demonstrator, in *J. Phys. Conf. Ser.*, **888**, 012035 (2017).
12. COHERENT Collab. (P. Coloma *et al.*), COHERENT enlightenment of the neutrino dark side, *Phys. Rev. D* **96**, 115007 (2017).
13. C. Buck for the CONUS collaboration, Latest results of the CONUS reactor neutrino experiment, Rencontres du Moriond (2019) `http://moriond.in2p3.fr/2019/EW/Program.html`.
14. M. Laubenstein *et al.*, Underground measurements of radioactivity, *Appl. Radiat. Isotop.* **61**, No. 2-3, 167 (2004).
15. G. Heusser *et al.*), GIOVE: a new detector setup for high sensitivity germanium spectroscopy at shallow depth, *J. Eur. Phys. J. C* **75**, Issue 11, p. 531 (2015).
16. G. Heusser, M. Laubenstein and H. Neder, Low-level germanium $\gamma$-ray spectrometry at the $\mu$Bq/kg level and future developments towards higher sensitivity, in *Radionuclides in the Environment - Int. Conf. On Isotopes in Env. Studies*, vol. 8, eds. P. Povinec, J.A. Sanchez-Cabeza, (Elsevier, Amsterdam, 2006), p. 495.

17. M Laubenstein, Screening of materials with high purity germanium detectors at the Laboratori Nazionali del Gran Sasso, *Int. J. Mod. Phys. A* **32**, N. 30, 1743002 (2017).
18. C. Arpesella, A low background counting facility at laboratori nazionali del Gran Sasso, *Appl. Radiat. Isot.* **47**, Issues 9-10, p. 991 (1996).
19. S. Agostinelli *et al.*, Geant4 — a simulation toolkit, *Nucl. Inst. Meth. A* **506**, No. 3, p. 250 (2003).
20. M. Boswell *et al.*, MaGe-a Geant4-Based Monte Carlo Application Framework for Low-Background Germanium Experiments, *IEEE Transact. Nucl. Sci.* **58**, Issue 3, p. 1220 (2011).
21. P.P. Povinec *et al.*, Cosmogenic radionuclides and mineralogical properties of the Chelyabinsk (LL5) meteorite: What do we learn about the meteoroid?, *Meteor. Planet. Sci.* **50**, Issue 2, p. 273 (2015).
22. M.E.I. Riebe *et al.*, Cosmic-ray exposure ages of six chondritic Almahata Sitta fragments, *Meteor. Planet. Sci.* **52**, Issue 11, p. 2353 (2017).
23. Yu. P. Bobrov *et al.*, Purification of vanadium by electron beam melting, *Probl. Atom. Sci. Techn.* **1**, p. 27 (2014).
24. C. Cattadori *et al.*, Observation of $\beta$ decay of $^{115}$In to the first excited level of $^{115}$Sn, *Nucl. Phys A* **748**, No. 1-2, p. 333 (2005).
25. E. Andreotti *et al.*, Half-life of the $\beta$ decay $^{115}\mathrm{In}(9/2^+) \rightarrow {}^{115}\mathrm{Sn}(3/2^+)$, *Phys. Rev. C* **84**, 044605 (2011).
26. P. Belli *et al.*, First observation of $\alpha$ decay of $^{190}$Pt to the first excited level ($E_{exc} = 137.2$ keV) of $^{186}$Os, *Phys. Rev. C* **83**, 034603 (2011).
27. M. Laubenstein *et al.*, A new investigation of half-lives for the decay modes of $^{50}$V, *Phys. Rev. C* **99**, 045501 (2019).
28. M. Haaranen *et al.*, $\beta$-decay half-life of $^{50}$V calculated by the shell model, *Phys. Rev. C* **90**, 044314 (2014).
29. S. Nisi *et al*, Comparison of inductively coupled mass spectrometry and ultra low-level gamma-ray spectroscopy for ultra low background material selection, *Appl. Radiat. Isot.* **67**, Issue 5, p. 828 (2009).
30. BOREXINO collaboration (C. Arpesella *et al.*), Measurements of extremely low radioactivity levels in BOREXINO, *Astropart. Phys.* **18**, No. 1, p. 1 (2002).
31. GERDA Collab. (M. Agostini *et al.*), Upgrade for Phase II of the GERDA experiment, *Eur. Phys. J. C* **78**, 388 (2018).
32. CUORE Collab. (C. Alduino *et al.*), The projected background for the CUORE experiment, *Eur. Phys. J. C* **77**, 543 (2017).
33. XENON1T Collab. (E. Aprile *et al.*), Lowering the radioactivity of the photomultiplier tubes for the XENON1T dark matter experiment, *J. Eur. Phys. J. C* **75**, Issue 11, p. 546 (2015).
34. XENON1T Collab. (E. Aprile *et al.*), Material radio assay and selection for the XENON1T dark matter experiment, *Eur. Phys. J. C* **77**, p. 890 (2017).

35. F. Tzika *et al.*, Coordinated underground measurements of gamma-ray emitting radionuclides for plasma physics research, *Appl. Radiat. Isot.* **126**, p. 121 (2017).
36. M. Hult *et al.*, Measurements of $^{60}$Co in spoons activated by neutrons during the JCO criticality accident at Tokai-Mura in 1999, *J. Environ. Radioact.* **73**, Issue 3, p. 307 (2004).
37. K. Komura and Y. Hamajima, Ogoya underground laboratory for the measurement of extremely low levels of environmental radioactivity: review of recent projects carried out at OUL, *Appl. Radiat. Isot.* **61**, p. 185 (2004).
38. J. Gasparro *et al.*, Measurements of $^{60}$Co in massive steel samples exposed to the Hiroshima atomic bomb explosion, *Health Phys.*, **102**, p. 400 (2012).

# On the dynamical instability of self-gravitating systems

Giuseppe Alberti*

*Living Systems Research,*
*Roseggerstraße 27/2, A-9020 Klagenfurt am Wörthersee, Austria*
**E-mail: giuseppe.alberti@ilsr.at*

We study the dynamical stability of self-gravitating systems in presence of anisotropy. In particular, we introduce a stability criterion, in terms of the adiabatic local index, that generalizes the stability condition $< \gamma > \geq 4/3$ of the isotropic regime. Also, we discuss some applications of the criterion.

*Keywords*: Anisotropy; Dynamical Instability; Adiabatic Local Index; Gravity.

## 1. Introduction

The object of this work consists in the study of the dynamical stability in self-gravitating systems, through the deduction of a stability criterion. In particular, our aim is to obtain a criterion that extends the validity of the stability condition $< \gamma > \geq 4/3$ to the anisotropic systems.

In Sec. 2 we briefly deduce the stability criterion in Newtonian gravity whereas, in Sec. 3, we discuss some applications of the criterion in order to quantitatively check how much the presence of the anisotropy can affect the onset of the instability. In Sec. 4 we draw some conclusions.

## 2. Stability criterion in Newtonian gravity

In Newtonian gravity the pulsation equation writes (see Ref. 1)

$$\frac{\partial^2\xi}{\partial t^2}+\frac{4\xi P_r'}{r\rho}-\frac{1}{\rho}\Big[\frac{\gamma_r P_r}{r^2}(r^2\xi)'\Big]'+\frac{6(P_r-P_t)\xi}{r^2\rho}+\frac{2(P_r-P_t)\xi'}{r\rho}+\frac{2(\gamma_t P_t-\gamma_r P_r)(r^2\xi)'}{r^3\rho}=0 \tag{1}$$

where $P_r$ and $P_t$ are the radial and tangential components of the pressure tensor, respectively, $\rho$ the density, $r$ the radial coordinate, $\xi$ the Lagrangian displacement. The adiabatic local indexes $\gamma_r$ and $\gamma_t$, along the radial and the trasverse components, are given by

$$\gamma_r=\frac{\rho}{P_r}\Big(\frac{\partial P_r}{\partial\rho}\Big)_S \qquad \text{and} \qquad \gamma_t=\frac{\rho}{P_t}\Big(\frac{\partial P_t}{\partial\rho}\Big)_S\,, \tag{2}$$

where the subscript $S$ indicates that the derivatives are performed by keeping the entropy constant. Developing the calculations[2], we easily get the stability criterion

$$\mathcal{Q}=\frac{\int_0^R(\gamma_r P_r+2\gamma_t P_t)r^2dr}{\int_0^R(P_r+2P_t)\,r^2dr}\geq\frac{4}{3}\,. \tag{3}$$

In the isotropic limit the foregoing expression reduces to $\mathcal{Q}=< \gamma > \geq 4/3$.

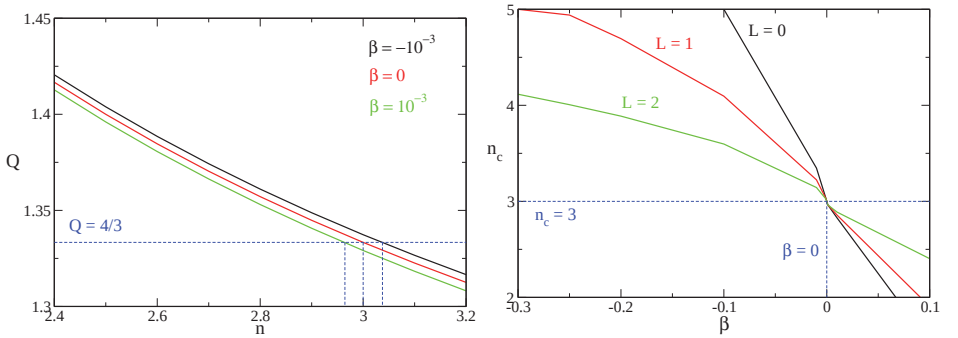



Fig. 1. **Left panel**: representation of the function $\mathcal{Q} = \mathcal{Q}(n)$ [see Eq. (3)], for $N = 2$, $L = 0$, $\beta = -10^{-3}$ (black line), 0 (red line) and $10^{-3}$ (green line). The plot shows that an unstable configuration in the isotropic limit can become stable if $\beta < 0$. Vice versa, when $\beta > 0$, a stable configuration in the isotropic limit can become unstable. **Right panel**: Critical value of the polytropic exponent $n_c$ as a function of the anisotropy parameter $\beta$, for $L = 0$, 1 and 2.

## 3. Applications

### 3.1. *Polytropes*

Let us apply Eq. (3) to the study of the stability of the polytropic models advanced by Herrera & Barreto[3]. Leaving out the details of the calculations[2], the anisotropic Lane-Emden equation writes

$$\frac{1}{\xi^2}\frac{d}{d\xi}\left(\xi^2\frac{d\theta}{d\xi}\right) = -\theta^n \mathcal{F}(\beta)\,, \tag{4}$$

where $n$ is the polytropic exponent, $\beta$ the (dimensionless) anisotropy parameter and $\mathcal{F}(\beta)$[a] is given by

$$\mathcal{F}(\beta) = 1 - \frac{\beta}{\xi^2\theta^n}\frac{d}{d\xi}\left[f(\xi)\xi^{N+1}\right]. \tag{5}$$

In the foregoing expression, the exponent $N$ and the function $f$ depend on the model. In the following, we consider $f(\xi) = (1+\xi)^L$ (with $L \in \mathbb{R}$). The isotropic Lane-Emden equation is recovered for $\beta = 0$ (implying $\mathcal{F} = 1$).

In Fig. 1 (left panel) we have represented the $\mathcal{Q}$-factor [see Eq. (3)] as a function of the polytropic exponent $n$. The diagram shows dissimilar behaviors according to the sign of the anisotropy parameter.

[a]To keep the decreasing behavior of $\theta$, we require that $\mathcal{F}(\beta) > 0$. This condition allows an upper limit for the anisotropy parameter $\beta$ (see Ref. 2).

If $\beta < 0$, which corresponds to the radial anisotropy (i.e. $P_t < P_r$, see Ref. 2), we observe a tendency towards the stability because the critical value $n_c$ of the polytropic exponent for the onset of the instability is larger than $n = 3$.[b] On the other hand, for $\beta > 0$, which corresponds to the tangential anisotropy (i.e. $P_t > P_r$), we observe that the rising of the dynamical instability is favored (from the plot, we note that $n_c < 3$).

In the right panel of Fig. 1 we have represented $n_c$ as a function of $\beta$, for three values of the exponent $L$. The plot shows that $n_c > 3$ for $\beta < 0$, in confirmation of the fact that the (presence of) radial anisotropy leads the system to the stability. If $\beta > 0$, by contrast, we find $n_c < 3$, in confirmation of the fact that the tangential anisotropy favors the rising of the instability.

### 3.2. *Anisotropic stars*

In this section we consider two models, advanced by Dev & Gleiser [5], conceived as deviations from the homogeneous model (i.e. $\rho = \rho_0$). In formulae

$$P_t = P_r + C\rho r^2 \,, \tag{6a}$$

$$P_t = P_r(1 + C\rho r^2) \,. \tag{6b}$$

In the foregoing expressions, $C$ estimates the strength of the anisotropy and can be both positive and negative, *a priori*. Similarly to the case of polytropes, $C > 0$ corresponds to the tangential anisotropy and $C < 0$ to the radial anisotropy. Integrating the equilibrium equations by using Eqs. (6a) and (6b), we obtain [5]

$$P_r = \rho_0^2 \Big( \frac{2\pi G}{3} - C \Big) (R^2 - r^2) \,, \tag{7a}$$

$$P_r = \frac{2\pi G \rho_0}{3C} \Big[ 1 - e^{C\rho_0 (r^2 - R^2)} \Big] \,. \tag{7b}$$

In the previous equations, $R$ and $G$ represent the radius of the star and the gravitational constant, respectively. Concerning the stability, for the ansatz (6a)-(7a), Eq. (3) yields $\mathcal{Q} = 2 > 4/3$, showing that the system is dynamically stable. For the ansatz (6b)-(7b), conversely, we find a more interesting situation. Eq. (3), indeed, takes the form

$$\mathcal{Q} = 1 + \frac{5}{2x^2} - \frac{15}{4x^4} + \frac{15 D_+(x)}{4x^5} \geq \frac{4}{3} \,, \tag{8}$$

where $x = R\sqrt{C\rho_0}$ and $D_+(x)$ is the Dawson function. In the limit $x \to 0$, Eq. (8) reduces to $\mathcal{Q} = 2 > 4/3$ that corresponds to the case previously analyzed. On the other hand, for $x \to +\infty$, we obtain $\mathcal{Q} = 1 < 4/3$ that corresponds to a loss of stability. Consequently, we expect to find a critical value of $x$ allowing the separation between stable and unstable configurations.

[b] $n = 3$ corresponds to the critical value in the isotropic case.

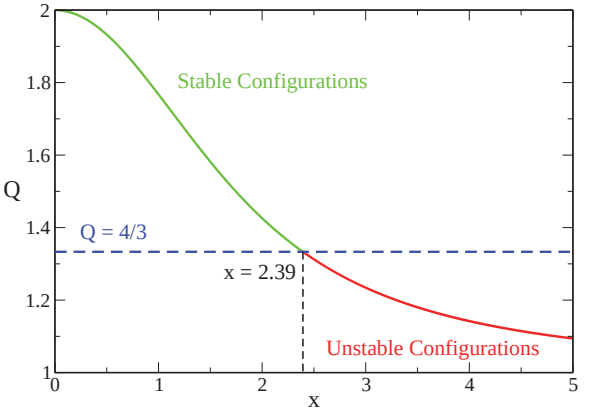



Fig. 2. Representation of the function $\mathcal{Q} = \mathcal{Q}(x)$, where $x$ is defined as $x = R\sqrt{C\rho_0}$. The plot shows the existence of a critical value $x = \bar{x}$, corresponding to the onset of the dynamical instability. Numerically we find $\bar{x} = 2.39$ or, equivalently, $\bar{C} = 5.73/(\rho_0 R^2)$.

In Fig. 2 we have represented Eq. (8) and, as we see, the existence of this critical value is confirmed. Numerically, we find that dynamical instability sets in if $x \geq \bar{x} = 2.39$, i.e. if $C \geq \bar{C} = 5.73/(\rho_0 R^2)$. It is interesting to notice that the stability of the star, rather than the anisotropy parameter, depends on the central density.

### 3.3. *Degenerate fermionic configurations*

In this section we focus on the degenerate fermionic configurations, a particular case of a more general study carried out in Ref. 6. The equation of state (EOS), in a parametric form, is

$$\rho = \frac{4\pi g m^4 \sigma^3}{3h^3}\sqrt{W^3}\left(1 + \frac{2r^2}{5r_a^2}W\right), \tag{9a}$$

$$P_r = \frac{4\pi g m^4 \sigma^5}{15h^3}\sqrt{W^5}\left(1 + \frac{2r^2}{7r_a^2}W\right), \tag{9b}$$

$$P_t = \frac{4\pi g m^4 \sigma^5}{15h^3}\sqrt{W^5}\left(1 + \frac{4r^2}{7r_a^2}W\right). \tag{9c}$$

In the foregoing expressions, $W$ is the cutoff energy (Fermi energy in this case[6]), $r_a$ the anisotropy radius, $\sigma$ the velocity dispersion and the other symbols have their usual meaning. The adiabatic local indexes $\gamma_r$ and $\gamma_t$ are given by

$$\gamma_r = \frac{5r_a^2 + 2r^2W}{7r_a^2 + 2r^2W}\,\frac{8rW^2 + 7W'(5r_a^2 + 2r^2W)}{8rW^2 + 5W'(3r_a^2 + 2r^2W)}, \tag{10a}$$

$$\gamma_t = \frac{5r_a^2 + 2r^2W}{7r_a^2 + 4r^2W}\,\frac{16rW^2 + 7W'(5r_a^2 + 4r^2W)}{8rW^2 + 5W'(3r_a^2 + 2r^2W)}, \tag{10b}$$

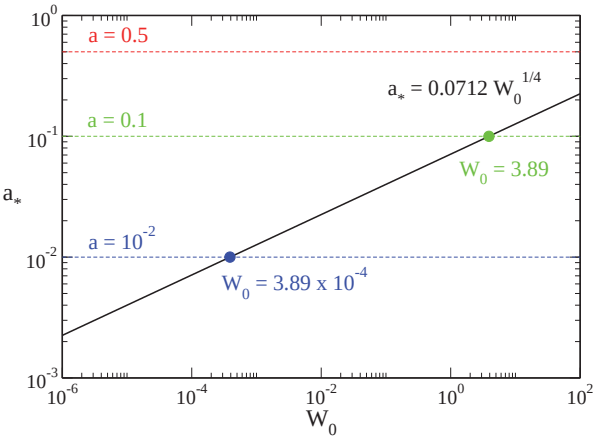



Fig. 3. Values of $a_*$ as a function of the central density $W_0$ [see Eq. (9a)]. The plot shows that, for a given value of $a$ (horizontal lines), there exists a critical value of the central density $W_0^c$ such that dynamical instabilities set in. For $a = 0.5$ (red line) we see that $W_0^c >> 1$ (the system is isotropic and thus stable) whereas, for a $a = 0.1$ (green line), we find $W_0^c = 3.89$ and for $a = 0.01$ (blue line) we have $W_0^c = 3.89 \times 10^{-4}$. This shows that $W_0^c \to 0$ for $a \to 0$ (the system is unstable).

where $W' = dW/dr$. In the isotropic limit $r_a \to +\infty$ the foregoing expressions yield $\gamma_t = \gamma_r = 5/3$. Consequently, according to Eq. (3), we get $\mathcal{Q} = 5/3 > 4/3$ confirming that, in the isotropic limit, degenerate fermionic configurations are stable[c]. Things change in the fully anisotropic regime $r_a \to 0$: we have indeed

$$\gamma_t = \gamma_r = \frac{4+7y}{4+5y}, \qquad \text{where} \qquad y = \frac{d\log(W)}{d\log(r)}. \tag{11}$$

As the reader can see, if $-4/5 < y < -4/7$, $\gamma_r$ and $\gamma_t$ are negative. In these conditions, the speed of the two sound waves along the radial and the tangential axis, given by

$$c_{sr} = \sqrt{\frac{\gamma_r P_r}{\rho}}, \quad c_{st} = \sqrt{\frac{\gamma_t P_t}{\rho}}. \tag{12}$$

become a complex number: the configuration, therefore, is unstable. This feature is a consequence of the trend of density and pressure profiles, which have become "hollow" [6].

Hollow configurations are characterized by the presence of a maximum achieved far from the center. Consequently, the profile is monotonic increasing until the maximum and monotonic decreasing from the maximum until the boundary.

A profile is hollow (and thus unstable) if, for a given value of $W_0$, the value of the anisotropy radius $r_a$ is below a critical one. To get this critical value, we compute

[c]We can also show that the EOS can be written as $P_t = P_r = P = K\rho^{5/3}$. This is the EOS of a polytrope of index $n = 3/2$ that, as it is known, represents a stable configuration (being $n = 3/2 < 3$).

$d^2\rho/dr^2 = 0$ at the center and solve for $a$. Defining the anisotropy parameter as $a = r_a/r_0$ ($r_0$ is a scaling length[6]) and indicating by $a_*$ the critical value, we have

$$a_* = \frac{r_a}{r_0}\Big|_{critical} = \frac{\sqrt[4]{W_0}}{2\pi\sqrt{5}}\,. \tag{13}$$

In Fig. 3 we have represented the function $a_* = a_*(W_0)$. According to the plot, degenerate fermionic configurations are stable if, for a given value of $W_0$, $a > a_*$ and unstable if $a \leq a_*$.

## 4. Concluding remarks

In this work we have studied the dynamical stability of anisotropic self-gravitating systems. The analysis carried out has shown that, according to the type of anisotropy, the onset of the instability is modified.

In prevalence of radial anisotropy ($P_r > P_t$), we have observed that the systems have the tendency to evolve towards stable configurations. Unstable configurations in the isotropic limit ($P_r = P_t$) can become stable. In prevalence of tangential anisotropy ($P_r < P_t$), by contrast, we have observed that the rising of the instability is favored. Stable configurations in the isotropic limit can become unstable.

In Ref. 2 we have studied the stability of other systems, such as Elliptical Galaxies.

We think that the stability criterion introduced and applied in this work represents a powerful tool for the study of the dynamical stability of a large class of astrophysical systems. The extension of the criterion to General Relativity and other applications will be adressed to forthcoming publications.

## References

1. G. Alberti & M. Merafina, "*Proceedings of the Fourteenth Marcel Grossmann Meeting*", 2485 - 2491 (2017)
2. G. Alberti & M. Merafina, *in preparation* (2019)
3. L. Herrera & W. Barreto, Phys. Rev. D **87**, 087303 (2013)
4. F. Shojai, M.R. Fazel, A. Stepanian & M. Kohandel, Eur. Phys. J. C **75**, 250 (2015)
5. K. Dev & M. Gleiser, Gen. Rel. Grav. **35**, 1435 (2003)
6. M. Merafina & G. Alberti, Phys. Rev. D **89**, 123010 (2014)

# Gravitational phase transition of self-gravitating systems of fermions in General Relativity

Giuseppe Alberti*

*Living Systems Research,*
*Roseggerstraße 27/2, A-9020 Klagenfurt am Wörthersee, Austria*
**E-mail: giuseppe.alberti@ilsr.at*

Pierre-Henri Chavanis*

*Laboratoire de Physique Théorique (UMR 5152), Université Paul Sabatier,*
*118 Route de Narbonne, F-31062 Toulouse, France*
**E-mail: chavanis@irsamc.ups-tlse.fr*

The Thomas-Fermi model is extended at finite temperature, to describe the gravitational phase transition occurring in massive fermionic systems in a general-relativistic framework. It is shown that, when a nondegenerate fermionic gas (for $N < N_{OV}$, where $N_{OV}$ is the Oppenheimer-Volkoff limit) is cooled down below a critical temperature, a condensed phase emerges and the gravitational collapse is prevented by quantum mechanics. If $N > N_{OV}$, by contrast, the system is destined to collapse towards a Black Hole because no equilibrium states exist.

*Keywords*: General Relativity; Thermodynamics; Fermions; Phase Transitions.

## 1. Introduction

The object of this work, intendend as a résumé of a previous publication[1], is to describe the occurrence of the gravitational phase transition that the Fermi gas at non-zero temperature experiences, in a general relativistic framework.

As it is known, gravitational phase transitions have been investigated since early '70s, when Hertel & Thirring[5] have shown the non-equivalence of statistical ensembles. Several steps forward have been made by Padmanabhan[7] and Chavanis[3] (for a detailed review see Ref. 4 and references therein) who have shown the occurrence of the phase transition in several systems (e.g. hard spheres,...).

In the case of fermions[3], when the system is cooled down below a critical temperature, a condensed phase emerges and the initial gaseous configuration evolves towards this condensed configuration.

Bilić & Viollier[2] firstly studied the occurrence of the gravitational phase transition in General Relativity (hereafter GR): however, they focused their attention only on a particular case. For this reason, in this work, we complete their previous investigations by describing the most general case.

The paper is organized as follows. In Sec. 2 we present the main equations. In Secs. 3 and 4 we discuss the occurrence of the phase transition in the canonical and microcanonical ensembles, respectively. In Sec. 5, finally, we draw some conclusions.

## 2. Theoretical framework

We consider a (static) self-gravitating[a] fermionic gas, formed by $N$ particles of mass $m$ at a temperature $T \neq 0$ and placed within a spherical box of dimension $R$. The equilibrium equations are given by

$$\frac{d\Phi}{dr} = -\frac{2G}{c^4}\frac{(\Phi+1)(M_r c^2 + 4\pi P r^3)}{r^2}\left(1 - \frac{2GM_r}{rc^2}\right)^{-1}, \tag{1a}$$

$$\frac{dM_r}{dr} = \frac{4\pi\epsilon r^2}{c^2}. \tag{1b}$$

with the conditions $\Phi(0) = \Phi_0$ and $M_r(0) = 0$. In the previous equations, $\Phi$ corresponds to the gravitational potential, $M_r$ represents the mass-energy contained within a sphere placed at distance $r < R$ from the center of the system ($\epsilon$ and $P$ are the mass-energy density and the pressure, respectively). The particle number $N$ is given by

$$N = N(\Phi_0) = \int_0^R 4\pi n r^2 \Big(1 - \frac{2GM_r}{rc^2}\Big)^{-1/2} dr\,, \tag{2}$$

being $n = \rho/m$ the particle number density ($\rho$ is the rest mass density). The thermodynamic analysis is performed by means of the caloric curve $T = T(E)$[b]. To this purpose, we define the following variables

$$\Lambda = -\frac{E_b R}{GN^2m^2} = \frac{(Nm - M)Rc^2}{GN^2m^2}\,, \qquad \eta = \frac{GNm^2}{k_B T R}\,. \tag{3}$$

## 3. Canonical instabilities

In this Section we discuss the case of the canonical ensemble. In Fig. 1 we have represented the occurrence of the phase transition for $R = 15\,R_{OV}$ and $N = 0.7277\,N_{OV}$ (left panel) and $R = 15\,R_{OV}$ and $N = 1.0012\,N_{OV}$[c] (right panel).

In both cases the system exhibits a phase transition because we observe the coexistence, at the same (transition) temperature ($\eta_t = 1.5722$ in the first case and $\eta_t = 1.3315$ in the second one), of several states identified by the points $P_1$, $P_2$, $P_3$ and $P_4$. The plots present different colors according to the stability of the solutions (black and green lines corresponding to stable states, red lines corresponding to unstable states). Let us consider the case $N < N_{OV}$ first.

[a] In this work we neglect the contribution of the other interactions that, in a more realistic situation, fermions certainly feel.

[b] For the details concerning the numerical procedure used to get the caloric curve see Ref. 1.

[c] $R_{OV}$ and $N_{OV}$ are the values of the radius and of the particle number at the Oppenheimer-Volkoff limit[6]. Numerically we have $R_{OV} = 9.162$ km and $N_{OV} = 8.752 \times 10^{56}$ particles.

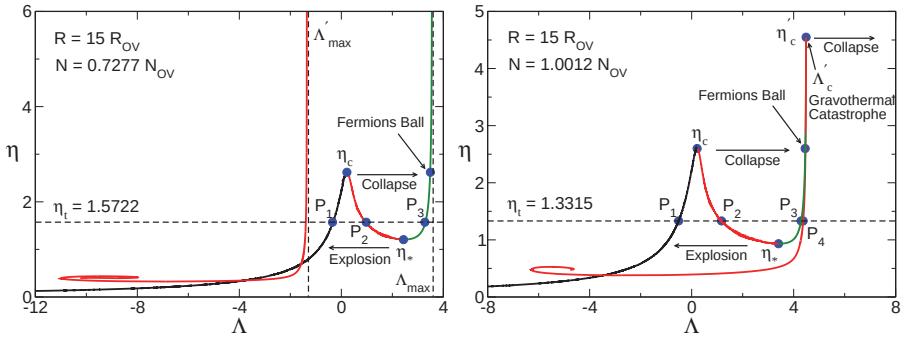



Fig. 1. Equilibrium phase diagrams for fermionic systems in GR. $\eta_t$ corresponds to the transition temperature. **Left Panel**: At the temperature $\eta_c$ the system evolves from the gaseous (black line) to the condensed phase (green line). The gravitational collapse is prevented by Pauli's exclusion principle. At the temperature $\eta_*$, the system evolves from the condensed phase to the gaseous one. The gravitational explosion is halted by the box. **Right Panel**: The condensed phase collapses towards a Black Hole at the temperature $\eta_c'$.

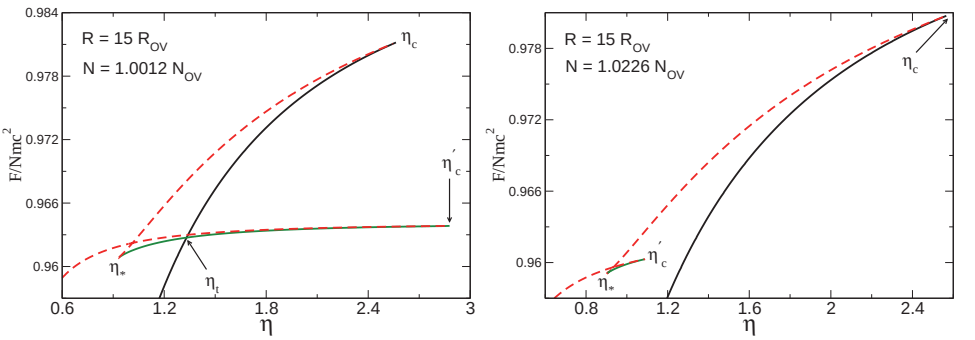



Fig. 2. Free energy as a function of the normalized inverse temperature $\eta$. **Left Panel**: The phase transition is identified by the intersection point between the gaseous (black full line) and the condensed branch (green full line). **Right Panel**: The intersection point between the gaseous and the condensed phase is missing: the phase transition is thus suppressed (the crossing point concerns unstable states).

Points $P_1$ represent the gaseous phase (black line): they are global minima of free energy (thus stable) for $\eta < \eta_t$ and local minima of free energy (thus metastable[d]) for $\eta > \eta_t$. When the temperature exceeds the critical value $\eta_c$, the gaseous phase undergoes a collapse and forms a compact object (condensed phase, green line, points $P_3$) containing all the mass. The stability of the solutions $P_3$ is inverted with respect to that of the solutions $P_1$. Consequently, the compact object represents

[d]The lifetime of metastable states can be longer than the age of the Universe. This implies that the phase transition does not occur in practice. However, metastable states can play a central role in astrophysics[4].

a stable configuration (it is a global minimum of free energy). We refer to this compact object as "fermion ball".

The condensed phase, similar to the gaseous one, evolves too. For $\eta < \eta_*$, indeed, the condensed undergoes an explosion (halted by the box). Points $P_2$, by contrast, correspond to unstable physical solutions (they are saddle points). Moreover, in this region of the diagram, the specific heat is negative (because $d\eta/d\Lambda < 0$).

Let us now turn our attention to the case $N > N_{OV}$ (right panel of Fig. 1). The difference, with respect to the former case, is represented by the presence of a second collapse (of general relativistic origin) at the temperature $\eta'_c$. In this case, the condensed phase undergoes a second collapse towards a Black Hole.

The temperature $\eta'_c$, which is a decreasing function of $N$, allows un upper limit for the extension of the condensed phase. The theory identifies a critical value of $N$, namely that corresponding to the extinction of the condensed phase and to the suppression of the phase transition[1].

The reason of this phenomenon is shown in Fig. 2. The left panel plots the free energy as a function of the (reverse) normalized temperature, for $N = 1.0012\, N_{OV}$. The phase transition is identified by the intersection point between the gaseous and the condensed phase. The right panel shows the case $N = 1.0226\, N_{OV}$. As we see, the plot does not display any intersection point between the two phases, implying that the phase transition is suppressed.

## 4. Microcanonical instabilities

Let us now discuss the microcanonical ensemble. In Fig. 3 we study the occurrence of the phase transition for $R = 179\, R_{OV}$ and $N = 3.2620\, N_{OV}$ (left panel) and $R = 179\, R_{OV}$ and $N = 4.2657\, N_{OV}$ (right panel). Similar to the canonical ensemble, the left panel of Fig. 3 shows the occurrence of the phase transition because of the coexistence, at the same (transition) energy $\Lambda_t = -0.025$, of several states identified by the points $P_1$, $P_2$, $P_3$ and $P_4$.

Points $P_1$ represent the gaseous phase (black line): they are global entropy maxima (thus stable) for $\Lambda < \Lambda_t$ and local entropy maxima (thus metastable) for $\Lambda > \Lambda_t$. When the energy exceeds the critical value $\Lambda_c$, the gaseous phase undergoes a collapse and forms a compact object (condensed phase, green line, points $P_3$) containing a fraction ($\sim 1/4$) of the total mass and surrounded by an atmosphere.

The stability of the solutions $P_3$ is inverted with respect to that of the solutions $P_1$. Consequently, the compact object represents a stable configuration (it is a global entropy maximum). We refer to this compact object as "fermion ball".

The condensed phase, similar to the gaseous one, evolves too. For $\Lambda < \Lambda_*$, it undergoes an explosion (halted by the box). Analogous to the canonical ensemble, points $P_2$ are unstable saddle points. In this region of the diagram the specific heat is negative (because $d\eta/d\Lambda < 0$).

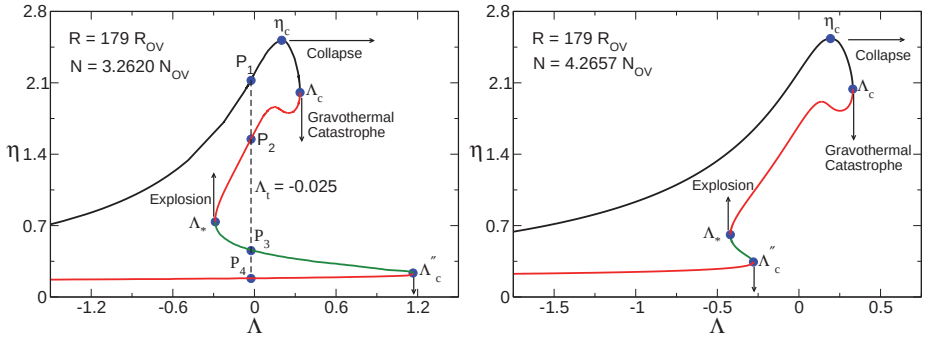



Fig. 3. Equilibrium phase diagrams for fermionic systems in GR. **Left Panel**: A phase transition from the gaseous to the condensed phase occurs at the energy $\Lambda_t$. At the energy $\Lambda_c$ the system evolves from the gaseous to the condensed phase. The gravitational collapse is prevented by quantum degeneracy. At the energy $\Lambda_*$ the system evolves from the condensed to the gaseous phase. The gravitational explosion is halted by the box. At the energy $\Lambda_c''$ the system collapses towards a Black Hole. **Right Panel**: The phase transition is suppressed.

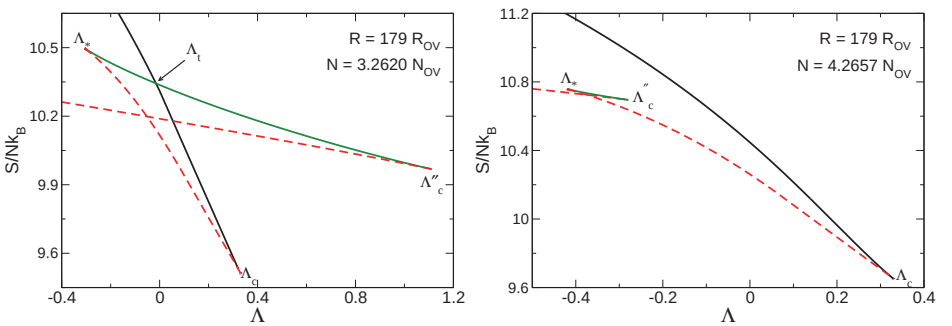



Fig. 4. Entropy as a function of the normalized energy $\Lambda$. **Left Panel**: The phase transition is identified by the intersection point between the gaseous (black full line) and the condensed branch (green full line). **Right Panel**: The intersection point between the gaseous and the condensed phase is missing, so the phase transition is suppressed (the crossing point concerns unstable states).

The novelty with respect to the non-relativistic regime[3] is the presence of a second collapse, occurring at the critical energy $\Lambda_c''$. The existence of this critical energy allows an upper limit to the extension of the condensed phase and, as a consequence, the suppression of the phase transition[1].

To better understand this phenomenon we have represented, in Fig. 4, the entropy as a function of the normalized energy $\Lambda$. The left panel shows the occurrence of the phase transition, because of the presence of a crossing point between the gaseous and the condensed phase. The right panel of Fig. 4, by contrast, does not display any crossing point between the two phases and the phase transition is thus suppressed (Fig. 3 plots the caloric curve associated).

## 5. Concluding remarks

In this work we have studied the nature of phase transitions of the Fermi gas in a general relativistic framework. The model takes both quantum mechanics and GR into account. As we have seen, the occurrence of the phase transition depends on the values of $N$ and $R$.

In the main paper[1] we have determined two critical values of the cavity radius, namely the canonical critical radius $R_{CCP} = 3.57\, R_{OV}$ and the microcanonical critical radius $R_{MCP} = 27.4\, R_{OV}$. If $R < R_{CCP}$, the system does not experience any phase transition whereas, if $R_{CCP} \leq R < R_{MCP}$, the system experiences the canonical phase transition. If $R \geq R_{MCP}$, both types of phase transition occur.

In Sec. 3 we have considered the case of the canonical phase transition. We have seen that, for $N < N_{OV}$, the result is similar to that obtained in the non-relativistic regime[3]. The result of the phase transition is, indeed, the formation of a compact object containing all the mass of the initial configuration. Things change when we consider the case $N \geq N_{OV}$, because the system exhibits a second collapse at the temperature $\eta_c'$. The result of this collapse is a Black Hole. However, if the value of $N$ is above a critical one (see Ref. 1), the phase transition is suppressed.

In Sec. 4 we have considered the case of the microcanonical phase transition. Similar to the canonical ensemble, the system exhibits a second collapse (at the energy $\Lambda_c''$) towards a Black Hole. Analogous to the canonical ensemble, we observe the phenomenon of the suppression of the phase transition.

The results obtained in this work call for further investigations. For example, the microcanonial phase transitions may be related to the onset of red-giant structure or to the Supernova phenomenon. Furthermore, to check the robustness of the model and have a first "experimental" proof of the occurrence of the phase transition, it would be interesting to perform $N$-body simulations in GR. This could represent a new challenge for Numerical Relativity.

## References

1. G. Alberti & P.H. Chavanis, arXiv:1808.01007 (2018)
2. N. Bilić & R.D. Viollier, Eur. Phys. J. C **11**, 173 (1999)
3. P.H. Chavanis, Phys. Rev. E **65**, 056123 (2002)
4. P.H. Chavanis, Int. J. Mod. Phys. B **20**, 3113 (2006)
5. P. Hertel & W. Thirring, Commun. Math. Phys. **24**, 22 (1971)
6. J.R. Oppenheimer & G.M. Volkoff, Phys. Rev. **55**, 374 (1939)
7. T. Padmanabhan, Phys. Rep. **188**, 285 (1990)

# The Schrödinger-Poisson equations as the $N$-body double of dark matter dynamics at large scales

Fabio Briscese

*SUSTech Academy for Advanced Interdisciplinary Studies, Southern University of Science and Technology, Shenzhen 518055, China, and*
*Istituto Nazionale di Alta Matematica Francesco Severi, Gruppo Nazionale di Fisica Matematica, Città Universitaria, P.le A. Moro 5, 00185 Rome, Italy*
*E-mail: briscese.phys@gmail.com, briscesef@sustech.edu.cn*

Nicolò Burzillà and Andrea Dosi

*Departement of Physics, Southern University of Science and Technology, Shenzhen 518055, China*

The formation of large scale structures in the universe can be described in terms of a Schrödinger-Poisson system. This fact is justified showing that the Schrödinger-Poisson equations can be obtained as the large $N$ limit of a Newtonian $N$-body system, that, in the cosmological context, describes the evolution of dark matter halos. This result is based on the Nelson stochastic quantization, and on the Calogero conjecture. The relevance of this finding for the study of LSS is discussed.

*Keywords*: Schrödinger-Poisson equations; large tructure formation; dark matter; cosmology, $N$-body simulations.

## 1. Introduction

It is widely accepted tat the formation of large scale structures (LSS) in the universe, as superclusters, sheets and filaments[1], is shaped by collisionless Dark Matter[2] (DM), see[3] for an historical review. In the standard $\Lambda$-CDM cosmological model[4], DM is assumed to be constituted of unknown particle species that interact (almost) only gravitationally, and DM is described as a cold fluid at cosmological scales. In facts, Cold Dark Matter (CDM) is in agreement with all cosmological data, including LSS[1], CMB[5], leansing[6], BAO[7], and Supernovae[8]. Even though DM particles are still elusive, there are currently many DM candidates, and the search for DM particles is an open issue[9].

At cosmological scales above $100\,Mpc$, CDM is successfully described as a pressureless dust fluid; but this approximation fails at smaller scales. Since, at such smaller scales, the typical velocities are non-relativistic, the Newtonian limit of the Einstein equations is sufficient to describe the time evolution of massive bodies within the universe. Indeed, CDM can be described as a classical $N$-body system, in which the individual particles represent huge agglomerates of mass $M$ of elementary dark matter particles of mass $m \ll M$.

The first and most natural way to treat this $N$-body system is to resort to numerical simulations[10–14]. Such simulations show that, if the mass $M$ is much smaller than the mass scale of the structures that we want to characterize, the exact value of $M$ does not affect the LSS formation. For instance, the MILLENNIUM

simulation[10] was carried out tracing the evolution of $N \sim 10^{10}$ identical particles of mass $M \simeq 10^9 M_\odot$, where $M_\odot \simeq 2 \times 10^{30} Kg$ is the solar mass. Although the choice $M \simeq 10^9 M_\odot$ in the $N$-body simulations is fit for purpose, it is still arbitrary. However, an indication on the plausibility of this value for DM halos comes from scalar field DM models[15].

In alternative, one can describe the $N$-body dynamics statistically, by means of the phase space distribution of the bodies $f(t, x, p)$, where the evolution of $f(t, x, p)$ is given by the Boltzmann equation. Furthermore, in the case of LSS formation, $N$ is large and collisions are suppressed, so that the Boltzmann equation reduces to a Vlasov equation[16].

## 2. The Schrödinger method

The so called Schrödinger method (ScM) has been proposed[17–32] as a numerical double of the $N$-body system describing the dynamics of CDM at large cosmological scales. The ScM is based on the hypothesis that, it is possible to associate to the $N$-body system a wave function $\psi$, such that the mass density of the system is given by $\rho = M\, N\, |\psi|^2$. Such a wave function $\psi$ obeys the coupled Schrödinger-Poisson equations (SPEs)

$$\begin{aligned} i\hbar \partial_t \psi &= -\tfrac{\hbar^2}{2a(t)M} \Delta\psi + MV\,\psi\,, \\ \Delta V &= 4\pi G\,\rho\,, \end{aligned} \tag{1}$$

where $V$ is the Newtonian potential, $\rho$ is the DM density, $M$ is the mass of the DM halos, $a(t)$ is a scale factor introduced to take into account the expansion of the universe, and $\hbar$ is a free parameter. The Newtonian potential $V$ is determined through the Poisson equation in (1). The form of the DM density used to run cosmological simulations in[17–32] is $\rho = (M\,|\psi|^2 - \rho_{crit})/a(t)$, where $\rho_{crit}$ is a parameter representing a *comoving critical density of the universe*, although some authors assume $\rho_{crit} = 0$, so that the SPEs reduce to the Schrödinger-Newton equation[33]. However, since we are not interested in discussing the explicit form of the DM density, we will use the generic expression $\rho$. Furthermore, here we are only focused on the validity of the ScM approach, and this is not related to the expansion of the universe, thus hereafter we set $a(t) = 1$.

The SPEs can be viewed as the non-relativistic limit of the Klein-Gordon and Dirac equations. In this case $\hbar$ coincides with the Planck constant $\sim 10^{-34} J\,s$, and numerical solutions of the SPEs point towards ultralight DM particles of mass $m \sim 10^{-23} eV$. However, when the SPEs is interpreted as a numerical double of the $N$-body system of DM halos, $M$ is huge and $\hbar$ is merely a free parameter chosen at will, that determines the phase space resolution in the ScM. To ensure the match with $N$-body simulations one must require $\hbar/M \sim 10^{-4} Mpc \cdot c$, so that $\hbar$ is huge comparison with the Planck constant.

### 2.1. *The Nelson stochastic quantization and the Calogero conjecture*

The agreement between the numerical simulations of DM dynamics based on the $N$-body Newtonian systems (when the bodies have the mass of DM halos, and $\hbar$ is a free parameter) and those based on the SPEs, is usually accepted as an empirical fact[17–32]. However, a deeper view on the relation between the ScM and the $N$-body DM system has been recently proposed in[34], where it has been argued that the SPEs can be obtained as the large $N$ limit of the Newtonian $N$-body system. This result is based on the so called Nelson stochastic quantization and on the Calogero conjecture, as resumed below.

To begin with, let us recall briefly the basis of the Nelson stochastic quantization[35]. Let us consider the classical trajectory of a particle by

$$M\,\ddot{\vec{x}} = -\vec{\nabla}\phi + \vec{B}(t) \tag{2}$$

where $-\vec{\nabla}\phi$ represents all the conservative forces, and $\vec{B}(t)$ is a random force with zero mean, representing a small random noise.

Nelson has shown[36] that, under these hypothesis, the motion of the particle can be described by means of a stochastic process, and the probability distribution $f(x)$ of the particle can be expressed as $f(x) = |\psi|^2$, in terms of a wave function $\psi$ satisfying the Schrödinger equation

$$i\hbar\partial_t\psi = -\tfrac{\hbar^2}{2M}\Delta\psi + \phi\,\psi\,. \tag{3}$$

In this picture the quantum behavior of the dynamics of the particle is induced by the random field $\vec{B}(t)$.

At this point, one can question about the nature of the random force $\vec{B}(t)$. One possibility, conjectured by Calogero[37], is that, in an $N$-body system with Newtonian interactions, $\vec{B}(t)$ is the resultant of the gravitational interactions of each particle with all the other particles, and its random behavior comes from the fact that the classical dynamics of a $N$-body system is chaotic. Exploiting this idea, Calogero has shown that the order of magnitude of the induced Planck constant is[37]

$$\hbar \sim M^{5/3}G^{1/2}(N/ < \rho >)^{1/6}\,, \tag{4}$$

where $N$ and $M$ are the number and the mass of the bodies, $< \rho >$ is the average density of the system, and $G$ the gravitational constant.

Indeed, due to the mechanism conjectured by Calogero, any particle in the $N$-body system undergoes a stochastic gravitational noise due to the classical gravitational interaction with all the other particles, which plays the role of the stochastic random noise $\vec{B}(t)$ in (2). Thus, the dynamics of each particle of the system can be described in terms of a wavefunction $\psi$ solution of (3), where $\hbar$ is given by (4). At that point, one can express the wavefunction of the entire system using the Hartree-Fock approximation, so that the number density of the $N$-body system will be $n(x) = N\,|\psi|^2$. This is the analog of the derivation of the famous Gross-Pitaevskii

equation[38] for a Bose-Einstein condensate by means of the Hartree-Fock approximation, see[39]. Finally, the potential in (3) is $\phi = M\,V$, where $V$ is the gravitational potential solution of the Poisson equation $\Delta V = 4\pi G\rho$, where $\rho = M\,n(x)$. Therefore, based on these arguments, one understands why the dynamics of an $N$-body system in the large N limit is well described by the SPEs (1).

Since the dynamics of DM at large scales is described by huge agglomerates of DM halos interacting via Newtonian potentials, the hypothesis at the basis of the Nelson quantization and of the Calogero conjecture are satisfied. This implies that the evolution of DM can be described statistically by means of the SPEs. The stochastic background responsible for the Nelson quantization is given, as in the Calogero conjecture, by the gravitational interaction between the DM halos, and its stochastic character is due to the chaotic behavior of the $N$-body dynamics. This explains the equivalence of numerical simulations of DM dynamics based on SPEs and $N$-body systems.

What is more, the Calogero conjecture also allows to estimate the order of magnitude of the effective Planck constant in (1). In facts, there is perfect agreement between the empirical value of $\hbar$ used in SPEs, and that obtained by Calogero estimate (4). For instance, in the case of the Millennium simulation[10], the $N$-body problem is solved for $N \simeq 10^{10}$ particles of mass $M \simeq 10^9 M_\odot$ where $M_\odot \simeq 2 \times 10^{30} Kg$ is the solar mass. Using $< \rho > \simeq 3H_0^2/8\pi G \simeq 4 \times 10^{-26} Kg/m^3$, where $H_0 \simeq h^{-1} \times 100 \times km/sMpc$, and $h \simeq 0.73$ is the Hubble constant, equation (4) gives $\hbar \simeq 2 \times 10^{66} Kg\,m^2/s$. This value corresponds to a ratio $\hbar/M \simeq 10^{-4} Mpc\,c$ in the range of values used in numerical simulations, e.g. $\hbar/M \sim 10^{-4} Mpc\,c$ in[28,30] or $\hbar/M \sim 10^{-6} Mpc\,c$ in[29,31].

## 3. Conclusions

In conclusion, the validity of the Schrödinger method as $N$-body double stays in the fact that the $N$-body dynamics is well described in terms of the SPEs in the large N limit. This is due to the the stochastic quantization induced by the random gravitational background produced by the N bodies, as in the Calogero conjecture. When applied to LSS formation, this gives a further argument in support of the validity of the Schrödinger method for the DM dynamics at cosmological scales. Moreover, this derivation of the SPEs is the first known realistic application of the Nelson quantization and of the Calogero conjecture.

## References

1. V. Springel, *et al.*, Nature (London) **435**, 629 (2005), astro-ph/0504097. E. Tempel,*et al.*, Mon. Not. R. Astron. Soc. **438**, 3465 (2014), arXiv:1308.2533 [astro-ph.CO]; M. Tegmark *et al.* [SDSS Collaboration], Phys. Rev. **D 69**, 103501 (2004) [arXiv:astro-ph/0310723]; U. Seljak *et al.* [SDSS Collaboration], Phys. Rev. **D 71**, 103515 (2005) [arXiv:astro-ph/0407372].

2. T. J. Sumner, Living Rev. Relativity **5** (2002), 4; G. Bertone, D. Hooper, J. Silk, Phys. Rept. **405** (2005) 279-390.
3. G. Bertone, D. Hooper, arXiv:1605.04909 [astro-ph.CO].
4. V. Mukhanov, *Physical Foundations of Cosmology*, Cambridge University Press, 2005.
5. Planck 2015 results, XIII, Cosmological parameters, Planck Collaboration: P. A. R. Ade *et al.*, arXiv:1502.01589; Planck 2015, XX, Constraints on inflation, Planck Collaboration: P. A. R. Ade *et al.*, arXiv:1502.02114.
6. B. Jain and A. Taylor, Phys. Rev. Lett. **91**, 141302 (2003) [arXiv:astro-ph/0306046].
7. D. J. Eisenstein *et al.* [SDSS Collaboration], Astrophys. J. **633**, 560 (2005) [arXiv:astro-ph/0501171].
8. S. Perlmutter *et al.* [SNCP Collaboration], Astrophys. J. **517**, 565 (1999) [arXiv:astro-ph/9812133]; A. G. Riess *et al.* [Supernova Search Team Collaboration], Astron. J. **116**, 1009 (1998) [arXiv:astro-ph/9805201].
9. F. S. Queiroz, arXiv:1605.08788 [hep-ph].
10. V. Springel *et al.* Nature, 435, (2005) 629; V. Springel, C. S. Frenk, S. D. M. White, Nature (London) **440**, 1137 (2006), astro-ph/0604561.
11. R. Teyssier, Astron. Astrophys. **385**, 337 (2002), astroph /0111367.
12. M. Boylan-Kolchin, *et al.* Mon. Not. R. Astron. Soc. 398, 1150 (2009), arXiv:0903.3041 [astro-ph.CO].
13. T. Abel, O. Hahn, R. Kaehler, Mon. Not. R. Astron. Soc. **427**, 61 (2012), arXiv:1111.3944 [astro-ph.CO].
14. O. Hahn, T. Abel, R. Kaehler, Mon. Not. R. Astron. Soc. **434**, 1171 (2013), arXiv:1210.6652 [astro-ph.CO].
15. A. Suarez, V. H. Robles, T. Matos, Astrophys. Space Sci. Proc. **38** (2014) 107-142, arXiv:1302.0903 [astro-ph.CO].
16. H. Andreasson, Living Rev. Relativity **14** (2011), 4.
17. L. M. Widrow, N. Kaiser, Astrophys. J. Letters **416**, L71 (1993).
18. G. Davies, L. M. Widrow, astro-ph/9607133.
19. P. Coles and K. Spencer, Mon. Not. Roy. Astron. Soc. **342**, 176 (2003)
20. C. J. Short, P. Coles, J. Cosmol. Astropart. Phys. **12**, 012 (2006).
21. D. Giulini, A. Groardt, Class. and Quant. Grav. **29**, 215010 (2012).
22. E. R. Arriola, J. Soler, J. Stat. Phys. **103**, 1069 (2001).
23. I. M. Moroz, R. Penrose, P. Tod, Class. and Quant. Grav. **15**, 2733 (1998).
24. R. Johnston, A. N. Lasenby, M. P. Hobson, arXiv:0904.0611 [astro-ph.CO].
25. I. Szapudi and N. Kaiser, Astrophys. J. Letters **583**, L1 (2003).
26. M. Schaller, *et al.*, Monthly Notices of the Royal Astronomical Society **442**, 4. 3073-3095, arXiv:1310.5102 [astro-ph.CO].
27. H.-Y. Schive, T. Chiueh, T. Broadhurst, Nature Phys. **10**, (2014) 496-499; H.-Y. Schive, T. Chiueh, T. Broadhurst, Astrophys. J. 818 (2016) no.1, 89.
28. C. Uhlemann, M. Kopp, T. Haugg, Phys. Rev. D 90, 023517 (2014).

29. J. Zhang, Y.-L. S. Tsai, K. Cheung, M.-C. Chu, arXiv:1611.00892.
30. A. Paredes, H. Michinel, Phys. of the Dark Univ. **12**, 50-55.
31. H.-Y. Schive, T. Chiueh, T. Broadhurst, Nature Physics **10**, 496–499 (2014).
32. L. A. Urena-Lopez, Phys.Rev. D **90** (2014) no.2, 027306, arXiv:1310.8601 [astro-ph.CO].
33. R. Ruffini, S. Bonazzola, Silvano, Phys. Rev. **187**, 5 (1969), 1767–1783.
34. F. Briscese, Eur. Phys. J. C **77** (2017) no.9, 623.
35. L. de la Peña, A. M. Cetto, A. Valdes Hernandez, *The Emerging Quantum*, Springer (2015).
36. E. Nelson, Phys. Rev 150, 4 (1966).
37. F. Calogero, Phys Lett. A **228** (1997) 335-346; F. Calogero, Int. J. Mod. Phys. B **18**, Nos 4 & 5 (2004) 519-525.
38. F. Dalfovo, S. Giorgini, L. P. Pitaevskii, and S. Stringari, Rev. Mod. Phys. **71**, 463 (1999).
39. Lev Pitaevskii and Sandro Stringari, *Bose-Einstein Condensation* (Clarendon Press, Oxford, 2003).

# Self–gravitating particles, entropy, and structure formation

Andrew J. Wren*

*London, United Kingdom*
* *E-mail: andrew.wren@ntlworld.com*

I outline a kinetic theory model of gravitational collapse due to a small perturbation. This model produces a pattern of entropy destruction in a spherical core around the perturbation, and entropy creation in a surrounding halo. Core–halo patterns are ubiquitous in the astrophysics of gravitational collapse, and are found here without any of the prior assumptions of such a pattern usually made in analytical models. Motivated by this analysis, I outline a possible scheme for identifying structure formation via data from observations or a simulation. This might aid exploration of hierarchical structure formation, supplementing the usual density–based methods for highlighting astrophysical and cosmological structure at various scales.

*Keywords*: Kinetic theory; Gravitational collapse; Entropy; Structure formation.

## 1. Introduction and Motivation

Many astrophysical contexts see gravitational collapse leading to structure formation. A simple model of gravitational collapse can be constructed based on the virial theorem, with an artificial division between a central core and surrounding halo.[1] Entropy rises within the halo, with at least a relative entropy decrease in the shrinking core. If we further assume that the core's density profile scales with its radius $R$, then its phase space volume varies like $R^{3/2}$ and it is easily seen that there is an absolute fall in entropy within the core.

Note that there may, at least in principle, be two distinguishable entropy–related effects: entropy transport from one volume element to another; and entropy creation/destruction. The distinction is somewhat arbitrary in the artificial virial theorem model just discussed. In kinetic theory treatments however, entropy creation is separately identifiable, and arises only from collisional effects.

This note is based on a full account in Ref. 2. I outline how to construct a kinetic theory model of gravitational collapse allowing analytical description of entropy creation (Sec. 2). A core–halo pattern emerges as a result (Sec. 3). This suggests an approach to identifying structure in simulations and observations (Sec. 4).

## 2. Outline of the Model

The model consists of a small central perturbation to an underlying uniform distribution of self–gravitating particles. Via the well–known "Jeans swindle," a system with a finite number of particles $N$ in a bounded volume can be used to model a uniform arrangement of particles which is unbounded in extent, and so stable if unperturbed.[1]

The perturbation evolves under truncated zeroth and first order BBGKY equations for the evolution of the distribution function (DF) and the correlation function. Write $f_0, f_1$ for the underlying and perturbation DFs, similarly $g_0, g_1$ for the

correlation functions, and set $\mathbf{a}_{1,2}$ to be $N$ times the acceleration of a particle at $1 = (\mathbf{x}_1, \mathbf{v}_1)$ due to a particle at $2 = (\mathbf{x}_2, \mathbf{v}_2)$. The full first order equation for the DF's evolution is[3]

$$\frac{\partial f_1(1)}{\partial t} + \mathbf{v}_1 \cdot \frac{\partial f_1(1)}{\partial \mathbf{x}_1} + \int \mathbf{a}_{1,2}\, f_1(2)\, \mathrm{d}(2) \cdot \frac{\partial f_0(1)}{\partial \mathbf{v}_1} = -\frac{1}{N} \frac{\partial}{\partial \mathbf{v}_1} \int \mathbf{a}_{1,2}\, g_1(1,2)\, \mathrm{d}(2), \quad (1)$$

where the right–hand side gives the "collisional" effects of two–body interactions.

Both $f_0$ and $f_1$ are assumed to have initial Maxwellian velocity distributions, with parameter $\sigma$, and a typical scale for the system is then given[1] by the "Jeans wave number," $k_\mathrm{J} = \sqrt{4\pi G m n}/\sigma$, where $m$ is the (identical) mass of each particle, $n$ is the average number density, and $G$ is Newton's gravitational constant. The model is valid for the beginning of gravitational collapse — the initial period during which the perturbation remains small.

The entropy creation rate comes entirely from collisional effects, and its density at $\mathbf{x}_1$ is

$$\left( \frac{\partial S_{\mathbf{x}_1}}{\partial t} \right)_{\text{creation}} = \int \ln\left[ f(1) \right] \frac{\partial}{\partial \mathbf{v}_1} \cdot \int \mathbf{a}_{1,2}\, g(1,2)\, \mathrm{d}(2)\, \mathrm{d}^3\mathbf{v}_1, \quad (2)$$

where $f = f_0 + f_1$ and $g = g_0 + g_1$. It is useful to coarse grain the entropy, both because it enables progress to be made with analytical calculations, and because identification of structure typically implies focusing on only a range of scales.

The approach of Ref. 2 coarse grains by choosing a parameter $\beta \ll 1$, and then in applying Eq. (2) considering only wave numbers $k < \beta k_\mathrm{J}$. Furthermore, Ref. 2 focuses on only the "asymptotically dominant" part of the DF, correlation, and entropy creation — the part that soon comes to grow fastest.

## 3. Result: a Core–Halo Pattern

I briefly describe the approach of Ref. 2 to calculating the resulting asymptotically–dominant coarse–grained (acg) entropy creation rate, and in particular its dependence on the distance $r$ from the initial central perturbation. Let $S^\circ_\mathrm{acg}$ be the acg entropy created within a sphere of radius $r$ in the time $t$ since the initial perturbation was introduced. For large $N$, Eq. (1)'s collisional term is highly suppressed, and, at leading order in $1/N$, we can ignore it in calculating the contribution of the DF to the first order correlation equation. The resulting equations can then be solved to give, to leading order in $1/N$, perturbation size $\epsilon$, and $\beta$,

$$\frac{\partial^2 S^\circ_\mathrm{acg}}{\partial t\, \partial r} = \frac{2 k_\mathrm{J}^2\, \sigma\, \mathrm{e}^{3 k_\mathrm{J} \sigma t}}{9\pi^2 \beta} \left( \frac{\epsilon N}{n B} \right)^2 \hat{S}^\circ_\mathrm{acg}(k_\mathrm{J} \beta r)\,, \quad (3)$$

with $B = 4\pi/(k_\mathrm{J}\beta)^3$ being the volume of a sphere associated with the coarse–graining scale, and $\hat{S}^\circ_\mathrm{acg}$ is a function shown in Fig. 1. Note that Eq. (3) shows the entropy creation rate's "shell density", expressing the entropy creation on a shell of radius $r$ around the initial perturbation.

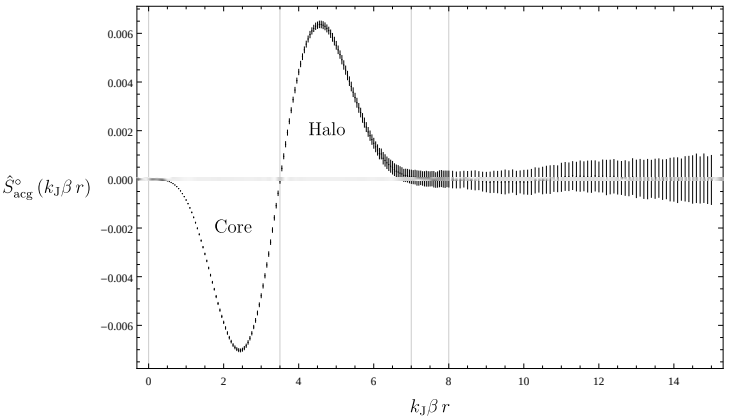



Fig. 1. The entropy–creation pattern function $S^{\circ}_{\mathrm{acg}}$, calculated by numerical integration of an analytical expression. The error bars show integration error estimates. Adapted from fig. 2 of A core–halo pattern of entropy creation in gravitational collapse, A. J. Wren, MNRAS, **477** (2018).

Entropy destruction occurs in a "core" around the central perturbation, with, at leading order, equal (see Ref. 2) and opposite entropy creation in a "halo" extending for a finite radius beyond that core, as shown in Fig. 1. The physical scale for the core–halo pattern depends on the coarse–graining parameter $\beta$: the coarser the graining, the bigger the pattern's physical scale. Entropy destruction (resp. creation) corresponds to collisional relaxation (resp. "de–relaxation") of the perturbation.

A core–halo pattern of gravitational collapse, well known from simulations and observations, is generally set "by hand" in analytical models. As far as the author has been able to determine, this is the first time an analytical kinetic theory model has produced a core–halo pattern.

## 4. Structure Formation in Simulations and Observations

It is well known that the Universe has a multi–scale hierarchical structure, in which core–halo patterns are ubiquitous. The identification of observed or simulated astrophysical structure typically involves considering features of especially high or low densities, in physical space, or phase space. There is no unambiguous definition of structure in this context, which can result in different methods giving different results — for example, in identifying sub–haloes near the centre of dark matter haloes (see Ref. 4), in major halo mergers (see Ref. 5) and in classifying elements of the cosmic web (see Ref. 6). This suggests that complementary methods for identifying structure, or structure formation, may be helpful.

Given data from observations or a simulation, the above analysis suggests we might construct a coarse–grained particle DF and correlation function, which then could give the entropy creation rate from Eq. (2). The pattern of coarse–grained

entropy creation might then give a way to identify structure formation, presumably associated with volumes of relatively lower entropy creation.

A key step is extracting a correlation function from the data. There are various methods for doing this.[7] At least in the kinetic theory model of Sec. 2, the correlation between two particles is small compared with their joint DF. In calculating correlation functions from data, this would imply a need to identify the difference between two quantities of relatively similar size.[7]

This means that robust identification of entropy creation, and hence structure formation, may need rather precise data. The outlook for data of sufficient precision is perhaps encouraging with the development of ever more powerful computer simulations, and the availability of detailed phase space observations from, for example, *Gaia*.[8]

**Acknowledgements**

My thanks to the anonymous referee of Ref. 2 for suggesting more physical discussion should be included, which prompted me to develop Sec. 4's scheme.

**References**

1. J. Binney and S. Tremaine, *Galactic Dynamics: Second Edition* (Princeton University Press, 2008).
2. A. J. Wren, A core–halo pattern of entropy creation in gravitational collapse, *MNRAS* **477**, 3983 (2018).
3. I. H. Gilbert, Collisional Relaxation in Stellar Systems, *ApJ* **152**, p. 1043 (1968).
4. J. Onions, A. Knebe, F. R. Pearce, S. I. Muldrew, H. Lux, S. R. Knollmann, Y. Ascasibar, P. Behroozi, P. Elahi, J. Han, M. Maciejewski, M. E. Merchán, M. Neyrinck, A. N. Ruiz, M. A. Sgró, V. Springel and D. Tweed, Subhaloes going Notts: the subhalo-finder comparison project, *MNRAS* **423**, 1200 (2012).
5. P. Behroozi, A. Knebe, F. R. Pearce, P. Elahi, J. Han, H. Lux, Y.-Y. Mao, S. I. Muldrew, D. Potter and C. Srisawat, Major mergers going Notts: challenges for modern halo finders, *MNRAS* **454**, 3020 (2015).
6. N. I. Libeskind, R. van de Weygaert, M. Cautun, B. Falck, E. Tempel, T. Abel, M. Alpaslan, M. A. Aragón–Calvo, J. E. Forero–Romero, R. Gonzalez, S. Gottlöber, O. Hahn, W. A. Hellwing, Y. Hoffman, B. J. T. Jones, F. Kitaura, A. Knebe, S. Manti, M. Neyrinck, S. E. Nuza, N. Padilla, E. Platen, N. Ramachandra, A. Robotham, E. Saar, S. Shandarin, M. Steinmetz, R. S. Stoica, T. Sousbie and G. Yepes, Tracing the cosmic web, *MNRAS* **473**, 1195 (2018).
7. M. Kerscher, I. Szapudi and A. S. Szalay, A Comparison of Estimators for the Two–Point Correlation Function, *ApJ* **535**, L13 (2000).
8. Gaia Collaboration, Brown, A. G. A., Vallenari, A., Prusti, T., de Bruijne, J. H. J. and et al., Gaia Data Release 2. Summary of the contents and survey properties, *A&A* (2018).

# Teach Newton's theory as the simplest model for the nonlocal organization of Cartesian matter-extension

Bulyzhenkov, Igor E.

*Lebedev Physics Institute RAS, Levich Institute of Time Nature Studies, and Moscow Institute of Physics & Technology*
*Moscow, 119991, Russia*
*E-mail: ibw@sci.lebedev.ru*

Blinov, Sergey V.

*Moscow Institute of Physics & Technology*
*Dolgoprudny, Moscow region, 141701, Russia*
*E-mail: serj_blinov@mail.ru*

Contrary to Cartesian vortexes, Newtonian point masses do not have internal structures and inside thermal energies. The Hamilton kinetic energy demonstrates independent variables for ordered translations and Umov's inside heat-chaos. The sum of these indivisible energies obeys the Lorentz transformations, while their difference drives adaptive self-organization of mechanical systems. The simplified Newtonian model with one consolidated degree of freedom for energies of order and chaos can be replaced by the Cartesian world paradigm. The latter can be learned quantitatively in high and higher schools due to Umov's thermomechanics for inertial heat and de Broglie's thermodynamics of the isolated particle.

*Keywords*: Inside Variables; Adaptive Nonlocality; Chaos-Order Self-Organization.

## 1. Umov's inertial heat

It is almost impossible to understand the nonlocal laws of quantum mechanics when primary education is based on Newton's point masses in empty space, and not on continuous distributions of Cartesian matter-extensions. High school students should learn the basic principles of Descartes and Newton alternatives before they begin formal calculations in the simplest approximation of point mechanical bodies. At the very beginning it is important to understand that "no theory is completely correct", and "experiments can only falsify theories, not confirm them" [1]. In particular, the Newton model erroneously supports independent conservations of mass and energy, while Einstein's relativistic formalism requires only a single conservation of energy, including thermal parts. Teachers of physics should also not conceal Kuhn's metaphysical ideas, such as "Einstein's theory is not merely a more complex version of Newton's (point masses)" and "there is no coherent direction of ontological development".

Cartesian mechanics of 'matter-extension' was developed by Nikolay Umov [3] quantitatively for the slow transfer of continuous energy densities long before the Lorentz transformations and Special Relativity. In 1873, Umov assigned to each mass $m$ the inside energy variable $kmc^2$ and pioneered in the inertial properties of heat after light emission/absorption. The modern relativistic theory can specify the Umov kinematic coefficient $k$ as the Fitzgerald-Lorentz contraction

factor $\sqrt{1-\beta^2} \Rightarrow k$. That time Umov estimated this variable for nonrelativistic ($\beta^2 \equiv v^2/c^2 \ll 1$) energy flows as $1/2 \leq k = (1-0.5\beta^2) \leq 1$. Today it is clear that the Umov's inside energy corresponds to the Lagrange function $L = -mc^2\sqrt{1-\beta^2}$ in the full Hamilton energy $E \equiv -L + \mathbf{v}\cdot\mathbf{P} = mc^2/\sqrt{1-\beta^2}$. The latter does not vanish in the co-moving frame of references, $E_o \equiv mc^2 \neq 0$, and obeys the Lorentz transformation rules,

$$E_o = \frac{E - \mathbf{vP}}{\sqrt{1-\beta^2}}, \quad \mathbf{P}_o = \frac{\mathbf{P} - E\mathbf{v}}{\sqrt{1-\beta^2}} = 0. \tag{1}$$

Hence the Hamiltonian structure of kinetic energies,

$$E = E_o\sqrt{1-\beta^2} + \mathbf{vP} = E_o\sqrt{1-\beta^2} + \frac{\beta^2 E_o}{\sqrt{1-\beta^2}} \equiv \frac{E_o}{\sqrt{1-\beta^2}}, \tag{2}$$

has two virtual (immeasurable) parts of different origins. The first part, $E_o\sqrt{1-\beta^2}$, is the Lagrange (1788) - Umov (1873) energy of inside heat-chaos and the second part, $\mathbf{vP}$, is the Hamilton (1833) energy of ordered spatial translations. Changes of the sum of this kinetic energy can be measured in practice. Newtonian mechanics with one degree of freedom for the summary kinetic energy (2) is well formulated for direct tests and mathematical descriptions of 'how' bodies move. But mechanics without inside degrees of freedom never explains 'why' free bodies move (or gravitate) in their closed system and 'what' burned Hiroshima.

In Umovian mechanics of inertial heat, the inside energy variable $E_o\sqrt{1-\beta^2}$ decreases with increasing Hamilton energy of ordered translations $\mathbf{vP} = \beta^2 E_o/\sqrt{1-\beta^2}$. From our Cartesian point of view, every gravitating body tends to a relativistic equipartition of its kinetic energies of inside chaos and external translational order. The principle of equipartition of relativistic kinetic energies behind the observed mechanical motion [4] forms the non-Newtonian ontology of thermomechanics and frees Einstein's adaptive dynamics from the thermal death problem.

Internal heat or latent degree of kinematic freedom cannot be reasonably introduced by the point mass model. Consequently, the inside chaos variable fundamentally distinguishes the Cartesian paradigm of 'matter-extension' from the empty space concept of Newton [5,6]. Indeed, the point particle of Newton has only kinematic degrees of freedom for spatial translations, while the continuous Cartesian vortex contains elementary internal energy even in the co-moving system. Planck-Laue relativistic cooling $Q(\beta^2) = E_o\sqrt{1-\beta^2} \approx (mc^2 - mv^2/2 - mv^4/8c^2 - ...)$ of thermal energy $E_o \equiv Q_o$ (or inside inertial heat) and Hamilton heating $p_i v^i \equiv H(\beta^2) + L(\beta^2) = E_o/\sqrt{1-\beta^2} - E_o\sqrt{1-\beta^2} \approx (mv^2 + mv^4/2c^2 + ...)$ due to ordered translations reiterate together pro-Newtonian changes $\Delta[Q(\beta^2) + p_i v^i] = \Delta(mv^2/2 + 3mv^4/8c^2 + ...)$ of the summary kinetic energy at the low speed limit of Cartesian transport.

## 2. Cartesian interpretation of relativistic thermomechanics

The slow spatial motion of a Cartesian extended body is primarily the energy transfer of the volume integral, $Q_o(1-\beta^2/2)$, over ordered densities of Umov's

inside chaos. The concomitant fraction of the translational kinetic energy $Q_o\beta^2$ is considered a small correction of inside thermal chaos due to the ordered spatial movement of chaotic densities. The Cartesian transport of two competing energies from chaos and order formally corresponds to the Newtonian transport of constant mass $Q_o/2c^2$ plus one variable energy $Q_o\beta^2/2$. Newton had no idea about the inside degree of kinematic freedom and modeled measurable changes $[Q_o(1-\beta^2/2)+Q_o\beta^2]-Q_o = Q_o\beta^2/2$ of the total kinetic energy through one constant coefficient $Q_o/2c^2 \equiv m/2$.

It is important to teach that the kinematic cooling of Umov-Planck-Laue $Q/Q_o = \sqrt{1-\beta^2} \approx 1-\beta^2/2$ of mechanical bodies coincides with a decrease in the frequency-energy $\hbar\omega/\hbar\omega_o = \sqrt{1-\beta^2} \approx 1-\beta^2/2$ of a quantum particle in wave dynamics. This cooling originates from the kinematic time dilation in the Lorentz transformations. Therefore, Cartesian thermomechanics of non-empty space with the inside heat-chaos of Umov, relativistic dynamics of the rest mass-energy of Einstein, the particle-wave dualism of quantum mechanics and the thermodynamics of the isolated particle of de Broglie have the common temporological nature. Such coherent convergence of nonlocal Cartesian mechanics with quantum mechanics of elementary material distributions encourages high school teachers to discuss Umov's thermomechanics of inertial inside energies in parallel with the Newtonian simplifications of reality. High and higher schools should start teaching inside variables of the Hamilton energy in the available nonrelativistic[3] and relativistic[7] terms. Then students can be prepared to understand the spatial transport of quantum energy and accept the $\hbar$-randomness[8] of the nonlocal Cartesian world.

## References


1. K. Popper, *The Logic of Scientific Discovery* (Hutchinson & Co, London and New York, 1959).
2. T.S. Kuhn, *The Structure of Scientific Revolutions*, 1st edition (University of Chicago Press, Chicago, 1962).
3. N. Umov, *Beweg-Gleich. d. Energie in contin. Korpern. Zeitschriff d. Math. und Phys.*, **19** (Schmolich, 1874).
4. I.E. Bulyzhenkov, *Gravitational Attraction until Relativistic Equipartition of Internal and Translational Kinetic Energies. Astr. Space Sci.*, 363:39 (2018).
5. D. Garber, *Descartes's Metaphysical Physics* (University of Chicago Press, Chicago, 1992).
6. I.E. Bulyzhenkov, *Cartesian Material Space with Active-Passive Densities of Complex Charges and Yin-Yang Compensation of Energy Integrals. Galaxies* **6**(2), 60 (2018).
7. L. de Broglie, *The Thermodynamics of the Isolated Particle (or the hidden thermodynamics of particles)* (Gauthier-Villars, Paris, 1964).
8. L. de Broglie, *The Reinterpretation of Wave Mechanics. Foundations of Physics* Vol. **1**, No. 1 (1970).

# How history and philosophy of science can inform teaching and learning of general relativity in upper secondary school

Magdalena Kersting

*Department of Physics, University of Oslo,*
*Postboks 1048, Oslo, 0316, Norway*
**E-mail: magdalena.kersting@fys.uio.no*
*www.mn.uio.no/fysikk/english/people/aca/magdak*

Science educators have to move beyond traditional content-focused instruction to teach concepts of Einsteinian physics. This work presents a design-based research project that introduced general relativity (GR) to upper secondary school students in Norway. The educational approach invited students to explore the historical development and philosophical aspects of GR within a digital learning environment. Results based on focus group interviews show that students were particularly motivated by such an approach because it emphasised the cultural and social relevance of physics and it linked GR to their previous knowledge of physics. Employing history and philosophy of science in the service of physics education can serve as a successful approach to making GR more accessible to young learners.

*Keywords*: Physics Education; History & Philosophy of Science; General Relativity.

## 1. Context and Background

Few scientific discoveries have had a bigger impact on our understanding of the universe than GR. Einstein's revolutionary theory of gravity did not only herald a new scientific age but fostered a new interest in the philosophy of space and time as well. Yet, despite its scientific, philosophical, and technological importance, GR and other topics of Einsteinian Physics[1] lack from most school curricula around the world. This work presents a design-based research project that introduced GR to upper secondary physics students in Norway.[2] The educational approach invited students to explore the historic development and philosophic aspects of GR within a digital learning environment. Research suggests that history and philosophy of science (HPS) can be fruitful in teaching and learning of physics: HPS perspectives can foster awareness of the nature of science and they help to promote deeper understanding of scientific ideas.[3] This work aims to study how such approaches might foster understanding of and motivation for GR.

## 2. History and Philosophy in General Relativity Education

Science educators promote the use of HPS as a fruitful teaching strategy in learning domains where students have to fundamentally revise their understanding of space and time.[4] This strategy can be traced back to Einstein who was guided by the wish to understand the intrinsic value of physics when developing his theories of relativity:[5] The physicist wished for the human mind to understand the reality of nature. Indeed, by presenting a new view towards space, time, and gravity, GR invites students to view reality from a new philosophical viewpoint. This philosophic

perspective can be complemented with historic approaches to teaching physics. According to Einstein, knowledge of the history of physics is important because it emphasises connections between a new theory and its rich environment:

> To use a comparison, we could say that creating a new theory is not like destroying an old barn and erecting a skyscraper in its place. It is rather like climbing a mountain, gaining new and wider views, discovering unexpected connections between our starting point and its rich environment. But the point from which we started out still exists and can be seen, although it appears smaller and forms a tiny part of our broad view gained by the mastery of the obstacles on our adventurous way up.[6]

In the next sections, this work explores how HPS can inform teaching and learning of GR in upper secondary schools. This exploration is guided by the following research question: **How can a historic and philosophic approach to teaching and learning of Einsteinian Physics foster understanding of and motivation for GR?**

## 3. Theoretical Framework and Methodology

To bridge the gap between educational research and the reality of physics classrooms, methods of design-based research informed the design and development of a digital learning environment about GR.[7] The goal was to find workable ways of making GR accessible at the upper secondary school level. In line with the iterative nature of the design-based research framework, two rounds of field studies were conducted - one in spring 2016 and one in spring 2017. In total, eleven physics classes in five upper secondary schools in Norway participated in the trials. To gain insight into students' experiences with the digital learning resources, the author of this paper conducted seven semi-structured focus group interviews with 5 to 8 participants per group and 46 students (18-19 years) in total. The teachers chose the students to allow for a balance of gender and to include both stronger and weaker students. The interviews were based on an interview guide that included the design of the learning activities, the use of history and philosophy, and challenges and motivation in the learning domain of GR. The interviews were transcribed and analysed using thematic analysis.[8] The themes of the thematic analysis focused on the use of history and philosophy; the individual codes unpacked students' motivation, interests, and challenges.

## 4. Results and Design Examples

In the focus group interviews, students expressed their curiosity towards learning about GR at the high school level and without advanced calculations. Students had a general expectation that the topic would be difficult and they showed respect towards Einstein's achievement. At the same time, they were looking forward to

learning more about relativistic ideas such as gravitational waves that they had encountered in popular culture and in the news before. Generally, the findings show that students approved of an HPS approach in GR education. Four factors turned out be particularly motivating or insightful for the participating students:

(1) Einstein is a popular figure in contemporary culture and students mentioned repeatedly that they enjoyed learning more about Einstein. The learning environment presented Einstein as a real person who struggled to find a new theory of gravity. Being able to follow some of these struggles helped students to identify with Einstein. One student in the interview expressed that relating to Einstein's reasoning helped him to understand GR more qualitatively:

**Student**: *You got into the same reasoning that made Einstein and his friend, it's the same that got into the thoughts of them that resonate within me a bit as well. That you could use this knowledge in a somehow different way than to calculate.*

(2) Physics was presented as an important part of our cultural heritage with social relevance for our daily lives. Students liked to get a broader and more comprehensive perspective on physics and to make connections to the society as a whole. In particular, they appreciated to learn that physics is an ongoing endeavour and that physicists do not have answers to everything.

(3) Students learned that physics was a modern field that has more to offer than just textbook knowledge. Many students appreciated these insights into perspectives of modern researchers as the following excerpt shows:

**Interviewer**: *So do you feel you have learned something?*
**Student**: *Yes, I think it was really nice that you got, like, the perspective of recent scientists who made their new discoveries so that you get a bit more recent perspectives. Because it is a bit easy to think that physics is just somehow in the book at school.*

(4) Finally, the historic contextualisation helped students to link new ideas in GR to their knowledge of classical physics and to their previous experiences. Students appreciated to be able to follow the reasoning that led to GR; this presentation helped them to understand why a theory actually works as the following excerpt from an interview illustrates:

**Interviewer**: *So you mentioned this historic approach. [...] What do the others think about it?*
**Student 1**: *I think it was good.*
**Student 2**: *But that it became clear why it is like that. And it is good that they explained how they ended up there, why it is actually like that.*

**Student 3**: *Yes, it gets maybe easier to make connections. Or the theories up to now, if you get, somehow, you know a bit about the history and what happened before. And you get something you think - because you think the history is fascinating - you make better connections this way than just getting the theory or the formula straight into your face without any evidence for understanding why it came out like that.*

**Student 4**: *And if you also learn about how they reached this, then it can be easier to understand why it actually works.*

The findings from the focus group interviews informed revisions of the learning environment that took place over the course of two consecutive years in 2016 and 2017. These revisions led to a series of activities that draw on historic events and philosophic questions (Figures 1-3). The learning environment *General Relativity* is accessible at the Norwegian open-access learning platform Viten: www.viten.no/relativity


Fig. 1. The learning environment contrasts the first experimental confirmation of GR to the recent breakthrough in gravitational wave astronomy. Students get to explore the gravitational bending of light observerd during the solar eclipse in 1919 and they can move between newspaper headlines from 1919 and 2016.

## 5. Discussion and Conclusion

Science educators have to move beyond traditional content-focused instruction to teach concepts of Einsteinian physics. This work explored how perspectives of history and philosophy of science can inform teaching and learning of GR at the upper secondary school level. Specifically, this work presented a design-based research project that invited students to explore the historic development and philosophic aspects of GR within a digital learning environment[2]. Previous research suggests

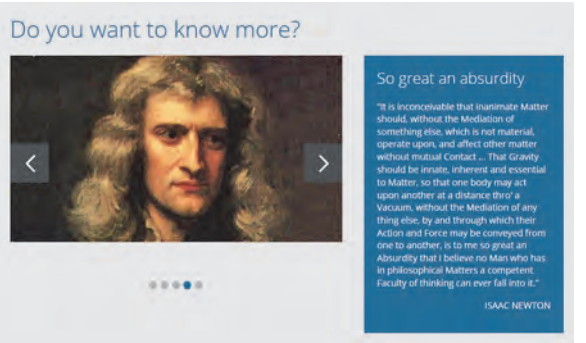



Fig. 2. Newton felt uncomfortable with the mysterious force of gravity that exhibits the principle of action at a distance. In a letter to a colleague, Newton called the gravitational force a 'great absurdity'.

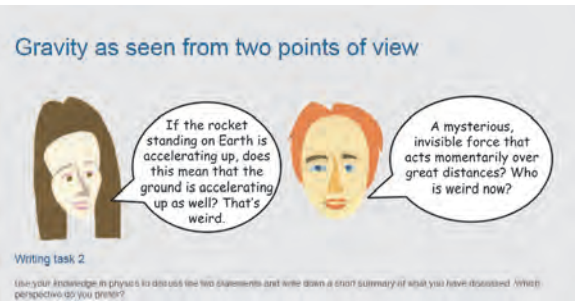



Fig. 3. In a discussion task, students are asked to discuss the epistemological implications of Einstein's and Newton's interpretation of gravity.

that HPS approaches can be an effective way to help students revise their understanding of space and time in line with relativistic ideas[4,9]. Findings of this study shift the focus away from novel conceptualisations of space and time towards more personal and societal aspects of GR: Guided by the question of how an HPS approach to teaching and learning of GR might foster understanding of and motivation for GR, the findings show that students were motivated by the presentation of Einstein as a real and struggling person instead of an all-knowing genius. Moreover, students appreciated perspectives that emphasised the cultural and social relevance of physics. Being able to gain insights into modern researchers' perspectives and seeing that there are still many unresolved questions in physics was another motivating factor. HPS perspectives did not only foster motivation among students; the historic contextualisation of GR helped students to link relativistic concepts to their previous knowledge of physics. The findings thus support an observation by Einstein that awareness of the history of physics can facilitate deeper understanding of physics[6]. In conclusion, the findings of this work suggests that employing history

and philosophy of science in the service of physics education can serve as a successful approach to making GR more accessible to learners at the upper secondary school level.

## References


1. Kaur, T., Blair, D., Stannard, W., Treagust, D., Venville, G., Zadnik, M., ... Perks, D. (2018). Determining the Intelligibility of Einsteinian Concepts with Middle School Students. *Research in Science Education*.
2. Kersting, M., Henriksen, E. K., Boe, M. V., Angell, C. (2018). General relativity in upper secondary school: design and evaluation of an online learning environment using the model of educational reconstruction. *Physical Review Physics Education Research*, 14(1), 010130-1-010130-18.
3. Hottecke, D., Henke, A., Riess, F. (2012). Implementing History and Philosophy in Science Teaching: Strategies, Methods, Results and Experiences from the European HIPST Project. *Science and Education*, 21(9), 1233–1261.
4. Levrini, O. (2014). The role of history and philosophy in research on teaching and learning of relativity. In M. R. Matthews (Ed.), International Handbook of Research in History, Philosophy and Science Teaching (pp. 157–181). Dordrecht: Springer.
5. Kim, H., Lee, G. (2018). Reflection and Outlook on Special Relativity Education from the Perspective of Einstein: Focusing on Research Papers Published in Korea. *Journal of the Korean Physical Society*, 73(4), 422–435.
6. Einstein, A., Infeld, L. (1938). The Evolution of Physics: From early concepts to relativity and quanta. Simon and Schuster.
7. Anderson, T., Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research?*Educational Researcher*, 41(1), 16–25.
8. Braun, V., Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3, 77–101.
9. Kersting, M., Steier, R. (2018). Understanding curved spacetime - the role of the rubber sheet analogy in learning general relativity. *Science and Education*, 27(7), 593–623.

# Teaching Einsteinian concepts in the first years of Italian secondary school: some preliminary results

Matteo Luca Ruggiero*

*IIS Russell-Moro-Guarini, Corso Molise 58, Torino Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy and INFN, Sezione di Pisa, Italy*
**E-mail: matteo.ruggiero@polito.it*

As a matter of fact, modern physics is not taught in the majority of secondary schools; however, recent researches suggest that there are several advantages in introducing modern physics early in school. This is, for instance, the approach of the Einstein-First project. In the context of this approach, we report here some preliminary results of an intervention in the first year of Italian secondary school.

*Keywords*: Einsteinian physics; relativity; teaching modern physics.

## 1. Introduction

Einstein's theory of General Relativity (GR) has had a renewed popularity due to recent discoveries, such as the observations of gravitational waves signals by interferometers[1], which also proved the existence of black holes, and the reconstruction of a black hole shadow[2]. These discoveries have had a widespread impact on the non-specialist public: indeed, they have been presented as "events" by the scientific community, and the results have been shared with great success in the media. These scientific events testify a renewed interest in science which could perhaps stimulate new enrollments in Science, Technology, Engineering and Mathematics (STEM) subjects: as a matter of fact (see e.g. Refs. 3, 4, 5) recent analyses show a decline of interest among students in these disciplines. The detection of gravitational waves required an enormous scientific endeavour in order to build and operate the interferometers such LIGO and VIRGO; the detection process exploits the quantum behavior of light and, consequently, we may say that these discoveries are a beautiful example where the two great theories of XX century are used together to discover the ultimate secrets of Nature. Notwithstanding the great interest in the new discoveries also among teenagers, who have a privileged and frequent access to the internet and social media, we notice that modern physics, i.e. Quantum Mechanics (QM) and GR are usually out of school curricula, except for some outline in the last year of secondary schools. The Italian education system makes no exception: in fact, in the syllabi for secondary school[a] the word "Einstein" appears just one time, in the syllabus of schools leaving qualification in scientific studies, while the word "quantum mechanics" is totally absent: the syllabus ends with the "wave nature of light, as postulated by De Broglie" and the "uncertainty principle". As a consequence, the great majority of students are completely unaware of the physical principles on which are based not only the recent discoveries but also their

[a]Indicazioni Nazionali, MIUR (2010): https : //tinyurl.com/indicazioni − nazionali

mobile phones and tablets. Teaching modern physics in view of its application in technology could help to motivate students: on one hand, this is important to face the decline in STEM enrollment and, on the other hand, this can help to shorten the distance between citizens and science, in order to have more awareness in political choices connected with scientific issues. Recent research has suggested that it is possible and also effective to introduce modern physics early in schools: the Einsten-First project[6–8] aims at introducing the basic ideas of modern physics from an early age, i.e. starting from primary school. The introduction of the language and the concepts of modern understanding of space, time, matter and radiation at early age could prevent conceptual conflicts between Newtonian physics and Euclidean geometry, that are commonly taught at school, and the completely different world view of modern physics. According to the latter, Euclidean geometry is just a good approximation in regions where the space-time curvature is small; similarly, space and time are quite different from the absolute entities of Newtonian physics, but they are woven into the fabric of space-time, which is a "living" thing according to the famous J.A. Wheeler quotation: *spacetime tells matter how to move, matter tells spacetime how to curve*. So, if the concepts of modern physics can be learnt from an early age, there will be no conflict associated to the transition from the classical to the modern paradigm of physics. Moreover, an early introduction of the fundamental ideas of modern physics can help also those students that will not study modern physics at secondary or tertiary level to have an understanding of the physics principles which current technology is based on. Of course, this new approach does not suggest to completely abandon the paradigm of classical physics, because it is important to understand many phenomena; it is rather suggested that classical physics can be introduced *after* learning the modern physics paradigm or, at least, its language and basic concepts. This paper reports the results of a short intervention, according to the approach of the Einstein-first project, in the first year of Italian secondary school, aimed at introducing some basic ideas of GR.

## 2. The context and the intervention

The intervention involved students in the first year of "ITES Russell-Moro" in Torino, a secondary school leaving qualification in economics. Students in this school are taught science (physics, Earth science, biology and chemistry) during the first two years only. In particular, physics is taught during the first year and its curriculum includes Newtonian mechanics, thermology and electricity. Modern physics is completely out of the syllabus, so that students leave this school without any knowledge of the XX century physics and its applications. In particular, 46 students, aged 14-16, from three classes took part in the intervention. It is important to emphasize that during the year, students were not taught Newton law of gravitational interaction: they only studied weight as a force. The development of the intervention, which consisted in four 1-hour lessons during four days for each class, is described in details below.

Fig. 1. Triangles on a flat (left) and curved surface (right).

Fig. 2. The space-time simulator. Frames on the left show the path of two steel balls on the flat surface of the simulator, which corresponds to propagation in vacuum: the two balls move along straight lines. Frames on the right show the path of the two steel balls on the curved surface of the simulator, around masses: the two balls undergo a deflection with respect to the vacuum case. This is a simple model of gravitational deflection.

**DAY 1.** Students were asked to answer a pre intervention questionnaire, which is an Italian translation of the one already used by the Einstein-First collaboration in the interventions described in[6,8]; we just removed questions about the quantum nature of matter and light, since these subjects were not discussed during the intervention. The key ideas we focused on are the non Euclidean features of geometry, that are important to describe the current model of space-time, and the way distant bodies interact. The relevant questions are:

Q1: Can parallel lines ever meet? Circle Yes or No. Please give reasons for your answer.

Q2: Can the sum of the angles in a triangle be different from 180 degrees? Circle Yes or No. Please give reasons for your answer.

After compiling the questionnaire, students were asked to measure the perimeter and the sum of the internal angles of triangles on a sheet of paper (see Figure 1, left).

**DAY 2.** Students were asked to draw triangles on the surface of balloons (using sticky tape to draw "straight" lines) and, then, to measure their perimeter and the sum of internal angles (see Figure 1, right). Moreover, they had to draw a plot of the sum of internal angles as a function of the perimeter.

**DAY 3.** Students were involved in activities at the "space-time simulator". Actually, models and analogies prove useful to introduce the concepts of modern physics and they are developed and successfully tested by the Einstein-First project. In particular, the space-time simulator is a table with a stretch lycra sheet, which

creates a visual representation of the curvature of space-time in presence of masses (see Figure 2). Using steel or billiard balls, the simulator allows to visualize the connection between mass and curvature. A ball moves along a straight line if no other masses are present on the sheet, while its path is curved in presence of other masses: in analogy, a mass move along a geodesic in space-time, that is a straight line if space-time is flat; in presence of other masses the ball moves along a curved geodesic. Moreover, it is manifest that the lycra sheet deformation is proportional to the masses on it. The simulator allows to visualize the action at distance between masses, which is explained in terms of curvature (see Figure 3). In summary, the space-time simulator enables to perform numerous experiments that simulate what is observed in the Universe, such as, for instance, planetary motion, tidal effects, gravitational deflection of light. Indeed, what is really visualised is the interplay between masses and the lycra sheet deformation, which can be described in terms of spatial curvature. Accordingly, the space-time simulator might be misleading, since in the Universe curvature is refereed both to the space and *time* dimensions. For instance, free fall in curved space-time can be explained as a consequence of the warping of time (see Ref. 9 and references therein). That being said, in our intervention the space-time simulator was used as an analogy, and it proved useful to emphasize the non Euclidean features of surfaces.

Fig. 3. The space-time simulator and action ad distance: the simulator can be used to simulate action at distance, since moving one ball over the lycra sheet provokes a deformation and, hence, the motion of the other ball.

**DAY 4.** After a brief summary of the activities performed, with emphasis on the common idea linking them (curvature of surfaces) and a short description of the analogy with what really happens in the Universe (planetary motion, tidal effects, gravitational deflection of light), students were asked to answer the post intervention questionnaire. In the latter we added a question, motivated by the experiments performed at the space-time simulator (see Figure 3):
QX: The gravitational attraction force is effective only when the two bodies are close to each other. Circle Yes or No. Please give reasons for your answer.

## 3. Analysis of the results

In Figure 4 we report the analysis of the answers to the questions Q1 and Q2. As for Q1 “Can parallel lines ever meet?”, before the intervention *all students* answered

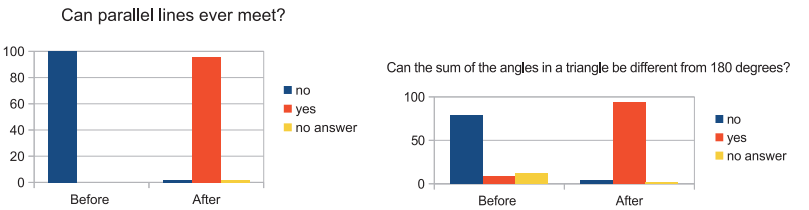



Fig. 4. Distributions of the answers Q1 (left) and Q2 (right) before and after the intervention

NO, which is an expected results: Euclidean geometry is based on the parallel postulate and students are taught that two lines are parallel if they *never* meet. However, during the intervention students experimented that things are quite different if the surface is curved, such as on a balloon; furthermore, the analogy with what happens on the surface of the Earth was discussed during the intervention. After the intervention, the number of students who answered YES is about 95%; it is relevant that in the reasons they mentioned the role of the surface: "on curved surfaces, such as the Earth, they can meet". Things are quite similar for Q2 "Can the sum of the angles in a triangle be different from 180 degrees?" Before the intervention about the 79% of students answered that the sum cannot be different from 180 degrees; 9% said that the sum can be different and the rest (12 %) did not answer. In this case, the percentage of students who said that the sum needs to be 180 is lower than 100% reported in Q1: we have to consider that perhaps some of them did not remember this rule of Euclidean geometry. After the intervention, those who say that the sum can be different from 180 degrees are about 94%, which is similar to the result of Q1. In the motivations they say "it depends on the surface", "it depends on the triangle", "it depends on the perimeter"; the latter motivations rely on to fact that they analyzed the mathematical relationship between the perimeter and the sum of internal angles on the balloons. As for the extra question QX "The gravitational attraction force is effective only when the two bodies are close to each other", 54% of the students correctly answered that bodies interact at distance; in the motivations they correctly point out that "masses influence each other even though they are not close", thus revealing that they correctly understood the mutual character of interaction, which has been emphasized during the experiment with the space-time simulator. The other students (46%) did not believe that interaction between distant bodies is possible, notwithstanding the experiments at the space-time simulator. Actually, the ideas of secondary schools students about action at distance were investigated in the past (see e.g. Ref. 10), and evidence was found that in order to realize the abstract notion of action at distance students need a connection: for instance, between a falling object and the Earth, they consider air as a medium that creates a connection. Also, their ideas about forces acting at distance are guided by the notion of a support: for instance,

objects at rest or moving horizontally are supported by other objects, and these supports prevent them by falling. Moreover, students are not concerned with mutual interactions, rather they think of the Earth that attracts the apple, or the Sun that attracts the Earth, thus emphasizing a dominant role in the interaction: a much higher level of abstraction is needed to understand gravitation as an interaction between any two bodies at any distance in space. Indeed, there is the possibility that even though students correctly understood the experiments with two masses on the space-time simulator, they thought that, in any case, masses are not "far away" on the simulator and they were not able to generalize their observations to a more abstract situation of truly distant objects. In conclusion, in this intervention we focused on the non Euclidean features of curved surfaces and, using the space time simulator, we described by analogy what happens in curved space-time. As a result of the intervention, the great majority of students was able to understand that the properties of Euclidean geometry are not true on curved surface. The analogue model at the space-time simulator enabled us also to introduce the concept of action at distance, which was correctly understood by more than one half of the sample. The results of the intervention are positive, if we consider the context and the students' attitude: of course further interventions on different contexts, with a larger number of students and with more educational content are needed to evaluate the effectiveness of this approach.

## References


1. B. P. Abbott *et al.*, Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, p. 061102 (Feb 2016).
2. K. Akiyama *et al.*, First m87 event horizon telescope results. i. the shadow of the supermassive black hole, *Astrophysical Journal Letters* **875**, p. L1 (2019).
3. C. Tanenbaum, Stem 2026: A vision for innovation in stem education, *US Department of Education, Washington, DC* (2016).
4. Australian Governement, Australia's national science statement (2017).
5. H. Shapiro, S. Østergård and K. Hougard, Does the eu need more stem graduates, *Publications Office of the European Union: Luxembourg* (2015).
6. T. Kaur *et al.*, Teaching einsteinian physics at schools: part 1, models and analogies for relativity, *Physics Education* **52**, p. 065012 (2017).
7. A. Foppoli *et al.*, Public and teacher response to einsteinian physics in schools, *Physics Education* **54**, p. 015001 (2018).
8. R. K. Choudhary *et al.*, Can a short intervention focused on gravitational waves and quantum physics improve students' understanding and attitude? *Physics Education* **53**, p. 065020 (2018).
9. M. Kersting, Free fall in curved spacetime—how to visualise gravity in general relativity, *Physics Education* **54**, p. 035008 (2019).
10. V. Bar, B. Zinn and E. Rubin, Children's ideas about action at a distance, *International Journal of Science Education* **19**, 1137 (1997).

# Historical, philosophical and theological framing of General Relativity for high school students

C. Sigismondi (1)* T. Pompa (2) and D. Impellizzeri (3)
*ICRA/Sapienza and ITIS G. Ferraris, Roma, Italy
(2) Liceo G. Galilei, Pescara
(3) ITAS G. Garibaldi, Roma
*E-mail: sigismondi@icra.it www.icra.it/solar

Motivating the students to study General Relativity from many points of view is a key for success in teaching GR. The risk of flattening GR to its mathematical contents is continuously present in the secondary school as well as in the undergraduate studies of Physics.

*Keywords*: History of Science, General Relativity, Philosophy of Science.

## 1. Motivations: from wave optics to quantum behavior

Most of Italian students do not choose STEM curricula after school. Among the reasons there is a lack of experimental activities especially for the last year of physics in the High Schools. An approach to Quantum Physics is recommended through spectroscopy, thanks to the wave-particle behavior of the photon, the availability of a LASER and the simplicity of many diffraction experiments. This not help to understand how interference and diffraction phenomena have been studied before the advent of LASER. The Michelson Interferometer was used to study the vectorial composition of the light with the orbital Earth velocity, or the cosmic ether existence. The result was that the modulus of velocity c was conserved, while the direction followed the vectorial composition traditional rule (Andrei, et al. 2014; Sigismondi, 2008).

Nevertheless the historical process through selection rules and suspected extraterrestrial elements like Helium and Coronium is also important. Quantum Mechanics verifies the Heisemberg principle for the matter, the diffraction of light is already the Heisemberg principle for photons (Vavilov, 1959).

At the Italian Ministry for Education a course in Spectroscopy is already organized for the high school teachers http://ls-osa.uniroma3.it/documents/98 (2018) but the general vision given by the historical perspective is still lacking. The selection rules found in spectroscopy $\Delta L, \Delta m = 0, \pm 1, \pm 2$...were interpreted later as permitted quantum gaps, and their violations as 'prohibited rules' upon which many fundamental discoveries have been done (e.g. 21 cm line of neutral Hydrogen and all galactic structure). The 'rarity' of some transition were obtained directly from quantum mechanics density function, as tunnel effect.

## 2. Gravitational waves and interferometry

In General Relativity the interferometers of Michelson e Morley (1881-1887) have been adapted to gravitational waves in LIGO/VIRGO experiments using

the same principles and the same geometry. The interference of Michelson experiment http://adsabs.harvard.edu/abs/1882USNAO...1..109M (1882) and http://adsabs.harvard.edu/abs/1887SidM....6..306M (1887) was realized with white light, a situation which is different from monochromatic LASER light, and much closer with the first experiences on diffraction made by the Jesuit father Francesco Maria Grimaldi in 1648, with very tiny Fresnel shadows occurring for the interplay between Planck's black body photons distribution and Fresnel integrals for each wavelength. Similarly at Mount Wilson Observatory the interference between the stellar light received at opposite parts of the diameter of the 100 inches reflecting Hooker telescope (the technological marvel of its time) permitted the evaluation of the stellar diameter of Betelgeuse and the first stellar surface images. (S. Metchev (2013) in http://www.astro.sunysb.edu/metchev/PHY517_AST443/solar_diameter.pdf and M. Richmond http://spiff.rit.edu/richmond/occult/bessel/bessel.html (2018))

Also this interferometer can be realized and studied at school, as in 2018 at Pescara Galilei Lyceum. The utilization of Magneto Optical Filter in Gravitational Waves detection was proposed by the late Alessandro Cacciani (1937-2007) (P. Rapex, 2008).

## 3. Relativistic light bending

Other effects like light gravitational bending are very difficult to be observed, even if the GAIA astrometry satellite is now able to measure it at 180 degrees from the Sun and the major planets (Crosta and Mignard, 2005). The gravitational light bending can be simulated only by the analogy photon-glass ball on the elastic carpet made by lycra, with a big mass at the center. This similarity (used e.g. at Rome 3 University, Physics dept.) is not completely satisfying because of the photon is massless. The apparatus is good for showing N-body simulations with friction, to show e.g. the onset of King profiles in galactic clusters... in a few seconds, instead of billions of years.

This effect shows clearly the difference from Newtonian and Einstein gravity upon light (sir James Jeans, The Growth of Physical Sciences, 1948). Unless it is not evident form Mitchell 1780 approach of escape velocity, which gives correctly the value of Schwarszschild Radius, but considers light gravitating as matter, even if we know the photon is massless.

The great project Gravity Probe B on the Lense-Thirring effect of Earth's frame dragging is also far from experimental level without using analogies (Ruffini and Sigismondi, 2002). The Solar figure, spherical within few parts in a million, was measured by Dicke (Damiani et al., 2007) to assess the possibility of a scalar-tensorial theory beyond Einstein and now is measured to understand the physics of solar secular variability, so important for our climate (Sigismondi, 2011 and 2005).

## 4. General Relativity and the view of the World

General Relativity was important to the view of the World as much as Quantum Physics. Bertrand Russell (1974) was one of the most relevant divulgators of Einstein's theory, he contributed to spread the basis of Special and General Relativity, but the relevance of the new cosmological view of the World is left too often to the University specialistic studies, to avoid also to fall into 'unpleasant metaphysical questions'. While right these questions can drive more students to STEM. The continuum model of the Universe has been the same at all times, in order to agree with cosmological expansion a continuum creation is required, which is not against the physical laws. This last part, being not comprised, is not present in textbooks like M. Parotto, E. L. Palmieri, il Globo Terrestre e la sua Evoluzione, Zanichelli Bologna 2017. The 'unpleasant metaphysical questions' are indeed more stimulating than *increasing precision for saying anything about the truth* [Martin Heidegger]. Russell in this respect says that 'the consequences of Relativity are neither grandiose nor surprising as believed' (p. 218-219). The differences between classical and relativistic physics are usually very small, and the mass of 4 weights of 1 Kg is smaller than 4 Kg, but the evidence of this effect of the equivalence of mass - binding energy in percentage four nuclei of H are lighter (0.7% less) than one nucleus of He. Russell explains that events instead of coordinates are the essence of Relativity, and the action [Energy*time] is its natural unit of measure (p. 156-7). The action which fulfilled some aesthetic, almost theological principle of symmetry with Maupertuis, Lagrange and Laplace, and become crucial in Quantum Mechanics with Planck constant in 1900 (J. Jeans, Il Cammino della Scienza, Bompiani, 1953, p. 318-328).

## 5. Cosmology, a new science with proper identity after Einstein

With Einstein the Physical Cosmology was born, because the Universe's behavior was comprised as a consequence of the matter's distribution Peebles, Principles of Physical Cosmology, 1990). The Galileian idea of same physics on the Earth and across all Universe reached in this way its maximum extent. Einstein by introducing the Cosmological Constant and later Hoyle, Bondi and Gold (Dubois and Fuezfa, 2019), and finally Hawking and Hartle tried to obtain a static or stationary Universe, conceptually not different from Aristotle's one, while Lemaitre proposed an evolutionary Universe, finding the observational proofs. The hot big bang theory was confirmed by the Cosmic Background Radiation discovery (1965) even if Zeldovich (1965) theorized a cold one, based upon Fermi energy (Aguirre, 1999).

## 6. Ether and eternity of the Universe from Aristotle to Hawking

The role of Cosmic Ether, object of the studies of Michelson and Morley, is now played by the Dark Energy, which behaves exactly as the Cosmological Constant, and produces an accelerating Universe. Framing philosophically the Physical

Fig. 1. Y. B. Zeldovich and Pope John Paul II.

Cosmology helps the students, also the ones far from STEM, to understand the present debate on the view of our World. Comparing the theories of expanding Universe, eternal Universe, oscillating Universe with the ancient ones, made by Aristotle and object of a condemnation by Etienne Tempier bishop of Paris in 1277, is also important. Instead of blocking science this act stimulated it, because always Christianity has considered God as Creator. He is not a lier and the man can know Him from Revelation, but also from the Creation itself (as S. Thomas of Aquinus few years before Etienne Tempier taught in Paris University). 'God does not play with dices' used to say Einstein and this debate is nowadays extremely far from the current main stream. [Alessandro Giostra http://www.upra.org/evento/pierre-duhem-la-nascita-della-scienza-nel-mondo-occidentale-cristiano/]

The beginning of science and the possibility of confutation for a theory According to Pierre Duhem (Système du Monde, 1913-58), the act of Etienne Tempier was the starting point of the modern science. The construction of the theory, moreover, is based upon some principles, and the speed of light in vacuum-the veritable "celeritas"- cannot be surpassed. The recent (2011) debate of OPERA results in Gran Sasso National Laboratories and CERN of Geneva can be studied in this optics. Provided this frame, there are plenty of motivations to know the essence of General Relativity and its contribution to the view of our World.

Einstein did know well that only one contradictory fact on the speed of light in vacuum would destroy all General Relativity theory. This characteristic of a scientific theory is the Fälschungsmöglichkeit, possibility of confutation, theorized by Karl Popper.

Nowadays we are facing more and more with dogmatic theories like the anthropogenic global warming or the OGM vs organic food with respect to human health or the gender theory on the sexual differentiation induced by the environment and the cultural background... whatever fact has an explanation pro that theory, and no disproofs are possible, and we are in a scientific, post-galileian era. This note may appear unappropriated either in a scientific congress either in the school, but the paralysis of autonomous thinking starts from the school and spread to all parts of the society. If the most influential theory of Physics is subjected naturally to the possibility of confutation, why there should be theories that cannot be falsified? A theory since Ptolemy and before is a representation of the reality, which can be changed with a better one, when it is available.

Fig. 2. The cover of "Sozein ta Phainomena" of Pierre Duhem.

## References


1. A. H. Andrei, C. Sigismondi, V. Regoli, eprint arXiv:1412.6096 (2014).
2. Sigismondi, C., Proc. Marcel Grossmann Meeting XI 2470 (2008).
3. Vavilov, S. I., Vavilov, "L'Occhio e il Sole", Feltrinelli, Milano (1959).
4. Rapex, P., http://www.icra.it/solar/alessandro_cacciani_memorial.htm (2011).
5. M. T. Crosta and F. Mignard, https://arxiv.org/abs/astro-ph/0512359 (2005).
6. R. Ruffini and C. Sigismondi, "The Lense-Thirring effect", World Scientific (2002).
7. C. Damiani et al. https://hal.archives-ouvertes.fr/hal-00519433/document A Brief History of Solar Oblateness. A review (2010).
8. C. Sigismondi, "High precision ground-based measurements of solar diameter in support of PICARD mission", PhD Thesis University of Nice-Sofia Antipolis http://adsabs.harvard.edu/abs/2011PhDT.....372S (2011).
9. B. Russell, "L'Abc della Relatività", Longanesi, Milano 1974, p. 173.
10. C. Sigismondi, Nuovo Cimento B 120, 1169 (2005).
11. Dubois, E. A. and A. Fuezfa, A. Gen. Relativ. Gravit. 51, 11 (2019).
12. Zeldovich, Ya. B. 1965, Adv. Astron. Astrophys., 3, 241.
13. A. Aguirre, 1999 http://iopscience.iop.org/article/10.1086/307534/pdf.

# Lie point symmetries of the geodesic equations of the Gödel's metric

Fatma Al-Kindi and Muhammad Ziad*

*Department of Mathematics, College of Science, Sultan Qaboos University, AlKhoudh, Muscat, Sultanate of Oman*

**E-mail: mziad@squ.edu.om*

Lie point symmetries of the geodesic equations of the Gödel's metric are found. These form a ten dimensional Lie algebra. The Lie algebra contains a maximal seven-dimensional solvable sub-algebra. It also contains five dimensional subalgebra of isometries of the metric. The isometries are used to reduce the order of the geodesic system by one. The time-like trajectories of the Gödel's metric are then derived and their graphs in the $(r, \phi)$ plane are displayed showing some interesting features of the dynamics in this universe.

*Keywords*: Lie Symmetries; Isometries; Geodesic Equations.

## 1. Introduction

Systems of ordinary differential equations frequently appear in applied sciences and engineering, and most of the time finding their solution appears a big challenge. Theory of continuous groups via Lie point symmetries provides an alternate approach for finding solutions of such systems.[1–3] Finding Lie point symmetries (using Einstein summation convention)

$$X = \xi(x, y^i)\frac{\partial}{\partial x} + \eta^i(x, y^i)\frac{\partial}{\partial y^i} \tag{1}$$

of a system of $k$ second order ordinary differential equations

$$H^a(x, y^i, \dot{y}^i, \ddot{y}^i) = 0, \quad a, i = 1, .., k, \tag{2}$$

means finding the general solution $\xi(x, y^i)$ and $\eta^i(x, y^i)$ of the determining equations obtained from the symmetry condition[2]

$$\hat{X}(H) = 0, \tag{3}$$

where $\hat{X}$ is the extension of the symmetry operator $X$ written as

$$\hat{X} = \xi(x, y^i)\frac{\partial}{\partial x} + \eta^i(x, y^i)\frac{\partial}{\partial y^i} + (\eta^i)'(x, y^i)\frac{\partial}{\partial \dot{y}^i} + (\eta^i)''(x, y^i)\frac{\partial}{\partial \ddot{y}^i}, \tag{4}$$

and $(\eta^i)'$ and $(\eta^i)''$ are obtained from the extension formula given by

$$(\eta^i)^{(n)} = \frac{d(\eta^i)^{(n-1)}}{dx} - (y^i)^{(n-1)}\frac{d\xi}{dx}. \tag{5}$$

Determining equations obtained from eq. (3) appear as identities in the derivatives of dependent variables, thus they split into $k$ systems of linear partial differential equations in the functions $\xi$ and $\eta^i$ which are to be solved simultaneously.
Gödel's metric in a cylindrical coordinate system $(t, r, \phi, z)$, where $t < \infty$, $0 \leq r \leq \infty$, $0 \leq \phi \leq 2\pi$, $-\infty < z < \infty$, given by

$$ds^2 = a^2 \left( [dt + \sqrt{2} \sinh^2 r d\phi]^2 - dr^2 - dz^2 - \sinh^2 r \cosh^2 r d\phi^2 \right), \tag{6}$$

ascertain the geodesic equations

$$\ddot{t} + \frac{4 \sinh r}{\cosh r} \dot{t}\dot{r} + \frac{2\sqrt{2} \sinh^3 r}{\cosh r} \dot{r}\dot{\phi} = 0, \tag{7a}$$

$$\ddot{r} + 2\sqrt{2} \sinh r \cosh r \dot{t}\dot{\phi} - \sinh r \cosh r (1 - 2 \sinh^2 r) \dot{\phi}^2 = 0, \tag{7b}$$

$$\ddot{\phi} - \frac{2\sqrt{2}}{\sinh r \cosh r} \dot{t}\dot{r} + \frac{2}{\sinh r \cosh r} \dot{r}\dot{\phi} = 0, \tag{7c}$$

$$\ddot{z} = 0, \tag{7d}$$

where dot over head the variables $t$, $r$, $\phi$ and $z$ denote the derivatives with respect to the arc length parameter $s$. The solution of these equations have been a topic of interest to many researchers following different approaches.[4–6] The Killing vectors from these symmetries are singled out to reduce the order of the system. In the following section, the Lie point symmetries of the system are found along with the analysis of the resulting Lie algebra. In section 3 we use the Killing vectors to reduce the order of the system. The solution of the system is then achieved and a graphic representation of the geodesics is given.

## 2. Lie point symmetries of the geodesic system

Applying the second prolongation of the symmetry generator of the system (7) produces four determining equations, which give a system of seventy PDEs. The solution of this system then gives the components of the infinitesimal generator $X$ which appears to be

$$\xi = c_1 s + c_2 z + c_3, \tag{8}$$

$$\eta^1 = -\sqrt{2} \tanh r \, (c_4 \sin \phi - c_5 \cos \phi) + c_7, \tag{9}$$

$$\eta^2 = c_4 \cos \phi + c_5 \sin \phi, \tag{10}$$

$$\eta^3 = -\frac{2 \cosh^2 r - 1}{\sinh r \cosh r} \, (c_4 \sin \phi - c_5 \cos \phi) + c_6, \tag{11}$$

$$\eta^4 = c_8 s + c_9 z + c_{10}. \tag{12}$$

This provides a basis of the Lie algebra given by

$$X_1 = s\frac{\partial}{\partial s},\ X_2 = z\frac{\partial}{\partial s},\ X_3 = \frac{\partial}{\partial s},$$
$$X_4 = -\sqrt{2}\tanh r\sin\phi\frac{\partial}{\partial t} + \cos\phi\frac{\partial}{\partial r} - \frac{2\cosh^2 r - 1}{\sinh r\cosh r}\sin\phi\frac{\partial}{\partial\phi},$$
$$X_5 = \sqrt{2}\tanh r\cos\phi\frac{\partial}{\partial t} + \sin\phi\frac{\partial}{\partial r} + \frac{2\cosh^2 r - 1}{\sinh r\cosh r}\cos\phi\frac{\partial}{\partial\phi},$$
$$X_6 = \frac{\partial}{\partial\phi},\ X_7 = \frac{\partial}{\partial t},\ X_8 = s\frac{\partial}{\partial z},\ X_9 = z\frac{\partial}{\partial z},\ X_{10} = \frac{\partial}{\partial z}, \tag{13}$$

forming a ten-dimensional Lie algebra $L_{10}$ which is neither solvable nor nilpotent. It is trivial to identify the Nöether gauge symmetries which appear to be $X_3,\ X_4,\ X_5,\ X_6 + \frac{\sqrt{2}}{2}X_7,\ X_7,\ X_8,\ X_{10}$ in eqs. (13). The maximal solvable subalgebra is given by

$$L_7 = \left\langle X_1 + X_9,\ X_1 - X_9,\ X_2,\ X_3, X_6 + \tfrac{\sqrt{2}}{2}X_7,\ X_7,\ X_{10}\right\rangle \subset L_{10}. \tag{14}$$

## 3. Reduction of the order of the system

One can use Cartan's theory,[2] according to which there exists a first integral, $X_a\dot{x}^a$ for each symmetry generator $X = \xi_a\partial_a$ obtained in eqs. (13) which satisfies the equations of Killing[2]

$$X_{a;b} + X_{b;a} = 0. \tag{15}$$

It is straight forward to check that the symmetry generators $X_i$ where $i = 4, ..., 7, 10$ satisfy eqs. (15). The corresponding first integrals are given by

$$a = -\sinh r\cosh r\sin\phi\left[2\sqrt{2}\dot{t} + \dot{\phi}(2\sinh^2 r - 1)\right] - \dot{r}\cos\phi, \tag{16}$$
$$b = \sinh r\cosh r\cos\phi\left[2\sqrt{2}\dot{t} + \dot{\phi}(2\sinh^2 r - 1)\right] - \dot{r}\sin\phi, \tag{17}$$
$$c_1 = \sqrt{2}\dot{t}\sinh^2 r + \dot{\phi}\sinh^2 r(\sinh^2 r - 1), \tag{18}$$
$$c_2 = \dot{t} + \sqrt{2}\sinh^2 r\dot{\phi}, \tag{19}$$
$$c_\circ = -\dot{z}, \tag{20}$$

where $a, b, c_\circ, c_1$ and $c_2$ are arbitrary constants. The tangent vector $\dot{x}^a$ and the associated underlying curve $x^a(s)$ are said to be timelike if and only if $g_{ab}\dot{x}^a\dot{x}^b > 0$. Fixing $g_{ab}\dot{x}^a\dot{x}^b = 1$ yields

$$\left(\dot{t} + \sqrt{2}\sinh^2 r\dot{\phi}\right)^2 - \sinh^2 r\cosh^2 r\dot{\phi}^2 - \dot{r}^2 = 1 + c_\circ^2 \equiv c_3 \tag{21}$$

or

$$c_2\dot{t} + c_1\dot{\phi} - \dot{r}^2 = c_3. \tag{22}$$

The sign of the term $\dot{t} + \sqrt{2}\sinh^2 r\dot{\phi}$ and hence of $c_2$ may be defined to identify the future or past directions of the tangent vector $\dot{x}^a$. One may be interested

in exploring $x^a(s)$ which are timelike and whose tangent vector fields are future directed. Thus necessarily we require $c_2 > 0$, whereas $c_3 > 0$ from eq. (21).

Eqs. (16)-(22) give explicit expression of $\dot{x}^a$ reducing the system as

$$\dot{t} = c_2 \left[1 - \frac{2\sinh^2 r}{\cosh^2 r}\right] + \frac{\sqrt{2}c_1}{\cosh^2 r}, \tag{23a}$$

$$\dot{\phi} = \frac{\sqrt{2}c_2}{\cosh^2 r} - \frac{c_1}{\sinh^2 r \cosh^2 r}, \tag{23b}$$

$$\dot{r}^2 = c_2^2 - c_3 - \left(\frac{\sqrt{2}c_2 \sinh r}{\cosh r} - \frac{c_1}{\sinh r \cosh r}\right)^2, \tag{23c}$$

$$\dot{r} = -\left(a\cos\phi + b\sin\phi\right). \tag{23d}$$

Using the transformation $u = \sinh^2 r$, and letting

$$4(c_2^2 + c_3) = \varepsilon^2, \quad \frac{(c_2^2 + 2\sqrt{2}c_1c_2 - c_3)}{2(c_2^2 + c_3)} = \alpha, \quad \frac{c_1^2}{(c_2^2 + c_3)} = \beta^2, \tag{24}$$

eq. (23c) can be written as

$$\dot{u}^2 = -\varepsilon^2(u^2 - \alpha u + \beta^2), \tag{25}$$

where $c_2^2 + c_3 > 0$ as $c_3 > 0$. The latter equation gives

$$u(s) = \alpha + \sqrt{\alpha^2 - \beta^2}\cos(\varepsilon s + s_\circ) = \sinh^2 r. \tag{26}$$

Now using eq. (26) in eqs. (23a)-(23b) after simplification one gets the time-like trajectories in the Gödel universe as

$$\begin{aligned}(t, r, \phi, z) = \Bigg(&\sqrt{2}\tan^{-1}\left(\sqrt{\frac{\alpha+1-\sqrt{\alpha^2-\beta^2}}{\alpha+1+\sqrt{\alpha^2-\beta^2}}}\tan\left(\sqrt{c_2^2+c_3}s + \frac{s_\circ}{2}\right)\right) - c_2 s + t_\circ,\\ &\sinh^{-1}\sqrt{\alpha + \sqrt{\alpha^2-\beta^2}\cos(\varepsilon s + s_\circ)},\\ &\tan^{-1}\left(\frac{\left(\sqrt{\frac{\alpha+1-\sqrt{\alpha^2-\beta^2}}{\alpha+1+\sqrt{\alpha^2-\beta^2}}} - \sqrt{\frac{\alpha-\sqrt{\alpha^2-\beta^2}}{\alpha+\sqrt{\alpha^2-\beta^2}}}\right)\tan\left(\sqrt{c_2^2+c_3}s + \frac{s_\circ}{2}\right)}{1 + \sqrt{\frac{\alpha+1-\sqrt{\alpha^2-\beta^2}}{\alpha+1+\sqrt{\alpha^2-\beta^2}}}\sqrt{\frac{\alpha-\sqrt{\alpha^2-\beta^2}}{\alpha+\sqrt{\alpha^2-\beta^2}}}\tan^2\left(\sqrt{c_2^2+c_3}s + \frac{s_\circ}{2}\right)}\right) + \phi_\circ, -c_\circ z + z_\circ\Bigg).\end{aligned} \tag{27}$$

For specific values of the constants $\alpha$ and $\beta$ given in eqs. (24) one can find the extrema of $u(s)$. Putting $\dot{u}(s) = 0$ of eq. (25) gives

$$u^2 - 2\alpha u + \beta^2 = 0, \tag{28}$$

which yields

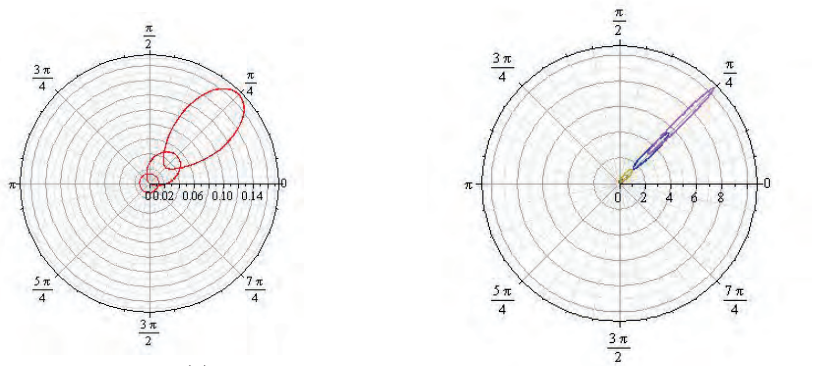
$$u(s) = \alpha \pm \sqrt{\alpha^2 - \beta^2}. \tag{29}$$

So $u(s)$ attains its maximum at $u(s) = \alpha + \sqrt{\alpha^2 - \beta^2}$ and its minimum at $u(s) = \alpha - \sqrt{\alpha^2 - \beta^2}$.

For $c_{1min} = \dfrac{-\sqrt{2}c_2 \pm \sqrt{c_2^2 + c_3}}{2}$, $\alpha^2 - \beta^2 = 0$, which means that $u(s) = \alpha$ is constant and the graph in $(u, \phi)$ plane is a circle with radius equal to $\alpha$. In this case

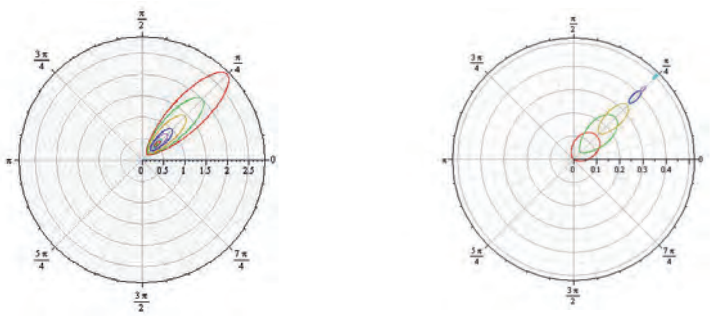
$$0 < u(s) < \frac{-1 + \sqrt{2}}{2}. \tag{30}$$

For $c_1 > c_{1min}$ the trajectories appear to be ellipses stretching and tending to be a line as $c_1 \to \infty$. If $c_3 \to c_2^2$ the ellipses shrinks and tends to be a point. Figures below illustrates the above analysis.




(a) (b)

Fig. 1. Trajectories in $(y, \phi)$ plane with increasing $c_1$, and $c_2, c_3$ are fixed.



(a) (b)

Fig. 2. (a) Trajectories with $c_1$ and $c_2$ are fixed, $0 < c_3 < c_2^2$. (b) Trajectories with $c_1$ is increasing as $c_3 \to c_2^2$.

## 4. Conclusion

It is known that a system of $n$ $kth$-order ODEs admitting a $kn$-dimensional transitive solvable Lie sub algebra, is solvable by quadratures.[2] In our case the system of four coupled ODEs admits a seven dimensional solvable sub-algebra, thus this result can not be applied for the reduction of the system. Other well known procedure is given by the Cartan theory according to which, if a system of ODEs admits a set of Killing vectors, then one obtains first integrals of the system corresponding to these Killing vectors.[2] In the present situation, the Gödel's metric serves as the Lagrangian of the system, which admits five Killing vectors and correspondingly gives five first integrals of the system given by eqs. (16)-(20). This led to a solution of the system in parametric form $x^a(s)$, $a = 1..4$ given by eq. (27). The graphs of the trajectories in $(r, \phi)$-plane for all possible values of the parameters appearing in the solution are also given at the end.

## References


1. Ibragimov, N. H., *Elementary Lie group analysis and ordinary differential equations*, John Wiley and Sons, Inc. Chichester, England, 1999.
2. Stephani, H., *Differential equations: their solution using symmetries*, Cambridge University Press, USA, 1989.
3. Hydon, P. E., *Symmetry Methods For Differential Equations*, Cambridge University Press, New York, USA, 2000.
4. Chandrasekhar, S. and Wright, J. P., *The Geodesics in Gödel Universe*, Proc. Natl. Acad. Sci. USA, **47**(1961)341.
5. Novello, M., Damiao, I. and Tiomno, J. (1983). *Geodesic motion and confinement in Gödel universe*. Phys. Rev. D, **27**(1983)779.
6. Camci, U., *Symmetries of geodesic motion in Gödel-type spacetimes*, JCAP**07**(2014)002.
7. Stephani, H., *General Relativity: An introduction to the theory of the gravitational field*, Cambridge University Press, Cambridge, New York, second edition 1990.
8. Misner C., Thorne, K. and Wheeler, J., *Gravitation*, W. H. Freeman and company, NewYork, 1973.
9. Cook, S., *Killing Spinors and Affine Symmetry Tensors in Gödel's Universe*, Oregon State University, 2010.
10. Grave, F., Müller, T., Wunner, G., Ertl, T., Buser, M. and Schleich, W., *Visualization of the Gödel Spacetime*, Germany.
11. Soh, C. Wafo and Mahomed, F. M., *Reduction of order for systems of Ordinary Differential Equations*, Journal of Nonlinear Mathematical Physics, **11**(2004)13-20.
12. Senthilvelan, M., Chandrasekar, V. K., and Mohanasubha, R. (2015). *Symmetries of nonlinear ordinary differential equations: The modified Emden equation as a case study*. Pramana, 85(5), 755-787.
13. Feroze, T., *Some aspects of symmetries of differential equations and their connection with the underlying geometry*, Ph.D thesis, Department of Mathematics, Quaid-i-Azam university, Pakistan, 2004.

# Simultaneous baldness and cosmic baldness and the Kottler spacetime

Valerio Faraoni

*Dept. of Physics & Astronomy and STAR Research Cluster, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*
*E-mail: vfaraoni@ubishops.ca*

Adriana M. Cardini

*Dept. of Physics & Astronomy, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*
*E-mail: acardini15@ubishops.ca*

Wen-Jan Chung

*Dept. of Physics & Astronomy, Bishop's University*
*Sherbrooke, Québec, Canada J1M 1Z7*
*E-mail: wchung13@ubishops.ca*

The Schwarzschild-de Sitter/Kottler geometry is the unique spherical solution of the vacuum Einstein equations with positive cosmological constant. Putative alternatives in the literature are shown to either solve different equations or to be the SdSK solution in disguise. No-hair and cosmic no-hair come together in a new simultaneous theorem for SdSK in the presence of an imperfect fluid.

*Keywords*: Schwarzschild-de Sitter, no-hair, cosmic no-hair.

## 1. Introduction

The Jebsen-Birkhoff theorem of general relativity (GR) states that the Schwarzschild geometry is the unique vacuum, spherical, asymptotically flat solution of the Einstein equations

$$G_{ab} \equiv R_{ab} - \frac{1}{2}\, g_{ab} R = 0\,. \tag{1}$$

The theorem is no longer true if matter is present, but what if a cosmological constant $\Lambda$ is included? The Einstein equations change to $G_{ab} = -\Lambda g_{ab}$. It is straightforward to extend the proof of the Jebsen-Birkhoff theorem to this case. In particular, *for $\Lambda > 0$, the Schwarzschild-de Sitter/Kottler (SdSK) geometry is the unique solution,*

$$ds^2 = -\left(1 - \frac{2m}{R} - H^2 R^2\right) dT^2 + \frac{dR^2}{1 - \frac{2m}{R} - H^2 R^2} + R^2 d\Omega^2_{(2)} \tag{2}$$

*(where $H = \sqrt{\Lambda/3}$).*

This fact is not mentioned in modern GR textbooks, even though de Sitter space is extremely important for inflationary and dark energy cosmology [1] and, for $\Lambda < 0$, anti-de Sitter space is fundamental for string theories and the AdS/CFT correspondence. A proof of the generalized Jebsen-Birkhoff theorem for $\Lambda \neq 0$ can be found in Synge's 1960 textbook.[2] A recent proof in null coordinates has

appeared in[3]. Complicated proofs appear in[4–7]. The proof in spherical coordinates is a very simple extension of the standard textbook proof of the Jebsen-Birkhoff theorem.[11] However, there is confusion in the literature: purported alternatives to SdSK have been reported, which would violate the theorem. More general solutions with FLRW "background" in the presence of matter do not seem to reduce to SdSK when the "background" becomes de Sitter. Enter modified gravity: there are claims that polytropic stars in $f(R)$ and scalar-tensor gravity cannot be matched to SdSK. If they are interesting for cosmology, these theories contain a time-dependent $\Lambda$. Perhaps SdSK is not the correct solution to match internal solutions to—the situation is unclear. It does not help if the situation is unclear even in GR, which is what we clarify here.

## 2. Putative Alternatives to SdSK

Here we compare putative alternative geometries to SdSK, using the areal radius as the radial coordinate.

### 2.1. *Abbassi-Meissner proposal*

It is claimed that an alternative to SdSK is the Abbassi-Meissner solution[8–10]

$$ds^2 = -f(t,r)dt^2 + \frac{\mathrm{e}^{2Ht}}{f(t,r)}\,dr^2 + \mathrm{e}^{2Ht}r^2 d\Omega^2_{(2)}\,, \tag{3}$$

where $H = \sqrt{\Lambda/3}$, $m = \text{const.}$, and

$$f(t,r) = h(t,r) + \sqrt{h^2(t,r) + H^2 r^2 \mathrm{e}^{2Ht}}\,, \tag{4}$$

$$h(t,r) = \frac{1}{2}\left(1 - H^2 r^2 \mathrm{e}^{2Ht} - \frac{2m}{r}\,\mathrm{e}^{-Ht}\right). \tag{5}$$

Let us change to the areal radius $R = \mathrm{e}^{Ht}r$ instead of $r$, and to the new time $T$ defined by

$$dT = dt + \frac{2HRdR}{A_0\left(A_0 + \sqrt{A_0^2 + 4H^2R^2}\right)}\,, \tag{6}$$

where

$$A_0(R) = 1 - \frac{2m}{R} - H^2R^2 = 2h(0,R)\,. \tag{7}$$

The line element becomes diagonal and locally static

$$ds^2 = -A_0(R)dT^2 + \frac{2}{A_0(R) + \sqrt{A_0^2(R) + 4H^2R^2}} \cdot \left[1 + \frac{2H^2R^2}{A_0\left(A_0 + \sqrt{A_0^2 + 4H^2R^2}\right)}\right] dR^2 + R^2 d\Omega^2_{(2)}\,, \tag{8}$$

and it is clearly not SdSK. It solves the field equations $G_{ab} = -\Lambda g_{ab} + T_{ab}$ with a radial flow.[11]

### 2.2. *Non-rotating Thakurta solution*

The Thakurta solution of GR[12] is conformal to Kerr and describes a rotating black hole embedded in a FLRW universe. The non-rotating, spherical subcase is the late time limit of generalized McVittie solutions[13] and it is also the $\omega \to \infty$ limit of a class of Brans-Dicke perfect fluid solutions describing inhomogeneous universes found by Ref. 14. The non-rotating Thakurta line element is

$$\begin{aligned} ds^2 &= a^2(\eta)\left[-\left(1-\frac{2m}{r}\right)d\eta^2 + \frac{dr^2}{1-2m/r} + r^2 d\Omega^2_{(2)}\right] \\ &= -\left(1-\frac{2m}{r}\right)dt^2 + \frac{a^2 dr^2}{1-2m/r} + a^2 r^2 d\Omega^2_{(2)}\,, \end{aligned} \tag{9}$$

where $a(\eta)$ is the scale factor of the FLRW "background". Changing coordinates to $(t, r) \to (T, a(t)r)$ with

$$dT = \frac{1}{F}\left(dt + \frac{HRdR}{A^2 - H^2R^2}\right)\,, \tag{10}$$

$$A(t,R) = 1 - \frac{2m}{r} = 1 - \frac{2M}{R}\,, \qquad M \equiv ma(t)\,, \tag{11}$$

the line element becomes

$$ds^2 = -\left(1 - \frac{2M}{R} - \frac{H^2R^2}{1-\frac{2M}{R}}\right)F^2 dT^2 + \frac{dR^2}{1-\frac{2M}{R} - \frac{H^2R^2}{1-2M/R}} + R^2 d\Omega^2_{(2)} \tag{12}$$

(where $F$ is an integrating factor) and it solves the field equations $G_{ab} = -\Lambda g_{ab} + q_a u_b + q_b u_a$, where $q^c$ (with $q^c u_c = 0$) describes a purely spatial radial flow.

### 2.3. *Castelo Ferreira metric*

The Castelo Ferreira line element[15] is

$$\begin{aligned} ds^2 = &-\left[1 - \frac{2m}{R} - H^2R^2\left(1-\frac{2m}{R}\right)^{\alpha}\right]dt^2 + \frac{dR^2}{1-\frac{2m}{R}} - 2HR\left(1-\frac{2m}{R}\right)^{\frac{\alpha-1}{2}} dtdR \\ &+R^2 d\Omega^2_{(2)} \end{aligned} \tag{13}$$

with $\alpha, m$ constants and $H = H(t)$. The coordinate change

$$dT = \frac{1}{F}\left(dt + \frac{HRdR\left(1-\frac{2m}{R}\right)^{\frac{\alpha-1}{2}}}{1-\frac{2m}{R} - H^2R^2\left(1-\frac{2m}{R}\right)^{\alpha}}\right) \tag{14}$$

turns the line element into

$$\begin{aligned} ds^2 = &-\left[1 - \frac{2m}{R} - H^2R^2\left(1-\frac{2m}{R}\right)^{\alpha}\right]F^2 dT^2 \\ &+\frac{dR^2}{1-\frac{2m}{R} - H^2R^2\left(1-\frac{2m}{R}\right)^{\alpha}} + R^2 d\Omega^2_{(2)}\,. \end{aligned} \tag{15}$$

One can choose the background to be de Sitter space, then one obtains the non-rotating Thakurta solution if $\alpha = -1$ and the SdSK solution if $\alpha = 0$. However, the general Castelo Ferreira metric is not SdSK.

## 3. A Simultaneous No-Hair/Cosmic No Hair Theorem

The Jebsen-Birkhoff no-hair theorem states that the Schwarzschild solution is generic; the cosmic no-hair theorem states that de Sitter space is generic. Then, does a simultaneous no-hair/cosmic no-hair theorem exist, stating that SdSK is generic in some sense?

One can write any spherical metric as

$$ds^2 = -A^2(t,R)dt^2 + B^2(t,r)dR^2 + R^2 d\Omega^2_{(2)} \tag{16}$$

without loss of generality. Now one needs some assumptions about the matter content. In the previous examples of putative SdSK solutions, a common ingredient was the imperfect fluid stress-energy tensor

$$T_{ab} = (P+\rho)\, u_a u_b + P g_{ab} + q_a u_b + q_b u_a \,, \quad u_c u^c = -1 \,, \quad q_c u^c = 0 \,, \tag{17}$$

so we assume 1) this $T_{ab}$ with barotropic and constant equation of state $P = w\rho$, with $w = \text{const.}$; 2) spherical symmetry; 3) the solution of $G_{ab} = -\Lambda g_{ab} + 8\pi T_{ab}$ is asymptotically de Sitter: there is a de Sitter-like cosmological horizon of radius $R_H$ and the solution reduces to SdSK as $R \to R_H$.

The Einstein equations become

$$\frac{\dot{B}}{BR} = 4\pi T_{01} \,, \tag{18}$$

$$A^2 \left( \frac{2B'}{B^3 R} - \frac{1}{B^2 R^2} + \frac{1}{R^2} \right) = \Lambda A^2 + 8\pi T_{00} \,, \tag{19}$$

$$\frac{2A'}{AR} - \frac{B^2}{R^2} + \frac{1}{R^2} = -\Lambda B^2 + 8\pi T_{11} \,, \tag{20}$$

$$\frac{A'B}{A} - B' - \frac{RB^2\ddot{B}}{A^2} + \frac{R\dot{A}\dot{B}B^2}{A^3} - \frac{RA'B'}{A} + \frac{RA''B}{A}$$
$$= \left(-\Lambda R^2 + 8\pi T_{22}\right) \frac{B^3}{R} \,. \tag{21}$$

The fluid 4-velocity and energy flux density are

$$u_\mu = (-|A|, 0, 0, 0) \,, \qquad q_\mu = \left(0, B^2 q, 0, 0\right) \,, \tag{22}$$

while the stress-energy tensor is

$$T_{00} = A^2 \rho \,, \qquad T_{01} = -|A| B^2 q \,, \tag{23}$$

$$T_{11} = B^2 P \,, \qquad T_{22} = \frac{T_{33}}{\sin^2\theta} = R^2 P \,, \tag{24}$$

where $T_{01} > 0$ and $q < 0$ correspond to radial inflow.

In the case of inflow $q < 0$ we have

$$(B^2)^{\cdot} = -8\pi|A|B^4 Rq > 0\,, \tag{25}$$

therefore $B^2 = g_{11}$ increases with time. Assuming the metric coefficients to be continuous and differentiable, there are then two possibilities:

- either $B^2(t,R) \to +\infty$ for any fixed $R$ as $t \to +\infty$, or
- $B^2(t,R)$ has an horizontal asymptote as $t \to +\infty$.

Consider the first case. The apparent horizons are located by

$$\nabla^c R \nabla_c R = 0 \quad \leftrightarrow \quad 1/B^2 = 0. \tag{26}$$

If $B^2 \to +\infty$ as $t \to +\infty$ or as $t \to t_{max}$, then at late times all points of space (at any $R < R_H$) lie arbitrarily close to an apparent horizon. This situation is familiar in cosmology: it corresponds to a phantom universe ending in a Big Rip singularity at $t_{max}$ and the apparent horizon shrinks around a comoving observer as the cosmic expansion super-accelerates. This phantom asymptotics contradict the assumption of de Sitter asymptotics and we discard this case.

In the other case in which $B^2(t,R) \to B_0^2(R)$ as $t \to +\infty$, we have $\dot{B} \to 0$ as $t \to +\infty$. Then the radial flow $q \to 0$ as $t \to +\infty$. Differentiate the (0,0) Einstein equation to obtain

$$8\pi\dot{\rho} = \frac{2}{R}\left(\frac{B'}{B^3}\right)^{\cdot} - \frac{1}{R^2}\left(\frac{1}{B^2}\right)^{\cdot} \to 0 \quad \text{ as } \; t \to +\infty\,. \tag{27}$$

Then the equation of state $P = w\rho$ yields $\dot{P} \to 0$ as $t \to +\infty$. The (2,2) (or (3,3)) equation gives

$$8\pi\dot{P} = \frac{2}{R}\left(\frac{A'}{AB^2}\right)^{\cdot} + \frac{1}{R^2}\left(\frac{1}{B^2}\right)^{\cdot} \approx \frac{2}{RB^2}\left(\frac{A'}{A}\right)^{\cdot} \to 0 \tag{28}$$

as $t \to +\infty$, then $A^2$ also becomes time-independent, and the metric becomes static. To make progress, use the covariant conservation equation $\nabla^b T_{ab} = 0$ for the imperfect fluid, obtaining

$$\begin{aligned} &u_a u^b \nabla_b (P+\rho) + [(P+\rho)\, u_a + q_a]\, \nabla^b u_b \\ &+ [(P+\rho)\, u_b + q_b]\, \nabla^b u_a + \nabla_a P + u^b \nabla_b q_a + u_a \nabla^b q_b = 0\,. \end{aligned} \tag{29}$$

Projecting onto the time direction $u^a$ and using $u^a \nabla_b u_a = 0$ leads to

$$-\dot{\rho} - (P+\rho)\, \nabla^b u_b + u^a q^b \nabla_b u_a + u^a u^b \nabla_b q_a - \nabla^b q_b = 0\,. \tag{30}$$

At late times $q^c$ and $\dot{\rho}$ disappear from this equation, leaving $(P+\rho)\, \nabla^b u_b \simeq 0$. In general $\nabla^b u_b \neq 0$ (this quantity reduces to $3H > 0$ for large $r$) and we are left with $P + \rho \to 0$ as $t \to +\infty$. Either the fluid reduces to a pure $\Lambda$ (then the vacuum uniqueness theorem for SdSK holds trivially), or else both $\rho$ and $P = w\rho$ become subdominant and $\Lambda$ dominates. In this case the solution also reduces to SdSK.

In the case of outflow $q > 0$, one has instead $(B^2)^{\cdot} = -8\pi|A|B^4Rq < 0$; since $B^2$ is bounded from below by zero and it decreases as $t \to +\infty$, it must have an horizontal asymptote with $B^2(t, R) \to B_0^2(R)^+$ as $t \to +\infty$. Then $\dot{B} \to 0$ and $q \to 0$ and we repeat the reasoning done for $q < 0$ from here.

A special case is that of a perfect fluid $q^a = 0$; then $T_{01} = 0$ and the (0,1) equation gives $B = B(R)$. It is then straightforward to prove that it must be $P = -\rho$ and that SdSK is the unique solution.[16]

## 4. Conclusions

A generalized Jebsen-Birkhoff theorem holds for $\Lambda \neq 0$ but, although its proof is straightforward, it does not appear in modern GR textbooks. In particular, for $\Lambda > 0$, SdSK is the unique spherical vacuum solution. Putative alternatives to SdSK in this situation are either non-vacuum solutions or SdSK in disguise. Going beyond the vacuum case, we have proved a simultaneous no-hair and cosmic no-hair theorem in the presence of $\Lambda$ and a radial purely spatial heat flow. Further generalization to other forms of matter will be pursued in the future.

## Acknowledgment

V.F. is supported by the Natural Sciences and Engineering Research Council of Canada (grant no. 2016-03803).

## References

1. L. Amendola and S. Tsujikawa, *Dark Energy, Theory and Observations* (Cambridge University Press, Cambridge, UK, 2010).
2. J.L. Synge, *Relativity: The General Theory* (North Holland, Amsterdam, 1960).
3. K. Schleich and D. Witt, *J. Math. Phys.* **51**, 112502 (2010).
4. M. Boucher, G.W. Gibbons, and G.T. Horowitz *Phys. Rev. D* **30**, 2447 (1984).
5. H. Kodama, *J. Korean Phys. Soc.* **45**, S68 (2004).
6. P.G. LeFloch and L. Rozoy, *Compt. Rendu. Acad. Sci. Paris Ser. I* **348**, 1129 (2010).
7. A.K.M. Masood-ul-Alam and W. Yu, *Comm. Analys. Geom.* **23**, 377 (2015).
8. A.H. Abbassi, *J. High Energy Phys.* **04**, 011 (1999).
9. A.H. Abbassi, S. Gharanfoli, and A.M. Abbassi, *Apeiron* **9**, 1 (2002).
10. K.A. Meissner, `arXiv:0901.0640`.
11. V. Faraoni, A.M. Cardini, and W.-J. Chung, *Phys. Rev. D* **97**, 024046 (2018).
12. S.N.G. Thakurta, *Indian J. Phys.* **55B**, 304 (1981).
13. V. Faraoni, C. Gao, X. Chen, and Y.-G. Shen, *Phys. Lett. B* **95**, 084044 (2017).
14. T. Clifton, D.F. Mota, and J.D. Barrow, *Mon. Not. Roy. Astron. Soc.* **358**, 601 (2005).
15. P. Castelo Ferreira, *Phys. Lett. B* **684**, 73 (2010).
16. V. Faraoni and A. Jacques, *Phys. Rev. D* **76**, 063510 (2007).

# Higher dimensional conformal-$U(1)$ gauge/gravity black holes

Seyed Hossein Hendi[1,2*]

[1] *Physics Department and Biruni Observatory, College of Sciences, Shiraz University, Shiraz 71454, Iran*
[2] *Research Institute for Astronomy and Astrophysics of Maragha (RIAAM), P.O. Box 55134-441, Maragha, Iran*

Motivated by quantum nature of gravitating black holes, higher dimensional exact solutions of conformal gravity with an abelian gauge field is obtained. It is shown that the obtained solutions can be interpreted as singular black holes.

## 1. Introduction

Considering the quantum effects in gravitational interaction, one may find that the higher-curvature modification of general relativity is inevitable. However, in order to have a physically ghost free theory of higher-curvature modifications, some special constraints should be applied. Fortunately, there are known higher-curvature interesting renormalizable actions with no ghosts under certain criterion. As an interesting example, we can regard the so-called Conformal Gravity (CG), which is defined by the square of the Weyl tensor [1, 2].

The CG is an interesting theory of modified general relativity with a remarkable property which is sensitive to angles, but not distances. In other words, it is invariant under local stretching of the metric which is called the Weyl transformation, $g_{\mu\nu}(x) \to \Omega^2(x) g_{\mu\nu}(x)$. It has been shown that CG is useful for constructing supergravity theories [3, 4] and can be considered as a possible UV completion of gravity [5–7]. It may be also arisen from twister-string theory with both closed strings and gauge-singlet open strings [8]. Moreover, CG can be appeared as a counterterm in $adS_5/CFT_4$ calculations [9, 10]. In addition to the motivations mentioned above, solving the dark matter and dark energy problems are two of the most important and interesting motivations of studying CG theory [7].

Since CG is renormalizable [5, 11] and the requirement of conformal invariance at the classical level leads to a renormalizable gauge theory of gravity, it seems interesting to consider black holes in CG which permits a consistent picture of black hole evaporation [12]. The first attempt to obtain the spherically symmetric black hole solutions in four dimensions has been done by Bach [13], and then, Buchdahl has considered a particular case of the conformal solutions in [14]. It is worthwhile to mention that the 4-dimensional solution of Einstein gravity is a solution of CG as well. In addition, it has been shown that the Einstein solutions can be obtained by considering suitable boundary condition on the metric in CG [15, 16].

CG can also be introduced in higher dimensions ($D > 4$), straightforwardly [17]. Nevertheless, unlike the 4-dimensional case, CG does not admit Einstein trivial

*email address: hendi@shirazu.ac.ir

solutions in higher dimensions. This nontrivial behavior is due to the fact that in contrast to the 4-dimensional action, the Kretschmann scalar $R^{\alpha\beta\gamma\delta}R_{\alpha\beta\gamma\delta}$ contributes dynamically in the higher dimensions [18]. Such a nontrivial behavior motivates one to investigate higher dimensional CG black hole solutions.

The outline of this paper is as follows. In the next section, we give a brief review on neutral and charged black holes of CG in 4-dimensional spacetime. We also construct $D$-dimensional topological static black hole solutions of CG gravity in the presence of generalized Maxwell theory. We finish our paper with some concluding remarks.

## 2. Four-dimensional exact solutions

At the first step, we consider a four-dimensional conformal action as

$$\begin{aligned} I_G &= -\alpha \int d^4x\sqrt{-g}C_{\lambda\mu\nu\kappa}C^{\lambda\mu\nu\kappa} \\ &\equiv -2\alpha \int d^4x\sqrt{-g}\left[R^{\mu\nu}R_{\mu\nu} - \frac{1}{3}(R^{\alpha}{}_{\alpha})^2\right], \end{aligned} \tag{1}$$

where the Weyl conformal tensor is

$$\begin{aligned} C_{\lambda\mu\nu\kappa} &= R_{\lambda\mu\nu\kappa} + \frac{1}{6}R^{\alpha}{}_{\alpha}\left[g_{\lambda\nu}g_{\mu\kappa} - g_{\lambda\kappa}g_{\mu\nu}\right] \\ &\quad - \frac{1}{2}\left[g_{\lambda\nu}R_{\mu\kappa} - g_{\lambda\kappa}R_{\mu\nu} - g_{\mu\nu}R_{\lambda\kappa} + g_{\mu\kappa}R_{\lambda\nu}\right]. \end{aligned} \tag{2}$$

Variation of action (1) with respect to the metric tensor leads to the following equation of motion

$$\begin{aligned} &W^{\mu\nu} = 2C^{\mu\lambda\nu\kappa}{}_{;\lambda\kappa} - C^{\mu\lambda\nu\kappa}R_{\lambda\kappa} = \\ &\frac{1}{2}g^{\mu\nu}(R^{\alpha}{}_{\alpha})^{;\beta}{}_{;\beta} + R^{\mu\nu;\beta}{}_{;\beta} - R^{\mu\beta;\nu}{}_{;\beta} - R^{\nu\beta;\mu}{}_{;\beta} - 2R^{\mu\beta}R^{\nu}{}_{\beta} \\ &+\frac{1}{2}g^{\mu\nu}R_{\alpha\beta}R^{\alpha\beta} - \frac{2}{3}g^{\mu\nu}(R^{\alpha}{}_{\alpha})^{;\beta}{}_{;\beta} + \frac{2}{3}(R^{\alpha}{}_{\alpha})^{;\mu;\nu} + \frac{2}{3}R^{\alpha}{}_{\alpha}R^{\mu\nu} - \frac{1}{6}g^{\mu\nu}(R^{\alpha}{}_{\alpha})^2 = 0. \end{aligned} \tag{3}$$

It was shown that the static spherically symmetric solution of conformal gravity in four dimensions can be written as

$$ds^2 = -f(r)dt^2 + \frac{dr^2}{f(r)} + r^2d\Omega^2, \tag{4}$$

where $d\Omega^2$ is the line element of a 2−sphere, $S^2$, and the metric function is [14]

$$f(r) = c_0 + \frac{d}{r} + \frac{c_0^2 - 1}{3d}r - \frac{1}{3}\Lambda r^2. \tag{5}$$

It is clear that for nonvanishing $\Lambda$, Eq. (5) is not a solution of Einstein gravity, while as long as $\Lambda = 0$, the metric becomes identical to the Schwarzschild solution of Einstein gravity. It is worth mentioning that although $\Lambda$ plays the role of the cosmological constant, it is arisen purely as an integration constant and is not put in the action by hand. Such a constant cannot be added to the action of CG because it would introduce a length scale and hence break the conformal invariance.

In order to add an action of matter, we should take care of its conformal transformation to keep the theory be conformally invariant. Fortunately, the Lagrangian of Maxwell theory is conformal invariant in four dimensions and we can add it to the gravitational sector of conformal theory as an appropriate matter field.

So, we can consider the static charged adS solutions of conformal$-U(1)$ gravity in four dimensions [13]. The appropriate action is

$$I = \alpha \int d^4x\sqrt{-g}\left(\frac{1}{2}C^{\mu\nu\rho\sigma}C_{\mu\nu\rho\sigma} + \frac{1}{3}F^{\mu\nu}F_{\mu\nu}\right), \tag{6}$$

where the unusual sign in front of the Maxwell term comes from the so-called critical gravity, which is necessary to recover the Einstein gravity from conformal gravity in IR limit. The static topological solution is found in [19] the same as Eq. (5) with the following gauge potential one form

$$A = -\frac{Q}{r}dt, \tag{7}$$

which leads to the following metric function [19]

$$f(r) = c_0 + \frac{d}{r} + c_1 r - \frac{1}{3}\Lambda r^2. \tag{8}$$

In order to have consistent solutions, three integration constants $c_0$, $c_1$, and $d$ should obey an algebraic constraint

$$3c_1 d + \varepsilon^2 + Q^2 = c_0^2, \tag{9}$$

where $\varepsilon = 1, -1, 0$ denotes spherical, hyperbolic, and planar horizons, respectively. Therefore, the metric function takes the following compact form

$$f(r) = c_0 + \frac{d}{r} + \frac{c_0^2 - \varepsilon^2 - Q^2}{3d}r - \frac{1}{3}\Lambda r^2. \tag{10}$$

## 3. Higher dimensional solutions

Here, we are going to generalize the conformal action of $U(1)$-gauge/gravity coupling in higher dimensions. As we know, the Maxwell action does not enjoy the conformal invariance properties in higher dimensions, and therefore, the higher dimensional solutions in CG cannot be produced in the presence of Maxwell field (and also the other electrodynamic fields that are not conformal invariant in higher dimensions).

So, we should consider a generalization of linear Maxwell action to the case that it respects the invariance of conformal transformation. To do so, we take into account the power Maxwell nonlinear theory, which its Lagrangian is a power of Maxwell invariant, $(-F_{\mu\nu}F^{\mu\nu})^s$. It is a matter of calculation to show that the power Maxwell action enjoys the conformal invariance, for $s = D/4$ ($D =$ dimension of spacetime) [20]. In other words, it is easy to show that as long as $s = D/4$, the energy-momentum tensor of power Maxwell invariant theory is traceless [20].

Regarding the mentioned issues, we find that the suitable action of higher dimensional conformal $U(1)$-gauge/gravity action can be written as

$$I = \frac{1}{16\pi}\int d^D x\sqrt{-g}\left(C^{\mu\nu\rho\sigma}C_{\mu\nu\rho\sigma} + \beta(-F^{\mu\nu}F_{\mu\nu})^{\frac{D}{4}}\right). \tag{11}$$

Hereafter, we can regard a higher dimensional static spacetime and look for exact solutions with black hole interpretation. Variation of this action with respect to the metric tensor $g_{\mu\nu}$ and the Faraday tensor $F_{\mu\nu}$ leads to the following field equations

$$\begin{aligned}\mathbf{E}_{\rho\sigma} &= \left(\nabla^\mu\nabla^\nu + \frac{1}{2}R^{\mu\nu}\right)C_{\rho\nu\mu\sigma} + \frac{\beta}{8}\left[g_{\rho\sigma}(-F^{\mu\nu}F_{\mu\nu})^{\frac{D}{4}} + D(-F^{\mu\nu}F_{\mu\nu})^{\frac{D}{4}-1}F_{\sigma\delta}F_{\rho}{}^{\delta}\right]\\ &= 0,\end{aligned} \tag{12}$$

$$\partial_\rho\left[\sqrt{-g}(-F^{\mu\nu}F_{\mu\nu})^{\frac{D}{4}-1}F^{\rho\sigma}\right] = 0. \tag{13}$$

Since we are looking for topological black hole solutions of mentioned field equations, we express the metric of a $D$-dimensional spacetime as follows

$$ds^2 = -f(r)dt^2 + f^{-1}(r)dr^2 + r^2 d\Sigma^2_{k,D-2}, \tag{14}$$

where $k$ denotes spherical ($k = 1$), hyperbolic ($k = -1$), and planar ($k = 0$) horizons of the $(D-2)$-dimensional manifold with the following line element

$$d\Sigma^2_{k,D-2} = \begin{cases} d\Omega^2_{D-2} = d\theta_1^2 + \sum\limits_{i=2}^{D-2}\prod\limits_{j=1}^{i-1}\sin^2\theta_j d\theta_i^2 & k=1 \\ d\Xi^2_{D-2} = d\theta_1^2 + \sinh^2\theta_1\left(d\theta_2^2 + \sum\limits_{i=3}^{D-2}\prod\limits_{j=2}^{i-1}\sin^2\theta_j d\theta_i^2\right) & k=-1 \\ dl^2_{D-2} = \sum\limits_{i=1}^{D-2} d\theta_i^2 & k=0 \end{cases}, \tag{15}$$

in which $d\Omega^2_{D-2}$ is the standard metric of a unit $(D-2)$-sphere, $d\Xi^2_{D-2}$ is the metric of a $(D-2)$-dimensional hyperbolic plane with unit curvature, and $dl^2_{D-2}$ is the flat metric of $R^{D-2}$.

Using this metric and a radial gauge potential ansatz $A_\mu = -qr^{(2s-D+1)/(2s-1)}\delta^0_\mu$, one can find the nonzero components of the theory as follows

$$\begin{aligned}\mathbf{E}_{tt} &= 2D_3r^4f(r)f^{(4)}(r)+D_3\left[rf'(r)+4D_{5/2}f(r)\right]r^3f'''(r)\\&+\left[\frac{D_3}{2}r^2f''(r)+\left(2D^2D_{23/2}+85D_{102/85}\right)f(r)+D_3D_4\left(\frac{3rf'(r)}{2}+k\right)\right]r^2f''(r)\\&-\left[3D_{10/3}D_3rf'(r)+\left(4D^2D_{49/4}+187D_{228/187}\right)f(r)+5kD_3D_{16/5}\right]rf'(r)\\&+2\left(2D^2D_{23/2}+84D_{99/84}\right)f^2(r)-4k\left(D^2D_{12}+45D_{54/45}\right)f(r)-2D_3^2k^2\\&-\frac{D_1}{2}r^4\beta(1-2s)\left(\frac{\sqrt{2}q(2s-D_1)}{(2s-1)r^{(D-2)/(2s-1)}}\right)^{2s},\end{aligned}\tag{16}$$

$$\begin{aligned}\mathbf{E}_{rr} &= D_3\left[2f(r)-rf'(r)\right]r^3f'''(r)-D_3\left[\frac{r^2}{2}f''(r)-3D_{10/3}f(r)+D_4\left(\frac{3}{2}rf'(r)+k\right)\right]\\&\times r^2f''(r)+D_3\left[3D_{10/3}rf'(r)-9D_{28/9}f(r)+5kD_{16/5}\right]rf'(r)\\&+2D_3^2\left[3f^2(r)-4kf(r)+k^2\right]+\frac{D_1}{2}r^4\beta(1-2s)\left(\frac{\sqrt{2}q(2s-D_1)}{(2s-1)r^{(D-2)/(2s-1)}}\right)^{2s},\end{aligned}\tag{17}$$

$$\begin{aligned}\mathbf{E}_{\theta\theta} &= D_3r^4f(r)f^{(4)}(r)+D_3\left[rf'(r)+2D_3f(r)\right]r^3f'''(r)\\&+\left[\frac{D_3}{2}r^2f''(r)+\left(D^2D_{13}+52D_{66/52}\right)f(r)+D_3D_4\left(\frac{3rf'(r)}{2}+k\right)\right]r^2f''(r)\\&-\left[3D_{10/3}D_3rf'(r)+\left(2D^2D_{29/2}+121D_{156/121}\right)f(r)+5kD_3D_{16/5}\right]rf'(r)\\&+2\left(D^2D_{13}+51D_{63/51}\right)f^2(r)-2k\left(D^2D_{14}+57D_{72/57}\right)f(r)-2D_3^2k^2\\&+\frac{D_1D_2}{4}r^4\beta\left(\frac{\sqrt{2}q(2s-D_1)}{(2s-1)r^{(D-2)/(2s-1)}}\right)^{2s},\end{aligned}\tag{18}$$

in which we used $D_i = D-i$ for convenience and prime refers to $d/dr$. However, we will use common $D-i$ in indices and powers for clarity of equations. Solving Eqs. (16)-(18), and keeping in mind that $s = D/4$, we can obtain the metric function for $D \geq 5$ as follows

$$f(r) = k - \frac{C_1}{r^{D-3}} - \frac{\beta\mu q^{D/2}}{C_1}r + C_2r^2,\tag{19}$$

where $C_1$ and $C_2$ are two integration constants, and $\mu = 2^{(D-4)/4}\left(D_2D_3\right)^{-1}$ is a dimensionful constant. Interestingly, we see that although the field equations are too complicated, the solutions are quite simple. It is noticeable to mention that the solution to $D = 4$ is given by (10) when we set $\beta = -2/3$ in (11), but it cannot be obtained by using the field equations (16)-(18). Therefore, the four-dimensional spacetime has one more integration constant compared to the other dimensions. In order to obtain a compact form of the metric function to be valid for all dimensions, one can consider a special case $c_0 = \varepsilon$ of the four-dimensional metric function (10).

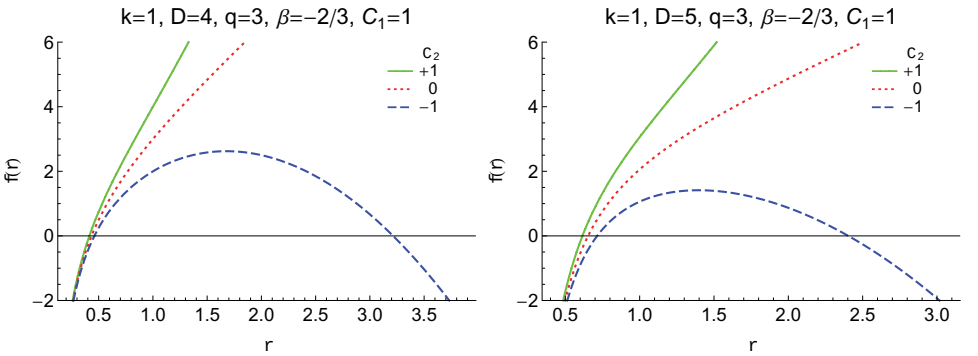



Fig. 1. The metric function versus radial coordinate. Positive $C_2$ corresponds to asymptotic adS solutions whereas the negative sign belongs to dS ones. In addition, for zero value of this parameter, the asymptotic behavior of these solutions are neither flat nor (a)dS. This is correct until the factor of $r$ in the metric function (19) is a nonzero value.

In this situation, the metric function of $D \geq 4$ is given by (19) and the field equations (16)-(18) can be used for $D = 4$ as well. We should note that since we considered a special case, $c_0 = \varepsilon$, the 4-dimensional solution given in (19) is also a solution of the field equations of (6).

Having solutions at hand, we are in a position to check that these solutions can be considered as a black hole or not. To do so, we first look for the singularities of the solutions. By calculating the Kretschmann scalar

$$R^{\lambda\tau\rho\sigma}R_{\lambda\tau\rho\sigma} = 2DD_1C_2^2 - \frac{4D_1D_2C_2\beta\mu q^{D/2}}{C_1 r} + \frac{2D_2^2\beta^2\mu^2q^D}{C_1^2r^2} + \frac{D_1D_2^2D_3C_1^2}{r^{2(D-1)}}, \quad (20)$$

one can easily find that the metric (14) with the metric function (19) has an essential singularity at the origin ($\lim\limits_{r\to 0}\left(R^{\lambda\tau\rho\sigma}R_{\lambda\tau\rho\sigma}\right) = \infty$). In addition, Fig. 1 shows that this singularity can be covered with an event horizon, and therefore, we can interpret the solution as a singular black hole.

## 4. Conclusions

Motivated by the importance of higher dimensional spacetime in high energy physics, we have constructed the conformal-$U(1)$ gauge/gravity black hole solutions in $D \geq 4$. Since higher-dimensional solutions in CG cannot be produced in the presence of Maxwell field (and also the other electrodynamic fields that are not conformal invariant in higher dimensions), we have used a class of nonlinear electrodynamics $(-F_{\mu\nu}F^{\mu\nu})^s$ (which enjoys the conformal invariance properties in higher dimensions as $s = D/4$) to obtain black hole solutions. In addition, we have seen that the obtained solutions enjoy an essential singularity at the origin and they can be considered as black holes.

## References


[1] R. J. Riegert, Phys. Rev. Lett. 53 (1984) 315.
[2] H. Lu and C. N. Pope, Phys. Rev. Lett. 106 (2011) 181302.
[3] E. Bergshoeff, M. de Roo and B. de Wit, Nucl. Phys. B 182 (1981) 173.
[4] B. de Wit, J. W. van Holten and A. Van Proeyen, Nucl. Phys. B 184 (1981) 77.
[5] S. L. Adler, Rev. Mod. Phys. 54 (1982) 729.
[6] G. 't Hooft, Found. Phys. 41 (2011) 1829.
[7] P. D. Mannheim, Found. Phys. 42 (2012) 388.
[8] N. Berkovits and E. Witten, JHEP 08 (2004) 009.
[9] H. Liu and A. A. Tseytlin, Nucl. Phys. B 533 (1998) 88.
[10] V. Balasubramanian, E. G. Gimon, D. Minic and J. Rahmfeld, Phys. Rev. D 63 (2001) 104009.
[11] K. S. Stelle, Phys. Rev. D 16 (1977) 953.
[12] B. Hasslacher and E. Mottola, Phys. Lett. B 99 (1981) 221.
[13] R. Bach, Math. Z. 9 (1921) 110.
[14] A. Buchdahl, Edinburgh Math. Soc. Proc. 10 (1953) 16.
[15] J. Maldacena, [arXiv:1105.5632].
[16] G. Anastasiou and R. Olea, Phys. Rev. D 94 (2016) 086008.
[17] S. Deser, H. Liu, H. Lu, C. N. Pope, T. C. Sisman and B. Tekin, Phys. Rev. D 83 (2001) 061502.
[18] C. Lanczos, Annals Math. 39 (1938) 842.
[19] J. Li, H. S. Liu, H. Lu and Z. L. Wang, JHEP 02 (2013) 109.
[20] M. Hassaine and C. Martinez, Phys. Rev. D 75 (2007) 027502.

# Kundt geometries in higher-derivative gravity

Ondřej Hruška*, Robert Švarc† and Jiří Podolský‡

*Institute of Theoretical Physics, Faculty of Mathematics and Physics, Charles University, Prague, V Holesovickach 2, 180 00 Praha 8, Czech Republic*
*E-mail: *hruskaondrej@seznam.cz, †robert.svarc@mff.cuni.cz*
*‡podolsky@mbox.troja.mff.cuni.cz*
*www.mff.cuni.cz*

Quadratic Gravity is one of the modified theories of gravity; its action contains additional terms quadratic in Riemann tensor and its contractions. We study Kundt metrics, defined geometrically by admitting non-expanding, non-twisting and non-shearing null geodesic congruence, in the framework of Quadratic Gravity. We examine the special cases, namely the *pp*-waves and VSI spacetimes. Equations of geodesic deviation are calculated to provide means of physical interpretation.

*Keywords*: Quadratic Gravity; Kundt spacetimes; Geodesic deviation.

## 1. Introduction

Despite many successes of general relativity, including the prediction of gravitational waves that were directly detected nearly four years ago, even such brilliant theory has its limits. Attempts to solve problems of general relativity (e.g., quantization, inflation, dark matter and energy) involve modifying the action.[1] Our focus is on the so-called Quadratic Gravity (QG), where the action contains additional terms that are quadratic combinations of Riemann tensor and its contractions

$$S = \int \left(\tfrac{1}{\kappa}(R - 2\Lambda) + \alpha\, R^2 + \beta\, R_{ab}^2 + \gamma(R_{abcd}^2 - 4R_{ab}^2 + R^2)\right) \sqrt{-g}\, \mathrm{d}^D x\,. \qquad (1)$$

The vacuum field equations derived from this action are then

$$\begin{aligned}&\tfrac{1}{\kappa}\left(R_{ab} - \tfrac{1}{2}R\, g_{ab} + \Lambda\, g_{ab}\right) + 2\alpha\, R\left(R_{ab} - \tfrac{1}{4}R\, g_{ab}\right) + (2\alpha + \beta)(g_{ab}\Box - \nabla_a\nabla_b)R\\ &+2\gamma\left(R\, R_{ab} - 2R_{abcd}\, R^{cd} + R_{acde}\, R_b^{\ cde} - 2R_{ac}R_b^{\ c} - \tfrac{1}{4}g_{ab}(R_{cdef}^2 - 4R_{cd}^2 + R^2)\right)\\ &+\beta\Box\left(R_{ab} - \tfrac{1}{2}R\, g_{ab}\right) + 2\beta\left(R_{abcd} - \tfrac{1}{4}g_{ab}R_{cd}\right) R^{cd} = 0\,. \qquad (2)\end{aligned}$$

We can easily see that for $\alpha = \beta = \gamma = 0$, we get famous Einstein's general relativity. In general, Eq. (2) contains second order derivatives of the Ricci scalar and tensor, which implies the fourth order derivatives of the metric. A special case of QG is Gauss–Bonnet theory ($\alpha = 0 = \beta$), that belongs to so called Lovelock's gravity.[2] Our aim is to study spacetimes known in general relativity, which are defined by purely geometrical properties independent of the specific field equations. Metric describing such spacetimes can then be considered as an ansatz and inserted into the QG field equations to observe restrictions on the metric functions and differences between QG and classic general relativity.

Specific representative of such spacetimes is the Kundt family whose line element can be written as[3–5]

$$\mathrm{d}s^2 = g_{pq}(u,x)\,\mathrm{d}x^p\mathrm{d}x^q + 2g_{up}(r,u,x)\,\mathrm{d}u\,\mathrm{d}x^p - 2\mathrm{d}u\,\mathrm{d}r + g_{uu}(r,u,x)\,\mathrm{d}u^2\,. \qquad (3)$$

Its *geometric definition* is that it admits non-expanding, non-twisting, and non-shearing null geodesic congruence. The optically privileged vector field generating the congruence is $\boldsymbol{k} = \partial_r$, where $r$ is an affine parameter along the congruence. The spacetime is foliated by the null surfaces $u = \text{const.}$, to which $\boldsymbol{k}$ is orthogonal (and tangential). The transverse $(D-2)$-space: $u = \text{const.}$ and $r = \text{const.}$ is described by the metric $g_{pq}(u, x)$, where $x \equiv x^k$, $k = 2, \ldots D-1$ are coordinates covering the transverse space. For simplicity, we will study Kundt metrics with $g_{up} = 0$, i.e., without so-called gyratons, that describe spinning null fluid.[6,7]

To algebraically classify these spacetimes, the Weyl tensor projected on a null tetrad $\boldsymbol{k}, \boldsymbol{l}, \boldsymbol{m}_i$ can be employed. These (invariant) projections are called Weyl scalars. Spacetime can be classified based on the number of Weyl scalars, sorted by their boost weights, that can be set to zero. Details on the algebraic classification can be found, e.g., in Refs. 8, 9. In Table 1, we list conditions for algebraic types of the Kundt metric.[10]

Table 1. Conditions for algebraic types of the Kundt metric (with $g_{up} = 0$)

| Type | Necessary and sufficient conditions[a] |
|---|---|
| II(ad) | $g_{uu} = -\frac{{}^S R(u,x)}{(D-2)(D-3)}\, r^2 + c(u,x)\, r + d(u,x)$ |
| II(bd) | ${}^S R_{pq} = \frac{1}{D-2}\, g_{pq}\, {}^S R$ |
| II(cd) | ${}^S C_{mpnq} = 0$ |
| III | II(abcd) |
| III(a) | ${}^S R_{,p} = 0 \quad$ and $\quad c_{,p} = \frac{2}{D-3}\, g^{mn} g_{m[n,u||p]}$ |
| III(b) | $g_{p[m,u||q]} = \frac{1}{D-3}\, g^{os} \left( g_{pm} g_{o[s,u||q]} - g_{pq} g_{o[s,u||m]} \right)$ |
| N | III(ab) |
| O | $g^{os} g_{o[s,u||p]||q} + \frac{1}{2(D-2)}\, {}^S R\, g_{pq,u} = \frac{g_{pq}}{D-2}\, g^{mn} \left( g^{os} g_{o[s,u||m]||n} + \frac{1}{2(D-2)}\, {}^S R\, g_{mn,u} \right)$ <br> $d_{||pq} + g_{pq,uu} - \frac{1}{2} g^{os} g_{op,u} g_{sq,u} - \frac{1}{2}\, c\, g_{pq,u}$ <br> $= \frac{g_{pq}}{D-2}\, g^{mn} \left( d_{||mn} + g_{mn,uu} - \frac{1}{2} g^{os} g_{om,u} g_{sn,u} - \frac{1}{2}\, c\, g_{mn,u} \right)$ |

*Note*: [a] The symbol "$||$" denotes spatial covariant derivative ($d_{||pq} = d_{,pq} - \Gamma^m_{pq}\, d_{,m}$).

## 2. Field Equations

We insert the (non-gyratonic) Kundt metric into the QG field equations and observe the conditions they impose upon the metric functions. Firstly, we look at general Kundt metric, but our main focus will be on more special cases. Results for the simpler case of the Gauss–Bonnet gravity can be found in Refs. 10, 11.

### 2.1. *General case*

The $rr$-component $(2\alpha + \beta)\, g_{uu,rrrr} = 0$ implies that either $(2\alpha + \beta) = 0$ or $g_{uu}$ is at most cubic in $r$, so it can be written as $g_{uu} = a\, r^3 + b\, r^2 + c\, r + d$, where $a, b, c, d$ are functions of $u$ and the transverse coordinates $x^p$. These functions will be constrained by the following components of the field equations.

The $rp$-component is $(2\alpha+\beta)\, g_{uu,rrrp}=0$. For $(2\alpha+\beta)\neq 0$, it implies that $a_{,p}=0$, so that $a=a(u)$.

The remaining components $(ru, pq, up, uu)$ get more complicated, and longer, therefor we will not show them here. They can be found for example in Ref. 10. The equations give relations between the functions $a, b, c, d$ and the metric $g_{pq}$.

### 2.2. *pp-waves with constant spatial curvature*

One of the most common examples of the Kundt spacetimes are plane-fronted waves with parallel rays, or *pp*-waves, see e.g., Ref. 12. They are defined as spacetimes admitting a covariantly constant vector field $\boldsymbol{k}=\partial_r$, which implies that the metric is $r$-independent. In the non-gyratonic case, we have

$$\mathrm{d}s^2 = g_{pq}(u,x)\,\mathrm{d}x^p\mathrm{d}x^q - 2\mathrm{d}u\,\mathrm{d}r + g_{uu}(u,x)\,\mathrm{d}u^2\,. \tag{4}$$

For further simplification, we assume that the transverse space is of a constant curvature, which means that its Riemann and Ricci tensors (denoted ${}^{S}R_{pqmn}$ and ${}^{S}R_{pq}$) are related to the Ricci scalar ${}^{S}R$ by

$${}^{S}R_{pqmn} = \frac{{}^{S}R}{(D-3)(D-2)}\,(g_{pm}g_{qn}-g_{pn}g_{qm})\,, \quad {}^{S}R_{pq} = \frac{{}^{S}R}{D-2}\,g_{pq}\,. \tag{5}$$

The $rr$, $rp$ and $up$-components of the field equations are satisfied identically. The $ru$ and $pq$-components give a quadratic equation for the scalar curvature ${}^{S}R$ with two solutions

(a) Trivial solution

$${}^{S}R=0 \Rightarrow g_{pq}=\delta_{pq}, \quad \Lambda=0\,. \tag{6}$$

(b) Non-trivial solution

$${}^{S}R = -\frac{1}{2\kappa}\,\frac{(D-2)(D-3)}{(D-2)(D-3)\alpha+(D-4)(D-5)\gamma+(D-3)\beta} = 4\Lambda_0\,. \tag{7}$$

Since ${}^{S}R$ is constant, necessarily $g_{pq,u}=0$.

For both cases (6), (7), the $uu$-component can be written as

$$k\Delta g_{uu} + \beta\Delta\Delta g_{uu} = 0\,, \tag{8}$$

where the constant $k$ differs for the two solutions mentioned above

$$\text{(a) } k=\kappa^{-1}\,, \quad \text{(b) } k=\frac{(D-3)\beta-2(D-4)\gamma}{\omega\kappa}+\frac{(D-2)(D-3)^2\beta}{2\omega^2\kappa^2}\,,$$
$$\omega \equiv (D-2)(D-3)\alpha+(D-4)(D-5)\gamma+(D-3)\beta\,. \tag{9}$$

This means that we need to solve Helmholtz-like equation for $g_{uu}$ [13]

$$\Delta g_{uu} + \frac{k}{\beta}\,g_{uu} = f\,, \qquad \text{where} \qquad \Delta f = 0\,. \tag{10}$$

The interesting fact is that it is possible to find solution of (10) that does not solve $\Delta g_{uu} = 0$, so such a solution does not exist in Einstein's general relativity.

From Table 1, we see that the non-trivial solution (b) is of type II(bcd), and the non-zero Weyl scalars are

$$\Psi_{2S} = \frac{{}^{S}R}{(D-1)(D-2)}\,, \qquad \Psi_{4^{ij}} = \left[-\frac{1}{2}g_{uu||pq} + \frac{g_{pq}}{2(D-2)}\Delta g_{uu}\right] m_i^p\, m_j^q\,. \quad (11)$$

### 2.3. *VSI spacetimes*

The definition of VSI spacetimes[14] is that all scalar curvature invariants of all orders vanish. Their (non-gyratonic) line element has the form

$$\mathrm{d}s^2 = \delta_{pq}\,\mathrm{d}x^p\mathrm{d}x^q - 2\mathrm{d}u\,\mathrm{d}r + \big(c(u,x)\,r + d(u,x)\big)\mathrm{d}u^2\,, \quad (12)$$

so $g_{uu}$ is at most linear in $r$. The $rr$, $rp$, $ru$, $pq$-components are satisfied identically, the only non-trivial components are $up$ and $uu$. The $up$-component implies $(1/\kappa)c_{,p} + \beta\Delta c_{,p} = 0$, which in general relativity ($\beta = 0$) restricts $c = c(u)$.

The $uu$-component gives two constraints with respect to $r$, namely

$$\Delta\left(\tfrac{1}{\kappa}c + \beta\Delta c\right) = 0 \quad \text{and} \quad \Delta\left[-\tfrac{1}{\kappa}d + \beta\left(2c_{,u} - \tfrac{1}{2}c^2 - \Delta d\right)\right] = 0\,. \quad (13)$$

Looking at Table 1, we observe that VSI spacetimes are of type III(b) with non-zero scalar components of the Weyl tensor

$$\Psi_{3T^i} = -m_i^p\frac{D-3}{2(D-2)}c_{,p}\,,$$
$$\Psi_{4^{ij}} = -\frac{1}{2}m_i^p\, m_j^q\left[r\left(c_{||pq} - \frac{g_{pq}}{D-2}\Delta c\right) + d_{||pq} - \frac{g_{pq}}{D-2}\Delta d\right]. \quad (14)$$

## 3. Geodesic Deviation

Possible way of physical and geometrical interpretation of spacetimes is to investigate local effects of curvature on freely falling test particles.[15,16] Such particles move along geodesics and due to the gravitational field, they can move closer or further away from each other. Let us have a geodesic congruence with central geodesic $\gamma(\tau)$. Its tangent vector $\boldsymbol{u}$ is a four-velocity of the particle moving along geodesic. The vector separating geodesics in the congruence is $\boldsymbol{Z}$. The equation of geodesic deviation then takes the form

$$\frac{\mathrm{D}^2 Z^a}{\mathrm{d}\tau^2} = R^a{}_{bcd}\,u^b\,u^c\,Z^d\,. \quad (15)$$

We choose an orthonormal frame $\boldsymbol{e}_{(a)}$ with $\boldsymbol{e}_{(0)} = \boldsymbol{u}$. The round brackets denote tetrad indices. The projection of Eq. (15) on spatial tetrad vectors becomes

$$\ddot{Z}^{(i)} = R^{(i)}{}_{(0)(0)(j)}Z^{(j)}, \quad i,j = 1,\dots D-1\,. \quad (16)$$

In the Kundt case $\boldsymbol{u} = \dot{r}\,\partial_r + \dot{u}\,\partial_u + \dot{x}^p\,\partial_p$ and the vector $\boldsymbol{e}_{(1)}$ is defined as $\boldsymbol{e}_{(1)} = \sqrt{2}\,\tilde{\boldsymbol{k}} - \boldsymbol{u}$, where $\tilde{\boldsymbol{k}} = 1/(\sqrt{2}\,\dot{u})\,\partial_r$. We call it longitudinal direction. The remaining vectors $\boldsymbol{e}_{(i)}$, $i = 2,\dots D-1$, are called transverse directions.

### 3.1. *Geodesic deviation for pp-waves*

The equations of geodesic deviations for *pp*-waves (with $g_{up} = 0$) are

$$\begin{aligned}\ddot{Z}^{(1)} &= 0\,,\\ \ddot{Z}^{(i)} &= \Big(\tfrac{1}{2}\dot{u}^2\left(g_{uu||pq} + g_{pq,uu} - \tfrac{1}{2}g^{os}g_{op,u}\,g_{sq,u}\right)m_i^p\,m_j^q \\ &\qquad +2\dot{u}\,g_{m[p,u||n]}m^m_{(i}\,m^p_{j)}\,\dot{x}^n - {}^{S}R_{mpnq}m_i^m m_j^n\,\dot{x}^p\dot{x}^q\Big)Z^{(j)}\,.\end{aligned} \tag{17}$$

These were calculated without any use of field equations. Due to the absence of the terms $R^{(1)}_{\ (0)(0)(i)}$, the effect of the waves on particles is only transverse, i.e., in the $\boldsymbol{e}_{(i)}$ directions. Unlike general relativity, the deformations caused by the waves are not always traceless. Since it is possible to choose a frame where $\dot{x}^i = 0$, the conditions for the waves to be traceless become

$$g^{pq}\left(g_{uu||pq} + g_{pq,uu} - \tfrac{1}{2}g^{os}g_{op,u}g_{sq,u}\right) = 0\,. \tag{18}$$

If we now restrict ourselves to the QG and transverse space with constant curvature, the equations simplify into

$$\ddot{Z}^{(1)} = 0, \tag{19}$$

$$\ddot{Z}^{(i)} = \left(\tfrac{1}{2}\dot{u}^2\,g_{uu||pq}m_i^p\,m_j^q + \frac{{}^{S}R}{(D-2)(D-3)}g_{pq}\big(g_{kl}\,\dot{x}^p\,\dot{x}^k\,m_i^q\,m_j^l - \dot{x}^p\,\dot{x}^q\,\delta_{ij}\big)\right)Z^j,$$

and the waves are traceless when $\Delta g_{uu} = 0$, which corresponds to the solution of general relativity. If we solve Eq. (10) for $g_{uu}$, we can insert the resulting metric component into Eq. (19).

### 3.2. *Geodesic deviation for VSI-spacetimes*

For the VSI spacetimes (12), the equations of geodesic deviation are

$$\begin{aligned}\ddot{Z}^{(1)} &= \tfrac{1}{2}\dot{u}\,c_{,p}\,m_j^pZ^{(j)},\\ \ddot{Z}^{(i)} &= \tfrac{1}{2}\dot{u}\,c_{,p}\,m_i^pZ^{(1)} + \left(\tfrac{1}{2}\dot{u}^2\,g_{uu||pq}m_i^p\,m_j^q + m^p_{(i}m^q_{j)}\dot{u}\,\dot{x}^k\,c_{,p}\,g_{qk}\right)Z^{(j)}\,.\end{aligned} \tag{20}$$

We can see that in general, there is a longitudinal effect due to the term $1/2\,\dot{u}\,c_{,p}$. This would be impossible in general relativity, because of $c_{,p} = 0$.

## 4. Conclusion

We studied Kundt spacetime in QG by an explicit analyses of the constraints imposed by the field equations. We investigated the special cases, the *pp*-waves and VSI spacetimes, and obtained solutions that are not allowed in general relativity. The equations of geodesic deviation shown that the effect of *pp*-waves is always transverse, but *generally not traceless* and the VSI spacetimes cause *longitudinal effect*. Both effects are not possible in general relativity.

**Acknowledgments**

OH was supported by the Charles University Grant GAUK 196516 and Mobility Fund FM/c/2018-2-002, and Austrian–Czech projects AÖCZ ICM-2018-12448 awarded by OeAD and financed by BMBWF, and Mobility Grant 8J18AT02. RŠ and JP were supported by the Czech Science Foundation Grant GAČR 17-01625S.

**References**

1. T. Clifton, P. G. Ferreira, A. Padilla and C. Skordis, Modified Gravity and Cosmology *Phys. Rep.* **513**, pp. 1–189 (2012).
2. D. Lovelock, The Einstein tensor and its generalizations, *J. Math. Phys.* **12**, pp. 498–502 (1971).
3. W. Kundt, The plane-fronted gravitational waves, *Z. Physik* **163**, pp. 77–86 (1961).
4. J. Podolský and M. Žofka, General Kundt spacetimes in higher dimensions, *Class. Quantum Grav.* **26**, 105008 (2009).
5. A. Coley, S. Hervik, G. Papadopoulos and N. Pelavas, Kundt spacetimes, *Class. Quantum Grav.* **26**, 105016 (2009).
6. W. B. Bonnor, Spinning null fluid in general relativity, *Int. J. Theor. Phys.* **3**, pp. 257–66 (1970).
7. P. Krtouš, J. Podolský, A. Zelnikov and H. Kadlecová, Higher-dimensional Kundt waves and gyratons, *Phys. Rev. D* **86**, 044039 (2012).
8. M. Ortaggio, V. Pravda and A. Pravdová, Algebraic classification of higher dimensional spacetimes based on null alignment, *Class. Quantum Grav.* **30**, 013001 (2013).
9. J. Podolský and R. Švarc, Explicit algebraic classification of Kundt geometries in any dimension, *Class. Quantum Grav.* **30**, 125007 (2013).
10. M. Karamazov, Exact spacetimes in modified theories of gravity, Master Thesis, Charles University, Faculty of Mathematics and Physics, Prague (2017).
11. O. Hruška and J. Podolský, Kundt Spacetimes in the Gauss–Bonnet Gravity, *WDS'16 Proceedings of Contributed Papers* (Prague: Matfyzpress), 108–15 (2016).
12. J. B. Griffiths and and J. Podolský, *Exact Space-Times in Einstein's General Relativity* (Cambridge University Press, Cambridge, 2009).
13. M. S. Madsen, The plane gravitational wave in quadratic gravity, *Class. Quantum Grav.* **7**, pp. 87–96 (1990).
14. V. Pravda, A. Pravdová, A. Coley and R. Milson, All spacetimes with vanishing curvature invariants, *Class. Quantum Grav.* **19**, pp. 6213–36 (2002).
15. J. Podolský and R. Švarc, Interpreting spacetimes of any dimension using geodesic deviation, *Phys. Rev. D* **85**, 044057 (2012).
16. J. Podolský and R. Švarc, Physical interpretation of Kundt spacetimes using geodesic deviation, *Class. Quantum Grav.* **30**, 205016 (2013).

# The double field theory algebroid as a projection of a large Courant algebroid

Athanasios Chatzistavrakidis, Larisa Jonke and Fech Scen Khoo

*Division of Theoretical Physics, Rudjer Bošković Institute, Bijenička 54, 10000 Zagreb, Croatia*
*E-mails: athanasios.chatzistavrakidis@irb.hr, larisa@irb.hr, Fech.Scen.Khoo@irb.hr*

Richard J. Szabo

*Department of Mathematics, Heriot-Watt University, Colin Maclaurin Building, Riccarton, Edinburgh EH14 4AS, U.K.*
*Maxwell Institute for Mathematical Sciences, Edinburgh, U.K.*
*The Higgs Centre for Theoretical Physics, Edinburgh, U.K.*
*E-mail: R.J.Szabo@hw.ac.uk*

Double field theory promotes the T-duality of closed string theory to a manifest symmetry, thus leading to a new perspective on the geometry experienced by stringy probes. In this contribution, we discuss the mathematical structure underlying the symmetries of double field theory, thus defining a DFT algebroid. We trace its origins in a large Courant algebroid defined over a doubled geometry, and show that after imposing a section condition the DFT algebroid reduces to a canonical Courant algebroid, as expected in generalized geometry.

*Keywords*: Differential Geometry; String Duality; Courant Algebroid; Supergravity.

## 1. Motivation and goal

Finding the ultimate theory that would describe quantum gravitational phenomena remains an open problem today. Nevertheless, theoretical arguments suggest that the geometry of quantum gravity should depart from Riemannian geometry.

One possible direction which goes beyond Riemannian geometry is the generalized geometry proposed by Ref. 1, see also Ref. 2 and Ref. 3. In generalized geometry one considers extensions of the tangent bundle over a smooth spacetime manifold; in the simplest case, the tangent bundle is extended by the cotangent bundle. From a physical perspective, we are familiar with the physics on the tangent bundle $TM$ of a manifold $M$, whose sections are vector fields which are closed under the Lie bracket and generate diffeomorphisms. In generalized geometry, there is a symmetry group $O(d,d)$, where each $d$ represents separately the rank of the tangent and cotangent bundles, and it contains not only diffeomorphisms but also 2-form gauge transformations. This highlights the relation to closed string theory, whose massless sector indeed includes the graviton and the 2-form Kalb-Ramond $B$-field.

On the other hand, double field theory (DFT) incorporates T-duality (or rather $O(d,d;\mathbb{R})$) as a manifest symmetry of a field theory which is based on a doubled space (Refs. 4, 5, 6, 7, 8). T-duality is a symmetry in string theory that relates the momentum modes of strings propagating in a compact space to their winding modes in a dual compact space. Promoting T-duality to a symmetry in field theory

comes with the cost of imposing a section condition, or strong constraint, that effectively reduces the doubled geometry to a half-slice. For instance, the standard supergravity frame in DFT is obtained by dropping all dependence on the dual winding coordinates of the doubled geometry, thus resulting in the standard low-energy action of the NS-NS (Neveu-Schwarz) sector of string theory in $d$ dimensions.

Understanding better the structure and the geometry of string duality is motivated at a conceptual level through the connections to the geometry of quantum gravity that we already mentioned, and also with regard to the implications of string theory for cosmology, where the winding modes included in DFT could play a crucial role. This is the case in the cosmological scenario of Ref. 9, which is recently explored from the point of view of DFT too, see e.g. Ref. 10.

In this contribution, our goal is to explain the mathematical structure which underlies the symmetries of double field theory. We shall show that this is an algebroid that lies in a certain precise way in between two Courant algebroids over different base manifolds. This contribution is based on Ref. 11.

## 2. Courant algebroid

A Courant algebroid is a quadruple $(E, [\cdot,\cdot]_E, \langle\cdot,\cdot\rangle, \rho)$ comprising the following, see Ref. 12. $E \to M$ is a vector bundle over a smooth manifold $M$, equipped with a binary operation $[\cdot,\cdot]_E : \Gamma(E) \otimes \Gamma(E) \to \Gamma(E)$ on its sections, and a symmetric $C^\infty(M)$-bilinear, non-degenerate form $\langle\cdot,\cdot\rangle : \Gamma(E) \otimes \Gamma(E) \to C^\infty(M)$ which takes smooth sections of the bundle and returns a function. Finally, a bundle map $\rho : E \to TM$, called the anchor, maps from $E$ to the tangent bundle $TM$.

What turns this quadruple to a Courant algebroid is a set of five properties. The first one is a Jacobi identity

$$[[A,B]_E,C]_E + cyclic = \tfrac{1}{3}\,\mathcal{D}\langle[A,B]_E,C\rangle + cyclic\ , \tag{1}$$

which tells us that for any smooth sections $A, B, C \in \Gamma(E)$, the Jacobiator of the binary operation is exact with respect to the differential operator $\mathcal{D} : C^\infty(M) \to \Gamma(E)$ which is defined by $\langle\mathcal{D}f, A\rangle = \frac{1}{2}\,\rho(A)f$, for functions $f \in C^\infty(M)$. The second property is a Leibniz rule

$$[A, fB]_E = f[A,B]_E + (\rho(A)f)B - \langle A,B\rangle\mathcal{D}f\ . \tag{2}$$

Moreover, the anchor is a homomorphism of bundles, $\rho[A,B]_E = [\rho(A),\rho(B)]$, and the image of the derivation $\mathcal{D}$ is in the kernel of the anchor, $\rho\circ\mathcal{D} = 0$, or equivalently $\langle\mathcal{D}f, \mathcal{D}g\rangle = 0$. Lastly, a compatibility condition for the bracket and bilinear form, the analog of ad-invariance of the Killing form of a Lie algebra, reads as

$$\rho(C)\langle A,B\rangle = \langle[C,A]_E + \mathcal{D}\langle C,A\rangle, B\rangle + \langle A, [C,B]_E + \mathcal{D}\langle C,B\rangle\rangle\ . \tag{3}$$

When we consider the generalized tangent bundle $E = TM \oplus T^*M$, its sections are generalized vectors $V + \lambda$, formal sums of a vector $V \in \Gamma(TM)$ and 1-form

$\lambda \in \Gamma(T^*M)$. The standard Courant bracket,

$$[V+\lambda, V'+\lambda']_E = [V,V'] + \mathcal{L}_V\lambda' - \mathcal{L}_{V'}\lambda - \frac{1}{2}d(i_V\lambda' - i_{V'}\lambda) \ , \tag{4}$$

is an antisymmetric operation, an extension of the Lie bracket of vector fields which is not a Lie bracket itself, that satisfies the properties of a Courant algebroid when $\rho = (\mathrm{id}, 0)$. The inner product, which is the fiber-wise metric on $E$, is

$$\langle V+\lambda, V'+\lambda'\rangle = \tfrac{1}{2}(\lambda'(V)+\lambda(V')) = \tfrac{1}{2}\begin{pmatrix} V & \lambda \end{pmatrix}\begin{pmatrix} 0 & \mathbf{1}^d \\ \mathbf{1}^d & 0 \end{pmatrix}\begin{pmatrix} V' \\ \lambda' \end{pmatrix} \ . \tag{5}$$

The bilinear form explicitly exhibits an indefinite $(d,d)$ signature. The $O(d,d)$ transformations are respectively the diffeomorphisms, $B$-transforms and $\beta$-transforms,

$$\begin{pmatrix} N & 0 \\ 0 & N^{-T} \end{pmatrix} \ , \ e^B = \begin{pmatrix} 1 & 0 \\ -B & 1 \end{pmatrix} \ , \ e^\beta = \begin{pmatrix} 1 & -\beta \\ 0 & 1 \end{pmatrix} \ , \tag{6}$$

where $N \in GL(d)$, $B$ is a 2-form and $\beta$ is a 2-vector. The inner product is indeed invariant under the $B$-transform,

$$\langle e^B(V+\lambda), e^B(V'+\lambda')\rangle = \langle V+\lambda, V'+\lambda'\rangle \ . \tag{7}$$

More details can be found in Ref. 3.

## 3. Double field theory

In this section we give a short introduction to double field theory. It is a theory that doubles the space of coordinates: $(x^i, \tilde{x}_i)$ where $i$ is $d$-dimensional, and it contains both coordinates conjugate to the momentum modes and the dual coordinates conjugate to the winding modes of closed strings. Interested readers may refer to Ref. 13 and references therein.

The double field theory action is formulated in doubled geometry (Ref. 14),

$$S_{DFT}(\mathcal{H}_{MN}, \Phi) = \int d^dx \ d^d\tilde{x} \ e^{-2\Phi}\mathcal{R} \ , \tag{8}$$

where its Ricci scalar is given by

$$\begin{aligned}\mathcal{R} = {} & \frac{1}{8}\mathcal{H}^{MN}\partial_M\mathcal{H}^{KL}(\partial_N\mathcal{H}_{KL} - 4\partial_K\mathcal{H}_{NL}) + 4(\partial_N\Phi\,\partial_M\mathcal{H}^{MN} - \mathcal{H}^{MN}\partial_M\Phi\,\partial_N\Phi) \\ & +4\mathcal{H}^{MN}\partial_M\partial_N\Phi - \partial_M\partial_N\mathcal{H}^{MN} \ ,\end{aligned} \tag{9}$$

expressed in terms of the generalized metric $\mathcal{H}_{MN}$,

$$\mathcal{H}_{MN} = \begin{pmatrix} g^{ij} & -g^{ik}B_{kj} \\ B_{ik}g^{kj} & g_{ij} - B_{ik}g^{kl}B_{lj} \end{pmatrix} \tag{10}$$

and a scalar field $\Phi$. The generalized metric combines the 2-form $B$ and the Riemannian metric $g$. We can raise or lower its indices as $\mathcal{H}^{MN} = \eta^{MP}\eta^{NQ}\mathcal{H}_{PQ}$, using the $O(d,d)$ metric $\eta = \eta^{-1} = \begin{pmatrix} 0 & \mathbf{1}^d \\ \mathbf{1}^d & 0 \end{pmatrix}$. The scalar field $\Phi$ in the action is

related to the dilaton $\phi$ by $e^{-2\Phi} = \sqrt{|g|}e^{-2\phi}$. Note that the capital indices are $2d$-dimensional, $M, N, K, L = 1, \ldots, 2d$, while $i, j, k, l = 1, \ldots, d$.

The DFT action $S_{DFT}$ reduces to the Neveu-Schwarz sector of supergravity action in $d$ dimensions,

$$S_{NS} = \int d^d x \sqrt{g} e^{-2\phi} \left( R + 4\partial_i \phi \, \partial^i \phi - \frac{1}{12} H_{ijk} H^{ijk} \right) , \tag{11}$$

where $H_{ijk}$ is the field strength of the 2-form $B$, after imposing the strong constraint, that is $\widetilde{\partial}^i = \frac{\partial}{\partial \tilde{x}_i} = 0$ (supergravity frame). The strong constraint is inherent in DFT. It is related to the closure of the gauge transformations. DFT is therefore an $O(d,d)$-covariant formulation of the low-energy sector of string theory on a compact space. See Ref. 6 for further details.

## 4. An algebroid for double field theory

Now we present our proposal of an algebroid structure which is underlying DFT. The strategy we follow is to double the canonical Courant algebroid, decompose the bundle and subsequently make a suitable projection.

Consider a $2d$-dimensional target space $T^*M$ with local coordinates

$$\mathbb{X} = (\mathbb{X}^I) = (\mathbb{X}^i, \mathbb{X}_i) =: (X^i, \widetilde{X}_i) \ , \tag{12}$$

where $I = 1, \ldots, 2d$ and $i = 1, \ldots, d$. The generalized tangent bundle is promoted from $TM \oplus T^*M$ with $O(d,d)$ symmetry to

$$\mathbb{E} = \mathbb{T}(T^*M) := T(T^*M) \oplus T^*(T^*M) \tag{13}$$

with $O(2d, 2d)$ symmetry. The sections of the bundle are $\mathbb{A}^{\hat{I}}$, $\hat{I} = 1, \ldots, 4d$, where $(\mathbb{A}^{\hat{I}}) = (\mathbb{A}^I, \widetilde{\mathbb{A}}_I) = (\mathbb{A}^i, \mathbb{A}_i, \widetilde{\mathbb{A}}_i, \widetilde{\mathbb{A}}^i)$. Here the generalized vector over the doubled target space is similarly a sum of vector and 1-form, $\mathbb{A} = \mathbb{A}_V + \mathbb{A}_F := \mathbb{A}^I \, \partial_I + \widetilde{\mathbb{A}}_I \, \mathrm{d}\mathbb{X}^I$, where $(\partial_I) = (\partial/\partial X^i, \partial/\partial \widetilde{X}_i) =: (\partial_i, \tilde{\partial}^i)$ and $(\mathrm{d}\mathbb{X}^I) := (\mathrm{d}X^i, \mathrm{d}\widetilde{X}_i)$. As for the anchor $\rho^I{}_{\hat{J}}$ of the bundle $\mathbb{E}$, its components are $(\rho^I{}_J, \widetilde{\rho}^{IJ})$. With that, we can define a 'large' Courant algebroid: $(\mathbb{E}, [\,\cdot\,,\,\cdot\,]_{\mathbb{E}}, \langle\,\cdot\,,\,\cdot\,\rangle, \rho)$.

Before we proceed, note that the sections of the bundle in the large Courant algebroid carry a $4d$-dimensional index. In order to relate to the DFT data which are $2d$-dimensional, we need to resolve this excess of $2d$. First we introduce a decomposition in the sections of the bundle and their corresponding basis,

$$\mathbb{A}^I_\pm = \tfrac{1}{2} \left( \mathbb{A}^I \pm \eta^{IJ} \, \widetilde{\mathbb{A}}_J \right) \ , \quad e^\pm_I = \partial_I \pm \eta_{IJ} \, \mathrm{d}\mathbb{X}^J \ . \tag{14}$$

For the anchor components $\rho^I{}_{\hat{J}}$, we define $(\rho_\pm)^I{}_J = \rho^I{}_J \pm \eta_{JK} \, \widetilde{\rho}^{IK}$. Consequently, the large generalized tangent bundle is decomposed as $\mathbb{E} = L_+ \oplus L_-$, where $L_\pm$ is the bundle with the space of sections spanned by $e^\pm_I$ locally.

The DFT data are obtained from the large Courant algebroid upon a suitable projection to $L_+$. This projection,

$$\mathsf{p}_+ : \mathbb{E} \longrightarrow L_+ \ , \quad (\mathbb{A}_V, \mathbb{A}_F) \longmapsto \mathbb{A}_+ := A \ , \tag{15}$$

reduces a generalized vector of the large Courant algebroid $\mathbb{E}$ to a DFT vector, which may be written as $A = A_i \left(\mathrm{d}X^i + \tilde{\partial}^i\right) + A^i \left(\mathrm{d}\widetilde{X}_i + \partial_i\right)$.

For elements of $L_+$ to be treated as elements of $\mathbb{E}$, we define an inclusion map $\iota : L_+ \hookrightarrow \mathbb{E}$ and the composition map $\mathsf{a} := \iota \circ \mathsf{p}_+ : \mathbb{E} \longrightarrow \mathbb{E}$. By applying the map $\mathsf{a}$ and a projection to the standard Courant bracket of the large Courant algebroid, we obtain precisely the C-bracket of DFT vectors, which is the operation related to generalized diffeomorphisms of DFT,

$$\mathsf{p}_+\big([\mathsf{a}(\mathbb{A}), \mathsf{a}(\mathbb{B})]_{\mathbb{E}}\big) = (A^K\, \partial_K B^J - \tfrac{1}{2}\, A^K\, \partial^J B_K - (A \leftrightarrow B)) e_J^+ := [\![A, B]\!]_{L_+}. \quad (16)$$

Similarly, projecting the generalized Lie derivative[a] of the large Courant algebroid yields the generalized Lie derivative in DFT,

$$\mathsf{p}_+\big(\mathbb{L}_{\mathsf{a}(\mathbb{A})}\mathsf{a}(\mathbb{B})\big) = (A^I\, \partial_I B^J - B^I\, \partial_I A^J + B_I\, \partial^J A^I) e_J^+ := \mathsf{L}_A B\ . \quad (17)$$

Note that $[\mathsf{L}_C, \mathsf{L}_A] = \mathsf{L}_{[\![C,A]\!]_{L_+}}$ only if we impose the condition $\eta^{IJ}\, \partial_I f\, \partial_J g = 0$, for all fields $f, g$ of DFT. This condition is the strong constraint or section condition.

The next step in our strategy is to examine which of the properties of a Courant algebroid survive the projection and which get obstructed. The result allows us to give the following definition.

**Definition.** Let $L_+ \to T^*M$ be a vector bundle of rank $2d$ over $T^*M$. Let $[\![\cdot\,,\cdot]\!]_{L_+} : \Gamma(L_+) \otimes \Gamma(L_+) \to \Gamma(L_+)$ be a skew-symmetric bracket (C-bracket). Let $\langle\cdot\,,\cdot\rangle_{L_+} : \Gamma(L_+) \otimes \Gamma(L_+) \to C^\infty(T^*M)$ be a non-degenerate symmetric form. Lastly, a smooth bundle map $\rho_+ : L_+ \to T(T^*M)$. We define a DFT algebroid: $(L_+, [\![\cdot\,,\cdot]\!]_{L_+}, \langle\cdot\,,\cdot\rangle_{L_+}, \rho_+)$, such that it satisfies three properties, one related to the strong constraint, one being the compatibility condition between the C-bracket and the bilinear form, and a Leibniz rule for the C-bracket respectively,

$$\langle \mathcal{D}_+ f, \mathcal{D}_+ g\rangle_{L_+} = \frac{1}{4}\, \langle \mathrm{d}f, \mathrm{d}g\rangle_{L_+}\ , \quad (18)$$

$$\begin{aligned}\rho_+(C)\langle A, B\rangle_{L_+} &= \langle [\![C, A]\!]_{L_+} + \mathcal{D}_+\langle C, A\rangle_{L_+}, B\rangle_{L_+} \\ &\quad + \langle A, [\![C, B]\!]_{L_+} + \mathcal{D}_+\langle C, B\rangle_{L_+}\rangle_{L_+}\ , \end{aligned} \quad (19)$$

$$[\![A, f\, B]\!]_{L_+} = f\, [\![A, B]\!]_{L_+} + \big(\rho_+(A) f\big)\, B - \langle A, B\rangle_{L_+}\, \mathcal{D}_+ f\ , \quad (20)$$

for sections of the bundle $A, B, C \in \Gamma(L_+)$ and functions $f, g \in C^\infty(T^*M)$. The derivative $\mathcal{D}_+ : C^\infty(T^*M) \to \Gamma(L_+)$ is defined through $\langle \mathcal{D}_+ f, A\rangle_{L_+} = \frac{1}{2}\, \rho_+(A) f$.

More generally, a DFT algebroid is a special case of a structure called pre-DFT algebroid, which is related to the Vaisman algebroid of Ref. 15. An interesting formulation of such structures in the language of graded geometry is found in Refs. 16 and 17.

[a] Antisymmetrization of the generalized Lie derivatives gives the Courant bracket: $[\mathbb{A}, \mathbb{B}]_{\mathbb{E}} = \mathbb{L}_{\mathbb{A}}\mathbb{B} - \mathbb{L}_{\mathbb{B}}\mathbb{A}$.

## 5. Conclusions

Using a method of doubling-splitting-projecting, the geometric origin of DFT can be traced to a large Courant algebroid. To summarize our finding schematically,

$$\text{Large Courant algebroid} \xrightarrow{\mathsf{p}_+} \text{DFT algebroid} \overset{\langle \mathcal{D}_+ f, \mathcal{D}_+ g \rangle_{L_+}=0}{\longrightarrow} \text{Courant algebroid}.$$

Additionally in Ref. 11, we propose a DFT membrane sigma-model which describes both geometric and non-geometric fluxes of string theory, thus completing earlier proposals in Refs. 18 and 19.

## Acknowledgments

We acknowledge support by COST (European Cooperation in Science and Technology) in the framework of the Action MP1405 QSPACE. The work of A.Ch., L.J. and F.S.K. was supported by the Croatian Science Foundation under the Project IP-2014-09-3258, and the European Union through the European Regional Development Fund - the Competitiveness and Cohesion Operational Programme (KK.01.1.1.06). The work of R.J.S. was supported by the Consolidated Grant ST/P000363/1 from the U.K. Science and Technology Facilities Council.

## References

1. N. Hitchin, *Quart. J. Math.* **54**, 281 (2003).
2. D. Roytenberg, *Ph.D. Thesis*, arXiv:math/9910078 [math.DG], 1999.
3. M. Gualtieri, *Ph.D. Thesis*, arXiv:math/0401221 [math.DG], 2003.
4. W. Siegel, *Phys. Rev. D* **47**, 5453 (1993).
5. W. Siegel, *Phys. Rev. D* **48**, 2826 (1993).
6. C. Hull and B. Zwiebach, *JHEP* **09**, 099 (2009).
7. C. Hull and B. Zwiebach, *JHEP* **09**, 090 (2009).
8. O. Hohm, C. Hull and B. Zwiebach, *JHEP* **07**, 016 (2010).
9. R. Brandenberger and C. Vafa, *Nucl. Phys. B* **316**, 391 (1989).
10. R. Brandenberger, R. Costa, G. Franzmann and A. Weltman, *Phys. Rev. D* **99**, 023531 (2019).
11. A. Chatzistavrakidis, L. Jonke, F. S. Khoo and R. J. Szabo, *JHEP* **1807**, 015 (2018).
12. Z.-J. Liu, A. Weinstein and P. Xu, *J. Diff. Geom.* **45**, 547 (1997).
13. G. Aldazabal, D. Marqués and C. Núñez, *Class. Quant. Grav.* **30**, 163001 (2013).
14. O. Hohm, C. Hull and B. Zwiebach, *JHEP* **08**, 008 (2010).
15. I. Vaisman, *J. Math. Phys.* **53**, 033509 (2012).
16. A. Deser and J. Stasheff, *Commun. Math. Phys.* **339**, 1003 (2015).
17. A. J. Bruce and J. Grabowski, arXiv:1608.01585 [math-ph], 2016.
18. D. Mylonas, P. Schupp and R. J. Szabo, *JHEP* **09**, 012 (2012).
19. A. Chatzistavrakidis, L. Jonke and O. Lechtenfeld, *JHEP* **11**, 182 (2015).

# Clarifying spatial distance measurement

Colin MacLaurin*

*School of Mathematics and Physics*
*University of Queensland, Australia*
**E-mail: colin.maclaurin@uqconnect.edu.au*

We examine length measurement in curved spacetime, based on the 1+3-splitting of a local observer frame. This situates extended objects within spacetime, in terms of a given coordinate which serves as an external reference. The radar metric is shown to coincide with the spatial projector, but these only give meaningful results on the observer's 3-space, where they reduce to the metric. Examples from Schwarzschild spacetime are given.

## 1. Introduction and motivation

Recall the textbook "radial proper distance" in Schwarzschild spacetime:

$$ds = \left(1 - \frac{2M}{r}\right)^{-1/2} dr, \tag{1}$$

which follows from setting $dt = d\theta = d\phi = 0$ in the line element in Schwarzschild coordinates. But what is the physical motivation for choosing a slice Schwarzschild $t = \text{const}$, rather than some other time coordinate? Special relativity stresses length is relative to the observer, so which observers measure Equation 1, and what do others measure? The claim the $dt = 0$ slice is "measurement by an observer at infinity" is problematic, because there are many ways to extend a local frame, or to choose a simultaneity convention between distant frames.

Consider the same procedure repeated for Gullstrand-Painlevé coordinates:

$$ds = dr. \tag{2}$$

Painlevé hence concluded relativity is self-contradictory.[1] Instead, as we shall see, these correspond to measurements by different observers. (Mathematically, note the differing expressions called "$ds$" are really restrictions to different subspaces.) We apply four complementary theoretical tools: suitably chosen coordinates, the spatial projector, the radar metric, and adapted frames.

## 2. Well-suited coordinates

An intuitive and pedagogical approach to length measurement is to provide coordinates suited to a given congruence of observers, if possible. Consider for example radial geodesic motion in Schwarzschild spacetime, with 4-velocity field $\mathbf{u}$ parametrised by the Killing energy per mass $e := -\mathbf{g}(\mathbf{u}, \partial_t)$. A generalisation of Gullstrand-Painlevé coordinates has metric:[2–6]

$$ds^2 = -\frac{1}{e^2}\left(1 - \frac{2M}{r}\right)dT^2 \mp \frac{2}{e^2}\sqrt{e^2 - 1 + \frac{2M}{r}}\,dT\,dr + \frac{1}{e^2}dr^2 + r^2(d\theta^2 + \sin^2\theta\, d\phi^2), \tag{3}$$

where $T \equiv T_e$ is the Einstein-synchronised proper time of the observers, which all share the same $e \in \mathbb{R}\backslash\{0\}$. For details see Ref. 6. Setting $dT = d\theta = d\phi = 0$ gives:

$$dL = \frac{1}{e} dr, \tag{4}$$

where we write $dL$ in place of $ds$ for the length element, and the sign choice is mere convention. The physical justification behind $dT = 0$ is that the $T = \text{const}$ hypersurface is orthogonal to the 4-velocities, so coincides with their local 3-spaces.

Note a static observer at $r = r_0$ is identical to an observer falling from rest at $r_0$, in the sense their velocities and hence local 3-spaces coincide instantaneously. Both have $e = \sqrt{1 - 2M/r_0}$, so Equation 4 reduces to the usual quantity (Equation 1). Some textbooks set up a false dichotomy that $dr$ is not the distance but $(1 - 2M/r)^{-1/2} dr$ *is*. Instead, for radial observers with $e = \pm 1$ the $r$-coordinate is precisely proper distance.

Equation 4 is remarkably little known. Gautreau & Hoffmann derived it, using a different parameter corresponding to the $0 < e < 1$ case.[2] Taylor & Wheeler present the $e = 1$ case clearly, which is the only textbook coverage apparently.[7] Finch showed the 3-volume inside the horizon is $1/e$ times its Euclidean value $\frac{4}{3}\pi(2M)^3$, for the $e > 0$ case. (Precedents include Lemaître, who pointed out 3-space is Euclidean for $e = 1$, and mentioned measurement.[8] Painlevé made the same observations, but mistakenly saw contradiction.[1])

In general, consider a 4-velocity field $\mathbf{u}$. Define a new coordinate $T$ by:

$$dT := -N^{-1} \mathbf{u}^\flat, \tag{5}$$

where $\mathbf{u}^\flat$ is the 1-form dual to the 4-velocity, and $N$ is a lapse. $T$ exists locally iff the velocity gradient is vorticity-free, a consequence of Frobenius' theorem. Then the $T = \text{const}$ hypersurfaces are orthogonal to the congruence, since $dT(\boldsymbol{\xi}) = 0$ for any vector $\boldsymbol{\xi}$ orthogonal to $\mathbf{u}$.[a] Now express the metric in coordinates including $T$, and set $dT = 0$.

## 3. Spatial projector

Given a 4-velocity $\mathbf{u}$, the metric splits into parts parallel and orthogonal to $\mathbf{u}$ as $g_{\mu\nu} \equiv -u_\mu u_\nu + (g_{\mu\nu} + u_\mu u_\nu)$, assuming metric signature $-+++$. The latter term is the *spatial projection tensor* $\mathbf{P}$, which extracts the spatial part of tensors via contraction. In particular, $P^\mu_\nu u^\nu = 0$, and $P^\mu_\nu \xi^\nu = \xi^\nu$ for any vector $\boldsymbol{\xi}$ in the 3-space orthogonal to $\mathbf{u}$. Furthermore $P_{\mu\nu}\xi^\mu\xi^\nu = g_{\mu\nu}\xi^\mu\xi^\nu$ for such a $\boldsymbol{\xi}$, so $\mathbf{P}$ is also called the *spatial metric*.

One may wonder if $dL'^2 := P_{\mu\nu}\xi^\mu\xi^\nu$ is meaningful as a length measurement for *any* $\boldsymbol{\xi}$, not necessarily orthogonal to $\mathbf{u}$. For radial motion in Schwarzschild

[a] If in addition the congruence is geodesic, we can set $N \equiv 1$, then $dT/d\tau = -\mathbf{u}^\flat(\mathbf{u}) = 1$, so $T$ measures proper time. This trick to derive $T$ was applied to relativity by Synge, and Lagrange's 3-velocity potential in Newtonian mechanics is an antecedent.

spacetime, the projector in Schwarzschild coordinates is

$$P_{\mu\nu} = \begin{pmatrix} e^2 - 1 + \frac{2M}{r} & e\big(1 - \frac{2M}{r}\big)^{-1}\sqrt{e^2 - 1 + \frac{2M}{r}} \\ e\big(1 - \frac{2M}{r}\big)^{-1}\sqrt{e^2 - 1 + \frac{2M}{r}} & \boxed{e^2\big(1 - \frac{2M}{r}\big)^{-2}} \end{pmatrix} \tag{6}$$

in the $t$-$r$ block, plus the usual 2-sphere metric components $\mathrm{diag}(r^2, r^2 \sin^2\theta)$ in the $\theta$-$\phi$ block. One might expect the radial direction to be the coordinate basis vector $\partial_r$, at least for $r > 2M$. Contracting with $\boldsymbol{\xi} = \partial_r$ picks out the $P_{rr}$ component (boxed), for a candidate spatial measurement:

$$dL = e\Big(1 - \frac{2M}{r}\Big)^{-1} dr. \tag{7}$$

But consider the "same" contraction of tensors expressed in the generalised Gullstrand-Painlevé coordinates. The projector is:

$$P_{\mu\nu} = \begin{pmatrix} \frac{1}{e^2}(e^2 - 1 + \frac{2M}{r}) & \frac{1}{e^2}\sqrt{e^2 - 1 + \frac{2M}{r}} \\ \frac{1}{e^2}\sqrt{e^2 - 1 + \frac{2M}{r}} & \boxed{\frac{1}{e^2}} \end{pmatrix}, \tag{8}$$

in the $T$-$r$ block, so the contraction with $\partial_r$ yields $dL = \frac{1}{e} dr$ as in Equation 4. But why the discrepancy with Equation 7? While Equations 6 and 8 represent the same tensor, it turns out the coordinate vectors $\partial_r$ are distinct. By definition of coordinate basis, ${}^{(Schw)}\partial_r$ is orthogonal to $dt$, whereas ${}^{(GP)}\partial_r$ is orthogonal to $dT$. (In contrast, the vector $(dr)^\sharp$ depends only on $r$.) This potential confusion about coordinate vectors is rarely discussed explicitly.[4,6] In the present context, it shows a potential pitfall for measurement, and the superficial contradiction motivates deeper study.

In fact Equation 7 does have physical meaning: it is the measurement of a falling ruler as determined in the local static frame. By this, we mean the comparison of the ruler's length-contracted tick marks with the $r$-coordinate. The two results are related by the Lorentz factor $\gamma = |e|(1 - 2M/r)^{-1/2}$, since the frames are in standard configuration. This will be examined in future work. For now we conclude $\sqrt{\mathbf{P_u}(\boldsymbol{\xi}, \boldsymbol{\xi})}$ is *not* a measurement in $\mathbf{u}$'s frame, if $\mathbf{g}(\boldsymbol{\xi}, \mathbf{u}) \neq 0$.

## 4. Radar metric

The sonar / radar method of distance measurement involves bouncing a signal off a distant object, and timing the return journey. In relativity a null signal is used, along with the proper time $\Delta\tau$ of the emitter, hence the one-way distance is $\Delta L := \Delta\tau/2$ (assuming an isotropic speed of light $c = 1$). While radar was promoted by Poincaré, Einstein, Milne, Bondi, and others, the following formula was derived by Landau & Lifshitz:[9]

$$\gamma_{ij} := g_{ij} - \frac{g_{0i} g_{0j}}{g_{00}}, \tag{9}$$

for $i, j = 1, 2, 3$, with infinitesimal length element $dL^2 = \gamma_{ij} dx^i dx^j$. This assumes a particular coordinate system is provided, and that the radar instrument is *comoving* in those coordinates.

To apply this to radial motion in Schwarzschild, we need comoving coordinates. One choice is a case of Lemaître-Tolman-Bondi coordinates, using $\rho_e$ given in Refs. 2 and 6, together with $T_e$ from above. The radar metric is:

$$\gamma_{ij} = \begin{pmatrix} \frac{1}{e^2}\left(e^2 - 1 + \frac{2M}{r}\right) & 0 & 0 \\ 0 & r^2 & 0 \\ 0 & 0 & r^2 \sin^2\theta \end{pmatrix}, \tag{10}$$

with radial distance $dL = \frac{1}{e}\sqrt{e^2 - 1 + 2M/r}\, d\rho$. The interval "$d\rho$" is unfamiliar, hence reinterpret the radar metric as 4-dimensional (which simply adds terms of 0), and transform into other coordinates. It turns out Equation 10 is identical to the spatial projector (Equations 6 and 8), and the coordinate vector $\partial_\rho$ is parallel to both ${}^{(GP)}\partial_r$ and the radial ruler seen later (Equation 11).[b]

In fact the spatial projector is the covariant generalisation of the radar metric. In any comoving coordinates, the observer has 4-velocity $u^\mu = ((-g_{00})^{-1/2}, 0, 0, 0)$, assuming $\partial_0$ is future-pointing. Then $P_{\mu\nu}$ reduces to $\gamma_{\mu\nu}$, taking the latter as 4-dimensional. Landau & Lifshitz interpret 3-space as spanned by the coordinate vectors $\partial_i$. However these are not necessarily orthogonal to $\partial_0$ and the 4-velocity. Equation 9 gives incorrect results for directions not orthogonal to $\mathbf{u}$, as discussed for the spatial projector. When restricted appropriately, both radar and the projector are simply $g_{\mu\nu}$.

## 5. Adapted frames

Our final technical tool is local reference frames. The observer 4-velocity splits the local tangent space into "time" $\mathbf{u}$, and 3-dimensional "space" orthogonal to $\mathbf{u}$. For our purposes a single spatial vector $\boldsymbol{\xi}$ is often sufficient. While it is well known that measurements relate to expressing tensors in an observer's frame, strangely this is not applied to spatial distance — with rare exceptions.[10] Likewise Rindler writes, "rigid scales [rulers] are of ill repute in relativity", but there is nothing wrong with " 'resilient' scales", within limits on acceleration and tidal forces.[11]

For radial motion in Schwarzschild, the obvious choice of radial vector is that orthogonal to $\mathbf{u}$, $d\theta$ and $d\phi$:

$$\xi^\mu := \Big( -\left(1 - \frac{2M}{r}\right)^{-1} \sqrt{e^2 - 1 + \frac{2M}{r}}, e, 0, 0 \Big). \tag{11}$$

Since the $r$-component is $e$ and the vector has unit length, this means a coordinate interval $\Delta r = e$ corresponds to a proper length $\Delta L = 1$. Hence $dL = \frac{1}{e} dr$ as before.[c]

[b] Taking instead Schwarzschild $t$ with $\rho$, the radar metric gives Equation 7 upon contraction with the alternative $\partial'_\rho$. As before, coordinate vector directions can be misinterpreted.

[c] Our "rulers" are vectors in a single tangent space, which approximate short rulers over spacetime.

In general, suppose a unit spatial vector $\boldsymbol{\xi}$, and a scalar $\Phi$ (for instance a coordinate) are provided. The change in $\Phi$ over the extent of the ruler is $\Delta\Phi = d\Phi(\boldsymbol{\xi})$, hence:

$$dL = \frac{1}{d\Phi(\boldsymbol{\xi})} d\Phi = \frac{1}{\xi^{\Phi}} d\Phi, \tag{12}$$

where the latter expression applies if $\Phi$ is taken as a coordinate. We require $d\Phi(\boldsymbol{\xi}) \neq 0$, meaning $\Phi$ is not constant along the ruler direction, so can demarcate length. In any spacetime, given a coordinate expression of a tetrad, one can simply read off a coordinate length interval by inverting the relevant component.

The "best" ruler direction to choose is typically not obvious. Given $\Phi$ and $\mathbf{u}$, one candidate is a certain maximal direction as follows. Recall the gradient vector $(d\Phi)^\sharp$ is the direction of steepest increase of $\Phi$ per unit length. However this is generally not *purely* spatial, according to $\mathbf{u}$. Hence, restrict $d\Phi$ to $\mathbf{u}$'s 3-space, then take the vector dual. This vector shows the fastest increase of $\Phi$ along any possible ruler of $\mathbf{u}$. The measurement turns out to be $(g^{-1}(d\Phi, d\Phi) + (d\Phi(\mathbf{u}))^2)^{-1/2} d\Phi$, or:

$$dL_{\text{max-}\Phi} = \frac{1}{\sqrt{g^{\Phi\Phi} + (u^{\Phi})^2}} d\Phi, \tag{13}$$

if $\Phi$ is a coordinate. Alternatively, given $\Phi$, we can ask which observers $\mathbf{u}$ make extremal measurements. If a $\Phi = \text{const}$ slice is spatial ($g^{\Phi\Phi} > 0$), then observers with $u^{\Phi} \equiv d\Phi/d\tau = 0$ are possible, these measure $(g^{\Phi\Phi})^{-1/2} d\Phi$ in the direction $(d\Phi)^\sharp$. No ruler orientation for observers with $u^{\Phi} \neq 0$ can achieve this.

Consider a Schwarzschild observer parametrised by both $e$ and the Killing angular momentum per mass $\ell$. Its maximum possible $r$ and $\phi$ measurements are:

$$dL_{r\text{-max}} = \left(e^2 - \left(1 - \frac{2M}{r}\right)\frac{\ell^2}{r^2}\right)^{-1/2} dr, \qquad dL_{\phi\text{-max}} = \frac{r\, d\phi}{\sqrt{1 + \ell^2/r^2}}, \tag{14}$$

where motion is in the plane $\theta = \pi/2$. Incidentally, the corresponding ruler vectors are not orthogonal to one-another. Note that while $r$ is known as the reduced circumference,[12] the Euclidean 2-sphere measurement $r\, d\phi$ only holds for zero angular momentum $\ell = 0$, under this natural ruler orientation.

Our orthonormal frame approach is trivially the zero-distance limit of Fermi coordinates, however very few exact Fermi coordinate expressions are known.[13]

## 6. Discussion

There are many potential questions and objections. It is non-trivial to start from textbook material such as $\int ds$ or first-principles radar. Despite some excellent books,[14,15] our results fill an independent niche.

Given two events, why not extremise the length of spatial geodesics between them? This is a 4-dimensional approach, whereas ours uses a 1+3-dimensional splitting, measuring within the rest space of a given observer.

Is the result coordinate-dependent? The coordinate $\Phi$ serves as an extrinsic standard, but the length element is simply the ruler $\boldsymbol{\xi}^\flat$, which is also the metric interval restricted to a 1-dimensional subspace.

In Schwarzschild, $r$ is timelike inside the horizon, so how can it describe spatial measurement? Indeed the normal vector $(dr)^\sharp$ is timelike, but the measurement direction $\boldsymbol{\xi}$ is spatial.

## References


1. P. Painlevé, La mécanique classique et la théorie de la relativité, *Comptes rendus de l'Académie des Sciences* (1921).
2. R. Gautreau and B. Hoffmann, The Schwarzschild radial coordinate as a measure of proper distance, *Physical Review D* (1978).
3. K. Martel and E. Poisson, Regular coordinate systems for Schwarzschild and other spherical spacetimes, *American Journal of Physics* (2001).
4. T. Finch, Coordinate families for the Schwarzschild geometry based on radial timelike geodesics, *General Relativity and Gravitation* (2015).
5. D. Bini, A. Geralico and R. Jantzen, Separable geodesic action slicing in stationary spacetimes, *General Relativity and Gravitation* (2012).
6. C. MacLaurin, Schwarzschild spacetime under generalised Gullstrand-Painlevé slicing, in *Einstein equations: Physical and mathematical aspects of general relativity*, eds. S. Cacciatori, B. Güneysu and S. Pigola (Springer, 2019).
7. E. Taylor and J. Wheeler, *Exploring black holes: Introduction to general relativity* (Addison-Wesley, 2000).
8. G. Lemaître, L'Univers en expansion, *Publication du Laboratoire d'Astronomie et de Géodésie de l'Université de Louvain* (1932), translation in General Relativity and Gravitation, 1997.
9. L. Landau and E. Lifshitz, *Field theory* (GITTL, 1941).
10. R. Klauber, Toward a consistent theory of relativistic rotation, in *Relativity in rotating frames*, eds. G. Rizzi and M. Ruggiero (Springer, 2004).
11. W. Rindler, *Essential relativity: Special, general, and cosmological* (Springer-Verlag, 1977).
12. J. Droste, The field of a single centre in Einstein's theory of gravitation, and the motion of a particle in that field, *Koninklijke Nederlandse Akademie van Wetenschappen Proceedings Series B Physical Sciences* (1917), republished in General Relativity and Gravitation, 2002.
13. D. Bini, L. Lusanna and B. Mashhoon, Limitations of radar coordinates, *International Journal of Modern Physics D* (2005).
14. F. de Felice and D. Bini, *Classical measurements in curved space-times* (Cambridge, 2010).
15. R. Jantzen, P. Carini and D. Bini, *GEM [Gravitoelectromagnetism] the user manual: Understanding spacetime splittings and their relationships* (online draft, 2013).

# A study of inhomogeneous massless scalar gauge fields in cosmology

Ben David Normann[1*], Sigbjørn Hervik[1], Angelo Ricciardone[1,2] and Mikjel Thorsrud[3]

[1] *Faculty of Science and Technology, University of Stavanger, 4036, Stavanger, Norway*
[2] *INFN, Sezione di Padova, via Marzolo 8, I-35131, Padova, Italy*
[3] *Faculty of Engineering, Østfold University College, 1757 Halden, Norway*

[*] Presenter. Correspondance via *ben.d.normann@uis.no*.

Why is the Universe so homogeneous and isotropic? We summarize a general study of a $\gamma$-law perfect fluid alongside an inhomogeneous, massless scalar gauge field (with homogeneous gradient) in anisotropic spaces with General Relativity. The anisotropic matter sector is implemented as a $j$-form (field-strength level), where $j \in \{1, 3\}$, and the spaces studied are Bianchi space-times of solvable type. Wald's no-hair theorem is extended to include the $j$-form case. We highlight three new self-similar space-times: the Edge, the Rope and Wonderland. The latter solution is so far found to exist in the physical state space of types I,II, IV, $\text{VI}_0$, $\text{VI}_h$, $\text{VII}_0$ and $\text{VII}_h$, and is a global attractor in I and V. The stability analysis of the other types has not yet been performed. This paper is a summary of [1], with some remarks towards new results which will be further laid out in upcoming work.

*Keywords*: $p$-form gauge fields, anisotropic space-times, Bianchi models, inflation, dynamical system, orthonormal frame.

## 1. Introduction

Why does the Universe seem so isotropic on large scales [2, 3, 4]? Standard cosmology invokes this observation as a principle, as it is hard to solve the Einstein equations without symmetry requirements [5]. In order to understand this high degree of isotropy within the paradigm of GR, it is necessary to relive the theory of such assumption, seeking to replace it instead by an *explanation*. To this end we have softened the isotropy requirement of the cosmological principle by studying the Bianchi models of solvable type (types I-$\text{VII}_h$).

A natural generalization of massless scalar fields to the anisotropic case, is the $p$-form field[a] with $p = 1$ or 3. Collectively, we shall refer to this form field as a $j$-form field, since the equations will be the same in either case (as explained in Section 3). The connection to massless scalar fields may be drawn from Eq. (4). In this study we summarize some results we have obtained in our study so far, and mention a few results to be further reported on in upcoming work(s)(Normann's Ph.D. thesis). We would like to refer the reader to [1] for further references to previous works on related topics. Also, note the recent works [6, 7] and many interesting references therein.

[a] We here and throughout reference the form field on the field strength level.

## 2. The general *p*-form action

A natural candidate for anisotropic matter sourcing is the one stemming from the general $p$-form action [8]

$$S_{\rm f} = -\frac{1}{2}\int \mathcal{P}\wedge\star\mathcal{P}\,, \tag{1}$$

where $\mathcal{P}$ is a $p$ -form constructed by the exterior derivative of a $(p-1)$-form $\mathcal{K}$.

The equations of motion and the Bianchi identity, both obtained from the action, Eq. (1), may now be given as

$$\mathbf{d}\mathcal{P}=0 \quad \rightarrow \quad \nabla_{[\alpha_0}\mathcal{P}_{\alpha_1\cdots\alpha_p]}=0 \qquad \text{Bianchi Identity.} \tag{2}$$

$$\mathbf{d}\star\mathcal{P}=0 \quad \rightarrow \quad \nabla_{\alpha_1}\mathcal{P}^{\alpha_1\cdots\alpha_p}=0 \qquad \text{Equations of motion.} \tag{3}$$

Here $\mathcal{P}_{\mu_1\cdots\mu_p}$ are the components of $\mathcal{P}$ in a general basis. As evident from the above equations[b], our study assume that there is no source to the $j$-form field.

**General properties of the *p*-form action:** Note that the theories derived from the general $p$-form action Eq. (1) respect the following properties: (i) gauge invariance $\mathcal{L}\rightarrow\mathcal{L}$ under $\mathcal{K}\rightarrow\mathcal{K}+\mathbf{d}\mathcal{U}$, where $\mathcal{U}$ is a $(p-2)$ -form; (ii) only up to second order derivatives in equations of motion; (iii) Lagrangian is up to second order in field strength $\mathcal{P}$; (iv) constructed by exterior derivative of a $p$-form and (v) minimally coupled to gravity.

## 3. The *j*-form fluid

In our study, $\mathcal{P}$ is required homogeneous: $\mathcal{P}(t,\mathbf{x})\Rightarrow\mathcal{P}(t)$. However, generally, the underlying gauge field $\mathcal{K}(t,\mathbf{x})$ is allowed to vary both with space and time. Hence we study a $j$-form fieldstrength with an underlying inhomogenous gauge field. This is different from [9], where the gauge potential is a function of time only. In order to classify the possible cases of $p$-form matter fields that can be constructed from the exterior derivative of a $(p-1)$-form, the following notation is introduced: $\{a,b\}$ where $a$ denotes the rank of the $p$-form $\mathcal{P}$ and $b$ the rank of its Hodge dual $\star\mathcal{P}$. In four dimensional space-time ($a+b=4$) there are three distinct cases to consider: (i)$\{2,2\}$, (ii) $\{3,1\}$ or $\{1,3\}$ and (iii) $\{4,0\}$. The degeneracy in (ii) is due to the symmetry of the equations Eq. (2) and Eq. (3)[c]. This symmetry can also be seen in the action Eq. (1), up to a prefactor. Since these two cases yield the same set of equations, we effectively study both cases, by studying one of them.

In the following analysis the cases $\{1,3\}$ and $\{3,1\}$ will be taken into account. In order to include both scenarios, notation shall here, and throughout the rest of the paper, be such that $\mathcal{J}$ denotes either (i) the Hodge dual of a 3-form field strength

[b]The Hodge dual $\star\mathcal{P}$ is closed.

[c]The reason why this degeneracy is not found in the case (iii) is because $\mathcal{P}\neq\mathbf{d}\mathcal{K}$ in the case $\{0,4\}$. Hence one is left only with $\{4,0\}$.

$\mathcal{C}(t) = \mathbf{d}\mathcal{B}$ (where $\mathcal{B}(t, \mathbf{x})$ is a 2-form) or (ii) the 1-form field strength $\mathcal{A}(t) = \mathbf{d}\phi$ (where $\phi(t, \mathbf{x})$ is a scalar field). That both cases give rise to the same equations is evident.

**Energy-momentum tensor:** The energy-momentum tensor of the $j$-form fluid can be shown to be

$$\mathcal{L}_{\rm jf} = -\frac{1}{2}\mathcal{J}_\mu \mathcal{J}^\mu \quad \rightarrow \quad T^{\rm jf}_{\mu\nu} = \mathcal{J}_\mu \mathcal{J}_\nu - \frac{1}{2} g_{\mu\nu} \mathcal{J}_\gamma \mathcal{J}^\gamma . \tag{4}$$

Equations Eq. (2) and Eq. (3) now take the component forms

$$\mathbf{d}\mathcal{J} = 0 \quad \rightarrow \quad \nabla_{[\mu}\mathcal{J}_{\nu]} = 0, \tag{5}$$

$$\mathbf{d}\star\mathcal{J} = 0 \quad \rightarrow \quad \nabla_\mu \mathcal{J}^\mu = 0. \tag{6}$$

These are the equations for a massless scalar field. Hence, our study can be viewed as a study of an inhomogeneous, massless scalar field with a homogeneous gradient.

**State parameter** The field strength $\mathcal{J}_\alpha$ may be decomposed according to

$$\mathcal{J}_\alpha = -w\, u_\alpha + v_\alpha \,, \tag{7}$$

where the 4-velocity $u_\alpha$ is time-like ($u_\alpha u^\alpha < 0$), whereas $v_\alpha$ is defined to be orthogonal to $u_\alpha$ and therefore space-like ($v_\alpha v^\alpha > 0$). The range of the equation of state parameter $\xi$ defined through $p_{\rm f} = (\xi - 1)\rho_{\rm f}$[d] comes as no surprise. Performing the calculations, one finds

$$\xi = \frac{w^2 - v^2/3}{w^2 + v^2} + 1 \qquad \rightarrow \qquad \frac{2}{3} \leq \xi \leq 2\,. \tag{8}$$

The range of $\xi$ follows directly from requiring that $\mathcal{J}_\alpha \in \mathbb{R}$. Note that Eq. (8) is a dynamical equation of state, since the components of $\mathcal{J}$ in general change with time. The lower bound ($\xi = 2/3$) is found for $w = 0$ and the upper bound ($\xi = 2$) is found for $v = 0$. Note also that $w = v$ gives $\xi = 4/3$, as in the case of electromagnetic radiation.

## 4. Sourcing anisotropy with a $j$-form in General Relativity

We take the evolution to be governed by the Einstein Field Equations. In particular

$$R_{\mu\nu} - \frac{1}{2} R\, g_{\mu\nu} = T^{\rm pf}_{\mu\nu} + T^{\rm f}_{\mu\nu} + T^{\rm 4f}_{\mu\nu}\,, \tag{9}$$

where $R_{\mu\nu}$ is the Ricci tensor components, $R = R^\mu{}_\mu$ is the Ricci scalar and $T^{\rm pf}_{\mu\nu}$ and $T^{\rm f}_{\mu\nu}$ the perfect fluid and form fluid energy-momentum tensor components, respectively. The constants $8\pi G$ and $c$ are fixed to 1. Note that a 4-form is also added, playing the role of a cosmological constant.

It is, for simplicity, assumed that the three fluids do not interact.

[d] In a standard irreducible notation where $p$ denotes pressure and $\rho$ is the energy density.

### 4.1. *Bianchi models and choice of frame*

In dimension three there are nine different (classes of) Lie algebras – these are the nine different Bianchi types I-IX. The line element of the Bianchi models can be written as

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + \delta_{ab}\,\omega^a\omega^b \qquad \text{where} \qquad \mathbf{d}\omega^a = -\frac{1}{2}\gamma^a{}_{bc}\omega^b \wedge \omega^c - \gamma^a{}_{0c}\mathrm{d}t \wedge \omega^c. \tag{10}$$

$\{\omega^a\}$ is here a triad of 1-forms, and $\gamma^a{}_{bc}$ are the spatial structure coefficients of the Lie algebra characterizing the corresponding Bianchi type. The tetrad $\{\omega^\alpha\}$ is dual to the vector basis $\{\mathbf{e}_\alpha\}$, which must satisfy the relation $[\mathbf{e}_\mu, \mathbf{e}_\nu] = \gamma^\rho{}_{\mu\nu}\mathbf{e}_\rho$. Refer to ([10], Chapter 15) for details. The time direction is chosen orthogonal to the orbits of the isometry subgroup (i.e.: orthogonal to the three-dimensional hypersurfaces of homogeneity), and the fundamental observer's 4-velocity is aligned with this direction. It is given by $\mathbf{u} = \frac{\partial}{\partial t}$ where $t$ is the cosmological time. We also define a dimensionless time quantity $\tau$ such that $\mathrm{d}t/\mathrm{d}\tau = 1/H$, where $H$ is the Hubble parameter. The deceleration parameter $q$ is now such that $\mathrm{d}H/\mathrm{d}t = -(1+q)H^2$.

**A convenient frame** in which to conduct the analysis is the orthonormal frame. Such a frame will give first order evolution equations alongside a set of constraints which are useful to simplify the analysis. The Bianchi space-times analyzed in the present paper (I-VII$_h$) admit an Abelian $G_2$ subgroup. This allows for a $1+1+2$ split of the four dimensional space-time. As will become clear later, this translates into a $1+1+2$ decomposition of the Einstein Field Equations, as well as the Jacobi and the Bianchi identities. When the orthonormal frame approach is applied to $G_2$ cosmologies, it is common to choose a *group-invariant orbit-aligned* frame, i.e. an orthonormal frame which is invariant under the action of $G_2$ [11].

## 5. No-hair theorems for the $j$-form

No-hair theorems that in previous literature has been established for the Bianchi space-times in the presence of a cosmological constant and a perfect fluid are in this section extended to the presence of the $j$-form in the equations. In particular, it is found that the cosmic no-hair theorem [12] is valid also in this case[e]. Refer to [1] for proofs.

**Theorem 5.1 (First no-hair theorem).** *All Bianchi space-times I-VII$_h$ with a $j$-form, a non-phantom perfect fluid*[f] *and a positive cosmological constant will be asymptotically de Sitter with $\Omega_\Lambda = 1$ in the case where $\gamma > 0$ (and similarly $\Omega_\Lambda + \Omega_{\mathrm{pf}} = 1$ in the case where $\gamma = 0$).*

A similar but less general theorem holds also in the case of a perfect fluid with $0 \leq \gamma < 2/3$:

[e]Note that an anisotropic fluid may sustain an inflationary phase of expansion if it violates the strong or dominant energy condition [13]. A $j$-form respects these energy conditions.
[f]A perfect fluid is said to be phantom if $\gamma < 0$.

**Theorem 5.2 (Second no-hair theorem).** *All Bianchi space-times I-VII$_h$ with a $j$-form, a non-phantom perfect fluid $\Omega_{\rm pf}$ with equation of state parameter $0 \leq \gamma < 2/3$ will be asymptotically quasi de Sitter with $q = \frac{3}{2}\gamma - 1 < 0$.*

## 6. New stable, anisotropic, self-similar space-times

We have also performed a dynamical systems analysis of certain Bianchi types with a $\gamma$-law perfect fluid and a $j$-form fluid. Due to the no-hair theorems, we necessarily remove the cosmological constant in the further study. Among extensions of known self-similar space-times, we interestingly find three new anisotropic space-times; *Wonderland*, *The Edge* and *The Rope* in Bianchi type I. Their global stability has been determined by monotone functions. Extensions to these into other Bianchi types are found, as summarized in the following.

**Type I:** Type I splits into two further invariant subspaces; a temporal and a spatial part (as referred to the components of the $j$-form fluid). All of the three new solutions presented below are in the subspace where the $j$-form fluid is purely spatial.

*Wonderland* is an LRS solution containing both a non-rotating vector and the perfect fluid. The field strength is aligned with the LRS axis and the expansion asymmetry is of prolate type. Its range of existence is the open interval $\gamma \in (2/3, 2)$. It approaches the flat FLRW solution $\Omega_{\rm pf}$ when $\gamma \to 2/3$ and the Kasner solution ($\Sigma_+ = -1$) when $\gamma \to 2$. Interestingly, it has a deceleration parameter $q = -1+3\gamma/2$ identical to the flat FLRW solution. The line element of Wonderland is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + t^2\mathrm{d}x^2 + t^{\frac{2-\gamma}{\gamma}}(\mathrm{d}y^2 + \mathrm{d}z^2). \tag{11}$$

Global attractor for the Bianchi type I state space for $\gamma \in (2/3, 6/5]$.

*The Rope* contains a rotating vector and the perfect fluid. Its range of existence is the open interval $\gamma \in (6/5, 4/3)$. It approaches Wonderland in the limit $\gamma \to 6/5$ and the Edge in the limit $\gamma \to 4/3$. Like Wonderland, it has a deceleration parameter $q = -1 + 3\gamma/2$ identical to the flat FLRW solution. The Rope is not an LRS solution, although it is "almost LRS" close to Wonderland. The line element of the Rope is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + t^2\left(\mathrm{d}x + \sqrt{\frac{2(5\gamma-6)}{(2-\gamma)}}t^{1-\frac{2}{\gamma}}\mathrm{d}z\right)^2 + t^{\frac{2(4-3\gamma)}{\gamma}}\mathrm{d}y^2 + t^{\frac{4(\gamma-1)}{\gamma}}\mathrm{d}z^2. \tag{12}$$

Global attractor for the Bianchi type I state space for $\gamma \in (6/5, 4/3)$.

*The Edge* contains only a rotating vector and has deceleration parameter $q = 1$, similar to a radiation dominated universe. Since $\Omega_{\rm pf} = 0$, it exists in the entire range of models, $\gamma \in [0, 2]$. The line element of the Edge is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + t^2\left(\mathrm{d}x + \sqrt{2}t^{-1/2}\mathrm{d}z\right)^2 + \mathrm{d}y^2 + t\mathrm{d}z^2. \tag{13}$$

Global attractor for the Bianchi type I state space for $\gamma \in [4/3, 2)$.

**Type V:** In the type V subspace, we find a one-parameter extension of the Wonderland equilibrium set. Generally, in this one-parameter family, the $j-$form may have both temporal and spatial components. We also find Plane Waves. For $\gamma \in (2/3, 2)$ these two are the only attractors.

**Types II, IV, $\mathbf{VI_0}$, $\mathbf{VI_h}$, $\mathbf{VII_0}$ and $\mathbf{VII_h}$:** In an upcoming paper (part of Normann's Ph.D. thesis, due in October), we intend to study at least the types I, II, IV, $\mathrm{VII}_0$ and $\mathrm{VII}_h$ in depth. Here we only report a few preliminary results.

*Wonderland.* This solution (alongside one parameter extensions) are found in the physical state space of the types II, IV, $\mathrm{VI}_0$, $\mathrm{VI}_h$, $\mathrm{VII}_0$ and $\mathrm{VII}_h$. This is promising in the sense that this solution might not be sensitive to choices of background geometry. There remains, however, some work in order to pin down its stability.

*Edge and Rope.* These two solutions show up also in the physical state space of type II.

*Plane Waves.* In type IV we find a two parameter extension of the plane wave solution. In type $\mathrm{VII}_h$, the plane waves is a three parameter family of solutions.

## References


1. B. D. Normann, S. Hervik, A. Ricciardone and M. Thorsrud, *Class. Quantum Grav.* **35**, 9 (2018) doi:10.1088/1361-6382/aab3a7
2. P. A. R. Ade *et al.* [Planck Collaboration], *Astron. Astrophys.* **571** (2014) A23 [arXiv:1303.5083 [astro-ph.CO]].
3. P. A. R. Ade *et al.* [Planck Collaboration], *Astron. Astrophys.* **594** (2016) A16 [arXiv:1506.07135 [astro-ph.CO]].
4. D. Saadeh, S. M. Feeney, A. Pontzen, H. V. Peiris and J. D. McEwen, *Phys. Rev. Lett.* **117** (2016) no.13, 131302 [arXiv:1605.07178 [astro-ph.CO]].
5. J. Barrow, Some Generalities About Generality. In *The Philosophy of Cosmology* (pp. 85-108). Cambridge: Cambridge University Press. doi:10.1017/9781316535783.006 (2017)
6. J. Almeida et al., arXiv:1810.05301 [astro-ph.CO] PI/UAN-2018-640FT (2018).
7. J. Almeida et al., JCAP **03** 025 (2019).
8. M. Thorsrud, *Class. Quantum Grav.* **35** 095011 [arXiv:1712.02778 [gr-qc]].
9. J. D. Barrow and K. E. Kunze, *Phys. Rev. D* **55** (1997) 623 [hep-th/9608045].
10. Grøn Ø and Hervik S 2007 *Einstein's General Theory of Relativity* (Springer).
11. 1997 *Dynamical Systems in Cosmology* ed J Wainwright and G F R Ellis (Cambridge: Cambridge University Press).
12. R. M. Wald, *Phys. Rev. D* **28** (1983) 2118.
13. A. Maleknejad and M. M. Sheikh-Jabbari, *Phys. Rev. D* **85**, 123508 (2012) [arXiv:1203.0219 [hep-th]].

# A new class of non-aligned Einstein-Maxwell solutions with a geodesic, shearfree and non-expanding multiple Debever-Penrose vector

Norbert Van den Bergh

*Department of Mathematical Analysis FEA16, Gent University, Gent, 9000, Belgium*
*E-mail: norbert.vandenbergh@ugent.be*

In a recent study of algebraically special Einstein-Maxwell fields[1] it was shown that, for non-zero cosmological constant, non-aligned solutions cannot have a geodesic and shearfree multiple Debever-Penrose vector $\boldsymbol{k}$. When $\Lambda = 0$ such solutions do exist and can be classified, after fixing the null-tetrad such that $\Psi_0 = \Psi_1 = \Phi_1 = 0$ and $\Phi_0 = 1$, according to whether the Newman-Penrose coefficient $\pi$ is 0 or not. The family $\pi = 0$ contains the Griffiths solutions[2], with as sub-families the Cahen-Spelkens, Cahen-Leroy and Szekeres metrics. It was claimed in Ref. 2 (and repeated in Ref. 1) that for $\pi = 0$ both null-rays $\boldsymbol{k}$ and $\boldsymbol{\ell}$ are necessarily non-twisting ($\bar{\rho} - \rho = \bar{\mu} - \mu = 0$): while it is certainly true that $\mu(\bar{\rho} - \rho) = 0$, the case $\mu = 0$ appears to have been overlooked. I reduce the sub-family in which $\boldsymbol{k}$ is non-expanding ($\rho + \bar{\rho} = 0$) to an integrable system of pde's and I present an explicit family of solutions.

*Keywords*: Non-aligned Einstein-Maxwell fields.

## 1. Introduction

In the quest for exact solutions of the Einstein-Maxwell equations

$$R_{ab} - \frac{1}{2}Rg_{ab} + \Lambda g_{ab} = F_{ac}F_b{}^c - \frac{1}{4}g_{ab}F_{cd}F^{cd}, \tag{1}$$

a large amount of research has been devoted to the study of so called *aligned* Einstein-Maxwell fields, in which at least one of the principal null directions (PND's) of the electromagnetic field tensor $\boldsymbol{F}$ is parallel to a PND of the Weyl tensor, a so called Debever-Penrose (DP) direction. One of the main properties in this respect is the Goldberg-Sachs theorem[3], stating that, if a space-time admits a complex null tetrad $(\boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{m}, \overline{\boldsymbol{m}})$ such that $\boldsymbol{k}$ is shear-free and geodesic and $R_{ab}k^ak^b = R_{ab}k^am^b = R_{ab}m^am^b = 0$ (as is the case when $\boldsymbol{k}$ is also a PND of $\boldsymbol{F}$), then the Weyl tensor is algebraically special, with $\boldsymbol{k}$ being a multiple Weyl-PND. One of the topics considered in Ref. 1, dealing with the reverse problem, enquired after the existence of algebraically special (non-conformally flat and non-null) Einstein-Maxwell fields with a possible non-zero cosmological constant for which the multiple Weyl-PND $\boldsymbol{k}$ is geodesic and shear-free ($\Psi_0 = \Psi_1 = \kappa = \sigma = 0$) and for which $\boldsymbol{k}$ is *not* parallel to a PND of $\boldsymbol{F}$ ($\Phi_0 \neq 0$). Choosing a null-rotation about $\boldsymbol{k}$ such that $\Phi_1 = 0$, it follows that $\Phi_2 \neq 0$: with $\Phi_2 = 0$ $\boldsymbol{\ell}$ would be geodesic and shear-free and the Goldberg-Sachs theorem would imply $\Psi_3 = \Psi_4 = 0$. The Petrov type would then be D, in which case[4,5] the only null Einstein-Maxwell solutions are given by the (doubly aligned) Robinson-Trautman metrics.

Using the GHP formalism[a] the Maxwell equations and Bianchi equations become then

$$\eth'\Phi_0 = -\pi\Phi_0, \tag{2}$$
$$\text{Þ}\Phi_0 = 0, \tag{3}$$
$$\text{Þ}'\Phi_0 = -\mu\Phi_0, \tag{4}$$
$$\eth\Phi_2 = -\nu\Phi_0 + \tau\Phi_2, \tag{5}$$
$$\text{Þ}\Phi_2 = -\lambda\Phi_0 + \rho\Phi_2, \tag{6}$$
$$\eth\Psi_2 = -\pi\Phi_0\overline{\Phi_2} + 3\tau\Psi_2, \tag{7}$$
$$\text{Þ}\Psi_2 = \mu|\Phi_0|^2 + 3\rho\Psi_2, \tag{8}$$

after which the commutators $[\eth', \eth], [\eth', \text{Þ}], [\eth, \text{Þ}'], [\eth', \text{Þ}']$ and $[\text{Þ}', \text{Þ}]$ applied to $\Phi_0$ give

$$\eth\pi = (3\rho - \overline{\rho})\mu - 2\Psi_2 + \frac{R}{12}, \tag{9}$$
$$\text{Þ}\pi = 3\rho\pi, \tag{10}$$
$$\eth\mu = \overline{\lambda}\pi + 3\mu\tau, \tag{11}$$
$$\text{Þ}'\pi - \eth'\mu = 2\nu\rho - 2\lambda\tau - \pi\overline{\mu} - \mu\overline{\tau} - 2\Psi_3, \tag{12}$$
$$\text{Þ}\mu = \pi(\overline{\pi} + 3\tau) + 2\Psi_2 - \frac{R}{12}. \tag{13}$$

Herewith one of the GHP equations becomes a simple algebraic equation for $\Psi_2$,

$$\Psi_2 = \rho\mu - \tau\pi + \frac{R}{12}, \tag{14}$$

the Þ derivative of which results in $\rho R = 0$.

As $\rho = 0$ would imply $\Phi_0 = 0$, this leads to the remarkable consequence that an algebraically special Einstein-Maxwell solution possessing a shear-free and geodesic multiple Weyl-PND which is not a PND of $\boldsymbol{F}$ necessarily has a vanishing cosmological constant. The corresponding class of solutions is non-empty, as it contains the Griffiths[2] metrics, encompassing as special cases the metrics of Refs. 7, 8, 9, 10.

In Ref. 2 it was claimed that for $\pi = 0$ both null-rays $\boldsymbol{k}$ and $\boldsymbol{\ell}$ are necessarily non-twisting ($\bar{\rho} - \rho = \bar{\mu} - \mu = 0$). As a consequence it was also claimed in Ref. 1 that the Griffiths metrics are uniquely characterised by the condition $\pi = 0$. However, when $\pi = 0$ the only conclusion to be drawn from (10, 14) is that $\mu(\bar{\rho} - \rho) = 0$. When $\rho$ is real this indeed leads to the metrics of Ref. 2, while the case $\mu = 0$ appears to have been overlooked and leads, as shown in the section below, to new classes of solutions.

[a] Throughout I use the sign conventions and notations of Ref. 6 §7.4, with the tetrad basis vectors taken as $\boldsymbol{k}, \boldsymbol{\ell}, \boldsymbol{m}, \overline{\boldsymbol{m}}$ with $-k^a\ell_a = 1 = m^a\overline{m}_a$. In order to ease comparison with the (more familiar) Newman-Penrose formalism, I will write primed variables, such as $\kappa', \sigma', \rho'$ and $\tau'$, as their NP equivalents $-\nu, -\lambda, -\mu$ and $-\pi$.

## 2. The missing class

When $\pi = 0$ and $\mu = 0$ the equations from the previous paragraph immmediately lead to $\Psi_0 = \Psi_1 = \Psi_2 = 0$ and $\Psi_3 = \rho\nu - \lambda\tau$. We try to make some progress by looking for solutions for which $\boldsymbol{k}$ is non-expanding ($\rho + \overline{\rho} = 0$). Acting on this condition with the $\eth$ and $\text{Þ}$ operators, the GHP equations lead to $\tau = 0$ and

$$\rho^2 + |\Phi_0|^2 = 0, \tag{15}$$

the $\eth$ derivative of which implies $\lambda = \Phi_2\overline{\Phi_0}\rho^{-1}$. We translate these results into Newman-Penrose (NP) language and fix a boost and spatial rotation in the $\boldsymbol{k},\boldsymbol{\ell}$ and $\boldsymbol{m},\overline{\boldsymbol{m}}$ planes such that $\Phi_0 = 1$ and $\rho = i$. It follows that the only non-0 spin coefficients are $\rho$, $\nu$ and $\lambda = -i\Phi_2$, with the only non-vanishing components of the Weyl spinor being $\Psi_3 = i\nu$ and $\Psi_4$. As $[D,\Delta] = 0$ coordinates $u, v$ and $\zeta, \overline{\zeta}$ exist such that $D = \partial_u$, $\Delta = \partial_v$ and

$$\delta = e^{-iu}(\xi\partial_\zeta + \eta\partial_{\overline{\zeta}} + P\partial_u + Q\partial_v),$$

$\xi, \eta, P, Q$ being arbitrary functions. The $e^{-iu}$ factor is included for convenience: applying the $[\delta, D]$ commutator to $u, v$ and $\zeta$ shows that $\xi, \eta, P, Q$ are functions of $v, \zeta, \overline{\zeta}$ only. Introducing new variables $n = e^{-iu}\nu$ and $\phi = e^{-2iu}\Phi_2$ it follows that also $n$ and $\phi$ depend on $v, \zeta, \overline{\zeta}$ only, with the full set of Jacobi equations and field equations reducing to the following system of pde's:

$$P_v + i\overline{P\phi} - \overline{n} = 0, \tag{16}$$
$$Q_v + i\overline{Q\phi} = 0, \tag{17}$$
$$\xi_v + i\overline{\eta}\overline{\phi} = 0, \tag{18}$$
$$\eta_v + i\overline{\xi}\overline{\phi} = 0, \tag{19}$$

$$e^{-iu}\overline{\delta}P - e^{iu}\delta\overline{P} - 2i|P|^2 = 0, \tag{20}$$
$$e^{-iu}\overline{\delta}Q - e^{iu}\delta\overline{Q} - 2i(\Re Q\overline{P} - 1) = 0, \tag{21}$$
$$e^{-iu}\overline{\delta}\xi - e^{iu}\delta\overline{\eta} - i(\xi\overline{P} + \overline{\eta}P) = 0, \tag{22}$$

$$e^{iu}\delta n = -iPn + 2|\phi|^2, \tag{23}$$
$$e^{iu}\delta\phi = -2iP\phi - n, \tag{24}$$

with the $\Psi_4$ component of the Weyl spinor given by $\Psi_4 = ie^{2iu}(\overline{n}\overline{P} + \Delta\phi) + e^{iu}\overline{\delta}n$. This system is integrable and a simple solution is obtained by assuming $\phi = H^2$ to be a positive constant: in terms of new coordinates $u, v, a, b$ the null tetrad is given by

$$\boldsymbol{\omega}^1 = \frac{e^{iu}}{2H(a+b)}[e^{i\frac{\pi}{4}-v}\mathrm{d}a - e^{-i\frac{\pi}{4}+v}\mathrm{d}b], \tag{25}$$

$$\boldsymbol{\omega}^3 = \frac{1}{H^2}\left[\mathrm{d}v - \frac{1}{2(a+b)}\mathrm{d}(a-b)\right], \tag{26}$$

$$\boldsymbol{\omega}^4 = \mathrm{d}u - \frac{1}{2(a+b)}[e^{-2v}\mathrm{d}a - e^{2v}\mathrm{d}b]. \tag{27}$$

The corresponding space-time metric is

$$\mathrm{d}s^2 = \frac{1}{H^2}\left[-2\mathrm{d}v + \frac{1}{a+b}\mathrm{d}(a-b)\right]\mathrm{d}u + \frac{1}{H^2(a+b)}(e^{-2v}\mathrm{d}a - e^{2v}\mathrm{d}b)\mathrm{d}v + \frac{\cosh 2v}{H^2(a+b)^2}\mathrm{d}a\mathrm{d}b \tag{28}$$

and the Maxwell field and energy-momentum tensor are obtained as

$$\boldsymbol{F} = iH^2(\boldsymbol{\omega}^1 - \boldsymbol{\omega}^2)\wedge\boldsymbol{\omega}^3 + i(e^{-2iu}\boldsymbol{\omega}^1 - e^{2iu}\boldsymbol{\omega}^2)\boldsymbol{\omega}^4, \tag{29}$$

$$\boldsymbol{T} = 2H^2(e^{-2iu}\boldsymbol{\omega}^{1^2} + e^{2iu}\boldsymbol{\omega}^{2^2} + H^2\boldsymbol{\omega}^{3^2}) + 2\boldsymbol{\omega}^{4^2}. \tag{30}$$

The Petrov type is III and the space-time admits three Killing vectors, $\partial_u, \partial_a - \partial_b$ and $a\partial_a + b\partial b$. At first sight it seems odd that the Weyl spinor components $\Psi_3$ and $\Psi_4$ depend on $u$, while the frame has been "invariantly" fixed. The explanation however is that the the frame was fixed by means of a null rotation putting $\Phi_0 = 1$, while the Maxwell field itself does not share the space-time symmetries: $\boldsymbol{F}$ is not Lie-propagated along the integral curves of the null Killing vector $\partial_u$ (this also shows that the family of solutions presented here is distinct from the Einstein-Maxwell solutions admitting null-Killing vectors of Ref. 11.

**Acknowledgment**

All calculations were done using the Maple symbolic algebra system, while the properties of the metric (28) were checked with the aid of Maple's Differential Geometry package.

**References**

1. N. Van den Bergh, *Gen. Rel. Grav.*, **49**, 1, 2017
2. J. Griffiths, *Gen. Rel. Grav.*, **18**, 4, 1986
3. J.N. Goldberg, R.K. Sachs, *Acta Phys. Pol. Suppl.*, **22**, 13, 1962
4. R. Debever, N. Van den Bergh, J. Leroy, *Class. Quantum Grav.*, **6**, 1373, 1989
5. N. Van den Bergh, *Class. Quantum Grav.*, **6**, 1871, 1989
6. H. Stephani, D. Kramer, M.A.H. MacCallum, C. Hoenselaers, E. Herlt, *Exact solutions of Einstein's field equations*, Cambridge University Press, 2003
7. M. Cahen, J. Leroy, *J. Math. Mech.*, **16**, 501, 1966
8. M. Cahen, J. Spelkens, *Bull. Soc. Math. Belg.*, **53**, 817, 1967
9. J.B. Griffiths, *J. Phys. A*, **16**, 1175, 1983
10. P. Szekeres, *J. Math. Phys.*, **7**, 751, 1966
11. B. Lucács, Z. Perjés, Á. Sebestyén, *J. Math. Phys.*, **22**, 1248, 1981

# Equal-field cylindrical electrovacuum universes

V. Balek

*Department of Theoretical Physics, Comenius University,*
*Mlynská dolina, 842 48 Bratislava, Slovakia*
*E-mail: balek@fmph.uniba.sk*

M. Čermák

*Institute of Physical Electronics, Masaryk University,*
*Kotlářská 267/2, 611 37 Brno, Czech Republic*
*E-mail: 63855@seznam.cz*

We explore a family of solutions to Einstein-Maxwell equations with cylindrical symmetry, in which both metric and electromagnetic field are expressed in terms of elementary functions. This is achieved by choosing the integration constants in (00) and (22) components of Einstein equations equal to each other, which is equivalent to putting electric field (transversal) and magnetic field (longitudinal), under suitable rescaling of coordinates, equal to each other. We discuss the connection between our family of solutions and previously known solutions summarized in the 1983 paper by MacCallum, analyze properties of spacetime described by our solutions and show how the equal-field condition can be relaxed perturbatively.

*Keywords*: Einstein-Maxwell equations; cylindrical symmetry; perturbation theory.

## 1. Einstein-Maxwell equations

Consider cylindrical version of Weyl-Lewis-Papapetrou metric,

$$ds^2 = \frac{1}{F}(dt + Ad\phi)^2 - F[B(dr^2 + dz^2) + w^2 d\phi^2], \tag{1}$$

with the electromagnetic vector potential of the form

$$\boldsymbol{A} = A_0 \boldsymbol{d}t + A_2 \boldsymbol{d}\phi, \tag{2}$$

where the functions $F$, $A$, $B$ and $w$, as well as $A_0$ and $A_2$, depend only on the coordinate $r$ [1].

Due to high symmetry of the problem, we can perform one integration of all Maxwell equations and all Einstein equations but one, reducing them to differential equations of first order. For Maxwell equations it is seen immediately and for Einstein equations we can use identity

$$\xi^{\mu\nu}{}_{;\nu} = 2R^{\mu}_{\nu}\xi^{\nu}, \tag{3}$$

where $\xi_{\mu\nu}$ is a bivector associated with Killing vector $\xi^{\mu}$, $\xi_{\mu\nu} = \xi_{\nu,\mu} - \xi_{\mu,\nu} = -2\xi_{\mu;\nu}$, and $R_{\mu\nu}$ is Ricci tensor. Integrated equations were obtained first with the help of Hamiltonian formalism [2], and then with the help of complex potentials [3] introduced in the series of papers [4–6]. Neither procedure used the identity (3) explicitly, although in the paper [6] the authors mentioned it with a reference to Synge's book [7].

First integral of Maxwell equations is

$$\sqrt{-^2 g} A'_A = g_{AB} C^B, \tag{4}$$

and first integral of Einstein equations (five of six) is

$$\sqrt{-{}^2g}g^{ac}g'_{cb} = 2\delta^a_A\delta^B_b C^A A_B - \delta^a_b C^C A_C + K_b{}^a, \tag{5}$$

where the indices $A, B, \ldots$ assume the values 0 and 2, the indices $a, b, \ldots$ assume the values 0, 2, 3 (the coordinates are identified in a usual way as $(x^0, x^1, x^2, x^3) = (t, r, \phi, z)$), ${}^2g$ is determinant of the $(t, \phi)$ part of the metric, ${}^2g = \det(g_{AB}) = -w^2$, and $C^A$, $K_b{}^a$ are constants. After the $\binom{1}{1}$ component of Einstein equations is added to the system, we have eight equations for six functions $F$, $A$, $B$, $w$, $A_0$, $A_2$; thus, two equations are just constraints on integration constants.

The trace of the $\binom{A}{B}$ part of (5) yields $w = (1/2)K_A{}^A r$, so that by rescaling $r$ (as well as $z$, in order to preserve the quadratic form $dr^2 + dz^2$) we obtain $w = r$. If we also gauge away the constants ${K_0}^2$ and ${K_2}^0$, in the $\binom{A}{B}$ part of (5) there remains a single constant $C = (1/2)({K_0}^0 - {K_2}^2)$. The constraint on $C$ can be turned into an independent equation by inserting expressions for $A_A$ in terms of $A$, $A'$, $F$, $F'$ into it, and equations (4) can be then reinterpreted as constraints. However, one of these equations is satisfied identically, so that we are left with one constraint only.

## 2. Equal-field solution

We are interested in solutions to equations (4), (5) with $C = 0$. Their subclass was found long ago by using transformation to rotating frame in which the metric was static[8]. Later it was noted[3,9] that the subclass belongs to a broader family with constant ratio $\mathcal{R} = A'_2/A'_0$. It has $\mathcal{R} = -1/k$, where $k = C^0/C^2$, while a class of solutions proposed in[10] has another distinguished value $\mathcal{R} = -k$. We construct solutions with arbitrary $\mathcal{R}$ by solving the equations directly; however, our solutions can be obtained also from solutions with $\mathcal{R} = -1/k$ by means of the scaling discussed below.

Introduce a new function $f = (Fr/A_+)^{1/2}$, where $A_+ = A + k$. Einstein equations yield a pair of coupled equations for the functions $f$ and $A$,

$$\frac{f''}{f} - \frac{p}{2A_+r^2}\Big(1 + \frac{3}{2}\frac{rA'}{A_+}\Big) - \frac{q + 1/2}{2r^2} = 0, \quad r[A_+(f^4 - 1)]' + pf^4 = 0, \tag{6}$$

where $p = 2kC$ and $q = K_3^3 - C$. For $C = 0$ we have $p = 0$, so that the first equation (6) decouples from the second equation and can be solved analytically. The solution is

$$f = r^{1/2} \times \begin{cases} (k_1 r^a + k_2 r^{-a}), \text{ if } a > 0 \\ (k_1 \ln r + k_2), \text{ if } a = 0 \\ [k_1 \cos(\hat{a} \ln r) + k_2 \sin(\hat{a} \ln r)], \text{ if } a = i\hat{a} \end{cases}, \tag{7}$$

where $a = \sqrt{(1 + q)/2}$. The function $A_+$ is obtained immediately from the second equation (6), the function $F$ appears in the definition of $f$, for the function $\mathcal{B} = BF$ we have separate equation and the functions $A_A$ are given algebraically in terms of $A$, $A'$, $F$, $F'$. In this way we find

$$A_+ = \frac{k_A}{f^4 - 1}, \quad Fr = \frac{k_A f^2}{f^4 - 1}, \quad \mathcal{B} = k_B r^q f^2, \tag{8a}$$

$$A_0 = \frac{2}{C^2 k_A}\frac{rf'}{f}, \quad A_2 = -(k_A + k)A_0. \tag{8b}$$

Finally, upon inserting expressions for $A_0$, $F$ and $A$ into the first equation (4), we find

$$\begin{cases} 4a^2 k_1 k_2 \\ -k_1^2 \\ -\hat{a}^2(k_1^2 + k_2^2) \end{cases} = \frac{1}{2}(C^2)^2 k_A, \text{ if } \begin{cases} a > 0 \\ a = 0 \\ a = i\hat{a} \end{cases}. \tag{9}$$

The solution preserves its form under the scaling $t \to \gamma t$ and $(r, z) \to \gamma^{-1}(r, z)$, if $f$ stays unchanged, which can be always achieved by suitably scaling $(k_1, k_2)$, and if the remaining constants scale as $(k, k_A) \to \gamma(k, k_A)$ and $k_B \to \gamma^{q+2} k_B$. In case the constant $\kappa = k_A + k$ is nonzero, we can use this procedure to normalize it to $\pm 1$ by fixing the unit of length and choosing $\gamma = |\kappa|^{-1}$. As a result, we make the potentials $A_0$ and $A_2$, as well as the field strengths $F_{01}$ and $F_{21}$, equal in absolute value. In that sense we call the solution (8) "equal-field solution". Note, however, that the field strengths are the same only when computed with respect to *coordinate* basis. The physical field strengths which refer to the orthonormal basis $e_{\hat{\alpha}}{}^{\mu}$ carried by the observer at rest, $E^{\hat{1}} = F_{\hat{0}\hat{1}} = \sqrt{F/\mathcal{B}}F_{01}$ and $B^{\hat{3}} = F_{\hat{2}\hat{1}} = (1/\sqrt{F\mathcal{B}})(F_{21} - AF_{01})/r$, differ from each other.

## 3. Properties of spacetime

The $(t, \phi)$ part of the metric can be expressed in two different ways,

$$ds_2^2 = \frac{1}{F}(dt + Ad\phi)^2 - Fr^2 d\phi^2 = \frac{1}{\mathcal{F}}dt^2 - \mathcal{F}r^2(d\phi - \omega dt)^2, \tag{10}$$

where the functions $\mathcal{F}$, $\omega$ are defined in terms of the function $\mathcal{A} = A^2 - F^2 r^2$ as $\mathcal{F} = -\mathcal{A}/(Fr^2)$, $\omega = -A/\mathcal{A}$. The cylinders with $1/F = 0$ are *static limits* (surfaces separating regions in which observers can stay at rest from regions in which they must rotate) and the cylinders with $\mathcal{F} = 0$ are *chronological limits* (surfaces separating regions in which time travel is forbidden from regions in which it is allowed). Suppose $a > 0$, so that $f = (r/r_1)^{a_+} \pm (r/r_2)^{-a_-}$, where $a_\pm = a \pm 1/2$ and the sign in front of the second term is given by the sign of $k_A$. For $k_A > 0$, $a > 1/2$, as well as for $k_A < 0$, $a < 1/2$, the function $f$ has one minimum, defining a cylinder composed of photon orbits; in the latter case it has also one zero, defining a singular "cylinder" (in fact, a line parallel to the $z$ axis).

The behavior of light cones in equal-field solution is depicted in fig. 1. In the left panel, regions in the plane $(a, r_2/r_1)$ with static (S) and rotating (R) neighborhood of photon orbits are shown; in the right panel, light cones in metric with $k_A > 0$, $a > 1/2$ (upper diagram) and $k_A < 0$, $a < 1/2$ (lower diagram) are displayed at various values of $r$ in the planes tangential to respective cylinders. The sign of $k_A$ is denoted by $+$ and $-$ in the left panel and singular "cylinder", photon cylinder,

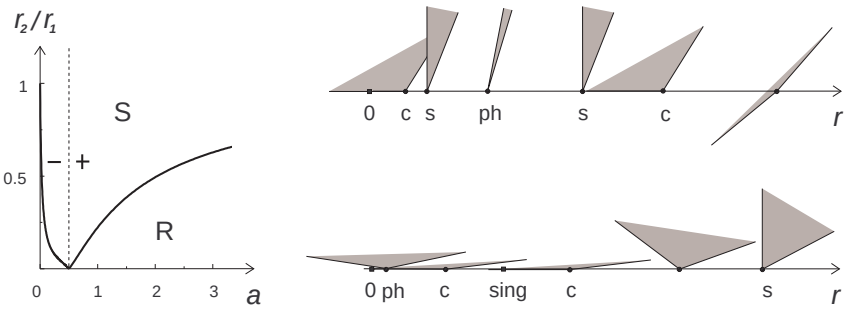



Fig. 1. Regions in parametric space with and without rotating domain in the neighborhood of photon orbits (left) and light cones in two typical spacetimes (right).

static limits and chronology limits are denoted by *sing*, *ph*, *s* and *c* in the right panel.

Circular orbits are given by the conditions $\dot{r}^2 = (\dot{r}^2)' = 0$, where $\dot{r}$ is radial component of 4-velocity. For a photon with impact parameter $b$ we have

$$\dot{r}^2 \propto (b - b_-)(b_+ - b), \quad b_\pm = -A \pm Fr, \tag{11}$$

so that on photon orbits it holds $b = b_\pm$ and $b'_\pm = 0$. From the explicit expression for the functions $b_\pm$, $b_\pm = k \pm k_A/(f^2 \pm 1)$, we find that the latter equation is equivalent to $f' = 0$. (Solution $f = 0$, which corresponds to the singular line parallel to the $z$ axis, does not count.) In the case $a > 0$ this implies that photon orbit exists only for $a > 1/2$ if $k_A > 0$ and $a < 1/2$ if $k_A < 0$, and that its radius is $r_{ph} = (\alpha r_1^{a_+} r_2^{a_-})^{1/(2a)}$, where $\alpha = |a_-|/a_+$. Condition of stability reads $(\dot{r}^2)'' < 0$, which yields, after including the prefactor into equation (11), $\pm b''_\pm < 0$, or equivalently, $k_A f f'' = k_A a_+ a_- f^2 r^{-2} > 0$. This is clearly satisfied for all orbits.

For massive particles, similar analysis leads to following conditions under which the orbits exist and are stable:

$$\text{existence: } k_A(\alpha - x)(x - \hat{\alpha}) > 0, \text{ stability: } (\alpha - x)X > 0, \tag{12}$$

where $x = \mathrm{sgn} k_A r^{2a}/(r_1^{a_+} r_2^{a_-})$, $\hat{\alpha} = (a-1)/(a+1)$ and $X = a_+(1+a)x^3 + a_+(1-3a)x^2 + a_-(1+3a)x + a_-(1-a)$. Location and character of orbits depending on the parameter $a$ is depicted in fig. 2. Blank regions contain stable ($S$) and unstable ($N$) orbits, in shaded regions there are no orbits at all. The line denoted by 'ph' represents photon orbits, the light line represents potential marginally stable orbits (circles on which $X = 0$), and the rugged line is the singularity at $f = 0$.

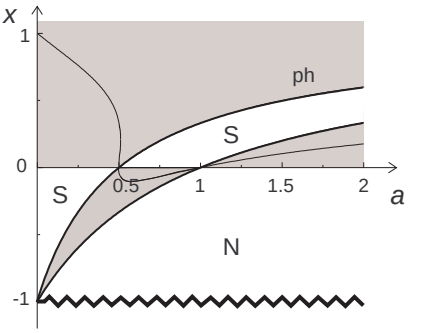



Fig. 2. Existence and stability of particle orbits.

## 4. Perturbing the metric

Equations for the functions $f$ and $\mu = A_+^{-1}$, rewritten into the variable $\tau = \ln r$, read

$$\ddot{f} - \dot{f} - Qf = \frac{p}{2}\Big(\mu - \frac{3}{2}\dot{\mu}\Big)f, \quad \Big(\frac{f^4-1}{\mu}\Big)^{\cdot} = -pf^4, \tag{13}$$

where $Q = (1/2)(q + 1/2)$ and the dot denotes differentiation with respect to $\tau$. Suppose the parameter $p$ is small. In the zeroth order in $p$ we have the previous expressions $f_0 = k_1 e^{a_+\tau} + k_2 e^{-a_-\tau}$ and $\mu_0 = k_A^{-1}(f_0^4 - 1)$, and if we write down the equations for the first corrections to $f$ and $A$,

$$\ddot{f}_1 - \dot{f}_1 - Qf_1 = \frac{p}{2k_A}(f_0 - 6\dot{f}_0)f_0^4, \quad \Big[(f_0^4-1)A_1 + \frac{4f_0^3}{\mu_0}f_1\Big]^{\cdot} = -pf_0^4 \tag{14}$$

(we have skipped the term proportional to $f_0$ from the first equation, since it can be always removed by redefining $Q$), we find

$$f_1 = \frac{p}{2k_A}f_0\Phi_0, \quad A_1 = -p\frac{f_0^4+1}{(f_0^4-1)^2}\Phi_0, \tag{15}$$

where $\Phi_0 = \int f_0^4 d\tau$. To complete the calculation of the perturbed metric, we can determine $F_1$ from the definition of $f$ and $B_1$ from the equation for $\mathcal{B}$. Finally, we can impose the first equation (4) as a constraint on the perturbed solution. With the notations $k_1 = k_{10}(1 + p\lambda_1)$, $k_2 = k_{20}(1 + p\lambda_2)$, $k_A = k_{A0}(1 + p\lambda_A)$ and $C^2 = C_0^2(1 + p\lambda_C)$, the resulting equation reads

$$\lambda_C = -\frac{1}{2}\Big(\lambda_1 + \lambda_2 - \lambda_A - \frac{1}{2a^2 k_{A0}}\Big). \tag{16}$$

## Acknowledgments

VB was supported by the grant VEGA 1/0985/16.

## References


1. H. Stephani, D. Kramer, M. MacCallum, C. Hoenselaers, E. Hertl: *Exact Solutions of Einstein's Field Equations*, CUP, Cambridge (2003).
2. D. M. Chitre, R. Güven, Y. Nutku, J. Math. Phys. **16**, 475 (1975).
3. N. Van den Bergh, P. Wils, J. Phys. **A**, 3843 (1982).
4. W. Kinnersley, J. Math. Phys. **18**, 1529 (1977).
5. W. Kinnersley, D. M. Chitre, J. Math. Phys. **18**, 1538 (1977).
6. W. Kinnersley, D. M. Chitre, J. Math. Phys. **19**, 1926 (1978).
7. J. Synge: *Relativity: The General Theory*, North-Holland, Amsterdam (1960).
8. N. Arbex, M. M. Som, Il Nuovo Cim. **13 B**, 49 (1973).
9. M. MacCallum, J. Phys. **A16**, 3854 (1983).
10. J. D. McCrea, J. Phys. **A15**, 1587 (1982).

# Kerr-Newman spinning particle meets "New Physics" beyond the Standard Model: Unification of gravitation with particle physics

Burinskii, Alexander*

*Nuclear Safety Institute, Russian Academy of Sciences,*
*B.Tulskaya 52, Moscow, 115191, Russia*
**E-mail: burinskii@mail.ru.*
*www.university_name.edu*

Standard model has to be generalized to a "New Physics" beyond the Standard Model. Main problem is the lack of consistency SM with gravity. We analyse Kerr-Newman spinning particle which is consistent with gravity by nature and, contrary to opinion that gravity conflicts with quantum theory, we obtain that spinning Kerr's gravity collaborates with quantum theory in the process of formation of spinning particle. The most dramatic is the shift of the fundamental scale from Planck to Compton distances.

*Keywords*: Kerr-Newman solution, spin, Compton scale, electron, supersymmetry, Higgs, Landau-Ginzburg, phase transition, bag model, strings.

## 1. Introduction

The Standard model (SM) gives at present the best understanding of the particle physics and explained successfully the most off experimental results. Meanwhile, it is not ultimate and has several unsolved problems that are supposed to be resolved in the "New Physics" beyond the SM. In particular,

(i) origin of the mechanism symmetry breaking is not defined up to corresponding nonperturbative solutions,

(ii) the extra dimensions and superpartners were not confirmed in the recent experiments: *"... the LHC experiments have found not even a hint of supersymmetry, despite many searches making different assumptions about the supersymmetric spectrum..."*, [1],

and finally, the main problem,

(iii) Quantum theory and Gravity cannot be combined in a unified theory.

Reason of this trouble is that GRAVITY REJECTS POINTLIKE particles following from stochastic interpretation of quantum theory, because the right side of Einstein equations, $G_{\mu\nu} = 8\pi T_{\mu\nu}$ requires an *extended nonperturbative* solution in configuration space.

One sees that roots of this conflict lie in stochastic interpretation of quantum theory, which had intensive objections from Einstein, de Broglie, Schrodinger and other famous founders of the modern theoretical physics. As is known, the stochastic Copenhagen interpretation won and becomes dominate, but as result of this victory, we have got incompatibility of the gravity with particle physics. People can accept this interpretation, but gravity definitely opposed to this!

Perfect old model of the consistent with gravity particle was the Kaluza-Klein (KK) theory with FIFTH CIRCLING DIMENSION. The corresponding 5d tensor

splits in 4d + 1,

$$^5G_{MN} = 0 \;\Rightarrow\; ^4G_{\mu\nu} = 8\pi T_{\mu\nu}, \tag{1}$$

and potential $A^\mu$, having extra fifth component – scalar field $\Phi$.

Main problem of this theory was invisibility of this extra dimension, and to avoid this problem, authors assumed that compactification occurs at the Planck scale, about $M_p \approx 10^{22} m_e$.

The Planck scale and the additional fifth dimension of the KK model were precursors of the modern superstring theory. The Planck scale was brilliantly matched with well-known "Weakness of gravity" – the opinion based on the known Schwarzschild solution which distorts space near gravitational radius

$$r \sim r_g = 2Gm, \tag{2}$$

which is small with respect to other interactions even at the Planck scale.

Superstring theory inherited Planck scale and higher dimensions together with other remarkable features of the KK model. Theoreticians of the string theory argued: *"... at such energies (Planck scale), strings can be accurately approximated by point particles. This explains why quantum field theory has been so successful in describing our world."* [2].

However, just the Planck scale was the main defect, because it is extremely far from the real scale of the particle physics. Considering electron as one of the principal particles of quantum theory, we should accept that unification must be defined by the Compton scale – electroweak scale of the Standard Model.

**"New physics" beyond the SM was created in 1963 by Kerr solution[3,4].** The spinning Kerr gravity changed the existing conception of gravitational interaction, resolving the principal problem – removal of the giant gap between the Planck scale and the Compton scale required for consistency gravity with physics of the Standard model[5,6].

**Kerr solution exhibited new effect of distortion**, showing that space distorts ***topologically*** at the distance

$$r \sim r_c = \frac{J}{m}, \tag{3}$$

which is proportional to angular momentum $J$ and inverse proportional to mass m. For spinning particles, which have the huge spin/mass ratio (about 20-22 orders in dimensionless units $G = c = \hbar = 1$) topological influence of gravity becomes very strong. The Kerr singular ring forms the branch line of the space, and space differs drastically from the expected flat background in the area of order radius of the Kerr ring. The black hole horizons disappear, opening the hole of another type – two-sheeted space forming the Einstein-Rosen bridge – the door to another sheet of the Kerr geometry. For parameters of an electron, size of this hole is

$$a = \hbar/2mc, \tag{4}$$

which corresponds to the reduced Compton wave length. It gives miraculous magnification of the effective gravitational interaction, scale of which is increased on 22 orders concerning the accepted Planck scale. Conflict between Gravity and Quantum theory increases dramatically, because the both theories become on the equal footing. The **nonperturbative bag-like solution** based on the supersymmetric Higgs model of symmetry breaking specifies the above problem (ii). Supersymmetric bag removes conflict without modification of the Einstein-Maxwell field equations. Bag model regularizes the original Kerr solution near the Kerr singular ring, closing the door to the second sheet. Bag is formed as a *supersymmetric Domain Wall solution (boundary of the bag) which separates external gravitational field from the supersymmetric and flat internal vacuum state* – the flat area which is necessary for work of quantum theory, see Fig. 1. As is known, *the bags are soft and flexible*, and shape of the bag, as well as its size are defined in the process of interaction between gravitational, electromagnetic and quantum fields (see Fig. 2). Problem (ii) is solved by use of the supersymmetric Landau-Ginzburg field model of phase transition instead of the "superpartner" models.

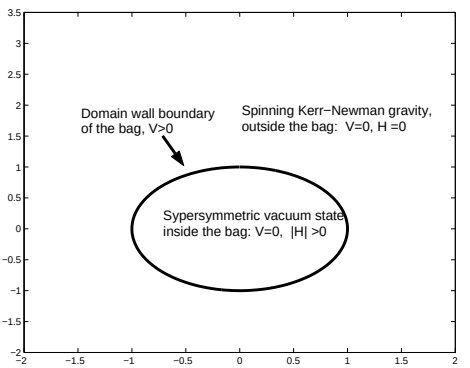



Fig. 1. Separation of the flat Quantum interior from zone of the spinning Gravity.

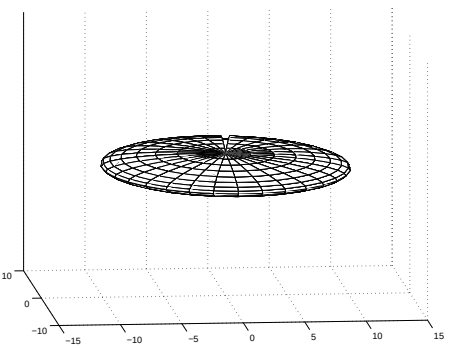


Fig. 2. Bag takes form of a thin disk.

## 2. The Kerr-Newman Spinning Particle

In 1968 Carter noticed [7], that Kerr-Newman (KN) solution has gyromagnetic ratio $g = 2$, as that of the Dirac electron [4]. The KN solution was considered as a classical model of electron [8–10]. Since for electron $a >> m$, the horizons disappear and topology becomes broken by naked singular ring. Metric of the KN solution in the Kerr-Schild form is [4]

$$g_{\mu\nu} = \eta_{\mu\nu} + 2Hk_\mu k_\nu, \tag{5}$$

where $\eta_{\mu\nu}$ is metric of auxiliary Minkowski space $M^4$, (signature $(-++{+})$), and scalar function

$$H = \frac{mr - e^2/2}{r^2 + a^2\cos^2\theta}, \tag{6}$$

is given in Kerr's oblate spheroidal coordinates $r$ and $\theta$. The null vector field $k_\mu(x)$, $(k_\mu k^\mu = 0)$ determines Kerr congruence which has the form of a vortex, dragged by the naked Kerr singular ring $r = 0, \cos\theta = 0$, see Fig. 1, Fig. 2, which is branch line of Kerr space into two sheets. The congruence $k^-_\mu$, goes to disk $r = 0$ as 'ingoing', and intersecting disk, it is extended analytically, turning into 'out-going' congruence $k^+_\mu$, which forms a second sheet of the metric on the same background. The space becomes two-sheeted, with two different metrics on the same background, $g^\pm_{\mu\nu} = \eta_{\mu\nu} + 2Hk^\pm_\mu k^\pm_\nu$. The Kerr singular ring $r = 0,\ \cos\theta = 0$, of the Compton radius $a = \hbar/2m$, breaks space topologically, forming the famous Einstein-Rosen bridge[11]. Studying of the KN solution was mainly in two directions:

A) Israel[8] suggested to truncate negative sheet of Kerr space along disk $r = 0$ and put at this disk a distribution of matter in accord with right hand side of the Einstein equations. He obtained a strange matter distribution and argued that this disk-like source of Compton radius can be considered as a classical model of electron. Hamity showed that source of Kerr solution can be considered as a relativistically rotating disk. Later on López generalized this model to model of a rotating bubble[10].

B) Another stringy version of the Kerr source was developed in our works with Prof. D. Ivanenko[9,12]. The Kerr singular ring was compared there with Nielsen-Olesen (NO) string model as a vortex in superconductor[13]. It was shown that Kerr's singular ring can be considered as a lightlike string which can bear the electromagnetic and spinor traveling waves. Much later, such pp-wave strings were considered as solitonic string solutions to low energy string theory[14,15].

Bag model merges directions A) and B). In the both these directions of investigation source of the Kerr geometry created a gravitational vortex of the Compton size, and there appeared the question: Why this vortex and singular ring were not observed experimentally, and electron looks like a point? Answer follows from mechanism of Lorentz contraction[16]. *For external observer Kerr's vortex shrinks to point by Lorentz contraction.*

The suggested in[5,17] supersymmetric bag model solves problem of unification with gravity by separation of their zones of influence. Quantum theory requires flat space, at least in the Compton zone, while spinning Kerr's gravity spoils space topologically. Supersymmetric bag model realizes expulsion of gravity and electromagnetic field from interior of the bag, forming tree zones: flat quantum interior **(I)**, external zone with exact KN solution **(E)**, and a thin zone of transition **(R)**, see Fig. 2. Surface $(R)$ is defined by the continuous transition of KN solution to Minkowski interior of the bag, (C. López[10]). According (5) and (6), zone **(R)** corresponds to $H_{KN}(r) = 0$, which gives $r = R = e^2/2m$, and shows that domain wall

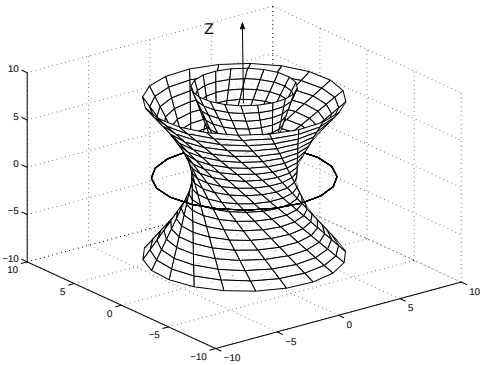



Fig. 3. The lightlike Kerr Null congruence $k^\mu$ determines space-dragging caused by mass and spin.

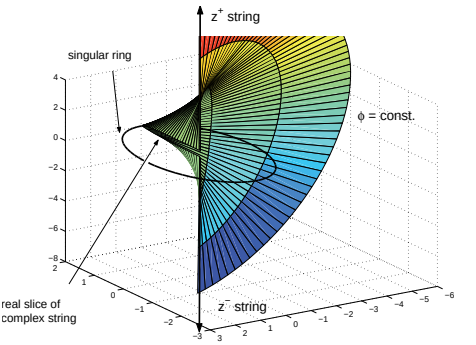



Fig. 4. Rotational dragging of vector-potential $A^\mu$ along border of bag creates a quantum Wilson loop.

boundary has an ellipsoidal form, the disk with radius $r_c = \sqrt{R^2 + a^2}$ and thickness $R$, see Fig. 2.

## 3. Landau-Ginzburg Field Model of Phase Transition

To satisfy **(I),(E),(R)**, it is natural to use Higgs mechanism of symmetry breaking which is used in many nonperturbative electroweak models and in the MIT and SLAC bag models[18,19]. The corresponding Lagrangian is also known as LG field model for superconducting phase transitions and used in famous Nielsen-Olesen (NO) model of vortex string in superconductor[13], $\mathcal{L}_{NO} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\mathcal{D}_\mu\Phi)(\mathcal{D}^\mu\Phi)^* - V(|\Phi|)$, where $\mathcal{D}_\mu = \nabla_\mu + ieA_\mu$ are $U(1)$ covariant derivatives and $F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu}$. Conditions **(I),(E),(R)** require to use generalized LG model with three Higgs-like fields,[20] $(H, Z, \Sigma) \equiv (\Phi_1, \Phi_2, \Phi_3)$. Lagrangian differs from $\mathcal{L}_{NO}$ only by summation over fields $\Phi_i, i = 1, 2, 3$, while potential $V$ is to be built in analogue with supersymmetric machinery from superpotential $W(\Phi_i)$, as follows $V(r) = \sum_i F_i F_i^*, \quad F_i = \partial W/\partial\Phi_i \equiv \partial_i W$, where $W(\Phi_i, \bar{\Phi}_i) = Z(\Sigma\bar{\Sigma} - \eta^2) + (Z + \mu)H\bar{H}$, The conditions $F_i = \partial_i W = 0$ determine two vacuum states with $V = 0$ : internal-**(I)**, $r < R - \delta$, condensation of Higgs field, $|H| = \eta$, and external- **(E)**, where $r > R + \delta$ and $H = 0$, separated by domain wall with $V > 0$, – zone **(R)**, in correspondence with **(I),(E),(R)**.

Higgs field $H$ is concentrated in zone **(I)**, creating superconducting vacuum state inside bag, leading to equations

$$\partial_\nu\partial^\nu A_\mu = J_\mu = e|H|^2(\chi_{,\mu} + eA_\mu), \tag{7}$$

where $\chi$ is oscillating phase of Higgs field $H = |H|e^{i\chi(t,\phi)}$. Vector potential of KN solution is dragged in $t, \phi$ direction by the Kerr congruence near the ring, (see Fig. 2) and concentrated in equatorial plane near the Kerr singular ring, forming a Wilson loop on the boundary of superconducting disk. At $r = R = e^2/2m$ it takes

maximal value $A^*_\mu dx^\mu = -Re\ \frac{2m}{e}(dt - ad\phi)$. Inside the disk it is compensated by phase of Higgs field, $\chi_{,\mu} + eA_\mu = 0$ so that $J_\mu = 0$ inside the disk. Thus, current $J_\mu$ is expelled to surface layer. Formation of the Wilson line on the rim of the disk leads to *quantization of angular momentum*[5,17], $J = n/2,\ n = 1, 2, 3, ...$

## 4. The Supersymmetric Bag Model

By nature bag models are assumed to be soft and flexible[21–23]. Under excitations bags are deformed and take the shape of a stringy flux-tube joining the quark-antiquark pair[18,21], or toroidal string[19,22,23]. In the KN gravitational field, bag takes ellipsoidal shape, see Fig. 2, and acquires circular string at the sharp boundary, closely to Kerr singular ring[17]. This string bears lightlike traveling wave and creates circulating singular pole, reproducing zitterbewegung of the naked electron[17]. Considering fields $\Phi_i$ and $\bar{\Phi}_i$ as independent chiral field, we double the charged chiral fields $\Sigma$ and $\Phi$, and obtain five chiral fields with identification $\Phi_i = (\Phi_+, \Phi_-, \Sigma_+, \Sigma_-, Z)$. It gives supersymmetric LG model with superpotential $W(\Phi_i, \tilde{\Phi}_i) = Z(\Sigma_+\Sigma_- - \eta^2) + (Z + \mu)\Phi_+\Phi_-$, which satisfies conditions **(I)**,**(E)**,**(R)**. Kinetic part of this model coincides with that of the Wess-Zumino SuperQED model, while potential part corresponds to the most general renormalizable supersymmetric generalization of the Wess-Zumino model. In the component form $\Phi_\pm(y) = H_\pm(y^\mu) + \sqrt{2}\theta\psi_\pm(y^\mu) + \theta\theta F_\pm(y^\mu)$, we obtain doublet of the Higgs fields $H_\pm$, corresponding to supersymmetric Higgs model in which two Weyl spinors $\psi^\pm$ are combined into one massive Dirac spinor of the electron – superpartner of the Higgs doublet[24].

## 5. Conclusion

Analysis of the compatible with gravity Kerr-Newman model of the spinning particle (electron) shows that spinning gravity cardinally increases scale of gravitational interaction, extending its influence up to Compton scale, where it starts interplay with quantum theory, forming nonperturbative model of the dressed electron based on the supersymmetric field model of symmetry breaking. Contrary to the assumed confrontation, *gravity collaborates with quantum theory*, fixing the size and shape of the dressed electron corresponding to the fine structure constant. Also, Kerr-Newman gravity fixes the Wilson loop of the vector potential, leading to quantization of angular momentum and defining the related stringy structure.[25]

## References


1. J. Ellis, *Int. J. Mod. Phys.* **A 32** 1746001 (2017).
2. Becker K, Becker M and Schwarz J 2007 *String Theory and M-Theory: A Modern Introduction.*
3. R.P. Kerr, *Phys. Rev. Lett.* **11** 237 (1963).
4. G. C. Debney, R. P. Kerr and A. Schild, *J. Math. Phys.* **10**, 1842 (1969).

5. A. Burinskii, JETP (Zh. Eksp. Teor. Fiz.) 2015. V. 148(8). P. 228. [arXiv:1505.03439].
6. A. Burinskii, *Int. J. of Mod. Phys. D* **26**, 1243022 (2017).
7. B. Carter, *Phys. Rev.* **174**, 1559 (1968).
8. W.Israel, *Phys. Rev. D* **2**, 64 (1969).
9. A. Burinskii, *Sov. Phys. JETP* **39**, 193 (1974).
10. C. A. López, *Phys. Rev. D* **30**, 313 (1984).
11. A. Burinskii, *J. Phys.: Conf. Ser.* **343**, 012019 (2012).
12. D. D. Ivanenko and A. Ya. Burinskii, *Izv. Vuz. Fiz.* **5**, 135 (1975).
13. H. B. Nielsen and P. Olesen, *Nucl. Phys. B* **61** 45 (1973).
14. A. Burinskii, *Phys. Rev. D* **52**, 5826 (1995). [arXiv:hep-th/9504139].
15. A. Dabholkar et al., *Nucl. Phys. B* **474** 85 (1996).
16. B. Punsly, *J. Math. Phys.* **26** 1728 (1985).
17. A. Burinskii, *JETP (Zh. Eksp. Teor. Fiz.)* **148(11)**, 937 (2015). [arXiv:1706.02979].
18. A. Chodos et al., *Phys. Rev. D* **9** 3471 (1974).
19. W. A. Bardeen et al., *Phys. Rev. D* **11**, 1094 (1974).
20. J. R. Morris, *Phys. Rev. D* **53** 2078 (1996). [arXiv:hep-ph/9511293].
21. K. Johnson and C. B. Thorn, *Phys. Rev. D* **13**, 1934 (1976).
22. S-H. H. Tye, *Phys. Rev. D* **13** 3416 (1976).
23. R. C. Giles and S-H. H. Tye, *Phys. Rev. D* **13** 1690 (1976).
24. A. Burinskii, *Phys. Part. Nucl.* **49**(5), 958 (2018). [arXiv:1701.01025].
25. A. Burinskii, Magic of the Kerr spinning gravity: new scale for unification of gravity with particle physics, *Essay 2019 GRF competition* (submitted for publication).

# Study of anisotropic compact stars

Sharif M.* and Sadiq S.

*Department of Mathematics, University of the Punjab, Lahore, 54590, Pakistan*
**E-mail: msharif.math@pu.edu.pk*
*www.pu.edu.pk*

We study the stability of anisotropic fluid configuration using a relation between density and pressure. For this purpose, we formulate the corresponding field equations, generalized Tolman-Oppenheimer-Volkoff equation and mass equation which are necessary to discuss the structure of compact stars and solve them through numerical technique. The stability of the compact object is analyzed through cracking and formulate the force distribution function after perturbing the matter quantities. Further, we plot the resulting force function and examine the occurrence of cracking/overturning.

*Keywords*: Electromagnetic field; Equation of state; Cracking.

## 1. Introduction

A stellar object is worthless if it is not stable towards perturbations in its physical variables. Herrera[1] introduced the notion of cracking as well as overturning to describe the behavior of fluid after its deviation from equilibrium state. The cracking appears when the radial force is positive in the inner regions but negative at outer ones while overturning occurs for the reverse situation. Di Prisco *et al.*[2] found that departure from equilibrium state leads to the cracking in sphere for local anisotropic perturbation. Gonzalez *et al.*[3] observed cracking in isotropic and anisotropic spheres with barotropic equation of state (EoS) using local density perturbations. Herrera *et al.*[4] observed the cracking and overturning for anisotropic polytropes by perturbing energy density as well as local anisotropy of the system.

The inclusion of charge in stellar configurations yields interesting results. Takisa and Maharaj[5] formulated exact solutions of the field equations with polytropic EoS to model charged anisotropic compact object. We have studied charged conformally flat polytropic sphere and checked its viability[6]. Azam *et al.*[7] observed cracking in charged sphere and found that stability regions increase by the increase of charge.

This paper investigates the cracking of charged spherical anisotropic polytropes. The next section is devoted to construct the generalized Tolman Oppenheimer Volkoff (TOV) equation as well as mass equation for charged anisotropic fluid distribution using two cases of polytropic EoS. In the section **3**, we formulate force function by perturbing density and anisotropy and plot the results numerically. Finally, we conclude our results in the last section.

## 2. Charged Anisotropic Matter Distribution

We consider static spherically symmetric spacetime as follows

$$ds^2 = -e^{\nu(r)}dt^2 + e^{\lambda(r)}dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right). \tag{1}$$

The energy-momentum tensor for charged anisotropic fluid is given by

$$T_{\mu\nu}=(\rho+P_{\perp})V_{\mu}V_{\nu}+P_{\perp}g_{\mu\nu}+(P_r-P_{\perp})S_{\mu}S_{\nu}+\frac{1}{4\pi}\left(F_{\mu}{}^{\alpha}F_{\nu\alpha}-\frac{1}{4}F^{\alpha\beta}F_{\alpha\beta}g_{\mu\nu}\right), \tag{2}$$

where $P_r,\ P_{\perp},\ V^{\mu}=e^{-\nu/2}\delta^{\mu}_{0},\ S^{\mu}=e^{-\lambda/2}\delta^{\mu}_{1},\ F_{\mu\nu}=\phi_{\nu,\mu}-\phi_{\mu,\nu}$ and $\phi_{\mu}$ are the radial pressure, tangential pressure, four-velocity, four-vector, the Maxwell field tensor and four potential, respectively. The Maxwell field equation $F^{\mu\nu}_{\ ;\nu}=\mu_0 J^{\mu}$ ($\mu_0$ is the magnetic permeability and $J^{\mu}$ is the four current) for our spacetime yields $q(r)=4\pi\int_0^r\zeta e^{\frac{\lambda}{2}}r^2dr$ indicating total charge inside the sphere. The corresponding Einstein-Maxwell field equations turn out to be

$$8\pi\rho+\frac{q^2}{r^4}=\frac{e^{-\lambda}\lambda'}{r}-\frac{e^{-\lambda}-1}{r^2},\quad 8\pi P_r-\frac{q^2}{r^4}=\frac{e^{-\lambda}\nu'}{r}+\frac{e^{-\lambda}-1}{r^2},$$
$$8\pi P_{\perp}+\frac{q^2}{r^4}=e^{-\lambda}\left(\frac{\nu''}{2}+\frac{\nu'^2}{4}-\frac{\nu'\lambda'}{4}+\frac{\nu'}{2r}-\frac{\lambda'}{2r}\right).$$

The corresponding Misner-Sharp mass leads to[9]

$$m(r)=\frac{r}{2}\left(1-e^{-\lambda}\right)+\frac{q^2}{2r}\ \Rightarrow\ m'=4\pi r^2\rho+\frac{qq'}{r}. \tag{3}$$

The conservation law, $T^{\mu}_{\nu;\mu}=0$, yields the generalized TOV equation as

$$P_r{}'+\frac{4\pi r^4P_r-q^2+rm}{r^3-2r^2m+rq^2}(\rho+P_r)-\frac{2\Delta}{r}-\frac{qq'}{4\pi r^4}=0, \tag{4}$$

where $\Delta=P_{\perp}-P_r$.

The polytropic EoS is a power-law relationship between energy density and pressure. It has two possible cases given as[8]

$$P_r=k\rho_0^{\gamma}=k\rho_0^{1+\frac{1}{n}},\quad \rho-\rho_0=nP_r, \tag{5}$$
$$P_r=k\rho^{\gamma}=k\rho^{1+\frac{1}{n}},\quad \rho\left(1-k\rho_0^{\frac{1}{n}}\right)^n=\rho_0, \tag{6}$$

where $\gamma,\ k,\ n$ and $\rho_0$ denote polytropic exponent, polytropic constant, polytropic index and baryonic density, respectively. For case **1**, we construct generalized TOV equation by taking

$$P_{rc}=\alpha\rho_c,\quad \xi=rA,\quad A^2=\frac{4\pi\rho_c}{\alpha(n+1)},\quad \Phi_0^n=\frac{\rho_0}{\rho_{0c}},\quad m(r)=\frac{4\pi\rho_c\eta(\xi)}{A^3}, \tag{7}$$

where subscript $c$ indicates the value at the center, $\alpha,\ \xi,\ \Phi_0,\ \eta$ are dimensionless variables and $A$ is constant. Using the variables along with Eq. (5) in (4) and (3), we obtain

$$\left\{\frac{1-\frac{2\eta\alpha(n+1)}{\xi}+\frac{4\pi\rho_cq^2}{\xi^2\alpha(n+1)}}{\alpha(n+1)\Phi_0+1-n\alpha}\right\}\left(\xi^2\frac{d\Phi_0}{d\xi}-2\Phi_0^{-n}\frac{\alpha^4(n+1)^2\xi^3\Delta+2\pi\alpha^2\rho_c^2q\frac{dq}{d\xi}}{\alpha^5\rho_c(n+1)^3\xi^2}\right)$$
$$+\alpha\xi^3\Phi_0^{n+1}+\eta-\frac{4\pi\rho_cq^2}{\xi\alpha^2(n+1)^2}=0, \tag{8}$$
$$\xi^2\Phi_0^n(1+n\alpha\Phi_0-n\alpha)+\frac{4\pi\rho_cq}{\xi\alpha^2(n+1)^2}\frac{dq}{d\xi}=\frac{d\eta}{d\xi}. \tag{9}$$

For case **2**, we construct the structure equations by taking $\Phi^n = \rho/\rho_c$. Consequently, the TOV equation as well as mass equation turns out to be

$$\left\{\frac{1-\frac{2\eta\alpha(n+1)}{\xi}+\frac{4\pi\rho_c q^2}{\xi^2\alpha(n+1)}}{1+\alpha\Phi}\right\}\left(\xi^2\frac{d\Phi}{d\xi}-2\Phi^{-n}\frac{\alpha^4(n+1)^2\xi^3\Delta+2\pi\alpha^2\rho_c^2 q\frac{dq}{d\xi}}{\alpha^5\rho_c(n+1)^3\xi^2}\right)$$

$$+\alpha\xi^3\Phi^{n+1}+\eta-\frac{4\pi\rho_c q^2}{\xi\alpha^2(n+1)^2}=0, \tag{10}$$

$$\xi^2\Phi^n+\frac{4\pi\rho_c q}{\xi\alpha^2(n+1)^2}\frac{dq}{d\xi}=\frac{d\eta}{d\xi}. \tag{11}$$

We see that Eqs. (8)-(11) form two systems of differential equations in three unknowns for case **1** and **2**, respectively. In order to reduce one unknown, we take $\Delta=\frac{B(4\pi r^4P_r-q^2+rm)}{r^2-2rm+q^2}(\rho+P_r)$ ($B$ is a constant).

## 3. Cracking of Anisotropic Polytrope

In order to observe cracking, the force function is given as

$$\mathcal{R}=\frac{dP_r}{dr}+\beta\left[\frac{4\pi r^4P_r-q^2+rm}{r^3-2r^2m+rq^2}\right](\rho+P_r)-\frac{qq'}{4\pi r^4}, \tag{12}$$

here $\beta=1-2B$ and $\mathcal{R}$ represents the force distribution function. We perturb matter variables for both cases through a set of parameters $(k,\beta)$ and $(n,\beta)$. We perturb the density and anisotropy via $k\rightarrow\tilde{k}=k+\delta k,\quad \beta\rightarrow\tilde{\beta}=\beta+\delta\beta$, yielding $\tilde{P}_r=\tilde{k}\rho_0^{1+\frac{1}{n}}=\omega P_r,\ \tilde{\rho}=\rho_0+n\omega P_r$ with $\omega=\frac{\tilde{k}}{k}$. Introducing the dimensionless variable $\hat{\tilde{\mathcal{R}}}=\frac{A}{4\pi\rho_c^2}\tilde{\mathcal{R}}$ and using the perturbed parameters alongwith Eq. (7) in (12), we have

$$\hat{\tilde{\mathcal{R}}}=\frac{\tilde{\beta}\Phi_0^n(1-n\alpha+\alpha\omega(n+1)\Phi_0)}{\xi^2}\left\{\frac{\tilde{\eta}+\alpha\omega\xi^3\Phi_0^{n+1}-\frac{4\pi\rho_c q^2}{\xi\alpha^2(n+1)^2}}{1-\frac{2\tilde{\eta}\alpha(n+1)}{\xi}+\frac{4\pi\rho_c q^2}{\xi^2\alpha(n+1)}}\right\}+\omega\Phi_0^n\frac{d\Phi_0}{d\xi}-\frac{qq'}{4\pi r^4}. \tag{13}$$

In equilibrium state, $\hat{\tilde{\mathcal{R}}}(\xi,1+\delta\omega,\beta+\delta\beta,\eta+\delta\eta)=0$. After expansion, we obtain

$$\delta\hat{\mathcal{R}}=\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\omega}|_{\omega=1,\tilde{\beta}=\beta,\tilde{\eta}=\eta}\delta\omega+\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{\beta}}|_{\omega=1,\tilde{\beta}=\beta,\tilde{\eta}=\eta}\delta\beta+\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{\eta}}|_{\omega=1,\tilde{\beta}=\beta,\tilde{\eta}=\eta}\delta\eta. \tag{14}$$

Making use of perturbed parameters in Eq. (9), we obtain

$$\tilde{\eta}=\int_0^{\xi}\left[\frac{(1-n\alpha+n\alpha\omega\Phi_0)}{\hat{\xi}^{-2}\Phi_0^{-n}}+\frac{4\pi\rho_c q\frac{dq}{d\hat{\xi}}}{\alpha^2\hat{\xi}(n+1)^2}\right]d\hat{\xi}\ \Rightarrow\ \delta\eta=\frac{f_1(\xi)\delta\omega}{(n\alpha)^{-1}}=\frac{\int_0^{\xi}\hat{\xi}^2\Phi_0^{n+1}d\hat{\xi}\delta\omega}{(n\alpha)^{-1}}.$$

Evaluating $\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\omega}, \frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\beta}, \frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{\eta}}$ from Eq. (13) and using the variable $x = \frac{\xi}{\bar{A}}$ in (14), we have

$$\delta\hat{\mathcal{R}}_1 = \Phi_0^n \left[ \frac{\frac{d\Phi_0}{dx}}{\bar{A}} + \frac{\alpha\beta}{\bar{A}^2x^2G_1} \left\{ \frac{2(n+1)\Phi_0 + 1 - n}{(\alpha\bar{A}^3x^3\Phi_0^{n+1})^{-1}} + (n+1)\Phi_0 \left\{ \eta - \frac{4\pi\rho_c q^2}{\bar{A}x\alpha^2(n+1)^2} \right\} \right.\right.$$
$$\left.\left. + \frac{G_2}{G_1} \left( 1 + \frac{2\alpha^2(n+1)}{(x^2\bar{A}^2\Phi_0^{n+1})^{-1}} - \frac{4\pi\rho_c q^2 x^{-2}}{\bar{A}^2\alpha(n+1)} \right) n f_1(x) \right\} \right] \delta\omega + \frac{\Phi_0^n G_2 G_3}{x^2\bar{A}^2 G_1}\delta\beta, \tag{15}$$

where

$$G_1 = 1 - \frac{2\alpha(n+1)\eta}{\bar{A}x} + \frac{4\pi\rho_c q^2}{\bar{A}^2x^2\alpha(n+1)},\ G_2 = 1 - n\alpha + \alpha\Phi_0(n+1),$$
$$G_3 = \eta + \alpha x^3\bar{A}^3\Phi_0^{n+1} - \frac{4\pi\rho_c q^2}{x\bar{A}\alpha^2(n+1)^2}.$$

In cracking, $\delta\hat{\tilde{\mathcal{R}}} > 0$ inside the sphere while $\delta\hat{\tilde{\mathcal{R}}} < 0$ for outer regions, so $\delta\hat{\tilde{\mathcal{R}}} = 0$ for some value of $\xi$ implying $\delta\omega = -\frac{\delta\beta}{\Gamma}$ with $\Gamma = \frac{\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\omega} + \frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{\eta}} f_1(\xi)}{\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\beta}}|_{\omega=1,\tilde{\beta}=\beta,\tilde{\eta}=\eta}$.

Now, we perturb fluid variables via $n \to \tilde{n} = n + \delta n,\ \beta \to \tilde{\beta} = \beta + \delta\beta$. In this scheme, Taylor's expansion yields

$$\delta\hat{\mathcal{R}}_2 = \Phi_0^n \left[ \frac{\ln\Phi_0}{\bar{A}} \frac{d\Phi_0}{dx} + \frac{\beta}{\bar{A}^2x^2G_1} \left\{ \ln\Phi_0 G_2 G_3 + \alpha(\Phi_0 - 1)G_3 + G_2 \left( \alpha A^3x^3 \right.\right.\right.$$
$$\left.\left.\left. \times\ \Phi_0^{n+1}\ln\Phi_0 + \frac{8\pi\rho_c q^2}{\bar{A}x\alpha^2(n+1)^3} \right) + \frac{G_2G_3}{G_1} \left( \frac{2\eta\alpha}{\bar{A}x} + \frac{4\pi\rho_c q^2}{\bar{A}^2x^2\alpha(n+1)^2} \right) \right.\right.$$
$$\left.\left. + \frac{G_2}{G_1} \left( G_1 + \frac{2\alpha(n+1)}{\bar{A}x} G_3 \right) f_2(x) \right\} \right] \delta n + \frac{\Phi_0^n}{\bar{A}^2x^2} \frac{G_2G_3}{G_1}\delta\beta, \tag{16}$$

where

$$f_2(x) = \int_0^x \frac{\alpha(\Phi_0^{1-n} - 1) + (1 - n\alpha)\ln\Phi_0}{(\bar{A}^3\hat{x}^2\Phi_0^n)^{-1}} d\hat{x},\ \delta n = \frac{-\delta\beta}{\Gamma = \frac{\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{n}} + \frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{\eta}} f_2(\xi)}{\frac{\partial\hat{\tilde{\mathcal{R}}}}{\partial\tilde{\beta}}}|_{\tilde{n}=n,\tilde{\beta}=\beta,\tilde{\eta}=\eta}}. \tag{17}$$

For case **2**, the perturbed parameters $(k, \beta)$ lead to $\tilde{\rho} = \rho + \delta\rho = \rho + \frac{\partial\tilde{\rho}}{\partial\omega}|_{\omega=1} = \rho + nP_r(1-\omega)$ yielding

$$\delta\hat{\mathcal{R}}_3 = \Phi^n \left[ \frac{d\Phi}{\bar{A}dx} + \frac{\alpha\beta}{\bar{A}^2x^2G_1} \left\{ \Phi(1-n) \left( \eta - \frac{4\pi\rho_c q^2}{x\bar{A}\alpha^2(n+1)^2} \right) + x^3\bar{A}^3\alpha(1-n)\Phi^{n+2} \right.\right.$$
$$\left.\left. - n(1+\alpha\Phi)\left(1 + 2\alpha^2\bar{A}^2x^2(n+1)\Phi^{n+1}\right)\frac{f_3(x)}{G_1} \right\} \right] \delta\omega + \frac{(1+\alpha\Phi)\,G_3\delta\beta}{x^2\bar{A}^2\Phi^{-n}G_1}, \tag{18}$$

where $f_3(x)=\int_0^x \bar{A}^3\hat{x}^2\Phi^{n+1}d\hat{x}$. Similarly, perturbation of $n$ and $\beta$ leads to

$$\delta\hat{\mathcal{R}}_4=\Phi^n\left[\frac{\ln\Phi\frac{d\Phi}{dx}}{\bar{A}}+\frac{\beta G_5}{\bar{A}^2x^2G_1}\left\{\ln\Phi(G_4+\alpha A^3x^3\Phi^{n+1})+\frac{8\pi\rho_cq^2}{\bar{A}x\alpha^2(n+1)^3}+\frac{2G_4}{\bar{A}xG_1}\right.\right.$$
$$\left.\left.\times\left(\eta\alpha+\frac{2\pi\rho_cq^2\alpha^{-1}}{\bar{A}x(n+1)^2}\right)+\left(1+\frac{2\alpha(n+1)G_4}{\bar{A}xG_1}\right)f_4(x)\right\}\right]\delta n+\frac{\Phi^nG_4G_5}{\bar{A}^2x^2G_1}\delta\beta, \tag{19}$$

where

$$G_4=\eta+\alpha x^3\bar{A}^3\Phi^{n+1}-\frac{4\pi\rho_cq^2}{x\bar{A}\alpha^2(n+1)^2},\ G_5=1+\alpha\Phi,\ f_4(x)=\int_0^x\bar{A}^3\hat{x}^2\Phi^n\ln\Phi d\hat{x}.$$

We examine the stability of polytropic models by evaluating $\Phi_0$ ,$\Phi$ and $\eta$ with boundary conditions $\eta(0)=0,\ \Phi_0(0)=1,\ \Phi_0(\xi_\Sigma)=0,\ \Phi(0)=1,\ \Phi(\xi_\Sigma)=0$, and use these results to plot force distribution functions. The graphical behavior of force distribution functions are shown in Fig. 1. The left graph of upper panel shows that $\delta\hat{\mathcal{R}}_1$ is positive in the inner regions and becomes negative for the outer ones for all considered values of $q$ thus ensuring the occurrence of cracking in the corresponding model. The right graph of upper panel shows that there is neither cracking nor overturning in $\delta\hat{\mathcal{R}}_2$ for all values of $q$. The left graph of lower panel indicates that $\delta\hat{\mathcal{R}}_3$ exhibits stable behavior for $q=0.2M_0$ while the strongest overturning appears for other two values of $q$. The plot of $\delta\hat{\mathcal{R}}_4$ for different values of charge shows stable configuration (right plot of lower panel).

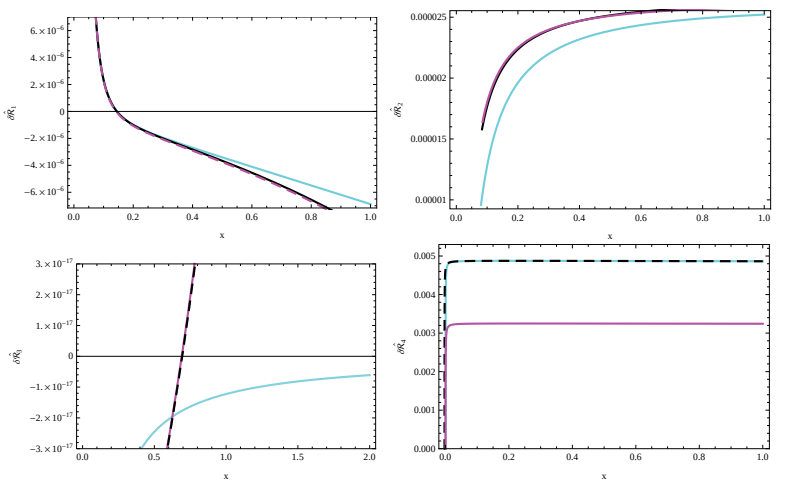


Fig. 1. Plots for force functions versus $x$ with $q=0.2M_0$ (cyan), $q=0.4M_0$ (black), $q=0.64M_0$ (magenta) and $n=1,\alpha=0.85,\beta=1.5,\Gamma=1.6$ for left plot (upper panel), $n=1.5,\alpha=0.90,\beta=0.5,\Gamma=1.4,q=0.2M_0$ for right plot (upper panel), $n=0.5,\alpha=0.98,\beta=1.5,\Gamma=0.6$ for left plot (lower panel) and $n=1.5,\alpha=0.7,\beta=1.5,\Gamma=-0.5$ for right plot (lower panel).

## 4. Concluding Remarks

This paper investigates the occurrence of cracking as well as overturning in charged anisotropic polytropic models. We have constructed the field equations for charged spherical anisotropic fluid obeying polytropic EoS. The TOV as well as mass equations in terms of dimensionless variables are formulated for two cases of polytropic EoS. We have perturbed energy density and pressure anisotropy of a system in two ways. Firstly, we have introduced perturbations through parameters $(k, \beta)$ and constructed the force distribution function $(\delta\hat{\mathcal{R}}_1)$ in terms of perturbed parameters describing total radial forces present in a system. The graphical analysis of $\delta\hat{\mathcal{R}}_1$ indicates the appearance of cracking for all choices of parameters thus leading to unstable configurations for this case. Secondly, we have used $(n, \beta)$ as perturbation parameters and constructed $\delta\hat{\mathcal{R}}_2$. It is found that the resulting models are stable towards perturbations. For case **2**, we have followed the similar procedure and constructed $\delta\hat{\mathcal{R}}_3$ as well as $\delta\hat{\mathcal{R}}_4$. It is found that polytropic models are unstable for perturbation in $(k, \beta)$, while the perturbation of $(n, \beta)$ leads to stable matter configuration representing relativistic polytrope. We have observed that charged matter distribution leads to stable configurations for $(n, \beta)$ perturbation while polytropic models remain unstable by perturbing $(k, \beta)$. The uncharged matter distribution exhibits cracking as well as overturning[4]. However, the inclusion of charge in anisotropic fluid distribution has a dominant effect on polytropes which leads to stable models.

## Acknowledgments

We would like to thank the Higher Education Commission, Islamabad, Pakistan for its financial support through the *Indigenous Ph.D. Fellowship, Phase-II, Batch-III.*

## References

1. L. Herrera, *Phys. Lett. A* **165**, 206 (1992).
2. A. Di Prisco, L. Herrera and V. Varela, *Gen. Relativ. Gravit.* **29**, 1239 (1997).
3. G.A. Gonzalez, A. Navarro and L.A. Núñez, *J. Phys. Conf. Ser.* **600**, 012014 (2015).
4. L. Herrera, E. Fuenmayor and P. León, *Phys. Rev. D* **93**, 024047 (2016).
5. P.M. Takisa and S.D. Maharaj, *Gen. Relativ. Gravit.* **45**, 1951 (2013).
6. M. Sharif and S. Sadiq, *Can. J. Phys.* **93**, 1420 (2015).
7. M. Azam, S.A. Mardan and M.A. Rehman, *Adv. High Energy Phys.* **2015**, 865086 (2015).
8. L. Herrera, A. Di Prisco, W. Barreto and J. Ospino, *Gen. Relativ. Gravit.* **46**, 1827 (2014).
9. C.W. Misner and D.H. Sharp, *Phys. Rev.* **136**, B571 (1964).

# Warped 5D spacetimes, cosmic strings and conformal invariance

Reinoud J. Slagter*
*ASFYON, Astronomisch Fysisch Onderzoek Nederland, Bussum, The Netherlands*
*and*
*Department of Physics, University of Amsterdam, The Netherlands*
**E-mail: info@asfyon.com*
*www.asfyon.com*

We investigate the conformal invariant Lagrangian of the self-gravitating U(1) scalar-gauge field on the time-dependent Bondi-Marder axially symmetric spacetime. By considering the conformal symmetry as exact at the level of the Lagrangian and broken in the vacuum, a consistent model is found with an exact solution of the vacuum Bondi-Marder spacetime, written as $g_{\mu\nu} = \omega^2 \bar{g}_{\mu\nu}$, where $\omega$ is the conformal factor and $\bar{g}_{\mu\nu}$ the 'un-physical' spacetime. Curvature could then be generated from Ricci-flat $\bar{g}_{\mu\nu}$ by suitable dilaton fields and additional gauge freedom. If we try to match this vacuum solution onto the interior vortex solution of the coupled Einstein-scalar-gauge field, we need, besides the matching conditions, constraint equations in order to obtain a topological regular description of the small-scale behaviour of the model. Probably, one needs the five-dimensional warped counterpart model, where the warp factor determines the large-scale behavior of the model. This warp factor is determined by the Einstein field equations for the five-dimensional warped space, where only gravity can propagate into the bulk. The warped five-dimensional model can be reformulated by considering the warp factor as a dilaton field conformally coupled to gravity and embedded in a smooth $M_4 \otimes R$ manifold. It is conjectured that the four-dimensional conformal factor is related to the dilaton field of the five-dimensional counterpart model. The dilaton field (alias warp factor), has a dual meaning. At very early times, when $\omega \to 0$, it describes the small-distance limit, while at later times it is a warp (or scale) factor that determines the dynamical evolution of the universe. However, as expected, the conformal invariance is broken (trace-anomaly) by the appearance of a mass term and a quadratic term in the energy-momentum tensor of the scalar-gauge field, arising from the extrinsic curvature terms of the projected Einstein tensor. These terms can be interpreted as a constraint in order to maintain conformal invariance and the tracelessness of the energy-momentum tensor could then be maintained by a contribution from the bulk. By considering the dilaton field and Higgs field on equal footing on small scales, there will be no singular behavior, when $\omega \to 0$ and one can deduce constraints to maintain regularity of the action. We also present a numerical solution of the model and calculate the (time-dependent) trace-anomaly. The solution depends on the mass ratio of the scalar and gauge fields, the parameters of the model and the vortex charge $n$.

*Keywords*: Warped Spacetimes; Conformal Invariance; Gauged Scalar Field; Gravitational Waves; Axial Symmetry.

## 1. Introduction

There is a urgent necessity for a description of gravity on small scales. Modifications of standard general relativity (GR) seems to be necessary in order to overcome the serious problems which one encounters when one decreases the scale closer to the Planck scale, where quantum effects come into play. This quantum approach has not yet been reconciled with the curved spacetime of GR. An elegant way to modify GR was given by Randall and Sundrum (RS), i.e., the warped five-dimensional spacetime[1–3]. The model can be regarded as the low-energy limit of

general higher-dimensional theories which more fully address the particle interactions. Gravity can propagate freely into the 5D bulk, whereas the standard model fields are constrained to the 4D hypersurface. The hierarchy problem in these models would be solved and the cosmological constant and dark matter can be emergent[4,5]. The characteristic warp factor in the RS-model can also be considered as a dilaton field in conformal invariant (CI) models[6–8].

CI in GR can be a promising formalism for disclosing the small-distance structure of GR. It can be considered as an exact local CI, spontaneously broken as in the case of the BEH mechanism. It is a controversial alternative method to describe canonical quantum gravity, because one is saddled with serious anomalies[9–11]. The key problem is perhaps the handling of asymptotic flatness of isolated systems in GR, specially when they radiate and the generation of the metric $g_{\mu\nu}$ from at least Ricci-flat spacetimes. Close to the Planck scale one should like to have Minkowski spacetime and somehow curvature must emerge. Curved spacetime will inevitable enter the field equations on small scales. The first task is then to construct a Lagrangian, where spacetime and the fields defined on it, are topological regular. This can be done by considering the scale factor (or warp factor in higher-dimensional models) as a dilaton field besides the conformally coupled scalar field. The same procedure can be applied to the 4D spacetime and one can try to generate from (Ricci)-flat spacetimes physical acceptable spacetimes in the non-vacuum case. It is known since the 70s[12], that quantum field theory combined with Einstein's gravity runs into serious problems. The Einstein-Hilbert(EH) action is non-renormalizable and it gives rise to intractable divergences at loop levels. On very small scales, due to quantum corrections to GR, one must modify Einstein's gravity by adding higher order terms in the Lagrangian like $R^2, R_{\mu\nu}R^{\mu\nu}$ or $R_{\mu\nu\sigma\tau}R^{\mu\nu\sigma\tau}$ (or combinations of them). However, serious difficulties arise in these higher-derivative models, for example, the occurrence of massive ghosts which cause unitary problems. A next step is then to disentangle the functional integral over the dilaton field from the ones over the metric fields and matter fields. Moreover, it is desirable that all beta-functions of the matter lagrangian, in combination with the dilaton field, disappear in order to fix all the coupling constants of the model. Further, conformal invariance of the action with matter fields implies that the trace of the energy-momentum tensor is zero. A theory based on a classical "bare" action which is conformally invariant, will lose it in quantum theory as a result of renormalization and the energy-momentum tensor acquires a non-vanishing trace (trace anomaly). We consider here the breaking of conformal invariance in conventional Einstein theory and will not enter into details of these quantum-gravity problems. It is conjectures that conformal symmetry is exact at the level of the Lagrangian and only broken in the vacuum, just as the BEH mechanism in standard model of particle physics. This approach can even be an alternative to supersymmetry and the dark energy problem. Because our axially symmetric model can easily transformed to spherical symmetry, it is clear that our conformal invariant study of the self-gravitating coupled scalar-gauge field on an

axially symmetric spacetime make sense in studying the small scale properties. We describe here an example of CI on an axially symmetric Bondi-Marder spacetime with a U(1) gauged scalar field in the interior.

## 2. The field equations of the model

We consider here the time-dependent axially symmetric spacetime suitable for the description of matter-filled regions of space[13–17]

$$ds^2 = e^{-2\psi}\Big[e^{2\gamma}(d\rho^2 - dt^2) + \rho^2 d\varphi^2\Big] + e^{2\psi+2\mu}dz^2, \tag{1}$$

where $\psi, \gamma$ and $\mu$ are functions of $t$ and $\rho$. The spacetime Eq. (1) is not suitable, because we would have $T_{tt}+T_{\rho\rho}$=0. This is not the case for cosmic strings. Further, we don't make yet a gauge choice, such as the usual one $e^{\mu} = \rho$. For the matter distribution we choose the scalar-gauge field

$$\Phi = \eta X e^{in\varphi}, \qquad A_\mu = [0, 0, 0, \frac{P-n}{e}], \tag{2}$$

with potential $V = \frac{1}{8}\lambda(\Phi\Phi^* - \eta^2)^2$. X and P are functions of $t$ and $\rho$. The vacuum expectation value of the Higgs field is $\eta$, n is the winding number, $\lambda$ the Higgs coupling constant and e the electric charge of a Cooper pair. For a detailed treatment of this $\Phi^4$ gauge model, which is heuristically equivalent with the Ginsburg-Landau theory of superconductivity, we refer to the text book of Vilenkin et al.[18]. The starting point is the action

$$\mathcal{S} = \int d^4x\sqrt{-g}\Big[\frac{1}{2\kappa^2}\Big(R - 2\Lambda) - \frac{1}{2}D_\alpha\Phi(D^\alpha\Phi)^* - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta} - \frac{1}{8}\lambda(\Phi^2 - \eta^2)^2\Big], \tag{3}$$

with $F_{ab} = \partial_a A_b - \partial_b A_a$ the electro-magnetic field tensor. The gauge-covariant derivative is $D_a = \partial_a + ieA_a$. The field equations become (for the moment with $\Lambda = 0$)

$$G_{\mu\nu} = \kappa^2 T_{\mu\nu}, \tag{4}$$

$$D_\mu D^\mu\Phi - 2\frac{\partial V}{\partial\Phi^*} = 0, \tag{5}$$

and

$$\nabla^\nu F_{\mu\nu} - \frac{1}{2}ie\Big[\Phi(D_\mu\Phi)^* - \Phi^*(D_\mu\Phi)\Big] = 0, \tag{6}$$

with $T_{\mu\nu}$ the energy momentum tensor. Written out in components, the equations become

$$\partial_{tt}\psi = \partial_{\rho\rho}\psi + \frac{1}{\rho}\partial_\rho\psi + \partial_\rho\psi\partial_\rho\mu - \partial_t\psi\partial_t\mu - \frac{1}{\rho}\partial_\rho\mu + \kappa^2\Big[e^{2\psi}\frac{\partial_t P^2 - \partial_\rho P^2}{2\rho^2 e^2} - \frac{1}{8}\lambda\eta^4 e^{2\gamma-2\psi}(X^2-1)^2 - \eta^2 e^{2\gamma}\frac{X^2P^2}{\rho^2}\Big], \tag{7}$$

$$\partial_{tt}\mu = \partial_{\rho\rho}\mu + \frac{2}{\rho}\partial_\rho\mu + \partial_\rho\mu^2 - \partial_t\mu^2 + \kappa^2\Big[\frac{1}{4}\lambda\eta^4 e^{2\gamma-2\psi}(X^2-1)^2 + \eta^2 e^{2\gamma}\frac{X^2P^2}{\rho^2}\Big], \tag{8}$$

$$\begin{aligned}\partial_{tt}\gamma = {} & \partial_{\rho\rho}\gamma + \partial_\rho\psi^2 - \partial_t\psi^2 + 2(\partial_\rho\psi\partial_\rho\mu - \partial_t\psi\partial_t\mu) - \frac{2}{\rho}\partial_\rho\mu \\ & + \frac{\kappa^2}{2}\Big[e^{2\psi}\frac{\partial_t P^2 - \partial_\rho P^2}{2\rho^2 e^2} + \eta^2(\partial_\rho X^2 - \partial_t X^2) \\ & - \frac{1}{4}\lambda\eta^4 e^{2\gamma-2\psi}(X^2-1)^2 - 3\eta^2 e^{2\gamma}\frac{X^2P^2}{\rho^2}\Big],\end{aligned} \tag{9}$$

$$\partial_{tt}X = \partial_{\rho\rho}X + \frac{\partial_\rho X}{\rho} + \partial_\rho X\partial_\rho\mu - \partial_t X\partial_t\mu - \frac{e^{2\gamma}XP^2}{\rho^2} - \lambda\eta^2 e^{2\gamma-2\psi}X(X^2-1), \tag{10}$$

$$\partial_{tt}P = \partial_{\rho\rho}P - \frac{\partial_\rho P}{\rho} + \partial_\rho P(2\partial_\rho\psi + \partial_\rho\mu) - \partial_t P(2\partial_t\psi + \partial_t\mu) + e^2\eta^2 e^{2\gamma-2\psi}X^2P. \tag{11}$$

These equation can be solved numerically[6,7] and the well-know solutions of the scalar field and gauge field are expected[19]. There is still a constraint equation from the $(t,r)$-component. In order to match this solution with the exterior vacuum solution, we need the exact solution of the exterior. Marder[14] already pointed out, from a physical point of view, that the concept of a smooth pulse wave of finite duration (apart from a residual "tail") is more acceptable than a wave motion that has to take an infinite time to its present state.

### 2.1. *The exact vacuum solution*

In the vacuum situation one can solve the equations for the metric components. We find

$$\mu = \ln\Big[\frac{\beta_1\sinh(\sqrt{c_1}\rho) + \beta_2\cosh(\sqrt{c_1}\rho)}{\sqrt{c_1}\rho}\Big] + \ln\Big[\frac{\beta_3 e^{2\sqrt{c_1}t} + \beta_4}{2\sqrt{c_1}}\Big] - \sqrt{c_1}t, \tag{12}$$

$$\psi = \mathcal{G}_1(\rho - t) + e^{-\sqrt{c_1}t}\mathcal{G}_2(\rho + t) + \ln(\rho), \tag{13}$$

$$\begin{aligned}\gamma = {} & \mathcal{G}_3(\rho - t) + \mathcal{G}_4(\rho + t) + \ln(\rho) - \frac{1}{4}e^{-2\sqrt{c_1}t}\sin^2(\rho + t) \\ & + 2e^{-\sqrt{c_1}t}\Big[\sin(\rho + t) + \frac{1}{c_1 + 4}(\cos(2t) - \cos(2\rho) + \sqrt{c_1}\sin(2\rho))\Big],\end{aligned} \tag{14}$$

with $\mathcal{G}_i$ arbitrary functions in the arguments and $\beta_i, c_1$ some constants. We took, for the time being, $\beta_1 = \beta_2 = \beta_3 = 1, \beta_4 = 0$ and $\mathcal{G}_1 = \sin(\rho - t)$ and $\mathcal{G}_2 = \cos(\rho + t)$.

## 3. The conformal invariant model

Conformal invariance(CI) is an approved property in string theory by the AdS/CFT correspondence, where a conformal field theory sits on the boundary of the AdS. In CI GR with (quantum) matter fields, the dilaton field plays a fundamental role. In the low-energy limit the dilaton field can act as a dynamical warp factor in 5D

warped spacetimes, while in the small-distance limit the dilaton fields can be used to generate from (Ricci-) flat spacetimes non-flat spacetimes. The Einstein-Hilbert (EH) action can be reformulated by focusing on local conformal symmetry as an exact but spontaneously broken symmetry. Performing a conformal transformation on a spacetime manifold, means that we change our standard measuring rods and clocks. This change is not the same in different points in spacetime. In other words, we multiply the spacetime metric by a kind of scalar field (or dilaton field).

### 3.1. *The interior*

Let us rewrite the metric of Eq. (1) as

$$ds^2 = \omega^2\Big[-dt^2 + d\rho^2 + e^{2\tau}dz^2 + \rho^2 e^{-2\gamma}d\varphi^2\Big], \tag{15}$$

where we wrote $e^{4\psi+2\mu-2\gamma} \equiv e^{2\tau}$ and $\omega^2 \equiv e^{2\gamma-2\psi}$. So we define a metric $\bar{g}_{\mu\nu}$ by $g_{\mu\nu} = \omega^2\bar{g}_{\mu\nu}$. Further, we define $\Phi = \frac{1}{\omega}\bar{\Phi}$. The action

$$\mathcal{S} = \int d^4x\sqrt{-\bar{g}}\Big\{-\frac{1}{12}\Big(\bar{\Phi}\bar{\Phi}^* + \bar{\omega}^2\Big)\bar{R} - \frac{1}{4}F_{\alpha\beta}F^{\alpha\beta}$$
$$-\frac{1}{2}\Big(D_\alpha\bar{\Phi}(D^\alpha\Phi)^* + \partial_\alpha\bar{\omega}\partial^\alpha\bar{\omega}\Big) - V(\bar{\Phi},\bar{\omega}) - \frac{1}{36}\kappa^2\Lambda\bar{\omega}^4\Big\}, \tag{16}$$

is then local conformal invariant under the transformation $g_{\mu\nu} \to \Omega^2 g_{\mu\nu}, \Phi \to \frac{1}{\Omega}\Phi$ and $\omega \to \frac{1}{\Omega}\omega$. We redefined $\bar{\omega}^2 \equiv -\frac{6\omega^2}{\kappa^2}$. Varying the Lagrangian, we obtain the field equations for the metric components, the scalar field, the gauge field and the "dilaton" field[7,8].

$$\bar{G}_{\mu\nu} = \frac{1}{(\bar{\omega}^2 + \bar{\Phi}\bar{\Phi}^*)}\Big(T^{(\bar{\omega})}_{\mu\nu} + T^{(\bar{\Phi},c)}_{\mu\nu} + \bar{T}^{(A)}_{\mu\nu} + \frac{1}{6}\bar{g}_{\mu\nu}\Lambda_{eff}\kappa_4^2\bar{\omega}^4 + \bar{g}_{\mu\nu}V(\bar{\Phi},\bar{\omega})\Big), \tag{17}$$

$$\bar{\nabla}^\alpha\partial_\alpha\bar{\omega} - \frac{1}{6}\bar{R}\bar{\omega} - \frac{\partial V}{\partial\bar{\omega}} - \frac{1}{9}\Lambda\kappa^2\bar{\omega}^3 = 0, \tag{18}$$

$$D^\alpha D_\alpha\bar{\Phi} - \frac{1}{6}\bar{R}\bar{\Phi} - \frac{\partial V}{\partial\bar{\Phi}^*} = 0, \qquad \bar{\nabla}^\nu F_{\mu\nu} = \frac{i}{2}\epsilon\Big(\bar{\Phi}(D_\mu\bar{\Phi})^* - \bar{\Phi}^* D_\mu\bar{\Phi}\Big), \tag{19}$$

with

$$\bar{T}^{(A)}_{\mu\nu} = F_{\mu\alpha}F^\alpha_\nu - \frac{1}{4}\bar{g}_{\mu\nu}F_{\alpha\beta}F^{\alpha\beta}, \tag{20}$$

$$\bar{T}^{(\bar{\Phi},c)}_{\mu\nu} = \Big(\bar{\nabla}_\mu\partial_\nu\bar{\Phi}\bar{\Phi}^* - \bar{g}_{\mu\nu}\bar{\nabla}_\alpha\partial^\alpha\bar{\Phi}\bar{\Phi}^*\Big)$$
$$-3\Big[D_\mu\bar{\Phi}(D_\nu\bar{\Phi})^* + (D_\mu\bar{\Phi})^* D_\nu\bar{\Phi} - \bar{g}_{\mu\nu}D_\alpha\bar{\Phi}(D^\alpha\bar{\Phi})^*\Big] \tag{21}$$

and

$$T^{(\bar{\omega})}_{\mu\nu} = \Big(\bar{\nabla}_\mu\partial_\nu\bar{\omega}^2 - \bar{g}_{\mu\nu}\bar{\nabla}_\alpha\partial^\alpha\bar{\omega}^2\Big) - 6\Big(\partial_\mu\bar{\omega}\partial_\nu\bar{\omega} - \frac{1}{2}\bar{g}_{\mu\nu}\partial_\alpha\bar{\omega}\partial^\alpha\bar{\omega})\Big). \tag{22}$$

Newton's constant reappears in the quadratic interaction term for the scalar field. The trace of Eq. (17) is$\sim \kappa^2\eta^2\lambda\bar{\omega}^2X^2$. It is not possible to isolate the equation for $\bar{\omega}$, as expected.

### 3.2. *The vacuum exterior*

Without matter fields one can solve the field equations exact. Varying the Einstein-Hilbert action

$$\mathcal{L}^{EH\omega} = \frac{\sqrt{-\bar{g}}}{2\kappa^2}\Big(\omega^2\bar{R} + 6\bar{g}^{\mu\nu}\partial_\mu\omega\partial_\nu\omega\Big), \tag{23}$$

with respect to $\bar{g}_{\mu\nu}$ and $\omega$, one obtains

$$\omega^2\bar{G}_{\mu\nu} = \bar{\nabla}_\mu\bar{\nabla}_\nu\omega^2 - \bar{g}_{\mu\nu}\bar{\nabla}_\alpha\bar{\nabla}^\alpha\omega^2 - 6\partial_\mu\omega\partial_\nu\omega + 3\bar{g}_{\mu\nu}\partial_\alpha\omega\partial^\alpha\omega \equiv T^{(\omega)}_{\mu\nu}, \tag{24}$$

$$\bar{\nabla}^\mu\partial_\mu\omega - \frac{1}{6}\omega\bar{R} = 0. \tag{25}$$

One can easily verify that $\mathbf{TR}[\bar{G}_{\mu\nu} - \frac{1}{\omega^2}T^{(\omega)}_{\mu\nu}] = 0$. So the trace of any matter field contribution must be zero. From the Einstein equations, Eq. (24) and the $\omega$ equation, Eq. (25), we can separate the equation for $\omega$

$$\partial_{tt}\omega = \partial_{\rho\rho}\omega + \frac{1}{\omega}(\partial_t\omega^2 - \partial_r\omega^2) + \upsilon\omega. \tag{26}$$

The solution is (we use the subscript E for exterior)

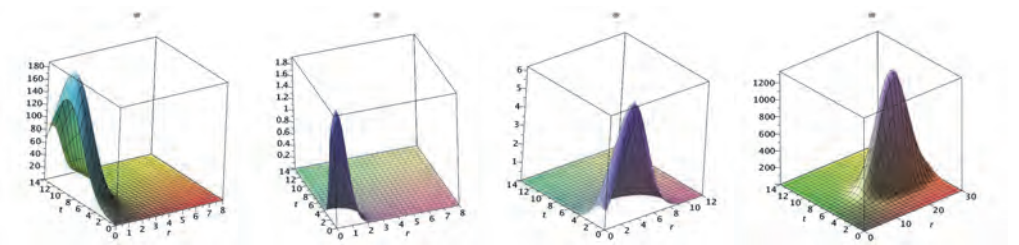


Fig. 1. Some example plots of $\omega_E$ of Eq. (27) for different values of the constants $c_i$ and $\upsilon$.

$$\omega_E = \mathcal{A}e^{\frac{1}{2}\zeta_1(\rho^2+t^2)-\frac{1}{2}\upsilon\rho^2+\zeta_2 t+r}, \tag{27}$$

with $\zeta_i$ and $\mathcal{A}$ integration constants. The constant $\upsilon$ enters the equation by the separation of variables in the Einstein equations. In Fig. 1 we plotted some typical solutions. Finally one finds the exterior metric

$$\begin{aligned} ds^2 = \mathcal{A}^2\Big\{ & e^{\zeta_1(\rho^2+t^2)-\upsilon\rho^2+2\zeta_2 t+2\rho}(-dt^2+d\rho^2) \\ & + \Big[\frac{(e^{\sqrt{\vartheta_1}t}+\delta_1 e^{-\sqrt{\vartheta_1}t})^2(e^{\sqrt{\vartheta_2}\rho}+\delta_2 e^{-\sqrt{\vartheta_2}\rho})^2}{\vartheta_1\vartheta_2}\Big]dz^2 \\ & + \Big[\frac{(e^{\sqrt{\vartheta_1^*}t}+\delta_3 e^{-\sqrt{\vartheta_1^*}t})^2(e^{\sqrt{\vartheta_2^*}\rho}+\delta_4 e^{-\sqrt{\vartheta_2^*}\rho})^2}{\vartheta_1^*\vartheta_2^*}\Big]d\varphi^2\Big\}, \end{aligned} \tag{28}$$

where $\vartheta_1 \equiv (\zeta_2^2 + c_1 + \zeta_1), \vartheta_2 \equiv (\zeta_2^2 + \upsilon + c_1 + \zeta_1), \vartheta_1^* \equiv (\zeta_2^2 - c_2 + \zeta_1), \vartheta_2^* \equiv (\zeta_2^2 + \upsilon - c_2 + \zeta_2)$. $c_i$ are integration constants. One can generate exact ($\Omega$-)conformal

equivalent spacetimes from $\bar{g}_{\mu\nu}$. One can easily verify that the Bianchi identities are fulfilled. Constraint equations can found for $\partial_t\omega$. Ricci-flat solutions will restrict the values for the several parameters. At the boundary, we have now exact values of the metric components, which can be used for the interior solution. The two line elements will be isometric at the boundary surface for the well-known junction conditions[20].

## 4. Connection with the 5D warped spacetime

In the 5D warped counterpart model[6] on the spacetime

$$ds^2 = \mathcal{W}(t,r,y)^2\Big[e^{2\gamma(t,r)-2\psi(t,r)}(-dt^2+dr^2)+e^{2\psi(t,r)}dz^2+\frac{r^2}{e^{2\psi(t,r)}}d\varphi^2\Big]+dy^2, \tag{29}$$

we found a typical solution for the warp factor

$$\mathcal{W}^2 = \frac{e^{-\frac{1}{6}\Lambda_5(y-y_0)^2}}{\tau r}\Big(d_1e^{(\sqrt{2\tau})t}-d_2e^{-(\sqrt{2\tau})t}\Big)\Big(d_3e^{(\sqrt{2\tau})r}-d_4e^{-(\sqrt{2\tau})r}\Big). \tag{30}$$

If we define

$${}^{(5)}g_{\mu\nu} = W_1(t,r)^2W_2(y)^2\tilde{g}_{\mu\nu}+n_\mu n_\nu\Gamma(y)^2, \tag{31}$$

then $W_1$ could be identified as dilaton field or "warp factor" coming from the bulk, while $W_2$ is equals the well-know warp factor in the RS brane world model[1,2]. The metric $\tilde{g}_{\mu\nu}$ is equals our $\tilde{g}_{\mu\nu}=\omega^2\bar{g}_{\mu\nu}$ of Eq. (15). In the former study[7], a numerical solution was found in the coupled Einstein-scalar-gauge field model, and the trace anomaly could possibly be solve due to contributions from the bulk. The solution for $W_1$ follows solely from the 5D Einstein equations with empty bulk (eventually with a cosmological constant $\Lambda_5$). The question is, can we in some way relate $W_1$ to the dilaton field $\bar{\omega}$ of section 3-1? After all, information about an extra dimension is visible as a curvature in a spacetime with one fewer dimension. The $W_1$-field need to be shifted to a complex contour to ensure that $W_1$ has the same unitary and positivity properties as the scalar field. One can easily check that $W_1$ has complex solutions. Further, the effective 4D Einstein equations of the warped model contain extra terms arising from the extrinsic curvature terms in the the projected Einstein tensor and the projected Weyl tensor[3]. When approaching smaller scales in the model and $W_1\to 0$, no singular behavior occurs due to the fact that the scalar field is present in the Lagrangian, which could be handled on the same footing as the dilaton field. For later times, on cosmological scale, $W_1$ behaves as a "standard" warp factor. The $\Omega$ transformation is necessary as a conformal "gauge" in order to make a renormalizable model. As 't Hooft stated[9], "We get a renormalizable gauge if we decide to choose our conformal factor $\Omega$ in such a way that the amount of activity in a given spacetime volume is fixed or at least bounded". It is then conjectured[9] that close to the Planck scale the conformal invariance is spontaneously broken due to the contributions of additional terms like $(R_{\mu\nu}R^{\mu\nu}-\frac{1}{3}R^2)$ and $\kappa^2\eta^2\lambda\bar{\omega}^2X^2$ in the action.

## 5. Conclusions

We investigated the coupled Einstein-scalar-gauge fields using a conformally invariant Lagrangian. We compare the exterior exact vacuum solution with the spacetime of an infinite line-mass and semi-infinite line-mass. We find a simple example of conformally equivalent (Ricci-) flat spacetimes $\omega^2 \bar{g}_{\mu\nu}$. This solution must then be matched on the non-vacuum solution of the cosmic string at the core. In order to obtain tracelessness of the energy momentum tensor, additional constraints are necessary. It is conjectured that a contribution from the bulk in the 5D counterpart model is needed to make the energy momentum tensor traceless. The conformal symmetry, if exact, can be broken spontaneously. This means that we need additional field transformations on the vacuum spacetime.

## References


1. L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 3370 (1999).
2. L. Randall and R. Sundrum, *Phys. Rev. Lett.* **83**, 4690 (1999).
3. T. Shiromizu, K. Maeda and M. Sasaki, *Phys. Rev. D* **62**, 024012 (2000).
4. P. D. Mannheim and D. Kazanas, *Astrph. Journ.* **341**, 635 (1989).
5. P. D. Mannheim, `arXiv:hep-th/161008907v2`, 2017.
6. R. J. Slagter and S. Pan, *Found. of Phys.* **46**, 1075 (2016).
7. R. J. Slagter, `arXiv:gr-qc/1711.08193`, 2017.
8. R. J. Slagter, `arXiv:gr-qc/1810.08793`, 2018.
9. G. 't Hooft, *Found. of Phys.* **41**, 1829 (2011).
10. G. 't Hooft, `arXiv:gr-qc/10110061v1`, 2010.
11. G. 't Hooft, `arXiv:gr-qc/151104427v1`, 2015.
12. L. E. Parker and D. J. Toms, *Quantum Field Theory in Curved Spacetime*, (Cambridge University Press, Cambridge, 2009).
13. H. Bondi, *Proc. R. Soc. Lond. A* **427**, 259 (1989).
14. L. Marder, *Proc. R. Soc. A* **244**, 524 (1958).
15. W. B. Bonner, (1991) *Gen. Rel. Gravit.* **24**, 551 (1991).
16. W. B. Bonner, J. B. Griffiths and M. A. H. MacCallum, *Gen. Rel. Gravit.* **26**, 687 (1994).
17. W. B. Bonner and M. A. P. Martins, *Class. Quant. Grav.* **8**, 727 (1991).
18. A. Vilenkin and E. P. S. Shellard, *Cosmic strings and other topological defects*, (Cambridge univiversity press: Cambrigde, 1994).
19. D. Garfinkle, *Phys. Rev. D.* **32**, 1323 (1985).
20. M. R. Anderson, *The Mathematical Theory of Cosmic Strings*, (IoP publishing, Bistol, 2003).

# Axisymmetric solution of Einstein's field equations with quadrupole

S. Toktarbay[1], M. E. Abishev[1], H. Quevedo[1,2,3], A. Mansurova[1] and A. Muratkhan[1]

[1] *Department of Theoretical and Nuclear Physics, Al-Farabi Kazakh National University, Almaty 050040, Kazakhstan*

[2] *Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México, AP 70543, México, DF 04510, Mexico*

[3] *Dipartimento di Fisica and ICRA, Università di Roma "La Sapienza", I-00185 Roma, Italy, City, State ZIP/Zone, Country*

[*] *E-mail: saken.yan@yandex.com, quevedo@nucleares.unam.mx*

*https://www.kaznu.kz/en*

We present a static axisymmetric solution of Einstein's field equations. Assuming that the deviation from spherically symmetry is small, we linearize the corresponding field equations with the perfect fluid source and up to the first order in quadrupole.

*Keywords*: Quadrupole moment; linearized field equations; q-metric; perfect fluid.

## 1. Introduction

Einstein's theory of General relativity (GR) is one of the best candidate to describe the gravitational field of astrophysical compact objects. According to the GR all the information about the gravitational field should be contained in the metric tensor which must be a solution of Einstein's equations. It is well known that the exterior field of an arbitrarily rotating mass can be described by the Kerr spacetime[1]. The kerr solution has not an internal counterpart, this is a major problem in classical general relativity[2]. In view of this situation, We consider alternative approaches that contains the quadrupole moment. The quadrupole moment is responsible for the deformation of any realistic mass distribution.

In a recent work[3], it was proposed to use the Zipoy-Voorhees transformation[4,5] to generate the quadrupolar metric ($q$-metric), which can be interpreted as the simplest generalization of the Schwarzschild metric. Consequently, this metric can be used to describe the exterior gravitational field of deformed distributions of mass. The question arises whether it is possible to find an interior metric that can be matched to the exterior one in such a way that the entire spacetime is described as a whole. One usually assumes that the matter inside the object can be described by a perfect fluid with two physical parameters, namely, energy density and pressure. In this work, we will concentrate on the problem of perfect-fluid solutions, without considering the rotation parameter.

This work is organized as follows. In Sec. 2, we consider the $q$-metric as describing the exterior gravitational field of a deformed source with mass and quadrupole moment. In Sec. 3, we present the exact equations for a perfect fluid source. In Sec. 4, we construct the approximate line element with a quadrupole moment and we present some particular vacuum solutions. Finally, Sec. 5 contains discussions of our results.

## 2. Exterior $q$-metric

If we start from the Schwarzschild solution and apply a Zipoy-Voorhees transformation[4 5] the new line element can be written as

$$ds^2 = h^{1+q}dt^2 - h^{-q}\left[\left(1+\frac{m^2\sin^2\theta}{r^2 h}\right)^{-q(2+q)}\left(\frac{dr^2}{h}+r^2 d\theta^2\right)+r^2\sin^2\theta d\varphi^2\right]. \tag{1}$$

where $h = 1 - 2m/r$.

In[3] it was shown that this is the simplest generalization of the Schwarzschild solution that contains the additional parameter $q$, which describes the deformation of the mass distribution. In the literature, this metric is known as the Zipoy-Voorhees metric, delta-metric, gamma-metric and $q$-metric[7]. A stationary generalization of the $q$-metric has been obtained in[8]. The physical interpretation of the parameters $m$ and $q$ can be clarified by calculating the invariant Geroch multipoles[9].

A study of the curvature of the $q$-metric shows that the outermost singularity is located at $r = 2m$, a hypersurface which in all known compact objects is situated inside the surface of the body. This implies that in order to describe the entire gravitational field, it is necessary to cover this type of singularity with an interior solution.

## 3. Interior metric

In this work, we will concentrate on the case of static perfect fluid spacetimes. Our experience with numerical perfect fluid solutions[11] indicates that for the case under consideration the line element

$$ds^2 = f dt^2 - \frac{e^{2\gamma}}{f}\left(\frac{dr^2}{h}+d\theta^2\right) - \frac{\mu^2}{f}d\varphi^2 \ , \tag{2}$$

is particularly convenient. Here $f = f(r,\theta)$, $\gamma = \gamma(r,\theta)$, $\mu = \mu(r,\theta)$, and $h = h(r)$. A redefinition of the coordinate $r$ leads to an equivalent line element which has been used to investigate anisotropic static fluids[12].

The Einstein equations for a perfect fluid with 4-velocity $U_\alpha$, density $\rho$, and pressure $p$ (we use geometric units with $G = c = 1$)

$$R_{\alpha\beta} - \frac{1}{2}R g_{\alpha\beta} = 8\pi\left[(\rho+p)U_\alpha U_\beta - p g_{\alpha\beta}\right]. \tag{3}$$

The field equations can be represented as two second-order differential equations for $\mu$ and $f$. Moreover, the function $\gamma$ is determined by a set of two partial differential equations which can be integrated by quadratures once $f$, $\mu$, $p$, and $h$ are known.

In[13] we discussed properties of the field equations and presented a new method for generating perfect fluid solutions of the Einstein equations, starting from a given seed solution.

## 4. Approximate metric with quadrupole

For understanding influences of quadrupole moment to the structure of spacetimes we consider the case of a slightly deformed mass. Our approach consists in postulating the interior line element and evaluating the energy-momentum tensor from the Einstein equations, a method which was first proposed by Synge and has been applied to find several approximate interior solutions [14,15].

To find the corresponding interior line element, we proceed as follows. Consider the case of a slightly deformed mass. This means that the parameter $q$ for the exterior $q$-metric can be considered as small and we can linearize the line element as

$$ds^2 = h\,(1+q\ln h)\,dt^2 - r^2\,(1-q\ln h)\sin^2\theta d\varphi^2 - \left[1+q\ln h - 2q\ln\left(h+\frac{m^2}{r^2}\sin^2\theta\right)\right]\left(\frac{dr^2}{h}+r^2d\theta^2\right)\ . \tag{4}$$

We will assume that the exterior gravitational field of the compact object is described to the first order in $q$ by the line element (4), which represents a particular approximate solution to Einstein's equations in vacuum.

To construct the approximate interior line element, we start from the exact line element (2) and use the approximate solution (4) as a guide. Following this procedure, an appropriate interior line element can be expressed as

$$ds^2 = e^{2\nu}(1+qa)dt^2 - \big[1+q(c+b)\big]\frac{dr^2}{1-\frac{2\tilde{m}}{r}} - \big[1+q(a+b)\big]r^2d\theta^2 - (1-qa)r^2\sin^2\theta d\varphi^2\ , \tag{5}$$

where the functions $\nu=\nu(r)$, $a=a(r)$, $c=c(r)$, $\tilde{m}=\tilde{m}(r)$, and $b=b(r,\theta)$. Although there exists quite a large number of approximate solutions such as [16 17 18] and in [14 19] it was shown that the different approaches under diverse circumstances turn out to be equivalent from a mathematical point of view.

### 4.1. *Vacuum solution*

To test the consistency of the linearized approach, we will derive explicitly the approximate vacuum $q$-metric (4). The details of this derivation will be presented elsewhere. The final expression for the metric fucntions

$$\nu = \frac{1}{2}\ln\left(1-\frac{2m}{r}\right)+\alpha_1\ , \quad a-c=\frac{\alpha_2 m^2}{(r-m)^2}\ , \tag{6}$$

where $\alpha_1$ and $\alpha_2$ are dimensionless integration constants. The remaining system of partial differential equations can be integrated in general and yields

$$a = -\frac{\alpha_2 m}{r-m}+\frac{1}{2}\,(\alpha_3-\alpha_2)\ln\left(1-\frac{2m}{r}\right)+\alpha_4\ , \tag{7}$$

$$c = -\frac{\alpha_2 mr}{(r-m)^2}+\frac{1}{2}\,(\alpha_3-\alpha_2)\ln\left(1-\frac{2m}{r}\right)+\alpha_4\ , \tag{8}$$

$$b = \frac{2\alpha_2 m}{r-m} - (\alpha_3 - \alpha_2)\left[\ln 2 + \ln\left(1 - \frac{2m}{r} + \frac{m^2 \sin^2\theta}{r^2}\right)\right] + \alpha_5 \ , \tag{9}$$

where we have replaced $\tilde{m} = m = const.$ in the remaining equations. $\alpha_3$, $\alpha_4$ and $\alpha_5$ are dimensionless integration constants.

Thus, we see that the general approximate exterior solution with quadrupole moment is represented by a 5-parameter family of solutions. The particular case

$$\alpha_1 = 0 \ , \quad \alpha_2 = 0 \ , \quad \alpha_3 = 2 \ , \quad \alpha_4 = 0 \ , \quad \alpha_5 = 2\ln 2 \ , \tag{10}$$

corresponds to the linearized $q$-metric as represented in Eq. (4). Another interesting particular case corresponds to the choice

$$\alpha_1 = 0 \ , \quad \alpha_3 = \alpha_2 \ , \quad \alpha_4 = 0 \ , \quad \alpha_5 = 0 \ , \tag{11}$$

which leads to the following line element

$$\begin{aligned} ds^2 = &\left(1 - \frac{2m}{r}\right)\left(1 - \frac{q\alpha_2 m}{r-m}\right) dt^2 - \left(1 + \frac{q\alpha_2 m}{r-m}\right) r^2 \sin^2\theta d\varphi^2 \\ &- \left[1 + \frac{q\alpha_2 m(r-2m)}{(r-m)^2}\right] \frac{dr^2}{1 - \frac{2m}{r}} - \left(1 + \frac{q\alpha_2 m}{r-m}\right) r^2 d\theta^2 \ . \end{aligned} \tag{12}$$

This is an asymptotically flat approximate solution with parameters $m$, $q$ and $\alpha_2$. The singularity structure can be found by analyzing the Kretschmann invariant $K = R_{\alpha\beta\gamma\delta}R^{\alpha\beta\gamma\delta}$ which in this case reduces to

$$K = \frac{48m^2}{r^6}\left(1 + q\alpha_2 \frac{r-4m}{r-m} + \mathcal{O}(q^2)\right) \ , \tag{13}$$

where the term proportional to $q^2$ has been neglected due to the approximate character of the solution. We see that there is a central singularity at $r = 0$ and a second one at $r = m$. We conclude that the solution (12) describes the exterior field of two naked singularities of mass $m$ and quadrupole $q$. The parameter $\alpha_2$ can be absorbed by redefining the constant $q$ and so it has no special physical meaning. In the general solution (7)–(9), the additive constants $\alpha_4$ and $\alpha_5$ can be chosen such that at infinity the solution describes the Minkowski spacetime in spherical coordinates. This means that non asymptotically flat solutions are also contained in the 5-parameter family (7)–(9). This is the most general vacuum solution which is linear in the quadrupole moment. To our knowledge, this general solution is new.

### 4.2. *Perfect fluid case*

We now apply the approximate line element (5) to the study of perfect fluid solutions. First, we note that in this case the conservation law reduces to

$$p_{,r} = -(\rho + p)\nu_{,r} \ , \qquad p_{,\theta} = 0 \ . \tag{14}$$

Calculating the second derivative $p_{,r\theta} = 0$, the above conservation laws lead to

$$\rho_{,\theta} = 0 \ , \tag{15}$$

implying that the perfect fluid variables can depend on the coordinate $r$ only. Notice that this does not imply that the source is spherically symmetric. In fact, due to the presence of the quadrupole parameter $q$ in the line element (5), the coordinate $r$ is no longer a radial coordinate and the equation $r =$constant represents, in general, a non-spherically symmetric deformed surface[20].

The corresponding linearized Einstein equations can be represented as

$$\overset{(0)}{G}{}^{\nu}_{\mu} + q\overset{(q)}{G}{}^{\nu}_{\mu} = 8\pi \left( \overset{(0)}{T}{}^{\nu}_{\mu} + q\overset{(q)}{T}{}^{\nu}_{\mu} \right), \tag{16}$$

where the (0)-terms correspond to the limiting case of spherical symmetry. As for the energy-momentum tensor, we assume that density and pressure can also be linearized as

$$p(r) = p_0(r) + q\ p_1(r), \quad \rho(r) = \rho_0(r) + q\rho_1(r), \tag{17}$$

in accordance with the conservation law conditions (14) and (15). Here, $p_0(r)$ and $\rho_0(r)$ are the pressure and density of the background spherically symmetric solution. If we now compute the linearized field equations (16) for the line element (5), we arrive at a set of nine differential equations for the functions $\nu,\ \tilde{m},\ a,\ b,\ c$, and $p_1$. After lengthy computations, it is then possible to isolate an equation that relates $p_1(r)$ and $b(r,\theta)$ from which it follows that

$$b_{,\theta} = 0\ . \tag{18}$$

This means that for the particular approximate line element (5), the field equations for a perfect fluid do not allow the metric functions to explicitly depend on the angular coordinate $\theta$. To search for concrete solutions which can be matched with an exterior metric with quadrupole, it is necessary to modify the exterior metric accordingly. We will see that this approximate exterior solution can be used together with the interior line element (5) to search for approximate solutions with a perfect fluid source. Taking into account that the conservation laws and the approximate field equations for a perfect fluid imply that the physical quantities $p$ and $\rho$ and the metric function $b$ depend only on the spatial coordinate $r$.

## 5. Conclusions

In this work, we have investigated approximate interior and exterior solutions of Einstein's equations in the case of static and axially symmetric perfect fluid spacetimes, which characterized by an arbitrary mass and a small quadrupole moment.

We analyze the corresponding linearized field equations and derive several classes of new vacuum and perfect fluid solutions.

We limit ourselves to the study of interior solutions that can be matched with the exterior $q$-metric. We only considered the simple and idealized case of a static mass distribution with a small quadrupole.

The resulting system of differential equations for the functions of the interior metric indicates that one can try to find analytical solutions, at least in the case of a slightly deformed mass distribution. To study more realistic configurations, it is necessary to take into account the rotation the exact quadrupole of the mass. We expect to investigate these problems in future works.

**Acknowledgments**

This work was partially supported by UNAM-DGAPA-PAPIIT, Grant No. 111617, and by the Ministry of Education and Science of RK, Grant No. AP05133630.

**References**

1. R. P. Kerr, *Phys. Rev* **11** 237–238 (1963).
2. H. Stephani, D. Kramer, M. MacCallum, C. Hoenselaers and E. Herlt, *Exact Solutions of Einstein's Field Equations* (Cambridge, UK: Cambridge University Press 2003).
3. H. Quevedo, *Int. J. Mod. Phys.* **20** 1779–1787 (2011).
4. D.M. Zipoy, *J. Math. Phys.* **7**, 1137–1143 (1966).
5. B. Voorhees, *Phys. Rev. D* **2**, 2119 (1970).
6. D. Malafarina, Conf. Proc. C0405132, 273 (2004).
7. A. Allahyari, H. Firouzjahi and B. Mashhoon (2018), https://arxiv.org/abs/1812.03376.
8. S. Toktarbay and H. Quevedo, *Gravit. Cosmol.* **20** 252–254 (2014).
9. R. Geroch, *J. Math. Phys.* **11** 1955–1961 (1970).
10. R. Geroch, *J. Math. Phys.* **11** 2580–2588 (1970).
11. H. Quevedo, Multipolar Solutions "*Proceedings of the XIV Brazilian School of Cosmology and Gravitation*" (1201.1608) (2012).
12. L. Herrera, A. Di Prisco, J. Ibáñez and J. Ospino, *Phys. Rev. D* **87** 024014 (2013).
13. H. Quevedo and S. Toktarbay, *J. Math. Phys.*, 1503.05300 (2015).
14. K. Boshkayev, H. Quevedo and R. Ruffini, *Phys. Rev. D* **86** 064043 (2012).
15. D. Malafarina, G. Magli and L. Herrera, *General Relativity and Gravitational Physics* (*American Institute of Physics Conference Series* vol 751), pp. 185–187 (2005).
16. J. B. Hartle, Astrophys. J. 150, 1005 (1967).
17. J. B. Hartle and K. S. Thorne, Astrophys. J. 150, 1005 (1968).
18. D. M. Sedrakyan and E. V. Chubaryan, Astrophysica 4, 239 (1968).
19. K. Boshkayev, H. Quevedo, S. Toktarbay et. al., *Gravit. Cosmol.* **22** 305 (2016).
20. Y. B. Zeldovich and I. D. Novikov, *Relativistic Astrophysics. Vol. 1: Stars and Relativity* (Chicago: University of Chicago Press 1971).

PART C

PROCEEDINGS OF THE FIFTEENTH

# MARCEL GROSSMANN MEETING ON GENERAL RELATIVITY

Editors

**Elia S. Battistelli**

**Robert T. Jantzen**

**Remo Ruffini**

World Scientific

# THE FIFTEENTH
# MARCEL GROSSMANN MEETING

**On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories**

This page intentionally left blank

PART C

# THE FIFTEENTH
# MARCEL GROSSMANN MEETING

## On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories

Proceedings of the MG15 Meeting
on General Relativity
University of Rome "La Sapienza", Italy 1–7 July 2018


Editors

**Elia S. Battistelli**
University of Rome "La Sapienza"
Rome, Italy

**Robert T. Jantzen**
Villanova University
Villanova, PA, USA

Series Editor

**Remo Ruffini**
International Center for Relativistic Astrophysics (ICRA)
University of Rome "La Sapienza"
Rome, Italy
International Center for Relativistic Astrophysics
Network (ICRANet)
Pescara, Italy



World Scientific

NEW JERSEY • LONDON • SINGAPORE • BEIJING • SHANGHAI • HONG KONG • TAIPEI • CHENNAI

*Published by*

World Scientific Publishing Co. Pte. Ltd.

5 Toh Tuck Link, Singapore 596224

*USA office:* 27 Warren Street, Suite 401-402, Hackensack, NJ 07601

*UK office:* 57 Shelton Street, Covent Garden, London WC2H 9HE

Library of Congress Control Number: 2022906140

**British Library Cataloguing-in-Publication Data**
A catalogue record for this book is available from the British Library.

**THE FIFTEENTH MARCEL GROSSMANN MEETING**
**On Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories**
**(In 3 Volumes)**



ISBN 978-981-125-824-4 (set_hardcover)
ISBN 978-981-125-825-1 (set_ebook for institutions)
ISBN 978-981-125-826-8 (set_ebook for individuals)

ISBN 978-981-125-661-5 (vol. 1_hardcover)

ISBN 978-981-125-662-2 (vol. 2_hardcover)

ISBN 978-981-125-663-9 (vol. 3_hardcover)

For any available supplementary material, please visit
https://www.worldscientific.com/worldscibooks/10.1142/12843#t=suppl

Desk Editor: Ng Kah Fee

Typeset by Stallion Press
Email: enquiries@stallionpress.com

Printed in Singapore

**THE MARCEL GROSSMANN MEETINGS**

**Series Editor:** REMO RUFFINI

**Publications in the Series of Proceedings**

Proceedings of the Fifteenth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2018)
Edited by E.S. Battistelli, R.T. Jantzen, R. Ruffini
World Scientific, 2021

Proceedings of the Fourteenth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2015)
Edited by M. Bianchi, R.T. Jantzen, R. Ruffini
World Scientific, 2017

Proceedings of the Thirteenth Marcel Grossmann Meeting on General Relativity
(Stockholm, Sweden, 2012)
Edited by K. Rosquist, R.T. Jantzen, R. Ruffini
World Scientific, 2015

Proceedings of the Twelfth Marcel Grossmann Meeting on General Relativity
(Paris, France, 2009)
Edited by T. Damour, R.T. Jantzen, R. Ruffini
World Scientific, 2012

Proceedings of the Eleventh Marcel Grossmann Meeting on General Relativity
(Berlin, Germany, 2006)
Edited by H. Kleinert, R.T. Jantzen, R. Ruffini
World Scientific, 2007

Proceedings of the Tenth Marcel Grossmann Meeting on General Relativity
(Rio de Janiero, Brazil, 2003)
Edited by M. Novello, S. Perez-Bergliaffa, R. Ruffini
World Scientific, 2005

Proceedings of the Ninth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 2000)
Edited by V.G. Gurzadyan, R.T. Jantzen, R. Ruffini
World Scientific, 2002

Proceedings of the Eighth Marcel Grossmann Meeting on General Relativity
(Jerusalem, Israel, 1997)
Edited by T. Piran
World Scientific, 1998

Proceedings of the Seventh Marcel Grossmann Meeting on General Relativity
(Stanford, USA, 1994)
Edited by R.T. Jantzen and G.M. Keiser
World Scientific, 1996

Proceedings of the Sixth Marcel Grossmann Meeting on General Relativity
(Kyoto, Japan, 1991)
Edited by H. Sato and T. Nakamura
World Scientific, 1992

Proceedings of the Fifth Marcel Grossmann Meeting on General Relativity
(Perth, Australia, 1988)
Edited by D.G. Blair and M.J. Buckingham
World Scientific, 1989

Proceedings of the Fourth Marcel Grossmann Meeting on General Relativity
(Rome, Italy, 1985)
Edited by R. Ruffini
World Scientific, 1986

Proceedings of the Third Marcel Grossmann Meeting on General Relativity
(Shanghai, People's Republic of China, 1982)
Edited by Hu Ning
Science Press – Beijing and North-Holland Publishing Company, 1983

Proceedings of the Second Marcel Grossmann Meeting on General Relativity
(Trieste, Italy, 1979)
Edited by R. Ruffini
North-Holland Publishing Company, 1982

Proceedings of the First Marcel Grossmann Meeting on General Relativity
(Trieste, Italy, 1975)
Edited by R. Ruffini
North-Holland Publishing Company, 1977

## SPONSORS

International Center for Relativistic Astrophysics Network (ICRANet)
International Center for Relativistic Astrophysics (ICRA)
International Centre for Theoretical Physics (ICTP)
Istituto Nazionale di Astrofisica (INAF)
International Union of Pure and Applied Physics (IUPAP)

## FREEDOM OF MOVEMENT FOR SCIENTISTS

The Marcel Grossman Meetings were founded with the premise that scientists of all nations have a right to meet to exchange knowledge independent of national borders. As such we affirm the IUPAP declaration: "To secure IUPAP sponsorship, the organisers have provided assurance that MG15 will be conducted in accordance with IUPAP principles as stated in the IUPAP resolution passed by the General Assembly in 2008. In particular, no bona fide scientist will be excluded from participation on the grounds of national origin, nationality, or political considerations unrelated to science."

## ACKNOWLEDGEMENTS

We acknowledge the outstanding job done before, during and after the meeting by the ICRANet/ICRA administrative and secretarial staff: Cristina Adamo, Silvia Latorre, Elisabetta Natale, and Cinzia di Niccolo and the conference agency support of Sistema Congressi (Marisa Santori, Ilaria Coccato, Lorena Moretti). Finally this meeting and its proceedings could not have functioned without the dedicated IT support of the ICRA system manager Vittorio Vannini for every MG meeting this century and the ICRANet system manager Gabriele Brandolini, with some temporary assistance from Valerio Antonucci and Damiano Verzulli.

## ORGANIZING BODIES OF THE FIFTEENTH MARCEL GROSSMANN MEETING:

### INTERNATIONAL ORGANIZING COMMITTEE

David Blair, Yvonne Choquet-Bruhat, Thibault Damour, Paolo De Bernardis, Francis Everitt, Riccardo Giacconi, Theodor Haensch, Stephen Hawking, Marc Henneaux, Christine Jones Forman, Roy Kerr, Hagen Kleinert, Jutta Kunz, Claus Laemmerzahl, Malcolm Longair, Tsvi Piran, Jorge Rueda, Remo Ruffini (Chair), Misao Sasaki, Humitaka Sato, Rashid Sunayev, Gerard 't Hooft, Steven Weinberg, Shing-Tung Yau

### LOCAL ORGANIZING COMMITTEE

Amati, L., Angelantonj, C., Barbiellini, G., Bassan, M., Battistelli, E. S. (chair), Belinski, V.A., Belli, L., Benedetti, R., Bernabei, R., Bianchi, M., Bianco, C.L. (co-chair), Bini, D., Buchert, T., Burgio, F., Capozziello, S., Chakrabarti, S.K., Chardonnet, P., Dall'Agata, G., De Angelis, A., De Bernardis, P., Della Valle, M., Di Virgilio, A., Frasca S., Fraschetti, F., Fré, P., Frontera, F., Giavalisco, M., Giommi, P., Gionti, G., Ingrosso, G., Jantzen, R.T., Jetzer, P., Lee, H.-W., Lerda, A., Liberati, S., Longo, R., Luzzi, G., Mandolesi, N., Marmo, G., Menotti, P., Merafina, M., Morselli, A., Pani, P., Pelster, A., Piacentini, F., Pian, E., Procesi, C., Quevedo, H., Ricci, F., Riccioni, F., Rosati, P., Ruffo, S., Scarpetta, E.V., Tavani, M., Titarchuk, L.G., Vereshchagin, G.V., Xue, S.-S., Zen Vasconcellos, C.

### INTERNATIONAL COORDINATING COMMITTEE

ALBANIA: Hafizi, M. ARGENTINA: Arguelles, C., Ghezzi, C.R., Mirabel, F., Romero, G.E. ARMENIA: Sahakyan, N. AUSTRALIA: Blair, D., Ju, L., Lun, A., Manchester, D., Melatos, A., Quinn, P., Scott, S.M., Steele, J.D., AUSTRIA: Aichelburg, P.C., Schindler, S. BELARUS: Kilin, S., Minkevich, A.V., Siutsou, I.A. BELGIUM: Henneaux, M. BOLIVIA: Aguirre, C.B. BOSNIA AND HERZEGOVINA: Pasic, V. BRAZIL: Aguiar, O., Barres de Almeida, U., Berkovits, N.J., Carneiro da Cunha, B., de Gouveia Dal Pino, E.M., de Souza Oliveira Filho, K., Guzzo, M., Hadjimichef, D., Klippert, R., Malheiro, M., Negreiros, R., Peres Menezes, D., Romero Filho, C.A, Shellard, R., Villela, T., Wuensche, C.A., Zen Vasconcellos, C. BULGARIA: Yazadjiev, S. CANADA: Singh, D., Smolin, L., Turok, N. CHILE: Bunster Weitzman, C., Giacomini, A., Reisenegger, A. CHINA: Cai, R.-G., Cao, Z., Chang, J., Chen, J.-S., Chen, X.-L, Dai, Z.-G, Feng, L.-L., Han, W.-B., Jing, Y.-P., Li, T.-P., Lin, W.-B., Lou, Y.-Q., Luo, J., Mei, J.-W., Wang, Y., Wu, X.-P., Wu, Y.-L., Yuan, Y.-F., Zhang, B., Zhang, S.-N., Zhao, G. CHINA (TAIWAN): Chen, C.-M., Chen, P., Lee, D.-S., Lee, W.-L., Ni, W.-T. COLOMBIA: Bargueño de Retes, P., Gonzalez, G., Granda Velasquez, L.N., Núñez, L., Pachón Contreras, L.A., Portilla, B., Valenzuela Toledo, C.A. CROATIA: Dominis Prester,

D., Milekovic, M., Smolčič, V., Smolić, I., Surić, T. CUBA: Pérez Martínez, A., Pérez Rojas, H., Quiros, I. CZECH REPUBLIC:Bicak, J. DENMARK: Naselsky, P. EGYPT: Tawfik, A.N., Wanas, M.I. ESTONIA: Einasto , J., Saar, E. FINLAND: Poutanen, J., Volovik, G. FRANCE: Brillet, A., Buchert, T., Chardonnet, P., Coullet, P., de Freitas Pacheco, J.A., Deruelle, N., Iliopoulos, J., Mignard, F., GEORGIA: Lavrelashvili, G., Machabeli, G. GERMANY: Biermann, P., Blumlein, J., Di Piazza, A., Fritzsch, H., Genzel, R., Gilmozzi, R., Hehl, F., Keitel, C., Kiefer, C., Mirzoyan, R., Neugebauer, G., Nicolai, H., Renn, J., Ringwald, A., Ruediger, A. GREECE: Batakis, N.A., Cotsakis, S., Vagenas, E.C. HUNGARY: Fodor, G., Levai, P. ICELAND: Bjornsson, G., Jakobsson, P. INDIA: Chakrabarti, S.K.,Iyer, B., Padmanabhan, T., Souradeep, T. IRAN: Mansouri, R., Mashhoon, B., Sobouti, Y. IRELAND: O'Murchada, N., Pe'er, A. ISRAEL: Milgrom, M., Nakar, E., Piran, T. ITALY: Belinski, V.A., Bianchi, M., Bianco, C.L., Cherubini, C., Della Valle, M., Falciano, S., Filippi, S., Menotti, P., Merafina, M., Pani, P., Ricci, F., Treves, A., Vereshchagin, G.V., Vitale, S., Xue, S.-S. JAPAN: Fujimoto, M.-K., Makishima, K., Nakamura, T., Sato, K., Shibata, M. KAZAKHSTAN: Abishev, M., Aimuratov, Y., Boshkayev, K., Mychelkin, E.G., Spitaleri, C. KOREA (PYEONGYANG): Kim, J.S. KOREA (SEOUL): Kim, S.P., Kim, S.-W., Lee, H.K., Lee, H.-W., van Putten, M.KYRGYZSTAN: Gurovich, V.Ts. LIBYA: Gadri, M. MEXICO: Breton, N., García-Diaz, A.A., Macías Alvarez, A., Mielke, E.W., Quevedo, H., Rodriguez, L.F. NEW ZEALAND: Visser, M., Wiltshire, D. NORWAY: Elgarøy, Ø., Fonseca Mota, D., Knutsen, H. POLAND: Belczynski, K., Demianski, M., Lewandowski, J., Nurowski, P., Sokolowski, L. PORTUGAL: Costa, M., da Silva, A., Lemos, J.P.S., Lobo, F., Moniz, P., Silva, L.O. ROMANIA: Visinescu, M. RUSSIA: Aksenov, A.G., Arkhangelskaja, I., Bisnovatyi-Kogan, G.S., Blinnikov, S.I., Chechetikin, V.M., Cherepaschuk, A.M., Khriplovich, I.B., Lipunov, V.M., Lukash, V.N., Novikov, I.D., Rudenko, V.N., Starobinsky, A.A. SERBIA: Djordjevic, G., Jovanovic, P., Knežević, Z., Pankov-Hzvojevic, M., Popovic, L., Prodanovic, T., M., Sijacki, D., Simic, S. SLOVAKIA: Balek, V. SLOVENIA: Cadez, A., Gomboc, A., Zavrtanik, D. SOUTH AFRICA: Colafrancesco, S., Larena, J., Maharaj, S., SPAIN: Elizalde, E., Ibanez, J., Perez Mercader, J., Verdaguer, E. SWEDEN: Abramowicz, M.A., Marklund, M., Ryde, F. SWITZERLAND: Durrer, R., Jetzer, P. TURKEY: Aliev, A., Gurses, M. UKRAINE: Novosyadlyj, B., Zaslavski, O.B., Zhuk, A. UK: Cruise, A.M., Frenk, C.S., Green, M., Mavromatos, N., Perry, M., Willingale, R. USA: Abel, T., Ashtekar, A., Bardeen, J.,Carlstrom, J., Cornish, N., Dermer, C., Fan, X., Flanagan, E., Fraschetti, F., Fryer, C., Incera, V., Jantzen, R.T. (Chairperson), Kolb, R., Laguna, P., Longair, M., Lousto, C., Madau, P., Mathews, G., Matzner, R., Melia, F., Mester, J., Michelson, P., Nordtvedt, K., Parker, L., Pretorius, F., Pullin, J., Shapiro, I., Shapiro, S., Shoemaker, D., Smoot, G., Stiavelli, M., Teukolsky, S., van Nieuwenhuizen, P., Zhang, B. UZBEKISTAN: Ahmedov, B., Zalaletdinov, R.M. VATICAN CITY: Gionti, G. VENEZUELA: Herrera, L. VIETNAM: Long, H.N.

## MARCEL GROSSMANN AWARDS

## FIFTEENTH MARCEL GROSSMANN MEETING

### Institutional Award

Planck Scientific Collaboration (ESA) (ESA)

*"for obtaining important constraints on the models of inflationary stage of the Universe and level of primordial non-Gaussianity; measuring with unprecedented sensitivity gravitational lensing of Cosmic Microwave Background fluctuations by large-scale structure of the Universe and corresponding Bpolarization of CMB, the imprint on the CMB of hot gas in galaxy clusters; getting unique information about the time of reionization of our Universe and distribution and properties of the dust and magnetic fields in our Galaxy."*
— presented to its Director General Johann-Dietrich Woerner

### Institutional Award

Hansen Experimental Physics Laboratory (HEPL) at Stanford University

*"for having developed interdepartmental activities at Stanford University at the frontier of fundamental physics, astrophysics and technology."*

### Individual Awards

Lyman Page
*"for his collaboration with David Wilkinson in realizing the NASA Explorer WMAP mission and who now leads the Atacama Cosmology Telescope as its project scientist."*

Rashid Alievich Sunyaev
*"for the development of theoretical tools in the scrutinising, through the CMB, of the first observable electromagnetic appearance of our Universe."*

Shing-Tung Yau
*"for the proof of the positivity of total mass in the theory of general relativity and perfecting as well the concept of quasi-local mass, for his proof of the Calabi conjecture, for his continuous inspiring role in the study of black holes physics ."*

Each recipient is presented with a silver casting of the TEST sculpture by the artist A. Pierelli. The original casting was presented to His Holiness Pope John Paul II on the first occasion of the Marcel Grossmann Awards.

**FOURTEENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

European Space Agency (ESA)

**Individual Awards**

Frank Chen Ny Yang, Tsung Dao Lee, Ken'ichi Nomoto, Martin Rees, Yakov G. Sinai, Sachiko Tsuruta

**THIRTEENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

ALBANOVA UNIVERSITY CENTER, STOCKHOLM

**Individual Awards**

David Arnett, Vladimir Belinski and Isaak M. Khalatnikov, Filippo Frontera

**TWELFTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Institut des Hautes Études Scientifique (IHES)

**Individual Awards**

Jaan Einasto, Christine Jones, Michael Kramer

**ELEVENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Freie Universität Berlin

**Individual Awards**

Roy Kerr, George Coyne, Joachim Trümper

**TENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

CBPF (Brazilian Center for Research in Physics)

**Individual Awards**

Yvonne Choquet-Bruhat, James W. York, Jr., Yval Ne'eman

**NINTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Solvay Institutes

**Individual Awards**

Riccardo Giacconi, Roger Penrose, Cecile and Bryce DeWitt

**EIGHTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Hebrew University of Jerusalem

**Individual Awards**

Tullio Regge, Francis Everitt

**SEVENTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Hubble Space Telescope Institute

**Individual Awards**

SUBRAHMANYAN CHANDRASEKHAR, JIM WILSON

**SIXTH MARCEL GROSSMANN MEETING**

**Institutional Award**

Research Institute for Theoretical Physics (Hiroshima)

**Individual Awards**

Minora Oda, Stephen Hawking

**FIFTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The University of Western Australia

**Individual Awards**

Satio Hayakawa, John Archibald Wheeler

**FOURTH MARCEL GROSSMANN MEETING**

**Institutional Award**

The Vatican Observatory

**Individual Awards**

William Fairbank, Abdus Salam

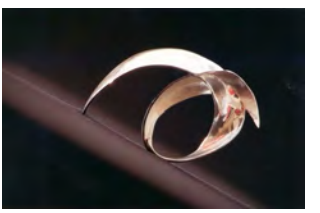


Fig. 1. TEST: sculpture by Attilio Pierelli

**PLANCK SCIENTIFIC COLLABORATION (ESA)**
presented to **Jean-Loup Puget**, the Principal Investigator of the High Frequency Instrument (HFI).

*"for obtaining important constraints on the models of inflationary stage of the Universe and level of primordial non-Gaussianity; measuring with unprecedented sensitivity gravitational lensing of Cosmic Microwave Background fluctuations by large-scale structure of the Universe and corresponding B-polarization of CMB, the imprint on the CMB of hot gas in galaxy clusters; getting unique information about the time of reionization of our Universe and distribution and properties of the dust and magnetic fields in our Galaxy"*

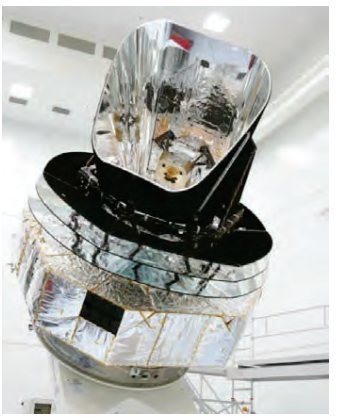
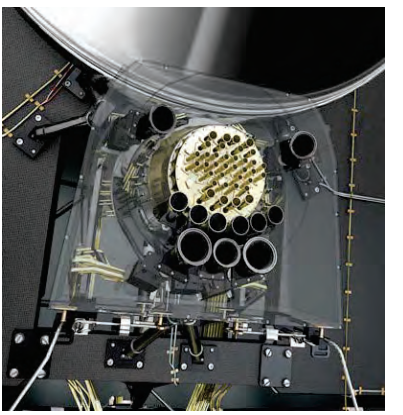


*Photo of the Planck satellite (Courtesy of ESA).*



*Planck focal plane (Courtesy of ESA)*

Planck ESA's mission, was designed to image the temperature and polarization anisotropies of the CMB over the whole sky, with unrivalled angular resolution and sensitivity, pushing the technology to unprecedent limits. In the framework of the highly precision experimental cosmology the legacy Planck results on testing theories of the early universe and the origin of cosmic structure, has provided a major source of information crucial to many cosmological and astrophysical issues. Planck carried out two instruments:

- the High Frequency Instrument (HFI), Principal Investigator: Jean Loup Puget;
- the Low Frequency Instrument (LFI), Principal Investigator: Nazzareno Mandolesi.

The instruments were complementary and using different technology to cross check independently final results and systematic errors. They worked together to produce the overall mission results. The Planck space mission (ESA) has been a wonderful example of Team effort in a large international collaboration, involving scientific, technical and managerial aspects. The unprecedented accuracy of the Planck measurements have established new standards in the determination of fundamental cosmological parameters, as well as new insight in Galactic and extragalactic astrophysics. The Planck full-sky maps in temperature and polarization will remain a lasting legacy for at least dozen years to come. More than 100 papers signed by Planck collaboration have already 30 000 citations in scientific literature. The success of Planck HFI and LFI would not have been possible without the contribution of a large number of talented and dedicated scientists and engineers from many countries of Europe, USA and Canada. HFI was designed to produce high-sensitivity, multi-frequency measurements of the diffuse radiation permeating the sky in all directions in the frequency range of 84 GHz to 1 THz cooled at 100 mK. The instrument consisted of an array of 52 bolometric detectors placed in the focal plane of the telescope. LFI, a microwave instrument, was designed to produce high-sensitivity, multi-frequency measurements of the microwave sky in the frequency range of 27 to 77 GHz. The instrument consisted of an array of 22 tuned radio receivers located in the focal plane of the telescope, cooled at 20 K.

*Jean Loup Puget - PI of the HFI.*

*Nazzareno Mandolesi - PI of the LFI.*

## HANSEN EXPERIMENTAL PHYSICS LABORATORY AT STANFORD UNIVERSITY

presented to **Leo Hollberg**, HEPL Assistant Director

*"to HEPL for having developed interdepartmental activities at Stanford University at the frontier of fundamental physics, astrophysics and technology"*

### Brief History of Stanford's HEPL and Ginzton Laboratories

1947: WW Hansen (right) and team with the Mark I linear accelerator

In 1947, working in the Stanford Physics Department's Microwave Lab, Physics Professor, William W. Hansen and his research team, along with Assistant Professor of Physics and microwave expert, Edward L. Ginzton, completed development on the world's first traveling wave electron linear accelerator. Dubbed the Mark I (see photo) it generated a 1.5 million electron volt (MeV) beam. Hansen's entire report to the U.S. Office of Naval Research (ONR) that funded the project was just one sentence: "We have accelerated electrons."

This successful first step in linear electron acceleration spawned the birth of Stanford's High Energy Physics Lab (HEPL) and Ginzton labs. In 1990, HEPL was renamed the WW Hansen Experimental Physics Lab (also HEPL). HEPL and Ginzton were setup as Stanford's first independent labs. They were organized to facilitate cross-disciplinary research, enabling scientists, engineers, staff and students to work towards common research goals using cutting edge lab equipment and technologies on medium-scale projects. For the past 70 years, the HEPL and Ginzton Labs have spearheaded Stanford's leadership in cross-disciplinary physics and become nurturing homes to a variety of physics-based, research projects: including the following examples:

### Robert Hofstadter's Nobel Prize & Later Work

In 1961, Stanford Professor Robert Hofstadter was awarded the Nobel Prize for his HEPL Mark III Linear Accelerator work on nuclear form factors (nucleons). In the 1980s, Hofstadter became interested in astrophysics and helped design the EGRET telescope in the NASA Compton Gamma Ray Observatory (CGRO).

### Gravity Probe B (GP-B)

2003: Francis Everitt (left) and Brad Parkinson holding a GP-B gyro

In 1959, Physics Department Chair, Leonard Schiff, became interested in using gyroscopes in a satellite to measure the Earth's geodetic effect and the miniscule frame-dragging effect predicted by Albert Einstein's general theory of relativity. Schiff discussed this project with Stanford cryogenic physicist, William Fairbank, and gyroscope expert, Robert Cannon (Aero-Astro department).

In 1962, Fairbank invited post-doc, Francis Everitt, to join the research effort. The team sent a proposal to NASA's Office of Space Sciences requesting funding to develop gyroscopes and a satellite to carry out this unprecedented test. It took 40 years of R&D at Stanford and other places to create and ready the cryogenic satellite and all of its cutting-edge technologies for launch. In 1975, Leonard Schiff moved the GP-B program to HEPL, breathing new life into the lab. In 1981, Francis Everitt became Principal Investigator, a position he still holds. In 1984, Brad Parkinson became Project Manager and a Co-PI, along with Co-PI's John Turneaure and Daniel DeBra.

On 20 April 2004, GP-B launched from Vandenberg AFB into a polar orbit. Data collection began on 28 August 2004 and lasted 50 weeks. Data analysis took five years in order to remove confounding factors in the data. The final results, published in PRL on 31 May 2011, yielded highly accurate geodetic and frame dragging measurements, with 0.28% and 19% margins of error, respectively.

### GPS Spinoffs from GP-B

In the 1990s, Brad Parkinson's research on GPS solutions for positioning the GP-B satellite led to two revolutionary spin-off projects: 1) Wide Area Augmentation System (WAAS) provides highly precise positioning accuracy and itegrity for navigation and the automatic landing of airplanes and 2) Precision Farming adding GPS technology to tractors has enabled the automation of many aspects of farming and has spawned a $1 billion/year industry.

### Fermi Gamma Ray Space Telescope (GLAST)

Peter Michelson, Fermi LAT Principal Investigator

Stanford Physics Professor, Peter Michelson, is a former HEPL Director and the Principal Investigator for the Large Area Telescope (LAT) on board NASA's Fermi Gamma Ray Space Telescope, the successor to CGRO/EGRET. Launched on 11 June 2008, Fermi has been highly successful mapping the gamma-ray sky. Under Michelson's guidance, HEPL's collaborations with Italy are noteworthy. The development of cryogenic bar detectors of gravitational waves, in collaboration with Edoardo Amaldi and his colleagues, established new stringent upper limits to the gravitational waves incident on the Earth. Likewise, the Fermi LAT was developed by a collaboration between Italian INFN and ASI, NASA, and international partners in France, Japan, and Sweden, and used tracking detectors developed, integrated, tested, and qualified for the mission by Italy. GP-B provided the first evidence of frame-dragging on a spinning, superconducting gyroscope. The Fermi detector offers the potential of seeing, through the GeV emission in the Binary Driven Hypernova subclass of long GRBs, the emission from a newly born Black Hole, originating in the induced gravitational collapse of a supernova hypercritically accreting on a binary neutron star companion.

### Robert Byer's LIGO and ACHIP Projects

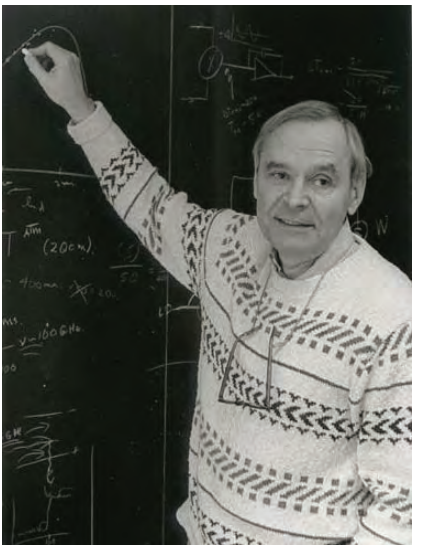
Professor Robert Byer

Robert Byer, former Stanford Dean of Research and former HEPL Director, nurtured the GP-B, GPS and Fermi programs to success during his tenure. He is currently an Applied Physics Professor specializing in lasers and optics. His LIGO Group provided seismic isolation, coatings and materials for the LIGO observatories. His ACHIP project is developing a particle accelerator on a microchip—bringing the HEPL/Ginzton Labs full circle to Hansen's 1947 research.

Professor **LYMAN PAGE**

*"for his collaboration with David Wilkinson in realizing the NASA Explorer WMAP mission and as founding director of the Atacama Cosmology Telescope"*

*Lyman Page*



*David Wilkinson*

This award is given in recognition of Lyman Page's pivotal role in transforming cosmic microwave background observations into a high-precision experimental science over the past two decades. In particular Page provided major contributions to the success of the Wilkinson Microwave Anisotropy Probe (WMAP) space mission, which delivered outstanding measurements of the CMB anisotropy and polarization pattern. He is now continuing his effort by promoting a new generation of experiments like the Atacama Cosmology Telescope to study CMB polarization to greater precision.

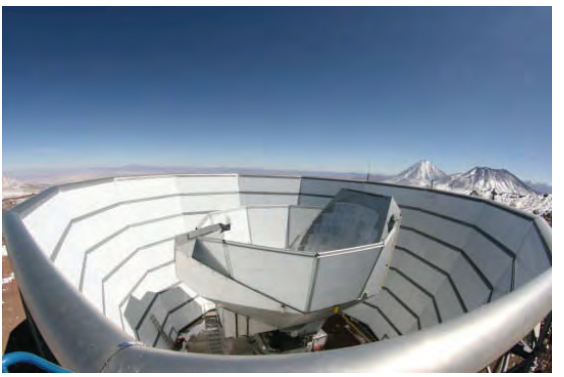


*Photo of the Atacama Cosmology Telescope*

The CMB, the faint afterglow of the Big Bang, is the most powerful probe of the early universe. From its study, we have learned the age of the universe, its major constituents, and have characterized the fundamental fluctuations in gravity that gave rise to cosmic structure. The desire to measure the CMB ever more precisely has driven the development of extraordinary detectors and techniques which will be reviewed in the Lectio Magistralis by Lyman Page. He will describe what we might hope to learn from the CMB in the next decade, including detecting gravitational waves from the birth of the universe if they exist at sufficient amplitude.

Professor **RASHID ALIEVICH SUNYAEV**
*"for the development of theoretical tools in the scrutinising, through the CMB, of the first observable electromagnetic appearance of our Universe"*

*Rashid Sunyaev*

*Rashid Sunyaev and Yakov Borisovich Zeldovich*

Rashid Sunyaev gave extraordinary contributions to the understanding of physical processes in the universe which identified new and uniquely informative targets for observational cosmology. In particular, the Sunyaev-Zeldovich effect, now observed in thousands of clusters of galaxies over the entire sky, has become a cornerstone of cosmology and extragalactic astrophysics, so much so that it is now considered a research field in its own right. Furthermore, Sunyaev's studies of processes in the early universe responsible for angular anisotropy and frequency distortions of the cosmic microwave background have left a profound and lasting legacy for cosmology. In particular, Sunyaev and Zeldovich predicted the presence of acoustic peaks in the CMB angular fluctuation power spectrum and the existence of baryonic acoustic oscillations.
He is currently the project scientist leading the scientific team of the international high-energy astrophysics observatory Spektr-RG being built under the direction of the Russian Space Research Institute.

*Yakov Zeldovich and Remo Ruffini at the audience with Pope John Paul II*

Professor **SHING-TUNG YAU**
*"for the proof of the positivity of total mass in the theory of general relativity and perfecting as well the concept of quasi-local mass, for his proof of the Calabi conjecture, for his continuous inspiring role in the study of black holes physics"*

*Shing-Tung Yau*

Shing-Tung Yau has made fundamental contributions to differential geometry which have influenced a wide range of scientific disciplines, including astronomy and theoretical physics. With Richard Schoen, Yau solved a longstanding question in general relativity by proving that the combined total energy of matter and gravitational field in an asymptotically flat universe is positive. In 1982 Yau was awarded the Fields Medal, the highest award in mathematics, and in 1994 he shared with Simon Donaldson the Crafoord Prize of the Royal Swedish Society in recognition of his development of nonlinear techniques in differential geometry leading to the solution of several outstanding problems.

Another outstanding achievement of Yau is his proof of the Calabi conjecture which allowed physicists to show that string theory is a viable candidate for a unified theory of nature. Furthermore in 2008 Yau (with M.T. Wang) introduced the concept of "quasi-local mass" in general relativity which can be of help to get around the old conundrum — the non-locality of the energy density in relativistic gravity.

During his scientific carrier Yau had more than 50 successful PhD students. At present he is a professor of mathematics at Harvard University where along with research he continues many pedagogical activities. For example, he has created the "Black Hole Initiative", an interdisciplinary center at Harvard University involving a collaboration between principal investigators from the fields of astronomy (Sheperd Doeleman, Avi Loeb and Ramesh Narayan), physics (Andrew Strominger), mathematics (Shing-Tung Yau) and philosophy (Peter Galison). This "Black Hole Initiative" is the first center worldwide to focus on the study of the many facets of black holes.

MG16 in 2021 will mark the $50^{th}$ anniversary of the mass-energy formula for black holes based on the Kerr metric. This timing is an omen that Yau and his school will soon further enlarge our knowledge of this formula with their powerful mathematical analysis.

# PREFACE

The Marcel Grossmann Meetings on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics and Relativistic Field Theories have always had the goal of bringing together scientists from all over the world enabling them to share recent developments in general relativity and cosmology, paying attention to the interplay between physical predictions and mathematical foundations.

More than 800 scientists met at the University of Rome "la Sapienza" for the Fifteenth Marcel Grossmann Meeting (MG15) during the week of July 1–7, 2018 to discuss theoretical topics and the status of the experimental testing and observations of Einstein's theory of gravitation together with the broad spectrum of gravitational physics related phenomena. The topics discussed ranged from classical and quantum aspects of gravity, relativistic astrophysics, branes and strings, inflation theory, the thermal history of the Universe, to more concrete astrophysics experiments, observations, and modeling, reviewing the current state of the art in theory, observations, and experiments related to relativistic gravitation.

The meeting featured 39 plenary talks during the six mornings of the meeting, plus 6 public lectures given by experts in gravitation and cosmology. 73 parallel sessions, some of which were split over multiple days, kept participants busy with the crucial support of 33 students from Sapienza who managed the coordination of each parallel session and offered technical help.

Three scientists and two collaborations were presented with Marcel Grossmann awards. These were: Lyman Page, Rashid Alievich Sunyaev, Shing-Tung Yau, the Planck Scientific Collaboration (ESA) represented by Jean-Loup Puget, and the Hansen Experimental Physics Laboratory (HEPL) at Stanford University, represented by Leo Hollberg,

The detailed program of plenary and public talks is as follows:

**Shing-Tung Yau** (Harvard University): Quasi-local mass at null infinity
**Malcolm J. Perry** (University of Cambridge): Black hole entropy and soft hair
**Thomas Hertog** (KU Leuven): A smooth exit form eternal inflation
**Jean-Luc Lehners** (Max Planck Institute for Gravitational Physics): No smooth beginning for spacetime
**Ivan Agullo** (Louisiana State University): Loop quantum cosmology and the cosmic microwave background
**Elena Pian** (IASF Bologna):Kilonovae: the cosmic foundries of heavy elements
**Nial Tanvir** (University of Leicester): A new era of gravitational-wave / electromagnetic multi-messenger astronomy
**Tsvi Piran** (Hebrew University of Jerusalem): Mergers and GRBs: past present and future
**Stephan Rosswog** (Stockholm University): Neutron star mergers as heavy element production site

**David Shoemaker** (MIT LIGO Laboratory): LIGO's past and future observations of black hole and neutron star binaries
**Yu Wang** (ICRANet): On the role of binary systems in GW170817/ GRB170817A/AT2017gfo
**Hao Liu** (University of Copenhagen): An independent investigation of gravitational wave data
**Stefano Vitale** (University of Trento): Gravitation wave astronomy in ESA science programme
**Takaaki Kajita** (University of Tokyo): Status of KAGRA and its scientific goals
**Masaki Ando** (University of Tokyo): DECIGO : Gravitational-Wave observations from space
**Jun Luo** (Sun Yat-Sen University): TianQin: a space-borne gravitational wave detector
**Jo Van Den Brand** (Dutch National Institute for Subatomic Physics Nikhef, and VU University Amsterdam): Gravitational wave science and Virgo
**Ernst Maria Rasel** (Leibniz Universität Hannover): Science fiction turns into reality: Interferometry with Bose-Einstein condensates on ground and in space
**Manuel Rodrigues** (Université Paris Saclay): The first results of the MICROSCOPE test of the equivalence principle in space
**Victoria Kaspi** (McGill University): Fast radio bursts
**Bing Zhang** (University of Nevada): From gamma-ray bursts to fast radio bursts: unveiling the mystery of cosmic bursting sources
**Jean-Loup Puget** (CNRS): The Planck mission
**Jorge Armando Rueda Hernandez** (ICRANet): Binary-driven hypernovae and the understanding of gamma-ray bursts
**Remo Ruffini** (ICRANet): The essential role of the nature of the binary progenitors for understanding gamma ray bursts
**Heino Falcke** (Radboud University Nijmegen): Imaging black holes now and in the future
**Luc Blanchet** (Institut d'Astrophysique de Paris): Post-Newtonian theory and gravitational waves
**Jean-Loup Puget** (Université Paris): Frome Planck to Atacama Cosmology Telescope
**Razmik Mirzoyan** (Max Planck Institute for Physics): Gamma-ray and multimessenger highlights with MAGIC
**Elisa Resconi** (Technical University Munich): Neutrino astronomy in the multimessenger era
**Francis Halzen** (University of Wisconsin-Madison): IceCube: opening a new window on the universe from the South Pole
**James Lattimer** (Stony Brook University): The history of R-process
**Ralph Engel** (Karlsruhe Institute of Technology): What have we learned about ultra-high-energy cosmic rays from the Pierre Auger Observatory?

**Paolo De Bernardis** (Sapienza - University of Rome):
**Fabio Gargano** (INFN Bari): DAMPE and its latest results
**Markus Arndt** (University of Vienna): Experiments to probe quantum linearity at the interface to gravity & complexity
**Tobias Westphal** (University of Vienna): Micro-mechanical measurements of weak gravitational forces
**Shu Zhang** (Institute of High Energy Physics, Chinese Academy of Sciences):
**Lorenzo Amati** (INAF - OAS Bologna): Cosmology and multi-messenger astrophysics with Gamma-Ray Bursts
**Elisabetta Cavazzuti** (ASI): Gev LAT observations from GRBs and active galactic nuclei
**Remo Ruffini** (ICRANet): Concluding Remarks
**Roy Kerr** (University of Canterbury): Towards MG16

**Public Lectures**

**Jeremiah Ostriker** (Columbia University): Ultra-light scalars as cosmological dark matter
**Malcolm Longair** (University of Cambridge): Ryle and Hewish: 50 and 100 year anniversaries [Radio astrophysics and the rise of high energy astrophysics]
**Lyman Page** (Princeton University): Measuring the Cosmic Microwave Background
**Marc Henneaux** (Université Libre de Bruxelles): The cosmological singularity
**Anne Archibald** (Newcastle University): Does extreme gravity affect how objects fall?

# Concluding remarks by Prof. Remo Ruffini

For over more than four decades the Marcel Grossmann Meetings have been fostering the interaction between mathematics, relativistic field theories and observations in physics and astrophysics and has witnessed the birth and exponential growth of various new subfields within astrophysics. This has also occurred in 2018 during the MG15 (see Fig. 1), which has seen the participation of approximately one thousand participants from 71 different countries (see Fig. 2). In MG15, as is clear from the above program, we have seen considerable progress in theory on topics ranging from models of quantum gravity to the mathematical structure of Einstein's equations.


Fig. 1. MG15 official Awards Ceremony on Monday, July 2, 2018.

Fig. 2. Group photo of MG15 participants in the Aula Magna, Sapienza University of Rome.

In particular, an MG15 individual award was given to Prof. Rashid Sunyaev, as a member of the relativistic astrophysics school of Zel'dovich-Sunyaev (see Fig. 3) and to Prof. Shing-Tung Yau, as a representative of the Chinese-American school of differential geometry and general relativity (see Fig. 4).

Fig. 3. Rashid Sunyaev receiving the MG15 individual award from Roy Kerr.

Fig. 4. Shing-Tung Yau receiving the MG15 individual award from Roy Kerr.

This has been accompanied by extensive developments in the exploration of the cosmic microwave background with announcements of the final results of the Planck satellite that followed earlier predecessors, like WMAP. In this sense, particularly meaningful have been the MG15

Awards presented to Prof. Jean-Loup Puget, who has successfully accomplished the ESA Plank mission (see Fig. 5) and to Prof. Lyman Page, one of the main participants of the glorious NASA WMAP mission, successfully lead by David Wilkinson (see Fig. 6). I was very fortunate to be in Princeton as an assistant professor at the time WMAP was initially conceived.

Fig. 5. Jean-Loup Puget accepting the MG15 institutional award on behalf of the Planck Scientific Collaboration (ESA) from Roy Kerr.

Fig. 6. Lyman Page accepting the MG15 individual award from Roy Kerr.

X-ray astrophysics gradually successfully evolved from the first rockets of Riccardo Giacconi discovering Scorpius X-1 using well tested Geiger counters, moving on to larger missions using X-ray mirrors in the Einstein Observatory, in XMM and on to the NASA Chandra mission, each new mission introducing new technologies based on previous successes. The

Chandra data on GRB 170817A were amply discussed in the MG15 meeting. Analogously, starting from the earlier gamma ray detectors on the Vela satellite, gamma ray astrophysics reached maturity with the Compton Observatory with the BATSE instruments on board. These two fields joined their separate expertise together in the hybrid Beppo SAX satellite. Equally impressive was the transition from Beppo SAX to the Neil Gehrels Swift Observatory and the Fermi Gamma-Ray Space Telescope that followed EGRET and opened up the field of high energy gamma-ray astrophysics, adding further successes to this ongoing story. One of the most significant contributions to the success of the Fermi mission has been the LAT detector, jointly led by the Hansen Experimental Physics Laboratory (HEPL) at Stanford University, which received the MG15 institutional award, presented to Prof. Leo Hollberg (see Fig. 7).

Fig. 7. Leo Hollberg accepting the MG15 institutional award from Roy Kerr on behalf of the Hansen Experimental Physics Laboratory at Stanford University.

More recently we are witnessing the birth of TeV astrophysics springing from the ground-based MAGIC, HESS and Whipple telescopes, as well as neutrino astrophysics associated with underground (ICE Cube) and underwater laboratories. New tantalizing results concerning the spectrum and composition of cosmic rays have been reported by the Alpha Magnetic Spectrometer experiment on the international space station. All of these

together mark the gradual development of the largest observational effort in history, well recorded in these MG meetings (see Fig. 8). Today relativistic astrophysics is offering us the possibility of utilizing all of these multi-wavelength observations simultaneously in the study of GRBs and active galactic nuclei. Indeed the description of the distinct temporally discrete episodes for the GRB 130427A, one of the most complex astrophysical objects ever observed, requires the simultaneous knowledge of observations from all of these various wavelengths. This exponential growth in all these areas of astrophysics has been accompanied by attempts at data acquisition of gravitational wave signals arriving on the Earth. After the failure of the initial attempt by Misner and Weber in 1972, the new attempt by LIGO-VIRGO was discussed extensively in this MG15 meeting.

Fig. 8. From the left to right: Leo Hollberg, Rashid Sunyaev, Shing-Tung Yau and Remo Ruffini during the MG15 official awards ceremony.

Indeed in addition to the many topics discussed in the plenary lectures, there was a special session dedicated to kilonova and gravitational waves chaired by Enrico Costa on July 3 and a session on gravitational waves chaired by Claus Lämmerzahl on July 4. Many intense discussions took place at the meeting, both in the plenary and parallel sessions, bringing together different points of view, at times controversial, which stimulated a large number of articles published in leading scientific journals in the months following the meetings. The impossibility of finding these discussions in these proceedings is simply explained: it is due to the incomprehensible decision by Sapienza University not to renew the collaboration agreement with ICRANet. What is incomprehensible is that both the research and teaching activities of ICRA and ICRANet are

recognized internationally at a worldwide level, in more than 70 nations. Locks were put on the ICRA Offices (see Fig. 9).

Fig. 9. Remo Ruffini shows to Barry Barish the locked offices of ICRA at Sapienza University.

The preparation of the proceedings was halted. It was not possible to follow up with post-conference interactions and the impossibility of access to our offices and interact with our students. The invited lecturers, clearly indicated in the conference program and some of the world leaders in their fields who had just presented their extremely interesting lectures stimulating widespread interest, mostly chose to publish their results in international journals giving rise to some of the most highly cited articles in recent times. Only a minority were published here. This was also the case for some of the ICRA-ICRANet lectures, which appeared in:

- J.F. Rodriguez, J.A. Rueda, R. Ruffini, *On the final gravitational wave burst from binary black hole mergers*, Astron. Rep. 62, 940 (2018).
- J.A. Rueda, R. Ruffini, Y. Wang, Y. Aimuratov, U. Barres de Almeida, C.L. Bianco, Y.C. Chen, R.V. Lobato, C. Maia, D. Primorac, R. Moradia, and J.F. Rodriguez, *GRB 170817A-GW170817-AT 2017gfo and the observations of NS-NS, NS-WD and WD-WD mergers*, JCAP 10, 006 (2018); and J.A. Rueda, R. Ruffini, Y. Wang, C.L. Bianco, J.M. Blanco-Iglesias, M. Karlica, P. Lorén-Aguilar, R. Moradi and N. Sahakyan, *Electromagnetic emission of white dwarf binary mergers*, JCAP 03, 044 (2019).

- R. Ruffini, R. Moradi, J. A. Rueda, L. Becerra, C. L. Bianco, C. Cherubini, S. Filippi, Y. C. Chen, M. Karlica, N. Sahakyan, Y. Wang, and S. S. Xue, *On the GeV Emission of the Type I BdHN GRB 130427A*, The Astrophysical Journal, 886, 82 (2019).

The dialogue we had started at MG15 undoubtedly had a very stimulating effect on the community which have been followed up and expanded upon in the subsequent MG16 meeting in 2021. Paradoxically, the moments of greatest difficulty for ICRA members have coincided with the flow of new ideas which promoted the steps forward towards the final identification of some fundamental laws of black hole electrodynamics. The very night ICRA offices were locked up and its electronic communications cut, we had submitted a paper to The Astrophysical Journal for publication containing the first fundamental steps for identifying the "inner engine" of GRBs as discussed in talks in this conference. After one year and four months a judge ordered Sapienza to return the ICRA offices.

I am looking forward to a renewed successful collaboration with Sapienza University, enjoyed for more than three decades, to pursue a new common "path" for the rapidly expanding knowledge of the largest distant objects in our Universe. There is still room for improvement: now that the differences have been identified in published articles, a joint effort on an attentive scientific analysis on both sides can lead to the understanding of the nature of these differences and to jointly converge to common solutions. For this to occur, dialogue and not obstruction of office space is needed. As soon as these clarifications will be achieved, we will be able to proceed in a broad scientific effort of common topics of interest in Einstein's theory, in relativistic astrophysics, ranging from the astrophysics of black holes, neutron stars, dark matter and primordial cosmology, observing the rules of basic academic relations.

I am grateful to all members of the IOC and to Carlo Luciano Bianco, Nathalie Deruelle, Rahim Moradi, Tsvi Piran, Jorge Rueda and Narek Sahakyan for advice and discussion.

Remo Ruffini
Chair of the MG International Organizing Committee
Director of ICRANet
President of ICRA

January 10, 2022

This page intentionally left blank

# CONTENTS



## PART A
## PLENARY AND REVIEW TALKS

## PARALLEL SESSIONS

## PART B

## PART C

# PART C

This page intentionally left blank

# Chiral symmetry breaking and the Unruh effect

Adrián Casado-Turrión and Antonio Dobado

*Departamento de Física Teórica and Instituto IPARCOS. Universidad Complutense, Madrid, 28040, Spain*

The possibility of chiral symmetry restoration by acceleration is considered. The Thermalization Theorem formalism and the large $N$ limit (with $N$ being the number of pions) are employed to solve the lowest-order approximation to QCD at low energies in Rindler spacetime. It is shown that chiral symmetry is restored for accelerations higher than the critical value $a_c = 4\pi f_\pi$, with $f_\pi$ being the pion decay constant. The results are completely analogous to those obtained in the inertial, finite-temperature case, evincing the ontic character of the Unruh effect.

*Keywords*: Unruh effect; chiral symmetry restoration.

## 1. Introduction

The decade of the 1970s saw the generalization of quantum field theory to curved spacetimes and arbitrary observers in Minkowski spacetime. This led to the establishment of several seminal results, such as the discovery by Hawking that black holes have an intrinsic temperature[1], and thus radiate until they evaporate.

Attempting to better understand Hawking radiation, Unruh found[2] that a similar result may be obtained even when gravity is not present: he proved that a uniformly accelerated observer in Minkowski spacetime perceives the vacuum state of a free quantum field theory as a thermal ensemble of particles at temperature $T = a/2\pi$ (in natural units), with $a$ being the observer's acceleration. Despite the fact that this phenomenon requires experimentally unattainable accelerations to be observed directly (for instance, accelerations of the order of $10^{22}$ times the Earth's mean surface gravity are required to produce an Unruh temperature similar to ambient temperature), it is nevertheless considered to be highly fundamental, since it has been derived in several ways, such as the study of the response of particle detectors along non-inertial trajectories, canonical quantization[3] and even axiomatic quantum field theory[4] (for a thorough review of the Unruh effect, see Ref. 5).

However, none of these methods allows to discern whether the temperature–acceleration analogy introduced by the Unruh effect is just a formal result or if, on the contrary, it is capable of triggering non-trivial dynamical effects, such as phase transitions. The answer to this question requires the use of a formalism which is based on a path-integral approach to field theory, which will therefore be ideally suited to address arbitrary interacting theories and the connection between quantum field theory and thermodynamics. This formalism was developed by Lee[6] in 1986, and is known as the Thermalization Theorem.

In this work, we shall make use of Lee's formalism in order to study whether accelerated observers are able to experience the spontaneous symmetry breaking or restoration of chiral symmetry, which is one of the main features of the fundamental

theory of strong interactions, QCD, at finite temperature in Minkowski spacetime. Indeed, it has already been shown that accelerating observers detect the restoration of continuous global symmetries in some systems featuring Spontaneous Symmetry Breaking (SSB), such as the Nambu-Jona-Lasinio model,[7] the $\lambda\Phi^4$ theory at the one-loop level[8] and the Linear Sigma Model (L$\sigma$M) in the large $N$ limit.[9]

## 2. Rindler spacetime and the Thermalization Theorem

Consider a uniformly accelerated observer with proper acceleration $a$ along the $X$ direction of Minkowski spacetime $\mathrm{d}s^2 = \mathrm{d}T^2 - \mathrm{d}X^2 - \mathrm{d}Y^2 - \mathrm{d}Z^2 \equiv \mathrm{d}T^2 - \mathrm{d}X^2 - \mathrm{d}X_\perp^2$, where $X^\mu = (T, X, Y, Z)$ are the usual Cartesian-like coordinates. The trajectory of such an observer is the hyperbola $T^2 - X^2 = -1/a^2$, whose asymptotes —the null lines $T = \pm X$— divide Minkowski spacetime into four quadrants: the regions $\pm X > |T|$, respectively known as the right ($R$) and left ($L$) Rindler wedges, each of which contains one of the branches of the hyperbola; and the regions $\pm T > |X|$, which are respectively the future ($F$) and past ($P$) of the origin. Notice that, while $R$ and $L$ are causally disconnected, both share a common past, and so correlations between them may exist. Hyperbolic motion is conveniently described using comoving coordinates $x^\mu = (t, x, y, z)$, which range in $(-\infty, \infty)$ and are defined by

$$T = a^{-1}\mathrm{e}^{ax}\sinh(at), \quad X = \pm a^{-1}\mathrm{e}^{ax}\cosh(at), \quad Y = y, \quad Z = z, \tag{1}$$

with the plus-sign choice covering only $R$ and the minus-sign choice covering only $L$. It is important to notice that this implies that the Euclidean Minkowski time $T_E \equiv \mathrm{i}T$ and $X$ are periodic functions of the Euclidean comoving time $t_E \equiv \mathrm{i}t$, with period $2\pi/a$. In terms of comoving coordinates, the Minkowski metric is

$$\mathrm{d}s^2 = \mathrm{e}^{2ax}(\mathrm{d}t^2 - \mathrm{d}x^2) - \mathrm{d}x_\perp^2 \tag{2}$$

in both $R$ and $L$. The worldline with $x = 0$ corresponds to the trajectory of constant proper acceleration $a$; in fact, each of the worldlines with constant $x$ is a trajectory with proper acceleration $a(x) = a\mathrm{e}^{-ax}$. Hence, it is also useful to define the so-called Rindler coordinates $\rho \equiv 1/a(x) = \mathrm{e}^{ax}/a \in [0, \infty)$ and $\eta \equiv at \in (-\infty, \infty)$.

Let us suppose now that there exists a certain quantum field theory defined over Minkowski spacetime, which describes fields of arbitrary spin and their possible interactions. Consider, without loss of generality, a Rindler observer in $R$ (the results generalize straightforwardly to $L$). Because this observer is causally disconnected from all the events in $L$, she will be insensitive to any vacuum fluctuations outside $R$. As a result, the Minkowski vacuum state of the theory, which is a pure state $|\Omega_M\rangle$ for any inertial observer, becomes a mixed state in the eyes of the accelerated observer, who needs to perform a partial trace over the degrees of freedom in $L$. Most importantly, the corresponding density matrix was shown by Lee to be

$$\rho_R = \mathrm{tr}_L|\Omega_M\rangle\langle\Omega_M| = \frac{\mathrm{e}^{-2\pi H_R/a}}{\mathrm{tr}\,\mathrm{e}^{-2\pi H_R/a}}, \tag{3}$$

where $H_R$ is the Rindler Hamiltonian (i.e. the generarator of $t$-translations). This density matrix describes a thermal ensemble at temperature $T = a/2\pi$, which is precisely the Unruh temperature originally found for free scalar field theories. However, the applicability of this result is virtually universal: it guarantees that the Minkowski vacuum state will be perceived by a Rindler observer as a thermal state, independently of the field theory considered. This is the reason why it is also known as the Thermalization Theorem.

## 3. Chiral symmetry and its restoration by acceleration

Two-flavor massless QCD in ordinary Minkowski spacetime is known to possess two distinct phases, one in which chiral $\mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$ symmetry is spontaneously broken into isospin symmetry $\mathrm{SU}(2)_{L+R}$, and one in which it is restored. Both phases are separated by a typical Landau-Ginzburg (i.e. second order) phase transition at a certain critical temperature $T_c$, with the order parameter being the quark condensate $\langle \bar{q}q \rangle_T$, where $q$ is the quark field and $\langle \cdot \rangle_T$ denotes the expectation value at temperature $T$ on the Minkowski QCD vacuum state. Thus, because of the temperature–acceleration analogy, it is natural to wonder whether a Rindler observer perceives a restoration of chiral symmetry if her acceleration $a$ is higher than a certain critical value $a_c$. This issue may be elucidated by functionally quantizing the low-energy effective theory for QCD[a] in the right Rindler wedge, and then computing the expectation $\langle \bar{q}q \rangle_a$ on the Minkowski QCD vacuum (which, as per the Thermalization Theorem, will be felt by the accelerated observer as a thermal ensemble at temperature $a/2\pi$). The lowest-order effective field theory for QCD is a non-linear sigma model (NL$\sigma$M) whose Euclidean partition function in $R$ is

$$Z = \int [\mathrm{d}\Phi][\mathrm{d}\lambda] \exp\left( - \int \mathrm{d}^4x \sqrt{g} \left( \frac{1}{2} \partial_\mu \Phi^{\mathrm{T}} \partial_\mu \Phi + \frac{\lambda}{2} (\Phi^{\mathrm{T}} \Phi - f_\pi^2) - M_\pi^2 f_\pi \sigma \right) \right), \quad (4)$$

where $\sqrt{g}$ is the determinant of the Euclidean Rindler metric, $M_\pi$ is the pion mass, $f_\pi$ is the pion decay constant, and $\Phi^{\mathrm{T}} = (\pi^a, \sigma)$ is a real quadruplet belonging to the fundamental representation of $\mathrm{SO}(4) \simeq \mathrm{SU}(2)_L \times \mathrm{SU}(2)_R$, consisting of the three pion fields $\pi^a$ (with $a = 1, 2, 3$) plus an additional field $\sigma$ related to them through the non-linear constraint of the model, $\Phi^{\mathrm{T}} \Phi = \pi^a \pi^a + \sigma^2 = f_\pi^2$, which is in turn enforced through the introduction of the non-dynamical Lagrange-multiplier field $\lambda$. The fields are also required to satisfy the thermal-like Rindler boundary conditions $\pi^a(\tau = 0, \mathbf{x}) = \pi^a(\tau = 2\pi/a, \mathbf{x})$, $\sigma(\tau = 0, \mathbf{x}) = \sigma(\tau = 2\pi/a, \mathbf{x})$, $\sigma(\tau, |\mathbf{x}| = \infty) = f_\pi$ and $\lambda(\tau = 0, \mathbf{x}) = \lambda(\tau = 2\pi/a, \mathbf{x})$. In addition, by comparing (4) with the standard QCD partition function (restricted to $R$) it is clear that $\langle \bar{q}q \rangle_a \propto \langle \sigma \rangle_a$.

[a] As it is well known, QCD is asymptotically free, and thus becomes non-perturbative at low energies. In this regime, quarks and gluons cease to be the relevant degrees of freedom and an effective description in terms of the next lightest particles in the hadronic spectrum, the pions, is needed. Coincidentally, the pions are also the (pseudo-)Nambu-Goldstone bosons associated to chiral symmetry breaking.

The NL$\sigma$M model described above may be solved non-perturbatively in the large $N$ limit, where $N$ is the number of pions. This limit is properly defined if $f_\pi^2$ is of order $N$, i.e. $f_\pi^2 \equiv NF^2$, with $F^2$ being $N$-independent. Bearing this in mind, the Euclidean partition function may be rewritten in the chiral limit $M_\pi \to 0$ as

$$Z = \int [\mathrm{d}\pi][\mathrm{d}\sigma][\mathrm{d}\lambda] \exp\left(-\int \mathrm{d}^4x\,\sqrt{g}\left(-\frac{1}{2}\pi^a\Box\pi^a - \frac{1}{2}\sigma\Box\sigma + \frac{\lambda}{2}(\pi^2+\sigma^2-f_\pi^2)\right)\right). \quad (5)$$

The functional integral over the $N$ pion fields is Gaussian, and thus may be readily performed. Defining $\Gamma[\sigma,\lambda]$ as the effective action in the exponent of the remaining integral, the fields can be expanded around some point $(\bar\sigma,\bar\lambda)$ in the functional space where the first functional derivative of $\Gamma[\sigma,\lambda]$ vanishes. Using the steepest descent method, we have $Z = \mathrm{e}^{-\Gamma[\bar\sigma,\bar\lambda]} + \mathcal{O}(N^{-1/2})$ and $\bar\sigma^2(x) = \langle\sigma(x)\rangle_a^2 = \langle\sigma^2(x)\rangle_a$ in the large $N$ limit. Thus $\bar\sigma(x)$ and $\bar\lambda(x)$ are chosen to be the solutions of

$$\frac{\delta\Gamma}{\delta\sigma(x)} = -\Box\sigma + \lambda\sigma = 0, \quad (6)$$

$$\frac{\delta\Gamma}{\delta\lambda(x)} = \frac{1}{2}(\sigma^2 - f_\pi^2) + \frac{N}{2}G(x,x;\lambda) = 0, \quad (7)$$

with boundary conditions $\bar\sigma = f_\pi$ and $\bar\lambda = 0$ at $x \to \infty$, and where

$$(-\Box + \lambda)_x\, G(x,x';\lambda) = \frac{1}{\sqrt{g}}\delta^4(x-x'). \quad (8)$$

Unfortunately, an exact solution of these equations cannot be easily found. However, they may be solved approximately for $ax << 1$ (i.e. close to the origin of the accelerating frame) by exploiting the peculiar properties of Rindler spacetime. In this region, it can be shown[10] that $\lambda \simeq 0$, and that the relevant Green function is

$$G(x,x;0) = \int_0^\infty \mathrm{d}\Omega\,\frac{\Omega\pi}{2\rho^2\tanh(\Omega\pi)}. \quad (9)$$

Introducing $\omega \equiv a\Omega$ and expanding $a\rho = 1 + ax + \ldots$, equation (7) becomes

$$\sigma^2 = f_\pi^2 - \frac{N}{4\pi^2}(1-2ax)\int_0^\infty \mathrm{d}\omega\,\omega\left(1 + \frac{2}{e^{2\pi\omega/a}-1}\right) + \mathcal{O}(a^2x^2). \quad (10)$$

While the second integral on this equation may be easily calculated, the first one clearly requires regularization (for example, by introducing an $x$-dependent ultraviolet cutoff $\Lambda\mathrm{e}^{-ax}$, which leads to a renormalization of $f_\pi$). Thus, if we now define the $N$-independent critical acceleration as $a_c^2 \equiv 48\pi^2 f_\pi^2/N$ we have

$$\bar\sigma^2(x) = f_\pi^2\left(1 - \frac{a^2}{a_c^2} + 2x\frac{a^3}{a_c^2} + \ldots\right), \quad (11)$$

which is also a solution of equation (6) at this order. Assuming $\bar\sigma(0)$ to be real, the evaluation of this expression at the origin $x = 0$ of the accelerating frame yields

$$\frac{\bar\sigma(0)\big|_a}{\bar\sigma(0)\big|_0} = \frac{\langle\sigma(0)\rangle_a}{\langle\sigma(0)\rangle_0} = \frac{\langle\bar q(0)q(0)\rangle_a}{\langle\bar q(0)q(0)\rangle_0} = \begin{cases}\sqrt{1-\dfrac{a^2}{a_c^2}} & \text{if } 0 \le a < a_c,\\ 0 & \text{if } a \ge a_c.\end{cases} \quad (12)$$

Thus, we clearly see that acceleration triggers the second-order chiral phase transition at the origin of the Rindler frame. For $N = 3$ pions (i.e. the physical case for two-flavour QCD), the critical acceleration $a_c = 4\pi f_\pi$ corresponds to an Unruh-like temperature $T_c = 2f_\pi$, which is precisely the temperature at which the QCD phase transition is found to take place if the NL$\sigma$M is solved in Minkowski spacetime[11] in the same large $N$ limit that we have consider in $R$. Also, the same behavior for $\langle\bar{q}q\rangle_T/\langle\bar{q}q\rangle_0$ is found in this case, but substituting $a/a_c \mapsto T/T_c$. However, the expression for $\langle\bar{q}q\rangle_T/\langle\bar{q}q\rangle_0$ is valid in all Minkowski space, due to its homogeneity and isotropy.

We can extend our discussion to other points close to the origin by considering two different accelerated observers at Rindler coordinates $\rho = 1/a$ and $\rho' = 1/a'$. From the point of view of the first observer, the second one is located at some point with $x$ coordinate given by $\rho' = 1/a' = \mathrm{e}^{ax}/a$, i.e. her acceleration is $a' = a\mathrm{e}^{-ax}$. Thus, the position-dependent condensate is given by

$$\frac{\langle\bar{q}(x)q(x)\rangle}{\langle\bar{q}(0)q(0)\rangle} = \sqrt{1 - \frac{a^2}{a_c^2}\mathrm{e}^{-2ax}}. \tag{13}$$

Therefore, for a Rindler observer with acceleration $a \in (0, a_c)$, the condensate ranges from $\langle\bar{q}(0)q(0)\rangle$ at infinity to zero at the critical value $x_c \equiv \ln(a/a_c)/a < 0$.

This means that the chiral phase transition takes place on the $(x_c, x_\perp)$ surface. The symmetry is also restored on the region close to the horizon $x < x_c$, where $\langle\bar{q}(x)q(x)\rangle = 0$. Thus, the boundary between the broken and restored phases is completely determined by $a_c$, which only depends on QCD parameters.

Finally, it is interesting to notice that (13) implies that, for points different to the origin, what the condensate feels is equivalent to a thermal bath with a space-dependent temperature[12] $T(x) \equiv a\mathrm{e}^{-ax}/2\pi$, which diverges at the horizon and goes to zero at infinity. This temperature is consistent with the Tolman and Ehrenfest rule[13] for thermal equilibrium in static spacetimes, since $T(x)\sqrt{g_{00}} = a/2\pi$ is an $x$-independent constant, as required by the rule.

## 4. Conclusions

The powerful Thermalization Theorem formalism has allowed us to demonstrate the ability of the Unruh effect to produce non-trivial dynamical effects. In particular, we have been able to study the restoration of the highly fundamental QCD chiral symmetry by acceleration. We have shown that analogous results are found in both the thermal and Rindler cases, with a typical second-order phase transition occurring for those accelerated observers whose accelerations are higher than the critical value $a_c = 4\pi f_\pi \simeq 1.6$ GeV for $N = 3$ pions. The behavior obtained for the spacetime-dependent temperature felt by the condensate is also compatible with the standard requirements for thermodynamic equilibrium in static spacetimes, of which Rindler space is a particular example. Our results may have applications in ultra-relativistic heavy collisions, in which the Unruh effect has been proposed

as a possible thermalization mechanism,[14] and also in black holes or cosmological scenarios, since local Rindler coordinates can always be defined close to any horizon.[15]

**Acknowledgments**

A. D. thanks Luis Álvarez-Gaumé and C. Pajares for very interesting comments and discussions. Work supported by the Spanish grant FPA2016-75654-C2-1-P.

**References**

1. S.W. Hawking, *Nature* **248**, 30 (1974); *Comm. Math. Phys.* **43**, 199 (1975); *Phys. Rev.* **D14**, 2460 (1976).
2. W.G. Unruh, *Phys. Rev.* **D14**, 870 (1976).
3. S. Fulling, *Phys. Rev.* **D7**, 2850 (1973): D.G. Boulware, *Phys. Rev.* **D11**, 1404 (1975); *Phys. Rev.* **D13**, 2169 (1976): L. Parker, *Phys. Rev.* **D12**, 1519 (1976): N.D. Birrell and P.C.W. Davies, *Quantum fields in curved space* (Cambridge University Press, 1982).
4. R. Haag, Berlin, Germany: Springer (1992) 356 p. (Texts and Monographs in Physics).
5. L. C. B. Crispino, A. Higuchi and G. E. A. Matsas, *Rev. Mod. Phys.* **80**, 787 (2008).
6. T. D. Lee, *Nucl. Phys.* **B264**, 437 (1986). R. Friedberg, T. D. Lee and Y. Pang, *Nucl. Phys.* **B276**, 549 (1986).
7. T. Ohsaku, Phys. Lett. B **599**, 102 (2004): D. Ebert and V. C. Zhukovsky, Phys. Lett. B **645** (2007) 267.
8. P. Castorina and M. Finocchiaro, J. Mod. Phys. **3** (2012) 1703.
9. A. Dobado, "Spontaneous symmetry breaking and the Unruh effect," in *General Relativity, 1916-2016* (Minkowski Institute Press, Montreal 2017). arXiv:1703.05675 [gr-qc].
10. A. Dobado, "Brout-Englert-Higgs mechanism for accelerating observers," Phys. Rev. D **96** (2017) no. 8, 085009.
11. S. Cortés, Á. Gómez Nicola and J. Morales, *Phys. Rev.* **D94**, no. 11, 116008 (2016).
12. P. Candelas and D. Deutsch, *Proc. Roy. Soc. Lond.* **A354**, 79 (1977).
13. R. Tolman and P. Ehrenfest, *Phys. Rev.* **36**, no. 12, 1791 (1930).
14. D. Kharzeev, K. Tuchin, *Nucl. Phys.* **A753**, 316-334 (2005).
15. T. Padmanabhan, *Phys. Rept.* **406**, 49-125 (2005); *Rep. Prog. Phys.* **73**, 046901 (2010).

# Beginnings of the helicity basis in the $(S,0)\oplus(0,S)$ representations of the Lorentz group

Valeriy V. Dvoeglazov*

*UAF, Universidad Autónoma de Zacatecas*
*Zacatecas, Zac. 98061, México*
**E-mail: valeri@fisica.uaz.edu.mx*
*fisica.uaz.edu.mx/˜valeri/*

We write solutions of relativistic quantum equations explicitly in the helicity basis for $S = 1/2$ and $S = 1$. We present the analyses of relations between Dirac-like and Majorana-like field operators. Several interesting features of bradyonic and tachyonic solutions are presented.

*Keywords*: Helicity Basis; QFT; Majorana.

## 1. Introduction

In Refs. 1, 2 we considered the procedure of construction of the field operators *ab initio* (including for neutral particles). The Bogoliubov-Shirkov method has been used. In the present article we investigate the helicity $h = 1/2$ and $h = 1$ cases in the helicity basis. We look for relations between the Dirac-like field operator and the Majorana-like field operator.

In the first part we refer to the previously found contradiction in the construction of the Majorana-like field operator for spin-1/2. In the 2nd part we analize the Majorana-like field operator in the $(1,0)\oplus(0,1)$ representation. It seems that the calculations in the helicity basis only give mathematically and physically reasonable results.

## 2. Helicity Basis in the $(1/2,0)+(0,1/2)$ Representation

The Dirac equation is:

$$[i\gamma^\mu\partial_\mu - mc/\hbar]\Psi(x) = 0\,, \quad \mu = 0,1,2,3\,. \tag{1}$$

The $\gamma^\mu$ are the Clifford algebra matrices

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2g^{\mu\nu}\,, \tag{2}$$

$g^{\mu\nu}$ is the metric tensor. Usually, everybody uses the definition of the field operator (in Ref. 3) in the pseudo-Euclidean metrics as given *ab initio*. After actions of the Dirac operator on $\exp(\mp ip_\mu x^\mu)$ the 4-spinors ( $u-$ and $v-$ ) satisfy the momentum-space equations: $(\hat{p} - m)u_h(p) = 0$ and $(\hat{p} + m)v_h(p) = 0$, respectively; the $h$ is the polarization index. It is easy to prove from the characteristic equations $Det(\hat{p} \mp m) = (p_0^2 - \mathbf{p}^2 - m^2)^2 = 0$ that the solutions should satisfy the energy-momentum relations $p_0 = \pm E_p = \pm\sqrt{\mathbf{p}^2 + m^2}$ for both $u-$ and $v-$ solutions.

However, the general scheme of construction of the field operator has been presented in Ref. 4. In the case of the $(1/2,0)\oplus(0,1/2)$ representation we have:

$$\begin{aligned}\Psi(x) &= \frac{1}{(2\pi)^3}\int d^4p\,\delta(p^2-m^2)e^{-ip\cdot x}\Psi(p)\\ &= \frac{1}{(2\pi)^3}\sum_h\int d^4p\,\delta(p_0^2-E_p^2)e^{-ip\cdot x}u_h(p_0,\mathbf{p})a_h(p_0,\mathbf{p})\\ &= \frac{1}{(2\pi)^3}\int\frac{d^4p}{2E_p}[\delta(p_0-E_p)+\delta(p_0+E_p)][\theta(p_0)+\theta(-p_0)]e^{-ip\cdot x}\\ &\quad\times\sum_h u_h(p)a_h(p) = \frac{1}{(2\pi)^3}\sum_h\int\frac{d^4p}{2E_p}[\delta(p_0-E_p)+\delta(p_0+E_p)]\\ &\quad\times\left[\theta(p_0)u_h(p)a_h(p)e^{-ip\cdot x}+\theta(p_0)u_h(-p)a_h(-p)e^{+ip\cdot x}\right]\\ &= \frac{1}{(2\pi)^3}\sum_h\int\frac{d^3\mathbf{p}}{2E_p}\theta(p_0)\left[u_h(p)a_h(p)|_{p_0=E_p}e^{-i(E_pt-\mathbf{p}\cdot\mathbf{x})}\right.\\ &\quad\left.+\,u_h(-p)a_h(-p)|_{p_0=E_p}e^{+i(E_pt-\mathbf{p}\cdot\mathbf{x})}\right] \qquad (3)\end{aligned}$$

During the calculations we had to represent $1=\theta(p_0)+\theta(-p_0)$ above in order to get positive- and negative-frequency parts. Moreover, we did not yet assume, which equation this field operator (namely, the $u$-spinor) satisfies, with negative- or positive-mass and/or $p^0=\pm E_p$. In general we should transform $u_h(-p)$ to the $v_h(p)$. The procedure is given below[1,2].

The explicit forms of the 4-spinors are very well known in the spinorial basis:

$$u_\sigma(\mathbf{p}) = \frac{N_\sigma^+}{2\sqrt{m(E_p+m)}}\begin{pmatrix}[E_p+m+\boldsymbol{\sigma}\cdot\mathbf{p}]\phi_\sigma(\mathbf{0})\\ [E_p+m-\boldsymbol{\sigma}\cdot\mathbf{p}]\chi_\sigma(\mathbf{0})\end{pmatrix}, \quad v_\sigma(\mathbf{p})=\gamma^5u_\sigma(\mathbf{p}), \qquad (4)$$

where $\phi_\uparrow(\mathbf{0})=\chi_\uparrow(\mathbf{0})=\begin{pmatrix}1\\0\end{pmatrix}$ and $\phi_\downarrow(\mathbf{0})=\chi_\downarrow(\mathbf{0})=\begin{pmatrix}0\\1\end{pmatrix}$. The transformation to the standard basis is produced with the $(\gamma^5+\gamma^0)/\sqrt{2}$ matrix. The normalizations, projection operators, propagators, dynamical invariants *etc* have been given in Ref. 5, for example.

We should assume the following relation in the field operator (3):

$$\sum_{h=\uparrow\downarrow}v_h(p)b_h^\dagger(p)=\sum_{h=\uparrow\downarrow}u_h(-p)a_h(-p)\,. \qquad (5)$$

We need $\Lambda_{\mu\lambda}(p)=\bar v_\mu(p)u_\lambda(-p)$. In the spinorial basis by direct calculations, we find $\Lambda_{\mu\lambda}=-im(\boldsymbol{\sigma}\cdot\mathbf{n})_{\mu\lambda}$, $\mathbf{n}=\mathbf{p}/|\mathbf{p}|$, provided that the normalization was chosen to the mass $m$. The indices $h$ and $\sigma,\mu,\lambda$ are the corresponding polarization indices. However, in the helicity basis with the helicity operator

$$\widehat{h}=\frac{1}{2}\boldsymbol{\sigma}\cdot\widehat{\mathbf{p}}=\frac{1}{2}\begin{pmatrix}\cos\theta & \sin\theta e^{-i\varphi}\\ \sin\theta e^{+i\varphi} & -\cos\theta\end{pmatrix} \qquad (6)$$

the 2-eigenspinors can be defined as follows[6,7]:

$$\phi_{\frac{1}{2}\uparrow} = \begin{pmatrix} \cos\frac{\theta}{2}e^{-i\varphi/2} \\ \sin\frac{\theta}{2}e^{+i\varphi/2} \end{pmatrix}, \quad \phi_{\frac{1}{2}\downarrow} = \begin{pmatrix} \sin\frac{\theta}{2}e^{-i\varphi/2} \\ -\cos\frac{\theta}{2}e^{+i\varphi/2} \end{pmatrix}, \tag{7}$$

for $\pm 1/2$ eigenvalues, respectively.

We can start from the Klein-Gordon equation, generalized for describing the spin-1/2 particles (i.e., two degrees of freedom); $c = \hbar = 1$:

$$(p_0 + \boldsymbol{\sigma}\cdot\mathbf{p})(p_0 - \boldsymbol{\sigma}\cdot\mathbf{p})\phi = m^2\phi\,. \tag{8}$$

If the $\phi_{\uparrow\downarrow}$ spinors are defined by the equation (7) then we can construct the corresponding $u-$ and $v-$ 4-spinors:

$$\begin{aligned} u_\uparrow(\mathbf{p}) &= N_\uparrow^+ \begin{pmatrix} \phi_\uparrow \\ \frac{E_p - p}{m}\phi_\uparrow \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{E_p+p}{m}}\phi_\uparrow \\ \sqrt{\frac{m}{E_p+p}}\phi_\uparrow \end{pmatrix}, \\ u_\downarrow(\mathbf{p}) &= N_\downarrow^+ \begin{pmatrix} \phi_\downarrow \\ \frac{E_p + p}{m}\phi_\downarrow \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{m}{E_p+p}}\phi_\downarrow \\ \sqrt{\frac{E_p+p}{m}}\phi_\downarrow \end{pmatrix}, \end{aligned} \tag{9}$$

$$\begin{aligned} v_\uparrow(\mathbf{p}) &= N_\uparrow^- \begin{pmatrix} \phi_\uparrow \\ -\frac{E_p - p}{m}\phi_\uparrow \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{E_p+p}{m}}\phi_\uparrow \\ -\sqrt{\frac{m}{E_p+p}}\phi_\uparrow \end{pmatrix}, \\ v_\downarrow(\mathbf{p}) &= N_\downarrow^- \begin{pmatrix} \phi_\downarrow \\ -\frac{E_p + p}{m}\phi_\downarrow \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \sqrt{\frac{m}{E_p+p}}\phi_\downarrow \\ -\sqrt{\frac{E_p+p}{m}}\phi_\downarrow \end{pmatrix}, \end{aligned} \tag{10}$$

where the normalization to the unit ($\pm 1$) was now used:

$$\bar{u}_h(\mathbf{p})u_{h'}(\mathbf{p}) = \delta_{hh'}\,, \bar{v}_h(\mathbf{p})v_{h'}(\mathbf{p}) = -\delta_{hh'}\,, \tag{11}$$

$$\bar{u}_h(\mathbf{p})v_{h'}(\mathbf{p}) = 0 = \bar{v}_h(\mathbf{p})u_{h'}(\mathbf{p})\,. \tag{12}$$

The commutation relations may be assumed to be the standard ones[3,4,8,9] (compare with[10])

$$\left[a_h(\mathbf{p}), a_{h'}^\dagger(\mathbf{k})\right]_+ = 2E_p\delta^{(3)}(\mathbf{p}-\mathbf{k})\delta_{hh'}\,, [a_h(\mathbf{p}), a_{h'}(\mathbf{k})]_+ = 0 = \left[a_h^\dagger(\mathbf{p}), a_{h'}^\dagger(\mathbf{k})\right]_+ \tag{13}$$

$$\left[a_h(\mathbf{p}), b_{h'}^\dagger(\mathbf{k})\right]_+ = 0 = \left[b_h(\mathbf{p}), a_{h'}^\dagger(\mathbf{k})\right]_+, \tag{14}$$

$$\left[b_h(\mathbf{p}), b_{h'}^\dagger(\mathbf{k})\right]_+ = 2E_p\delta^{(3)}(\mathbf{p}-\mathbf{k})\delta_{hh'}\,, [b_h(\mathbf{p}), b_{h'}(\mathbf{k})]_+ = 0 = \left[b_h^\dagger(\mathbf{p}), b_{h'}^\dagger(\mathbf{k})\right]_+ . \tag{15}$$

Other details of the helicity basis are given in Refs. 13, 11, 12. However, in this helicity case we construct

$$\Lambda_{\mu\lambda}(p) = \bar{v}_\mu(p)u_\lambda(-p) = i\sigma^y_{\mu\lambda}\,. \tag{16}$$

It is well known that "*particle=antiparticle*" in the Majorana theory[14]. So, in the language of the quantum field theory we should have

$$b_\mu(E_p,\mathbf{p}) = e^{i\varphi}a_\mu(E_p,\mathbf{p})\,. \tag{17}$$

Usually, different authors use $\varphi = 0, \pm\pi/2$ depending on the metrics and on the forms of the 4-spinors and commutation relations, *etc.* The application of the Majorana anzatz leads to the contradiction in the spinorial basis. Namely, it leads to existence of the preferred axis in every inertial system (only $p_y$ survives), thus breaking the rotational symmetry of the special relativity.

Next, we can use another Majorana anzatz $\Psi = \pm e^{i\alpha}\Psi^c$ with usual definitions

$$\mathcal{C} = e^{i\vartheta_c}\begin{pmatrix} 0 & i\Theta \\ -i\Theta & 0 \end{pmatrix}\mathcal{K}\,, \quad \Theta = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \tag{18}$$

Thus, on using $Cu^*_\uparrow(\mathbf{p}) = -iv_\downarrow(\mathbf{p})$, $Cu^*_\downarrow(\mathbf{p}) = +iv_\uparrow(\mathbf{p})$ we come to other relations between creation/annihilation operators

$$a^\dagger_\uparrow(\mathbf{p}) = \pm ie^{-i\alpha}b^\dagger_\downarrow(\mathbf{p})\,, \quad a^\dagger_\downarrow(\mathbf{p}) = \mp ie^{-i\alpha}b^\dagger_\uparrow(\mathbf{p})\,, \tag{19}$$

which may be used instead of (17). In the case of $\alpha = \pi/2$ we have similar relations as in (16), but for positive-energy operators. Due to the possible signs $\pm$ the number of the corresponding states is the same as in the Dirac case that permits us to have the complete system of the Fock states over the $(1/2,0)\oplus(0,1/2)$ representation space in the mathematical sense. Please note that the phase factors may have physical significance in quantum field theories as opposed to the textbook nonrelativistic quantum mechanics, as was discussed recently by several authors. However, in this case we deal with the self/anti-self charge conjugate quantum field operator instead of the self/anti-self charge conjugate quantum states. Please remember that it is the latter that answer for the neutral particles. The quantum field operator contains operators for more than one state, which may be either electrically neutral or charged.

## 3. Helicity Basis in the $(1,0)+(0,1)$ Representation

The solutions of the Weinberg-like equation

$$\left[\gamma^{\mu\nu}\partial_\mu\partial_\nu - \frac{(i\partial/\partial t)}{E}m^2\right]\Psi(x) = 0\,. \tag{20}$$

are found in Refs. 15, 16, 17, 18. Here they are:

$$u_\sigma(\mathbf{p}) = \begin{pmatrix} D^S(\Lambda_R)\xi_\sigma(\mathbf{0}) \\ D^S(\Lambda_L)\xi_\sigma(\mathbf{0}) \end{pmatrix}\,,\; v_\sigma(\mathbf{p}) = \begin{pmatrix} D^S(\Lambda_R\Theta_{[1/2]})\xi^*_\sigma(\mathbf{0}) \\ -D^S(\Lambda_L\Theta_{[1/2]})\xi^*_\sigma(\mathbf{0}) \end{pmatrix} = \Gamma^5 u_\sigma(\mathbf{p}), \tag{21}$$

$$\Gamma^5 = \begin{pmatrix} 1_{3\times3} & 0 \\ 0 & -1_{3\times3} \end{pmatrix}\,, \tag{22}$$

in the "spinorial" representation. The $D^S$ is the matrix of the $(S,0)$ representation of the spinor group $SL(2,c)$.

In the $(1,0)\oplus(0,1)$ representation the procedure of derivation of the creation operators (in the similar way as in the previous Section) leads to somewhat different situation:

$$\sum_{\sigma=0,\pm1} v_\sigma(p)b^\dagger_\sigma(p) = \sum_{\sigma=0,\pm1} u_\sigma(-p)a_\sigma(-p)\,. \tag{23}$$

We obtain

$$b^\dagger_\sigma(p) \equiv 0\,. \tag{24}$$

However, if we return to the original Weinberg equations $[\gamma^{\mu\nu}\partial_\mu\partial_\nu \pm m^2]\Psi_{1,2}(x) = 0$ with the field operators:

$$\Psi_1(x) = \frac{1}{(2\pi)^3}\sum_\mu \int \frac{d^3\mathbf{p}}{2E_p}[u_\mu(\mathbf{p})a_\mu(\mathbf{p})e^{-ip\cdot x} + u_\mu(\mathbf{p})b^\dagger_\mu(\mathbf{p})e^{+ip\cdot x}], \tag{25}$$

$$\Psi_2(x) = \frac{1}{(2\pi)^3}\sum_\mu \int \frac{d^3\mathbf{p}}{2E_p}[v_\mu(\mathbf{p})c_\mu(\mathbf{p})e^{-ip\cdot x} + v_\mu(\mathbf{p})d^\dagger_\mu(\mathbf{p})e^{+ip\cdot x}], \tag{26}$$

we obtain

$$b^\dagger_\mu(p) = [1 - 2(\mathbf{S}\cdot\mathbf{n})^2]_{\mu\lambda}a_\lambda(-p)\,, \tag{27}$$

$$d^\dagger_\mu(p) = [1 - 2(\mathbf{S}\cdot\mathbf{n})^2]_{\mu\lambda}c_\lambda(-p)\,. \tag{28}$$

The application of $\overline{u}_\mu(-p)u_\lambda(-p) = \delta_{\mu\lambda}$ and $\overline{u}_\mu(-p)u_\lambda(p) = [1 - 2(\mathbf{S}\cdot\mathbf{n})^2]_{\mu\lambda}$ prove that the equations are self-consistent (similarly to the consideration of the $(1/2,0)\oplus(0,1/2)$ representation). This situation signifies that in order to construct the Sankaranarayanan-Good field operator (which was used by Ahluwalia, Johnson and Goldman[17]) we need additional postulates. One can try to construct the left- and the right-hand side of the field operator separately each other. In this case the commutation relations may also be more complicated.

Is it possible to apply the Majorana-like anzatz to the $(1,0)+(0,1)$ fields? It appears that in this basis we also come to the same contradictions as before. We have two equations

$$a_\mu(p) = +e^{-i\varphi}[1 - 2(\mathbf{S}\cdot\mathbf{n})^2]_{\mu\lambda}a^\dagger_\lambda(-p)\,, \tag{29}$$

and

$$a^\dagger_\mu(p) = +e^{+i\varphi}[1 - 2(\mathbf{S}^*\cdot\mathbf{n})^2]_{\mu\lambda}a_\lambda(-p)\,. \tag{30}$$

In the basis where $S_z$ is diagonal the matrix $S_y$ is imaginary[6]. So, $(\mathbf{S}^*\cdot\mathbf{n}) = S_x n_x - S_y n_y + S_z n_z$, and $(\mathbf{S}^*\cdot\mathbf{n})^2 \neq (\mathbf{S}\cdot\mathbf{n})^2$ in the case of $S=1$. So, we conclude that there is the same problem in this point, in the aplication of the Majorana-like anzatz, as in the case of spin-1/2. Similarly, one can proceed with (28). What would we have in the basis where all $\mathbf{S}^i_{jk} = -i\epsilon_{ijk}$ are pure imaginary? Finally,

I just want to mention that the attempts of constructing the self/anti-self charge conjugate states failed in Ref. 19. Instead, the $\Gamma^5 S^c_{[1]}-$ self/anti-self conjugate states have been constructed therein.

Now we turn to the helicity basis. The helicity operator in the $(1/2,1/2)$ representation is frequently presented:

$$\frac{(\mathbf{S}^I\cdot\mathbf{p})}{p}=\frac{1}{p}\begin{pmatrix}0&0&0&0\\0&0&-ip^3&ip^2\\0&ip^3&0&-ip^1\\0&-ip^2&ip^1&0\end{pmatrix}\,,\ \frac{(\mathbf{S}\cdot\mathbf{p})}{p}\epsilon^\mu_{\pm1}=\pm\epsilon^\mu_{\pm1}\,,\ \frac{(\mathbf{S}\cdot\mathbf{p})}{p}\epsilon^\mu_{0,0_t}=0\,. \tag{31}$$

However, we are aware about some problems with the chosen basis. The helicity operator is (in the case of $\mathbf{S}^3$ diagonal):

$$\frac{(\mathbf{S}^{II}\cdot\mathbf{p})}{p}=\frac{1}{p}\begin{pmatrix}0&0&0&0\\0&p_z&\frac{p_l}{\sqrt{2}}&0\\0&\frac{p_r}{\sqrt{2}}&0&\frac{p_l}{\sqrt{2}}\\0&0&\frac{p_r}{\sqrt{2}}&-p_z\end{pmatrix}\,. \tag{32}$$

The unitary transformation, Ref. 6, p. 55

$$U=\begin{pmatrix}0&0&0&0\\0&-\frac{1}{\sqrt{2}}&\frac{i}{\sqrt{2}}&0\\0&0&0&1\\0&+\frac{1}{\sqrt{2}}&\frac{i}{\sqrt{2}}&0\end{pmatrix}\,,\ U\frac{(\mathbf{S}^I\cdot\mathbf{p})}{p}U^\dagger=\frac{(\mathbf{S}^{II}\cdot\mathbf{p})}{p} \tag{33}$$

can be perfomed to transfer operators and polarization vectors from one basis to another. The first-basis eigenvectors are:

$$\epsilon^\mu_{+1}=\frac{1}{\sqrt{2}}\frac{e^{i\alpha}}{p}\begin{pmatrix}0\\\frac{-p^1p^3+ip^2p}{\sqrt{(p^1)^2+(p^2)^2}}\\\frac{-p^2p^3-ip^1p}{\sqrt{(p^1)^2+(p^2)^2}}\\\sqrt{(p^1)^2+(p^2)^2}\end{pmatrix}\,,\ \epsilon^\mu_{-1}=\frac{1}{\sqrt{2}}\frac{e^{i\beta}}{p}\begin{pmatrix}0\\\frac{p^1p^3+ip^2p}{\sqrt{(p^1)^2+(p^2)^2}}\\\frac{p^2p^3-ip^1p}{\sqrt{(p^1)^2+(p^2)^2}}\\-\sqrt{(p^1)^2+(p^2)^2}\end{pmatrix} \tag{34}$$

$$\epsilon^\mu_0=\frac{1}{m}\begin{pmatrix}p\\\frac{E}{p}p^1\\\frac{E}{p}p^2\\\frac{E}{p}p^3\end{pmatrix}\,,\quad \epsilon^\mu_{0_t}=\frac{1}{m}\begin{pmatrix}E_p\\p^1\\p^2\\p^3\end{pmatrix}\,. \tag{35}$$

The eigenvectors $\epsilon^\mu_{\pm1}$ are not the eigenvectors of the parity operator $(\gamma_{00}R)$ of this representation. However, the $\epsilon^\mu_{1,0}$, $\epsilon^\mu_{0,0_t}$ are. Surprisingly, the latter have no well-defined massless limit. In order to get the well-known massless limit one should use the basis of the light-front form reprersentation, cf. Ref. 20. We also note that the polarization vectors have relations to the solutions of the $(1,0)\oplus(0,1)$ representation through the Proca equations or the Duffin-Kemmer-Petiau equations.

The corresponding helicity operator of the $(1,0)\otimes(0,1)$ representation is

$$\hat{h}=\begin{pmatrix}(\mathbf{S}_{3\times3}\cdot\mathbf{p}) & 0\\ 0 & (\mathbf{S}_{3\times3}\cdot\mathbf{p})\end{pmatrix} \tag{36}$$

The eigen 3-vectors are[6,7]

$$\phi_{\uparrow}=N\,e^{i\vartheta_{+}}\begin{pmatrix}\frac{1}{2}(1+\cos\theta)e^{-i\varphi}\\ \sqrt{\frac{1}{2}}\sin\theta\\ \frac{1}{2}(1-\cos\theta)e^{+i\varphi}\end{pmatrix},\ \phi_{\downarrow}=N\,e^{i\vartheta_{-}}\begin{pmatrix}-\frac{1}{2}(1-\cos\theta)e^{-i\varphi}\\ \sqrt{\frac{1}{2}}\sin\theta\\ -\frac{1}{2}(1+\cos\theta)e^{+i\varphi}\end{pmatrix} \tag{37}$$

$$\phi_{\rightarrow}=N\,e^{i\vartheta_{0}}\begin{pmatrix}-\sqrt{\frac{1}{2}}\sin\theta\,e^{-i\varphi}\\ \cos\theta\\ \sqrt{\frac{1}{2}}\sin\theta\,e^{+i\varphi}\end{pmatrix}. \tag{38}$$

Finally, some notes concerning with the tachyonic solutions of the Weinberg equations in the $(1,0)\oplus(0,1)$ representation space. While some authors, e.g. Ref. 21, argued recently that the tachyonic energy-momentum relation $E=\pm\sqrt{\mathbf{p}^2c^2-m_0^2c^4}$ may lead to some interpretational problems, we still consider it in this paper. The Weinberg equations $[\gamma^{\mu\nu}\partial_\mu\partial_\nu\pm m^2]\Psi_{1,2}(x)=0$ give us both bradyonic and tachyonic solutions, $E=\pm\sqrt{\mathbf{p}^2c^2\pm m_0^2c^4}$. We present them now in the helicity basis which may help us to overcome the difficulties in constructions of the Majorana(-like) field operators, as shown above. If $\phi_R(\mathbf{0})=\phi_L(\mathbf{0})$, the 6-objects can be normalized to the unit. The solutions of $[\gamma^{\mu\nu}p_\mu p_\nu-m^2]\Psi(x)=0$ are

$$u_{\uparrow}(\mathbf{p})=\frac{1}{\sqrt{2}}\begin{pmatrix}\frac{m}{E_p-p}\chi_{\uparrow}\\ \frac{E_p-p}{m}\chi_{\uparrow}\end{pmatrix},\ u_{\downarrow}(\mathbf{p})=\frac{1}{\sqrt{2}}\begin{pmatrix}\frac{m}{E_p+p}\chi_{\downarrow}\\ \frac{E_p+p}{m}\chi_{\downarrow}\end{pmatrix},\ u_{\rightarrow}(\mathbf{p})=\frac{1}{\sqrt{2}}\begin{pmatrix}\chi_{\rightarrow}\\ \chi_{\rightarrow}\end{pmatrix}. \tag{39}$$

In the case of tachyonic solutions ($E_p<p$) we shall be no able to normalize to 1. However, it is possible to normalize to $-1$. In this case we have in the helicity basis:

$$U_{\uparrow}(\mathbf{p})=\frac{1}{\sqrt{2}}\begin{pmatrix}\frac{E_p+p}{m}\chi_{\uparrow}\\ -\frac{m}{E_p+p}\chi_{\uparrow}\end{pmatrix},\ U_{\downarrow}(\mathbf{p})=\frac{1}{\sqrt{2}}\begin{pmatrix}\frac{m}{E_p+p}\chi_{\downarrow}\\ -\frac{E_p+p}{m}\chi_{\downarrow}\end{pmatrix},\ U_{\rightarrow}(\mathbf{p})=\frac{1}{\sqrt{2}}\begin{pmatrix}\chi_{\rightarrow}\\ -\chi_{\rightarrow}\end{pmatrix}. \tag{40}$$

Nevertheless, self/anti-self charge-conjugated 6-objects have not been constructed till now.

## 4. Conclusions

We conclude that the calculations in the helicity basis may be useful to give mathematically and physically reasonable results when dealing with the Majorana(-like) particles.

## Acknowledgments

I acknowledge discussions with colleagues at recent conferences. I am grateful to the Zacatecas University for professorship.

## References


1. V. V. Dvoeglazov, Hadronic J. Suppl. **18**, 239 (2003).
2. V. V. Dvoeglazov, Int. J. Mod. Phys. B**20**, 1317 (2006).
3. C. Itzykson and J.-B. Zuber, *Quantum Field Theory* (McGraw-Hill Book Co., 1980). Sec. 3-3.
4. N. N. Bogoliubov and D. V. Shirkov, *Introduction to the Theory of Quantized Fields.* 2nd Edition. (Nauka, Moscow, 1973).
5. L. H. Ryder, *Quantum Field Theory.* (Cambridge University Press, Cambridge, 1985).
6. D. A. Varshalovich, A. N. Moskalev and V. K. Khersonskiĭ, *Quantum Theory of Angular Momentum* (World Scientific, Singapore, 1988), §6.2.5.
7. V. V. Dvoeglazov, Fizika B**6**, 111 (1997).
8. S. Weinberg, *The Quantum Theory of Fields. Vol. I. Foundations.* (Cambridge University Press, Cambridge, 1995).
9. W. Greiner, *Field Quantization.* (Springer, 1996).
10. Z. Tokuoka, Prog. Theor. Phys. **37**, 603 (1967).
11. H. M. Rück and W. Greiner, J. Phys. G: Nucl. Phys. **3**, 657 (1977).
12. V. V. Dvoeglazov, Int. J. Theor. Phys. **43**, 1287 (2004).
13. V. B. Berestetskiĭ, E. M. Lifshitz and L. P. Pitaevskiĭ, *Quantum Electrodynamics.* (Pergamon Press, 1982, translated from the Russian), §16.
14. E. Majorana, Nuovo Cim. **14**, 171 (1937).
15. A. Sankaranarayanan and R. H. Good, Jr., Nuovo Cim. **36**, 1303 (1965).
16. Yu. V. Novozhilov, *Introduction to Elementary Particle Physics* (Pergamon Press, 1975).
17. D. V. Ahluwalia, M. B. Johnson and T. Goldman, Phys. Lett. B**316**, 102 (1993).
18. V. V. Dvoeglazov, Int. J. Theor. Phys. **37**, 1915 (1998).
19. D. V. Ahluwalia, Int. J. Mod. Phys. A**11**, 1855 (1996).
20. D. V. Ahluwalia and M. Sawicki, Phys. Rev. D**47**, 5161 (1993); Phys. Lett. B**335**, 24 (1994).
21. E. Kapuścik, in *Relativity, Gravitation, Cosmology: Beyond Foundations.* Ed. by V. V. Dvoeglazov. (Nova Science Pubs., Hauppauge, NY, USA, 2018).

# Sidetracked inflation

Sebastian Garcia-Saenz,† Sébastien Renaux-Petel and John Ronayne

*Sorbonne Université, UPMC Paris 6 and CNRS, UMR 7095, Institut d'Astrophysique de Paris, GReCO, Paris 75014, France*
†*E-mail: sebastian.garcia-saenz@iap.fr*

Scalar fields during inflation naturally define a curved field space through their kinetic interactions. When the curvature of this space is negative, the geodesic deviation of the inflationary trajectory may overcome the stabilizing force of the potential, and drive the system away from its original path in field space into a new attractor solution — a scenario known as *sidetracked inflation*. This note presents a brief introduction to the proposal along with a review of its main features and predictions. In particular, we highlight the strongly non-geodesic nature of the sidetracked inflationary phase and how this relates to some of the hallmarks of the model, such as the possibility of realizing inflation with steep potentials and a transient instability of perturbations that leads to unique observational signatures.

## 1. Introduction and set-up

Inflationary models with multiple scalar fields are attractive for a number of reasons. String theory for instance generically predicts a large number of fields in the form of moduli that arise upon compactifying to four spacetime dimensions. While single-field models of inflation are in strong tension with quantum gravity constraints, namely the eta problem and the swampland conjectures, multi-field theories can in principle be made consistent with such requirements. The presence of multiple scalars naturally leads one to consider non-linear sigma models of inflation characterized by a curved target space, a class of theories that can give rise to very interesting dynamical behaviors, such as strongly non-geodesic motion, non-standard inflationary attractors, and the phenomenon of geometrical destabilization.

Motivated by these considerations, we scrutinize a class of inflationary models defined by the action

$$S = \int d^4x\sqrt{-g}\left[\frac{M_P^2}{2}\,R[g] - \frac{1}{2}\,G_{IJ}(\phi)\nabla^\mu\phi^I\nabla_\mu\phi^J - V(\phi)\right]. \tag{1}$$

The gravitational part is the Einstein–Hilbert term for the metric tensor $g_{\mu\nu}$, and $\phi^I$ denotes the multiplet of scalar fields. The matrix $G_{IJ}$ has the natural interpretation of a metric in the internal *field space* in which the scalars $\phi^I$ perform as coordinates. Lastly $V(\phi)$ is the potential term and is a priori an arbitrary function of all the fields. An inflationary state is defined by a flat FLRW spacetime metric with scale factor $a(t)$ and Hubble parameter $H(t) \equiv \dot{a}/a$ (with $t$ the cosmic time), and by spatially homogeneous fields $\bar{\phi}^I(t)$. The latter define the *background trajectory* in the field space, which is in general not a geodesic due to the effects of gravity and the force from the potential.

As advertised, we are interested in the case where the field space has a non-trivial curvature. This means that there exists no choice of fields' parametrization for which $G_{IJ} = \delta_{IJ}$, and for this reason we refer to this class of models as *non-canonical*. Note that non-canonical models are necessarily multi-field as there is no intrinsic curvature in one dimension. The metric $G_{IJ}$, being a function of the fields, induces interactions among the scalars given by higher (mass) dimension operators,

$$G_{IJ}(\phi) = \delta_{IJ} + \frac{1}{M}\, a^{(1)}_{IJ,K}\phi^K + \frac{1}{M^2}\, a^{(2)}_{IJ,KL}\phi^K\phi^L + \cdots\,, \tag{2}$$

where the coefficients $a^{(n)}$ are dimensionless, and $M$ is the typical energy scale that characterizes the interactions. This implies that the curvature of the field space is of order $1/M^2$. We assume that $H \ll M \ll M_P$ so that the "new physics" associated to the scale $M$ may in principle have sizable effects on the physics of inflation.

Particularly interesting is the situation where the field space curvature is *negative*. Theories with a negatively curved target space are of prime importance in top-down constructions of inflation within supergravity, to wit the $\alpha$-attractor class of models. The geometrical destabilization of inflation alluded to before also relies on having a negative curvature, and which serves as a physical motivation for the scenario of sidetracked inflation that we introduce and review below. To this end, we consider linear perturbations about an inflationary solution of the action (1), parametrized by variables $Q^I$ corresponding to gauge invariant combinations of the scalar and metric perturbations (and they coincide with the scalar field fluctuations $\delta\phi^I \equiv \phi^I - \bar{\phi}^I$ in the spatially flat gauge). At this stage we restrict our attention to theories with two fields, as this is already enough to capture most of the interesting multi-field effects. We can then perform a projection of the linearized equation of motion along the adiabatic direction, i.e. the direction tangential to the background trajectory, and the entropic direction, which by definition is orthogonal to the adiabatic one, with the metric $G_{IJ}$ defining the inner product. The two projections of $Q^I$ are denoted by $Q_\sigma$ and $Q_s$ for the adiabatic and entropic perturbations, respectively. At linear level, the adiabatic mode is in fact related to the comoving curvature perturbation $\zeta$ via $Q_\sigma = \frac{\dot{\sigma}}{H}\,\zeta$, where $\dot{\sigma} \equiv \sqrt{G_{IJ}\dot{\bar{\phi}}^I\dot{\bar{\phi}}^J}$ is the speed of the background trajectory. The entropic mode on the other hand satisfies the equation of motion

$$\ddot{Q}_s + 3H\dot{Q}_s + \left(\frac{k^2}{a^2} + m_s^2\right) Q_s = -2\dot{\sigma}\eta_\perp\dot{\zeta}\,, \tag{3}$$

with the definitions

$$\eta_\perp \equiv -\frac{V_{,s}}{H\dot{\sigma}}\,, \qquad m_s^2 \equiv V_{;ss} - H^2\eta_\perp^2 + \epsilon M_P^2 H^2 R_{\rm fs}\,. \tag{4}$$

Here $\eta_\perp$ is known as the *bending parameter*,[a] which quantifies the deviation of the background trajectory from a geodesic path. We say that an inflationary solution

[a] To be more precise, the bending parameter is defined as the turning rate (in units of $H$) of the background trajectory. The relation $\eta_\perp = -\frac{V_{,s}}{H\dot{\sigma}}$ then follows from the background equation of motion rather than as a definition.

features *strongly non-geodesic motion* if $|\eta_\perp| \gg 1$. Lastly, $m_s^2$ corresponds to the effective *entropic mass* of the fluctuations, and crucially depends on the field space Ricci scalar $R_{\rm fs}$.

The interpretation of $m_s^2$ as an effective mass requires some care, as $Q_s$ is actually coupled to $\zeta$. In fact, on super-Hubble scales $k/a \ll H$ it can be shown that (4) reduces to $\ddot{Q}_s + 3H\dot{Q}_s + m_{s(\rm eff)}^2 Q_s = 0$, with $m_{s(\rm eff)}^2 \equiv m_s^2 + 4H^2\eta_\perp^2$. On the other hand, on sub-Hubble scales the perturbations' equations remain coupled and there is no clean notion of an effective mass for either field. Nevertheless, the quantity $m_s^2$ can still serve as a powerful probe of the dynamics of the fluctuations, in particular of their stability, and will play a crucial role in the scenario of inflation that we will describe below.

Returning to the super-Hubble evolution, we observe that the quantity $m_{s(\rm eff)}^2$ is a measure of the stability of large scale perturbations and hence of the background itself. As remarked above, in a negatively curved field space the curvature gives a contribution proportional to $-\epsilon|R_{\rm fs}|$ to the effective mass, which may eventually overcome the stabilizing effect of the potential provided that the slow-roll parameter $\epsilon$ grows fast enough. In a standard picture of inflation where the potential has a flat direction corresponding to the inflaton field, the effective mass $m_{s(\rm eff)}^2$ may thus transition from positive to negative at some critical time when the inflationary state becomes unstable. This is known as the *geometrical destabilization* of inflation[1]. The onset of the instability implies that perturbation theory breaks down and that the homogeneous inflationary vacuum can no longer be trusted. On physical grounds, however, it is plausible that the mayhem will eventually subside and that the system will settle down into a new homogeneous state,[b] i.e., into a second inflationary phase known as *sidetracked inflation*,[3] owing to the way in which the background trajectory deviates from its original path in field space (Fig. 1).

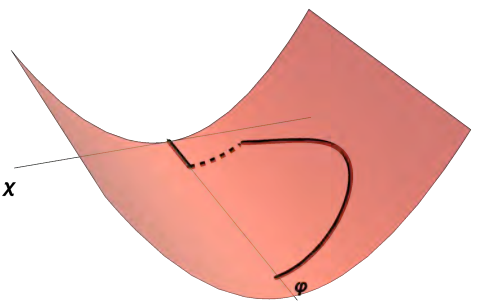



Fig. 1. Schematic plot of the field-space trajectory in sidetracked inflation.

[b] A recent analysis based on the stochastic formalism supports this assumption[2].

## 2. Sidetracked inflation

The story we have described relies on the existence of an attractor solution to the equations of motion that proceeds away from the minimum of the potential. Such an attractor may be heuristically understood as arising from the balance between the stabilizing force of the potential and the destabilizing geodesic deviation caused by the negative curvature of the field space.

For concreteness we focus on two-field models and write $\phi^1 \equiv \varphi$ and $\phi^2 \equiv \chi$. It will be useful to think of $\varphi$ as the inflaton and of $\chi$ as an additional heavy field. The potentials we consider will be of the form $V = \Lambda^4 \mathcal{V}(\varphi) + \frac{m_h^2}{2}\chi^2$, where $\mathcal{V}(\varphi)$ is the dimensionless inflaton potential (the scale $\Lambda$ being determined a posteriori by the normalization of the curvature power spectrum) and $m_h \gg H$ is the mass of $\chi$. Motivated by our goal of understanding the effects of the field space curvature in inflation, we restrict our attention to two simple field space metrics. A first one that we dub "minimal",

$$G^{(\mathrm{min})}_{IJ} d\phi^I d\phi^J = \left(1 + \frac{2\chi^2}{M^2}\right) d\varphi^2 + d\chi^2 \,, \qquad R^{(\mathrm{min})}_{\mathrm{fs}} = -\frac{4}{M^2(1 + 2\chi^2/M^2)} \,, \quad (5)$$

and a second corresponding to a hyperbolic plane metric,

$$G^{(\mathrm{hyp})}_{IJ} d\phi^I d\phi^J = \left(1 + \frac{2\chi^2}{M^2}\right) d\varphi^2 + \frac{2\sqrt{2}\,\chi}{M}\, d\varphi d\chi + d\chi^2 \,, \qquad R^{(\mathrm{hyp})}_{\mathrm{fs}} = -\frac{4}{M^2} \,. \quad (6)$$

The unusual parametrization of the hyperbolic plane given in (6) will be useful in order to compare the effects of the two geometries.

We begin by outlining the salient properties of the background dynamics. First, the sidetracked inflationary phase lasts in general much longer than it would in the absence of a geometrical destabilization. This is in fact common to other non-canonical theories and is due to the effective stretching of the inflaton potential as a result of the effect of the heavy field $\chi$ in the normalization of the inflaton's kinetic term. Second, sidetracked inflation is crucially different from standard slow-roll in that the covariant acceleration of $\chi$ is non-negligible. More explicitly, the equation of motion $D_t\dot{\chi} + 3H\dot{\chi} + V^{,\chi} = 0$ allows for two types of attractors (depending on the parameters of the model): a standard slow-roll attractor characterized by a negligible acceleration $D_t\dot{\chi}$, yielding the usual relation $\dot{\chi} \simeq -V^{,\chi}/3H$; and a "sidetracked attractor" where both $\dot{\chi}$ and $\ddot{\chi}$ are negligible, hence giving the relation $\Gamma^{\chi}_{\varphi\varphi}\dot{\varphi}^2 + V^{,\chi} \simeq 0$ which shows the balance between geometrical and potential effects that we alluded to above. Third, the expressions for the slow-roll parameters in sidetracked inflation are approximately given by

$$\epsilon \simeq \frac{1}{2\sqrt{2}} \frac{m_h}{H} \frac{M|V_{,\varphi}|}{V} \,, \qquad \eta \simeq 3\epsilon - \frac{1}{\sqrt{2}} \frac{m_h}{H} \frac{MV_{,\varphi\varphi}}{|V_{,\varphi}|} \,, \quad (7)$$

which differ from the usual single-field results. Given that $m_h/H$ is large (albeit not too much, for otherwise there will be no destabilization of the standard attractor), we conclude that the quasi-de Sitter approximation is valid provided that the

inflaton potential is flat in units of $M$, which should be contrasted to the single-field case where the potential is required to be flat in Planck units. This is a very welcome feature of the model in view of the swampland conjectures that constrain the derivatives of the potential in the context of inflation and that could rule out large classes of single-field theories.

Next we turn to the behavior of perturbations. A first remark is that the super-Hubble entropic mass $m^2_{s(\mathrm{eff})}$ is always large and positive, thanks to the contribution from the bending parameter $\eta_\perp$, which is large in our set-up. Sidetracked inflation is thus characterized by a strongly non-geodesic motion, and this is indeed at the heart of many of the peculiar properties of the proposal. The large value of $m^2_{s(\mathrm{eff})}$ also implies that entropic (or isocurvature) perturbations quickly decay after horizon crossing as well as the standard conservation law $\dot{\zeta} = 0$ for the adiabatic mode. On the other hand, the sub-Hubble evolution is quite model-dependent, although we can identify three qualitatively distinct behaviors based on the value of the effective mass $m_s^2$, as we now explain (see Ref. 3 for details).

**$|m_s^2| \ll H^2$.** For the inflaton potentials that we have scrutinized, this case arises for the choice of the minimal metric, Eq. (5). Naively, a light entropic mass would suggest strong multi-field effects around the time of horizon crossing. However, the strongly non-geodesic motion gives rise to a new scale $H^2\eta_\perp^2 \gg H^2$ in the problem, allowing for an effective single-field description featuring a non-linear dispersion relation $\omega(k) \propto k^2$ for the relevant degree of freedom. The large value of the bending typically also implies a reduced tensor-to-scalar ratio of order $r \sim \epsilon/\sqrt{\eta_\perp}$ (however with a time of evaluation that occurs later than Hubble crossing).

**$|m_s^2| \gg H^2$ and $m_s^2 > 0$.** This case occurs for the hyperbolic metric, eq. (6), whenever $V_{,\varphi} > 0$. The large entropic mass signifies that an effective theory for the adiabatic mode may be obtained upon integrating out the entropic perturbation. The resulting dispersion relation is given by $\omega^2 = c_s^2 k^2$, where $c_s$ is the effective sound speed, and we find $0 < c_s^2 \ll 1$ for the models we have looked at. This is the usual scenario of single-field inflation with a reduced speed of sound. The tensor-to-scalar ratio is correspondingly suppressed as $r \sim \epsilon c_s$, while non-Gaussianities are enhanced by a factor of $1/c_s^2$.

**$|m_s^2| \gg H^2$ and $m_s^2 < 0$.** This situation also arises when the field space is hyperbolic, but now with $V_{,\varphi} < 0$. The *negative* value of $m_s^2$ signals a tachyonic instability of the entropic mode (again with the caveat that the perturbations are coupled at linear order). Although counterintuitive, this situation allows one to integrate out the tachyonic field in a standard way, with the consequence that the resulting effective theory is characterized by an *imaginary speed of sound* — the low-energy "issue" given by the tachyonic instability has now become a high-energy phenomenon in the form of a gradient instability. This is not a priori catastrophic, as the instability is only transient, lasting a few e-folds around the time of horizon crossing. The predictions in this set-up are quite interesting, notably an exponentially reduced tensor-to-scalar ratio and non-Gaussianities of the flattened shape.[4]

## 3. Final remarks

We conclude this note with a brief discussion of some recent developments concerning the proposal of sidetracked inflation and other closely related topics. The non-standard attractor that defines the sidetracked inflationary phase has been understood to belong to a general class of attractor solutions characterized by a large degree of bending.[5–7] These analyses provide a unifying picture for several novel multi-field theories of inflation in which the curvature of the internal field space plays a crucial role. On the other hand, the possibility of having a transient tachyonic instability — which has been observed in the model of hyperinflation[8] and, as we have seen, in a class of models of sidetracked inflation — has been critically studied in Ref. 9. An important conclusion is that the exponential growth of the perturbations that results from the instability can easily lead to a breakdown of perturbation theory. Avoiding this can only be achieved by a strong restriction of the parameter space of these models, thus serving as a powerful criterion by which various recent proposals can be constrained and eventually ruled out.

## Acknowledgments

The work of SGS and his attendance to the MG15 Meeting were supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013 Grant Agreement no. 307934, NIRG project).

## References

1. S. Renaux-Petel and K. Turzyński, Geometrical Destabilization of Inflation, *Phys. Rev. Lett.* **117**, p. 141301 (2016).
2. O. Grocholski, M. Kalinowski, M. Kolanowski, S. Renaux-Petel, K. Turzyński and V. Vennin, On backreaction effects in geometrical destabilisation of inflation (2019).
3. S. Garcia-Saenz, S. Renaux-Petel and J. Ronayne, Primordial fluctuations and non-Gaussianities in sidetracked inflation, *JCAP* **1807**, p. 057 (2018).
4. S. Garcia-Saenz and S. Renaux-Petel, Flattened non-Gaussianities from the effective field theory of inflation with imaginary speed of sound, *JCAP* **1811**, p. 005 (2018).
5. T. Bjorkmo, The rapid-turn inflationary attractor (2019).
6. P. Christodoulidis, D. Roest and E. Sfakianakis, Attractors, Bifurcations and Curvature in Multi-field Inflation (2019).
7. P. Christodoulidis, D. Roest and E. I. Sfakianakis, Scaling attractors in multi-field inflation (2019).
8. A. R. Brown, Hyperbolic Inflation, *Phys. Rev. Lett.* **121**, p. 251601 (2018).
9. J. Fumagalli, S. Garcia-Saenz, L. Pinol, S. Renaux-Petel and J. Ronayne, Hyper non-Gaussianities in inflation with strongly non-geodesic motion (2019).

# Inflation driven by scalar field and solid matter

Peter Mészáros

*Department of Theoretical Physics, Comenius University, Bratislava, Slovakia*
**E-mail: peter.meszaros@fmph.uniba.sk*

This paper is a brief summary of work already published in Ref. 1. It is inspired by solid inflation which is a cosmological model with inflation driven by fields which enter the Lagrangian in the same way as body coordinates of a solid matter enter the equation of state. We construct a simple generalization of this model by adding a scalar field with standard kinetic term to the action, and focus on calculation of quantities which can be compared with observations of CMB anisotropies.

*Keywords*: Inflation; Solid Matter; Quantum Field Theory in Curved Spacetime.

## 1. Model Under Consideration

One of less standard example of multi-field inflationary models is *solid inflation*[2,3], driven by three-component scalar field $\phi^I$ which enter the Lagrangian in the same way as body coordinates of solid matter enter the equation of state, so that the matter action has to be invariant under internal translations and rotations,

$$\phi^I \to M^I{}_J \phi^J + C^I, \quad M^I{}_J \in SO(3), \quad C^I \in \mathbb{R}^3, \quad I, J = 1, 2, 3, \tag{1}$$

where the capital indices are raised and lowered by the Euclidean metric. The simplest possible background configuration, $\phi^I = \delta^I_i x^i$, with $x^i$ denoting spatial coordinates, breaks the spatial translational and rotational symmetry, but in a flat universe it is invariant under the combined spatial-internal transformations. As shown by Endlich et al.[3], in this model there appears anisotropic dependence of the scalar bispectrum on how the squeezed limit is approached. Further development of the theory includes[4–7].

In this paper we study a combined inflationary model including scalar field $\varphi$ with standard kinetic term and three-component scalar field $\phi^I$ with symmetries defined above. Similar approach can be found in Ref. 8 where the authors study a model with special form of equation of state of the solid but non-trivial coupling of scalar fields to gravity. In our Lagrangian

$$\mathcal{L} = \frac{1}{2} g^{\mu\nu} \partial_\mu \varphi \partial_\nu \varphi + F(\varphi, X, Y, Z), \quad X = \mathrm{Tr}B, \quad Y = \frac{\mathrm{Tr}B^2}{X^2}, \quad Z = \frac{\mathrm{Tr}B^3}{X^3}, \tag{2}$$

there is no non-trivial coupling of scalar fields to gravity, but we keep the form of the equation of state as general as possible, omitting derivative couplings only. Variables $X$, $Y$ and $Z$ are three independent quantities invariant under transformations (1), for which we have adopted definitions from Ref. 3, where $B^{IJ} = -g^{\mu\nu} \partial_\mu \phi^I \partial_\nu \phi^J$ is the body metric. (We have changed its sign in order to reconcile it with the signature of the metric tensor $(+---)$, which we use throughout the paper.) Our model represents a straightforward combination of the solid inflation and the basic

single-field models. It can be considered as, for instance, a simple toy model of interactions of fields driving the solid inflation with fields of an effective field theory of the standard model.

For different functions $F$ there are different solutions of the background equations and a useful quantity measuring the deviation from the de Sitter solution is the slow-roll parameter $\epsilon = -\dot{H}/H^2$, with $H = \dot{a}/a$, which can be written as

$$\epsilon = p + q - \frac{1}{3}pq, \quad p = \frac{\dot{\varphi}^2}{2M_{\rm Pl}^2 H^2}, \quad q = X\frac{F_X}{F}, \tag{3}$$

where $p$ and $q$ are the slow-roll parameters of the single-field inflation and the solid inflation respectively. In our combined model we have an additional degree of freedom, so that the slow-roll parameter can be small also for finite values of the parameters $p$ and $q$; however, in this work we restrict ourselves to the special case in which both $p$ and $q$ are small.

## 2. Cosmological Perturbations

For the perturbation theory we adopt the technical approach from Ref. 9. In the ADM parametrisation the metric is $ds^2 = N^2dt^2 - h_{ij}(dx^i + N^i dt)(dx^j + N^j dt)$. The flat Friedmann–Robertson–Walker–Lematre metric corresponds to $N = 1$, $N^i = 0$, and $h_{ij} = a^2\delta_{ij}$, where $a(t)$ is the scale factor. For analysing perturbations we will use spatially flat slicing gauge in which the scalar and vector perturbations of the three-dimensional metric are set to zero. The metric and matter perturbations are given by

$$\begin{aligned} N &= 1 + \delta N, \quad N^i = \xi_{,i} + N_T^i, \quad h_{ij} = a^2 e^{\gamma_{ij}}, \\ \varphi &= \varphi_0 + \delta\varphi, \quad \phi^I = x^I + \pi^I, \quad \pi^I = \rho_{,I} + \pi_T^I, \end{aligned} \tag{4}$$

where $N^i$ and $\pi^I$ are decomposed into scalar and vector parts (longitudinal and transversal parts in Helmholtz decomposition), $N_T^i$ and $\pi_T^I$ satisfying $N_{T,i}^i = 0$ and $\pi_{T,I}^I = 0$, and $\gamma_{ij}$ is traceless and transversal, i.e. $\gamma_{ii} = 0$ and $\gamma_{ij,j} = 0$. By varying the action consisting of the Einstein–Hilbert part and the matter Lagrangian with respect to $N^i$ and $\delta N$ we obtain the momentum and Hamiltonian constraints. By inserting solutions of constraints into the action and expanding it up to the second order of the perturbation theory we obtain the scalar quadratic action

$$\begin{aligned} S_S^{(2)} = \int \frac{d^3k dt}{(2\pi)^3} a^3 \Big[ & M_{\rm Pl}^2 a^2 H^2 (k^2 - 3a^2H^2Q)Q\dot{\rho}^2 + 2M_{\rm Pl}^2 a^2 H^3 k^2 Q^2 \dot{\rho}\rho \\ & - M_{\rm Pl}^2 H^2 c_L^2 k^4 Q\rho^2 + \frac{1}{2}\dot{\delta\varphi}^2 + \frac{1}{2}\left(F_{\varphi\varphi} - \frac{k^2}{a^2} + 3H^2 p\right)\delta\varphi^2 \\ & - Hp\delta\varphi\dot{\delta\varphi} \pm \sqrt{2}M_{\rm Pl} a^2 H^2 \sqrt{p} Q \left(\frac{k^2}{a^2}\delta\varphi\rho - \dot{\delta\varphi}\dot{\rho}\right) - 2F_{X\varphi}\frac{k^2}{a^2}\delta\varphi\rho \Big], \end{aligned} \tag{5}$$

where only terms up to the first order of the slow-roll approximation have been kept, and $Q = \epsilon - p$ must be positive in order to have the proper sign of action. The tensor quadratic action obtained in the same way is

$$S_T^{(2)} = \frac{1}{4}M_{\rm Pl}^2 \int d^4x a^3 \left( \frac{1}{2}\dot{\gamma}_{ij}\dot{\gamma}_{ij} - \frac{1}{2}a^{-2}\gamma_{ij,k}\gamma_{ij,k} + 2\tilde{h}c_T^2\gamma_{ij}\gamma_{ij} \right), \tag{6}$$

where $\tilde{h} = \dot{H} + \dot{\varphi}_0^2/(2M_{\rm Pl}^2) = H^2(p-\epsilon)$. The longitudinal sound speed appearing in the scalar quadratic action and the transverse sound speed in the tensor action are

$$c_L^2 = 1 + \frac{2}{3}\frac{XF_{XX}}{F_X} + \frac{8}{9}\frac{F_Y + F_Z}{XF_X}, \quad c_T^2 = 1 + \frac{2}{3}\frac{F_Y + F_Z}{XF_X}. \tag{7}$$

Unfortunately, equations of motion for scalar perturbations derived by variation of the action are coupled. Due to the effect of gravity, this occurs even if $F_{X\varphi}$ is zero. They are decoupled only in the special case such that

$$F_{X\varphi} = \pm\frac{\sqrt{p}F_X}{\sqrt{2}M_{\rm Pl}}\left(c_L^2 - 1\right). \tag{8}$$

Furthermore, the scalar perturbation $\rho$ must be replaced by $\mathcal{U} = H\delta u^{(s)} = a^2H(\dot{\rho} - \xi)$, where the $\delta u^{(s)}$ is the scalar part of the solid matter velocity, $u_i^{(s)} = \delta u_{,i}^{(s)} + u_i^{(s)T}$, $u_{i,i}^{(s)T} = 0$, and the term $-\xi$ in the brackets originates from lowering the index with use of the perturbed metric.

In order to solve equations of motion for perturbations it is useful to introduce the conformal time $\tau$ defined in the standard way as $\tau = \int a^{-1}dt$, $\tau \in (-\infty, 0)$. By replacing the cosmological time by it and considering assumptions imposed above including special form of $F_{X\varphi}$ given by (8), we find

$$\mathcal{U}'' - \frac{2 + 2\epsilon_c + \eta_{Q,c} - 2\eta_{L,c}}{\tau}\mathcal{U}' + \left( k^2c_L^2(\tau) + 3\frac{(1 + c_{L,c}^2)Q_c - 2\eta_{L,c}}{\tau^2} \right)\mathcal{U} = 0, \tag{9}$$

$$\delta\varphi_k'' - 2\frac{1 + \epsilon_c}{\tau}\delta\varphi_k' + \left( k^2 - \frac{6p_c + H_c^{-2}F_{\varphi\varphi,c}}{\tau^2} \right)\delta\varphi_k = 0, \tag{10}$$

where the prime denotes the differentiation with respect to the conformal time, $\mathcal{H} = a'/a$, both $\eta_Q = \dot{Q}/(\epsilon H)$ and $\eta_L = \dot{c}_L/(c_L H)$ have been considered to be of the first order in the slow-roll parameters, and the subscript $c$ stands for quantities evaluated at the reference time $\tau_c$ when the longest mode of observational relevance today with the wavenumber $k_{\rm min} \sim H_{\rm today}$ ($a_{\rm today} \equiv 1$) exits the horizon. The creation and annihilation operators defined when the perturbations are quantized obey the canonical commutation relations, and the correctly normalized classical modes obeying equations of motion (9) and (10) are

$$\begin{aligned} \mathcal{U}_k^{(\rm cl)} &= i\frac{\sqrt{\pi}}{2\sqrt{2}}\frac{H_c c_{L,c}}{M_{\rm Pl}\sqrt{Q_c}}(-\tau)^{\frac{3}{2}+\epsilon_c^{(\mathcal{U})}} H^{(1)}_{\frac{3}{2}+p_c^{(\mathcal{U})}}(-c_L(\tau)(1+\eta_{L,c})k\tau), \\ \delta\varphi_k^{(\rm cl)} &= -\frac{\sqrt{\pi}}{2}H_c(-\tau)^{\frac{3}{2}+\epsilon_c} H^{(1)}_{\frac{3}{2}+\epsilon_c^{(\delta\varphi)}}(-k\tau), \end{aligned} \tag{11}$$

where $H^{(1)}$ denotes Hankel functions of the first kind, $\epsilon^{(\mathcal{U})} = \epsilon + \frac{1}{2}\eta_Q - \eta_L$, $p^{(\mathcal{U})} = p - c_L^2 Q + \frac{1}{2}\eta_Q + \frac{5}{2}\eta_L$, and $\epsilon^{(\delta\varphi)} = \epsilon + 2p + \frac{1}{3}\frac{F_{\varphi\varphi}}{H^2}$.

From Ref. 10 we adopt definition of the scalar quantity $\zeta$ that parametrizes the curvature perturbations. In our model in the leading order in slow-roll parameters it is

$$\zeta_k = \frac{A_k}{2} - H\frac{\delta\rho_k}{\dot{\bar{\rho}}} = \frac{\pm\sqrt{p}}{\sqrt{2}M_{\rm Pl}\epsilon}\left(\frac{\delta\dot{\varphi}}{3H} - \delta\varphi\right) + \frac{Q}{c_L^2\epsilon}\left(\frac{\dot{\mathcal{U}}}{3H} + \mathcal{U}\right). \tag{12}$$

The corresponding power spectrum $\mathcal{P}_\zeta(k)$ is defined by the two-point function in the late time limit, $\langle 0|\,\zeta_{k_1}\zeta_{k_2}\,|0\rangle = \left[\mathcal{P}_\zeta(k_1)/(2k_1^3)\right](2\pi)^5\delta^{(3)}(\mathbf{k}_1 + \mathbf{k}_2)$. It is usually approximated by power-law function, $\mathcal{P}_\zeta(k) \propto k^{n_S - 1}$, where $n_S$ is the scalar spectral index, being close to one for a nearly flat spectrum. The spectral tilt up to the leading order of the slow-roll approximation is

$$n_S - 1 = -2\frac{c_{L,e}^5\sigma p_e\epsilon_c^{(\delta\varphi)} + (\epsilon_e - p_e)\,p_c^{(\mathcal{U})}}{\epsilon_e + \left(c_{L,e}^5\sigma - 1\right)p_e}, \quad \sigma = e^{2N_{\min}\left(\epsilon_c^{(\delta\varphi)} - p_c^{(\mathcal{U})}\right)} \tag{13}$$

where the subscript $e$ stands for quantities evaluated in the time when the inflation ends, $\tau_e \approx 0^-$, $N_{\min}$ is the minimal number of e-foldings ($N_{\min} \sim 60$), and $(k_{\max}/k_{\min})^{2\left(p_c^{(\mathcal{U})} - \epsilon_c^{(\delta\varphi)}\right)}$, $k_{\max} \sim 3000 k_{\min}$ being the maximal wavenumber corresponding to the highest observed multipole moment of the cosmic microwave background, and $c_{L,c}^{-2p_c^{(\mathcal{U})}}$ were replaced by one. (For example $3000^{0.01} \doteq 1.08$ and $0.1^{0.01} \doteq 0.98$.)

The tensor perturbations can be treated in a more straightforward way than scalar perturbations, since in contrast to them the tensor perturbations with two independent polarization are decoupled. The tensor spectral tilt in our model is $n_T - 1 = 2c_{L,c}^2\epsilon_c - 2(1 + c_{L,c}^2)p_c$ and tensor-to-scalar ratio is $r = \mathcal{P}_\gamma/\mathcal{P}_\zeta = 4c_L^5\epsilon^2/\left[\epsilon + \left(c_L^5 - 1\right)p\right]$.

## 3. Primordial Non-Gaussianity

In the linear order of the perturbation theory Gaussianity is preserved. Therefore, in order to compute bispectrum which encodes the non-Gaussianity, cubic terms in the action are needed. Since in our model we have $0 \le F_Y + F_Z \le -(3/8)XF_X = (9/8)M_{\rm Pl}^2H^2\,(\epsilon - p)$, $F_Y + F_Z$ can be neglected as well as $F_{XX}$, $F_{XXX}$,...$F_{X\varphi}$, $F_{XX\varphi}$,... However, it must be small, its time derivative may not be, because this restriction is just an inequality, and therefore there are no restrictions on $F_{XY} + F_{XZ}$ and $F_{Y\varphi} + F_{Z\varphi}$. On the other hand, small functions with not small derivative usually do not occur in physical problems, so that it is reasonable to restrict ourselves to the special case in which $F_Y + F_Z = (9/8)AM_{\rm Pl}^2H^2\,(\epsilon - p)$, where $A$ is a constant, or a slowly varying function, of order unity, $A \sim 1$. By differentiating this equation we obtain the restriction $X\,(F_{XY} + F_{XZ}) \mp M_{\rm Pl}\sqrt{p/2}\,(F_{Y\varphi} + F_{Z\varphi}) \sim \epsilon^2$, which is

satisfied if we put

$$\tilde{F} = X\left(F_{XY} + F_{XZ}\right) = \pm\frac{M_{\rm Pl}}{\sqrt{2}}\sqrt{p}\left(F_{Y\varphi} + F_{Z\varphi}\right), \tag{14}$$

and allow $\tilde{F}$ to be of arbitrary order in the slow-roll parameters, so that we have to keep this parameter when deriving the cubic action.

The scalar bispectrum is given by the three-point function of the scalar $\zeta$, which can be computed with the use of the in-in formalism[11] as

$$\langle\zeta_{\mathbf{k}_1}(\tau)\zeta_{\mathbf{k}_2}(\tau)\zeta_{\mathbf{k}_3}(\tau)\rangle = -i\int\limits_{-\infty}^{\tau} a(\tau')d\tau'\,\langle 0|\left[\zeta_{\mathbf{k}_1}(\tau)\zeta_{\mathbf{k}_2}(\tau)\zeta_{\mathbf{k}_3}(\tau), H_{\rm int}(\tau')\right]|0\rangle\,, \tag{15}$$

where only the first order term with a single integration and a simple commutator is considered. Using the relations above we find that the scalar bispectrum $B_\zeta(k_1,k_2,k_3)$, defined by relation $\langle\zeta_{\mathbf{k}_1}\zeta_{\mathbf{k}_2}\zeta_{\mathbf{k}_3}\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B_\zeta(k_1,k_2,k_3)$, consists of two parts,

$$B_\zeta(k_1,k_2,k_3) = F_Y B_\zeta^Y(k_1,k_2,k_3) + \mathcal{N}_\zeta c_{L,c}^2\tilde{F}\tilde{B}_\zeta(k_1,k_2,k_3), \tag{16}$$

parametrized by three independent parameters of the theory, $F_Y$, $\tilde{F}$ and $\mathcal{N}_\zeta c_{L,c}$, where $\mathcal{N}_\zeta$ is a number of of the order of number of $e$-foldings. Due to the delta-function in the tree-point correlation function, three wavenumbers $k_1$, $k_2$ and $k_3$ can be identified with the sides of a triangle, and all information about bispectrum is encoded in a function of two variables which characterize the shape of the triangle. Following conventions of Ref. 12, we define $x = k_2/k_1$ and $y = k_3/k_1$ and describe the bispectrum by the function $x^2y^2B_\zeta(1,x,y)$ defined in region $1-x \le y \le x$, $1/2 \le x \le 1$, $0 \le y \le 1$. Shapes of the functions $x^2y^2B_\zeta^Y(1,x,y)$ and $x^2y^2\tilde{B}_\zeta(1,x,y)$ are depicted in Fig. 1. All functions in the figure are normalized to have value 1 in the equilateral limit, $x = y = 1$.

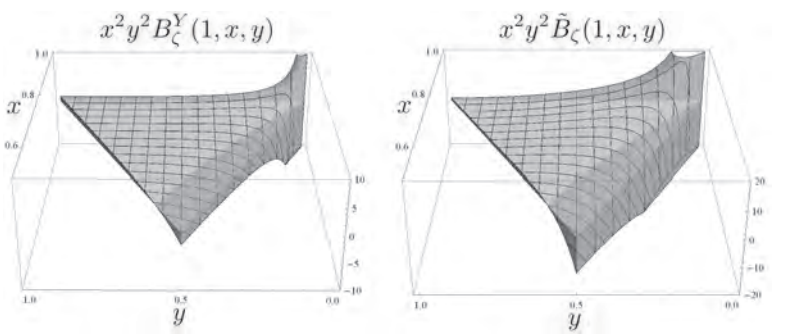



Fig. 1. Shapes of the scalar bispectrum. Flat triangles represent the zero plane.

Our model with the additional degree of freedom allows for a wider range of different shapes of the bispectrum than the solid inflation. The overall bispectrum

peaks in the squeezed limit, unless $\tilde{F}/F_Y = (5/6)\mathcal{N}_\zeta^{-1} c_{L,c}^{-2}$, when it peaks in the equilateral limit instead.

The size of the bispectrum is given by the non-linearity parameter $f_{\rm NL}$. Following the definition (4) in Ref. 13, we find

$$f_{\rm NL} = \frac{5}{72\pi^4}\frac{k^6 B_\zeta(k,k,k)}{\mathcal{P}_\zeta^2(k)} = \frac{\epsilon_c\left(\frac{19415}{13122}\frac{1}{c_{L,c}^2}\frac{F_Y}{F} - \frac{5}{18}\mathcal{N}_\zeta\frac{\tilde{F}}{F}\right)}{\left[\epsilon_c + \left(c_{L,c}^5 - 1\right)p_c\right]^2}. \tag{17}$$

We can see that if $\epsilon - p \sim \epsilon \sim p$, the non-linearity parameter is of the order $f_{\rm NL} \sim (F_Y/F)c_L^{-2}\epsilon^{-1}$, the same as for the solid inflation without the scalar field, or $f_{\rm NL} \sim \mathcal{N}_\zeta(\tilde{F}/F)\epsilon^{-1}$. Supposing that $c_L^5 \sim \epsilon$ we have $f_{\rm NL} \sim (F_Y/F)c_L^{-2}\epsilon^{-3}$ or $f_{\rm NL} \sim \mathcal{N}_\zeta(\tilde{F}/F)\epsilon^{-3}$ if $\epsilon - p$ is of the order $\epsilon^2$. The condition $\epsilon - p \lesssim \epsilon^2$ leading to an amplification of the non-linearity parameter can be rewritten as $q \ll p$, which means that the contribution of the solid matter to the overall stress-energy tensor is negligible in comparison to the contribution of the scalar field.

## Acknowledgments

I would like to thank Vladimír Balek for useful discussions and valuable comments. The work was supported by the grants VEGA 1/0985/16 and UK/36/2017.

## References

1. P. Mészáros, IJMPD **28**, 04, 1950072 (2019).
2. A. Gruzinov, Phys. Rev. **D70**, 063518 (2004).
3. S. Endlich, A. Nicolis, J. Wang, JCAP **1310**, 011 (2013).
4. S. Endlich, B. Horn, A. Nicolis, J. Wang, Phys. Rev. **D90**, 063506 (2014).
5. M. Akhshik, JCAP **1505**, 043 (2015).
6. N. Bartolo, S. Matarrese, M. Peloso and A. Ricciardone, JCAP **1308**, 022 (2013).
7. M. Sitwell, K. Sigurdson, Phys. Rev. **D89**, 123509 (2014).
8. A. Ricciardone, G. Tasinato, Phys. Rev. **D96**, 023508 (2017).
9. J. M. Maldacena, JHEP **0305**, 013 (2003).
10. S. Weinberg: *Cosmology*, Oxford University Press (2008).
11. S. Weinberg, Phys. Rev. **D72**, 043514 (2005).
12. D. Babich, P. Creminelli, M. Zaldarriaga, JCAP **0408**, 009 (2004).
13. P. Creminelli, A. Nicolis, L. Senatore, M. Tegmark, M. Zaldarriaga, JCAP **0605**, 004 (2006).

# Einstein anomaly for vector and axial-vector fields in six-dimensional curved space

Satoshi Yajima*

*Department of Physics, Faculty of Advanced Science and Technology, Kumamoto University, 2-39-1 Kurokami Chuo-ku, Kumamoto 860-8555, Japan*
**E-mail: yajima-s@kumamoto-u.ac.jp*

Kohei Eguchi and Makoto Fukuda

*Department of Physics, Graduate School of Science and Technology, Kumamoto University, 2-39-1 Kurokami Chuo-ku, Kumamoto 860-8555, Japan*

By applying the covariant Taylor expansion method of the heat kernel, Einstein anomaly associated with the Weyl fermion of spin-$\frac{1}{2}$ interacting with nonabelian vector and axial-vector fields in six-dimensional curved space are manifestly given. From the relation between Einstein and Lorentz anomalies, which are the gravitational anomalies, all terms of the Einstein anomaly should form total derivatives. It is shown before the trace operation of the gamma-matrices that the anomaly is expressed by the form expected.

Motivated by the quantum effects in supergravity, we study gravitational anomalies in higher dimensional curved space. In supergravity coupled with super Yang-Mills theory,[1,2] the Lagrangian contains four-fermion interactions, which are regarded as some two-fermion interactions with bosonic background fields expressed by odd-order tensors. The completely antisymmetric part of the highest order tensor should be rewritten as an axial-vector by contracting its tensor with the Levi-Civita symbol. The (polar-)vector and the axial-vector parts in the two-fermion interactions can be absorbed in the vector and the axial-vector gauge fields. The concrete form of the gravitational anomalies in the model may directly be calculated by using the heat kernel.[3]

The heat kernel $K^{\{d\}}(x,x')$ for a fermion of spin-$\frac{1}{2}$ in $d$ dimensions defined by

$$\frac{\partial}{\partial t}K^{(d)}(x,x';t) = -HK^{(d)}(x,x';t), \tag{1}$$

$$K^{(d)}(x,x';0) = \mathbf{1}|h(x)|^{-\frac{1}{2}}|h(x')|^{-\frac{1}{2}}\delta^{(d)}(x,x'), \tag{2}$$

where $\delta^{(d)}(x,x')$ is the $d$-dimensional invariant $\delta$-function, $\mathbf{1} = \{\delta^A{}_B\}$ the unit matrix for the spinor, and $h = \det h^a{}_\mu$, in which $h^a{}_\mu$ is a vielbein. Here $H$ is the second order differential operator, corresponding to the square of the Dirac operator $\not{D}$ in the case of the fermion $\psi$,

$$H = \not{D}^2 = D_\mu D^\mu + X, \quad \not{D} = \gamma^\mu\nabla_\mu + Y, \quad D_\mu = \nabla_\mu + Q_\mu, \quad Q_\mu = \frac{1}{2}\{\gamma_\mu, Y\},$$

$$X = Z - \nabla_\mu Q^\mu - Q_\mu Q^\mu, \quad \nabla_\mu\psi = \partial_\mu\psi + \frac{1}{4}\,\omega^{ab}{}_\mu\,\gamma_{ab}\,\psi, \quad \gamma_{a_1\cdots a_j} = \gamma_{[a_i}\cdots\gamma_{a_j]},$$

$$Z = \frac{1}{2}\,\gamma^{\mu\nu}\,[\nabla_\mu, \nabla_\nu] + \gamma^\mu\nabla_\mu Y + Y^2, \qquad [D_\mu, D_\nu]\psi = \Lambda_{\mu\nu}\,\psi, \tag{3}$$

where $\omega^{ab}{}_{\mu}$ is the Ricci's coefficient of rotation. When in $d = 2n$ dimensions the fermion interacts with vector and axial-vector fields which do not commute each other, the Dirac operator contains the coupling of these bosons in $Y$,

$$Y = \gamma^{\mu} V_{\mu} + \gamma_{2n+1} \gamma^{\mu} A_{\mu}, \quad V_{\mu} \equiv V_{\mu}^{a} T^{a}, \quad A_{\mu} \equiv A_{\mu}^{a} T^{a}, \quad \gamma_{2n+1} = i^{n} \gamma^{1} \gamma^{2} \cdots \gamma^{2n}. \tag{4}$$

Here the representation matrix $T^a$ of a gauge group, and $V_{\mu}^{a}$ ($A_{\mu}^{a}$) is pure imaginary (real), because of the hermiticity of the Dirac operator. The quantities $Q_{\mu}$, $X$ and $\Lambda_{\mu\nu}$ in (3) are expressed in the following tensorial form,

$$\begin{aligned}
&Q_{\mu} = V_{\mu} - \gamma_{2n+1}\, \gamma_{\mu\rho} A^{\rho}, \qquad F_{\mu\nu} = \partial_{\mu} V_{\nu} - \partial_{\nu} V_{\mu} + [V_{\mu}, V_{\nu}], \\
&X = -\frac{1}{4}\, R + 2(n-1)\, A_{\mu} A^{\mu} - \gamma_{2n+1}^{\prime} A^{\mu}{}_{;\mu} + \gamma^{\mu\nu} \left( \frac{1}{2}\, F_{\mu\nu} + \frac{2n-3}{2}\, [A_{\mu}, A_{\nu}] \right), \\
&\Lambda_{\mu\nu} = \frac{1}{4}\, \gamma^{\rho\sigma} R_{\rho\sigma\mu\nu} + F_{\mu\nu} - [A_{\mu}, A_{\nu}] - 2\, \gamma_{\mu\nu} A_{\rho} A^{\rho} + 2\, \gamma_{[\mu|}{}^{\rho} \{ A_{|\nu]}, A_{\rho} \} \\
&\qquad + 2\, \gamma_{2n+1}\, \gamma_{[\mu|\rho} A^{\rho}{}_{;|\nu]} - 2\, \gamma_{\mu\nu\rho\sigma} A^{\rho} A^{\sigma},
\end{aligned} \tag{5}$$

where $R_{\rho\sigma\mu\nu}$ denotes the curvature tensor, and the semi-colon '$;\mu$' means the Riemannian covariant differentiation $\nabla_{\mu} + V_{\mu}$ with respect to the vector gauge field. The completely antisymmetric product $\gamma_{\mu\nu\rho\sigma}$ of $\gamma$-matrices in the last term of $\Lambda_{\mu\nu}$ is rewritten by $-\epsilon_{\mu\nu\rho\sigma}\gamma_5$ and $-\frac{i}{2}\epsilon_{\mu\nu\rho\sigma\kappa\lambda}\gamma_7\gamma^{\kappa\lambda}$ in 4 and 6 dimensions, respectively.

The differential equation (1) of the heat kernel for the fermion interacting with the general boson fields is not solvable strictly. Therefore the heat kernel is usually calculated by using De Witt's ansatz[4], automatically satisfying (2),

$$K^{(2n)}(x, x'; t) \sim \frac{\Delta^{1/2}(x, x')}{(4\pi t)^{n}} \exp\left( \frac{\sigma(x, x')}{2t} \right) \sum_{q=0}^{\infty} a_q(x, x')\, t^{q}, \tag{6}$$

where $\sigma\,(x, x')$ is a half of square of the geodesic distance between $x$ and $x'$, $\Delta\,(x, x') = |h(x)|^{-1} |h(x')|^{-1} \det\{\nabla_{\mu} \nabla_{\nu'} \sigma\,(x, x')\}$, and $a_q(x, x')$ are bispinors. Note that the metric tensor in curved space is $g_{\mu\nu} = h^{a}{}_{\mu} h^{b}{}_{\nu} \eta_{ab}$ with $\eta_{ab} = -\,\delta_{ab}$ in flat tangent space, and that the coincidence limit of $a_0$ is $\lim_{x' \to x} a_0(x, x') \equiv [a_0](x) = \mathbf{1}$. The products of $\sigma_{;\mu}$ ($\equiv \nabla_{\mu}\sigma$) construct orthonormal bases $|n\rangle$ being the eigenfunctions for $\sigma^{;\nu} D_{\nu}$, and the bispinor $a_q$ can be expanded by the bases,[5]

$$\begin{aligned}
&a_q = \sum_{n=0}^{\infty} |n\rangle\langle n|a_q\rangle = \sum_{n} \frac{(-1)^{n}}{n!} \sigma^{;\mu_1'} \cdots \sigma^{;\mu_n'} \lim_{x \to x'} \left[ D_{(\mu_1} \cdots D_{\mu_n)} a_q \right], \\
&a_q\,(x, x') = \langle 0|a_q\rangle\,(x') - \langle \mu|a_q\rangle\,(x')\, \sigma^{;\mu'}\,(x, x') + \cdots .
\end{aligned} \tag{7}$$

The gravitational anomalies are obtained in the case of a massless Weyl fermion $\psi_L$ in $2n$ dimensions. The formal expressions of two gravitational anomalies, *i.e.* the general coordinate anomaly $\mathcal{A}_{\mu}^{(2n)}$ and the Lorentz anomaly $\mathcal{A}_{\mu\nu}^{(2n)}$, are given from the path integral measure.[6] They are expressed by using the heat kernel

$K^{(2n)}(x,x';t)$ after the Gaussian cut-off regularization,

$$
\begin{aligned}
&D^\mu\langle T_{\mu\nu}\rangle = \mathcal{A}_\nu^{(2n)}, \qquad \langle T_{\mu\nu}\rangle_{\rm A} \equiv \frac{1}{2}(\langle T_{\mu\nu}\rangle - \langle T_{\nu\mu}\rangle) = \mathcal{A}_{\mu\nu}^{(2n)},\\
&\mathcal{A}_\nu^{(2n)}(x) = -\frac{1}{2}\lim_{t\to 0}\lim_{x'\to x}{\rm Tr}\left\{\gamma_{2n+1}\left(D_\nu - D_{\nu'}\right)K^{(2n)}(x,x';t)\right\},\\
&\mathcal{A}_{\mu\nu}^{(2n)}(x) = -\frac{1}{4}\lim_{t\to 0}\lim_{x'\to x}{\rm Tr}\left\{\gamma_{2n+1}\gamma_{\mu\nu}K^{(2n)}(x,x';t)\right\},
\end{aligned}
\tag{8}
$$

where Tr runs over both indices of $\gamma$-matrices and representation matrices of the gauge group. Since these anomalies simultaneously appear and are related to each other, $\mathcal{A}_\nu^{(2n)} = 2D^\mu\mathcal{A}_{\mu\nu}^{(2n)}$,[7] it seems that both general covariance and local Lorentz symmetry break down.

We consider the "pure" general coordinate anomaly $G_\mu$ is given by redefining the energy-momentum tensor density so that the local Lorentz symmetry is preserved,

$$
D^\mu\langle T'_{\mu\nu}\rangle = D^\mu\langle T'_{\mu\nu}\rangle_S = G_\nu^{(2n)} = D^\mu\mathcal{A}_{\mu\nu}^{(2n)} = \frac{1}{2}\mathcal{A}_\nu^{(2n)}, \qquad \langle T'_{\mu\nu}\rangle_A = 0 \tag{9}
$$

with $\langle T'_{\mu\nu}\rangle = \langle T_{\mu\nu}\rangle - \mathcal{A}_{\mu\nu}^{(2n)}$, where $\langle T'_{\mu\nu}\rangle_S$ is the symmetric part of the expectation value of the energy-momentum tensor. The "pure" general coordinate anomaly in (9) is called as the Einstein anomaly. The "pure" Lorentz anomaly is also obtained by redefining the energy-momentum tensor density so that the general covariance is preserved,

$$
\langle T''_{\mu\nu}\rangle = \langle T_{\mu\nu}\rangle - 2\mathcal{A}_{\mu\nu}^{(2n)}, \qquad D^\mu\langle T''_{\mu\nu}\rangle = 0, \qquad \langle T''_{\mu\nu}\rangle_A = -\mathcal{A}_{\mu\nu}^{(2n)}. \tag{10}
$$

In order to perform the concrete calculation in $2n$ dimensions, the Einstein anomaly is rewritten by the expansion coefficients of $a_n$ in (7) and its derivatives,

$$
G_\nu^{(2n)}(x) = -\frac{1}{4(4\pi)^n}{\rm Tr}\left\{\gamma_{2n+1}(2\langle\nu|a_n\rangle - \langle 0|a_n\rangle_{!\nu})(x)\right\}. \tag{11}
$$

where the exclamation mark '$!\nu$' means the modified covariant differentiation $D_\nu$. The anomaly in 4-dimensional curved space had already been derived,[8,9]

$$
\begin{aligned}
G_\nu^{(4)} &= -\frac{1}{64\pi^2}{\rm Tr}\big\{\gamma_5(2\langle\nu|a_2\rangle - \langle 0|a_2\rangle_{!\nu})\big\} = \frac{1}{192\pi^2}{\rm Tr}\,\gamma_5\left(\Lambda_{\mu\nu}X\right)^{!\mu}\\
&= \frac{1}{64\pi^2}{\rm tr}\Big[\epsilon_{\mu\nu\rho\sigma}\Big(\frac{1}{6}R^{\rho\sigma}{}_{\kappa\lambda}F^{\kappa\lambda} - \frac{1}{6}RF^{\rho\sigma} + \frac{1}{3}F^{\rho\sigma;\lambda}{}_\lambda\\
&\quad + \frac{4}{3}\{A_\lambda, A^\rho\}F^{\lambda\sigma} + \frac{8}{3}A^\rho A^\sigma A_\lambda A^\lambda\Big)\\
&\quad - \frac{4}{3}(F_{\mu\nu}A^\sigma{}_{;\sigma} + 2F_{[\mu|\lambda}A^\lambda{}_{;|\nu]}) + 8A_{[\mu}A_\nu A^\sigma{}_{;\sigma]}\Big]^{;\mu},
\end{aligned}
\tag{12}
$$

where "tr" means a trace over the representation matrices of the gauge group. A derivative term in $G_\nu^{(4)}$ before the trace operation of $\gamma$-matrices becomes some terms in tensorial form after the operation, and the Lorentz anomaly $\mathcal{A}_{\mu\nu}^{(4)}$ may easily be given from the resultant form of $G_\nu^{(4)}$ by the relation (9). Such properties of $G_\nu^{(4)}$

is succeeded in the case of $G_\nu^{(6)}$. Indeed, the straightforward calculation gives the concrete form of $G_\nu^{(6)}$ as expected,

$$
\begin{aligned}
G_\nu^{(6)} &= -\frac{1}{256\pi^3}\mathrm{Tr}\big\{\gamma_7(2\langle\nu|a_3\rangle - \langle 0|a_3\rangle_{!\nu})\big\} \\
&= -\frac{1}{256\pi^3}\mathrm{Tr}\Big\{\gamma_7\Big[\frac{1}{6}\Lambda_{\mu\nu}\left(\frac{1}{6}R + X\right)^2 + \frac{1}{45}J_{[\mu}X_{!\nu]} - \frac{1}{60}J_{[\mu!\nu]}X \\
&\quad + \frac{1}{15}\Lambda_{\mu\nu}X_{!\rho}{}^{\rho} + \frac{2}{45}\Lambda_{\mu\nu!\rho}X^{!\rho} + \frac{1}{40}\Lambda_{\mu\nu!\rho}{}^{\rho}X + \frac{1}{180}[\Lambda_{\mu\rho},\Lambda_\nu{}^\rho]X \\
&\quad + \frac{1}{180}R^\rho{}_{[\mu}\Lambda_{\nu]\rho}X + \frac{17}{360}R_{\mu\nu\rho\sigma}\Lambda^{\rho\sigma}X + \frac{1}{36}\Lambda_{\mu\nu}\Lambda_{\rho\sigma}\Lambda^{\rho\sigma} \\
&\quad + \frac{1}{45}\Lambda_{[\mu}{}^\rho\Lambda_{\nu]}{}^\sigma\Lambda_{\rho\sigma} - \frac{1}{90}\Lambda_{\mu\nu!\rho}J^\rho + \frac{1}{45}\Lambda_{\rho[\mu}J_{\nu]}{}^{!\rho}\Big]^{;\mu}\Big\} \\
&= \frac{i}{32\pi^3}\,\mathrm{tr}\Big[\epsilon^{\alpha\beta\gamma\delta\kappa\lambda}\Big\{\frac{1}{96}R_{\mu\nu\alpha\beta}F_{\gamma\delta}F_{\kappa\lambda} + \frac{1}{2304}R_{\mu\nu\alpha\beta}R_{\rho\sigma\gamma\delta}R^{\rho\sigma}{}_{\kappa\lambda} \\
&\quad + \frac{1}{2880}R_{[\mu|\rho\alpha\beta}R_{|\nu]\sigma\gamma\delta}R^{\rho\sigma}{}_{\kappa\lambda}\Big\} + \frac{1}{15}A^{\rho;\sigma}{}_{\sigma[\mu}F_{\nu]\rho} - \frac{1}{15}A^\rho{}_{;\rho\sigma}{}^\sigma F_{\mu\nu} \\
&\quad - \frac{1}{45}A^\rho{}_{;\rho[\mu}F_{\nu]\sigma}{}^{;\sigma} + \frac{1}{90}A_{\rho;\sigma}{}^\sigma F_{\mu\nu}{}^{;\rho} + \frac{1}{45}A_{\rho;\sigma[\mu}F^{\rho\sigma}{}_{;\nu]} - \frac{2}{45}A_\sigma{}^{;\sigma\rho}F_{\mu\nu;\rho} \\
&\quad + \frac{4}{45}A^\alpha{}_{;[\mu}{}^\beta F_{\nu]\alpha;\beta} - \frac{1}{30}A_\sigma{}^{;\sigma}F_{\mu\nu;\rho}{}^\rho + \frac{2}{15}A^\sigma{}_{;[\mu}F_{\nu]\sigma;\rho}{}^\rho - \frac{29}{90}A_{\rho;[\mu}[F_{\nu]\sigma},F^{\rho\sigma}] \\
&\quad - \frac{1}{30}A_{\rho;\sigma}[F^{\rho\sigma},F_{\mu\nu}] + \frac{1}{90}A_\sigma{}^{;\sigma}[F_{[\mu}{}^\rho,F_{\nu]\rho}] + \frac{11}{90}A_{\rho;\sigma}[F_{[\mu}{}^\rho,F_{\nu]}{}^\sigma] \\
&\quad - \frac{1}{45}A_\rho[F_{\mu\nu;\sigma},F^{\rho\sigma}] - \frac{4}{45}A_\rho[F^{\rho\sigma}{}_\sigma,F_{\mu\nu}] - \frac{2}{45}A_\alpha[F_{\beta[\mu},F^{\alpha\beta}{}_{;\nu]}] \\
&\quad + \frac{4}{45}A_\rho[F_{[\mu|\sigma}{}^{;\sigma},F_{|\nu]}{}^\rho] + \frac{1}{180}R_{\mu\nu\alpha\beta}A^{\alpha;\sigma}{}_\sigma{}^\beta + \frac{1}{90}R_{\alpha\beta\rho[\mu}A^\alpha{}_{;\nu]}{}^{\beta\rho} \\
&\quad + \frac{1}{90}R_{\rho[\mu;\nu]}A^{;\rho;\sigma}{}_\sigma - \frac{1}{180}R_{\alpha\beta\mu\nu:\rho}A^{\alpha;\rho\beta} + \frac{1}{90}R_{\rho\sigma;[\mu}A^\rho{}_{;\nu]}{}^\sigma \\
&\quad - \frac{1}{90}R_{[\mu|}{}^{\rho;\sigma}A_{\sigma;|\nu]\rho} - \frac{1}{90}R_{[\mu|\rho;\sigma}{}^\rho A^\sigma{}_{;|\nu]} + \frac{1}{90}R_{[\mu|\sigma;\rho}{}^\rho A^\sigma{}_{;|\nu]} \\
&\quad - \frac{1}{90}R_{\rho[\mu;\nu]\sigma}A^{\rho;\sigma} + \frac{1}{90}R_{\alpha\beta\rho[\mu}R_{\nu]}{}^\alpha A^{\beta;\rho} - \frac{1}{180}R_{\alpha\beta\rho[\mu}R^{\alpha\beta\sigma}{}_{\nu]}A^\rho{}_{;\sigma} \\
&\quad + \frac{1}{360}R_{\alpha\beta\mu\nu;\rho}R^{\alpha\beta\rho\sigma}A_{;\sigma} - \frac{1}{180}R^{\alpha\rho}R_{\alpha\beta\mu\nu;\rho}A^\beta - \frac{1}{360}R^{;\rho}R_{\rho\sigma\mu\nu}A^\sigma \\
&\quad - \frac{1}{180}R^{\rho\alpha;\beta}R_{\alpha\beta\mu\nu}A_\rho + \frac{1}{180}(R_{\alpha\beta\rho[\mu}R^{\alpha\beta\sigma}{}_{\nu]})^{;\rho}A_\sigma + \frac{1}{90}(R_{\alpha\beta\rho[\mu}R_{\nu]}{}^\alpha)^{;\rho}A^\beta \\
&\quad + \frac{1}{36}RF_{\mu\nu}A^\rho{}_{;\rho} - \frac{1}{18}RF_{\rho[\mu}A^\rho{}_{;\nu]} + \frac{1}{36}R_{\mu\nu\alpha\beta}F^{\alpha\beta}A^\sigma{}_{;\sigma} - \frac{1}{45}R^\rho{}_{[\mu}F_{\nu]\rho}A^\sigma{}_{;\sigma} \\
&\quad + \frac{1}{15}R^{\alpha\beta}{}_{\mu\nu}F_{\rho\alpha}A^\rho{}_{;\beta} + \frac{1}{10}R^{\alpha\beta}{}_{\mu\nu}F_{\rho\alpha}A_\beta{}^{;\rho} + \frac{4}{45}R_{\alpha\beta\rho[\mu}F_{\nu]}{}^\alpha A^{\beta;\rho} \\
&\quad + \frac{1}{45}R_{\alpha[\mu}F_{\nu]\beta}A^{\alpha;\beta} + \frac{2}{45}R_{\alpha[\mu}F_{\nu]\beta}A^{\beta;\alpha} - \frac{1}{18}R_{\alpha\beta}F_{\mu\nu}A^{\alpha;\beta} \\
&\quad - \frac{1}{45}R_{\alpha\beta\rho[\mu}F_{\nu]}{}^\alpha A^{\rho;\beta} + \frac{1}{45}R_{\alpha\beta}F^\alpha{}_{[\mu}A^\beta{}_{;\nu]} + \frac{1}{18}R_{\alpha\beta\rho[\mu}F^{\alpha\beta}A^\rho{}_{;\nu]}
\end{aligned}
$$

$$
\begin{aligned}
&+\frac{1}{45}R_{\alpha[\mu}F^{\alpha\beta}{}_{;\nu]}A_\beta+\frac{2}{45}R^{\alpha\beta}{}_{\mu\nu}F_{\alpha\beta;\rho}A^\rho-\frac{4}{45}R^{\alpha\beta}{}_{\mu\nu}F_{\alpha\rho}{}^{;\rho}A_\beta\\
&+\frac{1}{90}R_{\alpha\beta\mu\nu;\rho}F^{\alpha\beta}A^\rho-\frac{1}{15}R_{\alpha[\mu;\nu]}F^{\alpha\beta}A_\beta\\
&-\frac{4}{45}R_{[\mu}{}^{\alpha;\beta}F_{\nu]\alpha}A_\beta+\frac{4}{45}R_{[\mu}{}^{\alpha;\beta}F_{\nu]\beta}A_\alpha\\
&-\epsilon_{\mu\nu\alpha\beta\gamma\delta}\Big\{\frac{1}{15}A^\alpha A^\beta F^{\gamma\delta}{}_{;\rho}{}^{\rho}-\frac{1}{30}[A^\alpha,A^{\beta;\gamma}]F^{\delta\rho}{}_{;\rho}\\
&+\Big(\frac{1}{60}[A^\alpha,A^{\rho;\beta}]-\frac{1}{60}[A^\rho,A^{\alpha;\beta}]+\frac{1}{36}[A^\alpha,A^{\beta;\rho}]\Big)F^{;\gamma\delta}{}_{;\rho}\\
&+\Big(\frac{1}{80}[A^\rho,A^{\alpha;\beta}{}_\rho]-\frac{1}{80}[A^\alpha,A^{\rho;\beta}{}_\rho]+\frac{1}{20}[A^\alpha,A^{\beta;\rho}{}_\rho]\\
&-\frac{1}{144}[A_\rho,A^{\alpha;\rho\beta}]+\frac{1}{144}[A^\alpha,A_\rho{}^{;\rho\beta}]\Big)F^{\gamma\delta}\\
&+\Big(\frac{1}{40}[A^\alpha,A^{\beta;\gamma}{}_\rho]-\frac{1}{72}[A^\alpha,A^\beta{}_{;\rho}{}^\gamma]-\frac{1}{45}[A^\alpha{}_{;\rho},A^{\beta;\gamma}]\Big)F^{\delta\rho}\\
&+\Big(-\frac{1}{90}[A^\rho{}_{;\rho},A^{\alpha;\beta}]+\frac{1}{90}[A^\alpha{}_{;\rho},A^{\rho;\beta}]+\frac{1}{30}A^\alpha{}_{;\rho}A^{\beta;\rho}\Big)F^{\gamma\delta}\\
&+\frac{1}{12}\{A^\alpha,A_\rho\}\{F^{\rho\beta},F^{\gamma\delta}\}+\frac{1}{240}[A^\alpha,A_\rho][F^{\rho\beta},F^{\gamma\delta}]-\frac{19}{60}A^\alpha A^\beta F^{\gamma\rho}F^\delta{}_\rho\\
&+\Big(\frac{1}{9}RA^\alpha A^\beta+\frac{13}{48}R^{\alpha\rho}[A^\beta,A_\rho]+\frac{19}{144}R^{\alpha\beta\rho\sigma}A_\rho A_\sigma\Big)F^{\gamma\delta}\\
&+\frac{5}{9}R^{\alpha\rho}A^\beta A^\gamma F^\delta{}_\rho-\frac{13}{48}R^{\alpha\beta}{}_{\rho\sigma}[A^\gamma,A^\rho]F^{\delta\sigma}+\frac{1}{6}R^{\alpha\beta\rho\sigma}A^\gamma A^\delta F_{\rho\sigma}\\
&-\frac{1}{72}RA^{\alpha;\beta}A^{\gamma;\delta}-\frac{1}{15}R^{\alpha\rho}A^{\beta;\gamma}A^\delta{}_{;\rho}+\frac{1}{18}R^{\alpha\rho}A^{\beta;\gamma}A_\rho{}^{;\delta}\\
&+\Big(-\frac{1}{30}A^{\rho;\sigma}A^{\alpha;\beta}+\frac{1}{40}A^{\alpha;\rho}A^{\beta;\sigma}+\frac{1}{72}A^{\rho;\alpha}A^{\sigma;\beta}\\
&-\frac{1}{30}A^{\alpha;\rho}A^{\sigma;\beta}\Big)R^{\gamma\delta}{}_{\rho\sigma}\Big\}\Big]^{;\mu}+O(A^3)
\end{aligned}
\tag{13}
$$

where $J^\rho\equiv\Lambda^{\sigma\rho}{}_{!\sigma}$. Some total derivative terms in $G^{(6)}_\nu$ before the trace operation yield many terms in tensorial form, by using (5), and the derivation is still in progress. The third order terms of $A$ in (13) are 157 terms, of which some terms contain the vector field strength and the curvature tensor, though these terms may be rewritten by the Bianchi and Jacobi identities and by symmetries of the curvature tensor and the vector field strength.

If all $A_\mu$ are abelian in (12), then $G^{(4)}_\nu$ corresponds to the anomaly in space with torsion, which is originally expressed by the third order antisymmetric tensor. The dual vector of the tensor in four dimensions behaves as the axial-vector.[8] Note that the dual tensor of torsion in six or higher dimensions is the third or higher order antisymmetric tensor. In supergravity, there appear the contributions of the vector, the axial-vector and the third order antisymmetric tensor fields, together, which do not commute. The anomaly with the vector and the axial-vector fields in

six-dimensional space with nonabelian torsion may have the new terms containing the third order torsion tensor.

## References


1. S. Ferrara, F. Gliozzi, J. Scherk and P. van Nieuwenhuizen, *Nucl. Phys.* **B117**, 333 (1976).
2. D. Z. Freedman and J. H. Schwarz, *Phys. Rev.* **D15**, 1007 (1977).
3. J. Schwinger, *Phys. Rev.* **82**, 664 (1951).
4. B. S. DeWitt, *Dynamical Theory of Groups and Fields* (Gordon and Breach, 1965).
5. I. G. Avramidi, *Nucl. Phys.* **B355**, 712 (1991); S. Yajima, Y. Higasida, K. Kawano, and S.-I. Kubota, *Class. Quantum. Grav.* **16**, 1389 (1999).
6. K. Fujikawa, *Phys. Rev.* **D21**, 2848 (1980).
7. K. Fujikawa, M. Tomiya and O. Yasuda, *Z. Phys.* **C28**, 289 (1985).
8. S. Yajima and T. Kimura, *Prog. Theor. Phys.* **74**, 866 (1985).
9. S.-I. Kubota, S. Yajima, Y. Higasida and K. Kawano, *Class. Quantum. Grav.* **18**, 433 (2000).

# Rigidly-rotating quantum thermal states in bounded systems

Victor E. Ambruș*

*Department of Physics, West University of Timișoara,*
*Bd. Vasile Pârvan No. 4, Timișoara, 300223, Romania*
**E-mail: victor.ambrus@e-uvt.ro*

We consider rigidly-rotating thermal states of a massless Klein-Gordon field enclosed within a cylindrical boundary, where Robin boundary conditions (RBCs) are imposed. The connection between the parameter of the RBCs and the energy density and four-velocity expressed in the Landau frame is revealed.

*Keywords*: Klein-Gordon field; Finite temperature field theory; Robin boundary conditions; Landau decomposition.

## 1. Introduction

In quantum field theory, the boundary conditions (b.c.s) are imposed at the level of the field operator $\Phi$, or equivalently, of the quantum modes. The interplay between the b.c. formulation and the ensuing operator expectation values in various states is far from obvious. In this paper, we consider the connection between the choice of b.c.s and the thermal expectation value (t.e.v.) of the stress-energy tensor (SET) operator in rigidly-rotating finite temperature states of the massless Klein-Gordon (KG) field. We show that the free parameter $\Psi$ in the Robin b.c.s (RBCs) can be related to the values of the Landau frame macroscopic four-velocity and energy density on the boundary.

The outline of this paper is as follows. In Sec. 2, the mode solutions of the KG equation inside a cylinder are reviewed. The procedure for constructing t.e.v.s in rigidly-rotating systems is summarised in Sec. 3. The analysis of the SET using the Landau frame decomposition and the connection between the Landau velocity $v$ and $\Psi$ is presented in Sec. 4. Section 5 concludes this paper.

## 2. Rigidly-rotating thermal expectation values

Let $\hat{\Phi}(x)$ be the field operator for a massless, neutral (real) scalar field which is confined within a cylinder of radius $R$, obeying the KG equation:

$$\Box\hat{\Phi}(x) = 0. \tag{1}$$

The mode solutions of Eq. (1) can be obtained as follows:[1]

$$f_j = \frac{N_j}{\sqrt{8\pi^2\omega_j}} e^{-i\omega_j t + ik_j z + im_j\varphi} J_{m_j}(q_j\rho), \tag{2}$$

where $(\rho, \varphi, z)$ are the usual cylindrical coordinates, while $\omega_j > 0$, $k_j$ and $m_j$ are the eigenvalues of the Hamiltonian $\hat{H}$, longitudinal momentum $\hat{P}_z$ and $z$ component of the angular momentum, $\hat{L}_z$. In order to fix the normalisation constant $N_j$, we

evaluate the KG inner product for $f_j$ and $f_{j'}$:

$$\begin{aligned}\langle f_j, f_{j'}\rangle &= i\int_{-\infty}^{\infty} dz \int_0^{\infty} \rho\, d\rho \int_0^{2\pi} d\varphi (f_j^* \partial_t f_{j'} - f_{j'} \partial_t f_j^*) \\ &= \frac{N_j^* N_{j'}(\omega_j + \omega_{j'})}{2\sqrt{\omega_j \omega_{j'}}} e^{i(\omega_j - \omega_{j'})t} \delta_{m_j, m_{j'}} \delta(k_j - k_{j'}) \frac{R}{q_j^2 - q_{j'}^2} \\ &\quad \times \left[ J_{m_j}(q_j R) q_{j'} J'_{m_j}(q_{j'} R) - J_{m_j}(q_{j'} R) q_j J'_{m_j}(q_j R) \right], \end{aligned} \tag{3}$$

where a standard identity involving integrals of Bessel functions was employed.[2] Orthogonality is ensured when the transverse momenta $q_j \to q_{m,\ell}$ are discretised according to the Robin boundary conditions:[3]

$$q_{m,\ell} R J'_m(q_{m,\ell} R) + \Psi J_m(q_{m,\ell} R) = 0, \tag{4}$$

where $\ell = 1, 2, \ldots$ indexes the non-negative solutions of Eq. (4) for fixed $m$ in ascending order, while $\Psi$ is considered to be a constant, real number. It is easy to see that $\Psi = 0$ corresponds to the von Neumann b.c.s $[J'_m(q_{m,\ell} R) = 0]$, while the Dirichlet b.c.s $[J_m(q_{m,\ell} R) = 0]$ can be recovered in the limit $\Psi \to \infty$. Imposing $\langle f_{km\ell}, f_{k'm'\ell'}\rangle = \delta(k - k')\delta_{m,m'}\delta_{\ell,\ell'}$ yields:[3]

$$N_{k,m,\ell} = \frac{q_{m,\ell}\sqrt{2}}{|J_m(q_{m,\ell} R)| \sqrt{q_{m,\ell}^2 R^2 + \Psi^2 - m^2}}. \tag{5}$$

The canonical expansion of the field operator with respect to the modes $f_j$ is:

$$\hat{\Phi}(x) = \sum_j \left[ f_j(x) \hat{a}_j + f_j^*(x) \hat{a}_j^\dagger \right], \tag{6}$$

where the one-particle creation ($\hat{a}_j^\dagger$) and annihilation ($\hat{a}_j$) operators obey the standard commutation relation $[\hat{a}_j, \hat{a}_{j'}^\dagger] = \delta(j, j')$.

## 3. Rigidly-rotating thermal states

We now consider rigidly-rotating thermal states, corresponding to an inverse temperature $\beta$ and an angular velocity $\Omega$. The thermal expectation value (t.e.v.) of an operator $\hat{A}$ is computed using the density operator $\hat{\rho}$ as follows:

$$\langle \hat{A} \rangle_\beta = Z^{-1} \mathrm{tr}(\hat{\rho} \hat{A}), \qquad \hat{\rho} = \exp\left[ -\beta(\hat{H} + \Omega \hat{L}_z) \right], \tag{7}$$

where $Z = \mathrm{tr}(\hat{\rho})$ is the partition function. It can be shown that:[4]

$$\langle \hat{a}_j^\dagger \hat{a}_{j'} \rangle_\beta = \frac{\delta(j, j')}{e^{\beta \widetilde{\omega}_j} - 1}, \qquad \widetilde{\omega}_j = \omega_j - \Omega m_j. \tag{8}$$

Equation (8) is not valid when the co-rotating energy $\widetilde{\omega}_j < 0$, since in this case, the vacuum limit (corresponding to $\beta \to \infty$) yields a non-vanishing value.[5] Moreover, modes with $\widetilde{\omega} < 0$ make infinite contributions to rigidly-rotating t.e.v.s.[1] It is noteworthy that finite quantum corrections can still be computed perturbatively.[6]

It is reasonable to expect that t.e.v.s should stay finite for all values of $\Omega$ provided that $\Omega R < 1$, which requires that $q_{m,\ell} R \geq m$. This property can be ensured only when $\Psi \geq 0$, thus we do not consider negative values of $\Psi$ in this paper.

Starting from the following expressions for the SET operator:[7,8]

$$\hat{T}_{\mu\nu} = \frac{2}{3}\nabla_{(\mu}\hat{\Phi}\nabla_{\nu)}\hat{\Phi} - \frac{1}{3}\hat{\Phi}\nabla_{(\mu}\nabla_{\nu)}\hat{\Phi} - \frac{1}{6}g_{\mu\nu}[(\nabla\hat{\Phi})^2 + \mu^2\hat{\Phi}^2], \tag{9}$$

the t.e.v. of the components of the SET can be obtained using the mode expansion (6) of the field operator. It is convenient to express the results with respect to the tetrad comprised of the vectors $e_{\hat{t}} = \partial_t$, $e_{\hat{\rho}} = \partial_\rho$, $e_{\hat{\varphi}} = \rho^{-1}\partial_\varphi$ and $e_{\hat{z}} = \partial_z$. Using the notation $T_{\hat{\alpha}\hat{\gamma}} \equiv \langle : \hat{T}_{\hat{\alpha}\hat{\gamma}} : \rangle_\beta$, the following results can be obtained:[9]

$$\begin{aligned}
T_{\hat{\alpha}\hat{\gamma}} &= \sum_{m=-\infty}^{\infty}\sum_{\ell=1}^{\infty}\int_{-\infty}^{\infty}\frac{N_{km\ell}^2 dk}{12\pi^2\omega_{km\ell}(e^{\beta\widetilde{\omega}_{km\ell}}-1)}F_{\hat{\alpha}\hat{\gamma}},\\
F_{\hat{t}\hat{t}} &= \left(6\omega^2 + \rho^{-2}m^2 - q^2\right)J_m^2 + q^2 {J'_m}^2,\\
F_{\hat{\varphi}\hat{t}} &= -6\omega\rho^{-1}mJ_m^2,\\
F_{\hat{\rho}\hat{\rho}} &= \left(-3\rho^{-2}m^2 + 3q^2\right)J_m^2 + 2q\rho^{-1}J_m J'_m + 3q^2 {J'_m}^2,\\
F_{\hat{\varphi}\hat{\varphi}} &= \left(5\rho^{-2}m^2 + q^2\right)J_m^2 - 2q\rho^{-1}J_m J'_m - q^2 {J'_m}^2,\\
F_{\hat{z}\hat{z}} &= \left(6k^2 - \rho^{-2}m^2 + q^2\right)J_m^2 - q^2 {J'_m}^2,
\end{aligned} \tag{10}$$

where it is understood that $\omega \equiv \omega_{m,\ell}$ and $q \equiv q_{m,\ell}$, while the Bessel functions and their derivatives take the argument $q_{m,\ell}\rho$. It can be shown that the components of the SET not displayed above vanish for all values of $\rho$, $\beta$ and $\Omega$.

## 4. Landau decomposition

The matrix structure of the SET given in Eq. (10) can be summarised as follows:

$$T_{\hat{\alpha}\hat{\gamma}} = \begin{pmatrix} T_{\hat{t}\hat{t}} & 0 & T_{\hat{t}\hat{\varphi}} & 0 \\ 0 & T_{\hat{\rho}\hat{\rho}} & 0 & 0 \\ T_{\hat{t}\hat{\varphi}} & 0 & T_{\hat{\varphi}\hat{\varphi}} & 0 \\ 0 & 0 & 0 & T_{\hat{z}\hat{z}} \end{pmatrix}. \tag{11}$$

The energy density $E$ and macroscopic four-velocity $u^{\hat{\alpha}}$ can be obtained in the Landau frame by solving the eigenvalue equation $T^{\hat{\alpha}}{}_{\hat{\gamma}}u^{\hat{\gamma}} = -Eu^{\hat{\alpha}}$.[10,11] The physically relevant solution for $E$ reads:

$$E = \frac{1}{2}\left[T_{\hat{t}\hat{t}} - T_{\hat{\varphi}\hat{\varphi}} + \sqrt{(T_{\hat{t}\hat{t}} + T_{\hat{\varphi}\hat{\varphi}})^2 - 4T_{\hat{t}\hat{\varphi}}^2}\right], \tag{12}$$

while the Landau velocity $u^{\hat{\alpha}} = \Gamma(1, 0, v, 0)^T$ can be characterised via:

$$v = -\frac{T_{\hat{t}\hat{\varphi}}}{E + T_{\hat{\varphi}\hat{\varphi}}}, \qquad \Gamma = \frac{1}{\sqrt{1-v^2}}. \tag{13}$$

Further manipulation of the above relations gives:

$$\frac{v}{1+v^2} = -\frac{T_{\hat{t}\hat{\varphi}}}{T_{\hat{t}\hat{t}} + T_{\hat{\varphi}\hat{\varphi}}}. \tag{14}$$

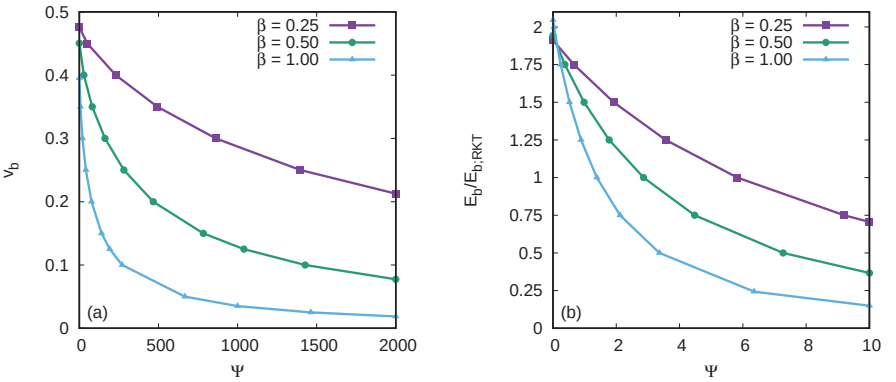



Fig. 1. The dependence of (a) $v_\mathrm{b}$ and (b) $E_\mathrm{b}/E_{\mathrm{b;RKT}}$ on $\Psi$ for various values of the inverse temperature $\beta$ when $\Omega = 0.5$ and $R = 1$.

When Dirichlet b.c.s are employed, it is easy to see that $v$ vanishes on the boundary. For finite values of $\Psi$, $v$ is in general non-vanishing on the boundary. Let $v_\mathrm{b}$ denote the value of $v$ on the boundary. We now ask what is the value of $\Psi$ which ensures $v(R) = v_\mathrm{b}$. Inverting Eq. (14) in order to obtain $\Psi$ as a function of $v_\mathrm{b}$ does not seem feasible. Instead, an iterative procedure can be established which allows the value of $\Psi$ to be computed numerically. Starting from:

$$
\begin{aligned}
T_{\hat{t}\hat{\varphi}}\rfloor_R &= -\sum_{m,\ell}\left(1+\frac{\Psi^2-m^2}{q^2R^2}\right)^{-1}\int_0^\infty dk\frac{2m}{\pi^2R^3(e^{\beta\widetilde{\omega}}-1)},\\
(T_{\hat{t}\hat{t}}+T_{\hat{\varphi}\hat{\varphi}})_R &= \sum_{m,\ell}\left(1+\frac{\Psi^2-m^2}{q^2R^2}\right)^{-1}\int_0^\infty dk\frac{2(m^2+\omega^2R^2+\frac{\Psi}{3})}{\pi^2R^4\omega(e^{\beta\widetilde{\omega}}-1)},
\end{aligned}
\tag{15}
$$

where the sums over $m$ and $\ell$ run between $-\infty$ and $\infty$ and 1 and $\infty$, respectively, it can be seen that $\Psi$ can be isolated from the last term of the second equality above:

$$
\Psi = \frac{3\displaystyle\sum_{m,\ell}\left(1+\frac{\Psi^2-m^2}{q^2R^2}\right)^{-1}\int_0^\infty \frac{dk}{\omega(e^{\beta\widetilde{\omega}}-1)}\left[m\omega R-\frac{v_\mathrm{b}(\omega^2R^2+m^2)}{1+v_\mathrm{b}^2}\right]}{\dfrac{v_\mathrm{b}}{1+v_\mathrm{b}^2}\displaystyle\sum_{m,\ell}\left(1+\frac{\Psi^2-m^2}{q^2R^2}\right)^{-1}\int_0^\infty\frac{dk}{e^{\beta\widetilde{\omega}}-1}}.
\tag{16}
$$

Equation (16) is solved iteratively. The value $\Psi^{(n)}$ corresponding to iteration $n$ is obtained by evaluating the right hand side of Eq. (16) after replacing $\Psi$ with the value $\Psi^{(n-1)}$ obtained at iteration $n-1$, while $v_\mathrm{b}$ is kept fixed at the desired value. Starting from $\Psi^{(0)} = 0$ yields the convergence value within a relatively small number of iterations and the process seems to be stable as long as $v_\mathrm{b}$ can be obtained using $\Psi \geq 0$. To illustrate the procedure, we consider a system with $R = 1$ and $\Omega = 0.5$. Figure 1(a) shows the variation of $v_\mathrm{b}$ with $\Psi$ for $\beta \in \{0.25, 0.5, 1\}$.

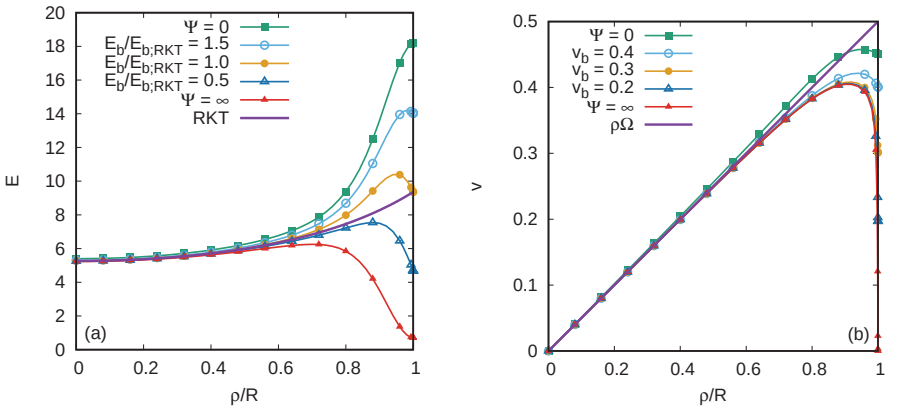



Fig. 2. Profiles of (a) $E$ and (b) $v$ for various values of $\Psi$ at $\beta = \Omega = 0.5$ and $R = 1$. The QFT results are shown using lines and points, while the solid purple line represents the RKT results.

It is natural to consider the relation between the Landau energy density $E_{\rm b}$ measured on the boundary and the energy density $E_{\rm b;RKT}$ expected for a rigidly-rotating Bose-Einstein gas, for which[12]

$$v_{\rm RKT} = \rho\Omega, \qquad E_{\rm RKT} = \frac{\pi^2 \Gamma_{\rm RKT}^4}{30\beta^4}. \tag{17}$$

An iterative scheme for finding $\Psi$ for a prescribed value of $E_{\rm b}$ involves working with quadratic functions with respect to the SET components. The stability and efficiency of such a scheme is questionable. Instead, we employ a bisection algorithm to find the value of $\Psi$ corresponding to $E_{\rm b}$. Typically, the ratio $E_{\rm b}/E_{\rm b;RKT}$ characterising the departure of the quantum state from the expected rigid-rotation profile ranges from $\sim 0.1$ for Dirichlet b.c.s to $\sim 2$ for von Neumann b.c.s. The dependence of $E_{\rm b}/E_{\rm b;RKT}$ on $\Psi$ for $\beta \in \{0.25, 0.5, 1\}$ is illustrated in Fig. 1(b).

Finally, we examine the profiles of the energy density and velocity when $\beta = 0.5$, $\Omega = 0.5$ and $R = 1$. Fig. 2(a) shows that the RBCs interpolate between the Dirichlet and von Neumann b.c.s. In the former case, the energy density exhibits a strong decreasing trend in the vicinity of the boundary, as also remarked in Ref. 1. For the von Neumann b.c.s, the energy density is amplified next to the boundary, as compared to the RKT prediction for a rigidly-rotating Bose-Einstein gas. The velocity $v$, shown in Fig. 2(b), shows small variations with respect to $\Psi$.

## 5. Conclusion

In this paper, a procedure to correlate the parameter $\Psi$ of the RBCs for the massless KG field enclosed within a cylinder and the boundary values of the rigidly-rotating t.e.v. of the SET operator was introduced. The restriction $\Psi \geq 0$ was imposed in order to eliminate modes with negative co-rotating energy $\widetilde{\omega}$ which would otherwise cause t.e.v.s to diverge. Employing the Landau frame decomposition to obtain the

macroscopic four-velocity of the state, the velocity on the boundary was shown to take values between 0 (Dirichlet limit) and a maximum value (von Neumann limit), which increases towards the value corresponding to rigid rotation as the temperature is increased. The energy density is strongly quenched compared to the relativistic kinetic theory prediction for a rigidly-rotating Bose-Einstein gas in the vicinity of the boundary when the Dirichlet b.c.s are employed. By contrast, it is amplified in the case of the von Neumann b.c.s.

## Acknowledgments

This work was supported by a grant of Ministry of Research and Innovation, CNCS-UEFISCDI, project number PN-III-P1-1.1-PD-2016-1423, within PNCDI III.

## References

1. G. Duffy and A. C. Ottewill, Rotating quantum thermal distribution, *Phys. Rev. D* **67**, p. 044002 (2003).
2. F. W. J. Olver, D. W. Lozier, R. F. Boisvert and C. W. Clark, *NIST handbook of mathematical functions* (Cambridge University Press, New York, NY, 2010).
3. A. Romeo and A. A. Saharian, Vacuum densities and zero-point energy for fields obeying robin conditions on cylindrical surfaces, *Phys. Rev. D* **63**, p. 105019 (2001).
4. A. Vilenkin, Quantum field theory at finite temperature in a rotating system, *Phys. Rev. D* **21**, 2260 (1980).
5. V. E. Ambruș and E. Winstanley, Rotating fermions inside a cylindrical boundary, *Phys. Rev. D* **93**, p. 104014 (2016).
6. F. Becattini and E. Grossi, Quantum corrections to the stress-energy tensor in thermodynamic equilibrium with acceleration, *Phys. Rev. D* **92**, p. 045037 (2015).
7. C. G. J. Callan, S. Coleman and R. Jackiw, A new improved energy-momentum tensor, *Ann. Phys.* **59**, 42 (1970).
8. P. B. Groves, P. R. Anderson and E. D. Carlson, Method to compute the stress-energy tensor for the massless spin $\frac{1}{2}$ field in a general static spherically symmetric spacetime, *Phys. Rev. D* **66**, p. 124017 (2002).
9. V. E. Ambruș, Quantum non-equilibrium effects in rigidly-rotating thermal states, *Phys. Lett. B* **771**, 151 (2017).
10. L. D. Landau and E. M. Lifshitz, *Fluid mechanics* (Pergamon Press, Oxford, UK, 1987).
11. L. Rezzolla and O. Zanotti, *Relativistic hydrodynamics* (Oxford University Press, Oxford, UK, 2013).
12. V. E. Ambruș and I. I. Cotăescu, Maxwell-Jüttner distribution for rigidly-rotating flows in spherically symmetric spacetimes using the tetrad formalism, *Phys. Rev. D* **94**, p. 085022 (2016).

# Electric-magnetic duality in the quantum theory

Adrian del Rio*

*Centro de Astrofísica e Gravitação (CENTRA), Departamento de Física, Instituto Superior Técnico (IST), Universidade de Lisboa - Lisboa 1049-001, Portugal*
**E-mail: adriandelrio@tecnico.ulisboa.pt*

It is known that an electric-magnetic duality transformation is a symmetry of the classical source-free Maxwell theory in generic spacetimes. This provides a conserved Noether charge, physically related to the polarization state of the electromagnetic field. We shall argue that this conservation law fails to hold at the quantum level in presence of a background classical gravitational field with non-trivial dynamics, as determined by the Chern-Pontryagin invariant. This is the spin-1 analog of the chiral anomaly for massless Dirac fermions.

*Keywords*: Quantum anomalies; quantum field theory in curved spacetimes; electromagnetic duality.

## 1. The classical electric-magnetic Noether symmetry

It is known from elementary courses in electrodynamics that Maxwell equations without electric charges and currents, as well as the electromagnetic stress-energy tensor, are manifestly invariant under and an "exchange" of the electric $\vec{E}$ and magnetic $\vec{B}$ fields, given by $\vec{E} \to \vec{B}$, $\vec{B} \to -\vec{E}$. In a manifestly covariant language, this amounts to say that

$$\nabla_\mu F^{\mu\nu} = 0, \qquad \nabla_\mu {}^\star F^{\mu\nu} = 0, \tag{1}$$

together with $T_{\mu\nu} = \frac{1}{2}\left[F_{\mu\rho}F^{\rho}{}_{\nu} + {}^\star F_{\mu\rho}{}^\star F^{\rho}{}_{\nu}\right]$, where $F$ denotes the electromagnetic 2-form field and ${}^\star F$ its Hodge-dual, remain invariant under the discrete transformation $F_{\mu\nu} \to {}^\star F_{\mu\nu}$, ${}^\star F_{\mu\nu} \to -F_{\mu\nu}$. Because this discrete transformation is nothing but the Hodge-duality mapping between 2-forms in a 4-dimensional spacetime, this invariance is usually regarded as the duality of Maxwell theory of electromagnetism.

Perhaps more interesting than this is that the invariance actually extends to *continuous* $SO(2)$ rotations between the fields, $F \to F\cos\theta + {}^\star F \sin\theta$, called electric-magnetic transformations. This continuous transformation not only leaves the above equations invariant, but it is also a symmetry of the source-free Maxwell action functional

$$S[A] = \int d^4x\sqrt{-g}F_{\mu\nu}F^{\mu\nu}\,, \tag{2}$$

as first found by Calkin[1]. Here we introduced the electromagnetic potential $A$, defined by $F = dA$. Some years later, this was reanalyzed in greater detail by Deser and Teitelboim[2] using the canonical approach, showing in particular that the symmetry holds even if the electromagnetic field propagates in a curved spacetime background. Noether's theorem enters the scene providing a conserved charge and current, with valuable physical information.

The Noether current takes the form (valid on-shell)[3]

$$j_D^\mu \approx \frac{1}{2}\left[A_\nu {}^\star F^{\mu\nu} - F^{\mu\nu} Z_\nu\right], \tag{3}$$

(where $Z_\nu$ is defined by $E^\mu = -\epsilon^{\mu\alpha\beta}\nabla_\alpha Z_\beta$) and it is conserved on-shell:

$$\nabla_\mu j_D^\mu = -Z_\nu \nabla_\mu F^{\mu\nu} \approx 0. \tag{4}$$

The electric-magnetic symmetry of the theory is generated by a conserved charge $Q_D$, which is the integral on a Cauchy surface of the "zero" component of the previous current:

$$Q_D = -\frac{1}{2}\int_{\Sigma_t} d\Sigma_t (A_\mu B^\mu - E^\mu Z_\mu), \tag{5}$$

and satisfies $\delta H = \{H, Q_D\} \approx 0$. This charge is gauge-invariant, it is a constant of motion $\dot{Q} \approx 0$, and it is not difficult to check that generates the correct transformation rule for the electric and magnetic fields in phase space:

$$\delta E = \{E, Q_D\} = B\,, \qquad \delta B = \{B, Q_D\} = -E\,. \tag{6}$$

In Minkowski spacetime, this charge shows interesting physics. Expanding the fields in Fourier modes with positive $h_+(k)$ and negative $h_-(k)$ helicity amplitudes, the equation above reduces to

$$Q_D = \int_{\mathbb{R}^3} \frac{d^3k}{(2\pi)^3\, k}\left[|h_+(\vec{k})|^2 - |h_-(\vec{k})|^2\right]. \tag{7}$$

As we can see, the Noether charge accounts for the net difference between right- and left-handed circularly polarized radiation (or photons, in the quantum theory). This is recognized as the usual V-stokes parameter, describing the polarization state of electromagnetic radiation. In some contexts this is also known as optical helicity[4]. One concludes that, as a direct consequence of the classical electric-magnetic duality symmetry, this physical magnitude is a constant of motion of source-free electrodynamics.

## 2. The anomaly in the quantum theory

An important question now is whether the symmetry still holds in the quantum theory, i.e. if $\langle \nabla_\mu j_D^\mu \rangle = 0$ for any given quantum vacuum state. Failing to satisfy this will imply that the electric-magnetic transformation is not a quantum symmetry. Given the well-known and diverse examples of anomalous symmetries in quantum field theory, this is not a trivial issue.

### 2.1. *Why is an anomaly expected?*

Anomalies in quantum field theory were first discovered by Adler, Bell, and Jackiw[5] for the Dirac chiral symmetry in the context spinor quantum electrodynamics, and

immediately after by Kimura[6] in the gravitational framework. The underlying reason for their appearance is renormalization. In short, the calculation of expectation values of quadratic field operators runs into problems due to UV divergences. They are ill-defined and as such we need renormalization to extract a well-defined physical quantity from them. The Noether currents are a prototype example of that. The key point is that renormalization subtractions do not necessarily respect the classical equations of motion. Then, quantum fluctuations lead to non-trivial off-shell contributions to the divergence of the current (4), causing the rupture of the classical invariance.

Of course, not every symmetry of a field theory is ruined by quantum fluctuations. One of the key points that makes the appearance of an anomaly very likely in this situation is the fact that the electric-magnetic transformation resembles an ordinary chiral transformation when working with self- and antiself-dual field variables. Indeed, if we define the self-dual fields $\vec{H}_\pm = \frac{1}{\sqrt{2}}\left[\vec{E} \pm i\vec{B}\right]$ then

$$\vec{H}_\pm \to e^{\mp i\theta}\vec{H}_\pm\,, \tag{8}$$

is an electric-magnetic rotation of the fields. Furthermore, under a Lorentz transformation both fields decouple, and their transformation rules are related to the two irreducible representations of the Lorentz group for fields of spin-1, written in standard terminology as $(0,1)$ and $(1,0)$. Compared with massless fermions of spin-1/2, $\vec{H}_+$ is the analog of a right-handed Weyl spinor — that transform under the $(0,1/2)$ Lorentz representation — and $\vec{H}_-$ is the analog of a left-handed Weyl spinor. So there is, indeed, full analogy with the fermion chiral transformation. Since this classical symmetry is known to be anomalous in the quantum theory if the Dirac field propagates in a curved spacetime, it becomes plausible that a similar feature will happen in the electromagnetic case too.

### 2.2. *Derivation of the quantum anomaly*

The strategy that we follow to compute $\langle \nabla_\mu j^\mu_D \rangle$ is to mimic as close as possible the chiral anomaly for Dirac fields.

The previous arguments strongly suggest to use self- and antiself-dual variables. The use of them actually makes the dynamics of the theory significantly more transparent. Source-free Maxwell equations (1) can be rewritten as

$$\alpha_I^{\mu\nu}\nabla_\mu H_+^I = 0\,, \tag{9}$$

with $\alpha_I^{\mu\nu}$ some matrices satisfying the $su(2)$ Lie algebra (again, this resembles the spin-1/2 theory of Weyl spinors). As in the standard formulation, this equation can be solved by introducing potentials. Defining complex potentials by $H_+^I = i\epsilon^{I\mu\nu}\nabla_\mu A_{+\,\nu}$ the equations of motion read now[3]

$$\bar{\alpha}_I^{\mu\nu}\nabla_\mu A_\nu^+ = 0\,. \tag{10}$$

The analogy with the Dirac theory goes even further. The standard Maxwell action functional (2) can be rewritten in terms of complex variables in the following form[3]

$$S[A^+, A^-] = -\frac{1}{4}\int d^4x\sqrt{-g}\ \bar{\Psi}\, i\beta^\mu \nabla_\mu \Psi\,, \tag{11}$$

where

$$\Psi = \begin{pmatrix} A^+ \\ H_+ \\ A^- \\ H_- \end{pmatrix}\,, \qquad \bar{\Psi} = (A^+, H_+, A^-, H_-)\,, \qquad \beta^\mu = i\begin{pmatrix} 0 & 0 & 0 & \bar{\alpha}^\mu \\ 0 & 0 & -\alpha^\mu & 0 \\ 0 & \alpha^\mu & 0 & 0 \\ -\bar{\alpha}^\mu & 0 & 0 & 0 \end{pmatrix}. \tag{12}$$

These new $\beta^\mu$ matrices satisfy the familiar Clifford algebra $\{\beta^\mu, \beta^\nu\} = 2g^{\mu\nu}$, and allow us to define a chiral matrix $\beta_5 = \frac{i}{4!}\epsilon_{\mu\nu\alpha\rho}\beta^\mu\beta^\nu\beta^\alpha\beta^\rho$. Then the chiral-type transformation of the fields

$$\Psi \to e^{i\beta_5\theta}\Psi\,, \tag{13}$$

defines the electric-magnetic rotation, and it is a symmetry of the above action. By applying Noether's theorem we immediately get the current $j^\mu_D = -\frac{1}{4}\bar{\Psi}\beta^\mu\beta_5\Psi$, which agrees with the one presented above.

Now we are equipped with sufficient tools to calculate the vacuum expectation value of (4). We took two different routes to obtain the result, each one illustrating different aspects of the anomaly.

The first method is a direct computation: we identify and regularize UV divergences, and subtract them in a covariant way. This is achieved by writing the divergence of $j^\mu_D$ in terms of the Wightman two point function $S(x,x') = \langle\bar{\Psi}(x)\Psi(x')\rangle$, subtract the corresponding UV asymptotic expansion up to 4 order in derivatives of metric (which contain UV divergences), and then taking the limit in which the two points merge:

$$\langle\nabla_\mu j^\mu_D\rangle_{ren} = \langle\nabla_\mu j^\mu_D\rangle - \langle\nabla_\mu j^\mu_D\rangle_{(4)} = \lim_{\vec{x}\to\vec{x}'\lambda\to 0}\frac{i}{2}\lambda\mathrm{Tr}\beta_5\left(S(x,x',\lambda) - S(x,x',\lambda)_{(4)}\right)\,.$$

The asymptotic expansion is evaluated using the Heat Kernel method[7]. After taking limits a non-vanishing term survives, yielding an anomalous non-conservation of the duality current[3]:

$$\langle\nabla_\mu j^\mu_D\rangle_{ren} = -\frac{\hbar}{96\pi^2}R_{\mu\nu\rho\sigma}{}^\star R^{\mu\nu\rho\sigma}\,. \tag{14}$$

The result comes entirely from the subtraction terms, which are independent of the vacuum state.

The second approach relies on Fujikawa's path integral viewpoint[8,9]. The partition function of the theory,

$$Z = \int d\mu[\Psi,\bar{\Psi}]e^{iS[\Psi,\bar{\Psi}]}\,, \tag{15}$$

remains invariant under a change of variables in phase space. An electric-magnetic duality rotation is a canonical transformation in phase space and as such it must leave $Z$ invariant. Noether's theorem states that the action changes as

$$S[\Psi', \bar{\Psi}'] = S[\Psi, \bar{\Psi}] - \int d^4x \sqrt{-g}\theta(x) \langle \nabla_\mu j^\mu_D \rangle_{ren} \ . \tag{16}$$

If everything else under the integral remained invariant, one would conclude that the current is conserved. However, as first pointed out by Fujikawa, the measure of the integral is not necessarily invariant and could change by a non trivial Jacobian, $d\mu[\Psi', \bar{\Psi}'] = Jd\mu[\Psi, \bar{\Psi}]$, $J \neq 1$. This is what yields the anomalous non conservation of the current in this framework. A detailed analysis following this approach gives back equation (14) again.

## 3. Conclusions and final comments

We argued that quantum fluctuations of the electromagnetic field spoil the classical conservation law of the Noether current associated to electric-magnetic rotations due to spacetime curvature, as calculated in (14). The result points out that $Q_D$ is no longer a constant of motion in the quantum theory: gravitational dynamics should distinguish between right and left photons. Physical implications of this quantum effect are still under investigation, but the analogy with the chiral anomaly [10] suggests that this is related to spontaneous electromagnetic circular polarization, as a result of asymmetric creation of right/left photons from the quantum vacuum due to background dynamics. If confirmed, our next goal will be to study the stimulated contribution to this asymmetric particle creation. This could be of interest in different astrophysical scenarios [11], ranging from stellar gravitational collapse to mergers of compact objects.

## Acknowledgments

The author is grateful to I. Agullo and J. Navarro-Salas for useful comments on the manuscript, and acknowledges financial support provided under the ERC Consolidator Grant "Matter and strong-field gravity: New frontiers in Einstein's theory", no. MaGRaTh-646597.

## References

1. M. G. Calkin, *Am. J. Phys.* **33**, 958 (1965).
2. S. Deser and C. Teitelboim, *Phys. Rev. D***13**, 1592 (1976). S. Deser, *J. Phys. A* **15**, 1053 (1982).
3. I. Agullo, A. del Rio, and J. Navarro-Salas. *Phys. Rev. Lett.*, **118**:111301, (2017); *Phys. Rev. D*, **98**(12):125001, (2018).
4. R. P. Cameron S. M. Barnett and A. M. Yao. *Phys. Rev.* , **A 86**:013845, (2012).
5. S. L. Adler, *Phys. Rev.* **177**, 2426 (1969). J. S. Bell and R. Jackiw, *Nuovo Cimento A* **60**, 47 (1969).

6. T. Kimura *Prog. Theor. Phys.* **42**, 1191 (1969).
7. L. Parker and D.J. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity* (Cambridge University Press, Cambridge, England, 2009).
8. K. Fujikawa, *Phys. Rev. Lett.* **42**, 1195 (1979); *Phys. Rev. D* **21**, 2848 (1980).
9. K. Fujikawa and H. Suzuki, *Path Integrals and Quantum Anomalies*, Oxford University Press, Oxford (2004).
10. N. H. Christ A. S. Blaer and J.-F. Tang. *Phys. Rev. Lett.*, **47**:1364–1367, (1981).
11. I. Agullo, A. del Rio, and J. Navarro-Salas. *Int. J. Mod. Phys. D*, **26**(12):1742001, (2017).

# Adiabatic regularization with a Yukawa interaction

Antonio Ferreiro

*Departamento de Fisica Teorica and IFIC, Universidad de Valencia-CSIC, Facultad de Fisica, Burjassot, 46100, Spain*
**E-mail: antonio.ferreiro@ific.uv.es*

Adrian del Rio

*Centro de Astrofísica e Gravitação - CENTRA, Departamento de Física, Instituto Superior Técnico - IST, Universidade de Lisboa - 1049 Lisboa, Portugal*

Jose Navarro-Salas and Silvia Pla

*Departamento de Fisica Teorica and IFIC, Universidad de Valencia-CSIC, Facultad de Fisica, Burjassot, 46100, Spain*

Francisco Torrenti

*Department of Physics, University of Basel, Klingelbergstr. 82, CH-4056 Basel, Switzerland*

We extend the adiabatic regularization method for an expanding universe to include the Yukawa interaction between a quantized Dirac field and a homogeneous time-dependent scalar field. We present the renormalized semiclassical equations that are needed in order to take into account the backreaction of the produced Dirac fermions in both gravitational and scalar background fields.

*Keywords*: Quantum field theory in curved spacetime, adiabatic regularization, preheating, Yukawa interaction, semiclassical gravity.

## 1. Introduction

Any inflationary model predicts an exponentially fast dilution of the number density of any particle species. To repopulate the Universe with matter and radiation, the energy density of the inflaton needs to be converted to the standard model fields. This epoch is called reheating. A preheating stage, where particles are created due to the classical behavior of the inflaton field before the thermal equilibrium has been proposed for both scalar[1] and fermion fields[2]. In some models, including new inflation and chaotic inflation, inflation is produced due to a classical scalar field, which at the end of inflation it starts to oscillate around the minimum of its potential.

Gravitational particle creation was first discovered in the expanding universe[3–5] and later extended to black holes[4–7]. Time-dependent gauge fields can also create particles[8]. Due to the Yukawa coupling between a time-varying classical scalar field $\Phi$ and a quantum fermionic field $\psi$, spontaneous particle production also occurs. In order to take into account the backreaction of the produced particles on the classical background fields, via the semiclassical equations, one need to compute the expectation values $\langle T_{\mu\nu}\rangle$ and $\langle\bar{\psi}\psi\rangle$. This requires a regularization and

renormalization method to deal with the ultraviolet divergences appearing in expectation values quadratic in fields. A very powerful method to tame the ultraviolet divergences for isotropically expanding universes was originally proposed by Parker and Fulling[4,5,9] for scalar fields. This renormalization method is based on an adiabatic expansion of the field modes living in the expanding universe, and it is known as adiabatic regularization. The adiabatic method has been extended to Dirac fermions in Refs. 10, 11.

In this contribution we further extend the adiabatic regularization method to incorporate the Yukawa interaction, which is expected to be of major importance in the preheating epoch. First, we present the semiclassical equations for fermions in an expanding universe and Yukawa-coupled to a scalar field and then use the adiabatic regularization scheme to cure the divergences of the vacuum expectation values of $\langle T_{\mu\nu}\rangle$ and $\langle\bar{\psi}\psi\rangle$. (For a extended analysis, see Ref. 12).

## 2. Semiclassical equations for a quantized Dirac matter field with Yukawa coupling

We consider the theory defined by the action functional $S = S[g_{\mu\nu}, \Phi, \psi, \nabla\psi]$, where $\psi$ represents a Dirac field, $\Phi$ is a scalar field, and $g_{\mu\nu}$ stands for the spacetime metric. We decompose the action as $S = S_g + S_m$, where $S_m$ is the matter sector

$$S_m = \int d^4x\sqrt{-g}\ \left\{\frac{i}{2}[\bar{\psi}\underline{\gamma}^\mu\nabla_\mu\psi - (\nabla_\mu\bar{\psi})\underline{\gamma}^\mu\psi)] - m\bar{\psi}\psi - g_Y\Phi\bar{\psi}\psi\right\}\ , \qquad (1)$$

and $S_g$ is the gravity-scalar sector. In (1), both the metric $g_{\mu\nu}(x)$ and the scalar field $\Phi(x)$ are regarded as classical external fields. The curved-space Dirac matrices $\underline{\gamma}^\mu$ obey the relations $\{\underline{\gamma}^\mu, \underline{\gamma}^\nu\} = 2g^{\mu\nu}$. The Dirac spinor $\psi(x)$ will be our quantized field, living in a curved spacetime and possessing a Yukawa coupling to the classical field $\Phi$. The Dirac equation is

$$(i\underline{\gamma}^\mu\nabla_\mu - m - g_Y\Phi)\psi = 0\ , \qquad (2)$$

and the stress-energy tensor is given by

$$T^m_{\mu\nu} := \frac{2}{\sqrt{-g}}\frac{\delta S_m}{\delta g^{\mu\nu}} = \frac{i}{2}\left[\bar{\psi}\underline{\gamma}_{(\mu}\nabla_{\nu)}\psi - (\nabla_{(\mu}\bar{\psi})\underline{\gamma}_{\nu)}\psi\right]\ . \qquad (3)$$

The complete theory, including the gravity-scalar sector in the action, can be described by

$$S = S_g + S_m = \frac{1}{16\pi G}\int d^4x\sqrt{-g}R + \int d^4x\sqrt{-g}\left\{\frac{1}{2}g^{\mu\nu}\nabla_\mu\Phi\nabla_\nu\Phi - V(\Phi)\right\} + S_m\ , \qquad (4)$$

where $S_m$ is the action for the matter sector given in (1).

The semiclassical Einstein and scalar equations are

$$G^{\mu\nu} + 8\pi G\left(\nabla^\mu\Phi\nabla^\nu\Phi - \frac{1}{2}g^{\mu\nu}\nabla^\rho\Phi\nabla_\rho\Phi + g^{\mu\nu}V(\Phi)\right) = -8\pi G\langle T^{\mu\nu}_m\rangle_{ren}\ , \qquad (5)$$

$$\Box\Phi + \frac{\partial V}{\partial\Phi} = -g_Y\langle\bar{\psi}\psi\rangle_{ren}\ . \qquad (6)$$

In order to obtain the renormalized quantities $\langle T_{\mu\nu}\rangle_{ren}$ and $\langle\bar{\psi}\psi\rangle_{ren}$ we extend the adiabatic regularization in the next section.

## 3. Fermion quantization and adiabatic regularization

In a spatially flat FLRW spacetime, we expand the field $\psi$ as

$$\psi(x)=\int d^3\vec{k}\sum_{\lambda}\left[B_{\vec{k}\lambda}u_{\vec{k}\lambda}(x)+D^{\dagger}_{\vec{k}\lambda}v_{\vec{k}\lambda}(x)\right]\ , \tag{7}$$

with

$$u_{\vec{k},\lambda}(x)=\frac{e^{i\vec{k}\vec{x}}}{\sqrt{(2\pi)^3a^3(t)}}\begin{pmatrix}h^I_k(t)\xi_\lambda(\vec{k})\\ h^{II}_k(t)\frac{\vec{\sigma}\vec{k}}{k}\xi_\lambda(\vec{k})\end{pmatrix} \tag{8}$$

and $v_{\vec{k},\lambda}(x)=Cu_{\vec{k},\lambda}(x)$. The modes $v_{\vec{k},\lambda}(x)$ can be obtained by applying a charge conjugate transformation $C\psi=-i\gamma^2\psi^*$. Here $\xi_\lambda$ with $\lambda=\pm1$ are two constant orthonormal two-spinors ($\xi^\dagger_\lambda\xi_{\lambda'}=\delta_{\lambda,\lambda'}$), eigenvectors of the helicity operator $\frac{\vec{\sigma}\vec{k}}{2k}\xi_\lambda=\frac{\lambda}{2}\xi_\lambda$. The time-dependent functions $h^I_k$ and $h^{II}_k$ satisfy the first-order coupled equations

$$h^{II}_k=\frac{ia}{k}\left(\frac{\partial h^I_k}{\partial t}+i(m+g_Y\Phi)h^I_k\right)\ ,\qquad h^I_k=\frac{ia}{k}\left(\frac{\partial h^{II}_k}{\partial t}-i(m+g_Y\Phi)h^{II}_k\right)\ . \tag{9}$$

The normalization condition for the above four-spinors reduces to

$$|h^I_k|^2+|h^{II}_k|^2=1\ . \tag{10}$$

For adiabatic regularization one can expand $h^I_k$ and $h^{II}_k$ as

$$\begin{aligned}h^I_k(t)&=\sqrt{\frac{\omega(t)+m}{2\omega(t)}}e^{-i\int^t\left(\omega(t')+\omega^{(1)}(t')+\ldots\right)dt'}\left(1+F^{(1)}(t)+\ldots\right)\\ h^{II}_k(t)&=\sqrt{\frac{\omega(t)-m}{2\omega(t)}}e^{-i\int^t\left(\omega(t')+\omega^{(1)}(t')+\ldots\right)dt'}\left(1+G^{(1)}(t)+\ldots\right)\ .\end{aligned} \tag{11}$$

Here, $F^{(n)}$, $G^{(n)}$ and $\omega^{(n)}$ are complex and real functions of $n$th adiabatic order. By substituting (11) into the equations of motion (9) and the normalization condition (10), one can obtain expressions for this functions by solving order by order in the adiabatic expansion. As usual, we consider $\dot{a}$ of adiabatic order 1, $\ddot{a}$ of adiabatic order 2, and so on. On the other hand, we consider the interaction term $s(t)$ of adiabatic order 1. For a recent discussion on a related issue see Ref. 13. We work out the first adiabatic order but the full computation until fourth adiabatic order can be found in Refs. 11, 12. In order to compute the three functions of the first adiabatic order contributions $F^{(1)}$, $G^{(1)}$ and $\omega^{(1)}$ we treat independently the real and imaginary parts by writing $F^{(1)}=f^{(1)}_x+if^{(1)}_y$ and $G^{(1)}=g^{(1)}_x+ig^{(1)}_y$.

From (11) and (10) we obtain for the real part

$$\begin{aligned}(\omega - m)(g_x^{(1)} - f_x^{(1)}) &= \omega^{(1)} - g_Y\Phi \ , \\ (\omega + m)(g_x^{(1)} - f_x^{(1)}) &= -\omega^{(1)} - g_Y\Phi \ , \\ (\omega + m)f_x^{(1)} + (\omega - m)g_x^{(1)} &= 0 \ , \end{aligned} \tag{12}$$

which has as solutions

$$f_x^{(1)} = \frac{g_Y\Phi}{2\omega} - \frac{mg_Y\Phi}{2\omega^2} \ , \qquad g_x^{(1)} = -\frac{g_Y\Phi}{2\omega} - \frac{mg_Y\Phi}{2\omega^2} \ , \qquad \omega^{(1)} = \frac{mg_Y\Phi}{\omega} \ . \tag{13}$$

On the other hand, the imaginary part of the system gives

$$\begin{aligned}(\omega - m)(g_y^{(1)} - f_y^{(1)}) &= \frac{1}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega + m} - \frac{1}{\omega}\right) \ , \\ (\omega + m)(g_y^{(1)} - f_y^{(1)}) &= -\frac{1}{2}\frac{d\omega}{dt}\left(\frac{1}{\omega - m} - \frac{1}{\omega}\right) \ . \end{aligned} \tag{14}$$

These two equations are not independent. The obtained solution for $g_y^{(1)}$ and $f_y^{(1)}$ is

$$f_y^{(1)} = A - \frac{m\dot{a}}{2a\omega^2} \ , \qquad g_y^{(1)} = A \ , \tag{15}$$

where $A$ is an arbitrary first-order adiabatic function. We will choose the simplest solution

$$f_y^{(1)} = -\frac{m\dot{a}}{4\omega^2 a} \ , \qquad g_y^{(1)} = \frac{m\dot{a}}{4\omega^2 a} \ , \tag{16}$$

obeying the condition $F^{(1)}(m, g_Y\Phi) = G^{(1)}(-m, -g_Y\Phi)$. Therefore, the adiabatic expansion will also preserve the symmetries of the equations (9) with respect to the change $(m, g_Y\Phi) \to (-m, -g_Y\Phi)$. We have checked that physical expectation values are independent to any potential ambiguity in this kind of choice. We want to remark that this ambiguity is originated by the possibility of adding a time derivative in the integral of the exponent in (11) and compensating it by redefining the corresponding functions $F$ and $G$. Therefore, the local subtraction terms for renormalization can be unambiguously fixed. Nevertheless, an unambiguous adiabatic expansion of the modes can be constructed following the procedure described in Ref. 11. The next adiabatic orders are obtained by iteration.

## 4. Renormalized backreaction equations

The classical stress-energy tensor in a FLRW spacetime has two independent components. For a Dirac field, they are (no sum on $i$),

$$T_0^0 = \frac{i}{2}\left(\bar{\psi}\gamma^0\frac{\partial\psi}{\partial t} - \frac{\partial\bar{\psi}}{\partial t}\gamma^0\psi\right) \ , \qquad T_i^i = \frac{i}{2a}\left(\bar{\psi}\gamma^i\frac{\partial\psi}{\partial x^i} - \frac{\partial\bar{\psi}}{\partial x^i}\gamma^i\psi\right) \ . \tag{17}$$

In the quantum theory, using (7) the vacuum expectation values of the stress-energy tensor take the form

$$\langle T_{00}\rangle = \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2 \rho_k(t)\ , \qquad \rho_k(t) \equiv 2i\left(h_k^I \frac{\partial h_k^{I*}}{\partial t} + h_k^{II}\frac{\partial h_k^{II*}}{\partial t}\right)\ , \tag{18}$$

and

$$\langle T_{ii}\rangle = \frac{1}{2\pi^2 a}\int_0^\infty dk k^2 p_k(t)\ , \qquad p_k(t) \equiv -\frac{2k}{3a}(h_k^I h_k^{II*} + h_k^{I*} h_k^{II})\ . \tag{19}$$

On the other hand we have:

$$\langle\bar{\psi}\psi\rangle = \frac{-1}{\pi^2 a^3}\int_0^\infty dk k^2 \langle\bar{\psi}\psi\rangle_k\ , \qquad \langle\bar{\psi}\psi\rangle_k \equiv |h_k^I|^2 - |h_k^{II}|^2\ . \tag{20}$$

In general, adiabatic renormalization proceeds by subtracting the adiabatic terms of the integrand that contains the ultraviolet divergences. For $\langle\bar{\psi}\psi\rangle$ we need to subtract until third adiabatic order and for $\langle T_{\mu\nu}\rangle$ until fourth adiabatic order. It is important to remark that all contributions of a given adiabatic term of fixed (adiabatic) order must be taken into account in the subtraction, otherwise general covariance is not maintained. Also, one subtracts only the minimum number of terms required to get a finite result. For example, for the first adiabatic order for $\langle\bar{\psi}\psi\rangle_k$ introducing (11) in (20) and selecting the corresponding first adiabatic order term of this expansion for the the expectation values we obtain

$$\begin{aligned}\langle\bar{\psi}\psi\rangle_k^{(1)} &= \frac{\omega+m}{2\omega}\left(|F|^2\right)^{(1)} - \frac{\omega-m}{2\omega}\left(|G|^2\right)^{(1)} \\ &= \frac{\omega+m}{2\omega}\left(2f_x^{(1)}\right) - \frac{\omega-m}{2\omega}\left(2g_x^{(1)}\right) = \frac{g_Y\Phi}{\omega} - \frac{m^2 g_Y\Phi}{\omega^3}\ .\end{aligned} \tag{21}$$

The complete renormalized expressions for the energy momentum tensor and $\langle\bar{\psi}\psi\rangle_{ren}$ are

$$\langle T_{00}\rangle_{ren} \equiv \langle T_{00}\rangle - \langle T_{00}\rangle_{Ad} = \frac{1}{2\pi^2 a^3}\int_0^\infty dk k^2(\rho_k - \rho_k^{(0-4)})\ , \tag{22}$$

$$\langle T_{ii}\rangle_{ren} \equiv \langle T_{ii}\rangle - \langle T_{ii}\rangle_{Ad} = \frac{1}{2\pi^2 a}\int_0^\infty dk k^2(p_k - p_k^{(0-4)})\ , \tag{23}$$

$$\langle\bar{\psi}\psi\rangle_{ren} = \langle\bar{\psi}\psi\rangle - \langle\bar{\psi}\psi\rangle_{Ad} = \frac{-1}{\pi^2 a^3}\int_0^\infty dk k^2(\langle\bar{\psi}\psi\rangle_k - \langle\bar{\psi}\psi\rangle_k^{(0-3)})\ . \tag{24}$$

where $^{(0-n)}$ denotes the substraction terms of adiabatic order $0, 1, ..., n$.

These finite renormalized quantities are the ones that are introduced in the semiclassical equations of (5) and (6) in order to take into account the backreaction of the produced fields in both the gravitational and the scalar background fields.

We want also to briefly recall that the adiabatic method serves not only as a regulator but also as a renormalization mechanism. Each adiabatic subtraction of $\langle\bar{\psi}\psi\rangle_{ren}$ and $\langle T_{\mu\nu}\rangle_{ren}$ can be understood as the contribution of a generic Lagrangian containing all possible counterterms having couplings with non-negative mass dimension, up to Newtons's coupling constant[12]. The associated running of

the couplings due to this adiabatic renormalization scheme has also been recently studied in Ref. 14 for a Maxwell-Einstein theory. This is still an open problem for a Yukawa interaction. Another issue is the correct fixing of the initial conditions in order to perform a numerical analysis of the full renormalized semiclassical equations. This has already been done for a scalar field in expanding universe[15] but a generalization for Dirac fields and also for Yukawa couplings is still missing[16].

## 5. Conclusions

In order to take into account the backreaction of the produced Dirac particles during the preheating stage one needs to compute $\langle T_{\mu\nu} \rangle$ and $\langle \bar{\psi}\psi \rangle$, which are plagued with UV divergences. To renormalize these observables we have extended the adiabatic regularization to include a Yukawa interaction between the quantized Dirac fields and the classical scalar field background.

## References


1. L. A. Kofman, A. D. Lindle, and A. A. Starobinsky, *Phys. Rev. Lett.* **73**, 3195 (1994); L. A. Kofman, A. D. Lindle, and A. A. Starobinsky, *Phys. Rev. Lett.* **76**, 1011 (1996).
2. J. Baacka, K. Heitmann and C. Ptzold, *Phys Rev. D* **58**, 125013 (1998); P. Greene and L. A. Kofman, *Phys. Lett.***B**448, 6 (1999).
3. L. Parker, *Phys. Rev. Lett.* **21**, 562 (1968); *Phys. Rev. D* **183**, 1057 (1969).
4. L. Parker and D. J. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity*, CUP, Cambridge, UK (2009).
5. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*, CUP, Cambridge, UK (1982).
6. S. W. Hawking, *Commun. Math. Phys.* **43**, 199 (1975).
7. A. Fabbri and J. Navarro-Salas, *Modeling black hole evaporation*, ICP-World Scientific, London, UK (2005).
8. E. Brezin and C. Itzykson, *Phys. Rev. D* **2**, 1191 (1970).
9. L. Parker and S. A. Fulling, *Phys. Rev. D* **9**, 341 (1974).
10. A. Landete, J. Navarro-Salas and F. Torrenti, *Phys Rev. D* **88**, 061501 (2013); *Phys. Rev. D* **89**, 044030 (2014). A. del Rio, J. Navarro-Salas and F. Torrenti, *Phys. Rev. D* **90**, 084017 (2014).
11. J. F. Barbero G., A. Ferreiro, J. Navarro-Salas, E. J. S. Villaseñor, *Phys.Rev. D* **98** 025016 (2018).
12. A. del Rio, A. Ferreiro, J. Navarro-Salas and F. Torrenti, *Phys. Rev. D* **95**, 105003 (2017).
13. A. Ferreiro, J. Navarro-Salas and S. Pla, *Phys. Rev. D***98**, 045015 (2018).
14. A. Ferreiro and J. Navarro-Salas, *Phys. Lett. B* **792**, 81 (2019).
15. I. Agullo, N. William and A. Ashtekar, *Phys. Rev. D* **91**, 06405 (2015).
16. A. Ferreiro, J. Navarro-Salas, S. Pla and F. Torrenti, work in progress.

# Interacting quantum field theories and topological defects

Antonino Flachi and Vincenzo Vitagliano

*Department of Physics & Research and Education Center for Natural Sciences, Keio University, 4-1-1 Hiyoshi, Kanagawa 223-8521, Japan*

We present here the results of a recent analysis of interacting quantum field theories on a curved manifold with a topological defect constructed by geometrically deforming a lattice. We discuss an explicit example where curvature and boundary conditions compete in altering the way the quantum vacuum is destabilized. We show that the competing action of the locally induced curvature and of boundary conditions generated by the non-trivial topology allows configurations where symmetries can be spontaneously broken close to the defect core. Inspired by this effect, we propose a novel mechanism to induce a superconducting phase by triggering particle condensation along cosmic strings.

As the energy scales of a physical system become higher, non-perturbative effects get into the game and the system is said to have entered an strongly coupled regime. Remarkable examples of strongly coupled systems include QCD, high temperature superconductors and the very primordial plasma filling the universe a few instants after the big bang. Interactions are also important in describing the propagation of conducting electrons in graphene.

Precise study of strong coupled dynamics is only possible at the cost of demanding numerical simulations. On the other hand, general phenomenological guidelines for the study of strongly coupled systems can be precisely draft exploiting mathematical considerations on the underlying symmetries. The physics of symmetry breaking, in fact, notably relies on exact mathematical statements which are intimately non-perturbative, as in the case of the Goldstone theorem.

A seminal paper by Coleman and Weinberg[1] showed how symmetries can be spontaneously broken due to quantum fluctuations. This mechanism of *dynamical* symmetry breaking, originally developed in the context of scalar field self-interactions, can be naturally extended to fermions. In a field theory with massless fermions, the interactions between particles preserve helicity: the polarised left- and right-handed sectors evolve separately. However, in the celebrated Nambu–Jona-Lasinio model ($\lambda$ is the coupling constant and $N$ the total number of fermion dof,

$$S_{\rm NJL} = \int d^4x\sqrt{g}\left\{\bar\psi i\gamma^\mu\nabla_\mu\psi + \frac{\lambda}{2N}(\bar\psi\psi)^2\right\}\,,$$

and in its progeny of models for strong interactions, such symmetry is spontaneously broken by a non-vanishing order parameter: when the composite operator $\phi\sim\bar\psi\psi$ acquires a non-zero vacuum expectation value, then a dynamical effective mass for the fermions, $M_{\rm eff}\sim\langle\bar\psi\psi\rangle$, is generated. This mechanism is responsible for most of the observed mass of hadrons.

Transitions between symmetry broken and symmetry restored phases are generally triggered by changes in the external conditions. In particular a nonzero temperature leads to temperature-dependent mass generation and to restoration of broken symmetries once a certain critical value of the temperature is reached. Other factors modifying the phase diagram of self-interacting theories span from finite density effects to the action of a non-vanishing chemical potential or of an external gauge field. Here we will instead concentrate on geometrical effects, and how these challenge the vacuum stability of a theory with four fermions interactions.

Many are the configurations in which geometry affects the symmetry breaking of strongly interacting systems: in flat spacetime with $R^3 \times S^1$ topology and periodic boundary conditions, for example, the consequences of the non-trivial topology are very similar to those of nonzero temperature[2]; on the other hand, in curved spacetime the effects of external gravitational fields resemble those of an effective extra mass[3]. The combination of these external factors, and in particular topology, nonzero temperature, and curvature, acting on self-interacting theories is likely to have been of considerable importance in the early stages of the evolution of the universe. During those eras a spontaneous breaking of an internal symmetry group results in the production of topological defects – the well-known Kibble–Zurek mechanism[4,5].

Suppose the dynamical symmetry breaking for some particle model to be at the origin of the formation of a static straight cosmic string[6,7] lying along the $z$-axis, namely an infinitely long thin tube of false vacuum generated in the sudden temperature-driven transition from a phase to another (here, the transverse size of the cosmic string is neglected while 'sudden transition' means a transition with a rate that is fast if compared with the size of the system). Away from the defect, the spacetime associated with the gravitating string is accurately described by the vacuum Einstein equations. It turns out that at large distance from the string, the geometry is locally flat, $ds^2_{\rm con} = dt^2 - dz^2 - dr^2 - r^2 d\theta^2$, but with an important *caveat*: it is not globally Euclidean, since the angular coordinate does not run on the entire $2\pi$ circle; instead, $0 \leq \theta < 2\pi - \Delta$, with $\Delta > 0$ ($\Delta < 0$) being the deficit (excess) angle: surfaces at constant $t$ and $z$ are cones, not planes.

This is a remembrance of defects insertion in crystal lattices: starting from a (locally) flat lattice, the subtraction (or addition) of atoms is equivalent to the extraction (or insertion) of sections of the lattice with a given angle. The procedure results in an ice-cream-shaped lattice (or a saddle) which is locally flat everywhere apart for the apex. Note yet that the spacetime surgery does not come for free: the price to pay is the implementation of some non-trivial new boundary conditions along the cut, which will be a reminder of the deficit (excess) angle at the apex.

To investigate the role of the background geometry in modifying the vacuum structure, we will consider a $(2+1)$–dimensional honeycomb lattice, whose hexagonal structure is obtained by the superposition of two triangular sublattices: the breakdown of this discrete symmetry is behind the phenomenon of condensation

we will describe here[8]. The dynamics of the delocalized electrons on such a lattice is often described in terms of a generalisation of the Hubbard model, that in the continuum limit is mapped onto a field theory with nine different couplings[9]; however, considering the limit for a large number of fermion flavours $N$ and after bosonization, the Hubbard model for the honeycomb lattice has the following form

$$\mathcal{L}_{\mathrm{GN}} = \bar{\psi}_\sigma \imath \partial\!\!\!/ \psi_\sigma + \sigma \bar{\psi}_\sigma \phi \psi_\sigma + \frac{\phi^2}{2\lambda} \,, \tag{1}$$

where $\sigma = \pm$ is a spin index on which one sums. The final goal is to study the behaviour of the order parameter $\phi$ when moving toward the apex of the cone.

Using a regularization of the space generated by the topological 'stringy' defect with a family of smoothed versions of the conical solution (in Euclidean time), $ds^2 = d\tau^2 + f_\epsilon(r) dr^2 + r^2 d\theta^2$, where $f_\epsilon(r)$ is a regularising function: in the limit $\epsilon \to 0$ one might recover the singular cone. The Schrödinger–Lichnerowicz–Weitzenböck formula allows us to write the effective action as

$$\Gamma[\phi] = -\int d^3x \sqrt{g} \frac{\phi^2}{2\lambda} + \frac{1}{2} \sum_{p=\pm} \log\det\left(\Box + \frac{R}{4} + \phi_p^2\right), \quad \phi_\pm^2 = \phi^2 \pm \sqrt{g^{rr}}\phi' \tag{2}$$

where the metric here employed is the smoothed one, $ds^2$, rather than $ds^2_{\mathrm{con}}$. As previously mentioned, the change in the topology has a prominent role in altering the boundary conditions on the glued side of the lattice; the response to this change is captured by the employment of a modified covariant derivative, $D_\mu = \nabla_\mu + \imath \mathcal{A}_\mu$, acting on spinors and encompassing an effective non-dynamical gauge field, $\mathcal{A}_\mu$, that depends on the deficit angle[10].

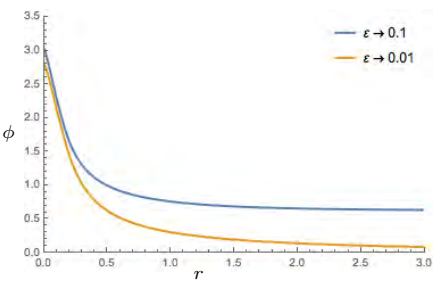



Fig. 1. Profile of the condensate $\phi$ as a function of the distance from the apex, $r$. Different colours refer to different values of the $\epsilon$ parameter regularising the cone.

Expressing the functional determinant of the (squared) Dirac operator in curved space in terms of its heat kernel expansion[11], and using zeta function regularization, it is possible to calculate the effective action (2) and find out the effective equations of motion of the order parameter $\phi$, whose solutions can be reconstructed numerically (See figure and see Ref. 12 for some examples). An alternative way to proceed could make use of the world-line formalism adapted to the present case (see Ref. 13 for a relevant calculation). Although the presence of curvature (which acts,

as for the chiral gap effect[14], as an effective mass term) is supposed to enhance a phase of symmetry restoration, the presence of the extra effective gauge field, another reminder of the geometry of the system, catalyse the formation of a bubble of condensed particles close the apex.

For relativistic cosmic strings, the possibility of particle condensation in the region surrounding the string core due only to configurational elements is a interesting way to induce a superconducting phase around the defect. However, it is worth mentioning that in a more realistic setup, phase transitions in the early universe did not seed the formation of a single straight string, but of a network of cosmic strings, which renders eventually even more striking and intriguing the connection with a crystal lattice, where a distribution of defects is more natural to occur. There is finally another interesting aspect: in order to simplify the discussion we have here considered the spontaneous symmetry breaking for the Gross–Neveu model catalysed by an external string defect, namely originated by the breakdown of some symmetry of a different field. Different would be the situation in which the responsible for the defect formation is the very same field, in which case one might take into account possible back-reaction effects.

## References


1. S. R. Coleman, E. J. Weinberg, "*Radiative Corrections as the Origin of Spontaneous Symmetry Breaking*," Phys. Rev. D **7**, 1888 (1973).
2. L. Dolan, R. Jackiw, "*Symmetry Behavior at Finite Temperature*," Phys. Rev. D **9**, 3320 (1974).
3. A. Flachi, T. Tanaka, "*Chiral Phase Transitions around Black Holes*," Phys. Rev. D **84**, 061503 (2011); A. Flachi, "*Deconfinement transition and Black Holes*," Phys. Rev. D **88**, no. 4, 041501 (2013); A. Flachi, "*Black holes as QCD laboratories*," Int. J. Mod. Phys. D **24**, 1542017 (2015).
4. T. W. B. Kibble, "*Topology of Cosmic Domains and Strings*," J. Phys. A **9**, 1387 (1976).
5. W. H. Zurek, "*Cosmological Experiments in Superfluid Helium?*," Nature **317**, 505 (1985).
6. A. Vilenkin, E. P. S. Shellard, "*Cosmic strings and other topological defects*", Cambridge University Press, (2000).
7. R. H. Brandenberger, T. Huang, K. Yang and X. Zhang, "*Cosmic strings in theories with dynamical symmetry breaking*," Phys. Rev. D **55**, 4511 (1997).
8. E. V. Castro, A. Flachi, P. Ribeiro and V. Vitagliano, "*Symmetry Breaking and Lattice Kirigami*," Phys. Rev. Lett. **121**, no. 22, 221601 (2018).
9. I. F. Herbut, "*Interactions and phase transitions on graphene's honeycomb lattice*," Phys. Rev. Lett. **97**, 146401 (2006).
10. Y. A. Sitenko, N. D. Vlasii, "*Electronic properties of graphene with a topological defect*," Nucl. Phys. B **787**, 241 (2007).

11. L. Parker, D. Toms, "*Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity*", Cambridge University Press, (2009).
12. A. Flachi and T. Tanaka, "*Chiral Modulations in Curved Space I: Formalism*," JHEP **1102**, 026 (2011); A. Flachi, "*Chiral Modulations in Curved Space II: Conifold Geometries*," JHEP **1201**, 023 (2012); A. Flachi, "*Dual Fermion Condensates in Curved Space*," Phys. Rev. D **88**, no. 8, 085011 (2013)
13. F. Bastianelli, O. Corradini, P. A. G. Pisani and C. Schubert, "*Scalar heat kernel with boundary in the worldline formalism*," JHEP **0810** (2008) 095; O. Corradini and C. Schubert, "*Spinning Particles in Quantum Mechanics and Quantum Field Theory*," arXiv:1512.08694 [hep-th].
14. A. Flachi, K. Fukushima, "*Chiral Mass-Gap in Curved Space*," Phys. Rev. Lett. **113**, 091102 (2014); A. Flachi, K. Fukushima and V. Vitagliano, "*Geometrically induced magnetic catalysis and critical dimensions*," Phys. Rev. Lett. **114**, no. 18, 181601 (2015).

# Primordial gravitational waves originating from an anisotropic pre-inflationary stage

Yu Furuya*, Yuki Niiyama and Yuuiti Sendouda

*Graduate School of Science and Technology, Hirosaki University, Hirosaki, Aomori 036-8561, Japan*
**E-mail: furuya@tap.st.hirosaki-u.ac.jp*

We discuss quantisation and time evolution of cosmological tensor perturbations in the Kasner–de Sitter space-time as a model of (pre-)inflation. We propose a prescription for reading off the quantum energy levels in the presence of the initial Kasner-like singularity.

*Keywords*: Cosmological perturbation theory, inflation, primordial gravitational waves.

## 1. Introduction

Unveiling the origin and mechanism of cosmic inflation [1,2] by means of observations is one of the ultimate challenges in modern cosmology. If there were preceding, 'pre-inflationary' stages before the onset of inflation, it is well anticipated that the space-time then was fairly anisotropic and/or inhomogeneous with a magnitude of the order of the scale of fundamental physics. For this purpose, primordial gravitational waves (PGWs) are expected to play an essential role as a probe for such an extremely high-energy universe.

Unfortunately, however, quantisation in cosmological backgrounds with a pre-inflationary, Kasner-like singularity has been a theoretical challenge, as the diverging spatial anisotropy renders the standard quantisation procedure inexecutable. [3,4] This is the central issue we have addressed in the present work.

One of the essential aspects of the problem of quantisation in an anisotropic early universe is that the ability in reading off the quantum spectra is restricted due to lack of a set of appropriate mode functions near the singularity. We argue that this difficulty may be simply circumvented if only there is a period, even if temporal, apart from the singularity where certain adiabatic conditions for an appropriately transformed variable are met. [5] Then the standard procedure of second quantisation can be carried out within the framework of the WKB approximation during such a period. It should be, however, stressed here that we are not addressing the issue of determining the quantum state of the pre-inflationary universe, which would need the knowledge of the more fundamental physics.

We demonstrate that our prescription for determining the quantum energy spectrum is useful in making physically meaningful predictions for PGWs in triaxially anisotropic Kasner–de Sitter (KdS) backgrounds. With a conservative assumption about the quantum state, we confirm that, on short wave-length scales, the resulting spectrum and directional distribution of the PGWs share the same feature with de Sitter inflation, namely, scale invariance and isotropy.

## 2. Setup

We begin with describing the basic setup for the calculation of PGWs in a KdS background.

### 2.1. *Background*

The line element of the KdS space-time is

$$g_{\mu\nu}\,\mathrm{d}x^\mu\,\mathrm{d}x^\nu = a(\eta)^2\,\left[-\mathrm{d}\eta^2 + \sum_{i=1}^{3}\left(\mathrm{e}^{\beta_i(\eta)}\,\mathrm{d}x^i\right)^2\right], \tag{1}$$

where

$$a(\eta) = a_{\rm iso}\,\sinh^{1/3}(3H_\Lambda t)\,, \quad \mathrm{e}^{\beta_i(\eta)} = \tanh^{q_i}\left(\frac{3H_\Lambda t}{2}\right) \quad (i=1,2,3) \tag{2}$$

with $a_{\rm iso}$ being a normalisation constant, $t$ the cosmic time defined by $\mathrm{d}t = a(\eta)\,\mathrm{d}\eta$, and $H_\Lambda \equiv \sqrt{\Lambda/3}$. The three exponents $q_i$ $(i=1,2,3)$ are usefully parametrised by an angular parameter $\Theta$ as

$$q_1 = \frac{2}{3}\,\sin\left(\Theta - \frac{2\pi}{3}\right), \quad q_2 = \frac{2}{3}\,\sin\left(\Theta - \frac{4\pi}{3}\right), \quad q_3 = \frac{2}{3}\,\sin\Theta\,. \tag{3}$$

In this work, we will take a particular combination $(q_1,q_2,q_3) = (1/\sqrt{3},0,-1/\sqrt{3})$ $(\Theta = 8\pi/6)$ as a typical background with a triaxial anisotropy. A KdS space-time with a triaxial anisotropy has an initial singularity at $t=0$, where a Weyl-curvature invariant $C_{\mu\nu\rho\sigma}\,C^{\mu\nu\rho\sigma}$ diverges.

The spatial metric is defined as

$$\gamma_{ij} = \mathrm{diag}(\mathrm{e}^{2\beta_1}, \mathrm{e}^{2\beta_2}, \mathrm{e}^{2\beta_3}), \quad \gamma^{ij} = \mathrm{diag}(\mathrm{e}^{-2\beta_1}, \mathrm{e}^{-2\beta_2}, \mathrm{e}^{-2\beta_3}), \tag{4}$$

and is used to raise and lower the indices of spatial tensors. The average Hubble rate $\mathcal{H}$ and the shear tensor $\sigma_{ij}$ are defined respectively by

$$\mathcal{H} \equiv \frac{a'}{a} = a_{\rm iso}\,H_\Lambda\,\frac{\cosh(3H_\Lambda t)}{\sinh^{2/3}(3H_\Lambda t)}\,, \tag{5}$$

$$\sigma_{ij} \equiv \frac{1}{2}\,\gamma'_{ij} = 3q_i\,a_{\rm iso}\,H_\Lambda\,\frac{\tanh^{2q_i}\left(\frac{3H_\Lambda t}{2}\right)}{\sinh^{2/3}(3H_\Lambda t)}\,\delta_{ij}\,, \tag{6}$$

where the prime denotes differentiation with respect to $\eta$.

The KdS universe has two regimes distinguished by the magnitude of the shear. In the earlier, 'Kasner-like' period for $t \ll H_\Lambda^{-1}$, the average scale factor $a$ and the spatial metric $\gamma_{ij}$ are approximated as

$$a(\eta) \sim a_{\rm iso}\,(3H_\Lambda t)^{1/3}\,, \quad \gamma_{ij}(\eta) \sim \left(\frac{3H_\Lambda t}{2}\right)^{2q_i}\,\delta_{ij}\,. \tag{7}$$

In the later, 'de Sitter-like' period for $t \gg H_\Lambda^{-1}$, $a$ and $\gamma_{ij}$ are approximated as

$$a(\eta) \sim 2^{-1/3}\,a_{\rm iso}\,\mathrm{e}^{H_\Lambda t}\,, \quad \gamma_{ij}(\eta) \sim \delta_{ij}\,. \tag{8}$$

The time $t = t_{\rm iso} \equiv H_\Lambda^{-1}$ appoints a moment around which the transition from the anisotropic to the isotropic regime occurs.

### 2.2. *Tensor perturbations*

We employ a gauge-invariant formalism for general Bianchi type-I cosmology developed by Pereira *et al.*[6] Here we only show the final form of the equations of motion for the '+'- and '×'-mode components of the gauge-invariant tensor perturbations:

$$\mu_+'' + \omega_+^2\,\mu_+ + \xi\,\mu_\times = 0\,, \tag{9}$$

$$\mu_\times'' + \omega_\times^2\,\mu_\times + \xi\,\mu_+ = 0\,, \tag{10}$$

with

$$\omega_+^2 \equiv \gamma^{ij}\,k_i\,k_j - \frac{a''}{a} - \frac{\left(a^2\,\sigma^{(\rm S)}\right)'}{a^2} - 2\,\left(\sigma_\times^{(\rm T)}\right)^2 - \frac{2}{a^2}\left(\frac{a^2\,\left(\sigma_+^{(\rm T)}\right)^2}{2\mathcal{H} - \sigma^{(\rm S)}}\right)'\,, \tag{11}$$

$$\omega_\times^2 \equiv \gamma^{ij}\,k_i\,k_j - \frac{a''}{a} - \frac{\left(a^2\,\sigma^{(\rm S)}\right)'}{a^2} - 2\,\left(\sigma_+^{(\rm T)}\right)^2 - \frac{2}{a^2}\left(\frac{a^2\,\left(\sigma_\times^{(\rm T)}\right)^2}{2\mathcal{H} - \sigma^{(\rm S)}}\right)'\,, \tag{12}$$

$$\xi \equiv 2\sigma_+^{(\rm T)}\,\sigma_\times^{(\rm T)} - \frac{2}{a^2}\left(\frac{a^2\,\sigma_+^{(\rm T)}\,\sigma_\times^{(\rm T)}}{2\mathcal{H} - \sigma^{(\rm S)}}\right)'\,, \tag{13}$$

where $k_i$ is the wave-vector and $\sigma^{(\rm S)}$ and $\sigma_\lambda^{(\rm T)}\,(\lambda = +, \times)$ are the polarisation components of the shear tensor.[6] Since the full expressions for the equations of motion are fairly involved, we refer the readers to Ref. 6, 7 for the details.

Due to the presence of the initial Kasner-like singularity at $a = 0$, the frequency squared $\omega_\lambda^2$ tends to blow up and the adiabatic conditions for the variables $\mu_\lambda$ are generally violated in the anisotropic era.[3,4,7] Now, let us focus on the wave-number term $\gamma^{ij}\,k_i\,k_j$ in $\omega_\lambda^2$ and, in order to make its divergent behaviour more transparent, introduce the following three functions of time:

$${}^{(i)}f = \tanh^{-q_i/2}\left(\frac{3H_\Lambda t}{2}\right) \quad (i = 1, 2, 3)\,, \tag{14}$$

with which the wave-number term can be written as

$$\gamma^{ij}\,k_i\,k_j = {}^{(1)}f^4\,k_1^2 + {}^{(2)}f^4\,k_2^2 + {}^{(3)}f^4\,k_3^2\,. \tag{15}$$

In a triaxially anisotropic background, the factors ${}^{(i)}f^4$ depend on time to the distinctive powers, so only one of the above three terms should dominate at any moment. Since $q_1 > 0$, the term proportional to $k_1^2$ necessarily diverges as $t \to 0$.

## 3. Quantisation and Time Evolution

The implications of the previous analysis urge us to seek for alternative quantities suitable for quantisation.[4,5] In what follows, we assume that the interaction terms in

the equations of motion, $\xi\,\mu_\lambda$, can be ignored, which can be justified on sub-horizon scales.[6]

For the purpose, let us consider a transformation using ${}^{(i)}f$:

$$ {}^{(i)}\chi_\lambda = {}^{(i)}f\,\mu_\lambda\,, \quad \mathrm{d}^{(i)}\tau = {}^{(i)}f^2\,\mathrm{d}\eta\,. \tag{16} $$

After this transformation, the equation of motion for the new variable ${}^{(i)}\chi_\lambda$ reads

$$ {}^{(i)}\ddot{\chi}_\lambda + {}^{(i)}\Omega_\lambda^2\,{}^{(i)}\chi_\lambda = 0\,, \tag{17} $$

where the dot denotes derivative with respect to the new time coordinate ${}^{(i)}\tau$ and the frequency squared for the new variable ${}^{(i)}\chi_\lambda$ is

$$ {}^{(i)}\Omega_\lambda^2 = \frac{\omega_\lambda^2}{{}^{(i)}f^4} + \frac{({}^{(i)}f^{-1})''}{{}^{(i)}f^3} = k_i^2 + \sum_{j\neq i} k_j^2\,\tanh^{-2(q_j-q_i)}\left(\frac{3H_\Lambda t}{2}\right) + \cdots\,, \tag{18} $$

where only the terms relevant to the norm of wave vector were shown in the last expression. Now, suppose there is some moment $t = t_*$ when a condition

$$ k_i^2 \gtrsim \sum_{j\neq i} k_j^2\,\tanh^{-2(q_j-q_i)}\left(\frac{3H_\Lambda t_*}{2}\right) \tag{19} $$

is satisfied, under which a variable ${}^{(i)}\chi_\lambda$ has an approximately constant frequency squared ${}^{(i)}\Omega_\lambda^2 \sim k_i^2$. In this case, we regard such a variable as 'suitable for quantisation' and perform its quantisation in a standard, canonical way. Hereafter, when unnecessary, we drop the superscript '$(i)$' for brevity.

In the canonical quantisation procedure, the positive-frequency modes are approximated by the (zeroth-order) WKB solutions

$$ u_\lambda^{\mathrm{WKB}}(k_i,\tau) = \frac{1}{\sqrt{2\,\Omega_\lambda(k_i,\tau)}}\mathrm{e}^{-\mathrm{i}\int^\tau \mathrm{d}\tau'\,\Omega_\lambda(k_i,\tau')}\,, \tag{20} $$

as long as the so-called WKB parameter

$$ Q_\lambda(k_i,\tau) = -\frac{1}{2\,\Omega_\lambda^2}\left[\frac{1}{\Omega_\lambda}\frac{\mathrm{d}^2\Omega_\lambda}{\mathrm{d}\tau^2} - \frac{3}{2}\left(\frac{1}{\Omega_\lambda}\frac{\mathrm{d}\Omega_\lambda}{\mathrm{d}\tau}\right)^2\right] \tag{21} $$

is tiny.[a] It is now clear from the previous analyses that modes with sufficiently large wave number $(k_1,k_2,k_3)$ have approximately constant $\Omega_\lambda^2$, hence $|Q_\lambda| \ll 1$. Also, for a wide range of wave number $(k_1,k_2,k_3)$, the ${}^{(1)}f$-transformed variable ${}^{(1)}\chi_\lambda$ behaves much like a harmonic oscillator with ${}^{(1)}\Omega_\lambda^2 \simeq k_1^2$ as long as it does not approach too close to $t = 0$ or $t = t_{\mathrm{iso}}$.

Figure 1 shows the evolution of the WKB parameter $|{}^{(1)}Q_\lambda|$ for ${}^{(1)}\chi_\lambda$ and its waveform for $k_1 = k_2 = k_3 = k/\sqrt{3} = (100/\sqrt{3})\,a_{\mathrm{iso}}\,H_\Lambda$. The initial condition for the numerical calculation at $t = t_* = 10^{-5}\,H_\Lambda^{-1}$ is so taken as to match the analytic

[a] The function $u_\lambda^{\mathrm{WKB}}$ satisfies $\frac{\mathrm{d}^2 u_\lambda^{\mathrm{WKB}}}{\mathrm{d}\tau^2} + [1-Q_\lambda]\,\Omega_\lambda^2\,u_\lambda^{\mathrm{WKB}} = 0$.

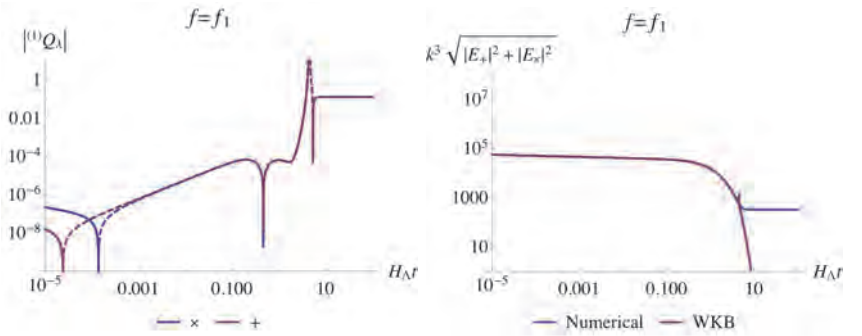



Fig. 1. Time evolution of the WKB parameter ${}^{(1)}Q_\lambda$ (left) and comparison of the WKB solution with the numerical solution (right) for a mode with $k_1 = k_2 = k_3 = k/\sqrt{3} = (100/\sqrt{3})\, a_{\rm iso}\, H_\Lambda$ . The lines for ${}^{(1)}Q_\lambda$ are dashed when the values are negative.

WKB mode function. As $t$ approaches to the time of isotropisation $t_{\rm iso}$ , the WKB mode function is isotropised as

$$|u_\lambda^{\rm WKB}|^2 \sim \frac{1}{2\,\Omega_\lambda} \to \frac{1}{2\,\sqrt{k_1^2+k_2^2+k_3^2}}\,. \tag{22}$$

Then, after the transition to ordinary, isotropic de Sitter inflation at $t = t_{\rm iso}$ , the tensor perturbations decay as $|E_\lambda| \propto a^{-1}$ until they exit the horizon when $a = k/H_\Lambda$ .

So far, we have not mentioned in which quantum state the universe is. Indeed, this is beyond the scope of this study, and we here simply assume that the quantum state is the ground state at $t = t_*$ . Then the power spectrum of PGWs after inflation is estimated as[5]

$$P_{\rm T} \equiv \sum_{\lambda=+,\times} 4\,|E_\lambda(a=k/H_\Lambda)|^2 = 64\pi\, G\, \left.\frac{|u_\lambda|^2}{a^2\, f^2}\right|_{a=k/H_\Lambda} \approx \frac{8\pi\, G\, H_\Lambda^2}{k^3}\,, \tag{23}$$

where the last expression depends only on the norm of wave number, $k = \sqrt{k_1^2+k_2^2+k_3^2}$ , and agrees with the one in de Sitter inflation, that is, isotropic and scale-invariant.

## 4. Sky Map of PGWs

A result of PGWs in a triaxially anisotropic KdS is shown in Fig. 2. One can see that the prediction in the triaxial KdS coincides with that of isotropic de Sitter inflation for the wave length considered and wide range of direction. This is consistent with our analytic evaluation of the power spectrum (23).

There are some features, though: In the polar regions where $|k_3| \gg |k_1|\,, |k_2|\,,$ there is an apparent enhancement of the PGWs. At this stage, we do not claim this should be a real signature of the initial anisotropy because there vast violations of the WKB approximation occur. The regions without value (in black) correspond

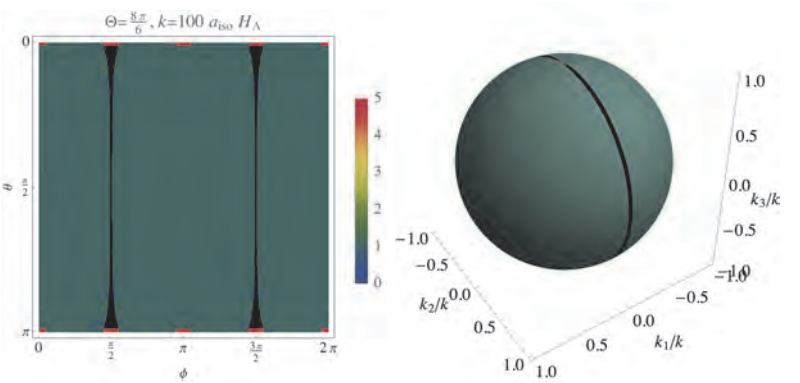



Fig. 2. An all-sky map of PGW intensity for a wave number $k = 10^2\, a_{\rm iso}\, H_\Lambda$ in the presence of a background anisotropy characterised by the indices $(q_1, q_2, q_3) = (1/\sqrt{3}, 0, -1/\sqrt{3})$ $(\Theta = 8\pi/6)$. The values are normalised by those of de Sitter inflation and they are approximately unity in the vast region. In regions without a value (black), all ${}^{(i)}\Omega_\times^2$ $(i = 1, 2, 3)$ are negative and we do not try to compute values.

to where all of ${}^{(i)}\Omega_\lambda^2$ $(i = 1, 2, 3)$ have negative values. We do not try to compute anything in such regions since our approximation method lacks a power of prediction. Note, however, that as we have shown in Ref. 7, the intensity of classical PGWs is enhanced in regions corresponding to the 'uncomputable' regions. This may imply that there is a possibility that such regions could play an important role in detecting the imprint of the anisotropic universe in future observations.

## Acknowledgments

The work of YS is in part supported by JSPS KAKENHI Grant Number 16K17675.

## References

1. A. H. Guth, *Phys. Rev. D* **23**, 347 (1981).
2. K. Sato, *Mon. Not. Roy. Astron. Soc.* **195**, 467 (1981).
3. A. E. Gümrükçüoğlu, C. R. Contaldi and M. Peloso, *J. Cosmol. Astropart. Phys.* **0711**, 005 (2007).
4. C. Pitrou, T. S. Pereira and J.-P. Uzan, *J. Cosmol. Astropart. Phys.* **0804**, 004 (2008).
5. Y. Furuya, Y. Niiyama and Y. Sendouda, *Class. Quant. Grav.* **36**, 085007 (2019).
6. T. S. Pereira, C. Pitrou and J.-P. Uzan, *J. Cosmol. Astropart. Phys.* **0709**, 006 (2007).
7. Y. Furuya, Y. Niiyama and Y. Sendouda, *J. Cosmol. Astropart. Phys.* **1701**, 009 (2017).

# Gravitational entropy production in a quantum low-energy gravity model free from causality violation problems

Filippo Maimone and Giovanni Scelza

*Associazione Culturale "Velia Polis", via Capo di Mezzo, Vallo della Lucania (Salerno), 84078, Italy*

Adele Naddeo

*INFN, Sezione di Napoli, C. U. Monte S. Angelo, Via Cinthia, Napoli, 80126, Italy*
*E-mail: anaddeo@na.infn.it*

Wave-function collapse following a measurement process is a longstanding controversial issue of quantum physics. It introduces an element of strong non-linearity and irreversibility in an otherwise unitary and reversible dynamics. Several proposals of modification of Quantum Mechanics have been put forward in the past few decades in order to solve such a dichotomy. Among them, some approaches and explicit models considered the possible role of gravity in the wave-function collapse as a result of the incompatibility of general relativity and unitary time evolution of Quantum Mechanics. In this contribution we present some results based on one of such models, De Filippo's Nonunitary Newtonian Gravity, which shows several appealing features: while reproducing at a macroscopic level the ordinary Newtonian interaction, it presents a mass threshold for gravitational localization. In particular, it provides a mechanism for the evolution of macroscopic coherent superpositions of states into ensembles of pure states. On one hand, we show the results of a numerical simulation of a simple system, i.e. two particles in a harmonic trap interacting via an 'electrical' delta-like potential and gravitational interaction. Starting from an energy eigenstate within the ordinary setting, we find that, while energy expectation remains constant, a slow net variation of the von Neumann entropy for the system as a whole takes place, with a small modulation induced on the relative entanglement entropy of the two particles. On the other hand, we explicitly show how a one-parameter generalization of the model, reproducing the nonlinear Newton-Schrödinger equation as the parameter goes to infinity, is free from any causality-violation problem for any finite value of it.

*Keywords*: Gravity; Entanglement entropy; Newton-Schrödinger; Causality violation.

## 1. Introduction

The possible role of gravity in producing a fundamental nonunitary time evolution of the state vector in quantum mechanics (QM) has been invoked in the past few decades by a number of authors, on different grounds[1–3]. The long-standing difficulties posed by the two basic processes of QM, *i.e.* the Schrödinger deterministic evolution and the nonunitary process associated with the act of measurement, lead to the quest of how to reconcile these radically different processes and explain the transition to classicality. These foundational issues appear to be strictly connected with quantum gravity and, in particular, with the way in which gravitational fields are generated by quantum matter. It is still not even clear if gravity has to be quantized at all or it is intrinsically classic and should be properly treated. In the latter case a proposed route is the semi-classical gravity, based on the requirement

that the energy-momentum tensor appearing in Einstein equations is taken as an expectation on the quantum state. This prescription, together with the Newtonian limit, leads to the Newton-Schrödinger equation[4]. While its solutions for a lump of matter show interesting properties, this equation cannot be a valid candidate to give self-gravity at low energies because the nonlinearity allows for superluminal communications. Within this context, in the past years De Filippo introduced a nonunitary non-Markovian model of Newtonian gravity, NNG from now on (see e.g. Ref. 5, references therein). On the one hand this model can be seen as the non-relativistic limit of a classically stable version of higher derivative gravity, whose heuristic application to black hole singularity is in agreement with Bekenstein-Hawking entropy[6,7], given the smoothed singularity of the black hole[5]. On the other hand, it presents several appealing features to become a natural candidate as an effective low-energy model of gravity. For example, while reproducing at a macroscopic level the ordinary Newtonian interaction, it presents a mass threshold for localization, which for ordinary matter densities is about $10^{11}$ proton masses[5,8]. As a key result, it provides a mechanism for the evolution of macroscopic coherent superpositions of states into ensembles of pure states. Furthermore the evolution of the density matrix is compatible with the expectations leading to the phenomenological spontaneous localization models[3], as it was argued that they should be both nonlinear and nonunitary. However, at variance with them, it does not present obstructions consistent with its special-relativistic extension[9]. As to the non-Markovian character of the model, it is interesting to mention the works of Unruh, who studied some nonunitary toy models mimicking black hole evaporation, pointing out the quest for non-Markovianity in order to avoid net dissipation while keeping decoherence[10]. In the NNG model such a feature is built-in.

On the black hole entropy side, in the last decades it has been shown how the decoherence associated to black hole formation and evaporation[11] appears to be a fundamental one, requiring a modification of the unitary time evolution of quantum mechanics in a nonunitary sense: that would allow a non vanishing probability for the evolution from pure states to mixed states. In this sense black hole entropy has been hypothesized[12] as due to the entanglement between matter-fields inside and outside the event horizon. Indeed, starting from suitable initial conditions, only a nonunitary quantum dynamics could allow for a microscopic derivation of the Second Law of thermodynamics for a closed system by resorting to the concept of von Neumann entropy, as pointed out in Ref. 13. In this way it should be possible in principle to shed new light on the long standing and still open problem of quantum foundations of the Second Law of thermodynamics[14–16] without resorting to the ambiguous and subjective procedure of coarse graining[17].

## 2. A two-particle simulation

In this Section we make a first step in demonstrating the ability of NNG to reproduce a gravity-induced relaxation towards thermodynamic equilibrium even for a

perfectly isolated system[18]. To do so, a simple system of two interacting particles is considered. While of course localization mechanism is not acting in this case, due to the far from threshold mass of the particles, some of the peculiar characteristic of the model can be verified in a more general context with respect to the single-lump problem.

More specifically we consider the two particles in an harmonic trap, interacting with each other through 'electrostatic' and gravitational interaction, whose 'physical' Hamiltonian, in the ordinary (first-quantization) setting, is

$$H_P(\mathbf{x}_1,\mathbf{x}_2)=\sum_{i=1}^{2}\left(-\frac{\hbar^2}{2\mu}\Delta_{\mathbf{x}_i}+\frac{\mu}{2}\omega^2\mathbf{x}_i^2\right)+\sum_{i<j=1}^{2}\left(\frac{4\pi\hbar^2 l_s}{\mu}\delta^{(3)}(\mathbf{x}_i-\mathbf{x}_j)-\frac{G\mu^2}{|\mathbf{x}_i-\mathbf{x}_j|}\right); \tag{1}$$

here $\mu$ is the mass of the particles, $\omega$ is the frequency of the trap and $l_s$ is the s-wave scattering length. We are considering a 'dilute' system, such that the electrical interaction can be assumed to have a contact form with a dominant s-wave scattering channel. We take numerical parameters that make the electrical interaction at most comparable with the oscillator's energy, while gravity enters the problem as an higher order correction. Our model, in the (first-quantization) ordinary setting, is defined by the following general meta-Hamiltonian:

$$H_{TOT}=H_P(\mathbf{x}_1,\mathbf{x}_2)+H_P(\widetilde{\mathbf{x}}_1,\widetilde{\mathbf{x}}_2)+H_{NNG}(\mathbf{x}_1,\mathbf{x}_2;\widetilde{\mathbf{x}}_1,\widetilde{\mathbf{x}}_2),$$

with $H_{NNG}=G\mu^2\sum_{i<j}\left(-\frac{1}{|\mathbf{x}_i-\widetilde{\mathbf{x}}_j|}+\frac{1}{2|\mathbf{x}_i-\mathbf{x}_j|}+\frac{1}{2|\widetilde{\mathbf{x}}_i-\widetilde{\mathbf{x}}_j|}\right)$.

Then the time dependent physical density matrix is computed by tracing out the hidden degrees of freedom and the corresponding von Neumann entropy is derived as the entanglement entropy with such (fictitious) hidden degrees of freedom. This is obtained via a numerical simulation, by choosing as initial condition an eigenstate of the physical Hamiltonian. As a result, we find that entropy fluctuations take place, owing to the (nonunitary part of) gravitational interactions, with the initial pure state evolving into a mixture[18]. The behavior of one- and two-particle von Neumann entropy as a function of time is depicted in Fig. 1 for an initial state $|\phi_2\rangle$, equal to the eigenstate of the physical energy associated with the 2nd highest energy eigenvalue $E_2$ of the two-particle system under study. The following values of the physical parameters have been chosen: $l_s, \mu$ and $\omega$: $l_s = 5.5\cdot 10^{-8}m$, $\mu = 1.2\cdot 10^{-24}Kg$ and $\omega = 4\pi\cdot 10^3 s^{-1}$ (which are compatible with current experiments with trapped ultracold atoms[19] and complex molecules[20]) together with an artificially augmented 'gravitational constant' $G = 6.67408\times 10^{-6}m^3kg^{-1}s^{-2}$ ($10^5$ times the real constant).

Due to gravity-induced fluctuations the system entropy shows a net variation over very long times, at variance with the case $G = 0$, in which it would have been a constant of motion. At the same time, single particle entropy, which in the ordinary setting is itself constant, shows now a modulation by the system entropy itself. Incidentally, it has been verified that the expectation of physical energy is

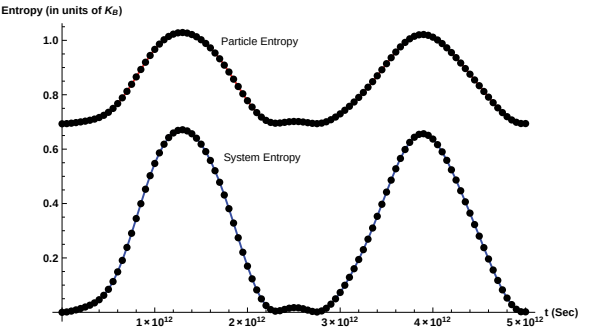



Fig. 1. Time evolution of one-particle and two-particle entropies (respectively, curve above and below) for the initial state chosen.

a constant, meaning that the non-unitary term has no net energy associated with itself, but is purely fluctuational. This constitutes a first necessary step: an extension of this work can be envisaged, which could show explicitly within a realistic setting the way in which thermodynamics could emerge in the framework of NNG, shedding new light on the possible role of gravity in the quantum foundations of thermodynamics in ordinary low-energy physics[5,21,22].

## 3. NNG as a natural cure for causality violations in Newton-Schrödinger equation

In this Section we look at the single particle Newton-Schrödinger equation as the mean-field approximation of an equation of $N$ identical copies of the particle, interacting via usual gravitational interaction, when $N$ goes to infinity (see Ref. 23 for details). The general $N$-copy model is a fully consistent quantum theory, in which superluminal communications are automatically avoided. This feature can be shown to be a consequence of the intrinsic mechanism of spontaneous state reduction, built in in our NNG model and completely suppressed in the Newton-Schrödinger limit.

Here we briefly discuss a specific (ideal) EPR-like experiment involving the superposition of two distinct Center of Mass position states of a massive body. This amounts to consider a sphere of matter of ordinary density of radius $R$, where the state $|Z\rangle$ denotes the sphere with center on the z-axis at $z = Z$, and to take the state of the sphere equal to $(|+Z\rangle+|-Z\rangle)/\sqrt{2}$ (within the ordinary QM setting). A probe mass moving along the $x-$ axis will, according to the non-relativistic quantum theory of gravity, become entangled with the state of the sphere, resulting in the state vector $(|+Z\rangle|up\rangle + |-Z\rangle|down\rangle)/\sqrt{2}$, where $|up\rangle$ ($|down\rangle$) means that the probe is deflected in the positive (negative) $z-$direction. Now, according to semi-classical gravity, the probe mass should be undeflected. A serious theoretical objection to semiclassical gravity is that it allows superluminal communication and then causality violation. To see this, consider the entangled state $(|+Z\rangle|0\rangle + |-Z\rangle|1\rangle)/\sqrt{2}$, where the state $|0\rangle$ and $|1\rangle$ denote orthogonal states of a two state system (qbit)

which is at a large distance from the sphere, but close to a "sender". A probe mass is then used as before. If the sender chooses not to measure the system, the "receiver", who is close to the sphere and uses the probe mass as described above, finds it undeflected. If, on the other hand, the sender chooses to measure the system, thereby finding it to be in the state $|0\rangle$ or $|1\rangle$, the sphere will immediately be in the state $|+Z\rangle$ or $|-Z\rangle$ respectively. Then the receiver will be able to see this because the probe mass will now be deflected up or down.

Let's translate now this conceptual experiment into our model of replicas. The entangled state of the *sphere+q-bit* is:

$$\begin{aligned}\|\Psi^{(N)}\rangle\rangle &= \bigotimes_{i=1}^{N} \frac{1}{\sqrt{2}}(\sqrt{p}|+Z^i\rangle|0^i\rangle + \sqrt{q}|-Z^i\rangle|1^i\rangle) \\ &= \sum_{k=0}^{N} \sqrt{\binom{N}{k}} \left(\frac{1}{2}\right)^{N/2} |+Z;0\rangle^{\otimes k}|-Z;1\rangle^{\otimes(N-k)}. \end{aligned} \tag{2}$$

Consider now the probe particle (supposed to have a mass much more smaller than the lump) shut just over the superposition. Including the probe in the system's description, we have that an initially unentangled (global) state of the probe and of the composite *sphere+q-bit* system evolves towards an entangled one,

$$\|\varphi_I^{(N)}\rangle\rangle \bigotimes \|\Psi^{(N)}\rangle\rangle \to \sum_{k=0}^{N} \sqrt{\binom{N}{k}} \left(\frac{1}{2}\right)^{N/2} \|\varphi_k^{(N)}\rangle\rangle|+Z;0\rangle^{\otimes k}|-Z;1\rangle^{\otimes(N-k)}, \tag{3}$$

where $\|\varphi_I^{(N)}\rangle\rangle = \bigotimes_{m=1}^{N} |\varphi_I^m\rangle$ is the product of identical copies of the probe's state (it is assumed that probe's mass is so small that gravity-induced internal entanglement among copies is irrelevant).

For simplicity let's consider the simplest case with $N = 2$, for which we have only three branches in the superposition (downward, central and upward trajectories). Generalization to a generic $N$ is straightforward. In order to detect a whatever deflection in the probe particle's trajectory, the size of the wave packets describing the particle's states in the superposition should be smaller than the deflection itself. Moreover, their spreading along the path have to be taken into account before the position measurement of the particle's position along the $z$-axis. Denoting with $v_x$ the velocity of the particle, $m$ its mass, and $M$ the mass of the sphere, the time during which the sphere's gravitational attraction is effective is of the order of $T \sim R/v_x$. Then it should be $\frac{GMT}{R^2} \gtrsim \frac{\hbar}{mZ}$.

On the other hand, a peculiar dynamical signature of the NNG model[5] is that, for a lump of mass $M$ above a threshold of $10^{11} m_p$, a superposition of the CM wave packets separated by a distance much smaller than the body's size undergoes a rapid state reduction after a characteristic time $\tau_g \sim \hbar G^{-1} M^{-5/3} \rho^{-1/3}$[24], leading from the initial superposition to an ensemble of localized states. The additional condition for the detection of a significant deflection of the probe is that the measurement

time should be smaller than this reduction time, which together with the constraint above gives the set of conditions that must be satisfied simultaneously:

$$\begin{cases} T \gtrsim \left(\frac{\hbar}{m}\right)^{1/3} \left[\frac{1}{2}\left(\frac{4}{3}\pi\rho\right)^{\frac{2}{3}} G M^{\frac{1}{3}}\right]^{-\frac{2}{3}} \\ T \ll \tau_g = \hbar G^{-1} M^{-5/3} \rho^{-1/3}. \end{cases} \tag{4}$$

These conditions cannot be satisfied together, as shown in Ref. 24. Thus we can conclude that gravity-induced state reduction is so rapid to forbid a sufficiently long measurement, which would otherwise permit a deflection discrimination with respect to the spreading of the wave packet. In other words the 'peaceful coexistence' between (deterministic) QM and special relativity indeed *imply* linearity, unless a certain amount of predictability loss is present.

## References


1. R. Penrose, *Gen. Rel. Grav.* **28**, 581 (1996).
2. L. Diosi, *Phys. Rev. A* **40**, 1165 (1989).
3. A. Bassi, G. C. Ghirardi, *Phys. Rept.* **379**, 257 (2003), references therein.
4. M. Bahrami, A. Grobardt, S. Donadi, A. Bassi *New J. Phys.* **16**, 115007 (2014).
5. S. De Filippo, F. Maimone, *Phys. Rev. D* **66**, 044018 (2002).
6. J. D. Bekenstein, *Phys. Rev. D* **7**, 2333 (1973).
7. S. W. Hawking, *Comm. Math. Phys.* **43**, 199 (1975).
8. F. Maimone, G. Scelza, A. Naddeo, V. Pelino, *Phys. Rev. A* **83**, 062124 (2011).
9. S. De Filippo, F. Maimone, *Phys. Lett. B* **584**, 141 (2004).
10. W. G. Unruh, R. M. Wald, *Phys. Rev. D* **52**, 2176 (1995).
11. S. W. Hawking, *Phys. Rev. D* **14**, 2460 (1976).
12. L. Bombelli, R. k. Koul, J. Lee, R. D. Sorkin, *Phys. Rev. D* **34**, 373 (1986).
13. R. M. Wald (2001) `http://www.livingreviews.org/articles/Volume4/2001-6wald`, references therein.
14. L. Landau, M. Lifshitz, *Statistical Physics*, (Pergamon Press, Oxford, 1978).
15. E. Schroedinger, *Statistical Thermodynamics*, (Dover, New York, 1989).
16. W. Zurek, J. Paz, *Phys. Rev. Lett.* **72**, 2508 (1994).
17. M. Gell-Mann, J. B. Hartle, *Phys. Rev. A* **76**, 022104 (2007).
18. G. Scelza, F. Maimone, A. Naddeo, *J. Phys. Comm.* **2**, 015014 (2018).
19. I. Bloch, J. Dalibard, W. Zwerger, *Rev. Mod. Phys.* **80**, 885 (2008).
20. K. Hornberger, S. Gerlich, P. Haslinger, S. Nimmrichter, M. Arndt, *Rev. Mod. Phys.* **84**, 157 (2012).
21. S. De Filippo, F. Maimone, *Entropy* **6**, 153 (2004).
22. S. De Filippo, F. Maimone, A. L. Robustelli, *Physica A* **330**, 459 (2003).
23. S. De Filippo, `arXiv:gr-qc/0106057` *The Schrödinger-Newton model as $N \to \infty$ limit of a N color model*, 2001.
24. F. Maimone, G. Scelza, A. Naddeo, *Physica Scripta* **94**, 075001 (2019).

# Radiation from a receding mirror: Can an observer distinguish fermions from bosons?

Wan Mohamad Husni Wan Mokhtar

*School of Mathematical Sciences, University of Nottingham, Nottingham NG7 2RD, UK*
*School of Physics, Universiti Sains Malaysia, 11800 USM, Penang, Malaysia*
*E-mail: wanhusni@usm.my*

It is well known that a receding mirror in Minkowski spacetime can model the formation of a black hole, producing Hawking-like radiation at late times. We ask what an observer would need to do to discern whether the radiation is fermionic or bosonic. Specialising to massless fields in 1+1 dimensions, we find that an Unruh–DeWitt detector accomplishes this: the late time transition rate of a detector coupled linearly to the scalar density of a spinor field is proportional to the Helmholtz free energy density of a fermionic thermal bath, hence showing a clear sign of Fermi–Dirac statistics, with no counterpart in the response of a detector coupled linearly to a scalar field or its derivative. By contrast, an observer examining just the stress-energy tensor sees no difference between a fermion and a boson, neither at late times nor early.

*Keywords*: Fermions; Hawking radiation; Moving mirror; Particle detector.

## 1. Introduction

In 1975, Hawking showed that an observer far away from a star undergoing gravitational collapse to form a black hole will see, at late times, thermal radiation with temperature[1]

$$T_{BH} = \frac{\kappa}{2\pi}, \tag{1}$$

where $\kappa$ is the surface gravity of the black hole. Two years later, Davies and Fulling showed that a receding mirror in Minkowski spacetime can produce similar radiation at late times.[2] This is provided the mirror follows a trajectory $z(t)$ satisfying

$$z(t) = -t - Ae^{-2at} + B \quad \text{as} \quad t \to \infty, \tag{2}$$

where $(t, z)$ are the standard Minkowski coordinates and $a, A, B$ are positive constants. The parameter $a$ plays the role of $\kappa$. An example of such a trajectory, illustrated in Fig. 1, is given by[3,4]

$$z(t) = \frac{1}{a} \ln \left( -\frac{e^{at}}{2} + \sqrt{\frac{e^{2at}}{4} + 1} \right). \tag{3}$$

## 2. Bosonic Radiation

A static observer at the finite distance $z = d$ to the right of the mirror may analyse the radiation by examining the stress-energy tensor of the quantum field. After subtracting the divergent Minkowski vacuum contribution, the renormalised expectation value of $T_{\mu\nu}$, evaluated at $(t, d)$ in the early times vacuum of a massless scalar

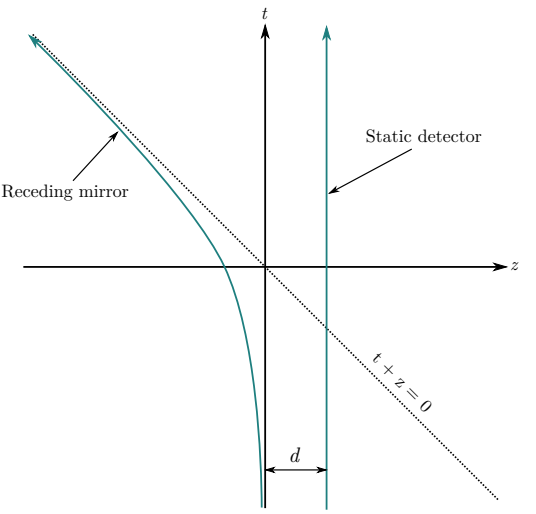



Fig. 1. An illustration of a mirror following the trajectory given in Eq. (3) and an Unruh–DeWitt detector resting at $z = d$. The mirror approaches an asymptote given by $t + z = 0$ at late times.

field, is given by[2]

$$\langle T_{tt}\rangle = \langle T_{zz}\rangle = -\langle T_{tz}\rangle = -\langle T_{zt}\rangle = -\frac{1}{12\pi}\frac{(1-\dot{z}^2)^{\frac{1}{2}}}{(1-\dot{z})^2}\frac{\mathrm{d}}{\mathrm{d}t_d}\left(\frac{\ddot{z}}{(1-\dot{z}^2)^{\frac{3}{2}}}\right). \quad (4)$$

The dot indicates a derivative with respect to the time coordinate $t_d$ of the mirror when it intersects the null line $t - d$.

Alternatively, the observer may probe the radiation using an Unruh–DeWitt detector.[5,6] The simplest version of such a detector is a two-level quantum system coupled to some scalar function $f$ of the field. For a single scalar field $\phi$, for instance, $f$ could be $\phi$ itself or its derivative $\partial_\tau\phi$ with respect to the detector's proper time $\tau$. Up to some detector-specific proportionality constant, the detector's transition probability from energy 0 to $\omega$, after the interaction is switched on and off at proper times $\tau_0$ and $\tau$ respectively, is given by the response function $F(\omega,\tau)$. The quantity that we are interested in is the detector's transition rate[7] $\dot{F}(\omega,\tau) := \partial_\tau F(\omega,\tau)$. Intuitively, $\dot{F}(\omega,\tau)$ measures how the transition probability would change if the detector interacts with the field at a constant coupling strength $c$ for an additional infinitesimal period of detector proper time.

When the interaction is switched on in the asymptotic past (i.e. $\tau_0 \to -\infty$), the late time (i.e. $\tau \to \infty$) transition rate of a static detector at $z = d$, coupled linearly to a scalar field $\phi$ and to its time derivative $\partial_\tau\phi$, is respectively given by[3,4]

$$\dot{F}(\omega,\tau) = -\frac{1}{\omega}\left[\frac{1}{2} - \cos(\omega|R|)\right]\Theta(-\omega) + \frac{a}{2\pi\omega^2} + \frac{1}{2\omega}\frac{1}{e^{2\pi\omega/a}-1} + o(1)\,, \quad (5)$$

$$\dot{F}(\omega,\tau) = -\frac{\omega}{2}\Theta(-\omega) + \frac{\omega}{2(e^{2\pi\omega/a}-1)} + o(1)\,. \quad (6)$$

The parameter $R$ actually diverges in this limit but we shall not concern with the issue here — see Ref. 3 for details. In each expression, we can see a term that is proportional to the Bose–Einstein statistics, indicating that the radiation is bosonic in nature.

## 3. Fermionic Radiation

Our new results[8] concern the radiation from a receding mirror when the field is a massless spinor field $\psi$. How do the local observations made by a static observer differ from those for the scalar field?

We model our totally reflecting mirror by a moving boundary that imposes the MIT bag boundary condition[9]

$$in_\mu \gamma^\mu \psi(t, z(t)) = \psi(t, z(t)), \tag{7}$$

where $n_\mu$ is the inward-directed unit normal to the mirror's worldline and $\gamma^\mu$ are the Dirac gamma matrices. After solving the equation of motion and quantising the field, we evaluate the renormalised expectation value of $T_{\mu\nu}$ via the point-splitting method.[10] We find that the result is exactly identical to Eq. (4). Hence, an observer analysing the radiation just by examining the stress-energy tensor will not be able to deduce that the radiation is made up of fermions, neither at late times nor early.

We then consider an Unruh–DeWitt detector coupled linearly to the scalar density $\overline{\psi}\psi$ of the field where $\overline{\psi}$ is the usual Dirac conjugate. This type of coupling introduces a divergent contribution to the response function[11,12] and hence the transition rate is undefined *a priori*. However, for a massless field in $1+1$ Minkowski spacetime, the divergent term can be dropped by a certain choice of operation ordering. The remaining terms then give a finite transition rate except when the interaction period coincides with the time taken for the field to be reflected back, backwards in time, to the observer. This, however, is not an issue in calculating the late time limit, $\tau \to \infty$.

When the interaction is switched on at a finite proper time $\tau_0$, we find that the late time transition rate is given by

$$\dot{F}(\omega, \tau) = \frac{a}{2\pi^2} \ln(1 + e^{-\frac{2\pi\omega}{a}}) + o(1)\,. \tag{8}$$

At first glance, it is not obvious that Eq. (8) is associated to fermions since there is no term proportional to the Fermi–Dirac statistics. However, recall that the Helmholtz free energy of fermions in a thermal bath at temperature $T = a/2\pi$ is given by[13]

$$F = -2L \int_0^\infty \frac{\mathrm{d}\omega}{2\pi} \left[\omega + \frac{a}{\pi} \ln(1 + e^{-\frac{2\pi\omega}{a}})\right], \tag{9}$$

where $L$ is the length of the system of fermions. Ignoring the first term in Eq. (9), which corresponds to the vacuum energy, we see that Eq. (8), when $\omega > 0$, is proportional to the Helmholtz free energy density of mode $\omega$ of the fermionic thermal

bath. Hence, an observer analysing the late times radiation with our detector will be able to deduce that the radiation consists of fermions.

## 4. Conclusion

We have analysed a receding mirror producing Hawking-like radiation at late times by calculating the renormalised stress-energy tensor and the transition rate of an Unruh–DeWitt detector. We find that the former is identical for massless scalar and spinor fields. Hence, an observer examining only the stress-energy tensor will see no difference between bosons and fermions, neither at late times nor early. In contrast, an observer equipped with an Unruh–DeWitt detector observes that the late time transition rate contains a term proportional to the Bose–Einstein statistics when the radiation is bosonic but is proportional to the Helmholtz free energy density of thermal fermions when the radiation is fermionic. The observer with a detector will hence be able to discern whether the late times radiation consists of bosons or fermions. Further technical details can be found in Ref. 8.

## Acknowledgements

The author thanks Jorma Louko for helpful discussions. Thanks also to the organisers of the Fifteenth Marcel Grossmann Meeting (MG15) in Rome, Italy. The author is financially supported by the Ministry of Education Malaysia (Higher Education) and Universiti Sains Malaysia.

## References

1. S. W. Hawking, Particle creation by black holes, *Commun. Math. Phys.* **43**, 199 (1975).
2. P. C. W. Davies and S. A. Fulling, Radiation from moving mirrors and from black holes, *Proc. Roy. Soc. Lond.* **A356**, 237 (1977).
3. L. Hodgkinson, Particle detectors in curved spacetime quantum field theory, PhD thesis, University of Nottingham, (Nottingham, UK, 2013).
4. B. A. Juárez-Aubry and J. Louko, Onset and decay of the 1 + 1 Hawking-Unruh effect: what the derivative-coupling detector saw, *Class. Quant. Grav.* **31**, 245007 (2014).
5. W. G. Unruh, Notes on black hole evaporation, *Phys. Rev.* **D14**, 870 (1976).
6. B. S. DeWitt, *Quantum Gravity: The New Synthesis*, in *General Relativity: An Einstein Centenary Survey*, eds. S. W. Hawking and W. Israel (Cambridge University Press, 1979), pp. 680–745.
7. A. Satz, Then again, how often does the Unruh-DeWitt detector click if we switch it carefully?, *Class. Quant. Grav.* **24**, 1719 (2007).
8. W. M. H. Wan Mokhtar, Radiation from a receding mirror: Unruh-DeWitt detector distinguishes a Dirac fermion from a scalar boson, *Class. Quant. Grav.* **37**, 075011 (2020).

9. A. Chodos, R. L. Jaffe, K. Johnson, C. B. Thorn and V. F. Weisskopf, New extended model of hadrons, *Phys. Rev.* **D9**, 3471 (1974).
10. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, UK, 1984).
11. D. Hümmer, E. Martin-Martinez and A. Kempf, Renormalized Unruh-DeWitt particle detector models for boson and fermion fields, *Phys. Rev.* **D93**, 024019 (2016).
12. J. Louko and V. Toussaint, Unruh-DeWitt detector's response to fermions in flat spacetimes, *Phys. Rev.* **D94**, 064027 (2016).
13. J. I. Kapusta and C. Gale, *Finite-Temperature Field Theory: Principles and Applications* (Cambridge University Press, Cambridge, UK, 2006).

# A self-consistent solution of evaporating black holes

Hikaru Kawai

*Department of Physics, Kyoto University, Kitashirakawa, Kyoto 606-8502, Japan*
*E-mail: hkawai@gauge.scphys.kyoto-u.ac.jp*

Yuki Yokokura

*iTHEMS, RIKEN, Wako, Saitama 351-0198, Japan*
*E-mail: yuki.yokokura@riken.jp*

What is the black hole in quantum mechanics? We try this problem in a self-consistent manner. First, we analyze time evolution of a 4D spherically symmetric collapsing matter including the back reaction of particle creation that occurs in the time-dependent spacetime. As a result, a compact high-density star with no horizon or singularity is formed and eventually evaporates. This is a quantum black hole. We can construct a self-consistent solution of the semi-classical Einstein equation showing this structure. In fact, we construct the metric, evaluate the expectation values of the energy momentum tensor, and prove the self-consistency. Large pressure appears in the angular direction to support this black hole, which is consistent with 4D Weyl anomaly. When the black hole is formed adiabatically in the heat bath, integrating the entropy density over the interior volume reproduces the area law.

*Keywords*: Black hole evaporation; semi-classical Einstein equation; self-consistent solution; Weyl anomaly; black hole entropy.

## 1. The basic idea

In quantum mechanics, black holes evaporate by nature. We consider this property seriously and try to understand what the quantum black hole is. A physically reasonable way to do it is to investigate the time evolution of a collapsing matter and see what will be formed. In the spacetime of the formation process, which is time dependent, particle creation occurs, and its back reaction also needs to be considered. Here, it has been believed for a long time that the back-reaction effect should be negligible because the time scale of the evaporation ($\sim a^3/l_p^2$) is much longer than that of the collapse ($\sim a$), where $a \equiv 2GM$ is the Schwarzschild radius and $l_p \equiv \sqrt{\hbar G}$ the Planck length. However, it should be noted that $\Delta t \sim a^3/l_p^2$ is measured by a standing observer far from the matter while $\Delta\tau \sim a$ by a comoving observer along with the matter, and that it is not meaningful to compare two time scales measured by such two different clocks. Rather, we need to analyze time evolution of both the matter and spacetime in a single time coordinate.

Motivated by this, we provide our basic idea [1–5]. Suppose that a spherical "black hole" with mass $M = \frac{a}{2G}$ is evaporating, and the spacetime changes slowly in time. We consider what happens if we add a thin shell to it. The answer is remarkable: if we consider the time evolution of both the shell and spacetime together, the shell will never cross the Schwarzschild radius.

To see this, we assume, for simplicity, that Hawking radiation goes to infinity without reflection by such as the gravitational potential, and then describe the

spacetime outside the black hole by the outgoing Vaidya metric

$$ds^2 = -\frac{r - a(u)}{r}du^2 - 2dudr + r^2 d\Omega^2, \tag{1}$$

where $M(u) = \frac{a(u)}{2G}$ is the Bondi mass and $a(u)$ satisfies

$$\frac{da}{du} = -\frac{\sigma}{a^2}. \tag{2}$$

(See the right of Fig. 1.) $\sigma = kNl_p^2$ is the intensity of the Hawking radiation, where

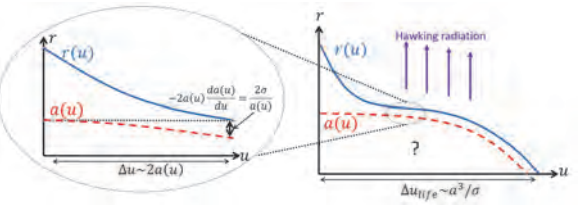



Fig. 1. Motion of a shell or a particle near the evaporating black hole.

$N$ is the degrees of freedom of fields in the theory and $k$ is an $O(1)$ constant. Note that at this stage we don't specify anything about the interior.

If the shell comes close to $a(u)$, the motion is governed by the equation for ingoing radial null geodesics:

$$\frac{dr(u)}{du} = -\frac{r(u) - a(u)}{2r(u)}, \tag{3}$$

no matter what mass and angular momentum the particles constituting the shell have[3]. Here, $r(u)$ is the radial coordinate of the shell. This reflects the fact that any particle becomes ultra-relativistic near $r \sim a$ and behaves like a massless particle. We can obtain the solution of (3)[1–4] as

$$\begin{aligned} r(u) &\approx a(u) - 2a(u)\frac{da}{du}(u) + Ce^{-\frac{u}{2a(u)}} \\ &= a(u) + \frac{2\sigma}{a(u)} + Ce^{-\frac{u}{2a(u)}} \longrightarrow a(u) + \frac{2\sigma}{a(u)} =: R(a(u)). \end{aligned} \tag{4}$$

This means the following (see the left of Fig. 1): The shell approaches the radius $a(u)$ in the time scale of $O(2a)$, but during this time, the radius $a(u)$ itself is shrinking as (2). Therefore, $r(u)$ is always apart from $a(u)$ by $-2a\frac{da}{du} = \frac{2\sigma}{a}$. Thus, the shell never crosses the radius $a(u)$ as long as the black hole evaporates in a finite time.

We then assume that, after the shell comes close to $r = a + \frac{2\sigma}{a}$, the total system composed of the black hole and shell behaves like an ordinary black hole with mass $M + \Delta M$, where $\Delta M$ is the mass of the shell. This means that the radiation emitted from the total system agrees with that from a black hole with mass $M + \Delta M$.

Now, we consider a spherical collapsing matter with a continuous distribution and regard it as many concentric shells. We can apply the above argument to each

shell recursively because its evolution is not affected by the outside shells due to the spherical symmetry. Then, we can see that the matter will shrink as a whole and become a dense object. (See Left of Fig. 2.) In the formation, no horizon will

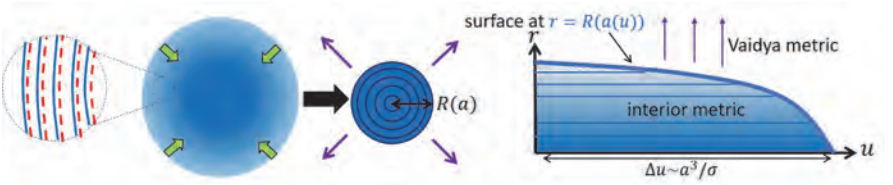



Fig. 2. A spherical matter becomes a dense object (left). The evolution of the object (right).

appear, but a surface will be formed at $r = a + \frac{2\sigma}{a}$, which is a boundary inside which the matter is distributed. From the outside, the object looks like an evaporating black hole in the ordinary picture. Thus, this should be a quantum black hole.

## 2. A self-consistent solution

In order to prove this idea, we have to analyze the dynamics of the coupled quantum system of matter and gravity. Here, we consider the semi-classical Einstein equation

$$G_{\mu\nu} = 8\pi G \langle T_{\mu\nu} \rangle, \tag{5}$$

where $\langle T_{\mu\nu} \rangle \equiv \langle \psi | T_{\mu\nu} | \psi \rangle$ contains the contribution from the collapsing matter and Hawking radiation, and $|\psi\rangle$ is any state that represents the matter at $u = -\infty$.

We solve this equation in a self-consistent manner[4]. First, we consider a physical model which assumes no reflection of Hawking radiation, and construct a candidate metric $g_{\mu\nu}$. Second, we evaluate $\langle T_{\mu\nu} \rangle$ on the obtained $g_{\mu\nu}$. More precisely, we show that all the components of $\langle T_{\mu\nu} \rangle$ can be expressed in terms of $\langle T_\mu{}^\mu \rangle$, assume conformal matter for simplicity, and use 4D Weyl anomaly formula

$$\langle T_\mu{}^\mu \rangle = \hbar c_W C_{\mu\nu\alpha\beta} C^{\mu\nu\alpha\beta} - \hbar a_W (R_{\mu\nu\alpha\beta} R^{\mu\nu\alpha\beta} - 4 R_{\mu\nu} R^{\mu\nu} + R^2). \tag{6}$$

Finally, by imposing the equality in (5), we can determine the self-consistent solution. Furthermore, to remove the artificial assumption, we introduce a phenomenological function $f$ that characterises how amount radiation is reflected by the gravitational potential or interaction with the other matters. It is defined by

$$\langle T_{UV} \rangle = f \langle T_{UU} \rangle, \tag{7}$$

where $(U, V)$ is a Kruskal-like coordinate, and $\langle T_{UU} \rangle$ is a purely radial outgoing energy flow. We expect that in the case of conformal matters the dimensionless function $f$ is constant. Thus, we can obtain the self-consistent solution of (5) as[1–4]

$$ds^2 = \begin{cases} -\frac{r - a(u)}{r} du^2 - 2 dr du + r^2 d\Omega^2, & \text{for } \ r \geq R(a(u)) \\ -\frac{2\sigma}{r^2} e^{-\frac{R(a(u))^2 - r^2}{2\sigma(1+f)}} du^2 - 2 e^{-\frac{R(a(u))^2 - r^2}{4\sigma(1+f)}} du dr + r^2 d\Omega^2, & \text{for } \ r \leq R(a(u)) \ . \end{cases} \tag{8}$$

See Right of Fig. 2. Here, the radius $a(u)$ decreases as (2) with the intensity

$$\sigma = \frac{8\pi l_p^2 c_W}{3(1+f)^2}, \tag{9}$$

where $c_W$ plays a role of $N$ in the basic idea.

This metric has interesting properties. First, the interior metric is static, which can be seen by choosing a time coordinate properly. The time-dependence of the system comes from the decreasing radius $R(a(u))$. Second, the exponential redshift factor freezes the interior in time, and the evolution occurs only in the surface region, which indicates that this object evaporates from the outside as if one peels off an onion. From a local point of view, however, the matter itself keeps shrinking by emitting the energy. In this sense, the interior will not be in equilibrium even if it is put into the heat bath. Finally, a classical limit $\hbar \to 0$ cannot be taken in the interior metric, and therefore, it is a non-perturbative solution of (5) w.r.t. $\hbar$.

## 3. Consistency check

We examine the consistency of the solution (8). First, the curvature $\mathcal{R}$ in the macroscopic region ($l_p < r \leq R(a)$) is $\mathcal{R} \sim \frac{1}{c_W l_p^2}$. Therefore, if the theory satisfies $c_W \gg 1$, the curvature is less than the Planck scale, and there is no singularity. This means that no quantum gravity effect is needed to describe the evaporating black hole as long as $a > l_p$. The Penrose diagram is given as in Fig. 3.

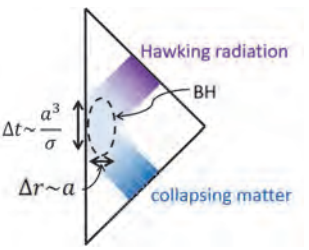



Fig. 3. The Penrose diagram for the metric (8).

The resolution of the singularity is a non-perturbative effect. The key equation in our self-consistent analysis is the trace part of (5), $G^\mu{}_\mu = 8\pi G \langle T^\mu{}_\mu \rangle$, where the r.h.s. is given by (6). This equation can be expressed schematically as $\mathcal{R} \sim N l_p^2 \mathcal{R}^2$. The interior metric of (8) has $\mathcal{R} \sim \frac{1}{N l_p^2}$, which satisfies this equation. Therefore, the metric is a result of solving (5) in a non-perturbative way w.r.t. $\hbar$.

Next, we can reproduce Hawking radiation in this picture. We can consider the collapsing process approaching to the metric (8), determine evolution of massless scalar fields in the eikonal approximation, and show that particle creation occurs in a Planck-like distribution of the time-dependent temperature $T = \frac{\hbar}{4\pi a(u)}$.

We check energy condition of the energy-momentum tensors in the interior:

$$-\langle T^t{}_t\rangle = \frac{1}{8\pi G r^2}, \quad \langle T^r{}_r\rangle = \frac{1}{8\pi G r^2}\frac{1-f}{1+f}, \quad \langle T^\theta{}_\theta\rangle = \frac{1}{8\pi G}\frac{3}{16\pi c_W l_p^2}. \tag{10}$$

The energy density is positive everywhere, and the radial pressure is so for $0 \le f < 1$. The angular pressure is much larger than the radial one, which means that the interior is anisotropic and not a fluid, and than the energy density, which breaks the dominant energy condition. This large pressure works as a outward radial pressure because of the spherical symmetry, which stabilizes the object against the gravitational attractive force.

We can see the reason why such a large pressure is produced. In (10), $\langle T^\theta{}_\theta\rangle \approx \frac{1}{2}\langle T^\mu{}_\mu\rangle$ holds, where the r.h.s. is given by (6). When derives the formula (6) in a spherically symmetric background spacetime, one expands fields into partial waves and sums up the modes with large angular momenta. Therefore, the large pressure $\langle T^\theta{}_\theta\rangle$ is a result of 4D dynamics.

We study the entropy of the object[3]. Imagine that in the heat bath we grow up the stationary black hole adiabatically. We can evaluate the entropy density inside the black hole as $s = \frac{1}{4\pi r^2}\frac{2\pi\sqrt{2\sigma}}{l_p^2}$, which means that $O(\sqrt{N})$ bits information are contained per the proper length. By summing it up over the interior volume, we can reproduce the area law:

$$S = \int_V dV\, s = \frac{A}{4l_p^2}. \tag{11}$$

This implies that the information is stored inside the black hole, not on the surface.

Finally, we make a comment on the information problem[3]. In this picture, Hawking radiation occurs inside the collapsing matter. We can expect that there is a small interaction between them, which reflects the initial information of the matter on the emitted radiation. This is reminiscent of the fact that in thermodynamics interactions bring a system into equilibrium states. Indeed, we can consider a simple model of collision of the radiation and matter, and evaluate the scattering time scale as $\Delta t \sim a \log \frac{a}{\lambda N l_p}$, where $\lambda$ is a dimensionless coupling constant.

## Acknowledgements

Y.Y. was supported by JSPS KAKENHI (No.18K13550 and 17H01148) and by RIKEN iTHEMS Program.

## References

1. H. Kawai, Y. Matsuo and Y. Yokokura, Int. J. Mod. Phys. A **28**, 1350050 (2013).
2. H. Kawai and Y. Yokokura, Int. J. Mod. Phys. A **30**, 1550091 (2015).
3. H. Kawai and Y. Yokokura, Phys. Rev. D **93**, no. 4, 044011 (2016).
4. H. Kawai and Y. Yokokura, Universe **3**, no. 2, 51 (2017).
5. P. M. Ho, H. Kawai, Y. Matsuo and Y. Yokokura, JHEP **1811**, 056 (2018).

# Inequivalent vacua in quantum field theory

Daniele Colosi

*Escuela Nacional de Estudios Superiores Unidad Morelia,*
*Universidad Nacional Autónoma de México,*
*Campus Morelia, C.P. 58190, Morelia, Michoacán, Mexico*
*E-mail: dcolosi@enesmorelia.unam.mx*

Implementing a path integral quantization of real massive scalar fields in a wide class of curved spaces we derive a general expression for the vacua in the Schrödinger representation. Using the general boundary formulation of quantum theory, we show how a symmetry of the space of classical solutions leads to a one-parameter family of quantum vacua. In this way we offer a new derivation of the alpha-vacua in de Sitter space.

*Keywords*: Vacuum, Curved spaces.

## 1. Introduction

Some aspects of the observed cosmic microwave background radiation[1] can found a description in terms of quantum fluctuations in the early universe. The study of properties of the quantum vacua for fields in curved spacetime may account for the large scale structure of our universe and has indeed received a great amount of attention.[2] In this contribution we consider a real massive scalar field in a wide class of curved spaces and construct a one parameter family of vacua. By expressing the solutions of the Klein-Gordon equation in a spacetime region $M$ in terms of the boundary configuration $\varphi$ satisfied by the field, i.e. $\phi|_{\partial M} = \varphi$, we point out the existence of two global symmetries. At the quantum level, both symmetries are broken since the resulting vacua, parametrised by a constant $\alpha$, lead to inequivalent quantum theories. In certain cases, the value of the parameter $\alpha$ is uniquely fixed according to general criteria such as demanding invariance under the group of isometry of the spacetime or other requirement the theory must satisfies.

After presenting the classical theory in Section 2, we quantize the field in Section 3 according to the general boundary formulation (GBF)[3] prescriptions. Section 4 is devoted to examples: In Minkowski and in Rindler space the parameter $\alpha$ is fixed to zero thanks to the symmetry of space. In contrast all the $\alpha$-vacua in de Sitter result to be invariant under the de Sitter group. Among these vacua the one defined by $\alpha = 0$ corresponds to the so called Bunch-Davies or Euclidean vacuum. Finally, Section 5 summarises our results and indicates future line of research.

## 2. Classical theory

The dynamics of a real scalar field with mass $m$ in a 4d Lorentzian spacetime with line element $\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu \mathrm{d}x^\nu$, in a spacetime region $M$ is described by the action $S_M[\phi] = \frac{1}{2}\int_M \mathrm{d}^4x\,\sqrt{-g}\left[g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - (m^2 + \xi R)\phi^2\right]$, where $g = \det(g_{\mu\nu})$ and $\partial_\mu$ denotes the partial derivative w.r.t. to $x^\mu$, $R$ is the scalar curvature of the spacetime, $\xi$ a dimensionless constant. We assume the following:[4]

(*i*) The region $M$ is foliated by hypersurfaces which are *not* required to be Cauchy surfaces. The foliation is described by the coordinates $(x^0, \underline{x})$ where $\underline{x} = (x^1, x^2, x^3)$. The leaves are parametrised by $x^0$ which in general is *not* a timelike coordinate.
(*ii*) The metric takes a block-diagonal form, i.e. $g^{0i} = 0 = g^{i0}$ for all $i \in \{1, 2, 3\}$.
(*iii*) Solutions of e.o.m. are obtainable by the method of separation of variables as

$$\phi(x^0, \underline{x}) = \big(X_a(x^0)Y_a\big)\,(\underline{x}) + \big(X_b(x^0)Y_b\big)\,(\underline{x}), \tag{1}$$

where $X_k(x^0)$ $(k = a, b)$ is a linear operator acting on the "initial" data $Y_k$ such that $\big(X_k(x^0)Y_k\big)\,(\underline{x})$ is a solution evaluated on the hypersurface of fixed $x^0$. We assume moreover that these operators commute among each other and are invertible.

Although restrictive, these conditions are encompassed by the most studied QFTs. If $M$ is the region $[t_1, t_2] \times \mathbb{R}^3$, solution (1) can be expressed in the form

$$\phi(x^0, \underline{x}) = \left(\frac{\mathcal{W}^{00}(x^0, t_2)}{\mathcal{W}^{00}(t_1, t_2)}\varphi_1\right)(\underline{x}) + \left(\frac{\mathcal{W}^{00}(t_1, x^0)}{\mathcal{W}^{00}(t_1, t_2)}\varphi_2\right)(\underline{x}), \tag{2}$$

where $\varphi_i = \phi|_{x^0=t_i}$ and $\mathcal{W}^{00}(t, t') = X_a(t)X_b(t') - X_a(t')X_b(t)$. The quotients have to be understood as operator acting on a mode decomposition of the boundary field configurations $\varphi_i$, $i = 1, 2$. The action of the field can be written in terms of $\varphi_i$ as

$$S_M(\varphi_1, \varphi_2) = \frac{1}{2}\int \mathrm{d}^3\underline{x}\,(\varphi_1\ \varphi_2)\,W\begin{pmatrix}\varphi_1\\ \varphi_2\end{pmatrix} \tag{3}$$

where the matrix $W$ has elements: $W^{11} = -\sqrt{\gamma|_{x^0=t_1}}\frac{\mathcal{W}^{10}(t_1,t_2)}{\mathcal{W}^{00}(t_1,t_2)}$, $W^{12} = -\sqrt{\gamma|_{x^0=t_1}}\frac{\mathcal{W}^{01}(t_1,t_1)}{\mathcal{W}^{00}(t_1,t_2)}$, $W^{21} = +\sqrt{\gamma|_{x^0=t_2}}\frac{\mathcal{W}^{10}(t_2,t_2)}{\mathcal{W}^{00}(t_1,t_2)}$ and $W^{22} = +\sqrt{\gamma|_{x^0=t_2}}\frac{\mathcal{W}^{01}(t_1,t_2)}{\mathcal{W}^{00}(t_1,t_2)}$, where $\gamma = \left|g^{(3)}g^{00}\right|$ with $g^{(3)}$ being the determinant of the induced metric on the hypersurface of constant $x^0$, and $\mathcal{W}^{10}(t, t') = (\partial_0 X_a)\,(t)X_b(t') - X_a(t')\,(\partial_0 X_b)\,(t)$ and $\mathcal{W}^{01}(t, t') = X_a(t)\,(\partial_0 X_b)\,(t') - (\partial_0 X_a)\,(t')X_b(t)$. The space of solutions of e.o.m. is equipped with the symplectic structure $\Omega(\phi_1, \phi_2) = \frac{1}{2}\int \mathrm{d}^3\underline{x}\sqrt{\left|g^{(3)}g^{00}\right|}$ $(\phi_1\partial_0\phi_2 - \phi_2\partial_0\phi_1)$, which is independent of the leaf of the foliation.

The key observation of this contribution consists in noticing that the classical solution (2) and the action (3) are left invariant by the global transformations $\Pi(X_a, X_b) = (\beta X_a, \beta^{-1}X_b)$ and $\tilde{\Pi}(X_a, X_b) = (\beta X_b, \beta^{-1}X_a)$ with $\beta \in \mathbb{R}\setminus\{0\}$. Since the parameter $\beta$ is real, both transformations project solutions of the Klein-Gordon equation into solutions. $\Pi$ preserves the symplectic structure $\Omega$, which in contrast changes sign under $\tilde{\Pi}$. In what follows we will consider only the transformation $\Pi$. Notice that the invariance of solutions, action and symplectic structure are a direct consequence of the representation of the field and the action in terms of the boundary configurations $\varphi_i$, $i = 1, 2$. This invariance will be broken at the quantum level, and the QFTs associated to different value of $\beta$ will result to be inequivalent.

## 3. Vacum state in the Schrödinger representation

The quantization of the theory is implemented in terms of the Feynman path integral combined with the Schrödinger representation in accordance with the axioms of

the GBF. In this formulation, as in TQFT, a Hilbert space is associated to each hypersurface. It must be emphasised that the novelty of the GBF is the possibility to implement this association for arbitrary hypersurfaces, not only Cauchy ones. Thus, for a hypersurface $\Sigma$ we denote the corresponding Hilbert space with $\mathcal{H}_\Sigma$. Quantum states are given by functional of field configurations, i.e. $\psi_\Sigma(\varphi) \in \mathcal{H}_\Sigma$ where $\varphi = \phi|_\Sigma$. The inner product between two states $\psi_\Sigma$ and $\eta_\Sigma$ in $\mathcal{H}_\Sigma$ is given by the integral $\int \mathcal{D}\varphi\, \overline{\psi_\Sigma(\varphi)}\eta_\Sigma(\varphi)$, where the overline indicates complex conjugation.

The quantum dynamics that unfolds in the spacetime region $M$ of interest is encoded in the so called field propagator: $Z_M(\varphi) = \int_{\phi|_{\partial M}=\varphi} \mathcal{D}\phi\, e^{\mathrm{i}S_M[\phi]}$, where the integration is extended over all spacetime configurations $\phi$ that reduce to $\varphi$ on the boundary $\partial M$. For the interval region presented in Section 2, the Hilbert space associated to the boundary decomposes as $\mathcal{H}_{\partial M} = \mathcal{H}_{t_1} \otimes \mathcal{H}^*_{t_2}$, where the star denotes the dualization of the Hilbert space due to the change in the orientation of the hypersurface at $x^0 = t_2$. The field propagator in this region reads $Z_M(\varphi_1, \varphi_2) = N_M e^{\mathrm{i}S_M(\varphi_1,\varphi_2)}$, where $N_M = \int \mathcal{D}\phi\, e^{\mathrm{i}S_M[\phi]}$ with $\phi|_{\partial M} = 0$, the action $S_M(\varphi_1, \varphi_2)$ is (3). Then, a state $\psi_{t_1} \in \mathcal{H}_{t_1}$ is evolved to a state in $\mathcal{H}_{t_2}$ by the field propagator as $\int \mathcal{D}\varphi_1\, \psi_{t_1}(\varphi_1) Z_M(\varphi_1, \varphi_2) \in \mathcal{H}_{t_2}$. In the GBF the vacuum state is obtained by requiring the functional invariance under such evolution; its general structure[4] is $\psi_{t,0}(\varphi) = C_t \exp\left(\frac{\mathrm{i}}{2}\int \mathrm{d}^2\underline{x}\, \varphi(\underline{x})\sqrt{\left|g^{(3)}g^{00}\right|}\frac{\partial_0 \Upsilon(x^0)}{\Upsilon(x^0)}\Big|_{x^0=t} \varphi(\underline{x})\right)$, where $C_t$ is a normalisation factor and the operator $\Upsilon$ is a linear combination of $X_a$ and $X_b$ with complex coefficients: $\Upsilon(x^0) = c_a X_a(x^0) + c_b X_b(x^0)$, where the coefficients $c_a$ and $c_b$ must satisfy the condition: $\Im(\overline{c_a} c_b) \neq 0$. In general the vacuum state is not invariant under the action of $\Pi$ since it transforms the vacuum in the new state

$$\psi^\beta_{t,0}(\varphi) = C_t \exp\left(\frac{\mathrm{i}}{2}\int \mathrm{d}^2\underline{x}\, \varphi(\underline{x})\sqrt{\left|g^{(3)}g^{00}\right|}\frac{\partial_0 \Upsilon_\beta(x^0)}{\Upsilon_\beta(x^0)}\Big|_{x^0=t} \varphi(\underline{x})\right), \tag{4}$$

with $\Upsilon_\beta(x^0) = c_a \beta X_a(x^0) + c_b \beta^{-1} X_b(x^0)$. The transformation $\Pi$ applied to $\Upsilon$ can be interpreted as a rescaling of the coefficients $c_a$ and $c_b$. This allows us to conclude that the state (4) is an acceptable vacuum state for the field. Moreover, the sign of the parameter $\beta$ does not play any role for this state and consequently its value can be restricted to be positive, i.e. $\beta = e^\alpha$ with $\alpha \in \mathbb{R}$. From now on we write $\Upsilon_\alpha$.

Notice that the theories defined by the vacuum (4) are inequivalent for different values of the parameter $\alpha$. Indeed, the Feynman propagator determined by (4) reduces to the standard Feynman propagator only if $\alpha = 0$. We conclude that for non vanishing values of the parameter $\alpha$ we obtain inequivalent quantum theories.

## 4. Fixing $\alpha$

For certain QFTs, the value of $\alpha$ can be fixed uniquely in terms of some general criteria. We present some examples, giving in particular the explicit expression of the vacuum states for the field in Minkowski, Rindler and de Sitter spaces. In the case of Minkowski and Rindler, the vacuum is unique thanks to the symmetry of the space. In de Sitter we specify the criteria that distinct the different vacua.

### 4.1. *Minkowski space*

Minkowski space is invariant under the Poincaré group. Requiring the same invariance for the vacuum state fixes the value of $\alpha$ to be zero. In the case of an interval region where $x^0$ is the Minkowski time $t$, we have that $X_a(t) = \cos(\omega t)$ and $X_b(t) = \sin(\omega t)$, with $\omega = \sqrt{-\partial^2_{\underline{x}} + m^2}$. Thus, the state (4) is left unchanged under the transformation $t \to t' = t + t_0$ only if $\alpha = 0$ and $c_a = 1, c_b = \mathrm{i}$. With this choice (4) reduces to the standard Minkowski vacuum[5–7], $\psi_0 = C \exp\left(-\frac{1}{2}\int \mathrm{d}^3\underline{x}\,\varphi(\underline{x})\omega\varphi(\underline{x})\right)$.

### 4.2. *Rindler space*

The line element of 2d Rindler space is $\mathrm{d}s^2 = \rho^2\mathrm{d}\eta^2 - \mathrm{d}\rho^2$ where $x^0 = \rho\sinh\eta$ and $x^1 = \rho\cosh\eta$, $x^0$ and $x^1$ being Minkowski coordinates. The same result as in Minkowski is recovered in Rindler space since, for the region $M = [\eta_1, \eta_2] \times \mathbb{R}^+$, $X_a(\eta) = \cos(\omega\eta)$ and $X_b(\eta) = \sin(\omega\eta)$ with $\omega = \sqrt{-(\rho\partial_\rho)^2 + m^2}$.[8] Invariance of the vacuum under a temporal translation is achieved only for $\alpha = 0$, and the vacuum on the hypersurface of constant $\eta$ results to be $\psi_0 = C\exp\left(-\frac{1}{2}\int \mathrm{d}\rho\,\frac{1}{\rho}\varphi(\rho)\omega\varphi(\rho)\right)$.

### 4.3. *de Sitter space*

The de Sitter group does not fix the value of the parameter $\alpha$. To see this we consider the de Sitter metric in the following coordinate system $\mathrm{d}s^2 = \frac{R^2}{(x^0)^2}\left((\mathrm{d}x^0)^2 - (\mathrm{d}\underline{x})^2\right)$, where $x^0 \in ]0, \infty[$, $\underline{x} \in \mathbb{R}^3$, and $R$ denotes the inverse of the Hubble constant. This coordinate system covers half of de Sitter space and the other half is covered by extending $x^0$ from $-\infty$ to 0. Solutions of e.o.m are parametrized by $X_a(x^0) = (x^0)^{3/2}J_\nu(kx^0)$ and $X_b(x^0) = (x^0)^{3/2}Y_\nu(kx^0)$, where $k = |\underline{k}|$,[a] and $J_\nu(z)$ and $Y_\nu(z)$ are the Bessel functions of the first and second kind respectively, with index $\nu = \sqrt{\frac{9}{4} - (mR)^2}$.[9] Vacuum states on a surface of constant $x^0$ read

$$\psi_0(\varphi) = C\exp\left(\frac{\mathrm{i}}{2}\int \mathrm{d}^3\underline{x}\,\varphi(\underline{x})\frac{R^2}{(x^0)^2}\frac{\partial_0\Upsilon_\alpha(x^0)}{\Upsilon_\alpha(x^0)}\varphi(\underline{x})\right), \tag{5}$$

with $\Upsilon_\alpha(x^0) = (x^0)^{3/2}\left(e^\alpha J_\nu(kx^0) + \mathrm{i}e^{-\alpha}Y_\nu(kx^0)\right)$, $\alpha \in \mathbb{R}$.

#### 4.3.1. Isometry of space

The de Sitter metric is invariant under 10 transformations: 3 spatial rotations, 3 spatial translations and the following transformation ($H = R^{-1}$, $x^0_0$ is a constant)

$$x^0 \to (x^0)' = e^{-Hx^0_0}x^0, \qquad \underline{x} \to \underline{x}' = e^{-Hx^0_0}\underline{x}. \tag{6}$$

[a] In order for the solution to be well defined in the whole spacetime, as well as in the interval type region, the components of the 3-vector $\underline{k}$ have to be real, so $k > 0$.

A vacuum state on the surface defined by $(x^0)' = e^{-Hx_0^0}x^0$ depends on the configuration $\varphi'(\underline{x}) = \int \frac{\mathrm{d}^3\underline{k}}{(2\pi)^{3/2}}\varphi'(\underline{k})e^{\mathrm{i}\underline{k}\cdot\underline{x}}$ and is left invariant by (6) if the coefficients $\varphi'$ satisfy the equality $\varphi'(\underline{k}e^{-Hx_0^0}) = \varphi(\underline{k})$. Therefore the invariance under the transformation (6) does not fix the value of $\alpha$. The one parameter family of vacua (5) are the so called $\alpha$-vacua in de Sitter.[10]

4.3.2. Minkowski limit

de Sitter space reduces to Minkowski space in the limit where the curvature is infinite, corresponding to the limits $R \to \infty$ and $x^0 \to \infty$ such that $\frac{R}{x^0}$ is kept constant. Using the asymptotic expression of the Bessel functions[11] we see that the Minkowski vacuum is recovered only if $\alpha = 0$. This is the Bunch-Davies vacuum for de Sitter.[12] Moreover, only for $\alpha = 0$ the singularities of the Feynman propagator in the limit $x \to x'$ are of the Hadamard form.

4.3.3. Energy expectation value

Another criterium to select a vacuum state is requiring that expectation value of the energy is the minimum when evaluated on the vacuum state. Denoting with $\rho^{T_{00}}(\psi_0)$ the expectation value of the 00-component of the energy momentum tensor on a vacuum state defined at a constant $x^0$ hypersurface in the region $M$, we have, in the case of a minimally coupled field, $\rho^{T_{00}}(\psi_0^\alpha) = \frac{1}{2\mathrm{i}}\lim_{x\to x'}(\partial_0\partial_0' + \partial_{\underline{x}}\partial_{\underline{x}'} + m^2)G_M^\alpha(x,x')$, where $\partial_0' = \partial/\partial(x_0)'$ and $G_M^\alpha$ is the Feynman propagator. This criterium select the Bunch-Davies vacuum in the asymptotic region $x^0 \gg 1$.

4.3.4. Wick rotation

The Bunch-Davies vacuum is also recovered through a Wick rotation, i.e., by performing an analytic continuation of the variable in terms of which the dynamics is described, namely $x^0 \to \mathrm{i}x^0$, in the field propagator, then taking the limit $x^0 \to \infty$ and setting the boundary configuration at infinity (say $\varphi_2$) to zero, in order to obtain the vacuum state[13], $\lim_{x^0\to\infty} Z_M(\varphi_1,\varphi_2)\big|_{\mathrm{i}x^0} = c\psi_0(\varphi_1)$, where $c$ is a constant.

## 5. Summary and outlook

For a wide class of field theories in curved spaces, we recognise the existence of two transformations under which classical solutions of the equation of motion are left invariant when expressed in terms of boundary field configurations. Also the action evaluated on these solutions are invariant under these transformations. Finally, one of the two transformations results to be a symmetry also for the symplectic structure associated with the space of solutions. Using the Schrödinger-Feynman quantization scheme of the GBF, we show that these classical symmetries are broken at the quantum level. Indeed, they generate a one parameter family of infinite vacua. For de Sitter space, these states are the well known $\alpha$-vacua.

It is important to underline the novelty of this approach for the derivation of the de Sitter $\alpha$-vacua. In literature theses states are usually obtained in terms of Bogoliubov transformation of the basis where the field is expanded. Moreover they are interpreted as squeezed states over the Bunch-Davies vacuum. An alternative, but equivalent derivation is based on the antipodal map of de Sitter that send a field evaluated at a point $x$ into a field evaluated at the antipodal point to $x$. Although our approach is similar from a mathematical point of view to the functional formalism[14], we underline the classical symmetries of the space of solutions, when expressed in terms of the boundary field configurations, from which the $\alpha$-vacua arise. This representation is the more adequate for the quantization of the theory according to the Schrödinger-Feynman prescription within the GBF. Moreover, the result presented here constitutes a successful test for the GBF itself.

In the GBF another quantization scheme has been implemented and the representation used there is the holomorphic one.[15] Moreover a one-to-one correspondence exists between vacuum states in the Schrödinger representation and complex structures used in the holomorphic representation.[16] A natural generalization of the result presented here is to investigate the representation of the transformation $\Pi$ in terms of the complex structure.

**Acknowledgments**

This work was partially supported by UNAM-PAPIIT project grant IA-106418.

**References**

1. Planck collaboration, Astron.Astrophys. 594 (2016) A20, arXiv:1502.02114.
2. N. D. Birrell and P. C. W. Davies, Quantum Fields in Curved Space, Cambridge Univ. Press, Cambridge, 1982; V. F. Mukhanov and S. Winitzki, Introduction to Quantum Effects in Gravity, Cambridge Univ. Press, Cambridge, 2007; L. Parker and D. Toms, Quantum field theory in curved spacetime, Cambridge Univ. Press, Cambridge, 2009.
3. R. Oeckl, Phys. Lett. **B 575** (2003), 318–324, arXiv:hep-th/0306025; R. Oeckl, Adv. Theor. Math. Phys. **12** (2008), 319–352, arXiv:hep-th/0509122.
4. D. Colosi, M. Dohse, J. Geom. Phys. 114 (2017) 65-84, arXiv:1011.2243.
5. R. Oeckl, Phys. Rev. D **73** (2006), 065017, arXiv:hep-th/0509123.
6. D. Colosi and R. Oeckl, Phys. Lett. **B 665** (2008), 310–313, arXiv:0710.5203.
7. D. Colosi and R. Oeckl, Phys. Rev. D **78** (2008), 025020, arXiv:0802.2274.
8. D. Colosi and D. Raetzel, Phys. Rev. D **87** (2013), 125001, arXiv:1303.5873.
9. D. Colosi, arXiv:1010.1209.
10. N. A. Chernikov and E. A. Tagirov, Ann. Inst. Henri Poincaré, A9 (1968), 109; C. Schomblond and P. Spindel, Ann. Inst. Henri Poincaré, A25 (1976), 1, 67; E. Mottola, Phys. Rev. D **31** (1985), 754; B. Allen, Phys. Rev. D **32** (1985), 3136–3149; J. de Boer, V. Jejjala and D. Minic, Phys. Rev. D **71** (2005) 044013, [hep-th/0406217].

11. I. S. Gradshteyn and I. M. Ryzhik, Table of Integrals, Series, and Products, Academic Press, New York, 2007.
12. T. S. Bunch and P. C. W. Davies, Proc. R. Soc. Lond. A **360** (1979), 117–134.
13. D. Colosi and R. Oeckl, *Vacuum states as Lagrangian subspaces*, in preparation.
14. C. J. C. Burges, Nucl. Phys. B **247**, 533 (1984.); R. Laflamme, Phys. Lett. B **198** (1987), 156; J. J. Halliwell, Phys. Rev. D **43** (1991), 2590; R. Floreanini, C. T. Hill and R. Jackiw, Annals Phys. **175** (1987), 345.
15. R. Oeckl, SIGMA **8** (2012) 050, arXiv:1009.5615v3.
16. R. Oeckl, J. Math. Phys. **53** (2012), 072301, arXiv:1109.5215. D. Colosi and M. Dohse, Int. J. Theor. Phys. **56** (2017) no.11, 3359, arXiv:1701.06053 [math-ph].

# Primordial gravitational waves in bouncing universe

Asuka Ito and Jiro Soda

*Department of Physics, Kobe University, Kobe 657-8501, Japan*

Both inflationary and ekpyrotic scenarios can account for the origin of the large scale structure of the universe. It is often said that detecting primordial gravitational waves is the key to distinguish both scenarios. We show that this is not true if the gauge kinetic function is present in the ekpyrotic scenario. In fact, primordial gravitational waves sourced by the gauge field can be produced in an ekpyrotic universe. We also study scalar fluctuations sourced by the gauge field and show that it is negligible compared to primordial gravitational waves. This comes from the fact that the fast roll condition holds in ekpyrotic models.

## 1. Introduction

Inflation has succeeded in solving several issues in big bang cosmology and explaining the temperature anisotropy of the cosmic microwave background radiation (CMB) and the large scale structure of the universe. However, it is known that bouncing universe models[1] such as the ekpyrotic scenario[2] based on superstring theory[3] can do the same job[4].[a] Therefore, it is important to clarify which scenario is actually realized in the early stage of the universe.

In the ekpyrotic scenario, the primordial fluctuations are produced in a slowly contracting (ekpyrotic) phase. The spectrum of the scalar and tensor vacuum fluctuations becomes blue-tilted in the phase. We therefore need an additional scalar field to explain the temperature anisotropy of the CMB[6]. Moreover, in the ekpyrotic scenario, the amplitude of primordial gravitational waves[7] is quite small and practically unobservable[8]. Hence, it is often said that, if we could detect the primordial gravitational waves, we would be able to disprove the ekpyrotic scenario. However, if there could exist another mechanism for producing gravitational waves in the ekpyrotic scenario, the story would be completely different. Indeed, we show that there exists a mechanism for producing abundant gravitational waves in the ekpyrotic phase.

The key is the presence of magnetic fields in the early universe. Observationally, there are several evidences for magnetic fields to exist on various cosmological scales[9]. Although the origin of primordial magnetic fields is unknown, the presence of magnetic fields on extra galactic scales[10] implies that the seed of magnetic fields must be produced in the early universe. Notably, there are attempts to make primordial magnetic fields with the gauge kinetic function in an inflationary universe[11] or in a bouncing universe[12].

This report is a review of our previous paper[13]. We first show that scale invariant magnetic fields can be produced in the ekpyrotic phase in the presence of the gauge kinetic function. Next, we show that the magnetic fields can be a source of abundant

[a] The pre-big bang scenario is also a kind of the models[5]. Our conclusion could apply to it too.

gravitational waves (such mechanism works also in inflation[14]). It turns out that the gravitational wave spectrum is nearly scale invariant (slightly blue) at the end of the ekpyrotic phase. Hence, it is difficult to discriminate between inflation and the ekpyrotic scenario by merely detecting primordial gravitational waves. We also study scalar fluctuations induced by the magnetic fields and show that the tensor to scalar ratio should be more than unity, which implies that scalar fluctuations in the CMB should be dominated by quantum fluctuations produced by an additional scalar field as is often assumed in the ekpyrotic scenario.

## 2. Ekpyrotic phase

The ekpyrotic scenario can be described by a four-dimensional effective theory with a scalar field $\phi$ moving in an effective potential $V(\phi)$ specified below. The action reads

$$S = \int d^4x\sqrt{-g}\left[\frac{M_{pl}^2}{2}R - \frac{1}{2}(\partial_\mu\phi)(\partial^\mu\phi) - V(\phi)\right] , \tag{1}$$

where $M_{pl}$ represents the reduced Planck mass, $g$ is the determinant of the metric $g_{\mu\nu}$ , and $R$ is the Ricci scalar. The scalar field represents the separation $l$ between two branes $l \sim e^{\phi}$. The contracting universe ($\dot{\phi} < 0$) is connected to the expanding universe ($\dot{\phi} > 0$) through a bounce (a collision of two branes). The scalar and tensor vacuum fluctuations are produced in the contracting phase where the scalar field rolls down a negative steep potential

$$V(\phi) \simeq V_0 e^{\lambda\frac{\phi}{M_{pl}}} , \tag{2}$$

where $V_0$ is a negative constant. Note that $\lambda$ is also negative and satisfies the fast roll condition $|\lambda| \gg 1$ to keep isotropy of the universe. Thus, we can take an isotropic metric ansatz in this phase as

$$ds^2 = a(\tau)\left[-d\tau^2 + dx^2 + dy^2 + dz^2\right] , \tag{3}$$

where we used a conformal time $\tau$. It is straightforward to derive scaling solutions from Eqs. (1)–(3):

$$a(\tau) = a_{end}\left(\frac{-\tau}{-\tau_{end}}\right)^{\frac{2}{\lambda^2-2}} , \quad \frac{\phi(\tau)}{M_{pl}} = \phi_0 - \frac{2\lambda}{\lambda^2-2}\ln(-M_{pl}\tau) , \tag{4}$$

where $\tau_{end}$ ($< 0$) and $a_{end}$ represent the moment and the scale factor at the end of the ekpyrotic phase, respectively. The obtained vacuum scalar and tensor power spectrums are blue-tilted, so that we need an additional scalar field to explain the CMB observation[6]. Then, the ekpyrotic scenario predicts the nearly scale invariant scalar power spectrum and the blue-tilted tensor power spectrum. The situation is different from inflation where both spectra are nearly scale invariant.

## 3. Scale Invariant Magnetic Fields

Let us consider a gauge field coupled with $\phi$ in the contracting phase as

$$-\frac{1}{4}f^2(\phi)F_{\mu\nu}F^{\mu\nu} \ , \tag{5}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength of the gauge field and $f(\phi)$ represents the gauge kinetic function. Now, we take the gauge kinetic function as exponential type functional form which is ubiquitous in models obtained from dimensional reduction

$$f(\phi) = f_0 e^{\rho\frac{\phi}{M_{pl}}} \ , \tag{6}$$

where it has been set to be unity at the end of the ekpyrotic phase, and then there is no strong coupling problem. As is discussed in Ref. 13, one can obtain a expression of Fourier coefficient of magnetic fields defined by $\vec{B}(\tau,\boldsymbol{x}) \equiv \frac{f}{a^2}\left(\nabla\times\vec{A}(\tau,\boldsymbol{x})\right)$ on the background (4) as

$$\mathcal{B}_k(\tau) = \frac{3\sqrt{2}}{8}(\lambda^2-2)^2 k^{-3/2}\left(\frac{-\tau}{-\tau_{end}}\right)^{-\frac{2\lambda^2}{\lambda^2-2}} H_{end}^2 \ . \tag{7}$$

Here, we have eliminated $\rho$ by requiring the scale invariance of magnetic fields at the end of the ekpyrotic phase. $\tau_{end}$ and $H_{end}$ are the conformal time and the Hubble parameter at the end of the ekpyrotic phase, respectively. For example, if we set $H_{end} = 10^{-5}M_{pl}$ and $\lambda = -17$, where back reaction from electromagnetic fields is negligible, the amplitude of the magnetic field at the end of the ekpyrotic phase is about $10^{49}$ G. Thus, the cosmological magnetic fields observed at present can be produced in the ekpyrotic scenario[9]. Remarkably, such magnetic fields can also induce abundant primordial gravitational waves.

## 4. Gravitational waves from Magnetic Fields

As is shown in Refs. 13, 14, one can get the tensor sector of the action (1) in the presence of the gauge field (5) as

$$S_{GW} = \int d\tau d^3x \left[\frac{M_{pl}^2}{8}a^2\left(h'_{ij}h'^{ij} - \partial_k h_{ij}\partial_k h^{ij}\right) + \frac{1}{2}a^4\left(E_iE_j + B_iB_j\right)h^{ij}\right] \ , \tag{8}$$

where $h_{ij}$ is the transverse traceless tensor, namely, gravitational waves. We see that electric and magnetic fields work as a source of gravitational waves. From above action with scale invariant magnetic fields (7), one can estimate the power spectrum of gravitational waves:

$$P_s(k) \simeq \frac{27}{16\pi^4}\lambda^8\left(\frac{H_{end}}{M_{pl}}\right)^4 \ln\left[\frac{k}{k_{in}}\right] \ . \tag{9}$$

There is a factor $\left(\frac{H_{end}}{M_{pl}}\right)^4$ in the spectrum (9) because of the nonlinear contribution of the magnetic fields (7). One can see that sourced gravitational waves have a

nearly scale invariant spectrum. This conclusion is different from the well-known blue-tilted spectrum in the ekpyrotic scenario[8]. Most importantly, there appears a factor $\lambda^8$ in $P_s(k)$. For example, if we set $H_{end} = 10^{-5}M_{pl}$ and $\lambda = -17$ to produce the observed magnetic field, the amplitude of the power spectrum is about $10^{-11}$. This is comparable with the gravitational waves in the inflationary universe $\sim \left(\frac{H_{end}}{\pi M_{pl}}\right)^2$. Therefore, we can not discriminate between inflation and the ekpyrotic scenario just by detecting primordial gravitational waves.

## 5. Scalar Fluctuations from Magnetic Fields

The power spectrum of scalar fluctuations induced by magnetic fields can be calculated as same as tensor fluctuations. The result is

$$\mathcal{P}_s(k) \simeq \frac{243}{1024\pi^4}\lambda^8 \left(\frac{H_{end}}{M_{pl}}\right)^4 \ln\left[\frac{k}{k_{in}}\right] \ . \tag{10}$$

From Eqs. (9) and (10), the tensor to scalar ratio $r_{source}$ is given by

$$r_{source} \simeq 7 \ . \tag{11}$$

Since the tensor to scalar ratio becomes larger than unity, we can say that the scalar fluctuations sourced by the scale invariant magnetic field are negligible in the ekpyrotic scenario. Therefore, the ekpyrotic model with a gauge field is compatible with the CMB data.

## 6. Conclusion

We studied the role of the gauge kinetic function in the ekpyrotic scenario and showed that abundant gravitational waves sourced by the gauge field can be produced. As a demonstration, we first showed that scale invariant magnetic fields can be produced in the ekpyrotic phase. It turned out that the magnetic fields induce nearly scale invariant gravitational waves (slightly blue) and the amplitude could be comparable with that of the inflationary universe. It turned out that it is difficult to disprove the ekpyrotic scenario by detecting primordial gravitational waves. In order to distinguish both scenarios, it is necessary to look at the details of the spectrum such as the tilt of the spectrum or the non-gaussianity[15]. Observing the distinction of higher order scalar perturbations is also important[16]. We should mention that the idea of finding an ekpyrotic model with observable gravitational waves on CMB scales using sourced fluctuations was put forward for the first time in Ref. 17 by investigating a different model with explicit parity violation. Our model has no explicit parity violation. Moreover, we also showed that the scalar fluctuations induced by the magnetic field are smaller than the sourced gravitational waves. Generally, as far as the fast roll condition is satisfied, the tensor to scalar ratio becomes more than unity in any ekpyrotic models with the gauge kinetic function. Therefore, our scenario would be compatible with the CMB data

provided that nearly scale invariant scalar fluctuations are produced in a standard way with an additional scalar field[6].

It should be noted that we must check the non-gaussianity of the primordial scalar fluctuations in the present model[18]. Moreover, we should consider a bounce process from contracting to expanding to connect the spectrum at the end of the ekpyrotic phase with observables. We have not looked into this issue in this paper since the mechanism is model dependent and the detailed analysis is beyond the scope of this paper[1]. However, actually, although we fixed the parameters such as $\rho, \lambda, H_{end}$ for simplicity in this paper, we can tune these parameters in our scenario so that our conclusion becomes valid for any ekpyrotic bouncing models. Therefore, our conclusion is robust.

## Acknowledgments

A. I. was supported by Grant-in-Aid for JSPS Research Fellow and JSPS KAKENHI Grant No.JP17J00216. J. S. was in part supported by JSPS KAKENHI Grant Numbers JP17H02894, JP17K18778, JP15H05895, JP17H06359, JP18H04589. J. S. is also supported by JSPS Bilateral Joint Research Projects (JSPS-NRF collaboration) String Axion Cosmology.

## References

1. D. Battefeld and P. Peter, Phys. Rept. **571** (2015) 1 [arXiv:1406.2790 [astro-ph.CO]]; R. Brandenberger and P. Peter, arXiv:1603.05834 [hep-th].
2. J. Khoury, B. A. Ovrut, P. J. Steinhardt and N. Turok, Phys. Rev. D **64** (2001) 123522 [hep-th/0103239].
3. P. Horava and E. Witten, Nucl. Phys. B **475** (1996) 94 [hep-th/9603142]; P. Horava and E. Witten, Nucl. Phys. B **460** (1996) 506 [hep-th/9510209].
4. A. Ijjas and P. J. Steinhardt, Class. Quant. Grav. **33** (2016) no.4, 044001 [arXiv:1512.09010 [astro-ph.CO]].
5. M. Gasperini and G. Veneziano, Phys. Rept. **373** (2003) 1 [hep-th/0207130]; M. Gasperini and G. Veneziano, Astropart. Phys. **1** (1993) 317 [hep-th/9211021]; R. Brustein, M. Gasperini, M. Giovannini and G. Veneziano, Phys. Lett. B **361** (1995) 45 [hep-th/9507017].
6. M. Li, Phys. Lett. B **724** (2013) 192 [arXiv:1306.0191 [hep-th]]; A. M. Levy, A. Ijjas and P. J. Steinhardt, Phys. Rev. D **92** (2015) no.6, 063524 [arXiv:1506.01011 [astro-ph.CO]]; A. Notari and A. Riotto, Nucl. Phys. B **644** (2002) 371 [hep-th/0205019].
7. A. A. Starobinsky, JETP Lett. **30** (1979) 682 [Pisma Zh. Eksp. Teor. Fiz. **30** (1979) 719].
8. L. A. Boyle, P. J. Steinhardt and N. Turok, Phys. Rev. D **69** (2004) 127302 [hep-th/0307170].
9. K. Subramanian, arXiv:1504.02311 [astro-ph.CO]; R. Durrer and A. Neronov, Astron. Astrophys. Rev. **21** (2013) 62 [arXiv:1303.7121 [astro-ph.CO]].

10. W. Chen, J. H. Buckley and F. Ferrer, Phys. Rev. Lett. **115** (2015) 211103 [arXiv:1410.7717 [astro-ph.HE]]; W. Chen, B. D. Chowdhury, F. Ferrer, H. Tashiro and T. Vachaspati, Mon. Not. Roy. Astron. Soc. **450** (2015) no.4, 3371 [arXiv:1412.3171 [astro-ph.CO]]; K. Takahashi, M. Mori, K. Ichiki, S. Inoue and H. Takami, Astrophys. J. **771** (2013) L42 [arXiv:1303.3069 [astro-ph.CO]]; A. M. Taylor, I. Vovk and A. Neronov, Astron. Astrophys. **529** (2011) A144 [arXiv:1101.0932 [astro-ph.HE]]; W. Essey, S. Ando and A. Kusenko, Astropart. Phys. **35** (2011) 135 [arXiv:1012.5313 [astro-ph.HE]]; K. Dolag, M. Kachelriess, S. Ostapchenko and R. Tomas, Astrophys. J. **727** (2011) L4 [arXiv:1009.1782 [astro-ph.HE]]; S. Ando and A. Kusenko, Astrophys. J. **722** (2010) L39 [arXiv:1005.1924 [astro-ph.HE]]; F. Tavecchio, G. Ghisellini, L. Foschini, G. Bonnoli, G. Ghirlanda and P. Coppi, Mon. Not. Roy. Astron. Soc. **406** (2010) L70 [arXiv:1004.1329 [astro-ph.CO]]; A. Neronov and I. Vovk, Science **328** (2010) 73 [arXiv:1006.3504 [astro-ph.HE]].
11. B. Ratra, Astrophys. J. **391** (1992) L1; M. Giovannini, Phys. Lett. B **659** (2008) 661 [arXiv:0711.3273 [astro-ph]]; K. Bamba and M. Sasaki, JCAP **0702** (2007) 030 [astro-ph/0611701]; J. Martin and J. Yokoyama, JCAP **0801** (2008) 025 [arXiv:0711.4307 [astro-ph]]; V. Demozzi, V. Mukhanov and H. Rubinstein, JCAP **0908** (2009) 025 [arXiv:0907.1030 [astro-ph.CO]]; S. Kanno, J. Soda and M. a. Watanabe, JCAP **0912** (2009) 009 [arXiv:0908.3509 [astro-ph.CO]]; R. J. Z. Ferreira, R. K. Jain and M. S. Sloth, JCAP **1310** (2013) 004 [arXiv:1305.7151 [astro-ph.CO]]; T. Kobayashi, JCAP **1405** (2014) 040 [arXiv:1403.5168 [astro-ph.CO]]; T. Fujita and R. Namba, arXiv:1602.05673 [astro-ph.CO].
12. F. A. Membiela, Nucl. Phys. B **885** (2014) 196 [arXiv:1312.2162 [astro-ph.CO]]; L. Sriramkumar, K. Atmjeet and R. K. Jain, JCAP **1509** (2015) no.09, 010 [arXiv:1504.06853 [astro-ph.CO]]; D. Chowdhury, L. Sriramkumar and R. K. Jain, arXiv:1604.02143 [gr-qc].
13. A. Ito and J. Soda, Phys. Lett. B **771**, 415 (2017) [arXiv:1607.07062 [hep-th]].
14. A. Ito and J. Soda, JCAP **1604** (2016) no.04, 035 [arXiv:1603.00602 [hep-th]].
15. M. Tsuneto, A. Ito, T. Noumi and J. Soda, JCAP **1903**, no. 03, 032 (2019) [arXiv:1812.10615 [gr-qc]].
16. X. Chen, M. H. Namjoo and Y. Wang, arXiv:1601.06228 [hep-th]; X. Chen, M. H. Namjoo and Y. Wang, JCAP **1602** (2016) no.02, 013 [arXiv:1509.03930 [astro-ph.CO]].
17. I. Ben-Dayan, arXiv:1604.07899 [astro-ph.CO].
18. A. Fertig, J. L. Lehners and E. Mallwitz, Phys. Rev. D **89** (2014) no.10, 103537 [arXiv:1310.8133 [hep-th]]; A. Ijjas, J. L. Lehners and P. J. Steinhardt, Phys. Rev. D **89** (2014) no.12, 123520 [arXiv:1404.1265 [astro-ph.CO]].

# Casimir effect and free fall in a Schwarzschild black hole

Francesco Sorge*

*I.N.F.N. sezione di Napoli*
*Complesso Universitario Monte S. Angelo, via Cintia, ed.6*
*Napoli, I-80126, Italy*
** E-mail: sorge@na.infn.it*

Justin H. Wilson†

*Institute of Quantum Information and Matter*
*and Department of Physics*
*CALTECH, CA*
*† E-mail: jwilson@caltech.edu*

Scalar field vacuum energy and particle creation in a 3D Casimir apparatus, freely falling in the Schwarzschild spacetime, are considered in the reference frame of a comoving observer. Following Schwinger's proper time approach, Casimir energy is evaluated from the effective action, resulting in a small correction to the flat spacetime case. Besides, a tiny amount of quanta excited out from the vacuum is found. Both effects are discussed, drawing attention to the role of the underlying spacetime dimensionality.

*Keywords*: Casimir effect; quantum fields in curved space-times; particle creation.

## 1. Introduction

Casimir effect[1–3] is one of the most intriguing as well as fascinating effects in the microscopical quantum world. Roughly speaking, it originates from a distortion in the modes of a quantum field constrained in a finite region of space by some boundaries. The latter can be material as well as due to the geometrical properties of the background spacetime. In the latter case, Casimir effect becomes an exciting arena in which General Relativity (GR) and Quantum Field Theory (QFT) face each other.

As we will see, an observer comoving with a Casimir cavity, freely falling in a Schwarzschild black hole, measures a small reduction in the (absolute) value of the (negative) Casimir energy[4–12] as the black hole horizon is approached. At a first glance, this may seem rather puzzling, as one would expect no change with respect to the usual flat spacetime result $\langle \epsilon_{Cas} \rangle_{stat} = -\frac{\pi^2}{1440 L^4}$, due to the Equivalence Principle. Actually, the *local* measurements performed by the comoving observer, are related to a *non-local* renormalized field stress-energy tensor $T^{\rm ren}_{\mu\nu}$. The latter is determined (by means of some regularization technique) by the low-energy contribution of $T_{\mu\nu}$, thus probing the *global* structure of the surrounding spacetime geometry. This, in turn, allows for a local measurement to be sensitive to the cavity fall.

The plan of the paper is as follows. In section 2 Lemaître coordinates are introduced. Section 3 is devoted to Schwinger's proper-time approach and to the evaluation of the effective action $W$. Section 4 deals with the static Casimir effect,

while particle creation is discussed in section 5. The last section is devoted to some concluding remarks. Throughout the paper use has been made of natural geometrized units. Greek indices take values from 0 to 3; latin ones take values from 1 to 3. The metric signature is $-2$, with determinant $g$.

## 2. Lemaître Coordinates

The Schwarzschild metric for a black hole of mass $M$ in the standard Schwarzschild coordinates $\{t, r, \theta, \phi\}$ reads

$$ds^2 = \left(1 - \frac{r_g}{r}\right) dt^2 - \left(1 - \frac{r_g}{r}\right)^{-1} dr^2 - r^2 d\Omega^2, \tag{1}$$

where $r_g = 2M$ is the gravitational (Schwarzschild) radius of the black hole and $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$. In such coordinates, we meet a *coordinate* singularity at the horizon. Being interested in the behaviour of a Casimir cavity falling into a black hole, we need a chart which is regular at the horizon, so the form (1) of the metric is not suitable. Among the various coordinate systems well behaved at the horizon, we will adopt the Lemaître chart. In the Lemaître coordinates, a freely falling test body has a constant value of the radial $\rho$ coordinate.

Consider a Casimir cavity, freely falling from spatial infinity. Adjust the cavity clock so that the proper time $\tau = 0$ when the cavity is at the radial horizon coordinate $r_0 = r_g$. Then the constant value is $\rho_0 = \frac{2}{3} r_g$ and the Schwarzschild metric in the Lemaître coordinates $\{\tau, \rho, \theta, \phi\}$ reads[13,14]

$$ds^2 = d\tau^2 - \frac{r_g}{r(\tau)} d\rho^2 - r^2(\tau) d\Omega^2, \qquad r(\tau) = r_g \left(1 - \frac{3\tau}{2r_g}\right)^{2/3}. \tag{2}$$

Notice that the travel from the infinity to the horizon is described by *negative* values of the proper time: $-\infty < \tau \leq 0$.

## 3. The Casimir Cavity and the Comoving Frame

The measurement of Casimir energy inside the falling cavity is performed by a *comoving* observer. The cavity plates (of area $A$ and separated by a distance $L$, such that $L \ll \sqrt{A}$) are taken orthogonal to the radial falling direction.[a] Furthermore

- the cavity is taken to fall from spatial infinity with zero initial velocity and zero angular momentum;
- the typical cavity size is much smaller than the gravitational radius of the black hole, so that, in particular, $L \ll r_g$, with $L$ being the plate separation
- the cavity is *rigid*; its dimensions and shape do not suffer any distortion, in spite of external tidal forces. Such assumption holds true provided $L \ll r_g$. Differently stated, we neglect tidal effects[15] inside the falling cavity. Tidal effects will be extensively studied in a lower-dimensional model in Ref. 15.

[a] Such a choice has been made only for the sake of definiteness.

- the cavity follows a true geodesic motion; hence we neglect other non-gravitational external effects.

## 4. Proper-time Schwinger's approach

For the sake of simplicity we will consider a massless scalar field. We also assume the field to obey the Dirichlet boundary conditions at the plates. In the tetrad frame adapted to the Lemaître chart the Klein-Gordon equation reads[16]

$$\left[\Box + \frac{1}{4}\frac{\xi^2}{(1-\xi\tau)^2}\right]\varphi = 0, \qquad \xi = \frac{3}{2r_g}. \tag{3}$$

From (3), the proper-time Hamiltonian $\hat{H}$ reads $\hat{H} = \hat{H}_0 + \hat{V}$, where $\hat{H}_0 = \partial_\tau^2 - \vec{\nabla}^2 \equiv -\hat{p}_0^2 + \hat{\vec{p}}^2$. As usual, we write the effective action $W = \lim_{\nu\to 0} W(\nu)$, where[17,18]

$$W(\nu) = -\frac{i}{2}\int_0^\infty ds\, s^{\nu-1}\mathrm{Tr}\, e^{-is\hat{H}}, \tag{4}$$

and the limit $\nu \to 0$ has to be taken at the end of calculations. The trace has to be evaluated all over the continuous as well the discrete degrees of freedom, including those of spacetime. After a quite long algebra we get

$$W(\nu) = -\frac{iA}{32\pi^{5/2}}\int_0^\infty ds \int_{-\infty}^T d\tau \sum_n \frac{s^{\nu-3/2-1}}{\beta^{1/2}} e^{-is(n\pi/L)^2} \times\left[\pi^{3/2} e^{-i/(2\beta)} H_0^{(1)}\big(1/(2\beta)\big) + 2G_{23}^{31}\left(-\frac{i}{\beta}\,\middle|\,\begin{matrix}0 & 1/2\\ 0 & 0 & 0\end{matrix}\right)\right], \tag{5}$$

where $\beta = \frac{s\xi^2}{(1-\xi\tau)^2}$ is a small adimensional parameter which will be used for further power expansion. $H_0^{(1)}$ and $G_{23}^{31}$ are a zero-order Hankel and a Meijer G-function, respectively. The $H_0^{(1)}$-dependent (real) part of $W(\nu)$ is responsible for the vacuum polarization (namely, the static Casimir effect), while the $G$-dependent (imaginary) part of $W(\nu)$ is responsible of the vacuum persistence amplitude (i.e., particle creation).

## 5. The Static Casimir Effect

From the real part of (5) we obtain $\langle\epsilon_{Cas}\rangle = -\lim_{\nu\to 0}\frac{1}{AL}\frac{\partial}{\partial\tau}\Re\mathrm{e}\, W(\nu)$, hence

$$\langle\epsilon_{Cas}\rangle = -\frac{\pi^{3/2}}{16L^4}\sum_{k=0}^\infty \frac{2^k\xi^{2k}a_k}{(1-\xi\tau)^{2k}}\left(\frac{L}{\pi}\right)^{2k}\Gamma\left(-\frac{3}{2}+k\right)\zeta(-3+2k). \tag{6}$$

Taking the leading ($k=0$) and the next to leading order ($k=1$) term, we get

$$\langle\epsilon_{Cas}\rangle = -\frac{\pi^2}{1440L^4} + \frac{1}{384L^2}\frac{\xi^2}{(1-\xi\tau)^2} + O(\xi^4). \tag{7}$$

The first term is the (expected) usual flat Casimir energy density. At the horizon crossing ($\tau \to 0^-$), we have (recall that $\xi = 3/(2r_g)$)

$$\langle \epsilon_{Cas} \rangle_{hor} = -\frac{\pi^2}{1440 L^4}\left[1 - \frac{135}{(4\pi)^2}\left(\frac{L}{r_g}\right)^2\right]. \tag{8}$$

Eq. (7) tells us how the corrections to the Casimir energy density change with the proper time as the cavity approaches the black hole horizon.

## 6. Bunch-Davies Vacuum and Particle Creation

Particle creation is related to the vacuum persistence amplitude, i.e., the imaginary part of the effective action $W$. In the in-out formalism we have

$$|\langle 0\,\mathrm{out}|0\,\mathrm{in}\rangle|^2 = e^{2i\Im \mathrm{m} W}, \tag{9}$$

so that the (usually small) number density of created particles inside the falling cavity is $\langle n \rangle \simeq \frac{2\,\Im \mathrm{m}\, W}{AL}$. Expanding the imaginary pert of $W(\nu)$ we obtain a divergent result; namely

$$\Im \mathrm{m} W = \frac{A}{24\pi^3 L^2}\left[-\frac{\xi^2 L^2}{(1-\xi\tau)^2}\zeta(1) + \frac{2\xi^4 L^4}{15(1-\xi\tau)^4} + \cdots\right]. \tag{10}$$

Inspection of (10) reveals that the first term in the square brackets is *divergent*. The origin of such divergence is likely to be related to the implicitly assumed *infinite* extension of the plates, as we will see below. We will avoid the difficulties stemming from the appearance of infinities in the imaginary part of the effective action exploiting the relationship between the Schwinger theory and the in-out formalism, based upon the Bogolubov approach.

The field modes related to (3) are proportional to

$$\chi_k(\eta) = \frac{1}{2}\sqrt{\frac{\pi}{\xi}\eta} H_0^{(1)}\left(\frac{\omega_k}{\xi}\eta\right), \qquad \eta = 1 - \xi\tau, \tag{11}$$

which have the required minkowskian (plane wave) behaviour at $\eta \to \infty$ (i.e. $\tau \to -\infty$), when the cavity is at the spatial infinity with respect to the black hole. The above modes satisfy the Bunch-Davies vacuum requirements[19–21]. Also, in the far past the field modes admit a plane wave solution

$$\chi_k(\tau) = \frac{\alpha}{\sqrt{2\omega_k}} e^{-i\omega_k \tau} + \frac{\beta}{\sqrt{2\omega_k}} e^{i\omega_k \tau}. \tag{12}$$

Matching the above solutions at the black hole horizon ($\eta \geq 1$), we easily compute the Bogolubov coefficients $\alpha$ and $\beta$[b]

$$|\alpha_k|^2 = 1 + \frac{\xi^2}{16\omega_k^2(1-\xi\tau)^2}, \qquad |\beta_k|^2 = \frac{\xi^2}{16\omega_k^2(1-\xi\tau)^2}, \tag{13}$$

[b](An interesting approach, requiring no detailed knowledge of state normalization, based upon the paper by Hamilton et al. (A. Hamilton, D. Kabat and M. Parikh, JHEP 0407, 024 (2004)), may be used as well to obtain the same result.)

satisfying $|\alpha_k|^2 - |\beta_k|^2 = 1$. The $\beta$ coefficient is related to particle creation. Note that, as $\tau \to -\infty$, $|\alpha_k|^2 \sim 1$ and $|\beta_k|^2 \sim 0$, i.e., we have no particle creation in the far past, as expected, meanwhile at the horizon crossing ($\tau = 0$) we have $|\beta_k|^2 = \frac{\xi^2}{16\omega_k^2}$. Although the number of created quanta is a divergent quantity, we can get a *finite* result for the energy density $\langle \epsilon_{\rm dyn} \rangle$ of the created quanta, writing

$$\langle \epsilon_{\rm dyn} \rangle = \frac{1}{AL}\left[\frac{A}{(2\pi)^2}\sum_n \int d^2k_\perp \frac{\xi^2}{16\omega_k^2\eta^2}\omega_k\right] = \frac{\xi^2}{384L^2(1-\xi\tau)^2}. \qquad (14)$$

Comparing the above result with (7), describing the vacuum energy density pertaining to the Casimir effect, we see, quite interestingly, that the small reduction observed in the static Casimir energy value *just* corresponds to the amount of energy of created field particles. This could suggest a close relationship between the two considered effects. Nevertheless, some care is required when speculating about such coincidence, as both the results have been obtained as first-order approximations.

## 7. Concluding remarks

We have considered the Casimir energy density corrections in a small cavity freely falling from the spatial infinity into a Schwarzschild black hole.

At a first glance, one could wonder that corrections to the static Casimir effect as well as particle creation are detected by an observer in a freely falling *inertial* frame. However, as anticipated at the end of section V, this is not so surprising. The Equivalence Principle (EP), deeply rooted in the theory of General Relativity (GR), applies well in the context of a *local* theory, just as GR is. On the other hand, when quantum fields are taken into account, the *non-local* character of the underlying quantum theory conflicts with the EP, causing the latter to be not straightforwardly applicable.

In the present scenario, the quantum field stress-energy tensor $T_{\mu\nu}$ behaves as a *non-local* object, thus probing the global spacetime structure, through the long wavelength field modes. The adopted renormalization procedure (whatever it may be) does transfer the spacetime details into the renormalized $T^{\rm ren}_{\mu\nu}$, which is basically the locally *measured* object. In such a way, information contained in the spacetime geometry surrounding the cavity bypasses - so to say - the EP, appearing both in the form of a small correction to the expected static Casimir energy and a tiny flux of created field quanta.

In deriving the above results several assumptions have been made. In particular, we have neglected other possible contributions related to from the cavity extension. Tidal effects, e.g., are expected to give rise to anisotropies in the energy density distribution inside the cavity; such aspect will be considered in detail (in the case of a 1+1D model) in a forthcoming paper[15].

It maybe that some of the divergencies met throughout the paper are due to the (in)finiteness of the Casimir plates. In that respect, it seems likely that a more

physically consistent analysis, based upon a *finite* 3D cavity, could help to remove the residual infinities encountered in the present approach.

Finally, an obvious improvement of the present research would be to extend the analysis of Ref. 15 to the 3+1D case, also including both *tidal* and *3D-finite-size* effects in evaluating the corrections to the Casimir effect. This we hope will be our next goal.

## References


1. H. Casimir, *Proc. K. Ned. Akad. Wet.* **51** 793 (1948).
2. H. Casimir and D. Polder, *Phys.Rev.* **73** 360 (1948).
3. K. A. Milton, *The Casimir effect: Physical Manifestations of Zero-Point Energy* (World Scientific, Singapore, 2001).
4. E. Calloni, L. Di Fiore, G. Esposito, L. Milano and L Rosa, *Phys. Lett. A* **297**, 328 (2002).
5. E. Calloni, L. Di Fiore, G. Esposito, L. Milano and L Rosa, *Int. J. Mod. Phys. A* **17**, 804 (2002).
6. G. de A. Marquez and V.B. Bezerra, *Mod. Phys. Lett. A* **19**, 49 (2004).
7. G. de A. Marquez, S. G. Fernandez and V.B. Bezerra, *Brazil. J. Phys.* **35**, 1110 (2005).
8. F. Sorge, *Class. Quantum Grav.* **22**, 5109 (2005).
9. S. A. Fulling, K. A. Milton, P. Parashar, A. Romeo, K. V. Shajesh and J. Wagner, *Phys. Rev. D* **76**, 025004 (2007).
10. K. A. Milton, S. A. Fulling, P. Parashar, A. Romeo, K. V. Shajesh and J. Wagner, *J. Phys. A: Math. Theor.* **41**, 164052 (2008).
11. F. Sorge, *Class. Quantum Grav.* **26**, 235002 (2009).
12. F. Sorge, *Phys. Rev. D.* **90**, 084050 (2014).
13. G. Lemaître *Ann. Soc. Sci.* I **A53**, 51 (1933).
14. D. Kramer, H. Stephani, M. MacCallum and E. Herlt, *Exact Solutions of Einstein's Field Equations*, (VEB Deutscher Verlag der Wissenschaften, Berlin, 1980).
15. J. H. Wilson, F. Sorge and S. A. Fulling, to be submitted (2018).
16. Birrell N D and Davies P C W 1982 *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, 1982).
17. J. Schwinger, *Phys. Rev.* **82**, 664 (1951).
18. J. Schwinger, *Lett. Math. Phys.* **24**, 59 (1992).
19. L. Parker, *Phys. Rev.* **183**, 1057 (1969).
20. S. A. Fulling, *Phys. Rev. D* **7**, 2850 (1973).
21. P. C. W. Davies, *J. Phys. A: Gen. Phys.* **8**, 609 (1975).

# Supersymmetry and singularity in a dynamical M-brane background

K. Maeda

*Faculty of Engineering, Shibaura Institute of Technology,*
*Saitama 330-8570, Japan*
*E-mail: maeda302@sic.shibaura-it.ac.jp*

K. Uzawa

*Department of Physics, School of Science and Technology, Kwansei Gakuin University,*
*Sanda, Hyogo 669-1337, Japan*
*E-mail: kunihito.uzawa@gmail.com*

The supersymmetry arises in certain theories of fermions coupled to gauge fields and gravity in a spacetime of 11 dimensions. The dynamical brane background has mainly been studied for the class of purely bosonic solutions only, but recent developments involving a time-dependent brane solution have made it clear that one can get more information by asking what happens on supersymmetric systems. In this work, we present an exact supersymmetric solution of a dynamical M-brane background in the 11-dimensional supergravity and investigate supersymmetry breaking, the geometric features near the singularity.

*Keywords*: Dynamical brane background; Supersymmetry; 11-dimensional supergravity; Singularity.

## 1. Introduction

The dynamical $p$-brane solutions in a higher-dimensional gravity theory were studied by Refs. 1–9 and have been widely discussed ever since. However, some aspects of the physical properties, such as supersymmetry and its breaking in the context of string theory, have remained slightly unclear. The motivation for this work is to improve this situation. For this purpose, it is first necessary to construct supersymmetric brane solutions depending on the time as well as space coordinates.

In this work, we will find the supersymmetric dynamical M-brane as an exact solution of the supergravity field equations. What we will construct is a time-dependent M2-brane solution preserving supersymmetry in the 11-dimensional supergravity theory. The dynamical M2-brane background preserving supersymmetry is a kind of natural extension of the static M2-brane system, which can be described by an analogous Reissner-Nordström solution in the four-dimensional Einstein-Maxwell theory. We present an exact solution having a quarter of a full supersymmetry for a dynamical M-brane in an 11-dimensional supergravity and discuss how to break supersymmetries in Sec. 2. We describe some applications of the result, which are the behavior of the geodesic, the analysis of the geometrical structure in dynamical M2-brane background. In Sec. 3, we start our discussion of supersymmetric M2-brane solution by examining the basic features of the background geometry. By solving the radial null geodesic equations, we show that the naked strong curvature singularity appears. Section 4 contains some discussions and concluding remarks.

## 2. Supersymmetry in a dynamical M2-brane

In this section, we present the exact solution to the field equations of an 11-dimensional supergravity corresponding to a dynamical M2-brane configuration. The 11-dimensional gravitino (Killing spinor field) equation gives the time-dependent solution with the particular ansatz of fields. We find that the supersymmetric solution depends on the null coordinate along the M2-brane world volume, as well as the coordinates of the transverse space to the M2-brane. The 11-dimensional supersymmetric solution is given by

$$ds^2 = h^{-2/3}(x,r)\eta_{\mu\nu}dx^\mu dx^\nu + h^{1/3}(x,r)\left(dr^2 + r^2 u_{ab}dz^a dz^b\right)\,, \tag{1}$$

where the function $h(x,r)$ is given by

$$h(x,r) = c_\mu x^\mu + \bar{c} + \frac{M}{r^6}\,, \tag{2}$$

$\mu$, $\nu = 0$, 1, 2, and $i$, $j = 3, 4, \ldots, 10$, $\chi = \pm 1$, $c_\mu$, $\bar{c}$, and $M$ are constant parameters, $u_{ab}$ denotes the metric of the seven-sphere. It follows that if $c_\mu = \partial_\mu h_0$ is not null, there exists only a trivial solution to the Killing spinor equation, and the supersymmetry is completely broken. On the other hand, when $c_\mu$ is a null vector, the Killing spinor equation leads to

$$c_\mu \gamma^\mu \varepsilon = 0\,, \tag{3}$$

where $\gamma^\mu$ gives the SO(2, 1) $\gamma$-matrices satisfying

$$\{\gamma^\mu\,,\ \gamma^\nu\} = \eta^{\mu\nu}\,. \tag{4}$$

Then, one quarter of the possible rigid supersymmetries in the maximal case survives. Here, we check the degree of supersymmetry for the case of $M = 0$. The dynamical M2-brane to the plane wave background if we consider the following special case of vanishing M2-brane charge[10,11]. The integrability condition reduces to $c_\mu\,\gamma^\mu\varepsilon = 0$. Then, the dynamical M2-brane solution with $c_\mu \neq 0$, preserves a half of the maximal supersymmetries. Since the number of unbroken spacetime supersymmetries in the present background must be a half of the full supersymmetries,

Next we comment on the degree of the supersymmetry breaking for the dynamical M2-brane background. The measure of the supersymmetry breaking for the dynamical background is obtained from the consistency condition. The mass scale corresponds to $h^{-2}c_\mu c^\mu$, which could be identified with a kind of induced effective mass scale for the spinor field. The divergence at $h = 0$ means that the degree of the supersymmetry breaking increases as the background approaches the curvature singularity. On the other hand, the supersymmetry breaking becomes negligible near the M2-brane region $r \to 0$, as $h$ diverges there.

Let us consider the relation between the dynamics of the background and supersymmetry breaking in more detail. Introducing a new time coordinate $\tau$, which is defined by $\tau/\tau_0 = (c_0\,t)^{2/3}$, with constant $\tau_0 = (3/2c_0)$, we find the 11-dimensional

metric (1) as

$$ds^2 = \left[1+\left(\frac{\tau_0}{\tau}\right)^{3/2}\left(c_i x^i + \frac{M}{r^6}\right)\right]^{-2/3}\left[-d\tau^2 + \left(\frac{\tau_0}{\tau}\right)\delta_{ij}dx^i dx^j\right] + \left[1+\left(\frac{\tau_0}{\tau}\right)^{3/2}\left(c_i x^i + \frac{M}{r^6}\right)\right]^{1/3}\left(\frac{\tau}{\tau_0}\right)^{1/2}\left[dr^2 + r^2 d\Omega^2_{(7)}\right], \qquad (5)$$

where $x^i$ $(i = 1, 2)$ denotes the space coordinates of the world volume spacetime, the metric $\delta_{ij}$ is the spatial part of the three-dimensional Minkowski metric $\eta_{\mu\nu}$, and $d\Omega^2_{(7)}$ is the line element of the seven-sphere. When we set $c_1 = c_2 = 0$, the spacetime is an isotropic and homogeneous universe with respect to the world volume coordinates, whose supersymmetry is completely broken. On the other hand, the 11-dimensional spacetime becomes inhomogeneous and preserves supersymmetry if parameters $c_\mu$ satisfy $c_\mu c^\mu = 0$, and $c_\mu \gamma^\mu \varepsilon = 0$. Thus, in the limit when the terms $c_i x^i$ are negligible, which is realized in the limit $(\tau/\tau_0) \to \infty$, for small $r$, we find an 11-dimensional universe without supersymmetry. For concreteness, we discuss the dynamics in the region where the term $c_i x^i$ in the function $h(\tau, x, r)$ is smaller compared to the contribution of the M2-brane charge $M/r^6$. In the case of $(\tau/\tau_0) > 0$, we have found that the domains near the M2-brane has the supersymmetry. As the time increases, the background satisfies $(\tau/\tau_0)^{3/2} \gg c_i x^i$. Then, the contribution of the term $c_i x^i$ in the function $h(\tau, x, r)$ eventually becomes negligible in the 11-dimensional metric such that supersymmetries are completely broken, which is guaranteed by the region $c_i x^i \ll M/r^6$. Then, the dynamical M2-brane solution also behaves as a nonsupersymmetric cosmological solution in the asymptotic future.

## 3. Geometry of the supersymmetric dynamical M2-brane solution

As one may expect from the dynamical M2-brane solution, the spacetime with (1) has curvature singularity. For a fixed $x$, the spacetime asymptotically approaches the anisotropic solution at a large $r$, while the metric becomes approximately $AdS_4 \times S^7$ near the M2-brane region[10,11] (at $r \to 0$). Now we investigate the geometric feature near the curvature singularity. We consider the following time dependent M2-brane solution with the 11-dimensional metric

$$ds^2 = \left(cu + \frac{M}{r^6}\right)^{-2/3}\left(-2du\,dv + dy^2\right) + \left(cu + \frac{M}{r^6}\right)^{1/3}\left[dr^2 + r^2 d\Omega^2_{(7)}\right], \qquad (6)$$

$$u = \frac{1}{\sqrt{2}}(t - x), \qquad v = \frac{1}{\sqrt{2}}(t + x), \qquad (7)$$

where $c$, $M$ are constants. Since the function $h_1(r)$ dominates near $r \to 0$, the background geometry describes the extremal Reissner-Nordström solution with an infinite throat. The geometry of the dynamical M2-brane is not asymptotically flat while the extremal Reissner-Nordström solution gives the asymptotically Minkowski spacetime. Near the M2-brane, the metric becomes $AdS_4 \times S^7$. Since the square of

the four-form field strength diverges at the zeros of the function $h(u,r) = 0$, the curvature of the metric (7) can be singular at $h(u,r) = 0$.

Now we discuss the cosmological evolution of the spatial geometry in the region $h > 0$ and assume $c < 0$, in the function $h(u,r)$. For $u < 0$, the function $h$ is positive everywhere and the spatial surfaces are nonsingular unless we treat the negative charge of the M2-brane $M < 0$. They are asymptotically anisotropic spacetime for a fixed $x$ coordinate. The spatial metric is still regular for $u = 0$ besides the region $r \to \infty$. As time increases slightly, a singularity appears at $r = \infty$ and moves in from spatial infinity. As $u$ evolves further, the singularity eventually wraps the horizon completely. We start by solving radial null geodesic equations for the affine parameter $s$ on the background (7). We now discuss the radial null geodesic near the timelike singularity. For the supersymmetric M2-brane background (7), $h = 0$ hypersurface corresponds to a timelike curvature singularity[12] because $g^{MN}\ell_M\ell_N > 0$ for $\ell_M = \nabla_M h$ near the singularity. Let us then consider the past directed null geodesics which can hit the curvature singularity within a finite affine parameter length. The geodesic equation gives

$$r(s) \simeq r_0 + r_1\,(s_0 - s)^{6/5}\,, \qquad h(s) = -\left(\frac{5cf}{9}\right)(s_0 - s)^{9/5} - (s_0 - s)^{6/5}\,, \tag{8}$$

where $c$, $f$ are constants, $r_0$ and $s_0$ denote the value of $r$ and $s$ at singularity.

We now turn our attention to calculate a geometrical quantity in a parallelly propagated frame along the null geodesic,

$$\Gamma \equiv C_{MPNQ}E_2{}^M E_2{}^N k^P k^Q\,, \tag{9}$$

where $C_{MNPQ}$ is the Weyl tensor, $k^M$ denotes the tangent vector of null geodesic, and $E_2{}^M$ is a parallelly propagated spacelike unit vector orthogonal to $k^M$. These are defined by

$$k = (du/ds)\partial_u + (dv/ds)\partial_v + (dr/ds)\partial_r\,, \quad E_2 = h^{1/3}\partial_y\,. \tag{10}$$

In terms of the metric (7) and 11-dimensional null vectors, we find

$$\Gamma \sim (s_0 - s)^{-2}\,. \tag{11}$$

The shear $\sigma$ and the expansion rate $d\theta/ds$ of the congruence along the null vector $k^M$ diverge near the singularity as

$$\sigma \sim \int^s \Gamma ds \sim (s_0 - s)^{-1}\,, \qquad \left(\frac{d\theta}{ds}\right) \sim -\sigma^2 \sim -(s_0 - s)^{-2}\,. \tag{12}$$

Then, we obtain

$$\int^s \theta\, ds = \int^s \left(\frac{d\ln A}{ds}\right) ds \sim \ln(s_0 - s)\,, \tag{13}$$

where $A$ is the volume element of the null geodesic congruence. This implies that the timelike singularity is a strong type of curvature singularity, as the volume element of any congruence along the radial null geodesic vanishes there.

## 4. Discussions

In the present work, we have constructed the dynamical supersymmetric M2-brane solution for the warped compactification of an 11-dimensional supergravity. The solution is given by an extension of a static supersymmetric M2-branes solution. In the case of a dynamical M2-brane background, a quarter of maximal supersymmetries exists. If the M2-brane charge vanishes, our solution gives a plane wave background which preserves a half of the full supersymmetry. Therefore, in the far region from the M2-brane, the background changes from the dynamical M2-brane to time-dependent plane wave background. This means that one quarter of the maximal supersymmetry is enhanced to a half of the possible rigid supersymmetries in the maximal case when one moves in the transverse space to the M2-brane. Although we have mainly discussed the single M2-brane solution in this work, it is possible to generalize it to the solution which describes an arbitrary number of extremal M2-branes in an expanding universe. We have found that the degree of the supersymmetry breaking is strongly related to the dynamics of the background. Then, the time evolution of the geometry is deeply connected with the hierarchy and supersymmetry breaking while the inhomogeneity of the M2-brane world volume coordinates makes preserving the supersymmetry. In the region where the effect of the inhomogeneity of the M2-brane world volume coordinates is smaller compared to the contribution of the M2-brane charge, our supersymmetric solution describes the breaking of the supersymmetry, which is the transition from the supersymmetric universe to a nonsupersymmetric one as time evolves.

The dynamical M2-brane solutions can always take a form in the function $h(x,r) = h_0(x) + h_1(r)$, where the function $h(x,r)$ depends on the linear function of the M2-brane world volume coordinates $x^\mu$ as well as coordinates of the transverse space to the M2-brane. Since the existence of the function $h_0(x)$ implies the dynamical instability in the moduli of internal space[13], it would be useful to study the stability of a solution.

Motivated by the construction of a new supersymmetric solution, we have studied the global structure of the dynamical M2-brane background. We have found that the time dependence changes the causal structure of a static M2-brane solution. Since the volume element of any congruence along the radial null geodesic vanishes at the curvature singularity, it turns out that this is a strong version of a timelike singularity. We have studied null geodesics which terminate a coordinate singularity in terms of an analytic extension across there and showed that there is a regular null hypersurface (or horizon) generated by a null Killing vector field. In particular, this null Killing vector field describes the generator of the horizon even if the bulk metric is asymptotically anisotropic geometry at a constant $x$ coordinate. Hence, the near horizon geometry in this solution gives the regular spacetime, and thus becomes $\mathrm{AdS}_4 \times \mathrm{S}^7$.

It is important to explore another analytic solution describing a supersymmetric M-brane or D-brane in the expanding Universe. One may present whether

supersymmetric dynamical brane solutions affect the formation of the naked singularity. Upon setting an appropriate initial condition, these solutions may allow us to violate the cosmic censorship[12,14].

We can also discuss a dynamical black hole solution whose spacetime gives a time dependent universe. The near M2-brane region of this black hole in the expanding Universe is the same as the static solutions while the asymptotic structures are completely different, giving the anisotropic spacetime at a fixed $x$ coordinate with scale factors for a dynamical universe. The effective M2-brane charge for the supersymmetric background depends on the world volume coordinates of the M2-brane. The supersymmetric solutions can contain the function depending on null coordinates of the M2-brane world volume direction. The results we have obtained are not unnatural because studies of the supersymmetric plane wave background showed that it is possible to obtain time-dependent supersymmetric solutions with a nontrivial dependence on spacetime coordinates[15]. Although this may be a limitation on the applications of our solution, it is interesting to explore if similar more general dynamical and supersymmetric solutions can be obtained by relaxing or extending some of our assumptions for the 10-, 11-, or lower-dimensional backgrounds. We will study this subject in the near future.

## Acknowledgments

This work is supported by Grants-in-Aid from the Scientific Research Fund of the Japan Society for the Promotion of Science, under Contracts No. 17K05451 (K. M.) and No. 16K05364 (K. U.).

## References

1. G. W. Gibbons, H. Lu and C. N. Pope, *Phys. Rev. Lett.* **94**, 131602 (2005).
2. P. Binetruy, M. Sasaki and K. Uzawa, *Phys. Rev. D* **80**, 026001 (2009).
3. K. i. Maeda, N. Ohta and K. Uzawa, *JHEP* **0906**, 051 (2009).
4. M. Minamitsuji, N. Ohta and K. Uzawa, *Phys. Rev. D* **82**, 086002 (2010).
5. K. i. Maeda, *et al.*, Phys. Rev. D **82**, 046007 (2010).
6. M. Minamitsuji and K. Uzawa, *Phys. Rev. D* **83**, 086002 (2011).
7. M. Minamitsuji and K. Uzawa, *Phys. Rev. D* **84**, 126006 (2011).
8. M. Minamitsuji and K. Uzawa, *Phys. Rev. D* **87**, no.4, 046010 (2013).
9. K. Uzawa, *Phys. Rev. D* **90**, no.2, 025024 (2014).
10. K. Maeda and K. Uzawa, *Phys. Rev. D* **96**, no.8, 084053 (2017).
11. K. Maeda and K. Uzawa, *Galaxies* **6**, no.1, 11 (2018).
12. K. Maeda and K. Uzawa, *Phys. Rev. D* **93**, no.4, 044003 (2016).
13. H. Kodama and K. Uzawa, *JHEP* **0507**, 061 (2005).
14. J. H. Horne and G. T. Horowitz, *Phys. Rev. D* **48**, R5457 (1993).
15. M. Blau, *et al.*, *JHEP* **0201**, 047 (2002).

# MAGIC follow-up of Gamma-Ray Bursts at very high energies

Alessio Berti*, Alessandro Carosi, Pierre Colin, Satoshi Fukami, Susumu Inoue, Francesco Longo, Elena Moretti, Koji Noda and Michele Palatiello

for the MAGIC collaboration

*Department of Physics, University of Torino and INFN Torino*
*via Pietro Giuria, 10125 Italy*
**E-mail: Alessio.Berti@to.infn.it*

Gamma-Ray Bursts (GRBs) are probably one of the most enigmatic sources in current astrophysics, especially when their high-energy (HE, 100 MeV $\lesssim E \lesssim$ 100 GeV) emission is considered. Fermi-LAT observations unveiled many properties of GRBs emission in the GeV band, but its physical origin is not clearly understood yet. In this context, observations of GRBs at very-high-energies (VHE, $E \gtrsim$ 100 GeV) could provide a crucial contribution in the understanding of the physical processes driving GRB emission at the highest energies. For this purpose, ground facilities performing fast-follow up of GRBs above $\sim$ 100 GeV with high sensitivity are needed. In particular the MAGIC telescopes were designed to have a low energy threshold ($\sim$ 50 GeV) and to perform fast follow-up of GRBs thanks to their fast slewing movement. Up to date, 101 GRBs were followed-up by MAGIC in different observational conditions, providing a large data sample to search for VHE emission. This contribution will focus on the MAGIC GRBs follow-up campaign and highlight the observation of the short and nearby ($z = 0.16$) GRB 160821B, an interesting event showing a hint of VHE gamma-ray emission.

*Keywords*: Gamma-Ray Bursts; MAGIC; VHE emission.

## 1. Introduction

Gamma-Ray Bursts (GRBs) are transient events occurring at cosmological distances, releasing a huge amount of energy in electromagnetic radiation in a small time period.[1] After more than 50 years from their discovery, many aspects of GRBs are still puzzling. Among them, the origin of their high-energy (HE) emission and the possible presence of a very-high-energy (VHE) component are still debated topics in GRB physics. In the HE range, Fermi-LAT provides most of the information, from which it is seen that the GeV emission is delayed with respect to the lower energy emission and with a long lasting duration. Generally the HE spectra are well fitted using a power-law model, even if in some cases an additional spectral component is needed.[2] The latter could be linked with a process producing photons at VHE, pointing out to the importance of performing GRB observations in the energy range above $\sim$ 100 GeV. In this context, ground-based instruments like Imaging Atmospheric Cherenkov Telescopes (IACTs) can provide a good photon statistics in the VHE range. Among IACTs, MAGIC has some important figures of merit which are crucial for the follow-up of GRBs, namely a low energy threshold, a high sensitivity and a lightweight structure to allow fast repositioning.

In the following contribution we describe the MAGIC GRB follow-up program, focusing on the observational challenges of such a project and on the results obtained in the last years.

## 2. The MAGIC experiment

MAGIC is an IACT experiment with two identical telescopes of 17 m diameter.[3] It is located in one of the Canary Islands, La Palma (28.8° N, 17.8° W), at an altitude of 2200 m a.s.l. The MAGIC telescopes were designed to observe astrophysical sources of gamma-rays in stereoscopic mode above an energy threshold of $\sim 50$ GeV at trigger level and at zenith. The sensitivity which can be achieved by MAGIC is as low as 0.7% of the Crab Nebula flux above 220 GeV in 50 h.[4] For short timescales, which are important in the case of the follow-up of transient sources like GRBs, the flux sensitivity is at the level of 10% of the Crab Nebula flux above 100 GeV in 1 hour.

### 2.1. *GRBs follow-up with MAGIC*

The detection of Gamma-Ray Bursts in the VHE band is one of the primary scientific goals of MAGIC and each year a considerable amount of hours are devoted to their follow-up. The transient nature of GRBs makes their follow-up challenging for facilities like MAGIC, characterized by a limited field of view. For this reason, MAGIC was designed to be able to repoint in short time[a] in response to GRB alerts coming from the GCN (*Gamma-ray Coordinates Network*[b]). These alerts are processed by the MAGIC Automatic Alert System, which checks the visibility of the targets from the MAGIC site. If a GRB is observable, an automatic procedure repoints the telescopes in fast mode towards its sky position, while the different subsystems are prepared for data taking. This procedure was revisited in 2013 in order to be more robust against possible hardware failures. Beside the automatic follow-up, MAGIC performs so called *late-time observations* of GRBs detected by *Fermi*-LAT: the source is observed even after 24 h or more with respect to the trigger time of the event. The goal of such strategy is to detect a possible long-lasting VHE emission from GRBs.

In 15 years of operation, up to June 2018, MAGIC followed-up 101 GRBs in good observational conditions and not affected by any major hardware problem. See Fig. 1 for a skymap of such events. Despite the huge effort, MAGIC could not detect any VHE gamma-ray signal from GRBs (status up to June 2018). In one case though, described below, MAGIC data shows a hint of signal in the VHE band.

### 2.2. *GRB 160821B*

*Swift*-BAT detected the short GRB 160821B ($T_{90} = 0.48$ s in the 15-350 keV energy range) at 22:29:13 UT[6] and MAGIC received the alert at 22:29:26 UT. Using

[a] 25 s for a 180° rotation.
[b] `https://gcn.gsfc.nasa.gov/`.

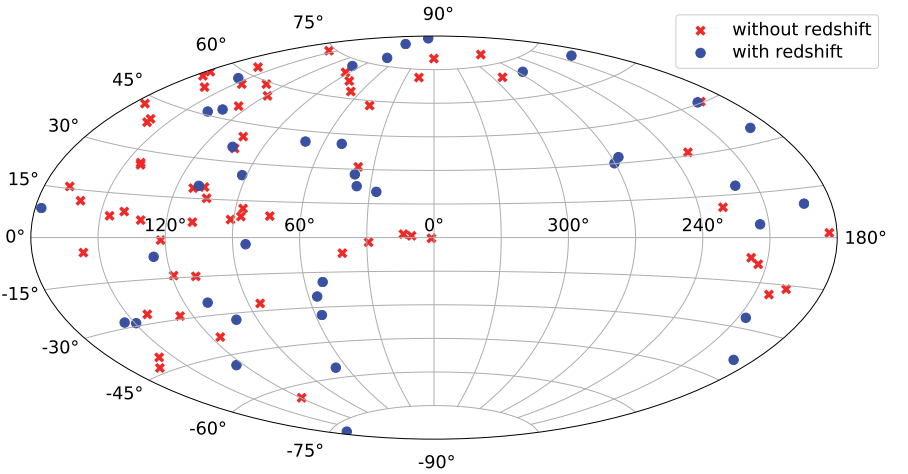



Fig. 1. Sky locations of the 101 GRBs followed-up by MAGIC up to June 2018. Filled blue dots are GRBs with redshift estimation, while red crosses are GRBs without redshift.

the automatic GRB procedure, MAGIC was on target just 24 s after the onset of the GRB. Thanks to optical observations by the William Herschel Telescope, the redshift of the GRB was estimated to be $z = 0.16$.[7]

MAGIC observation started under moderate moonlight conditions at a zenith of ∼34 degrees, with a night sky background (NSB) level between three and five times the one under dark moonless conditions. The first 1.5 hours of data were affected by a non optimal atmospheric transmission, due to the presence of clouds. For the rest of the observation the weather conditions were good and the NSB increased at a level between five and nine times with respect to dark conditions. The observation stopped after 4 hours from the trigger time, with the GRB reaching a zenith of ∼55 degrees.

A dedicated analysis was performed on GRB 160821B, for two reasons in particular: the presence of the Moon during observations and the low atmospheric transmission in the first part of data taking. The former issue is handled as as described in Ref. 5. In the latter case, the data are corrected for the low atmospheric transmission using the data collected by the LIDAR instrument installed on top of the MAGIC counting house building.[8] The dedicated analysis resulted in a hint of gamma-ray signal detection with a significance of $3.1\sigma$, as shown in Fig. 2.

If we make the assumption that this signal is real, this GRB would be the first to be detected at VHE energies and it would open a new window in the understanding of (short) GRBs emission. A detailed modeling and interpretation of such event will be the main topic of an upcoming paper, currently in preparation.

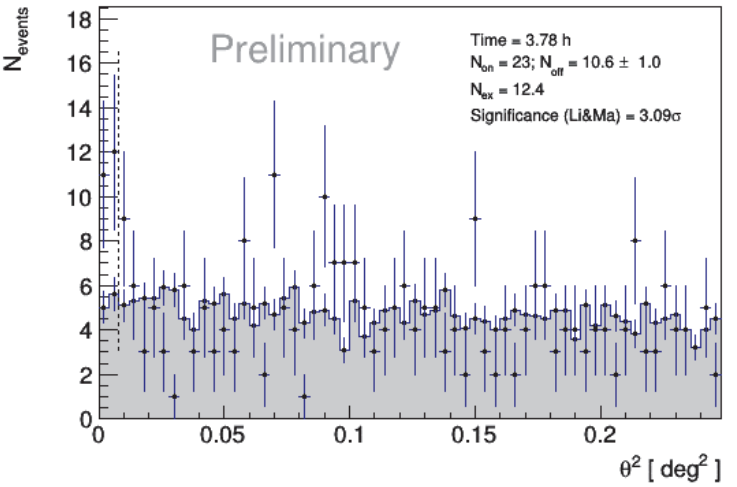



Fig. 2. $\theta^2$ distribution for data taken on GRB 160821B (points) and for the background events (grey shaded area). $\theta^2$ is defined as the squared angular distance between the nominal position of the source and the reconstructed arrival direction of the events.

## 3. Future prospects for GRB follow-up with MAGIC

MAGIC is currently one of the main actors in the follow-up of GRBs in the VHE band. The automatic procedure ensures to observe GRBs as fast as possible, without causing hardware failures. The efforts in pushing the capabilities of the telescopes under moonlight or non optimal weather conditions provide MAGIC with a higher chance to detect a GRB at very high energies, as indicated by the case of GRB 160821B. Even in this case, there is room for improvement. A change in the automatic procedure is foreseen in the nearby future, namely the inclusion of the Sum Trigger.[9] This novel trigger system, used successfully when MAGIC operated as a single telescope, is now in an advanced and stable state, providing an energy threshold as low as 30 GeV for low zenith angles. This feature is crucial for the observation of distant sources like GRBs.

## References

1. Kumar, P. and Zhang, B., *Physics Reports* **561**, 1-109 (2015).
2. Ackermann, M. et al., *ApJS* **209**, 11 (2013).
3. Aleksic, J. et al., *Astroparticle Physics* **72**, 61 (2016).
4. Aleksic, J. et al., *Astroparticle Physics* **72**, 76 (2016).
5. Ahnen, M. L. et al., *Astroparticle Physics* **94**, 29 (2017).
6. Siegel, M. H. et al., *GRB Coordinates Network* **19833** (2016).
7. Levan, A. J. et al., *GRB Coordinates Network* **19846** (2016).
8. Fruck, C. and Gaug, M., *Proc. of AtmoHEAD 2014* **02003** (2015).
9. Garcia, J. R. et al., *arXiv e-prints*, arXiv:1404.4219 (2014).

# Astrophysical tests of Lorentz invariance: Towards multi-gamma-ray bursts analyses

Michele Ronco*, Giacomo D'Amico

*Dipartimento di Fisica, Università di Roma "La Sapienza",*
*P.le A. Moro 2, 00185 Roma, Italy*
*INFN, Sez. Roma1,*
*P.le A. Moro 2, 00185 Roma, Italy*
*E-mail: michele.ronco@roma1.infn.it*

Giovanni Amelino-Camelia

*Dipartimento di Scienze Fisiche, Universitá "Federico II" di Napoli,*
*Complesso Universitario di Monte S. Angelo Edificio 6 via Cinthia, 45 – 80126 Napoli, Italy*
*INFN, Sez. Napoli,*
*Complesso Universitario di Monte S. Angelo Edificio 6 via Cinthia, 45 – 80126 Napoli, Italy*

Lorentz symmetries represent one of the cornerstones of modern physics, and yet independent approaches aiming at combining general relativistic with quantum effects often imply some form of departures from them. According to the simplest models, particles having different energies emitted at the same time from a given source should be detected at different times from a far-away detector, thereby producing a phenomenon of in-vacuo dispersion with a linear correlation between the time of observation and particles' energy. Given that, the search for energy-dependent time lags in gamma-ray bursts (GRB) has gradually become a standard way to make tests of fundamental physics and also look for the first signatures of the sought-after quantum theory of gravity. Most of the current studies, considering a single GRB or just the most energetic photon for each GRB analysed, allowed to set very tight constraints on the relevant scale, usually believed to be close to the Planck mass. However, due to the rather poor understanding of the spectral evolution of GRBs, statistical analyses over collections of GRBs would provide more reliable outcomes. Here we test in-vacuo dispersion by analysing all the photons with energy at the emission greater than 5 GeV emitted from 7 GRBs observed by Fermi-LAT. Remarkably, we find preliminary evidence of in-vacuo-dispersion-like spectral lags consistently with what has been noticed by some recent studies which, though, had focused only on the energy range above 40 GeV.

*Keywords*: Gamma-ray bursts.

## 1. Introduction

Over the last 15 years there has been considerable interest (see *e.g.* Refs. 1–9 and references therein) in quantum-gravity (QG) induced in-vacuo dispersion, the possibility that spacetime itself might behave essentially like a dispersive medium for particle propagation: there might be an energy dependence of the travel times of ultrarelativistic particles from a given source to a given detector.

The most studied[1–9] modelization of quantum-gravity-induced in-vacuo dispersion is

$$\Delta t = \eta_X \frac{E}{M_P} D(z) \pm \delta_X \frac{E}{M_P} D(z)\,, \tag{1}$$

where $z$ is the redshift of the relevant GRB and

$$D(z) = \int_0^z d\zeta \frac{(1+\zeta)}{H_0\sqrt{\Omega_\Lambda + (1+\zeta)^3\Omega_m}}\,. \tag{2}$$

$\Omega_\Lambda$, $H_0$ and $\Omega_m$ denote, as usual, respectively the cosmological constant, the Hubble parameter and the matter fraction[16]. $M_P$ denotes the Planck scale ($\simeq 1.2 \cdot 10^{28} eV$) and the values of the parameters $\eta_X$ and $\delta_X$ in (1) are to be determined experimentally. Here "$\pm\delta_X$" accounts for quantum-uncertainty (fuzziness) effects, while $\eta_X$ characterizes systematic effects. Finally, the label $X$ intends to allow for a possible dependence[1,9] on the type of particles and/or on their spin/helicity. We shall not consider either fuzziness or particle/spin dependent effects.

Eq. (1) tells us that, if we wish to test the in-vacuo-dispersion hypothesis, then we need far away transient sources emitting very-high energy particles. Given that, gamma-ray bursts (GRBs) perhaps represent the most suitable sources[1–4]. Many studies have been able to set very tight constraints on the QG scale (i.e. $E_{QG} \equiv M_P/\eta_X$ if $\delta_X = 0$) close to or, in some analyses, even beyond the Planck scale (see, e.g., Ref. 17 and references therein). However, the main challenge for this type of analyses consists in the difficulty to disentangle the QG effect from the intrinsic spectral lags. In absence of a satisfactory astrophysical mechanism to take into account source effects, there are two natural ways to face this problem: use multiple kind of sources and/or messengers, and increase the size of the data.

Some of us were involved in the first studies using IceCube data for searching for GRB-neutrino in-vacuo-dispersion candidates[8,10–12]. Analogous investigations were performed in a series of studies[13–15] focusing on the highest-energy GRB photons observed by the Fermi telescope. As summarized in Fig. 1 these studies provided rather strong statistical evidence of in-vacuo-dispersion-like spectral lags. For each point in Fig. 1 (black points are "GRB-neutrino candidates" [10], while the blue points are GRB photons with energy at emission greater than 40 GeV) we denote by $\Delta t$ the difference between the time of observation of the relevant particle and the time of observation of the first low-energy peak in the GRB, while $E^*$ is the redshift-rescaled energy of the relevant particle defined as $E^* \equiv (E \times D(z))/D(1)$. The linear correlation between $\Delta t$ and $E^*$ visible in Fig. 1 is just of the type expected for quantum-gravity-induced in-vacuo dispersion, and it has been estimated[10] that such a high level of correlation would occur accidentally (in absence of in-vacuo dispersion) only in less than 1% of cases, while GRB photons could produce such high correlation only in less than 0.1% of cases[12].

This "statistical evidence" motivated us to explore whether or not the in-vacuo-dispersion-like spectral lags persist at lower energies. Thus, we here contribute to the attempt to reduce the impact of intrinsic delays by extending the window of the statistical analysis down to 5 GeV. This increases the number of photons analyzed by more than an order of magnitude (only 11 photons are considered in Fig. 1, whereas the analysis we here report involves 148 photons). Indeed, given the poor

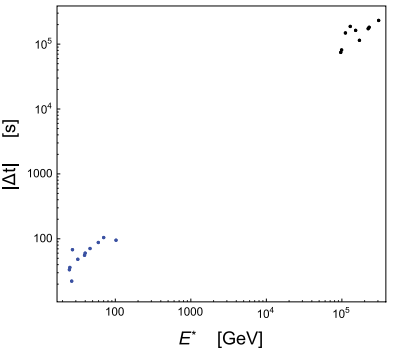



Fig. 1. Values of $|\Delta t|$ versus $E^*$ for the IceCube GRB-neutrino candidates discussed in Refs. 10, 12 (black points) and for the GRB photons discussed in Refs. 12, 15 (blue points). The photon points in figure also factor in the result of a one-parameter fit estimating the average magnitude of intrinsic time lags (details in Refs. 12, 15).

understanding of the GRB time spectra, analyses based on the time of observation of a single photon[18] may not uncover a feature, which though could be revealed by statistical analyses.

## 2. Data analysis: Widening the energy window

Our analysis focuses on the same GRBs whose photons took part in the analyses which led to the picture here summarized in Fig. 1, i.e. GRB080916C, GRB090510, GRB090902B, GRB090926A, GRB100414A, GRB130427A, GRB160509A, but includes all the photons with energy at the source greater than 5 GeV. Since we cannot assume all the GRB photons were emitted in coincidence with the first GRB peak as in Fig. 1, we consider a $\Delta t_{pair}$, which gives for each pair of photons in our sample their difference of time of observation. Thus, each pair of photons (from the same GRB) gives us an an estimated value of $\eta_\gamma$

$$\eta_\gamma^{[pair]} \equiv \frac{M_P \Delta t_{pair}}{D(1) E^*_{pair}}\,, \tag{3}$$

where $E^*_{pair}$ is the difference in values of $E^*$ for the two photons in the pair. Of course the $\Delta t_{pair}$ for many pairs of photons in our sample could not possibly have anything to do with in-vacuo dispersion: if the two photons were produced from different phases of the GRB (different peaks) their $\Delta t_{pair}$ will be dominated by the intrinsic time-of-emission difference. Those values of $\eta_\gamma^{[pair]}$ will be spurious, they will be "noise" for our analysis. However we also of course expect that some pairs of photons in our sample were emitted nearly simultaneously, and for those pairs the $\Delta t_{pair}$ could truly estimate $\eta_\gamma$. Form Fig. 1 one gets $\eta_\gamma = 30 \pm 6$, then we would expect that values of $\eta_\gamma^{[pair]}$ of about 30 are more frequent than expected without a relationship between arrival times and energy of the type produced by in-vacuo dispersion.

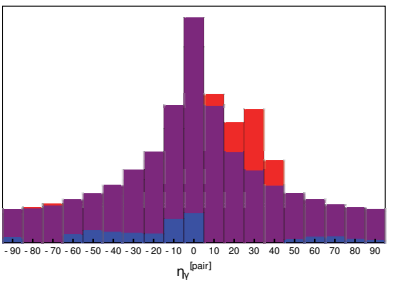



Fig. 2. Normalized distribution of $\eta_\gamma^{[pair]}$ for all pairs of photons (from the same GRB) within our data set. For bins where the observed population is higher than expected we color the bar in purple up to the level expected, showing then the excess in red. For bins where the observed population is lower than expected the bar height gives the expected population, while the blue portion of the bar quantifies the amount by which the observed population is lower than expected.

This is just what we find, as shown perhaps most vividly by the content of Fig. 2. The main point to be noticed in Fig. 2 is that we find in our sample a frequency of occurrence of values of $\eta_\gamma^{[pair]}$ between 25 and 35 which is tangibly higher than one would have expected in absence of a correlation between $\Delta t_{pair}$ and $E^*_{pair}$. Following a standard strategy of analysis (see, *e.g.*, Ref. 17) we estimate how frequently $25 \leq \eta_\gamma^{[pair]} \leq 35$ should occur in absence of correlation between $\Delta t_{pair}$ and $E^*_{pair}$ by producing $10^5$ sets of simulated data, each obtained by reshuffling randomly the times of observation of the photons in our sample. We also performed some variants of our analysis, first by dividing our data sample in three different energy ranges and considering only those pairs made of photons belonging to different groups (or excluding the photons with energy at the emission greater then 40 GeV, i.e. the only ones contributing to Fig. 1). Following a different procedure, we also estimated $\eta_\gamma$ with a best-fit technique performed for every triplet of photons from the same GRB. Remarkably, the excess of results for $\eta_\gamma^{[pair]}$ between 25 and 35 shows up in all these analyses with an overall significance of about 0.5%.

## 3. Discussion and Outlook

In summary we found rather striking indications in favor of values of $\eta_\gamma$ of about 30 in GRB data for all photons with energy at emission greater than 5 GeV. On the basis of our exploration, on future similar-size GRB data samples one should find again at least some partial manifestation of the same feature. We are of course much further from establishing whether this feature truly is connected with QG-induced in-vacuo dispersion, rather than being some intrinsic property of GRB signals. The imprint of in-vacuo dispersion is coded in the $D(z)$ for the distance dependence and, while that does give a good match to the data, one should keep in mind that only a few redshifts (a few GRBs) were relevant for our analysis. Moreover, given the very tight constraints on systematic QG delays[18], in-vacuo dispersion should most likely be of statistical ("fuzzy") nature.

## References


1. G. Amelino-Camelia, Living Rev. Rel. **16**, 5 (2013).
2. U. Jacob and T. Piran, arXiv:hep-ph/0607145, Nature Phys. **3**, 87 (2007).
3. G. Amelino-Camelia and L. Smolin, arXiv:0906.3731, Phys.Rev. D **80**, 084017 (2009).
4. G. Amelino-Camelia, J. Ellis, N.E. Mavromatos, D.V. Nanopoulos and S. Sarkar, arXiv:astro-ph/9712103, Nature **393**, 763 (1998).
5. R. Gambini and J. Pullin, Phys. Rev. **D59**, 124021 (1999).
6. J. Alfaro, H.A. Morales-Tecotl and L.F. Urrutia, arXiv:gr-qc/9909079, Phys. Rev. Lett. **84**, 2318 (2000).
7. G. Amelino-Camelia and S. Majid, arXiv:hep-th/9907110, Int. J. Mod. Phys. A **15**, 4301 (2000).
8. G. Amelino-Camelia, D. Guetta and T. Piran, Astrophys. J. **806** no.2, 269 (2015).
9. F. W. Stecker, S. T. Scully, S. Liberati and D. Mattingly, arXiv:1411.5889, Phys. Rev. D **91**, 045009 (2015).
10. G. Amelino-Camelia, L. Barcaroli, G. D'Amico, N. Loret and G. Rosati, arXiv:1605.00496, Phys. Lett. B **761**, 318 (2016).
11. G. Amelino-Camelia, L. Barcaroli, G. D'Amico, N. Loret and G. Rosati, arXiv:1609.03982 [gr-qc].
12. G. Amelino-Camelia, G. D'Amico, G. Rosati and N. Loret, arXiv:1612.02765, Nat.Astron. 1 (2017) 0139.
13. S. Zhang and B. Q. Ma, Astropart. Phys. **61** (2014) 108.
14. H. Xu and B. Q. Ma, Astropart. Phys. **82** (2016) 72.
15. H. Xu and B. Q. Ma, Phys. Lett. B **760** (2016) 602.
16. Planck Collaboration: P. A. R. Ade *et al.*, arXiv:1502.01589v2.
17. V. Vasileiou *et al.*, Phys. Rev. D **87** (2013) 122001.
18. A.A. Abdo *et al.* [Fermi LAT/GBM Collaborations], Nature **462**, (2009) 331.

# Constraints on the nuclear equation of state and r-process nucleosynthesis from numerical calculations of the chirp from binary neutron-star merger GW170817

G. J. Mathews, I.-S. Suh and N. Q. Lan

*Center for Astrophysics, Department of Physics, University of Notre Dame*
*Notre Dame, IN 46556, USA*
**E-mail: gmathews@nd.edu*
*https://physics.nd.edu/people/faculty/grant-j-mathews*

The first detection of gravitational waves a binary neutron star merger GW170817 by the LIGO-Virgo Collaboration has provided fundamental new insights into the astrophysical site for r-process nucleosynthesis and on the nature of dense neutron-star matter. The detected gravitational wave signal depends upon the tidal distortion of the neutron stars as they approach merger. We examine how the detected chirp depends the adopted equation of state. This places new constrains on the properties of nuclear matter. The detected evidence of heavy-element nucleosynthesis also provides insight into the nature of the r-process and the fission properties of the heaviest nuclei. Parametrically, one can divide models for the r-process into three scenarios roughly characterized by the number of neutron captures per seed nucleus (n/s). In addition to neutron-star mergers, these include magneto-hydrodynamic jets from supernovae and the neutrino heated wind above the proto neutron star in core-collapse supernovae. Insight from GW170817 allows one to better quantify the relative contributions of each astrophysical site and the fission termination of the r-process recycling.

*Keywords*: Binary neutron stars - Gamma-ray bursts: Kilonova - Gravitational Waves - Nucleosynthesis - Stars: neutron.

## 1. Introduction

On August 17, 2017, Nature revealed herself in a most spectacular way, both from the gravitational waves detected by the LIGO and VIRGO collaborations [1], and by the multitude of follow up observations [2] of the GRB170817A kilonova, along with IR and optical ground-based observations. In this work we consider two aspects of what may have been learned from this event. On the one hand, the GW chirp implies a possible dilemma regarding the Nuclear EOS deduced from the Post-Newtonian tidal-polarizability vs. observed neutron star properties. Here we discuss our analysis of the chirp in numerical general relativity [3] as a means to better constrain the EoS and clarify the current dilemma. The second aspect considered here is that the kilonova EM spectrum indicates evidence of r-process nucleosynthesis that has implications for the fission barriers and termination of fission recycling during the r-process.

In the LIGO analysis [1] and in follow up analysis [4,5] the tidal polarizability was deduced from post-Newtonian expansion. The tidal polarizability (or deformability) $\Lambda$ is an intrinsic neutron-star property highly sensitive to the compactness parameter $[\Lambda \sim (M/R)^5]$ that describes the tendency of a neutron star to develop a mass quadrupole as a response to the tidal field induced by its companion. [6,7] Based upon the most recent analysis by LIGO [5] it has been deduced that the reduced tidal

polarizability is $\tilde{\Lambda} = 190^{+390}_{120}$ implying that the radius of the stars of 1.4 $M_\odot$ is in the range 10.5 km $\leq R \leq$ 13.3 km.

However, in this work we consider the possibility that as the stars approach merger during the chirp they are neither well described by post-Newtonian physics, not as a simple quadrupole deformation. Indeed, one desires to trace the evolution of the binary from when they first enter the LIGO window until some 30 sec later when the stars merge. For an orbit period of $\sim 100$ ms, this would require evolving the system for $\sim 3000$ orbits. This is indeed an daunting task from the standpoint of numerical relativity for which it is difficult to follow more than a few orbits due to limitations of computational resource.

In this work, however, we report on the first analysis of a general relativistic hydrodynamic simulation [3,8] that is capable of stably numerically integrating thousands of orbits. The solution of the field equations and hydrodynamic equations of motion is summarized in Refs. 9, 10. One starts with the slicing of spacetime into the usual one-parameter family of hypersurfaces separated by differential displacements in a time-like coordinate as defined in the (3+1) ADM formalism [11,12].

In Cartesian $x, y, z$ isotropic coordinates, proper distance is expressed as

$$ds^2 = -(\alpha^2 - \beta_i\beta^i)dt^2 + 2\beta_i dx^i dt + \phi^4 \delta_{ij} dx^i dx^j, \tag{1}$$

where the lapse function $\alpha$ describes the differential lapse of proper time between two hypersurfaces. The quantity $\beta_i$ is the shift vector denoting the shift in space-like coordinates between hypersurfaces. The curvature of the metric of the 3-geometry is described by a position dependent conformal factor $\phi^4$ times a flat-space Kronecker delta ($\gamma_{ij} = \phi^4 \delta_{ij}$). This conformally flat condition on the metric provides a numerically valid initial solution to the Einstein equations. The vanishing of the Weyl tensor for a stationary system in three spatial dimensions guarantees that a conformally flat solution to the Einstein equations exists.

To solve for the fluid motion of the system in curved spacetime it is convenient to use an Eulerian fluid description [9]. By introducing the usual set of Lorentz contracted state variables it is possible to write the relativistic hydrodynamic equations in a form which is reminiscent of their Newtonian counterparts [9]. The hydrodynamic state variables are: the coordinate baryon mass density, $D$, the internal energy density $E$ and the covariant spatial momentum density, $S_i$;

$$D = W\rho; \quad E = W\rho\epsilon; \quad S_i = (D + E + PW)U_i. \tag{2}$$

Key to maintaining is to evolve the the spatial three velocity in the rotating frame,

$$V^i = \alpha \frac{U_i}{\phi^4 W} - \beta^i; \tag{3}$$

In terms of these state variables, the hydrodynamic equations in the CFA are as follows: The equation for the conservation of baryon number takes the form,

$$\frac{\partial D}{\partial t} = -6D\frac{\partial \log\phi}{\partial t} - \frac{1}{\phi^6}\frac{\partial}{\partial x^j}(\phi^6 D V^j). \tag{4}$$

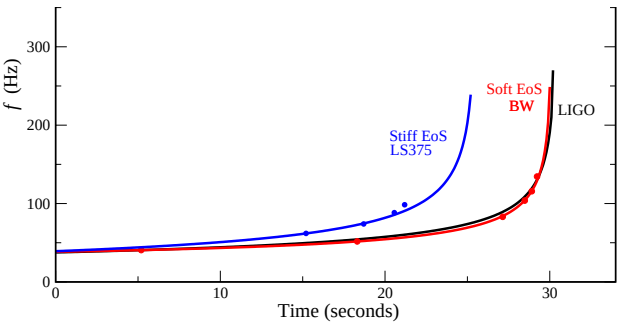



Fig. 1. Reconstructed LIGO gravitational wave chirp compared with the chirp calculated here based upon two equations of state for neutron stars.

The equation for internal energy evolution becomes,

$$\frac{\partial E}{\partial t} = -6(E+PW)\frac{\partial \log\phi}{\partial t} - \frac{1}{\phi^6}\frac{\partial}{\partial x^j}(\phi^6 E V^j) - P\left[\frac{\partial W}{\partial t} + \frac{1}{\phi^6}\frac{\partial}{\partial x^j}(\phi^6 W V^j)\right]. \tag{5}$$

Momentum conservation takes the form,

$$\frac{\partial S_i}{\partial t} = -6S_i\frac{\partial \log\phi}{\partial t} - \frac{1}{\phi^6}\frac{\partial}{\partial x^j}(\phi^6 S_i V^j) - \alpha\frac{\partial P}{\partial x^i} + 2\alpha(D+E+PW)\left(W - \frac{1}{W}\right)\frac{\partial \log\phi}{\partial x^i} + S_j\frac{\partial \beta^j}{\partial x^i} - W(D+E+PW)\frac{\partial \alpha}{\partial x^i} - \alpha W(D+\Gamma E)\frac{\partial \chi}{\partial x^i}. \tag{6}$$

where the last term in Eq. (6) is the contribution from the radiation reaction potential $\chi$ as defined in Refs. 3, 9, 8. Including this term allows for a calculation of the orbital evolution via gravitational wave emission in the CFA.

As a first test and calibration we have constructed quasi-stable orbits for which this term is set to zero. We then reconstruct the time sequence of orbits by fitting a modified form of the chirp,

$$\frac{96}{5}\pi^{8/3}\mathcal{M}^{5/3}t + \frac{3}{8}f^{5/3(at^2+bt+c)} + C = 0 \tag{7}$$

where $a$, $b$ and $c$ are deduced from fits to the numerical simulation and correspond to corrections due to the inclusion of higher moments in mass-energy and momentum in the chirp. This correction is motivated by the deviation of a logrithmic $f$ vs. $t$ plot from a straight line in numerical simulations.

An example of a reconstructed chirp is shown in figure 1 based upon orbits described in Refs. 3, 8. These show a stronger sensitivity to the nuclear EoS than that deduced from the PN analysis of tidal polarizability. [1,4]

## 2. Constraints on r-process nucleosynthesis and the termination of fission recycling

Observation of the multi-messenger followup[2] of the GRB170817A make a strong case for the existence newly synthesized $r$-process material in the ejecta from the binary neutron star merger. The light curve[16] and spectrum[17] are consistent with heating and opacity from $\sim 0.02 - 0.04$ $M_\odot$ of heavy-elements produced during $r$-process nucleosynthesis.

Ejecta from such neutron-star mergers can be characterized by extremely neutron neutron rich nucleosynthesis for which the r-process proceeds until beta- or neutron-induced fission occurs along the $r$-process path. Figure 2 from Ref. 13 highlights a key dilemma in nuclear physics. If the $r$-process path ends among light isotopes with $A \approx 285$ asymmetric fission occurs leading to a prominent peak around $A = 130$. On the other hand if the $r$-process proceeds until $A \approx 300$ as indicated by the fission barriers and yields from Refs. 14, 15, then a more unimodal

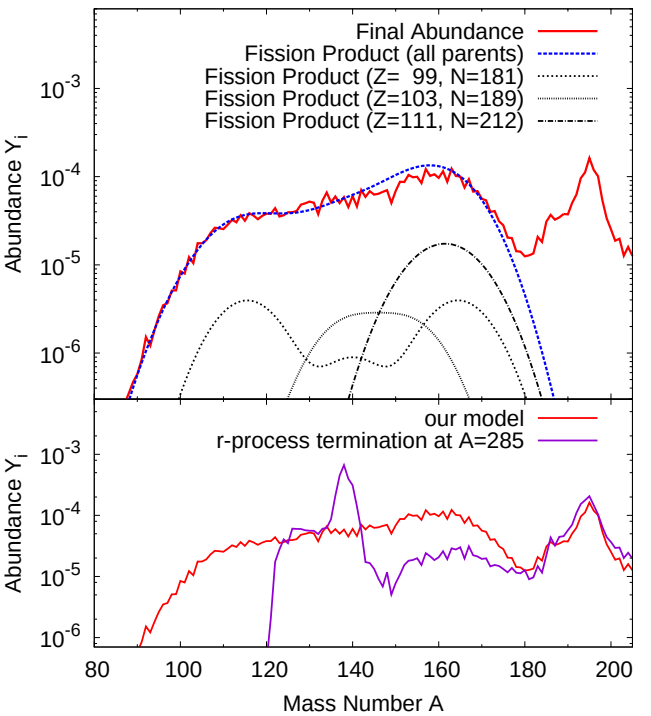



Fig. 2. Illustration from Ref. 13 of the impact of fission yields and fission recycling on the final r-process abundances. Upper panel shows the relative contributions for 3 representative nuclei compared with the final abundance distribution. The lower panel shows the same final r-process yields compared with the distribution that would result if fission recycling were only to occur from parent nuclei at the termination of the r-process path at $A = 285$.

fission distribution leads to the formation of the rare-earth peak, but little yield around the A=130 peak. Hence, identification of elements near the A=130 peak in the ejecta from the GRB170817A kilonova would answer a very important question regarding the fission barriers and yields of the heaviest elements formed in nature. Indeed, the tentative identification [16] of prominent absorption features due to neutral Te-I and Xe-I in the spectrum of the kilonova appears to have answered this question. However, this is controversial since it requires cool material and minimal line broadening, neither of which are expected in this environment. Obviously, confirmation of this detection is necessary in future observations.

## Acknowledgment

This work was supported by the U.S. Department of Energy under grant DE-FG02-95-ER40934.

## References

1. B. Abbott et al. (LIGO Scientific and Virgo Collaboration), Phys. Rev. Lett. 119, 161101 (2017).
2. B. P. Abbott et al. (Virgo, Fermi-GBM, INTEGRAL, and LIGO Scientific Collaboration), Astrophys. J. 848, L12 (2017).
3. I.-S. Suh, G. J. Mathews, J. R. Haywood, and N. Q. Lan, Adv. in Astr., 2017, 612703, (2017) arXiv:1601.01460.
4. F. J. Fattoyev, J. Piekarewicz, and C. J. Horowitz, Phys. Rev Lett. 120, 172702 (2018).
5. B. Abbott et al. (LIGO Scientific and Virgo Collaboration), Phys. Rev. Lett. 121, 161101 (2018).
6. T. Damour, M. Soffel, and C. Xu, Phys. Rev. D 45, 1017 (1992).
7. E. E. Flanagan and T. Hinderer, Phys. Rev. D 77, 021502 (2008).
8. N. Q. Lan., I.-S. Suh, G. J. Mathews, and J. R. Haywood, Comm. in Phys., 25, 299, (2015).
9. Wilson J R and Mathews G J *Relativistic Numerical Hydrodynamics*, (Cambridge University Press, Cambridge, United Kingdom) (2003).
10. Mathews G J and Wilson J R, *Astrophys. J.* **482** 929 (1997).
11. Arnowitt R, Deser S, and Misner C W *Gen. Rel. Grav.* **40** 1997 (2008).
12. York J W Jr *Sources of Gravitational Radiation* ed Smarr L L (Cambridge University Press, Cambridge, UK) p. 83 (1979).
13. S. Shibagaki, T.Kajino, G. J. Mathews, S. Chiba, S. Nishimura, and G. Lorusso, Astrophys. J., 816, 79 (2016).
14. M. Ohta et al., Proc. Int. Conf. on NDST, Nice, France, (2007).
15. S. Chiba et al., AIP Conf. Proc. 1016, 162 (2008).
16. Smartt et al. Nat., 551, 75 (2017).
17. Chornock, R. et al. ApJL, 848, L18 (2017).

# Gravitational waves from pulsars due to their magnetic ellipticity

José C. N. de Araujo*

*Divisão de Astrofísica, Instituto Nacional de Pesquisas Espaciais, S. J. Campos, SP 1227-010, Brazil*
** E-mail: jcarlos.dearaujo@inpe.br*

Jaziel G. Coelho**

*Departamento de Física, Universidade Tecnológica Federal do Paraná Medianeira, PR 85884-000, Brazil*
*** E-mail: jazielcoelho@utfpr.edu.br*

Samantha M. Ladislau† and César A. Costa‡

*Divisão de Astrofísica, Instituto Nacional de Pesquisas Espaciais, S. J. Campos, SP 1227-010, Brazil*
*† E-mail: samantha.ladislau@inpe.br*
*‡ E-mail: cesar.costa@inpe.br*

We discuss some aspects of de Araujo, Coelho and Costa[1,2] concerning the role of a time dependent magnetic ellipticity on the pulsars' braking indices and on the putative gravitational waves (GWs) these objects can emit. Since only nine of more than 2000 known pulsars have accurately measured braking indices, it is of interest to extend this study to all known pulsars, in particular as regards GW generation. In summary, our results show a pessimistic prospect for the detection of GWs generated by these pulsars, even for advanced detectors such as aLIGO and AdVirgo, and the planned Einstein Telescope, would not be able to detect these pulsar, if the ellipticity has magnetic origin.

*Keywords*: Pulsars; Gravitational Waves; braking index.

## 1. Ellipticity of Magnetic Origin and Gravitational Waves from Pulsars

If the magnetic field and (or) the angle between the axes of rotation and the magnetic dipole of the pulsars are independent of time, the combination of magnetic dipole and gravitational wave (GW) brakes could only explain braking index (n) in the interval $3 < n < 5$. The observations, however, show that only PSR J1640-4631 has braking index in this interval, as can be seen in Table 1. In particular, we consider this issue in the context of magnetic ellipticity[1]. It is worth stressing that the magnetic field and the angle between the axes of rotation and the magnetic dipole of the pulsars are dependent on time.

Recall that the equatorial ellipticity is given by

$$\epsilon = \frac{I_{xx} - I_{yy}}{I_{zz}}, \tag{1}$$

where $I_{xx}$, $I_{yy}$, $I_{zz}$ are the moment of inertia with respect to the rotation axis, $z$, and along directions perpendicular to it.

The pulsar is deformed by its own dipole magnetic field. Such deformation associated with the fact that the axes of rotation and of the magnetic dipole are

misaligned generates an ellipticity given by (see, e.g., Bonazzola and Gourgoulhon[3]; Konno et al.[4]; de Freitas Pacheco and Regimbau[5]):

$$\epsilon_B = \kappa \frac{B_0^2 R^4}{GM^2} \sin^2 \phi, \tag{2}$$

where $B_0$ is the dipole magnetic field, $R$ and $M$ are the radius and the mass of the star respectively, $\phi$ is the angle between the rotation and magnetic dipole axes, whereas $\kappa$ is the distortion parameter, which depends on both the star equation of state (EoS) and the magnetic field configuration[5]. We consider that $\kappa = 10 - 1000$, as suggested by numerical simulations[3,5].

Table 1. The periods ($P$) and their first derivatives ($\dot{P}$) for pulsars with known braking indices ($n$) (see also ATNF catalog[6,7]).

| Pulsar | $P$ (s) | $\dot{P}$ ($10^{-13}$ s/s) | $n^{\diamond}$ |
|---|---|---|---|
| PSR J1734-3333 | 1.17 | 22.8 | $0.9 \pm 0.2$[8] |
| PSR B0833-45 (Vela) | 0.089 | 1.25 | $1.4 \pm 0.2$[9] |
| PSR J1833-1034 | 0.062 | 2.02 | $1.8569 \pm 0.0006$[10] |
| PSR J0540-6919 | 0.050 | 4.79 | $2.140 \pm 0.009$[11] |
| PSR J1846-0258 | 0.324 | 71 | $2.19 \pm 0.03$[12] |
| PSR B0531+21 (Crab) | 0.033 | 4.21 | $2.51 \pm 0.01$[13] |
| PSR J1119-6127 | 0.408 | 40.2 | $2.684 \pm 0.002$[14] |
| PSR J1513-5908 | 0.151 | 15.3 | $2.839 \pm 0.001$[11] |
| PSR J1640-4631 | 0.207 | 9.72 | $3.15 \pm 0.03$[15] |

*Note*: $^{\diamond}$ $n \equiv f_{\rm rot}\, \ddot{f}_{\rm rot}/\dot{f}_{\rm rot}^2$, where $f_{\rm rot} = 1/P$ is the rotating frequency, $\dot{f}_{\rm rot}$ and $\ddot{f}_{\rm rot}$ are their time derivatives.

Recall that the power emitted by a rotating magnetic dipole is given by[16]

$$\dot{E}_{\rm d} = -\frac{16\pi^4}{3} \frac{B_0^2 R^6 \sin^2 \phi}{c^3} f_{\rm rot}^4, \tag{3}$$

and the power loss via GW emission reads[17]

$$\dot{E}_{\rm GW} = -\frac{2048\pi^6}{5} \frac{G}{c^5} I^2 \epsilon^2 f_{\rm rot}^6. \tag{4}$$

Also, the total energy of the pulsar is provided by its rotational energy, $E_{\rm rot} = 2\pi^2 I f_{\rm rot}^2$, and any change on it is given by $\dot{E}_{\rm d}$ and $\dot{E}_{\rm GW}$, namely

$$\dot{E}_{\rm rot} \equiv \dot{E}_{\rm GW} + \dot{E}_{\rm d}. \tag{5}$$

Now, from the definition of the braking index (see, e.g., the note in Table 1), one can easily obtain that[a]

$$n = 3 + 2\eta - 2\frac{P}{\dot{P}}(1+\eta) \left[ \frac{\dot{B}_0}{B_0} + \dot{\phi} \cot \phi \right], \tag{6}$$

[a]The detailed derivation of Eq. (6) can be found in de Araujo, Coelho and Costa[1].

where $\eta$ is defined in such a way that $\dot{E}_{\rm GW} = \eta \dot{E}_{\rm rot}$, which is interpreted as the efficiency of GW generation. In de Araujo, Coelho and Costa[1] it is also shown that with Eq. (6) one can explain, in principle, the braking indices of the pulsars of Table 1.

Recall that the GW amplitude generated by a pulsar reads

$$h^2 = \frac{5}{2}\frac{G}{c^3}\frac{I}{r^2}\frac{\mid \dot{f}_{\rm rot} \mid}{f_{\rm rot}}. \tag{7}$$

This equation considers that the spindown is due to gravitational waves only, i.e., $n = 5$ (spindown limit - SD).

From the definition of $\eta$ one obtains that $\dot{\bar{f}}_{\rm rot} = \eta \dot{f}_{\rm rot}$, i.e., the part of the spindown related to the GW emission brake. Thus, one can obtain an equation for the GW amplitude that holds for n $< 5$, namely

$$\bar{h}^2 = \frac{5}{2}\frac{G}{c^3}\frac{I}{r^2}\frac{\mid \dot{\bar{f}}_{rot} \mid}{f_{rot}} = \frac{5}{2}\frac{G}{c^3}\frac{I}{r^2}\frac{\mid \dot{f}_{\rm rot} \mid}{f_{\rm rot}}\,\eta. \tag{8}$$

Recall that the GW amplitude also reads

$$h = \frac{16\pi^2 G}{c^4}\frac{I\epsilon f_{\rm rot}^2}{r}, \tag{9}$$

(see, e.g, Shapiro and Teukolsky[17]). Combining both equations for the GW amplitude one obtains

$$\epsilon = \sqrt{\frac{5}{512\pi^4}\frac{c^5}{G}\frac{\dot{P}P^3}{I}\eta}. \tag{10}$$

Now, for a purely magnetic brake we have

$$\bar{B}_0 \sin^2\phi = \frac{3Ic^3}{4\pi^2 R^6} P\dot{P}, \tag{11}$$

where $\bar{B}_0$ would be the magnetic field whether the break were purely magnetic. If there is also a GW brake contribution we have that $B_0 < \bar{B}_0$. Combining the definition of $\eta$ and Eq. (10) one obtains after some algebraic manipulation the following equation for the efficiency $\eta$

$$\eta = 1 - \left(\frac{B_0}{\bar{B}_0}\right)^2, \tag{12}$$

which is obviously lower than one, as it should be. Substituting this last equation into Eq. (2) we obtain

$$\epsilon = \frac{3Ic^3}{4\pi^2 GM^2R^2} P\dot{P}\,(1-\eta)\,\kappa. \tag{13}$$

Finally, substituting this last equation into Eq. (10), we obtain

$$\eta = \frac{288}{5}\frac{I^3 c}{GM^4R^4}\frac{\dot{P}}{P}\,(1-\eta)^2\,\kappa^2. \tag{14}$$

Notice that with Eqs. (13) and (14) we obtain $\epsilon$ and $\eta$ in terms of $M$, $R$, $I$, $P$ and $\dot{P}$ for a given value of $\kappa$. Since in practice $\eta \ll 1$, the following useful equations are obtained

$$\epsilon \simeq \frac{3Ic^3}{4\pi^2 GM^2R^2} P\dot{P}\kappa \tag{15}$$

and

$$\eta \simeq \frac{288}{5}\frac{I^3c}{GM^4R^4}\frac{\dot{P}}{P}\kappa^2. \tag{16}$$

We now calculate $\epsilon_B$ and $\eta$ for the pulsars of Table 1. We then adopt fiducial values for M, R and I. We adopt $\kappa = 10$ and 1000, which have the same orders of magnitude of the values considered by, e.g., Regimbau and de Freitas Pacheco[5].

In Table 2 we present the result of these calculations. Even for the extremely optimistic case, the value of the ellipticity is at best $\epsilon_B \sim 10^{-5}$ (for PSR J1846-0258) and the corresponding efficiency $\eta \sim 10^{-8}$. Therefore, the amplitude of the GW in this case would be four orders of magnitude lower than the spindown limit ($\eta = 1$). Thus, even advanced detectors such as aLIGO and AdVirgo, and the planned Einstein Telescope, would not be able to detect these pulsars.

Table 2. $\epsilon$ and $\eta$ for $\kappa = 10$ (1000) for the Pulsars of Table 1.

| Pulsar | $\epsilon$ | $\eta$ |
|---|---|---|
| PSR J1734-3333 | $1.2 \times 10^{-7(-5)}$ | $1.1 \times 10^{-13(-9)}$ |
| PSR B0833-45 (Vela) | $4.9 \times 10^{-10(-8)}$ | $8.3 \times 10^{-14(-10)}$ |
| PSR J1833-1034 | $5.5 \times 10^{-10(-8)}$ | $1.9 \times 10^{-13(-9)}$ |
| PSR J0540-6919 | $1.1 \times 10^{-9(-7)}$ | $5.7 \times 10^{-13(-9)}$ |
| PSR J1846-0258 | $1.0 \times 10^{-7(-5)}$ | $1.3 \times 10^{-12(-8)}$ |
| PSR B0531+21 (Crab) | $6.1 \times 10^{-10(-8)}$ | $7.5 \times 10^{-13(-9)}$ |
| PSR J1119-6127 | $7.2 \times 10^{-8(-6)}$ | $5.8 \times 10^{-13(-9)}$ |
| PSR J1513-5908 | $1.0 \times 10^{-8(-6)}$ | $6.0 \times 10^{-13(-9)}$ |
| PSR J1640-4631 | $8.9 \times 10^{-9(-7)}$ | $2.8 \times 10^{-13(-9)}$ |

Notice that Eqs. (15) and (16) do not depend on the braking index n. Consequently, we can calculate such quantities for the pulsars of the ATNF Pulsar Catalog. We refer the reader to the paper by de Araujo, Coelho and Costa[2] for details. In Fig. 1 we show an interesting histogram with the data of the ATNF Catalog, namely, the number of pulsars for $\log \epsilon_B$ bin. Note the high number of pulsars concentrated around $\sim 10^{-10}$ ($10^{-8}$) for $k = 10$ (1000). The values of $\eta$ are also extremely small, a histogram can be found in de Araujo, Coelho and Costa[2], where can be seen a peak at $10^{-16} - 10^{-15}$.

These extremely small values of $\epsilon_B$ and $\eta$ imply that the GW amplitudes are at best seven orders of magnitude smaller than those obtained by assuming the spindown limit (SD), being therefore hardly detected (see Fig. 2).

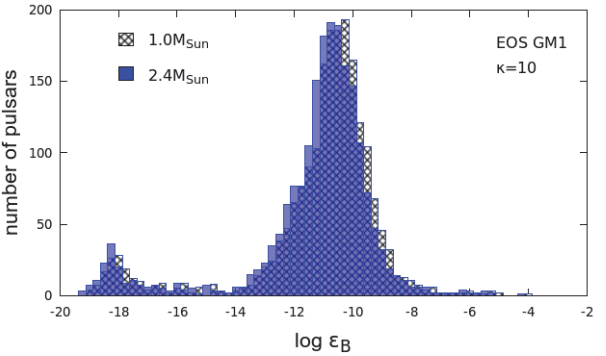



Fig. 1. Ellipticity histogram for the pulsars of ATNF Catalog for $\kappa = 10$.

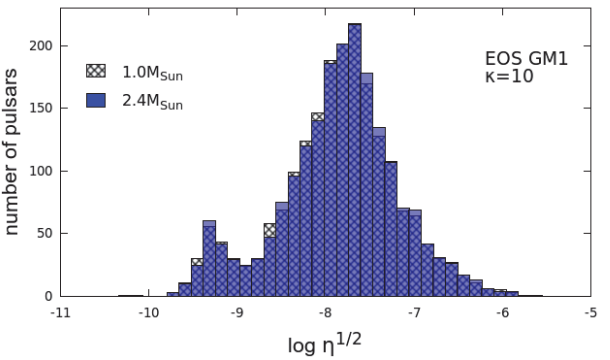



Fig. 2. Histogram of $\eta^{1/2} = h/h^{SD}$ (spin-down ratio) for the pulsars of ATNF Catalog for $\kappa = 10$.

## 2. Final Remarks

We present an expression for the braking index considering that the ellipticity is of magnetic dipole origin and time dependent. In this context, we model the braking indices of the 9 pulsars that have such measured quantities accurately. Then we calculate the amplitudes of the GWs generated by these 9 pulsars. Summing up, we conclude that these amplitudes are too small to be detected. For example, the pulsar PSR J1846-0258 would need to be observed for over 1000 years to be detected by the Einstein Telescope.

Since the equations for $\eta$, $\epsilon_B$ and $h$ are independent of n, we extend our study for most of the pulsars of the "ATNF Pulsar Catalog". Regarding detectability, the prospects remain pessimistic, since the ellipticity generated by the magnetic dipole is extremely small, the corresponding amplitude of GWs is much smaller than the amplitude obtained via the spindown limit.

**Acknowledgments**

J.C.N.A thanks FAPESP (2013/26258-4) and CNPq (307217/2016-7) for partial support. J.G.C. is likewise grateful to the support of CNPq (421265/2018-3 and 305369/2018-0). S.M.L. and C.A.C. acknowledge CAPES for financial support.

**References**

1. J.C.N.de Araujo, J.G. Coelho and C.A. Costa, *ApJ* **831** 35 (2016).
2. J.C.N.de Araujo, J.G. Coelho and C.A. Costa, *EPJC* **77**, 350 (2017).
3. S. Bonazzola and E. Gourgoulhon, *A&A* **312**, 675 (1996).
4. K. Konno, T. Obata and Y. Kojima, *A&A* **356**, 234 (2000).
5. T. Regimbau and J.A. de Freitas Pacheco, *A&A* **447**, 1 (2006).
6. CSIRO. ATNF Pulsar Catalogue. `http://www.atnf.csiro.au/people/pulsar/psrcat/` (2003).
7. R.N. Manchester, G.B. Hobbs, A. Teoh and M. Hobbs, *The Astronomical Journal* **129**, 1993 (2005).
8. C.M. Espinoza, A.G. Lyne, M. Kramer, R.N. Manchester and V.M. Kaspi, *ApJL* **741**, L13 (2011).
9. A.G. Lyne, R.S. Pritchard, F. Graham-Smith and F. Camilo, *Nature* **381**, 497 (1996).
10. J. Roy, Y. Gupta and W. Lewandowski, *MNRAS* **424**, 2213 (2012).
11. M.A. Livingstone, V.M. Kaspi, F.P. Gavriil, R.N. Manchester, E.V.G. Gotthelf and L. Kuiper, *Astrophys Space Sci* **308**, 317 (2007).
12. R.F. Archibald, V.M. Kaspi, A.P. Beardmore, N. Gehrels and J.A. Kennea, *ApJ* **810**, 67 (2015).
13. A.G. Lyne, R.S. Pritchard and F. Graham-Smith, *MNRAS* **265**, 1003 (1993).
14. P. Weltevrede, S. Johnston and C.M. Espinoza, *MNRAS* **411**, 1917 (2011).
15. R.F. Archibald, E.V. Gotthelf, R.D. Ferdman, V.M. Kaspi, S. Guillot, F.A. Harrison, E.F. Keane, M.J. Pivovaroff, D. Stern, S.P. Tendulkar and J.A. Tomsick, *ApJl* **819**, L16 (2016).
16. T. Padmanabhan, in *Theoretical Astrophysics - Volume 2, Stars and Stellar Systems* (Cambridge University Press, 2001).
17. S.L. Shapiro and S.A. Teukolsky, in *Black Holes, White Dwarfs, and Neutron Stars: The Physics of Compact Objects* (New York, Wiley-Interscience, 1983).

# A strategy for detecting non-gaussianity of stochastic gravitational waves

Makoto Tsuneto[1], Asuka Ito[2], Toshifumi Noumi[2,3] and Jiro Soda[2]

[1] *SUNY Stony Brook Department of Physics and Astronomy, Stony Brook 11794-3800, USA*
[2] *Department of Physics, Kobe University, Kobe 657-8501, Japan*
[3] *Department of Physics, University of Wisconsin-Madison, Madison, WI 53706, USA*

We study how to probe non-gaussianity of stochastic gravitational waves with pulsar timing arrays. The non-gaussianity is a key to probe the origin of stochastic gravitational waves. In particular, the shape of the bispectrum which characterize the non-gaussianity carries valuable information of inflation models. We show that an appropriate filter function for three point correlations enables us to extract a specific configuration of momentum triangles in bispectra.

## 1. Introduction

Stochastic gravitational waves (GWs), the GW analog for the cosmic microwave background, are going through us from all directions. They contain primordial GWs[1,2] produced during inflation[3–6] in addition to GWs of cosmological/astrophysical origin[7]. As the name suggests, stochastic GWs are characterized by statistics such as the power spectrum, bispectrum and higher order correlation functions.

The detection of the power spectrum of stochastic GWs is a clue to probe the early universe. Moreover, the bispectrum of stochastic GWs, which represents the non-gaussianity, is a powerful tool to discriminate astrophysical and primordial origin since the former has a gaussian distribution as long as event rates are high enough to create continuous GWs[7,8].[a] Therefore, the bispectrum of stochastic GWs enables us to probe the early universe. Indeed, the bispectrum of primordial GWs contains the detail of inflation models like nonlinear interactions of the graviton. The shape of bispectra, which depends on inflation models[15–18], allows us to discriminate inflation models.[b] Furthermore, the imprint of new particles with the mass comparable to the Hubble scale during inflation can potentially appear in the squeezed limit of momentum triangles[20–26]. Therefore, the bispectrum is a powerful probe of the early universe and beyond the standard model.

Now, GW detectors are in operation to probe stochastic GWs, although no signal of stochastic GWs has been detected yet. The sensitive frequency band of interferometers like LIGO[27] and Virgo[28] is around $10^2$ Hz, while pulsar timing

[a] Alternatively, if the event rate is too low to produce continuous GWs, the distribution is not gaussian. Detectability of such feature is explored in Refs. 9–14.
[b] Recently, it was shown that the detection of power spectrum of primordial GWs is not enough to exclude bouncing universe models[19]. Therefore, the bispectrum is also important to distinguish inflation and bouncing universe models.

arrays such as EPTA[29] and NANOGrav[30] are searching for stochastic GWs with a frequency range $10^{-9}$-$10^{-7}$ Hz. The constraints on the energy density of stochastic GWs are $\Omega_{GW} < 1.2 \times 10^{-9}$ (EPTA), $\Omega_{GW} < 3.4 \times 10^{-10}$ (NANOGrav), and $\Omega_{GW} < 1.1 \times 10^{-11}$ (PPTA), respectively. In future, the space interferometers, LISA[31] and DECIGO[32], will be launched in a few decades. The pulsar timing array project SKA[33] will start in 2020 and significantly improve the current sensitivity. Its possible upper limit is $\Omega_{GW} < 1.0 \times 10^{-13}$.[34] Therefore, it is worth exploring a new theoretical research area for forthcoming observations.

In this paper, we investigate a method for detecting the bispectrum of stochastic GWs with pulsar timing arrays[35]. In particular, we explain how to utilize a filter function not only to maximize the signal to noise ratio (SNR), but also to extract a specific configuration of momentum triangles in the bispectrum.

## 2. GW signal in pulsar timing arrays

In the Minkowski spacetime, GWs as tensor perturbations of the metric can be expanded with plane waves:

$$h_{ij}(t,\vec{x}) = \sum_A \int_{-\infty}^{\infty} \mathrm{d}f \int \mathrm{d}\hat{\Omega}\, e^{2\pi \mathrm{i}\left(ft - |f|\hat{\Omega}\cdot\vec{x}\right)}\, \tilde{h}_A(f,\hat{\Omega}) e^A_{ij}(\hat{\Omega})\,, \tag{1}$$

where $\hat{\Omega}$ is the direction of propagation of GWs. Polarization tensors, which satisfy $e^A_{ij}(\hat{\Omega}) e^{A'}_{ij}(\hat{\Omega}) = 2\delta^{AA'}$, can be defined by

$$e^+_{ij}(\hat{\Omega}) = \hat{m}_i\hat{m}_j - \hat{n}_i\hat{n}_j\ , \quad e^\times_{ij}(\hat{\Omega}) = \hat{m}_i\hat{n}_j + \hat{n}_i\hat{m}_j\ . \tag{2}$$

Here, $\hat{m}$ and $\hat{n}$ are unit vectors perpendicular to $\hat{\Omega}$ and one another.

GW detectors have their specific response to GWs. For instance, pulsars can be utilized as a detector. A pulsar is a neutron star which emits periodic electromagnetic fields very accurately. If gravitational waves $h_A(f,\hat{\Omega})$ exist continuously between the Earth and a pulsar (the direction $\hat{p}$), we observe the redshift of an emitted pulse as[36]

$$\tilde{Z}(f,\hat{\Omega}) = \left(e^{-2\pi \mathrm{i} L\left(f + |f|\hat{\Omega}\cdot\hat{p}\right)} - 1\right) \sum_A \tilde{h}_A(f,\hat{\Omega}) F^A(\hat{\Omega},\hat{p})\ , \tag{3}$$

where

$$F^A(\hat{\Omega},\hat{p}) \equiv e^A_{ij}\,\frac{1}{2}\,\frac{\hat{p}^i\hat{p}^j}{1+\hat{\Omega}\cdot\hat{p}} \tag{4}$$

is the pattern function. It represents a geometrical factor, namely, dependence of the sensitivity on the configuration of the detector and GWs. Furthermore, we integrate Eq. (3)

$$\tilde{z}(f) = \int \mathrm{d}\hat{\Omega}\, \tilde{Z}(f,\hat{\Omega})\ , \tag{5}$$

because stochastic GWs propagate toward all directions. The quantity that is actually measured is the residual defined by

$$R(t) \equiv \int_0^t \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}f\, e^{2\pi \mathrm{i} f t'}\, \tilde{z}(f) \tag{6}$$

$$= \int_0^t \mathrm{d}t' \int_{-\infty}^{\infty} \mathrm{d}f \int \mathrm{d}\hat{\Omega}\, e^{2\pi \mathrm{i} f t'} \left( e^{-2\pi \mathrm{i} L\left(f + |f| \hat{\Omega}\cdot\hat{p}\right)} - 1 \right) \sum_A \tilde{h}_A(f, \hat{\Omega}) F^A(\hat{\Omega}, \hat{p}). \tag{7}$$

Since, for pulsar timing measurement, the minimum frequency is about $0.1\,\mathrm{yr}^{-1}$ and the shortest distance between the Earth and a pulsar is $\sim 100\,\mathrm{ly}$, we have $fL \gtrsim 10$. In this range, the exponential term in the parenthesis of Eq. (7) can be approximated to zero because it oscillates rapidly. Hence, Eq. (5) can be approximated as

$$\tilde{z}(f) \simeq -\sum_A \int \mathrm{d}\hat{\Omega} \tilde{h}_A(f, \hat{\Omega}) F^A(\hat{\Omega}, \hat{p})\ . \tag{8}$$

We find that the correlation of residuals is directly related with that of stochastic GWs. Therefore, observing appropriate correlation function of the signals, we can probe the statistic of stochastic GWs.

## 3. Probing bispectrum with pulsar timing arrays

Let us define the bispectrum as

$$\langle \tilde{h}_A(f_1, \hat{\Omega}_1) \tilde{h}_{A'}(f_2, \hat{\Omega}_2) \tilde{h}_{A''}(f_3, \hat{\Omega}_3) \rangle = B_{AA'A''}(|f_1|, |f_2|, |f_3|) \delta(f_1 + f_2 + f_3) \times \delta^{(3)}(|f_1|\hat{\Omega}_1 + |f_2|\hat{\Omega}_2 + |f_3|\hat{\Omega}_3)\ . \tag{9}$$

The first delta function denotes that correlations of stochastic GWs are time-independent. The second delta function shows that the momenta form a closed triangle due to homogeneity of the Minkowski spacetime. The bispectral shape varies depending on the inflation model, so that its measurement is a key to probe the early universe.

We define a correlation function of signals in three detectors with a filter function as follows:

$$S_{123} = \iiint_{-T/2}^{T/2} \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t_3\ s_1(t_1) s_2(t_2) s_3(t_3) Q(t_1, t_2, t_3)\,, \tag{10}$$

where the signal $s_i(t_i)$ in each detector contains the GW signal $z_i(t_i)$ and noises $n_i(t_i)$. We introduce a filter function in a form $Q(t_1, t_2, t_3) = Q(at_1 + bt_2 + ct_3)$, where $a, b$ and $c$ are positive constants. It will turn out that the filter function enables us to extract a specific configuration of the momentum triangle determined by these constants. Moving on to Fourier space, we have

$$S_{123} = \iiint_{-T/2}^{T/2} \mathrm{d}t_1 \mathrm{d}t_2 \mathrm{d}t_3 \iiiint_{-\infty}^{\infty} \mathrm{d}f_1 \mathrm{d}f_2 \mathrm{d}f_3 \mathrm{d}f\ \tilde{s}_1(f_1) \tilde{s}_2(f_2) \tilde{s}_3(f_3) \tilde{Q}(f) \times e^{2\pi \mathrm{i} f_1 t_1}\, e^{2\pi \mathrm{i} f_2 t_2}\, e^{2\pi \mathrm{i} f_3 t_3}\, e^{-2\pi \mathrm{i} f (a t_1 + b t_2 + c t_3)}\ , \tag{11}$$

where we have assumed $\tilde{Q}(-f) = \tilde{Q}(f)$. As in the case of the power spectrum, taking $T \to \infty$, one can carry out the integration:

$$S_{123} = \iiint_{-\infty}^{\infty} \mathrm{d}f_1 \mathrm{d}f_2 \mathrm{d}f_3 \mathrm{d}f\, \tilde{s}_1(f_1)\tilde{s}_2(f_2)\tilde{s}_3(f_3)\tilde{Q}(f)\delta(f_1 - af)\delta(f_2 - bf)\delta(f_3 - cf)$$
$$= \int_{-\infty}^{\infty} \mathrm{d}f\, \tilde{s}_1(af)\tilde{s}_2(bf)\tilde{s}_3(cf)\tilde{Q}(f)\,. \tag{12}$$

The ensemble average of $S_{123}$ becomes

$$\langle S_{123}\rangle = \int_{-\infty}^{\infty} \mathrm{d}f\, \langle \tilde{z}_1(af)\tilde{z}_2(bf)\tilde{z}_3(cf)\rangle \tilde{Q}(f)\ . \tag{13}$$

It should be noted that Eq. (13) is valid even for a single pulsar case as long as the noise is gaussian, namely, $\langle n_i(t)n_i(t)n_i(t)\rangle = 0$. From Eqs. (8), (9) and (13), one can deduce

$$\langle S_{123}\rangle = 2T \sum_{A,A',A''} \int_0^{\infty} \mathrm{d}f \frac{1}{f^3} B_{AA'A''}(af, bf, cf)\tilde{Q}(f) \frac{\sin\left(\pi(a+b+c)fT\right)}{\pi(a+b+c)fT}$$
$$\times \frac{(4\pi)^2}{abc}\Gamma^{AA'A''}(a, b, c; \hat{p}_1, \hat{p}_2, \hat{p}_3)\ , \tag{14}$$

where the ORF for the three point correlation is defined by [c]

$$\Gamma^{AA'A''}(a, b, c; \hat{p}_1, \hat{p}_2, \hat{p}_3) = -\frac{abc}{(4\pi)^2} \iiint \mathrm{d}\hat{\Omega}_1 \mathrm{d}\hat{\Omega}_2 \mathrm{d}\hat{\Omega}_3 \delta^{(3)}(a\hat{\Omega}_1 + b\hat{\Omega}_2 + c\hat{\Omega}_3)$$
$$\times F^{A}(\hat{\Omega}_1, \hat{p}_1)F^{A'}(\hat{\Omega}_2, \hat{p}_2)F^{A''}(\hat{\Omega}_3, \hat{p}_3)\ . \tag{15}$$

In Eq. (14), we see that the frequency and the angular integrals are separated due to the filter function, although there appears a suppression factor, $\int_{-T/2}^{T/2} e^{2\pi i(a+b+c)ft}\mathrm{d}t = \frac{\sin\left(\pi(a+b+c)fT\right)}{\pi(a+b+c)f}$, in the first line.[d] Eq. (15) shows that a specific configuration of the momentum triangle determined by $a, b$ and $c$ is extracted.[e] Therefore, we can probe the shape of the bispectrum with changing those parameters. For this purpose, we need to carry out the angular integration to evaluate the ORF. Although we do not study it in this paper, a full discussion is found in the paper[35].

## 4. Conclusion

In this paper, we investigate a method for detecting the bispectrum of stochastic GWs with pulsar timing arrays[35]. We showed that an appropriate filter function in

[c] Note that we have defined the ORF to be scale invariant with respect to $a$,$b$ and $c$.
[d] Allowing $a, b$ and $c$ to be negative, one can remove the suppression factor when $a + b + c = 0$, i.e., in the collinear limit. It implies that pulsar timing arrays is more sensitive to collinear limits of the bispectrum. However, we only focus on positive constants case to probe general momentum triangles in this paper.
[e] In an equal-time three point correlation function, all momentum triangles are integrated. Such case is well studied in the context of LISA[37–39].

three point correlations enables us to extract a specific configuration of momentum triangles in the bispectrum of stochastic GWs. Therefore, one can probe the bispectral shape, which carries important information of the early universe, by adjusting the filter function. Although we did not evaluate the ORF in this paper, a full discussion is found in the paper[35].

## Acknowledgments

A. I. would like to thank N. Bartolo, S. Matarrese, E. Dimastrogiovanni, M. Fasiello and S. Räsänen for useful comments and discussions. A. I. was supported by Grant-in-Aid for JSPS Research Fellow and JSPS KAKENHI Grant No.JP17J00216. T. N. was supported in part by JSPS KAKENHI Grant Numbers JP17H02894 and JP18K13539, and MEXT KAKENHI Grant Number JP18H04352. J. S. was in part supported by JSPS KAKENHI Grant Numbers JP17H02894, JP17K18778, JP15H05895, JP17H06359, JP18H04589. J. S and T. N are also supported by JSPS Bilateral Joint Research Projects (JSPS-NRF collaboration) "String Axion Cosmology."

## References

1. L. P. Grishchuk, Sov. Phys. JETP **40**, 409 (1975) [Zh. Eksp. Teor. Fiz. **67**, 825 (1974)].
2. A. A. Starobinsky, JETP Lett. **30**, 682 (1979) [Pisma Zh. Eksp. Teor. Fiz. **30**, 719 (1979)].
3. A. A. Starobinsky, Phys. Lett. B **91**, 99 (1980) [Phys. Lett. **91B**, 99 (1980)] [Adv. Ser. Astrophys. Cosmol. **3**, 130 (1987)].
4. K. Sato, Mon. Not. Roy. Astron. Soc. **195**, 467 (1981).
5. A. H. Guth, Phys. Rev. D **23**, 347 (1981) [Adv. Ser. Astrophys. Cosmol. **3**, 139 (1987)].
6. A. D. Linde, Phys. Lett. **108B**, 389 (1982) [Adv. Ser. Astrophys. Cosmol. **3**, 149 (1987)].
7. T. Regimbau, Res. Astron. Astrophys. **11**, 369 (2011) [arXiv:1101.2762 [astro-ph.CO]].
8. D. Coward and T. Regimbau, New Astron. Rev. **50**, 461 (2006) [astro-ph/0607043].
9. S. Drasco and E. E. Flanagan, Phys. Rev. D **67**, 082003 (2003) [gr-qc/0210032].
10. D. M. Coward and R. R. Burman, Mon. Not. Roy. Astron. Soc. **361**, 362 (2005) [astro-ph/0505181].
11. E. Racine and C. Cutler, Phys. Rev. D **76**, 124033 (2007) [arXiv:0708.4242 [gr-qc]].
12. N. Seto, Astrophys. J. **683**, L95 (2008) [arXiv:0807.1151 [astro-ph]].
13. N. Seto, Phys. Rev. D **80**, 043003 (2009) [arXiv:0908.0228 [gr-qc]].

14. E. Thrane, Phys. Rev. D **87**, no. 4, 043009 (2013) [arXiv:1301.0263 [astro-ph.IM]].
15. D. Baumann, arXiv:0907.5424 [hep-th].
16. D. Babich, P. Creminelli and M. Zaldarriaga, JCAP **0408**, 009 (2004) [astro-ph/0405356].
17. N. Bartolo, E. Komatsu, S. Matarrese and A. Riotto, Phys. Rept. **402**, 103 (2004) [astro-ph/0406398].
18. G. Goon, K. Hinterbichler, A. Joyce and M. Trodden, arXiv:1812.07571 [hep-th].
19. A. Ito and J. Soda, Phys. Lett. B **771**, 415 (2017) [arXiv:1607.07062 [hep-th]].
20. E. Dimastrogiovanni, M. Fasiello, G. Tasinato and D. Wands, arXiv:1810.08866 [astro-ph.CO].
21. R. Saito and T. Kubota, JCAP **1806**, no. 06, 009 (2018) [arXiv:1804.06974 [hep-th]].
22. X. Chen and Y. Wang, JCAP **1004**, 027 (2010) [arXiv:0911.3380 [hep-th]].
23. D. Baumann and D. Green, Phys. Rev. D **85**, 103520 (2012) [arXiv:1109.0292 [hep-th]].
24. T. Noumi, M. Yamaguchi and D. Yokoyama, JHEP **1306**, 051 (2013) [arXiv:1211.1624 [hep-th]].
25. N. Arkani-Hamed and J. Maldacena, arXiv:1503.08043 [hep-th].
26. H. Lee, D. Baumann and G. L. Pimentel, JHEP **1612**, 040 (2016) [arXiv:1607.03735 [hep-th]].
27. B. P. Abbott *et al.* [LIGO Scientific and Virgo Collaborations], Phys. Rev. Lett. **118**, no. 12, 121101 (2017) Erratum: [Phys. Rev. Lett. **119**, no. 2, 029901 (2017)] [arXiv:1612.02029 [gr-qc]].
28. J. Abadie *et al.* [LIGO Scientific and VIRGO Collaborations], Phys. Rev. D **85**, 122001 (2012) [arXiv:1112.5004 [gr-qc]].
29. L. Lentati *et al.*, Mon. Not. Roy. Astron. Soc. **453**, no. 3, 2576 (2015) [arXiv:1504.03692 [astro-ph.CO]].
30. Z. Arzoumanian *et al.* [NANOGRAV Collaboration], Astrophys. J. **859**, no. 1, 47 (2018) [arXiv:1801.02617 [astro-ph.HE]].
31. N. Bartolo *et al.*, JCAP **1612**, no. 12, 026 (2016) [arXiv:1610.06481 [astro-ph.CO]].
32. S. Sato *et al.*, J. Phys. Conf. Ser. **840**, no. 1, 012010 (2017).
33. G. Janssen *et al.*, PoS AASKA **14**, 037 (2015) [arXiv:1501.00127 [astro-ph.IM]].
34. W. Zhao, Y. Zhang, X. P. You and Z. H. Zhu, Phys. Rev. D **87**, no. 12, 124012 (2013) [arXiv:1303.6718 [astro-ph.CO]].
35. M. Tsuneto, A. Ito, T. Noumi and J. Soda, JCAP **1903**, no. 03, 032 (2019) [arXiv:1812.10615 [gr-qc]].
36. M. Anholm, S. Ballmer, J. D. E. Creighton, L. R. Price and X. Siemens, Phys. Rev. D **79**, 084030 (2009) [arXiv:0809.0701 [gr-qc]].
37. N. Bartolo *et al.*, arXiv:1806.02819 [astro-ph.CO].

38. N. Bartolo, V. De Luca, G. Franciolini, M. Peloso and A. Riotto, arXiv:1810.12218 [astro-ph.CO].
39. N. Bartolo, V. De Luca, G. Franciolini, M. Peloso, D. Racco and A. Riotto, arXiv:1810.12224 [astro-ph.CO].

# Correlation between GWs and neutrinos from core-collapse supernovae

Takami Kuroda[1,*], Kei Kotake[2], Kazuhiro Hayama[2], and Tomoya Takiwaki[3]

[1] *Institut für Kernphysik, Technische Universität Darmstadt, Schlossgartenstrasse 9, D-64289 Darmstadt, Germany*

[2] *Department of Applied Physics, Fukuoka University, 8-19-1, Jonan, Nanakuma, Fukuoka, 814-0180, Japan*

[3] *Division of Theoretical Astronomy, National Astronomical Observatory of Japan (NAOJ), 2-21-1, Osawa, Mitaka, Tokyo, 181-8588, Japan*

[*] *E-mail: takami.kuroda@physik.tu-darmstadt.de*

We present results from general-relativistic (GR) three-dimensional (3D) core-collapse simulations with approximate neutrino transport for three non-rotating progenitors (11.2, 15, and 40 $M_{\odot}$) using different nuclear equations of state (EOSs). We find that the combination of progenitor's higher compactness at bounce, that is a consequence of the use of softer EOS, leads to stronger activity of the standing accretion shock instability (SASI). We confirm previous predication that the SASI produces characteristic time modulations both in neutrino and gravitational-wave (GW) signals. Our results indicate that the correlation of the neutrino and GW signals, if detected, would provide a new signature of the vigorous SASI activity in the supernova core.

*Keywords*: Supernovae; hydrodynamics; gravitational waves; neutrinos.

## 1. Introduction

Core-collapse supernovae (CCSNe) have been attracting attentions of theoretical and observational astrophysicists for many decades. From multi-wavelength electromagnetic (EM) wave signals, a wide variety of observational evidence have been reported so far, which have all pointed toward CCSNe being generally aspherical. However the EM signals could only provide an indirect probe of the explosion mechanism of CCSNe, because they snapshot images of optically thin regions far away from the central engine. On the contrary, neutrinos and gravitational waves (GWs) are expected to provide direct probes of the inner-workings of CCSNe[1] as they propagate from the inner-core almost without interacting stellar mantle.

From a theoretical point of view, neutrino radiation-hydrodynamics simulations of CCSNe are converging to a point that multi-dimensional (multi-D) hydrodynamics instabilities including neutrino-driven convection and the Standing-Accretion-Shock-Instability (SASI[2]) play a crucial key role in facilitating the neutrino mechanism of CCSNe. Conventionally the GW and neutrino signatures reflecting such CCSNe mechanisms have been studied rather separately. For the neutrino signals, Ref. 3 was the first to find the SASI-induced modulations in the neutrino signals. They found that the SASI-induced modulation is clearly visible for two high-mass progenitors where high SASI activity was observed. Regarding the GW emissions from CCSNe, there are several emission processes in the postbounce phase. Among them, the most distinct GW emission process generically seen in recent self-consistent CCSN models is the one from the PNS surface oscillation[4,5].

The characteristic GW frequency increases monotonically with time due to an accumulating accretion to the PNS, which ranges from $\sim 100$ Hz to $\sim 1000$ Hz. As for another recent finding from self-consistent 3D models, it becomes clear that the SASI also produces a characteristic signature in the GW emission[5–7]. In this emission process, the GW frequency appears in relatively low frequency range of $\sim 100$ to 250 Hz, that is consistent with the SASI frequency, and persists when the SASI dominates over neutrino-driven convection[5,6].

These findings may raise a simple question whether there is some correlation between the SASI-induced neutrino and GW signals. Spotted by the neutrino and GW astronomy in the advanced era, the time is ripe to study in detail what we can learn about the explosion mechanism from the future *simulatenous* detection of neutrinos and GWs using outcomes of multi-D CCSN models. We will estimate neutrino event rates in both Hyper-K and IceCube from our 3D-GR models. We perform a correlation analysis between the GW and neutrino signals.

## 2. Numerical Methods

The numerical schemes for our 3D-GR models are essentially the same as those in Ref. 5. For the metric evolution, we employ the standard BSSN variables ($\tilde{\gamma}_{ij}$, $\phi$, $\tilde{A}_{ij}$, $K$, and $\tilde{\Gamma}^i$)[8,9]. Solving the evolution equations of metric, hydrodynamics, and neutrino radiation in an operator-splitting manner, the system evolves self-consistently as a whole satisfying the Hamiltonian and momentum constraints. The total stress-energy tensor is $T^{\alpha\beta}_{(\mathrm{total})} = T^{\alpha\beta}_{(\mathrm{fluid})} + \sum_\nu T^{\alpha\beta}_{(\nu)}$, where $T^{\alpha\beta}_{(\mathrm{fluid})}$ and $T^{\alpha\beta}_{(\nu)}$ are the stress-energy tensor of fluid and the neutrino radiation field, respectively. We consider three flavors of neutrinos ($\nu \in \nu_e, \bar{\nu}_e, \nu_x$) with no energy dependence. Here $\nu_x$ represents heavy-lepton neutrinos (i.e. $\nu_\mu, \nu_\tau$ and their anti-particles).

We use three EOSs based on the relativistic-mean-field theory with different nuclear interaction treatments, which are DD2 and TM1 of Ref. 10 and SFHx of Ref. 11. SFHx is the softest EOS followed in order by DD2, and TM1. We study frequently used solar-metallicity models of a 15 $M_\odot$ star[12], 11.2, and 40 $M_\odot$ star[13]. The 3D computational domain is a cubic box with 15000 km width and nested boxes with 8 refinement levels are embedded. Each box contains $128^3$ cells and the minimum grid size near the origin is $\Delta x = 458$m. Our 3D-GR models are named by the progenitor mass with the EOS in parenthesis like S15.0(SFHx) which represents the progenitor mass of 15.0 $M_\odot$ and the EOS SFHx are used. GWs are extracted from our simulations using the conventional quadrupole formula.

## 3. Results

We first present a short overview of hydrodynamics features in Fig. 1 (see Ref. 14 for more details). In S15.0(SFHx), the SASI activity was observed in the non-linear phase with global scale hot entropy structures. While in a lighter progenitor model

S11.2(SFHx), convection is observed by formation of small-scale convective bubbles behind the roundish stalled shock and dominates over the SASI motion.

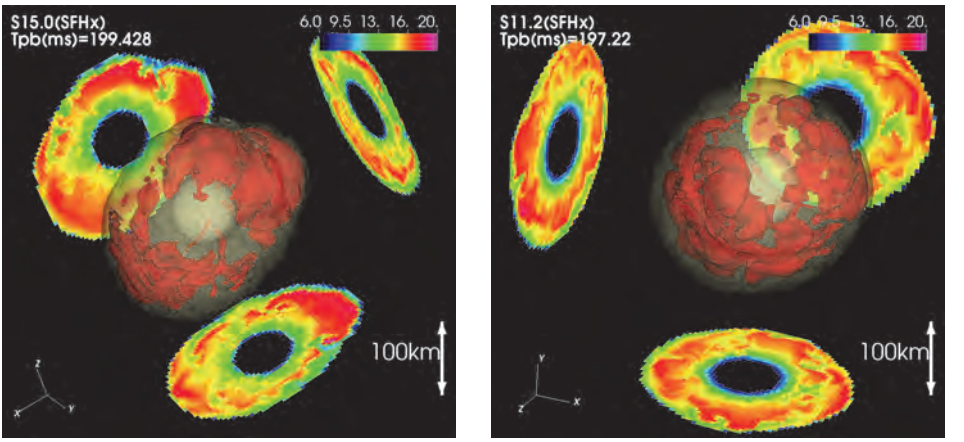



Fig. 1. Snapshots showing hydrodynamics features of represented models. Shown are the isentropic surfaces for $s = 7\ k_{\rm B}$ baryon$^{-1}$ (transparent shell) and for $s = 17\ k_{\rm B}$ baryon$^{-1}$ (red bubbles). The left and right panels are for S15.0(SFHx) and S11.2(SFHx), respectively. $T_{\rm pb}$ denotes the postbounce time. The contours on the cross sections in the $x = 0$, $y = 0$, and $z = 0$ planes are projected on the sidewalls. The left panel shows a typical (sloshing-)SASI structure, while the right one does many small scale convection bubbles.

To see how the SASI, that is more pronounced in S15.0(SFHx) than in S11.2(SFHx), makes a correlation between GW and neutrino signals, we show in Fig. 2 the GW amplitude $A_+$ [cm] and the neutrino event rate $N_\nu$ [ms$^{-1}$] in the top panel and their spectrum in the bottom. Here $N_\nu$ [ms$^{-1}$] is for Hyper-K (fiducial mass 440 kton[15]) and we show two extreme cases where the detector measures the original $\bar{\nu}_e$ (red line) or $\nu_x$ (green line) flux. The latter case corresponds to the complete flavor conversion through the Mikheyev-Smirnov-Wolfenstein (MSW) effect. In the bottom panel, the contours (red curves) correspond to the Fourier-decomposed anti-electron type neutrino event rates (two arbitrary chosen values of $dN_\nu/dF = 0.4$ (thin red line) and 0.8 (thick red line)). The left and right columns are for S15.0(SFHx) and S11.2(SFHx), respectively.

The bottom left panel (S15.0(SFHx)) shows a clear overlap between the neutrino modulation (see red contours in the spectrogram) and the GW modulation at $T_{\rm pb} \gtrsim$ 150 ms in the frequency range of $F \sim$ 100-150 Hz. Since the GW emission in the frequency range of $F \sim$ 100-150 Hz seen at $T_{\rm pb} \gtrsim 150$ ms is considered to be originated from strong SASI motion[5–7], the overlap naturally indicates that the SASI motion disturbs both the neutrino spheres and PNS core surface. The latter is the main emission region of the low frequency GWs[5,6]. On the other hand, in the smaller mass progenitor of S11.2(SFHx), we do not find any remarkable simultaneous oscillation of the neutrino and GW signals.

When the complete flavor conversion between $\bar{\nu}_e$ and $\nu_x$ is assumed (green line at $T_{\rm pb} \geq 150$ ms of model S15.0(SFHx) in Fig. 2), the time modulation is significantly suppressed as already reported in Ref. 3. This is because that the neutrino spheres of heavy-lepton neutrinos are located much deeper inside compared to those of anti-electron neutrinos. Consequently they are less affected by the SASI activity and the correlation between the GW and the neutrino event rate are expected to become weaker in the case of the complete flavor swap.

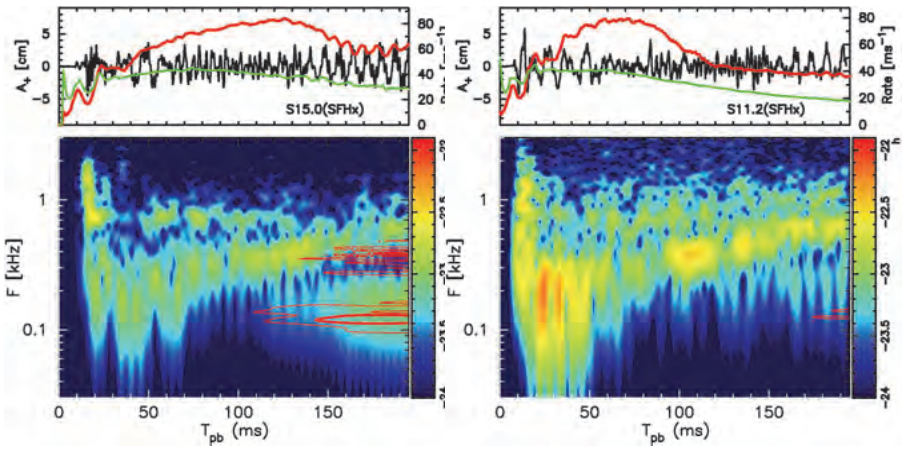



Fig. 2. For models S15.0(SFHx) (left) and S11.2(SFHx) (right), the top panel shows the neutrino event rate $N_\nu$ [ms$^{-1}$] (red and green lines are for $\bar{\nu}_e$ and $\nu_x$, respectively) for Hyper-K and the GW amplitude $A_+ = Dh_+$ [cm] (black line), whereas in the bottom panel we plot contours (red curves, only for $T_{\rm pb} \geq 100$ ms) of the anti-electron type neutrino spectra that are superimposed on the color-coded GW spectrum. Here $D$ and $h_+$ are the source distance of $D = 10$ kpc and the GW strain with plus mode, respectively. The observer's direction is fixed along the $z$-axis.

## 4. Summary and Discussion

We have presented results from our 3D-GR core-collapse simulations with approximate neutrino transport for two representative non-rotating progenitors (11.2 and 15 $M_\odot$). The SASI activity was seen for $15M_\odot$ star, while the entropy driven convection dominates over the SASI in $11.2M_\odot$ star. The combination of progenitor's higher compactness at bounce leads to the stronger SASI activity, that is consistent with previous studies, e.g., Ref. 16. Our 3D-GR models have confirmed previous predications that the SASI produces characteristic time modulations both in the neutrino and GW signals. Among the computed models, a 15.0 $M_\odot$ model exhibited more violent SASI motion, where the SASI-induced modulation in both GWs and neutrinos were clearly observed. The typical modulation frequency is in the range of ∼100-200 Hz, which is consistent with the oscillation period of the SASI motion.

Our results suggest that the correlation of the neutrino and GW signals, if detected, could provide a new signature of the vigorous SASI activity in the supernova core, which can be barely seen (like for the 11.2 $M_\odot$ model) if neutrino-convection dominates over the SASI.

In order to clarify whether we can or cannot detect the SASI-induced modulation in the GW and neutrino signals, we primarily need to perform a GW signal reconstruction study (e.g., Ref. 17). This is the most urgent task that we have to investigate as a sequel of this work. For a Galactic event, we apparently need third-generation detectors for observing the SASI-modulated GW signals[5–7], whereas the neutrino signals could be surely detected by IceCube and Super-K[3]. The neutrino burst can be used to determine the core bounce time, which raises significantly the detection efficiency of the GWs[18]. Our current study extends the horizon of previous prediction such as, when we would succeed the simultaneous detection of neutrino and GW signals from future nearby CCSN event, we could infer the supernova triggering dynamics (e.g., the SASI) from the following specific features (1) the low frequency ($F \sim 100$ Hz) modulation in both GW and neutrino signal and (2)[14] a few 10 ms time delay of the SASI-modulated GW signal from the SASI-modulated neutrino event rate.

## Acknowledgments

This research was supported by the European Research Council (ERC; FP7) under ERC Advanced Grant FISH-321263 (TK), ERC Starting Grant EUROPIUM-677912 (TK), JSPS KAKENHI Grant Number (JP15KK0173 (KK), JP17H05206, JP17K14306, JP15H00789, JP15H01039 (TK), and JP17H01130, JP17H06364 (KK and TT)), and JICFuS as a priority issue to be tackled by using the Post 'K' Computer. Numerical computations were carried out on Cray XC30 at CfCA, NAOJ.

## References

1. K. Kotake, Multiple physical elements to determine the gravitational-wave signatures of core-collapse supernovae, *Comptes Rendus Physique* **14**, 318 (April 2013).
2. J. M. Blondin, A. Mezzacappa and C. DeMarino, Stability of Standing Accretion Shocks, with an Eye toward Core-Collapse Supernovae, *ApJ* **584**, 971 (February 2003).
3. I. Tamborra, F. Hanke, B. Müller, H.-T. Janka and G. Raffelt, Neutrino Signature of Supernova Hydrodynamical Instabilities in Three Dimensions, *Physical Review Letters* **111**, p. 121104 (September 2013).
4. B. Müller, H.-T. Janka and A. Marek, A New Multi-dimensional General Relativistic Neutrino Hydrodynamics Code of Core-collapse Supernovae. III. Gravitational Wave Signals from Supernova Explosion Models, *ApJ* **766**, p. 43 (March 2013).

5. T. Kuroda, K. Kotake and T. Takiwaki, A New Gravitational-wave Signature from Standing Accretion Shock Instability in Supernovae, *ApJL* **829**, p. L14 (September 2016).
6. H. Andresen, B. Müller, E. Müller and H.-T. Janka, Gravitational wave signals from 3D neutrino hydrodynamics simulations of core-collapse supernovae, *MNRAS* **468**, 2032 (June 2017).
7. E. P. O'Connor and S. M. Couch, Exploring Fundamentally Three-dimensional Phenomena in High-fidelity Simulations of Core-collapse Supernovae, *ApJ* **865**, p. 81 (October 2018).
8. M. Shibata and T. Nakamura, Evolution of three-dimensional gravitational waves: Harmonic slicing case, *Phys. Rev. D* **52**, 5428 (November 1995).
9. T. W. Baumgarte and S. L. Shapiro, Numerical integration of Einstein's field equations, *Phys. Rev. D* **59**, p. 024007 (January 1999).
10. M. Hempel and J. Schaffner-Bielich, A statistical model for a complete supernova equation of state, *Nuclear Physics A* **837**, 210 (June 2010).
11. A. W. Steiner, M. Hempel and T. Fischer, Core-collapse Supernova Equations of State Based on Neutron Star Observations, *ApJ* **774**, p. 17 (September 2013).
12. S. E. Woosley and T. A. Weaver, The Evolution and Explosion of Massive Stars. II. Explosive Hydrodynamics and Nucleosynthesis, *ApJS* **101**, p. 181 (November 1995).
13. S. E. Woosley, A. Heger and T. A. Weaver, The evolution and explosion of massive stars, *Reviews of Modern Physics* **74**, 1015 (November 2002).
14. T. Kuroda, K. Kotake, K. Hayama and T. Takiwaki, Correlated Signatures of Gravitational-wave and Neutrino Emission in Three-dimensional General-relativistic Core-collapse Supernova Simulations, *ApJ* **851**, p. 62 (December 2017).
15. K. Abe, Hyper-Kamiokande Design Report, *KEK-PREPRINT-2016-21* (2016).
16. F. Hanke, A. Marek, B. Müller and H.-T. Janka, Is Strong SASI Activity the Key to Successful Neutrino-driven Supernova Explosions?, *ApJ* **755**, p. 138 (August 2012).
17. K. Hayama, T. Kuroda, K. Kotake and T. Takiwaki, Coherent network analysis of gravitational waves from three-dimensional core-collapse supernova models, *Phys. Rev. D* **92**, p. 122001 (December 2015).
18. S. E. Gossan, P. Sutton, A. Stuver, M. Zanolin, K. Gill and C. D. Ott, Observing gravitational waves from core-collapse supernovae in the advanced detector era, *Phys. Rev. D* **93**, p. 042002 (February 2016).

# Multi-peaked signatures of primordial gravitational waves from multi-step electroweak phase transition

Roman Pasechnik

*Department of Astronomy and Theoretical Physics, Lund University, 221 00 Lund, Sweden*
*E-mail: Roman.Pasechnik@thep.lu.se*

António P. Morais

*Departamento de Física, Universidade de Aveiro and CIDMA, Campus de Santiago, 3810-183 Aveiro, Portugal*
*E-mail: a.morais.physics@gmail.com*

Thibault Vieu

*Magistère de Physique Fondamentale, Université Paris-Saclay, Bât. 470, F-91405 Orsay, France*
*International Centre for Fundamental Physics, Ecole Normale Supérieure, 24 rue Lhomond, 75005 Paris, France*
*E-mail: vieu@apc.in2p3.fr*

The first-order electroweak phase transition in the early universe could occur in multiple steps leading to specific multi-peaked signatures in the primordial gravitational wave (GW) spectrum. We argue that these signatures are generic phenomena in multi-scalar extensions of the Standard Model. In a simple example of such an extension, we have studied the emergence of reoccurring and nested vacuum bubble configurations and their role in the formation of multiple peaks in the GW spectrum.

Despite of the great success of measurements at the Large Hadron Collider (LHC), the persistent absence of new physics evidence is driving an increasing discomfort among the particle physics community. The current void of new phenomena either indicates that new physics can only be manifest at a larger energy scale than previously thought, or results from a lack of sensitivity of the current experiments measuring rare events. In fact, the weaker the interaction strength between the SM and new physics, the greater the challenge to probe it.

On the other hand, the recent discovery of a binary neutron star merger, firstly observed by the gravitational waves (GW) interferometers of the LIGO-Virgo collaboration[1], a new era of multi-messenger astronomy has begun. Furthermore, the reach of GW observatories is by no means exhausted and larger sensitivities are designed for future space-based interferometers such as those of the LISA[2], DECIGO[3] and BBO[4] collaborations. This opens up the door for a plethora of new studies including connections with both cosmology and particle physics (see e.g. Refs. 5, 6 and references therein). In particular, the potential observation of a stochastic GW background produced by violent processes in the early universe, e.g. by expanding vacuum bubbles associated with strong cosmological phase transitions, may well become a gravitational probe for beyond-the-SM (BSM) physics and a complement for collider measurements. In this contribution, we discuss the key implications of such successive transitions for GW signals.

In cosmology, thermal evolution of the EW-breaking vacuum as the universe cools down is determined by the temperature-dependent part of the one-loop effective potential. Given the field content and its quantum numbers of an underlying multi-scalar fundamental theory, the shape of the potential can be determined at any temperature $T$. In a configuration of two minima of the effective potential coexisting at the critical temperature $T_c$, by using CosmoTransitions[7] one computes numerically the Euclidean action $\hat{S}_3$ describing transitions between the corresponding phases. The temperature $T_n$, at which the nucleation of vacuum bubbles effectively occurs, is estimated by setting the probability to nucleate one bubble per horizon volume to unity, which translates into $\hat{S}_3/T_n \sim 140$. The sphaleron suppression criterion $v_c/T_c \gtrsim 1$, with $v_c = v(T_c)$ the Higgs vacuum expectation value (VEV), defines a strong first-order phase transition. This typically implies that the transition produces strong GW signals detectable by the next generation of interferometers and may be used for constraining BSM scenarios. Here we study successive strong first-order EWPTs which we refer to as multi-step transitions. As a result one has more than a single transition pattern for a particular point in the parameter space, which results in sequential nucleation of bubbles of different vacua.

A generic BSM scenario typically contains a large number of scalar degrees of freedom which can be advantageous e.g. for EW baryogenesis. Even reducing the scalar sector to a few fields, new unexplored possibilities of transition patterns arise, in particular, transitions in several successive first-order steps. Therefore, a non-trivial EWPT is expected and multi-step transitions may have occurred in the early universe[5]. The basic characteristics of such multi-peak spectra may also be affected by dynamics of other sectors, in particular, by the neutrino sector suggesting the use of primordial GW data for probing the neutrino mass generation mechanisms[6].

In order to illustrate the generic features of multi-step first-order EW transitions, consider for instance a minimal extension of the SM scalar sector inspired by the high-scale Grand-unified trinification theory[8]. Besides the SM Higgs field $\mathcal{H}_1$, it contains an additional EW doublet $\mathcal{H}_2$ and singlet $\varphi$ fields which are charged under a U(1) family symmetry. The resulting potential possesses an approximate descrete $\mathbb{Z}_2$ symmetry acting as $\mathcal{H}_j \to -\mathcal{H}_j$ $(j = 1, 2)$ and $\varphi \to -\varphi$ which significantly simplifies the vacuum structure of the model. An expansion of the scalar fields in terms of real components

$$\mathcal{H}_j = \frac{1}{\sqrt{2}} \begin{pmatrix} \chi_j + i\chi'_j \\ \phi_j + h_j + i\eta_j \end{pmatrix}, \qquad \varphi = \frac{1}{\sqrt{2}} \left(\phi_s + S_R + iS_I\right), \tag{1}$$

defines the quantum fluctuations $h_1$, $h_2$ and $S_R$ about the classical configurations $\phi_\alpha = \{\phi_1, \phi_2, \phi_s\}$, respectively. With this expansion, the classical potential reads

$$V_{\rm cl}(\phi_\alpha) = \tfrac{1}{2} m_\alpha^2 |\phi_\alpha|^2 + \tfrac{1}{8}\lambda_\alpha |\phi_\alpha|^4 + \tfrac{1}{4}\lambda_{\alpha\beta} |\phi_\alpha|^2 |\phi_\beta|^2 . \tag{2}$$

A comprehensive analysis of the tree-level vacuum structure was performed recently in Ref. 8. It was shown that the basic characteristics of EWPTs in this model, in

particular, sequential first-order transitions, are generic for multi-Higgs extensions of the SM. Thus, this model, due to the simplicity of its potential (2), could serve as a good benchmark scenario for further in-depth explorations of cosmological implications of multi-scalar BSM theories.

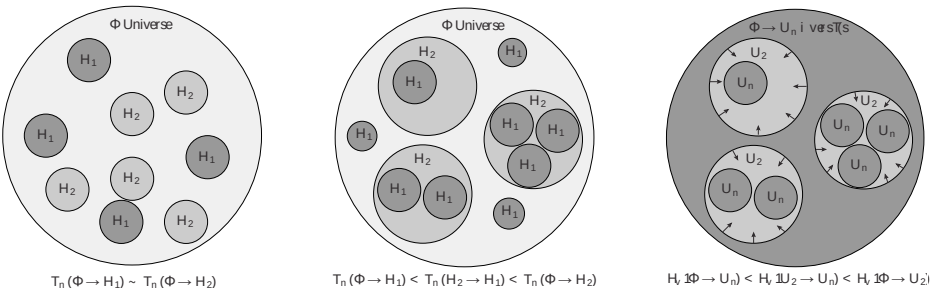


Fig. 1. An illustration of a universe in the Φ-phase filled with coexisting bubbles of new $H_1$ and $H_2$ phases emerging simultaneously (left panel), and also with nested bubbles when $H_1$-bubbles are born inside of $H_2$ ones (middle panel). In the right panel, the nucleation of the $H_1$-bubbles in the Φ phase causes the previously produced $H_2$-bubbles to contract simultaneously with nucleation of smaller $H_1$-bubbles inside them (reoccuring bubbles).

For simplicity, let us now consider a representative configuration of the parameter space[8] where the only existing phases given in terms of the VEVs of the scalar fields $v_\alpha \equiv \langle \phi_\alpha \rangle_{\rm vac} = \{v_1, v_2, v_s\}$ are $(0,0,0)$, $(v_1,0,0)$, $(0,v_2,0)$ and $(0,0,v_s)$, which we recast as $[0]$, $H_1$, $H_2$ and $\Phi$, respectively. The possible first-order transitions were found to be $H_1 \leftrightarrow H_2$, $H_1 \leftrightarrow \Phi$, $H_2 \leftrightarrow \Phi$, which take place readily in the leading $(m/T)^2$ order of the thermal expansion and are expected to be strong. Here, without loss of generality we identify the stable phase at $T = 0$ with $H_1$ which is simple but represents the basic features of a generic EW-breaking vacuum $\{v_1, v_2, 0\}$.

In a particular sequence of transitions to the true vacuum $H_1$ with the following two patterns

$$\Phi \to H_1\,, \qquad \Phi \to H_2 \to H_1\,, \tag{3}$$

one could expect several nucleation processes occurring in the same range of temperatures, e.g. $\Phi \to H_1$ and $\Phi \to H_2$. In this case, different sequences could be realized during the same cosmological evolution time leading to a universe where *coexisting bubbles* of different broken phases expand simultaneously (left panel in Fig. 1). In addition to the coexisting bubbles, more exotic cosmological objects may emerge from multi-step phase transitions. In particular, consider the second and third steps in the pattern $[0] \to \Phi \to H_2 \to H_1$, occurring at typical nucleation temperatures $T_n(\Phi \to H_2) \gtrsim T_n(H_2 \to H_1)$. Between $T_n(\Phi \to H_2)$ and $T_n(H_2 \to H_1)$, the $H_2$-bubbles nucleate and expand in a universe filled with the Φ-phase. Then at $T_n(H_2 \to H_1)$, while they are still expanding, the $H_1$-bubbles emerge and nucleate inside the $H_2$-bubbles. As such, the Φ-phase becomes populated with the $H_2$-bubbles containing the $H_1$-bubbles inside. We denote such objects as *nested bubbles* shown in Fig. 1 (middle panel).

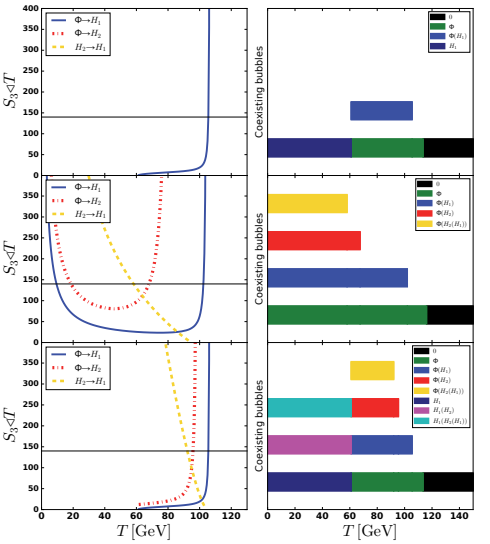



Fig. 2. Left column: evolution of $\hat{S}_3/T$ for all possible transitions. Right column: diagrams representing all types of bubbles (co)existing at a given temperature.

Since the scalar potential keeps evolving as the universe cools down below $T_n(H_2 \to H_1)$, the initial phase $\Phi$ becomes unstable also along the $H_1$ direction, so both transitions towards $H_1$ and $H_2$ can occur (coexisting bubbles scenario). In particular, if the potential barrier between the phases $\Phi$ and $H_1$ disappears, the new $H_1$-bubbles would nucleate in the parts of the universe that still remain in the $\Phi$-phase i.e. the direct $\Phi \to H_1$ transition quickly eliminates the $\Phi$-phase outside of the $H_2$-bubbles formed at an earlier time. Such a mixed situation with the coexistence of the ordinary $H_1$ and nested $H_2 \to H_1$ bubbles is depicted in Fig. 1 (middle panel). In the end of this process, one ends up with the $H_1$-bubbles inside the $H_2$-bubbles which exist in a universe filled with the $H_1$-phase. We denote these exotic cosmological objects as *reoccurring bubbles*. Since the $H_2$-bubbles cannot expand in a universe filled with the stable $H_1$-phase, they are pushed inwards and collapse while the $H_1$-bubbles nucleate inside them as illustrated in Fig. 1 (right panel).

In Fig. 2 we show three realistic cosmological scenarios where the objects discussed above are expected to occur. In the left column, we plot the evolution of the action $\hat{S}_3/T$ as a function of temperature for all possible transitions. Whenever a curve corresponding to a transition $i \to j$ crosses the horizontal line $\hat{S}_3/T = 140$, a bubble of phase $j$ is nucleated inside the phase $i$, which is denoted as $i(j)$ in what follows. In the right column, we show a diagrammatic representation displaying all types of bubbles (co)existing at a given temperature and corresponding to the plots in the left column. For example, the top panel describes a first-order phase transition $\Phi \to H_1$ when $H_1$-bubbles nucleate in a universe filled with the $\Phi$-phase (i.e. $\Phi(H_1)$). In particular, a universe in the symmetric phase [0] first collapses to

the Φ-phase through a second-order phase transition without generating any bubbles. Then, the $H_1$-bubbles are nucleated at $T \sim 105$ GeV and expand until the Φ-phase becomes unstable at around $T \sim 60$ GeV leaving a universe entirely filled by the true vacuum $H_1$.

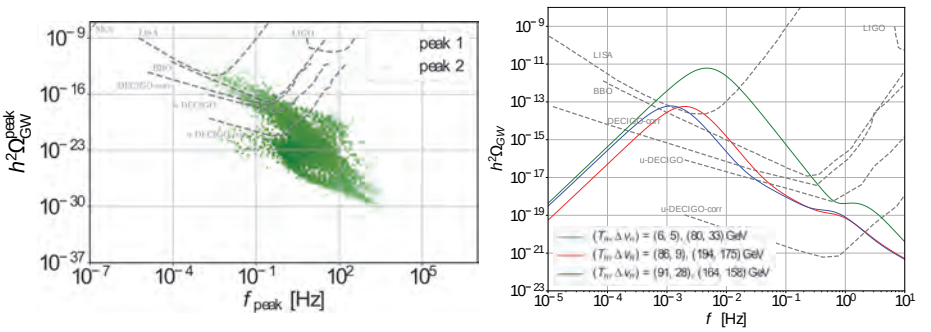



Fig. 3. **Left panel**: The results of the inclusive parameter scan in the considering 2HDSM where each point corresponds to a FOPT with the corresponding GW peak amplitude and frequency. Double-peak signatures are shown by red dots. **Right panel**: Benchmark double-peak GW spectra with the largest peaks entering the sensitivity domain of the LISA experiment with main characteristics specified in Table 1.

While the graphs on the second line of Fig. 2 represent a scenario where three successive steps lead to the nucleation of a nested bubbles $\Phi(H_2(H_1))$, those on the last line describe reoccurring bubbles that emerges from a nested ones. For instance, after nucleation of nested bubbles $\Phi(H_2(H_1))$, the potential barrier between the phases Φ and $H_1$ disappears (around $T \sim 60$ GeV), such that the parts of the universe in the Φ-phase collapse to the $H_1$-phase transforming the nested bubbles $\Phi(H_2(H_1))$ into the reoccurring ones $H_1(H_2(H_1))$. This type of cosmological objects is only possible if the nucleation temperatures of the corresponding steps are not too different (percolation typically occurs in the range of $\Delta T < 10$ GeV), and further likely to occur when symmetries in the potential enforces them to be identical.

We analyse the multi-peaked signatures in the power spectrum of GWs using the well known formalism of Ref. 9 which describes the energy density per logarithmic frequency of the GW radiation, $h^2\Omega_{\rm GW}$. The net GW signal is typically considered to be produced by three different sources due to bubble wall collisions, sound waves (SW) generated by the phase transitions, as well as magnetohydrodynamics (MHD) turbulences in the plasma. The bubble wall collisions typically do not contribute to the GWs production processes in the considering class of multi-scalar extensions of the SM (see also Ref. 5) so we account for the dominant SW and MHD contributions only.

The results of the inclusive parameter scan searching for FOPTs in the considering 2HDSM scenario are illustrated in Fig. 3 (left panel). Here, each point corresponds to a particular FOPT found in the scan, with calculated values of the

Table 1. Characteristics of selected benchmark double phase transitions whose GW signals emerge in the sensitivity domains of planned measurements and whose GW spectra are illustrated in Fig. 3 by green, red and blue lines, respectively. Here, the nucleation temperature, $T_n$, the scalar VEVs before $v_\alpha^i$ and after $v_\alpha^f$ the respective phase transition are given in units of GeV, while the peak-frequency, $f_{\rm peak}$, is given in Hz.

| line | $T_n$ | $\alpha$ | $\beta/H$ | $v_b$ | $v_1^i$ | $v_1^f$ | $v_2^i$ | $v_2^f$ | $v_s^i$ | $v_s^f$ | $f_{\rm peak}$ | $h^2\Omega_{\rm GW}^{\rm peak}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| green | 80 | $9.6{\cdot}10^{-3}$ | $6.1{\cdot}10^{4}$ | 0.65 | 0 | 87 | 86 | 81 | 0 | 0 | 1.6 | $4{\cdot}10^{-19}$ |
| | 6 | 0.4 | $3.4{\cdot}10^{3}$ | 0.88 | 240 | 246 | 23 | 0 | 0 | 0 | $5{\cdot}10^{-3}$ | $6.1{\cdot}10^{-12}$ |
| red | 194 | $7.1{\cdot}10^{-3}$ | $1.1{\cdot}10^{4}$ | 0.64 | 0 | 0 | 0 | 175 | 0 | 0 | 0.6 | $6.8{\cdot}10^{-20}$ |
| | 86 | 0.1 | 96 | 0.79 | 0 | 240 | 231 | 0 | 0 | 0 | $2{\cdot}10^{-3}$ | $5.7{\cdot}10^{-14}$ |
| blue | 164 | $6.6{\cdot}10^{-3}$ | $1.1{\cdot}10^{4}$ | 0.64 | 0 | 0 | 0 | 158 | 0 | 0 | 0.5 | $1.1{\cdot}10^{-19}$ |
| | 91 | $8.6{\cdot}10^{-2}$ | 51 | 0.77 | 0 | 235 | 207 | 0 | 0 | 0 | $1.3{\cdot}10^{-3}$ | $6{\cdot}10^{-14}$ |

induced GW peak amplitude and frequency. Scenarios corresponding to double-peak GW spectra are highlighted by red color. Here, dashed grey lines indicate sensitivities of the LISA and LIGO interferometers, as well as proposed DECIGO, BBO and SKA missions. We have selected three example scenarios whose largest peaks fall into the LISA sensitivity domain and whose GW spectra are shown in Fig. 3 (right panel). The basic characteristics of the corresponding transitions and GW signals for each green, red and blue line are summarised in Table 1. Such scenarios can be further considered as benchmarks for further explorations at GW interferometers. For a more detailed description of these and other benchmark scenarios, see Ref. 5.

## References


1. B. P. Abbott *et al.*, Observation of Gravitational Waves from a Binary Black Hole Merger, *Phys. Rev. Lett.* **116**, p. 061102 (2016).
2. N. Bartolo *et al.*, Science with the space-based interferometer LISA. IV: Probing inflation with gravitational waves, *JCAP* **1612**, p. 026 (2016).
3. S. Kawamura *et al.*, The Japanese space gravitational wave antenna: DECIGO, *Class. Quant. Grav.* **28**, p. 094011 (2011).
4. V. Corbin and N. J. Cornish, Detecting the cosmic gravitational wave background with the big bang observer, *Class. Quant. Grav.* **23**, 2435 (2006).
5. A. P. Morais and R. Pasechnik, Probing multi-step electroweak phase transition with multi-peaked primordial gravitational waves spectra, *JCAP* **2004**, p. 036 (2020).
6. A. Addazi, A. Marcianò, A. P. Morais, R. Pasechnik, R. Srivastava and J. W. F. Valle, Gravitational footprints of massive neutrinos and lepton number breaking (2019).
7. C. L. Wainwright, CosmoTransitions: Computing Cosmological Phase Transition Temperatures and Bubble Profiles with Multiple Fields, *Comput. Phys. Commun.* **183**, 2006 (2012).

8. T. Vieu, A. P. Morais and R. Pasechnik, Electroweak phase transitions in multi-Higgs models: the case of Trinification-inspired THDSM, *JCAP* **1807**, p. 014 (2018).
9. C. Caprini *et al.*, Science with the space-based interferometer eLISA. II: Gravitational waves from cosmological phase transitions, *JCAP* **1604**, p. 001 (2016).

# Gravitational wave luminosity and net momentum flux in head-on mergers of black holes: Radiative patterns and mode-mixing

Rafael F. Aranha[a], Ivano D. Soares[b] and Eduardo V. Tonini[c]

[a] *Universidade do Estado do Rio de Janeiro, Rio de Janeiro, RJ, 20550-900, Brazil*
[b] *Centro Brasileiro de Pesquisas Físicas, Rio de Janeiro, RJ, 22290-180, Brazil*
[c] *Instituto Federal do Espírito Santo, Vitória, ES, 29040-333, Brazil*
*E-mail:* [a] *rafael.aranha@uerj.br;* [b] *ivano@cbpf.br;* [c] *tonini@cefetes.br;*

We show that gravitational wave radiative patterns from a point test particle falling radially into a Schwarzschild black hole, as derived by Davis, Ruffini, Press and Price[1], are present in the nonlinear regime of head-on mergers of black holes. We use the Bondi-Sachs characteristic formulation and express the gravitational wave luminosity and the radiated energy in terms of the *news* functions. Our treatment is made in the realm of Robinson-Trautman dynamics, with characteristic initial data corresponding to the head-on merger of two black holes. We consider mass ratios in the range $0.01 \leq \alpha \leq 1$. We obtain the exponential decay with $\ell$ of the total energy contributed by each multipole $\ell$. The total rescaled radiated energy $E_W^{\rm total}/m_0\alpha^2$ decreases linearly with decreasing $\alpha$, yielding for the point particle limit $\alpha \to 0$ the value $\simeq 0.0484$, about five times larger than the result of Davis et al.[1] We also analyze the mode decomposition of the net momentum flux and the associated impulse of the gravitational waves emitted, resulting in an adjacent-even-odd mode-mixing pattern with the dominant contribution coming from the mixed mode $(2, 3)$. We obtain the exponential decay with $\ell$ of the total gravitational wave impulse contributed by each $(\ell, \ell + 1)$ mixed mode.

*Keywords*: Black hole head-on mergers; Gravitational wave luminosity; Net momentum flux; Radiative patterns.

## 1. Introduction

The collision and merger of two black holes are among the astrophysical sources which produce gravitational waves in the strong field regime and are therefore of crucial interest for the present direct observations made by the LIGO/VIRGO consortium[2].

In the realm of general relativity the production and extraction of gravitational waves in processes involving black holes have been investigated basically within three complementary approaches, most of them connected to binary black hole inspirals: Post-Newtonian (PN) approximations[3], numerical relativity[4] and the close-limit-approximation (CLA) supplemented with PN calculations[5], as well as combinations of these approaches.

Our treatment[6] is based on the Bondi-Sachs (BS) energy-momentum conservation laws in the characteristic formulation[7], that regulate the gravitational wave radiative transfer processes of the system, in the realm of Robinson-Trautman (RT) spacetimes[8]. The characteristic initial data constructed for the RT dynamics already present a global apparent horizon so that the dynamics covers the post-merger phase of the system, which represents one of the most dynamic parts of the evolution, up to the final configuration of the remnant black hole[9].

The cornerstone of our approach is the dependence of the Bondi-Sachs net four-momentum wave flux[10] on the *news* functions – which are the basic quantities characterizing the gravitational wave degrees of freedom of the system and are, by definition, quantities of spin-weight[11] $s = -2$. This decomposition is exact in the nonlinear regime, leading to an accurate evaluation of the even parity signals and of their relative contribution to the physical quantities involved in the radiative processes of the system. All these features are discussed in the following sections.

Throughout the paper we use geometrical units $G = c = 1$.

## 2. Robinson-Trautman Spacetimes

Robinson-Trautman (RT) spacetimes[8] are asymptotically flat solutions of Einstein's vacuum equations that describe the exterior gravitational field of a bounded system radiating gravitational waves. The RT metric can expressed as

$$ds^2 = \left(\lambda(u,\theta,\phi) - \frac{2m_0}{r} - 2r\frac{P_{,u}}{P}\right)du^2 + 2dudr - \frac{r^2}{P^2(u,\theta,\phi)}d\Omega^2, \tag{1}$$

where $r$ is an affine parameter defined along the shearfree null geodesics determined by the vector field $\partial/\partial r$. Here $d\Omega^2 = d\theta^2 + \sin^2\theta d\phi^2$, where $\lambda(u,\theta,\phi)$ is the Gaussian curvature of the surfaces ($u$ =const, $r$ =const) defined by

$$\lambda(u,\theta,\phi) = P^2 + \frac{P^2}{\sin\theta}\left(\sin\theta\ \frac{P_{,\theta}}{P}\right)_{,\theta} + \frac{P^2}{\sin^2\theta}\left(\frac{P_{,\phi}}{P}\right)_{,\phi}. \tag{2}$$

$m_0$ is the only dimensional parameter of the spacetime and fixes the energy and length scales of the system. For the stationary case $m_0$ corresponds to the rest mass of the black hole with respect to an asymptotic Lorentz frame at the future null infinity. Einstein's equations yield

$$12m_0\ P_{,u} + P^3\left(\frac{(\lambda_{,\theta}\sin\theta)_{,\theta}}{\sin\theta} + \frac{\lambda_{,\phi\phi}}{\sin^2\theta}\right) = 0. \tag{3}$$

In the above, the subscripts $u$, $\theta$ and $\phi$ preceded by a comma denote derivatives with respect to $u$, $\theta$, $\phi$, respectively. Eq. (3), denoted RT equation, governs the dynamics of the gravitational field (which is totally contained in the metric function $P(u,\theta,\phi)$) and propagates the initial data $P(u_0,\theta,\phi)$ from a given initial characteristic surface $u = u_0$.

An important feature of RT spacetimes that establishes its radiative character arises from the expression of its curvature tensor that, in a suitable semi-null tetrad basis, assumes the form

$$R_{ABCD} = \frac{N_{ABCD}}{r} + \frac{III_{ABCD}}{r^2} + \frac{II_{ABCD}}{r^3}, \tag{4}$$

where the scalar quantities $N_{ABCD}$, $III_{ABCD}$ and $II_{ABCD}$ are of the algebraic type $N$, $III$ and $II$, respectively, in the Petrov classification of the curvature tensor[12].

Considering axial symmetry, the curvature tensor components in the above basis that contribute to $N_{ABCD}$ are $R_{0303} = -R_{0202} = -D(u,\theta)/r + \mathcal{O}(1/r^2)$ where

$$D(u,\theta,\phi) = -P^2\partial_u\left(\frac{c_{,u}}{P}\right), \quad \text{with} \quad c_{,u}(u,\theta) = \frac{1}{2}\left(\partial^2_{\theta\theta} - \cot\theta\partial_\theta\right)P(u,\theta). \tag{5}$$

The *news* function $c_{,u}$ is one of the fundamental objects in our analysis. Generally speaking, it informs the presence of gravitational waves being emitted from a given bounded source and observed at the wave zone (Petrov type $N$).

## 3. The Bondi-Sachs Conservation Laws: The Gravitational Wave Luminosity and the Net Gravitational Wave Impulse

From the supplementary vacuum Einstein equations in the BS integration scheme together with the outgoing radiation condition, the BS four-momentum conservation laws for axisymmetric RT spacetimes are given by [10]

$$\frac{dP^\mu(u)}{du} = -\frac{1}{4\pi}\int_0^{2\pi} d\phi \int_0^\pi \frac{1}{P}\, l^\mu (c_{,u})^2 \sin\theta \; d\theta, \tag{6}$$

where $P^\mu(u)$ is the BS four-momentum. In the above the four vector $l^\mu = (1,0,0,\cos\theta)$ defines the generators of the translations of the BMS group in the temporal and Cartesian $z$ axes of an asymptotic Lorentz frame at future null infinity [13]. The luminosity of the gravitational waves emitted is given by the right-hand side of (6).

The integration in $u$ of Eq. (6) results for $\mu = 0$ and $\mu = z$, respectively, in the total energy $E_W(u)$ and the net impulse $I^z_W(u)$ of the gravitational wave emission, given as

$$E_W(u) = \frac{1}{2}\int_{u_0}^u du \int_0^\pi \frac{(c_{,u})^2}{P}\sin\theta \; d\theta, \quad I^z_W(u) = \frac{1}{2}\int_{u_0}^u du \int_0^\pi \frac{(c_{,u})^2}{P}\cos\theta\sin\theta \; d\theta, \tag{7}$$

where $u_0$ is the initial time. Due to the axisymmetry, the components $I^x_W(u) = 0 = I^y_W(u)$ for all $u$.

## 4. Initial data and numerical evolution

The initial data to be used was derived in Aranha et al. [9] and that can be interpreted as representing two instantaneously Schwarzschild black holes in head-on merger along the $z$ axis, at $u = u_0$,

$$P(u_0,\theta) = \left(\frac{\alpha_1}{\sqrt{\cosh\gamma + \cos\theta\sinh\gamma}} + \frac{\alpha_2}{\sqrt{\cosh\gamma - \cos\theta\sinh\gamma}}\right)^{-2}. \tag{8}$$

In the derivation of (8) it turns out that $\alpha = \alpha_2/\alpha_1$ is the mass ratio of the Schwarzschild masses of the initial data, as seen by an asymptotic observer. In the remaining of the paper we will take $\alpha_1 = 1$ and denote $\alpha_2 = \alpha$, the mass ratio. This data already has a single apparent horizon so that the evolution covers the post-merger regime up to the final configuration, when the gravitational wave emission

ceases. For $\alpha$ sufficiently small the data may be considered as a perturbation of a Schwarzschild black hole in the RT dynamics.

The initial data (8) is evolved numerically via the RT equation (3), which is integrated using a Galerkin method with a Legendre polynomial projection basis space adapted to the axisymmetric RT dynamics. The implementation of the Galerkin method, as well as its accuracy and stability for long time runs, is described in details in Section V of Aranha et al.[14] Exhaustive numerical experiments show that after a sufficiently long time $u \sim u_f$ all the modal coefficients of the Galerkin expansion become constant up to 12 significant digits, corresponding to the final time of computation $u_f$. At $u_f$ the gravitational wave emission is considered to effectively cease and we obtain $P(u_f, \theta)$ that, in all cases, can be approximated as

$$P(u_f, \theta) = P_f(\cosh\gamma_f + \cos\theta \sinh\gamma_f). \tag{9}$$

This final configuration corresponds to a Schwarzschild black hole boosted along the $z$ axis, with a final boost parameter $\gamma_f$ and a final Bondi rest mass $m_0/P_f^3$. In all cases $\gamma_f < \gamma$ and $P_f < 1$.

## 5. The mode decomposition of the radiative content of the gravitational wave emission: Energy patterns

We are now led to examine the total energy $E_W^{\rm total}$ carried out of the system by the gravitational waves emitted, expressed as

$$E_W^{\rm total} = \sum_{\ell\geq 2} E_{W\ell} = \sum_{\ell\geq 2} \frac{1}{4\pi}\int_{u_0}^{u_f} N_{\ell 0}^2(u)du. \tag{10}$$

where

$$N_{\ell 0}(u) = 2\pi \int_0^\pi \left(\frac{c_{,u}(\theta,u)}{\sqrt{P}}\right) {}_{-2}\mathcal{Y}_{\ell 0}(\theta)\sin\theta\, d\theta. \tag{11}$$

Here, ${}_{-2}\mathcal{Y}_{\ell 0}(\theta)$ is the spin-weight spherical harmonics[11] with $s = -2$ for the axial case.

Differentiating Eq. (10) with respect to $u$ gives the mode decomposition of the luminosity, corresponding to the exact RT equivalent of the Moncrief-Zerilli formula[15] for the radiated luminosity of a particle falling radially into a Schwarzschild black hole. For the mass ratios examined in our numerical simulations, we obtain that the radiated energy per multipole $\ell$ decays exponentially with $\ell$, as illustrated in Fig. 1 (left) for three values of $\alpha$ where we display the log-linear plot the points $E_{W\ell}/m_0$ versus $\ell$ for $\ell = 2...6$. The best fit curve to these points corresponds to the simple exponential law

$$E_{W\ell}/m_0 = \mathcal{A}\, e^{-\mathcal{B}\ell}. \tag{12}$$

The normalized rms error between the best fit straight lines of the log-linear plots and the points is of the order of, or smaller than 0.25%. This behavior is maintained up to $\alpha = 0.7$; a complete survey is given in Aranha et al.[6]

Its is remarkable that the exponential pattern of the plots – first observed in the computation by Davis et al.[1] of the gravitational radiation from a point test particle falling radially into a Schwarzschild black hole – extends to the nonlinear regime of head-on mergers for mass ratios at least up to $\alpha = 0.7$. In this sense, for the mass ratios considered in Fig. 1 (left), the initial data (8) may be considered to actually correspond to a perturbed Schwarzschild black hole. A complete survey is given in Aranha et al.[6]

The total radiated energy $E_W^{\rm total}/m_0$, for several mass ratios up to $\alpha = 0.3$, exhibits a simple linear relation with $\alpha$,

$$E_W^{\rm total}/m_0\alpha^2 = 0.143275\ \alpha + 0.048462. \tag{13}$$

The values for $\alpha = 0.01$, not shown in Fig. 1 (left), correspond to $E_W^{\rm total}/(m_0\alpha^2) = 0.0500813$. The straight line is the best fit of the points with a normalized rms deviation $\simeq 0.35\%$. In the point particle limit ($\alpha \to 0$) we obtain $E_W^{\rm total}/m_0\alpha^2 = 0.048462$, about five times larger than the value $\sim 0.0104$ of Davis et al.[1]

## 6. The total impulse imparted to the system via emission of gravitational waves: The adjacent-even-odd mode mixing

In the same vein we now examine the net impulse applied to the system due to the emission of gravitational waves. Following the last section, it can be expressed by

$$I_W^z(u_f) \equiv \sum_{\ell\geq 2} I^z_{W(\ell,\ell+1)} = \sum_{\ell\geq 2} \frac{1}{2\pi}\sqrt{\frac{(\ell+3)\ (\ell-1)}{(2\ell+1)(2\ell+3)}}\int_{u_0}^{u_f} N_{\ell 0}(u)N_{\ell+1,0}(u)du \tag{14}$$

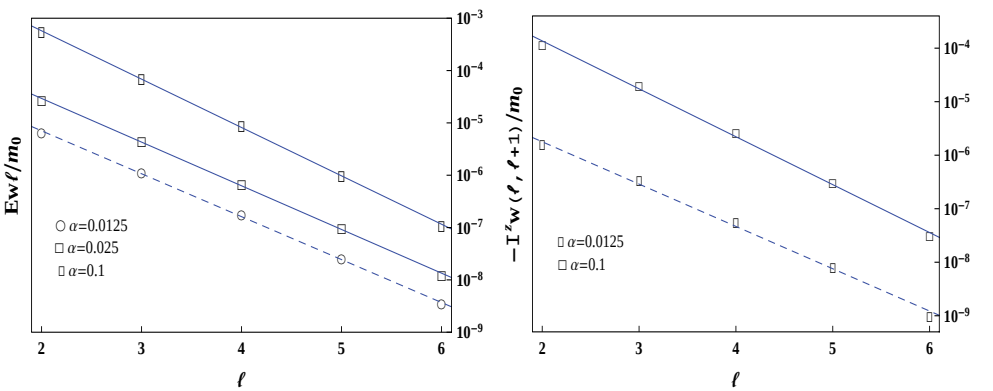



Fig. 1. *Left*: Plot of the points $E_{W\ell}/m_0$ versus $\ell$, for the several mass ratios $\alpha$ considered. This pattern – first observed by Davis et al.[1] in the gravitational radiation of a point test particle falling radially into a Schwarzschild black hole – is seen to be still maintained in the nonlinear regime of head-on merger of two black holes. *Right*: Log-linear plots of the mixed-mode impulses $-I^z_{W(\ell,\ell+1)}/m_0$ versus $\ell$, for the mass ratios $\alpha = 0.0125,\ 0.1$. The points are accurately fitted by straight lines, showing the exponential decay of the impulses with $\ell$.

that corresponds to the mode decomposition of the total net impulse into adjacent-even-odd mixed modes $I^z_{W(\ell,\ell+1)}$, which can be evaluated from the $N_{\ell 0}(u)$ given as in (11). We remark that the right-hand-side of (14) can be evaluated independently, so that we may obtain the impulse per mode $I^z_{W(\ell,\ell+1)}$ and the percentage of the contribution of each $(\ell, \ell+1)$ mixed-mode to the total net impulse imparted to the system. This can be seen in Fig. 1 (right).

This mode-mixing effect for the total momentum fluxes and the associated recoil velocities was first reported by Moncrief[15] for small odd-parity axisymmetric perturbations in the Oppenheimer-Snyder collapse models, and by Lousto and Price[16] for even-parity axisymmetric perturbations on a Schwarzschild black hole by a particle falling radially.

## Final Comments

The extension of this behavior for the whole mass ratio range $0 < \alpha \leq 1$ received a detailed examination in Ref. 6, both for the radiated energy per mode as well as for the mixed mode decomposition of the net momentum flux and impulse. In this reference we also examined the luminosity and the energy carried out by gravitational waves for increasing boost parameters $\gamma = [0.5, 0.6, 0.7, 0.8, 1.3]$. We obtain that the head-on mergers become more energetic as $\gamma$ increases, while the time duration of the GW bursts $\Delta u/m_0 \simeq 2$ remains approximately constant with increasing $\gamma$.

## Acknowledgments

The authors acknowledge the partial financial support of CNPq/MCTI-Brazil, through a Post-Doctoral Grant 151396/2013-2 (RFA), Research Grant 308728/2017-3 (IDS), and of FAPES-ES-Brazil (EVT).

## References

1. M. Davis, R. Ruffini. W. H. Press and R. H. Price, Phys. Rev. Lett. **27**, 1466 (1971).
2. LIGO detection papers at https://www.ligo.caltech.edu/page/detection-companion-papers.
3. L. Blanchet, Living Rev. Relativity **9**, 4 (2006).
4. T. W. Baumgarte and S. L. Shapiro, *Numerical Relativity: solving Einstein's equations in the computer*, (Cambridge University Press, Cambridge, 2010).
5. C.F. Sopuerta, N.Yunes, P. Laguna, Phys. Rev. D **74**, 124010 (2006) [Erratum: Phys. Rev. D **75**, 069903(E) (2007)].
6. R. F. Aranha, I. Damião Soares and E. V. Tonini, Phys. Rev. D **94**, 064017 (2016).
7. H. Bondi, M.G.J. van der Berg and A.W.K. Metzner, Proc. R. Soc. Lond. A **269**, 21 (1962).
8. I. Robinson, A. Trautman, Proc. R. Soc. Lond. A **265**, 463 (1962).

9. R. F. Aranha, I. Damião Soares and E. V. Tonini, Phys. Rev. D **81**, 104005 (2010).
10. R.F. Aranha, I. Damião Soares, E.V. Tonini, Class. Quantum Gravity **30**, 025014 (2013).
11. J. N. Goldberg, A. J. MacFarlane, E. T. Newman, F. Rohrlich and E. C. G. Sudarshan, J. Math. Phys. **8**, 2155 (1967).
12. A. Z. Petrov, *Sci. Nat. Kazan State University* **114**, 55 (1954).
13. R. K. Sachs , Phys. Rev. **128**, 2851 (1962).
14. R.F. Aranha, I. Damião Soares, E.V. Tonini. Phys. Rev. D **85**, 024003 (2012).
15. V. Moncrief, Ann. Phys. NY **88**, 323 (1974).
16. C. O. Lousto and R. H. Price, Phys. Rev. D **56**, 6439 (1997).

# Constraint on the equation of state via supernova gravitational waves

Hajime Sotani[1,*], Takami Kuroda[2], Tomoya Takiwaki[1], and Kei Kotake[3]

[1] *National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[2] *Technische Universität Darmstadt, Schlossgartenstrasse 9, 64289 Darmstadt, Germany*
[3] *Fukuoka University, 8-19-1, Jonan, Nanakuma, Fukuoka, 814-0180, Japan*
[*] *E-mail: sotani@yukawa.kyoto-u.ac.jp*

We focus on spacetime oscillations, the so-called $w$-modes, of gravitational waves emitted from a protoneutron star in the postbounce phase of core-collapse supernovae. By adopting numerical results from recent relativistic three-dimensional supernova models, we find that the $w_1$-mode frequency multiplied by the radius of the protoneutron star is expressed as a linear function with respect to the stellar compactness insensitively to the nuclear equation of state. Combining with another universal relation of the $f$-mode oscillations, it is shown that the time dependent mass-radius relation of the protoneutron star can be obtained by observing both the $f$- and $w_1$-mode gravitational waves simultaneously. That is, the simultaneous detection of the two modes could provide a new probe into finite-temperature nuclear equation of state that predominantly determines the protoneutron star evolution.

*Keywords*: Equation of state, protoneutron stars, gravitational waves.

## 1. Introduction

The gravitational waves from the mergers of binary black holes and a binary neutron star have been successfully detected[1,2], which leads to the beginning of a new age of the gravitational wave astronomy. In addition to the second-generation detecters, such as Advanced LIGO, Advanced VIRGO, and KAGRA, third-generation detecters like Einstein Telescope and Cosmic Explorer are already being proposed. Via observations with such detectors, one would obtain an imprint of open problems in compact objects. The most promising source for gravitational waves must be a merger of binary system of compact objects, while the core-collapse supernovae, which produce compact objects, could be a secondary candidate.

Up to now, in order to study the gravitational wave signals from core-collapse supernovae, extensive numerical simulations have been done. Through the results obtained by such simulations, it is considered that the $g$-mode oscillations excited around the protoneutron star surface are one of the most important gravitational wave emission in the postbounce phase[3–5]. In practice, the typical frequency of $g$-mode oscillations is expressed as $\sim M_{\rm PNS}/R_{\rm PNS}^2$ with the mass $M_{\rm PNS}$ and radius $R_{\rm PNS}$ of a protoneutron star. Thus, one may extract the information of protoneutron star properties by observing the $g$-mode gravitational waves from protoneutron stars.

On the other hand, as another approach, the gravitational-wave asteroseismology is a powerful technique for extracting interior information of compact objects. With such a technique one can constrain the equation of state (EOS), mass, and compactness of cold neutron stars (e.g., Refs. 6, 7, 8). With respect to the case of protoneutron stars, a few studies have been done[9–12]. In this study, by adopting

the results of numerical simulation of relativistic three-dimensional supernova models, we systematically examine the spacetime oscillations, the so-called $w$-modes, of gravitational waves from protoneutron stars. More details about this study can be seen in Ref. 10.

## 2. Protoneutron star models

To make a linear analysis, one has to prepare background protoneutron star models. In the present study, we adopt the numerical results of relativistic three-dimensional simulation[5] with the $15M_\odot$ progenitor models, adopting two different EOSs, i.e., SFHx and TM1. In particular, we consider very early phase up to $\sim$ 250 ms after core-bounce. In order to provide a static, spherically symmetric background model, the numerical data on each time step is averaged in the angular direction. Then, the protoneutron star's surface is determined at a fiducial rest-mass density of $10^{10}$ g/cm$^3$. In the left panel of Fig. 1, the evolution of protoneutron star models with two EOSs is shown in $M_{\rm PNS}$–$R_{\rm PNS}$ plane. With time after core-bounce, the mass increases due to the mass accretion, while the radius decreases due to the cooling. So, the mass and radius of protoneutron stars change with time from lower right to upper left in the left panel of Fig. 1. We remark that the sequence of mass and radius of protoneutron stars depends strongly on the adopted EOS.

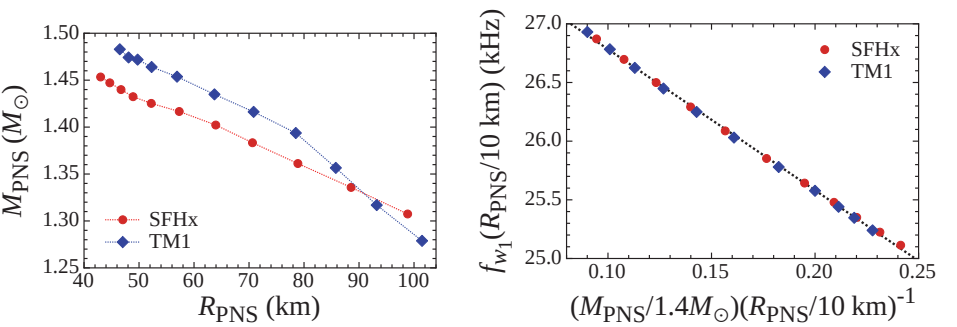



Fig. 1. Evolution of mass and radius of protoneutron stars in the left panel, and the $w_1$-mode multiplied by stellar radius as a function of stellar compactness in the right panel[10]. The circles and diamonds correspond to the results with SFHx and TM1 EOSs.

## 3. Asteroseismology in protoneutron stars

The perturbation equation derived from the linearized Einstein equations should be solved together with appropriate boundary conditions. First, we focus on the spacetime oscillations, the so-called $w$-modes, of gravitational waves from protoneutron star models. Since this mode is an oscillations of spacetime itself, the compactness, which is the ratio of $M_{\rm PNS}$ to $R_{\rm PNS}$, is an important parameter for describing

the frequencies[6]. In practice, we calculate the $w$-mode frequencies with the protoneutron star models at each time step and find that the evolution of frequencies depends on the adopted EOS. However, we also find that the frequencies multiplied with the radius can be expressed well as a linear function of the stellar compactness independently of the adopted EOS, as shown in the right panel of Fig. 1. In fact, we can derive the fitting formula of the $w_1$-mode frequencies as

$$f_{w_1}^{(\mathrm{PNS})}(\mathrm{kHz}) \approx \left[27.99 - 12.02 \left(\frac{M_{\mathrm{PNS}}}{1.4M_\odot}\right)\left(\frac{R_{\mathrm{PNS}}}{10\ \mathrm{km}}\right)^{-1}\right]\left(\frac{R_{\mathrm{PNS}}}{10\ \mathrm{km}}\right)^{-1}, \quad (1)$$

which is also plotted in the right panel of Fig. 1 with the dotted line.

On the other hand, we calculate the $f$-mode gravitational waves as in Ref. 9. Since the $f$-mode oscillations are a kind of acoustic oscillations, the frequencies can be characterized by the sound velocity or the stellar average density $M_{\mathrm{PNS}}/R_{\mathrm{PNS}}^3$[6]. In fact, we find that the time evolution of $f$-mode frequencies from protoneutron star depends on the adopted EOS, but the frequencies can be expressed as a linear function of the square root of the stellar average density independently of the adopted EOS as

$$f_f^{\mathrm{PNS}}(\mathrm{Hz}) \approx 14.48 + 4859 \left(\frac{M_{\mathrm{PNS}}}{1.4M_\odot}\right)^{1/2}\left(\frac{R_{\mathrm{PNS}}}{10\ \mathrm{km}}\right)^{-3/2}. \quad (2)$$

Now, we have two formulae for describing the $f$- and $w_1$-mode gravitational waves from protoneutron stars, which are independent of the adopted EOS for dense matter. So, by simultaneously observing the time evolution of two modes in the gravitational waves, one can obtain the time evolution of the stellar average density and compactness separately. That is, one can determine the mass and radius of protoneutron star at each time step after core-bounce as shown in the left panel of Fig. 1, which provides a new probe into the finite-temperature nuclear EOS.

## 4. Conclusion

The gravitational waves from supernovae are one of the most promising source. In the present analysis, we systematically examine the $w_1$- and $f$-mode gravitational waves after core-bounce, using the data of numerical simulations of relativistic three-dimensional supernova as a background model, and successfully derive universal formulae of $w_1$- and $f$-mode oscillations as a function of stellar compactness and average density independently of the adopted EOS. Thus, with the direct observations of gravitational waves from core-collapse supernovae, if one can identify the time evolution of $f$- and $w_1$-mode oscillations, one can in turn see the evolution of stellar average density and compactness, which tells us the evolution of mass and radius of a protoneutron star. As a result, one can in principle determine the finite-temperature nuclear EOS even through one event of gravitational waves from the core-collapse supernova.

**Acknowledgment**

This work was supported in part by Grant-in-Aid for Scientific Research (C) (No. 17K05458) provided by JSPS.

**References**

1. B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **116**, 061102 (2016).
2. B. P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration), Phys. Rev. Lett. **119**, 161101 (2017).
3. B. Müller, H. -T. Janka, and A. Marek, Astrophys. J. **766**, 43 (2013).
4. P. Cerdá-Durán, N. DeBrye, M. A. Aloy, J. A. Font, and M. Obergaulinger, Astrophys. J. Lett. **779**, L18 (2013).
5. T. Kuroda, K. Kotake, and T. Takiwaki, Astrophys. J. Lett. **829**, L14 (2016).
6. N. Andersson and K. D. Kokkotas, Mon. Not. R. Astron. Soc. **299**, 1059 (1998).
7. H. Sotani, K. Tominaga, and K. I. Maeda, Phys. Rev. D **65**, 024010 (2001).
8. H. Sotani, K. Nakazato, K. Iida, and K. Oyamatsu, Phys. Rev. Lett. **108**, 201101 (2012).
9. H. Sotani and T. Takiwaki, Phys. Rev. D **94**, 044043 (2016).
10. H. Sotani, T. Kuroda, T. Takiwaki, and K. Kotake, Phys. Rev. D **96**, 063005 (2017).
11. A. Torres-Forné, P. Cerdá-Durán, A. Passamonti, and J. A. Font, Mon. Not. R. Astron. Soc. **474**, 5272 (2018).
12. V. Morozova, D. Radice, A. Burrows, and D. Vartanyan, arXiv:1801.01914.

# Gravitoelectromagnetism: A way of Linking gravitational potential to the curvature

Dong Jiang*

*Yunnan Astronomical Observatories,*
*396 Yangfangwang, Guandu District, Kunming, 650216, P.R. China;*
*Key Laboratory for the Structure and Evolution of Celestial Objects, CAS;*
*Center for Astronomical Mega-Science, Chinese Academy of Sciences*
**E-mail: dj@ynao.ac.cn*
*https://github.com/JiangDong*

The relation between the physics and the difference geometry had been discussed many years, we think gravitoelectromagnetism is a way of Linking gravitational potential to the curvature in the theory and the experiments. We discuss the relation between Lense-Thirring Precession and Gravitomagnetic clock. And we report the new result of gravitational mass defect though still can not find gravitomagnetic mass. In the last, we discuss gravitomagnetic experiments in binary pulsars and near earth spacetime.

*Keywords*: General relativity experiments; gravitational potential; curvature; Gravitoelectromagnetism; difference geometry.

## 1. The Introduction

As we known that gravitoelectromagnetism's (GEM) tool that include Fermi-Walker transported and Lorenz group in the proper reference frame[1] had successful applied in the Thomas precession of quantum physics[2]. Considered the complex of difference geometry between an observer and the reference frame[3] the relation between Post-Newtonian approximation and gravitoelectromagnetism[4], and gravitoelectromagnetic analogy based on tidal tensors[5] just been discussed recently.

In the experiments, there are only three which inlcude Lense-Thirring Precession from X-ray astrophysics, Gravity probe B from gyroscope, and LARES from laser ranging that provide the blurry evidence. Here we think gravitoelectromagnetism can linking gravitational potential to the curvature in the theory and the experiments. After introducing and reviewing gravitoelectromagnetism in Sec. 2, we discuss gravitomagnetic experiments in binary pulsars and near earth spacetime in Sec. 3, in especial, we give the experiment plan to check Kerr character of Terrae spacetime in Sec. 3.1 and the reason that gravitoelectromagnetism can linking gravitational potential to the curvature from the field theory in Sec. 3.2.

## 2. Gravitoelectromagnetism

After Einstein had general relativity[6], the relation between the physics and the difference geometry had been discussed many years. Even if Einstein had field equation from difference geometry, he also doubt wherher it can describe the more physics for he believed in the experiments and equivalence principles. Whatever, gereral relativity had strong impact on compact physics[7], and many researcher want to merge it into other forces of interaction[8].

Gravitoelectromagnetism, refers to a set of formal analogies between the equations for electromagnetism and relativistic gravitation; specifically: between Maxwell's field equations and an approximation, valid under certain conditions, to the Einstein field equations for general relativity. The analogy and equations differing only by some small factors were first published in 1893, before general relativity, by Oliver Heaviside as a separate theory expanding Newton's law. Gravitomagnetism is a widely used term referring specifically to the kinetic effects of gravity, in analogy to the magnetic effects of moving electric charge. The most common version of GEM is valid only far from isolated sources, and for slowly moving test particles[9]. And from the classical theory of fields[10], we known the field of gravity and electromagnetism in the model of the particle in the central body field both could be divide to the field with source plus the vortex field. Therefor, we think GEM could describe the weak field of spacetime.

### 2.1. *Lense-Thirring Precession, and Gravitomagnetic clock*

In 1687, Newton publish his book[11], which include the famous discusson of inertial forces on a fiuid contained in a rotating vessel. This discussion was critically reexamined by Mach[12] in an attempt to understand better how inertial forces arise. He suggested that the shape of the water-surface may depend on the rotation of the vessel "if the sides of the vessel increased in thickness and mass till they were ultimately several leagues thick". A calculation of such effects became possible afetr Einstein formulated his general theory of relativity[6]. Lense and Thirring proposes the precession on the vertical direction of the motion[13].Lense-Thirring (LT) Precession (i.e. Frame dragging) aways was find in the astrophysical context and in Physics as gravitomagnetism effect.

Mashoon give the gravitomagnetic clock effect for the circular orbits[14]. Recently, generalized gravitomagnetic clock effect was calculated for the GPS near the earth[15].

As the above mentioned, if gravitational potential can be divided into gravitoelectric potential and gravitomagnetic potential by GEM field, When the potential difference of gravitoelectric potential is proportional to the gravitational frequency shift, gravitomagnetic potential also can naturally be reflected in time measurement by some form. If the precision of clock is enough to measure gravitomagnetic potential, that means we had the chance to explore gravitomagnetic clock effect.

### 2.2. *Gravitomagnetic mass and Gravitational mass defect*

In math-physics, after the work from Minkowski, Levi-Civita, Synge, Bondi, Pirani, Hawking, and Wald et al., general differential geometry becomes the mathematical foundations of relativity. Komar mass[16] and ADM mass[17] become the basis concept of general relativity, the early study show gravitomagnetic mass could be exist in the theory[18].

Here we report we think gravitational mass defect can be find in PSR B1913+16[19] though still can not find gravitomagnetic mass now.

## 3. Gravitomagnetic experiments in binary pulsars and near earth spacetime

In classical mechanics, the LaplaceRungeLenz vector (i.e. LRL vector) is a vector used chiefly to describe the shape and orientation of the orbit of one astronomical body around another, such as a planet revolving around a stellar. For two bodies interacting by Newtonian gravity, the LRL vector is a constant of motion, meaning that it is the same no matter where it is calculated on the orbit; equivalently, the LRL vector is said to be conserved. More generally, the LRL vector is conserved in all problems in which two bodies interact by a central force that varies as the inverse square of the distance between them; such problems are called Kepler problems.Therefor, the light bending and perihelion precession can have a unified approach which described by the LRL vector[20] as the evolution under the perturbed potentials for the conservation of the LRL vector origin from the inverse-square central force. For it, the model of two spin particle and the LRL vector[21] is well in most of binary pulsar for almost of it is the classical Kepler problems in the frame of classical Newtonian gravity.

For two neutron star system that is two compact sources system. In a *Gedanken-experiment*, to the extreme elliptic binary MSP system, we can have a hypothesis that the space-time of the orbit is flat in the apoapsis, BUT there is a new Curved space-time when two compact stars is in the periapsis for the masses of binary and the distance is very close. The Birth and death of this Curved space-time have related with Gravitational radiation. In this state, we can think it is the new Schwarzschild spacetime or Kerr spacetime in the periapsis. Therefore, LRL vector is not conservation and symmetry for it is not the inverse-square central force now. And the angular momentum vector of two stars is not conservation and symmetry for gravitional wave will bring the part of it. Therefor, we have the chance to mersure orbital symmetry violations in LT Precession of binary pulsar although it predicted by the particle in one body system[22,23]. and have the Angular Momentum of the Curved space-time in the periapsis by gravitomagnetic measurement[24]. Compared with Dr Wex, we think the frame of the particle in one body system is well in the periapsis. Whatever, $P_b$ and $\dot{P}_b$ are the directly measurement parameters to LT Precession of binary pulsar. But we still less the reasonable frame to estimate LT Precession and gravitomagnetic clock in the same time to pulsar - a natural clock and gyroscope.

Recently, the influence of the Coriolis Force in atom interferometry was find[25] in the earth spacetime, in fact, it was detect by torsion balance in 1919[26], and aways been study by geophysical and inertial navigation. In the frame of GEM, Semerk provide the way of spacetime dividing by five term[27,28] which include Coriolis term (one early reference that mentioned the apparent Coriolis force (Lense-Thirring

effect) and "induced Coriolis force", or "Lense-Thirring effect" by Pugh[29]). Here we think the Coriolis force can be seen the Coriolis term of spacetime in the order of gravitational potential for GEM field should extend to the Schwarzschild spacetime naturally. Then many relativistic phenomena are predicted and many Newtonian physics phenomenon are refresh in the viewpoint of spacetime geometry[19].

### 3.1. *Near earth spacetime as Kerr Spacetime*

Another interesting probelem is whether we can make a distinction between Schwarzschild and Kerr spacetime in the Terrae spacetime. To Kerr spacetime[30], it is the statinary spacetime which different with Schwarzschild spacetime that had the Static Observer[31]. In the procession of study the analogy between general relativity and electromagnetism for slowly moving particles in weak gravitational fields, some researcher think time variable term take place is the boundary of gravitoelectromagnetism application[9], here we think time variable term means that non-Static Observer take place which from Kerr spacetime. As the lunar tidal have impact on the free fall objects[2], and from Coriolis effect[25] and tidal effect was measured by the clock and the different type gyroscope (Optical, mechanical, vibration etc.), we think use two (or more) Full (or partly) Tensor Gradiometry (FTG) to interference in the large distance of the earth, if we can measure the time variable of the clock rate, that is enough to show that Kerr spacetime is existence for NO static observer in that scale spacetime. And the FTG Interferometer (Array) will help us in understand the large scale structure of spacetime, the partly tensor gradiometry only can find the time variable term. About the interferometric method of full tensor gravitational gradiometer array in measurement of gravitational theory and cosmology, we are applying the patent in China and in PCT[32].

### 3.2. *From the curvature to potential in the Gravitoelectromagnetism*

For the complex of definition about momentum (potential) of gravitational field, the local spread about the momentum of gravitational field, the momentum of gravitational wave, and the boundary between the local and non-local, all is the frontier of math-physics and theoretical physics. About gravitational wave, the objects just is in the state of fast than $C^{-5}$ which can provide Non-Local spread in the field theory[10]. Consider gravitational wave was find, we can think the $C^{-5}$ is the boundary between the local momentum interacting of gravitational field and non-local gravitational wave momentum spread. Consiered gravitomagnetic concept (potential and mass) cross the weak field and the drastic changes of the curvature, we think gravitoelectromagnetism as a way of linking gravitational potential to the curvature in the theory and the experiments is reasonable.

## 4. The Summary

We discuss gravitomagnetic experiments in binary pulsars and near earth spacetime, in particular, we give the experiment plan to check Kerr character of the earth spacetime and the reason that gravitoelectromagnetism can linking gravitational potential to the curvature after reviewed the theory of gravitoelectromagnetism.

## Acknowledgments

DJ thanks Project U1838104 supported by NSFC.

## References

1. R. T. Jantzen, P. Carini and D. Bini, The many faces of gravitoelectromagnetism, *Annals of Physics* **215**, 1 (April 1992).
2. C. Liang and B. ZHou, *Introduction to Differential Geometry and general relativity* (Science Press Pub., 2006).
3. W.-Q. Li and W.-T. Ni, Coupled inertial and gravitational effects in the proper reference frame of an accelerated rotating observer., *Journal of Mathematical Physics* **20**, 1473 (1979).
4. J. D. Kaplan, D. A. Nichols and K. S. Thorne, Post-Newtonian approximation in Maxwell-like form, *Phys. Rev. D* **80**, p. 124014 (December 2009).
5. L. F. O. Costa and C. A. R. Herdeiro, Gravitoelectromagnetic analogy based on tidal tensors, *Phys. Rev. D* **78**, p. 024021 (July 2008).
6. A. Einstein, Die Feldgleichungen der Gravitation, *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften (Berlin), Seite 844-847.* (1915).
7. T. D. Lee and Y. Pang, Stability of mini-boson stars, *Nuclear Physics B* **315**, 477 (March 1989).
8. C. N. Yang, Einstein and the Physics of the Second Half of the 20TH-CENTURY, in *The 2nd Marcel Grossmann Meeting: General Relativity*, ed. R. Ruffini1979.
9. E. G. Harris, Analogy between general relativity and electromagnetism for slowly moving particles in weak gravitational fields, *American Journal of Physics* **59**, 421 (May 1991).
10. L. D. Landau and E. M. Lifshitz, *The Classical Theory of Fields 4th edition (Translated from the Russian by M. Hamermesh)* 2008.
11. I. Newton, *Philosophiae Naturalis Principia Mathematica. Auctore Js. Newton* 1687.
12. E. Mach, *The science of mechanics: A critical and historical account of its development* 1893.
13. J. Lense and H. Thirring, Über den Einfluß der Eigenrotation der Zentralkörper auf die Bewegung der Planeten und Monde nach der Einsteinschen Gravitationstheorie, *Physikalische Zeitschrift* **19** (1918).

14. J. M. Cohen and B. Mashhoon, Standard clocks, interferometry, and gravitomagnetism, *Physics Letters A* **181**, 353 (October 1993).
15. E. Hackmann and C. Lämmerzahl, Generalized gravitomagnetic clock effect, *Phys. Rev. D* **90**, p. 044059 (August 2014).
16. A. Komar, Covariant Conservation Laws in General Relativity, *Physical Review* **113**, 934 (February 1959).
17. R. Arnowitt, S. Deser and C. W. Misner, Dynamical Structure and Definition of Energy in General Relativity, *Physical Review* **116**, 1322 (December 1959).
18. R. Ruffini and J. A. Wheeler, Relativistic cosmology and space platforms., *ESRO, SP, No. 52, p. 45 - 174* **52**, 45 (1971).
19. D. Jiang, Gravitational Mass Defect in the orbit of relativistic binary pulsars, *Submited* (2019).
20. D. R. Brill and D. Goel, Light bending and perihelion precession: A unified approach, *American Journal of Physics* **67**, 316 (April 1999).
21. B. M. Barker and R. F. O'Connell, Gravitational two-body problem with arbitrary masses, spins, and quadrupole moments, *Phys. Rev. D* **12**, 329 (July 1975).
22. T.-M. He and Y.-J. Wang, Frame dragging in the field of Kerr family, *Chinese Physics* **15**, 232 (January 2006).
23. A. K. Dubey and A. K. Sen, Frame-dragging from charged rotating body, in *Journal of Physics Conference Series*, , Journal of Physics Conference Series Vol. 759October 2016.
24. A. Tartaglia and M. L. Ruggiero, Gravitomagnetic Measurement of the Angular Momentum of Celestial Bodies, *General Relativity and Gravitation* **36**, 293 (February 2004).
25. S.-Y. Lan, P.-C. Kuan, B. Estey, P. Haslinger and H. Müller, Influence of the Coriolis Force in Atom Interferometry, *Physical Review Letters* **108**, p. 090402 (March 2012).
26. R. V. Eötvös, D. Pekár and E. Fekete, Beiträge zum Gesetze der Proportionalität von Trägheit und Gravität, *Annalen der Physik* **373**, 11 (1922).
27. O. Semerák, What forces drive the relativistic motion?, *Nuovo Cimento B Serie* **110**, 973 (August 1995).
28. O. Semerák, What forces act in relativistic gyroscope precession?, *Classical and Quantum Gravity* **13**, 2987 (November 1996).
29. G. E. Pugh, Proposal for a Satellite Test of the Coriolis Predictions of General Relativity, in *republished in Nonlinear Gravitodynamics: The Lense-Thirring Effect. 2003*, eds. R. Ruffini and C. SigismondiMay 1959.
30. R. P. Kerr, Gravitational Field of a Spinning Mass as an Example of Algebraically Special Metrics, *Phys. Rev. Lett.* **11**, 237 (September 1963).
31. R. K. Sachs and H.-H. Wu, *General Relativity for Mathematicians* 1977.
32. J. Dong, Interferometric method of full tensor gravitational gradiometer array in measurement of gravitational theory and cosmology, cn and pct applying (2019).

# Developing tools for multimessenger gravitational wave astronomy

Maria C. Babiuc Hamilton*

*Department of Physics, Marshall University, Huntington, WV 25755, US*
**E-mail: babiuc@marshall.edu*

The Marcel Grossmann triennial meetings are focused on reviewing developments in gravitation and general relativity, aimed at understanding and testing Einstein's theory of gravitation. The $15^{th}$ meeting (Rome, 2018) celebrated the $50^{th}$ anniversary of the first neutron star discovery (1967), and the birth of relativistic astrophysics. Another discovery of the same caliber is the detection of the binary neutron star GW170817 in 2017 – almost as if to celebrate the same jubilee – marking the beginning of multi-messenger gravitational wave astronomy. We present work in progress to craft open-sourced numerical tools that will enable the calculation of electromagnetic counterparts to gravitational waveforms: the `GiRaFFE` (General Relativistic Force-Free Electrodynamics) code. `GiRaFFE` numerically solves the general relativistic magnetohydrodynamics system of equations in the force-free limit, to model the magnetospheres surrounding compact binaries, in order (1) to characterize the nonlinear interaction between the source and its surrounding magnetosphere, and (2) to evaluate the electromagnetic counterparts of gravitational waves, including the production of collimated jets. We apply this code to various configurations of spinning black holes immersed in an external magnetic field, in order both to test our implementation and to explore the effects of (1) strong gravitational field, (2) high spins, and (3) tilt between the magnetic field lines and black hole spin, all on the amplification and collimation of Poynting jets. We will extend our work to collisions of black holes immersed in external magnetic field, which are prime candidates for coincident detection in both gravitational and electromagnetic spectra.

*Keywords*: Numerical relativity; black holes; gravitational waves, magnetosphere, jets.

## 1. Introduction

The 2017 detection of both gravitational waves (Ref. 1) and electromagnetic radiation (Ref. 2) from two colliding neutron stars is a very important event, akin to the discovery of neutron stars in 1967 by J. Bell and A. Hewish. This event – the rightfully-named Golden Binary – started the gold rush of multi-messenger gravitational wave astronomy. Observations from neutron star mergers provide deep insights into the highly-nonlinear interaction between strong gravity and the surrounding magnetosphere, which could explain the mechanism behind the emission of short gamma ray bursts and relativistic jets. Sources of strongly collimated astrophysical jets come at all scales, from pulsars to quasars, but they have in common one fact: they are compact objects rotating within their magnetospheres. Spinning neutron stars have a dipolar magnetic field and emit electromagnetic radiation due to the tilt between the spin axis and the magnetic field lines. Black holes don't have magnetic poles, and therefore their magnetic fields are external, being generated by the accretion disc. Although black holes tend to align their spin with the angular momentum of the accretion disc, the spin and orientation of the accretion disks usually change as the holes grow through mergers and gas accretion.[3]

Initially, merging supermassive black holes don't rotate in the same direction, and misalignment between the accretion disc and the black hole spin is expected.

Recent magnetohydrodynamic simulations show that for short timescales, jets are only partially aligned with the black hole spin.[4] The influence of the tilt between the black hole spin and the rotation axis of the accretion disc, and therefore the role the orientation of the magnetic field plays in the formation and quenching of jets, is not yet fully understood.[5] Another process that is not completely understood is what is under the hood of the main engine that powers the jet. Are high black hole spins essential in triggering jet production, or is the most important role played by the dynamics of the accretion disc?[6] Theoretical models show that relativistic jets can be generated from the rotational energy of a rapidly spinning black hole in strong magnetic fields through the Blandford-Znajek (BZ) mechanism.[7] The vacuum around a rotating black hole is electromagnetically active, and gives rise to strong electric fields when the black hole is surrounded by a magnetic field. Those gravito-rotationally induced electric fields cause a toroidal magnetic field, which plays an important role in collimating astrophysical jets. Astronomical observations report that the transition from radio-quiet to radio-loud quasars occurs for black hole spins around 0.7 to 0.9 times the speed of light.[8] If black hole spin is at the root of this transition, one should see a sharp delimitation in the value of the black hole spins necessary for the onset and quenching of the jet.[9] Another problem arises when two merging black holes, embedded in an external magnetic field, spiral towards each other. The gravitational waves produced in this astrophysical scenario are expected to exert a direct effect on the magnetic fields. This coupling between the magnetic field and the surrounding dynamical spacetime induces electromagnetic waves, coincidental with the production of gravitational waves, which provides a detection avenue for multimessenger gravitational wave astronomy.[10]

We engage in numerical explorations to study the interplay between black holes and the surrounding magnetosphere, in order to understand how spinning black holes drag the space-time around them, causing the collimation and amplification of jets. This is a problem not yet completely elucidated, although great strides have been made towards it's resolution.[11,12] We use the recently-released open-source `GiRaFFE` code (Ref. 13) to perform numerical simulation of single Kerr black holes evolving in external magnetosphere and to look at the effect of black hole spin on the output of electromagnetic luminosity in order to discern if there is a threshold spin for the production or extinction of collimated jets. Next we study the effect of tilted magnetospheres around a black hole with a fixed spin of 0.8 times the speed of light, in order to discern if the tilt angle plays any role in the amplification or reduction of the jet power.

## 2. Methods

*Theoretical Approach*: In order to explore the coupling between highly energetic gravitational and electromagnetic fields we use the General Relativistic Force-Free

Electrodynamics (`GRFFE` ) ansatz. `GRFFE` is a limiting case of General Relativistic MagnetoHydroDynamics (GRMHD), consisting on ideal plasma coupled to strong electromagnetic and gravitational fields, when the magnetic field energy dominates and the fluid pressure can be ignored. This theoretical model combines the assumption of perfect electric conductivity of the plasma with the fact that the Lorentz force in the comoving frame is immediately neutralized by induced electric currents, and it is commonly implemented in numerical codes to analyze the magnetized environment around pulsars or accreting black holes.[14] This `GRFFE` approach is implemented in our open-source `GiRaFFE` code (Ref. 15) designed to model the magnetosphere of highly-relativistic objects, such as neutron stars and black holes. The spacetime is evolved with the Einstein field equations, while the dynamics of the magnetized fluid are modeled using the conservation law $\Delta_\nu T^{\nu\mu}_{EM} = 0$ for the energy-momentum tensor $T^{\nu\mu}_{EM}$, and the Maxwell's equations $\Delta^*_\nu F^{\nu\mu} = 0$ for the electromagnetic tensor $F^{\nu\mu}$ in the force-free limit, $F^{\nu\mu} J_\nu = 0$. The updated variables are the magnetic field $B^i$ and the Poynting vector $S_i = -\gamma_{i\nu} n_\mu T^{\nu\mu}_{EM}$.[15] Here, $J_\nu$ is the 4-current, $\gamma_{i\nu}$ is spatial three-metric, and $n_\mu$ the timeline unit vector.

*Numerical Techniques*: The code evolves the vector potential $A_\mu = (\Phi, A_i)$ and Poynting vector $S_i$ one-forms, supplemented with an electromagnetic gauge evolution equation on a staggered grid, in order to keep the magnetic field divergenceless.[15] The evolution variables are linearly extrapolated to the outer boundary domain, which is causally disconnected from the interior by Adaptive Mesh Refinement (AMR). In order to prevent the force-free condition from breaking down during the evolution due to accumulation of numerical error, two supplementary conditions are imposed on $S_i$ during evolution to limit the direction and magnitude of the electric field so that it stays perpendicular to and smaller than the magnetic field: $E_i B^i = 0$ and $B^2 > E^2$. However, this condition does not account for the current sheet (CS), which inherently forms in magnetized plasmas and it's a critical component of the magnetospheres of rapidly spinning black holes.[11] We introduce a thin CS spatially localized at the equator by setting to zero the velocity perpendicular to the $z = 0$ plane.

*Initial Specifications*: We start with the magnetospheric Wald ansatz $A_t = 0, A_i = \frac{B_0}{2}(g_{i\theta} + 2a g_{it})$ in terms of the Kerr metric in spherical coordinates.[16] This vector potential creates a purely azimuthal magnetic field in Boyer-Lindquists coordinates, but in the horizon-penetrating Kerr-Schild coordinates there is an extra toroidal component $B_\phi$, due to the rotation of the geometry. The CS develops in the equatorial plane, forcing the anchored magnetic field lines to rotate with the black hole. There is no known exact solution for this magnetosphere, which is known as the "ultimate Rosetta Stone" for its close connection to the BZ process.[11,12]

The second initial configuration is a uniform azimuthal magnetic field created by a toroidal vector potential $A_\phi = \frac{B_0}{2} r^2 \sin^2\theta$, which in cartesian coordinates becomes $A_i = \frac{B_0}{2}(y, -x, 0)$ and produces a vertical magnetic field of magnitude $B_0$. With the CS in the equatorial plane, we tilt the magnetic field off the $z$-axis with an angle $\chi$ by a rotation $A_i \to \frac{B_0}{2}(y, -x\cos\chi, x\sin\chi)$ around the $x$-axis.

## 3. Results

The metric is described by the Kerr geometry in cartesian Kerr-Schild coordinates, radially shifted with $r_0 = \sqrt{1-a^2}$ such that $r \to 1 + r_0$, and the black hole inner horizon is at $r_- = 1$. We vary the spin from $a = 0.5c$ to $a = 0.95c$ and run a suite of 15 tests, with the constant $B_0 = 0.1M$ for the magnetic field. The numerical grid consists of $n = 5$ levels of refinement, with resolution scaling as $\Delta x_0/2^{n-1}$ where $\Delta x_0 = 2M$, and extends to $[-50M, +50M]^2$ in the $(x, y)$ plane and to $[-100M, 100M]$ in the $z$ direction. We visualize the EM luminosity: $L_{EM} = \int r^2 S d\Omega$, given by the Poynting flux across a sphere: $S = n^i \epsilon_{ijk} \sqrt{\gamma} E^j B^k$, $n^i = \frac{x^i}{r}$. In all the figures dark colors are low luminosity, bright colors are high luminosity.

*Magnetospheric Wald*: In Fig. 1 we show the time evolution of this testcase for $a = 0.95c$, at $r = 20M$. We distinguish two transient domains, between $t = \{0, 50\}M$ and $t = \{150, 200\}M$, when the electromagnetic field changes, and two stable phases: for $t = \{50, 150\}M$ with a collimated ring-like jet structure at the poles, and the second starting at $t = 200M$ when the jet broadens and precesses.

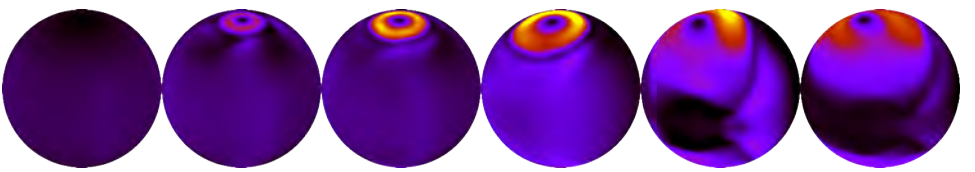


Fig. 1. Poynting luminosity of Wald magnetosphere with equatorial CS for Kerr black hole of spin $a = 0.95c$ and radius $r = 20M$. From left to right the time is $t = \{20, 60, 100, 140, 180, 200\}M$.

In Fig. 2 we plot the Poynting flux at $r = 20M$ and $t = 100M$ corresponding to the maximum amplitude and collimation of the jet. The intensity of the EM luminosity depends on the spin, as expected for the BZ process, however we discern the poloidal jet even at low spins, with no sign of spin threshold for jet production.

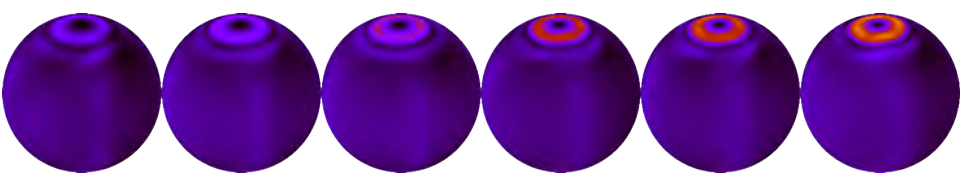


Fig. 2. Poynting luminosity of Wald magnetosphere with equatorial CS for Kerr black hole at $t = 100M$ and $r = 20M$. From left to right the spins are: $a = \{0.5, 0.6, 0.7, 0.8, 0.85, 0.9\}c$.

In Fig. 3 we plot the Poynting luminosity for a Kerr hole with spin $a = 0.95c$ for increasing radii $r = \{10, 20, 30, 40\}M$ at the time corresponding to the maximum

amplitude and collimation of the jet. We see that the power of the jet scale is inversely proportional to the radius, while the collimation is directly proportional.

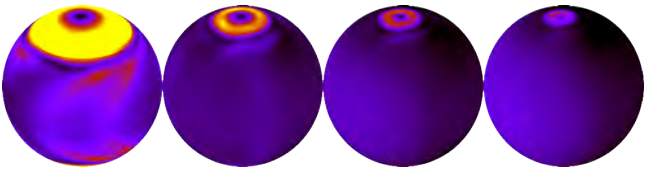


Fig. 3. Poynting luminosity of Wald magnetosphere with equatorial CS for Kerr black hole for $a = 0.95c$. From left to right $\{r, t\} = \{10M, 90M\}, \{20M, 100M\}, \{30M, 110M\}, \{40M, 120M\}$.

*Uniform Magnetosphere*: Fig. 4 shows that the time evolution of the luminosity for an initially vertical magnetosphere around a Kerr black hole is similar to the Wald magnetosphere. We compare the luminosity for four tilt angles at $t = 100M$ in Fig. 5 and at $t = 300M$ in Fig. 6, and we see that it increases with the tilt angle.

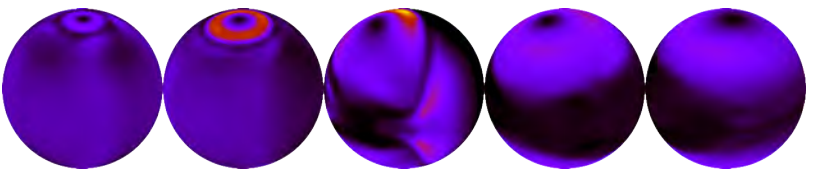


Fig. 4. Poynting luminosity of vertical magnetosphere with equatorial CS for Kerr black hole with $a = 0.8c$ and $r = 20M$. From left to right the time is $t = \{60, 120, 180, 240, 300\}M$.

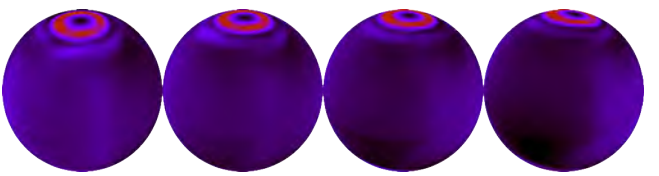


Fig. 5. Poynting luminosity of tilted magnetosphere with equatorial CS for Kerr black hole with $a = 0.8c$, $r = 20M$ and $t = 100M$. From left to right the angle is $\chi = \{0°, 15°, 30°, 45°\}$.

## 4. Discussion

We find that the geometry brings the black hole magnetosphere to a stability domain with a poloidal, precessing jet. We see two stable phases: early on, the jet forms a collimated, ring-like structure at the poles, while at late times, the jet broadens and precesses. Both the spin and the tilt angle affect the amplitude but not the shape of

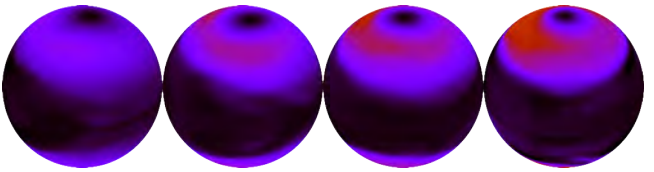


Fig. 6. Poynting luminosity of tilted magnetosphere with equatorial CS for Kerr black hole with $a = 0.8c$, $r = 20M$ and $t = 300M$. From left to right the angle is $\chi = \{0°, 15°, 30°, 45°\}$.

the jet, with no indication of high spin threshold for jet formation. Visualizations are available at `https://github.com/mbabiuc/KerrMagnetosphere`.

## Acknowledgments

This work supported by the National Science Foundation Award No. IIA-1458952. We gratefully acknowledge Z. Etienne and M. B. Wan for valuable discussions. The computer simulations were carried on the BigGreen Cluster at Marshall University.

## References

1. B. P. Abbott et al. (LIGO, Virgo), *Phys. Rev. Lett.* **119**, 161101 (2017).
2. B. P. Abbott et al. (LIGO, Virgo, Fermi, INTEGRAL, IceCube, AstroSat et al.), *Astrophys. J. Lett.* **848**, L12 (2017).
3. M. Volonteri, *M. Astron. Astrophys. Rev.* **10**, 279 (2010).
4. M. Liska, C. Hesp, A. Tchekhovskoy, A. Ingram, M. van der Klis, S. Markoff, *Mon. Not. R. Astron. Soc.* **474**, L81 (2017).
5. V. Mewes, J. A. Font, F. Galeazzi, P. J. Montero, N. Stergioulas, *Phys. Rev.* **D93**, 064055 (2016).
6. J. C. McKinney, A. Tchekhovskoy, R. D. Blandford, *Mon. Not. R. Astron. Soc.* **423**, 3083 (2012).
7. R. F. Penna, R. Narayan, A. Sadowski, *Mon. Not. R. Astron. Soc.* **436**, 3741 (2013).
8. E. Gardner and C. Done, *Mon. Not. R. Astron. Soc.* **473**, 2639-2654 (2017).
9. A. Schulze, C. Done, Y. Lu, F. Zhang, Y. Inoue, *Astrophys. J.* **849**, 1 (2017).
10. F. Cabral, F. S. N. Lobo, *Eur. Phys. J. Part. Fields* **C77(4)**, 237 (2017).
11. S. S. Komissarov, *Mon. Not. R. Astron. Soc.* **350**, 407 (2004).
12. C. Palenzuela, L. Lehner, S. Yoshida, *Phys. Rev.* **D81**, 084007 (2010).
13. Z.B. Etienne, M.C. Babiuc, "Open source code: GiRaFFE," `https://bitbucket.org/einsteintoolkit/` (2018).
14. V. Paschalidis and S. L. Shapiro, *Phys. Rev.* **D88**, 104031 (2013).
15. Z. B. Etienne, M. B. Wan, M. C. Babiuc, S. T. McWilliams, A. Choudhary, *Class. Quant. Grav.* **34**, 21 (2017).
16. R.M. Wald, *Phys. Rev*, **D10**, 1680 (1974).

# Development of position control system using thruster and thrust stand to measure $\mu$N-level force noise for drag-free control of DECIGO

Y. Hashimoto and S. Sato*

*Graduate School of Science and Engineering, Hosei University*
*3-7-2, Kajino-cho, Koganei, Tokyo, 184-8584, Japan*
** E-mail: sato.shuichi@hosei.ac.jp*

DECIGO (DECi-hertz Interferometer Gravitational wave Observatory) is a planned space mission targeting the detection of gravitational waves from very beginning of the universe, which no one has ever seen. As DECIGO utilizes freely flying Fabry-Perot cavities in orbit, a feasibility study of complex of its control systems including laser interferometer, which is tightly connected to drag-free control system together with formation-flying system, is a critical issue to realize DECIGO. As a first step to realize drag-free control required for DECIGO, we started from two experiments: the first one is a demonstration of position control using a thrust system, and the other is the development of a thrust stand to characterize the thrusters with $\mu$N-level force noise. In this paper, our progress and prospects for the next of these experiments are summarized.

*Keywords*: Gravitational wave; DECIGO; Drag-free control; Formation flying.

## 1. Introduction

The idea of DECIGO [1] is to build a laser interferometer having Fabry-Perot (FP) cavities in its arms with baseline of 1000 km, forming a triangular shape like LISA [2] in orbit. As test masses (TM), which work as mirrors for FP cavities also, probe for gravitational waves in spacetime, they are housed in spacecraft (S/C) and should be kept freely floating inside S/C using drag-free control. In addition to the control of three FP cavities, a drag-free control system tightly connected to formation-flying control to keep the three S/Cs in a triangular shape is essential to establish the control complex required for DECIGO. Using the FP cavity, instead of an optical transponder, for a laser link between S/Cs is a significant and also essential difference between DECIGO and LISA. To make these techniques feasible, sequential development of basic technologies are indispensable.

## 2. Demonstration of position control using thrust system

Drag-free control is a control system that S/C follow freely floating TMs inside S/C. The separation and the relative attitude between TMs and S/C are measured by local sensors, then by feeding these error signals back to the thrusters the S/C will ideally follow the TMs. Actually, as each S/C holds a couple of TM, common displacements and angle deviation of TMs relative to S/C are involved in drag-free control. On the other hand, error signals for differential degrees of freedom are fed back to the local actuators to provide forces on the TMs to maintain relative position and attitude between the TMs and the S/C. Therefore, the bandwidth of drag-free control should be below 10 Hz, in order not to degrade the signals in the observation band of the laser interferometer.

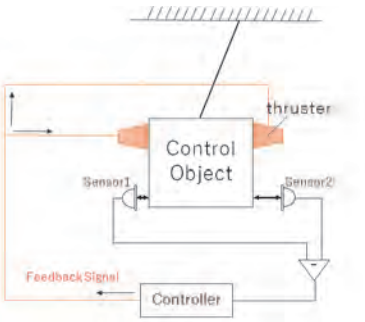



Fig. 1. The schematic diagram of position control experiment using thrust system.

The first step to demonstrate these complex systems is an experiment of position control using thruster-like propulsion as actuators, because one of most significant difference between terrestrial detectors and space missions is that S/C is to use thrusters as actuators instead of traditional electro-magnetic actuators. The object to be position-controlled, which is supposed to be a S/C thruster on board, is suspended with a single pendulum with one degrees of freedom for translation (Fig. 1). The position of the object is sensed using a photo sensor fixed on the optical table to give a differential displacement between the optical table and the object. We used a FPGA-based computer, cRIO (National Instruments), to provide digital filtering as a controller. A simple lag-lead filter was implemented on FPGA using LabVIEW software to stabilize the feedback control loop. We have composed a thruster-like propulsion system using a small $CO_2$ gas cartridge together with regulators and flow controller using piezo valves. Therefore, the error signals from the photo sensors were fed back to the input of piezo-controlled flow controller through the FPGA controller. The photo sensor and cRIO was located on the table top, so the signals from FPGA to the piezo valves were sent via a pair of very thin wires.

The behavior of error and feedback signals in the time domain is plotted in Fig. 2, which shows that the eigen mode of the pendulum at about 0.7Hz was certainly suppressed by engaging the control loop. The corresponding transfer function of the control loop is plotted in Fig. 2. The signals around this peak structure had been effectively fed back and suppressed as it was expected form time domain signals in Fig. 2. One of the next steps is to characterize and improve the current system to performance as well as DECIGO requires. The second is to extend the system to a demonstration of drag-free control for all degrees of freedom.

## 3. Development of thrust stand to measure $\mu$N-level force noise

The TMs are to be housed and freely floating inside S/C in an ideal situation. But in reality, TMs could be affected by external accelerations caused by many sources, including gravity, electro-magnetic interaction, cosmic rays, and so on. Especially,

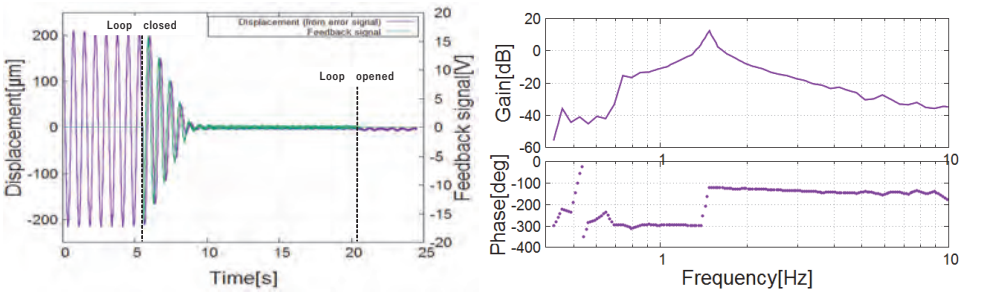



Fig. 2. The behavior of the error and the feedback signals in time domain (left) and the open-loop transfer function of control loop (right).

the displacement of S/C itself, which can act as a source of external acceleration on the TM through some coupling between TM and S/C, is thought to be one of most significant force noise sources. The force noise of the thrusters is an inevitable source for the displacement on S/C. Thus, in this sense, low force noise is an indispensable characteristic required for DECIGO. So far, the requirement for the force noise of the thruster is set to be $10^{-7}\mathrm{N}/\sqrt{\mathrm{Hz}}$ at 0.1 Hz, assuming some coupling between S/C displacements and external acceleration on the TM.

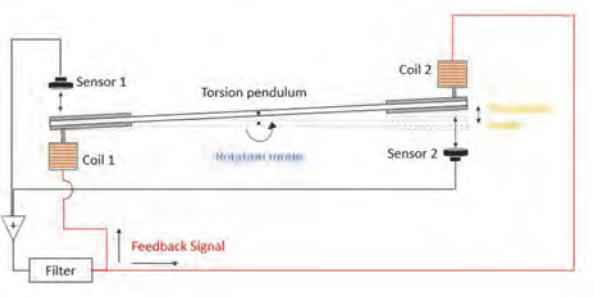



Fig. 3. The schematic diagram of the setup for the thrust stand system experiment.

We have started to develop a so-called thruster stand to measure the force noise spectrum of the thruster system. In order to extend measurement band width below 0.1 Hz, it is well-known, and also a beneficial idea to use a torsion pendulum for lowering resonant frequency, instead of single pendulum (Fig. 3). The measurement method of tiny force is based on a mechanism of a servo-type accelerometer, in which the error signals of position-controlled TM gives information on the external acceleration. Current thruster stands have no thrust systems on board, so the main objective of this setup is to evaluate the system noise of the measurement systems (Phase-1). The rotation of the torsion pendulum was sensed with a couple of photo

sensors located at opposite sides of the each end of the torsion bar. Adding these signals from photosensors can eliminate the common displacement signals from the translation mode of the pendulum, and thus gives an error signal for rotation degrees of freedom. This error signal was fed back to the coil-magnet actuator to maintain the rotation angle of the torsion bar.

The measured system noise of the thruster stand is plotted in Fig. 4 together with identified noise sources. The floor noise level below several Hz is already below our requirement except at around pendulum resonant frequency. We have also tried to modify our current system to have a thruster-like propulsion system on board (Phase-2). The concept of the measurement system is the same but the weight of the stand was increased to 2.4 kg due to the thrusters and a Michelson type laser interferometer was used as a displacement sensor instead of photo sensor. The floor noise level and the noise structure in the frequency range of our interest was changed according mainly to the change in the mechanical structure, as is shown in Fig. 4. The force measurement system using a torsion pendulum seems to work below our requirement noise level in terms of system noise. Thus, the next step is to actually measure the propulsion noise of the thrusters.

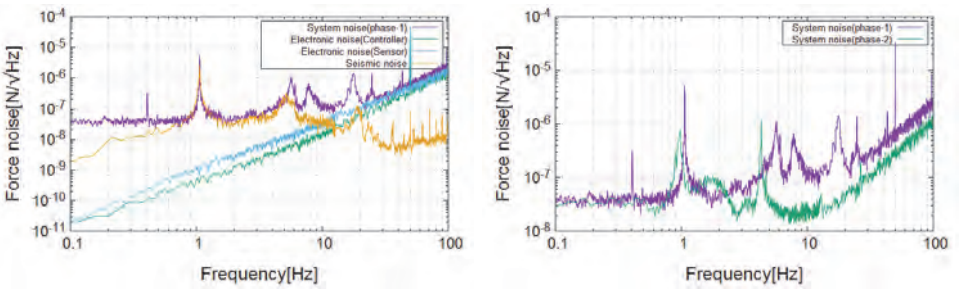



Fig. 4. The noise-equivalent sensitivity curve for the thruster stand system with identified noise sources (left) and the sensitivity curve for the thruster stand of phase-2 compared with phase-1 (right).

## 4. Conclusions

We have developed control and measurement systems closely related to a drag-free control of space gravitational wave antenna mission, DECIGO. The position control using a thrust system succeeded in showing the feasibility of the system for one degree of freedom of translation, and the thrust stand showed acceptable noise levels against requirement in terms of system noise.

## References

1. N. Seto, S. Kawamura and T. Nakamura, Phys. Rev. Lett. **87**, 221103 (2001) [astro-ph/0108011]. S. Kawamura *et. al*, Class. Quant. Grav. **23**, S125 (2006).
2. *The ESA-L3 GravitationalWaveMission Gravitational Observatory Advisory Team Final Report*, `http://www.cosmos.esa.int/web/goat` (2016).

# Constructing test bench for integration tests of components developed for DECIGO and B-DECIGO

Koji Nagano*

*KAGRA Observatory, Institute for Cosmic Ray Research, University of Tokyo, Kashiwa-city, Chiba 277-8582, Japan*
** E-mail: knagano@icrr.u-tokyo.ac.jp*

Tomofumi Shimoda, Yuta Michimura, and Masaki Ando

*Department of Physics, Graduate School of Science, University of Tokyo, Bunkyo-ku, Tokyo 113-0033, Japan*

DECIGO and B-DECIGO is space gravitational wave detecter project which aims decihertz band. In (B-)DECIGO, dual-pass differential Fabry-Pérot interferometer (DPDFPI) is planed to be used. Also, to reduce the noise from the space craft displacement, drag-free control is implemented. It is needed to demonstrate DPDFPI with drag-free control with a test bench. In the test bench, DPDFPI is constructed with softly suspended benches and cavities to mimic the motion in the space. In this article, the design of the optical and mechanical part in the test bench is reviewed. First, the optical design is shown. Second, mechanical suspension design is explained. All resonance modes of the important degree of freedom is designed to be less than $\sim 1$ Hz.

*Keywords*: Gravitational wave; laser interferometer; DECIGO; B-DECIGO.

## 1. Introduction

DECIGO and B-DECIGO is a proposed space-borne gravitational wave detector in decihertz frequency range.[1] (B-)DECIGO have many objectives, such as, observation of compact binary coalescences such as black holes and neutron stars, measurement of Hubble constant from long duration signal of neutron star inspirals, and, especially in DECIGO, characterization of inflation. The sensitivity of DECIGO will reach $\sim 10^{-24}\ 1/\sqrt{\mathrm{Hz}}$ since DECIGO has a Fabry-Pérot cavity in its arms which interact gravitational waves.

In DECIGO and B-DECIGO, the dual-pass differential Fabry-Pérot interferometer (DPDFPI) as shown in Fig. 1 is planed to be used. In DPDFPI, the laser lights from two satellites co-stand in one arm cavity and its length can be measured by two ways. This leads to the redundancy which is important in space mission. In addition, the number of test mass for one interferometer is reduced in DPDFPI.

Laser interferometers are operated in ground-based gravitational wave detectors, such as Advanced LIGO, Advanced Virgo, KAGRA.[2–5] However, specific control scheme of DPDFPI in (B-)DECIGO is still under consideration. In order to decide the control scheme of DPDFPI, some special techniques to space-borne detecters should be demonstrated, such as drag-free control with interferometer control. Drag-free control is a scheme where the distance and angle between the spacecraft and the test mass which is contained in the spacecraft are measured and controlled. This is necessary to reduce a noise coupled with spacecraft motion.

Before launch of (B-)DECIGO, DPDFPI with drag-free control has to be demonstrated on the ground. For the demonstration, construction of the test bench started. In this article, we reviewed the design of the test bench.

## 2. Design of the interferometer test bench

### 2.1. *Overview of the test bench*

The overview schematic of the test bench is shown in Fig. 1. The test bench is designed to have essential parts of (B-)DECIGO; DPDFPI and drag-free controlled spacecrafts. All main cavities and almost all input and output optics except laser sources and electro-optic modulators are placed in a vacuum chamber. The laser sources and its modulation systems for taking a length and angular signal of the cavity are placed outside of the chamber. The light is introduced though optical fibers into the chamber. The interferometer scale is about 1 m. This is limited by the size of the vacuum chamber. The detailed optical design is explained later. The mechanical system is divided into two components; test masses, i.e. cavity mirrors, and optical benches where the input and output optics are placed. The optical bench emulates the spacecraft in the space. To mimic low-gravity environment in space, the test masses and optical benches are softly suspended as explained later. As the demonstration of the drag free control, the relative motion and angle between the suspended optical bench and mirrors is measured and controlled with local sensors and actuators.

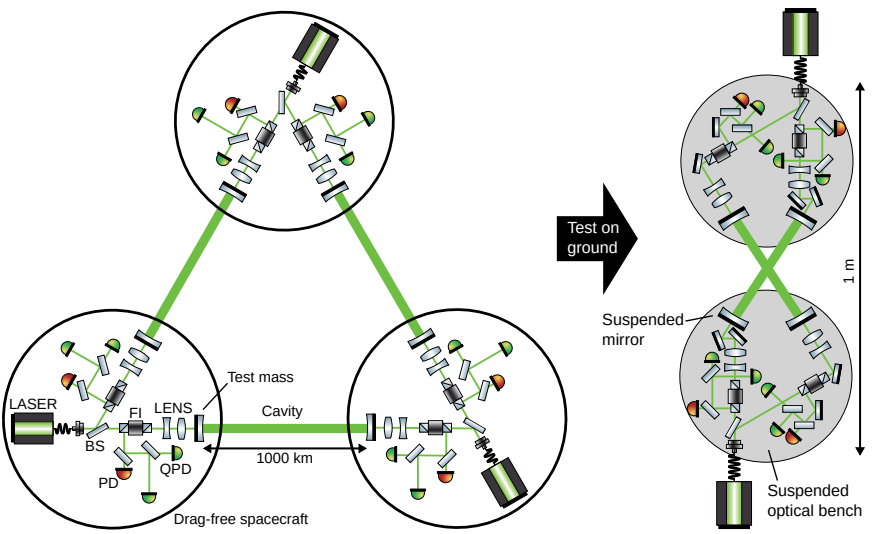



Fig. 1. Overview schematic of DECIGO and its the test bench on ground. BS, beam splitter; FI, Faraday isolator; PD, photodetector; QPD, quadrant photodetector.

### 2.2. *Optical design of the test bench*

The schematic of the optical design is shown in Fig. 2. The optical design of the test bench is made miniaturized and simplified compared with (B-)DECIGO design. The length of the cavity is 0.63 m and the radius of curvature of each mirror is 2 m. The cavity is critically coupled with finesse of 3100. The laser with wavelength of 1550 nm is introduced with fiber from the outside of the chamber. During the transfer, the laser is phase-modulated by electro-optic modulator for Pound-Drever-Hall technique and so-called wave front sensing technique for cavity length and alignment measurement.[6–8] The fiber output collimator is place on the suspended optical bench. The output light is splitted to two ways for two cavities. The light is isolated by Faraday isolators. The Faraday isolators can separate the reflected and transmitted light since the lights input from each side of the cavity have orthogonal linear polarization. The reflected light is received by radio-frequency photodetectors and quadrant photodetectors which are used to sensing cavity length and alignment measurement. The transmitted light is monitored with the photodetectors. In the test bench, the frequency and intensity stabilizing system is not implemented.

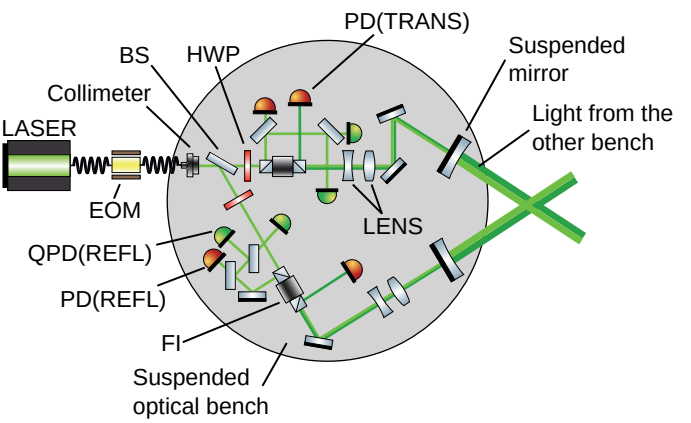



Fig. 2. Optical design of the test bench. EOM, electro-static modulator; HWP, half wave plate.

### 2.3. *Mechanical design of the test bench*

The schematic of the suspension system is shown in Fig. 3. In the test bench, there two types of suspension system: optical bench and cavity mirror suspension system. They are designed to have resonant frequency less than $\sim 1$ Hz for all considered mode.

The cavity mirrors are suspended with a double pendulum system. The first stainless beam is $10 \times 2 \times 2$ cm$^3$ and mass of 360 g. It is suspended by single tungsten wire with 10-cm long and 120-$\mu$m diameter. The mirror holder which contains the cavity mirror is suspended from one edge of the beam by two tungsten

wires with 40-$\mu$m diameter. The mirror holder is made of aluminum and its mass is about 220 g. The mirror holder has plus shape to increase moment of inertia for yaw and pitch motion. To compensate the mass of the mirror holder, a the counter weight is attached on the other edge of the beam.

The optical bench is suspended with a single piano wire whose diameter is 150 $\mu$m and length is 25 cm via a coil spring with spring constant of $\sim 10^3$ N/m. The optical bench is an aluminum plate with thick of 1.5 cm and diameter of 30 cm. The total mass of the optical bench including optic on it is estimated to be about 15 kg.

The estimated resonant frequencies of the mirror and the optical bench suspension for each impotant mode calculated with a toolkit of 3D rigid-body modeling coded in Mathematica® in Refs. 9, 10 are shown in Table 1. For the mirror and the optical bench suspension, the mode in terms of the center of the mirror and aluminum plate are considered, respectively. Table 1 shows that the requirement is almost achieved.

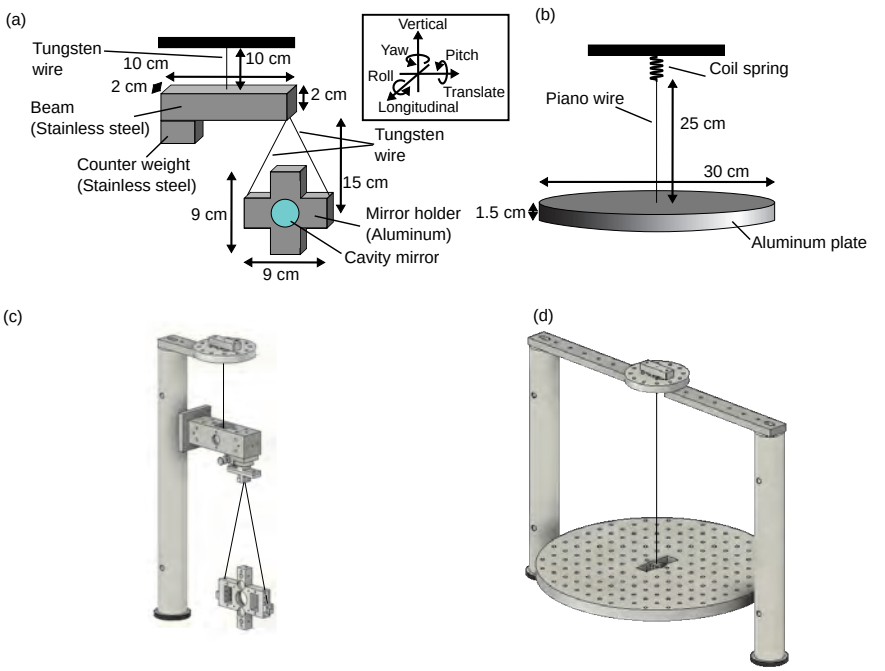



Fig. 3. Schematic (upper panel) and CAD image (lower panel) of the suspension system of the test bench. Left and right panels show the mirror and optical bench suspension system, respectively. In the CAD image of the optical bench (d), the coil spring is not shown. For the mirror and the optical bench suspension, the mode in terms of the center of the mirror and aluminum plate are considered, respectively.

Table 1. Resonant frequencies for each mode of two suspension types. Mode name: L, longitudinal; T, translation; V, vertical; Y, Yaw; P, pitch; R, roll. Since the roll mode of the mirror does not matter due to the symmetry of the mirror, its resonant frequency is shown just for reference.

| Suspension type | L | T | V | Y | P | R |
|---|---|---|---|---|---|---|
| Mirror | 0.02 Hz | 1 Hz | 0.5 Hz | 0.1 Hz | 0.05 Hz | (30 Hz) |
| Optical Bench | 1 Hz | 1 Hz | 1.4 Hz | 0.01 Hz | 0.09 Hz | 0.09 Hz |

### 2.4. *Local sensors and actuators*

To measure the relative motion and angle between the cavity mirrors and optical bench, reflective photosensors and optical levers are used. The motion and angle of the optical bench in the inertial system of a ground is also measured with photosensor and optical levers. As actuators, conventional coil-magnet actuators are used. They are used to apply force between mirror suspension and optical bench and apply force to the optical bench from the ground. The latter simulates thrusters of space craft, such as cold gas or field-emission electric propulsion thrusters. Note that these sensors and actuators are not ones which will be used in (B-)DECIGO. However, they can be replaced by the real ones for the integration test of the whole (B-)DECIGO system in future.

## 3. Summary and outlook

In this article, we reviewed the experimental design of the test bench for DECIGO and B-DECIGO. Optical and mechanical systems which is essential part of the test bench has been designed. Construction of the test bench is on going and will be followed by demonstration of the DPDFPI with drag-free control.

## Acknowledgments

This work is supported by JSPS Grant-in-Aid for Scientific Research (A) No. 15H02087. In addition, KN and YM are supported by the JSPS KAKENHI Grant No. JP17J01176, and JSPS Grant-in-Aid for Scientific Research (B) No. 18H01224. TS acknowledges financial support received from the Advanced Leading Graduate Course for Photon Science (ALPS) program at the University of Tokyo.

## References

1. S. Kawamura et al., *J. Phys.: Conf. Ser.* **120**, 032004 (2008).
2. J. Aasi et al., *Class. Quantum Grav.* **32**, 074001 (2015).
3. F. Acernese et al., *Class. Quantum Grav.* **32**, 024001 (2014).
4. K. Somiya, *Class. Quantum Grav.* **29**, 124007 (2012).
5. Y. Aso et al., *Phys. Rev. D* **88**, 043007 (2013).
6. R. W. P. Drever et al., *Appl. Phys. B* **31**, 97 (1983).
7. E. Morrison, B. J. Meers, D. I. Robertson, and H. Ward, *Applied Optics* **33**, 5037 (1994).

8. E. Morrison, B. J. Meers, D. I. Robertson, and H. Ward, *Applied Optics* **33**, 5041 (1994).
9. M. Barton, *LIGO Internal Document*, **LIGO-T020205**, "Models of the Advanced LIGO Suspensions in Mathematica" (2014).
10. T. Sekiguchi, *KAGRA Internal Document*, **JGW-T1503729**, "Suspension rigid-body modeling tool in Mathematica" (2015).

# Prospects for improving the sensitivity of KAGRA gravitational wave detector

Yuta Michimura,[a*] Masaki Ando,[a] Eleonora Capocasa,[b] Yutaro Enomoto,[a] Raffaele Flaminio,[b,c] Sadakazu Haino,[d] Kazuhiro Hayama,[e] Eiichi Hirose,[f] Yousuke Itoh,[g] Tomoya Kinugawa,[h] Kentro Komori,[a,i] Matteo Leonardi,[b] Norikatsu Mio,[j] Koji Nagano,[f] Hiroyuki Nakano,[k] Atsushi Nishizawa,[l] Norichika Sago,[m] Masaru Shibata,[n,o] Hisaaki Shinkai,[p] Kentaro Somiya,[q] Hiroki Takeda,[a] Takahiro Tanaka,[o,r] Satoshi Tanioka,[b,s] Li-Wei Wei,[t] Kazuhiro Yamamoto[u]

[a] *Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*
[b] *National Astronomical Observatory of Japan, Mitaka, Tokyo 181-8588, Japan*
[c] *Laboratoire d'Annecy de Physique des Particules (LAPP), Univ. Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy, France*
[d] *Institute of Physics, Academia Sinica, Nankang, Taipei 11529, Taiwan*
[e] *Department of Applied Physics, Fukuoka University, Nanakuma, Fukuoka 814-0180, Japan*
[f] *Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba 277-8582, Japan*
[g] *Department of Physics, Osaka City University, Sumiyoshi, Osaka 558-8585, Japan*
[h] *Department of Astronomy, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*
[i] *LIGO Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA*
[j] *Institute for Photon Science and Technology, University of Tokyo, Bunkyo, Tokyo 113-8656, Japan*
[k] *Faculty of Law, Ryukoku University, Fushimi, Kyoto 612-8577, Japan*
[l] *Research Center for the Early Universe (RESCEU), School of Science, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*
[m] *Faculty of Arts and Science, Kyushu University, Nishi, Fukuoka 819-0395, Japan*
[n] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Am Muhlenberg 1, Postdam-Golm 14476, Germany*
[o] *Center for Gravitational Physics, Yukawa Institute for Theoretical Physics, Kyoto University, Sakyo, Kyoto 606-8502, Japan*
[p] *Faculty of Information Science and Technology, Osaka Institute of Technology, Hirakata, Osaka 573-0196, Japan*
[q] *Department of Physics, Tokyo Institute of Technology, Meguro, Tokyo 152-8550, Japan*
[r] *Department of Physics, Kyoto University, Sakyo, Kyoto 606-8502, Japan*
[s] *The Graduate University for Advanced Studies (SOKENDAI), Mitaka, Tokyo 181-8588, Japan*
[t] *Max Planck Institute for Gravitational Physics (Albert Einstein Institute), Callinstraße, Hannover 30167, Germany*
[u] *Department of Physics, University of Toyama, Toyama, Toyama 930-8555, Japan*
[*] *E-mail: michimura@granite.phys.s.u-tokyo.ac.jp*

KAGRA is a new gravitational wave detector which aims to begin joint observation with Advanced LIGO and Advanced Virgo from late 2019. Here, we present KAGRA's possible upgrade plans to improve the sensitivity in the decade ahead. Unlike other state-of-the-art detectors, KAGRA requires different investigations for the upgrade since it is the only detector which employs cryogenic cooling of the test mass mirrors. In this paper, investigations on the upgrade plans which can be realized by changing the input laser power, increasing the mirror mass, and injecting frequency dependent squeezed vacuum are presented. We show how each upgrade affects to the detector frequency bands and also discuss impacts on gravitational-wave science. We then propose an effective progression of upgrades based on technical feasibility and scientific scenarios.

*Keywords*: Gravitational waves; Cryogenics; Underground; Laser interferometer; Optimization.

## 1. Introduction

The era of gravitational wave astronomy began with the first direct detections of gravitational waves from binary black holes and binary neutron star systems by Advanced LIGO and Advanced Virgo[1,2]. Improving the sensitivity of these detectors enables more frequent detections and more precise source parameter estimation. To this end, there have been extensive studies to improve the sensitivity beyond the detector's original design sensitivity.

Within LIGO Scientific Collaboration and Virgo Collaboration, there are ongoing effort to upgrade Advanced LIGO and Advanced Virgo detectors to $A+$[3] and $AdV+$[4], respectively, by around 2024[5]. The designed sensitivities of A+ and AdV+ are improved over that of Advanced LIGO and Advanced Virgo by roughly a factor of two. The improvement is in part realized by the coating thermal noise reduction either from the mechanical loss reduction of the coating material or from larger beam size. Also, broadband quantum noise reduction is expected by using a 300-m filter cavity to generate frequency dependent squeezed vacuum[6,7]. Twofold broadband sensitivity improvement leads to eightfold increase in the detection rate, and halves the parameter estimation error.

KAGRA is another laser interferometric gravitational wave detector which is being built in Japan[8,9] and plans to start observation jointly with Advanced LIGO and Advanced Virgo from late 2019. Compared with the other detectors, KAGRA has two technologically unique features: it is constructed at a seismically quiet underground site, and it uses sapphire mirrors at cryogenic temperatures to reduce thermal noise. Therefore, KAGRA has a unique potential to further improve its sensitivity, and upgrading KAGRA will require different approach compared with the other detectors.

In this paper, we discuss the prospects for the upgrade of the KAGRA detector. We start by describing possible technologies that can be applied for upgrading KAGRA and show that different technologies will improve the sensitivity in different frequency bands. We then discuss impacts on gravitational wave detections for each upgrade, and show possible strategy for the KAGRA upgrade in this decade.

## 2. Technologies for the KAGRA upgrade

The current design sensitivity of KAGRA is shown in Fig. 1. At low frequencies, the sensitivity is limited by the suspension thermal noise and the quantum radiation pressure noise. At high frequencies, the sensitivity is limited by the quantum shot noise. At the most sensitive band in the mid-frequencies, the sensitivity is limited by the mirror thermal noise, which manly comes from the coating Brownian noise. Thanks to cryogenic cooling of the sapphire test masses to 22 K, the mirror thermal noise is smaller than Advanced LIGO and Advanced Virgo although the size of the test mass is smaller. However, the suspension thermal noise is higher since the heat extraction is done by the sapphire fibers suspending the test mass and it requires thick and short fiber (1.6 mm diameter, 35 cm long) for efficient heat extraction. The

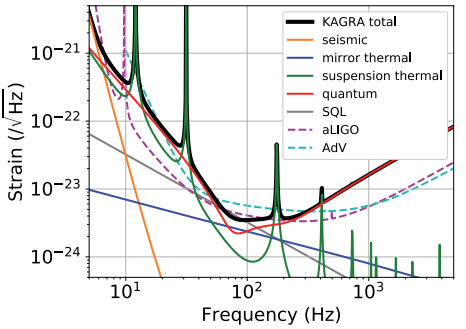



Fig. 1. The design sensitivity of KAGRA. The seismic noise shown includes the estimated Newtonian noise from the surface and bulk motion of the mountain containing KAGRA. The mirror thermal noise shown is the sum of the thermal noise from the test mass substrates and the coatings. Sensitivity curves for Advanced LIGO (aLIGO)[12] and Advanced Virgo (AdV)[5] are also shown for comparison.

quantum shot noise is also higher due to input laser power limitation for cryogenic cooling. Because of these features, KAGRA plans to use quantum non-demolition techniques such as the detuing of the signal recycling cavity and homodyne readout to reduce quantum noise in the most sensitive band at the cost of narrowing the detector bandwidth. Detailed discussion on the sensitivity optimization of KAGRA is given in Refs. 10, 11

To improve the sensitivity of KAGRA, retuning of laser power and suspension parameters will help at certain frequency bands. Increasing the mirror mass and injection of frequency dependent squeezed vacuum are also promising ways to improve the sensitivity. In the following subsections, we will discuss the effect of each technology for the upgrade of KAGRA. We will then discuss longer term prospects for the upgrade which can be realized by combining multiple technologies in this decade. Example sensitivity curves of KAGRA upgraded with different technologies discussed below are shown in Fig. 2 (Left). The interferometer parameters and the dimensions of the suspension fibers to calculate these sensitivity curves are optimized with particle swarm optimization method described in Ref. 11. The sensitivity curve data are available at Ref. 14.

### 2.1. *Laser power and heat extraction*

The input laser power and suspension thermal noise is closely related in KAGRA since heat extraction is done by the suspension fibers. To improve the sensitivity at low frequencies, reduction of suspension thermal noise is necessary. This can be done by changing the suspension fibers to thinner and longer ones since suspension thermal noise scales with $d_{\rm f}^2/l_{\rm f}$, where $d_{\rm f}$ and $l_{\rm f}$ are the diameter and the length of the fiber, respectively. However, this will result in larger shot noise because the heat extraction efficiency will be less and maximum input laser power allowed will be

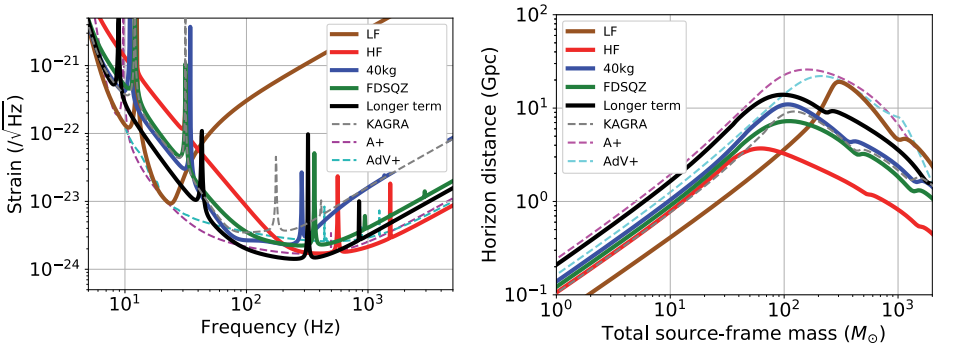



Fig. 2. (Left) Example sensitivity curves for the upgrade of KAGRA using different technologies. LF: Lower input power plan to focus on low frequency. HF: Higher power plan with frequency independent squeezed vacuum to focus on high frequency. 40kg: Sensitivity with increased mass of the test masses from 22.8 kg to 40 kg. FDSQZ: Sensitivity with the injection of frequency dependent squeezed vacuum generated with a 30-m filter cavity. Longer term: Example of longer term upgrade plan combining multiple technologies. Sensitivity curves for A+[13] and AdV+[5] are also shown for comparison. (Right) The horizon distance of example KAGRA upgrades for equal-mass, nonspinning binaries. The horizon distance shows the maximum distance at which gravitational waves can be detected with signal-to-noise ratio of more than 8.

less. Similarly, higher laser power to reduce shot noise at high frequencies require thicker and shorter suspension fibers, which will result in larger suspension thermal noise.

The *LF* curve shown in Fig. 2 is an example curve which the sensitivity at low frequencies is improved by lowering the laser power at the beam splitter from 673 W to 5 W. This plan requires higher detuning of the signal recycling cavity to reduce quantum noise at around 20-30 Hz. The suspension thermal noise peak at 31 Hz in the original KAGRA design sensitivity comes from the vertical motion of the intermediate mass suspension. Therefore, to remove this peak from the low frequency band, the *LF* plan also requires heavier intermediate mass with thinner and longer suspension wires. The interferometer parameters are optimized to maximize the inspiral range of 100 $M_\odot$-100 $M_\odot$ binary in the detector frame.

The *HF* curve shown in Fig. 2 on the other hand focuses on the high frequencies by increasing the laser power at the beam splitter to 3400 W. It also assumes the injection of frequency independent squeezed vacuum to further reduce the shot noise. Here, 6 dB of detected squeezing at high frequencies is assumed. The interferometer parameters are optimized to minimize the sky localization error of GW170814-like binary neutron stars[11].

### 2.2. *Increasing the mirror mass*

Increasing the mass of the test mass generally improves the sensitivity since the suspension thermal noise and quantum radiation pressure noise scales with $m^{-3/2}$ and $m^{-1}$, respectively. The coating thermal noise also can be reduced since larger

mirror allows larger beam size on the mirror. Assuming both the aspect ratio of the mirror and the ratio of the beam diameter to the mirror diameter to be the same, the coating thermal noise scales with $m^{-1/3}$.

The *40kg* curve shown in Fig. 2 is an example sensitivity with the mirror mass increased from 22.8 kg to 40 kg. Considering the design inside the current KAGRA cryostat, 40 kg would be the size limit without changing the cryostat drastically. The interferometer parameters are optimized to maximize the inspiral range of 1.4 $M_\odot$-1.4 $M_\odot$ binary. We note here that coating thermal noise reduction by larger beam size is assumed but smaller mechanical loss of the coating material is not assumed in the sensitivity calculation to show a feasible plan.

Interestingly, increasing the mirror mass result in the sensitivity improvement only at mid-frequencies where coating thermal noise dominates. This is because heavier mass requires higher laser power to keep the frequency $f_{\rm SQL}$ where quantum noise reaches the standard quantum limit to be the same. In case of KAGRA, $f_{\rm SQL}$ should be as high as possible until the quantum noise reaches the coating thermal noise, if we want to maximize the inspiral range. This is because the frequency dependence of the standard quantum limit ($f^{-1}$) is larger than that of the inspiral signal ($f^{-2/3}$). Therefore, the laser power scales with more than $m$. Higher laser power requires thicker suspension fibers and in the end the suspension thermal noise is not much dependent on the mirror mass.

### 2.3. *Frequency dependent squeezing*

Injection of frequency dependent squeezed vacuum is a promising way to reduce both radiation pressure noise and shot noise, which can be done without increasing the mirror mass or the laser power. The *FDSQZ* curve shown in Fig. 2 is an example curve which can be realized with 30-m filter cavity and 5 dB of detected squeezing at high frequencies. 30-m filter cavity can be constructed along the vacuum tubes of the signal recycling cavity. The interferometer parameters are optimized to maximize the inspiral range of 1.4 $M_\odot$-1.4 $M_\odot$ binary.

As discussed previously, the input laser power should be increased for higher $f_{\rm SQL}$ and this result in slightly worse suspension thermal noise. Also, injection of squeezed vacuum prefers no detuning of the signal recycling cavity. Therefore, injection of frequency dependent squeezed vacuum result in a sensitivity improvement at high frequencies.

### 2.4. *Longer term prospects*

As we have shown, applying only one of these technologies give sensitivity improvement at certain frequency bands. Combination of multiple technologies is necessary for broadband sensitivity improvement. The *Longer term* curve shown in Fig. 2 is an example sensitivity for 5 to 10-year upgrade plan which can be realized with 100 kg mirrors, 30-m filter cavity and 3500 W of the laser power at the beam splitter.

The interferometer parameters are optimized to maximize the inspiral range of 1.4 $M_\odot$-1.4 $M_\odot$ binary.

The situation is similar to *FDSQZ* plan, but because of larger test mass, suspension thermal noise and coating thermal noise are also reduced. In total, twofold broadband sensitivity improvement will be realized.

## 3. Science case study and disussion on strategic upgrade

Although combination of multiple upgrade components is necessary for the broadband sensitivity improvement, upgrades to the detector should be done in an incremental way. Which to be implemented at earlier stages depend on the technological feasibility and impact on gravitational-wave science.

Figure 2 (Right) shows the horizon distance of each example upgrade for compact binary coalescences. *LF* plan has the largest horizon distance above $\sim 200 M_\odot$ in total mass, whereas *40kg* plan has the largest horizon distance for smaller masses. We can say that *LF* has the highest probability of detecting the intermediate mass black holes (IMBHs).

Although the horizon distance is not great, *HF* plan gives the smallest sky localization error for binary neutron stars. The median of the sky localization error for GW170817-like binaries calculated with the same method described in Ref. 11 for *LF*, *HF*, *40kg* and *FDSQZ* are 0.507 deg$^2$, 0.105 deg$^2$, 0.156 deg$^2$ and 0.119 deg$^2$, respectively. For the sky localization of 30 $M_\odot$-30 $M_\odot$ binary black holes, *40kg* gives the smallest error. For constraining neutron star equation of state and for search for continuous waves from pulsars, *HF* and *FDSQZ* will be the best choices since the sensitivity from 500 Hz to 4 kHz is important for these studies. For the test of general relativity through inspiral-merger-ringdown waveform, broadband configuration such as *FDSQZ* and *40kg* would be preferred.

From the technical feasibility point of view, *LF* has the largest uncertainty since there are many kinds of low frequency excess noises other than the fundamental noises discussed above, such as scattered light noise, vibration noise from cryocoolers, interferometer controls noise etc. 40 kg test mass would be feasible in next few years, but even larger mirror is required for longer term upgrade. Considering that higher power laser source and squeezed vacuum source are required also for longer term upgrade, implementing these as a first step to focus on high frequency sensitivity improvement would be a strategy for the upgrade. *HF* plan is also attractive in that it might be able to do original science because *HF* has better sensitivity at high frequencies than A+ and AdV+.

## 4. Summary

Upgrading KAGRA requires simultaneous tuning of the parameters related to thermal noise and those related to quantum noise since the heat extraction is done through the fibers suspending the test mass mirrors. We showed that shifting the

detector frequency band of KAGRA is possible by changing the input laser power. We also showed that increasing the mirror mass and injection of frequency dependent squeezed vacuum will improve the sensitivity at mid-frequencies and high frequencies, respectively. Considering the technical feasibility and impact on the detection of gravitational waves, possible strategy for upgrading KAGRA would be to focus on high frequency improvement with higher laser power and squeezed vacuum injection for near term. In a longer term, broadband twofold improvement with frequency dependent squeezed vacuum injection and heavier mirror would be realized.

In this study, improvements in the coating, increased heat conductivity of the suspension sapphire fibers and reduced heat absorption of the sapphire mirror are not considered. More detailed investigations and other possibilities of the upgrade will be reported elsewhere.

## References


1. B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **116**, 061102 (2016).
2. B. P. Abbott *et al.* (LIGO Scientific Collaboration and Virgo Collaboration), *Phys. Rev. Lett.* **119**, 161101 (2017).
3. J. Miller *et al.*, *Phys. Rev. D* **91**, 062005 (2015).
4. J. Degallaix (the Virgo Collaboration), *Advanced Virgo+ preliminary studies*, Report No. VIR-0300A-18 (2018), `https://tds.virgo-gw.eu/?content=3&r=14287`.
5. B. P. Abbot *et al.* (KAGRA Collaboration, LIGO Scientific Collaboration and Virgo Collaboration), Living Rev. Relativity **21**, 3 (2018).
6. E. Oelker *et al.*, *Phys. Rev. Lett.* **116**, 041102 (2016).
7. E. Capocasa *et al.*, *Phys. Rev. D* **93**, 082004 (2016).
8. T. Akutsu *et al.* (KAGRA Collaboration), *Prog. Theor. Exp. Phys.* **2018**, 013F01 (2018).
9. T. Akutsu *et al.* (KAGRA Collaboration), arXiv:1901.03569.
10. K. Somiya (KAGRA Collaboration), *Class. Quantum Grav.* **29**, 124007 (2012).
11. Y. Michimura *et al.*, Phys. Rev. D 97, 122003 (2018).
12. L. Barsotti, S. Gras, M. Evans, P. Fritschel, *Updated Advanced LIGO sensitivity design curve*, Report No. LIGO-T1800044 (2018), `https://dcc.ligo.org/LIGO-T1800044/public`.
13. L. Barsotti, L. McCuller, M. Evans, P. Fritschel, *The A+ design curve*, Report No. LIGO-T1800042 (2018), `https://dcc.ligo.org/LIGO-T1800042/public`.
14. Y. Michimura, K. Komori, Y. Enomoto, K. Nagano, K. Somiya, *Example sensitivity curves for the KAGRA upgrade*, Report No. JGW-T1809537 (2018), `https://gwdoc.icrr.u-tokyo.ac.jp/cgi-bin/DocDB/ShowDocument?docid=9537`.

# Status of a cryogenic mirror suspension for KAGRA gravitational wave detector

Takafumi Ushiba* on behalf of the KAGRA Collaboration

*Institute for Cosmic Ray Research, The University of Tokyo*
*Kashiwa, 277-8582 Chiba, Japan*
**E-mail: ushiba@icrr.u-tokyo.ac.jp*

We demonstrated a cryogenic mirror suspension for a gravitational detector, KAGRA, and evaluated several performances such as cooling time and damping control. Cooling mirrors is promising technique for reducing the mirror thermal noise that fundamentally limits sensitivity of current ground-base gravitational-wave detectors around 100 Hz region. Cryogenic technology for mirror suspensions is therefore essential for further improvement of detector sensitivity beyond the 2nd generation detectors such as aLIGO, aVIRGO, and KAGRA. In this paper, current status of KAGRA cryopayloads is discussed.

*Keywords*: KAGRA, gravitational waves, laser interferometric gravitational-wave detector, cryogenics.

## 1. Introduction

The first direct detection of gravitational waves (GW) were achieved in 2015,[1] and more than 10 GWs are detected during LIGO-VIRGO observation run 1 and 2 by LIGO Scientific Collaboration and Virgo Collaboration.[2] These discoveries yield fruitful knowledge on our universe and open a new window of GW astrophysics and astronomy. For promoting them, it is crucial to construct a grobal network of GW detectors with sufficient high sensitivity such as KAGRA[3] and LIGO-India[4].

KAGRA is an interferometric GW detector located at Kamioka-mine in Japan. Two key features of KAGRA are using underground site for reducing seismic noise and cooling mirrors for mitigating thermal noise. KAGRA experienced two test runs called i-KAGRA operation in 2016 and bKAGRA phase1 operation in 2018, which are the first underground km-scale interferometer operation and km-scale cryogenic interferometer operation, respectively.[5,6]

It is a promising method to use cooled mirrors in order to reduce thermal noise, which fundamentally limits sensitivity of laser interferometric gravitational-wave detectors at room temperature such as aLIGO and aVIRGO. So, it is quite important to apply cryogenic technique to GW detectors for developing the further sensitive GW detectors like LIGO Voyager and Einstein Telescope.[7,8]

## 2. KAGRA cryogenic payload

### 2.1. *Overview*

Sapphire test masses of KAGRA are suspended by 13.5-meter 9-stage suspensions (Type-A suspensions), which consist of 5-stage room-temperature suspension (Type-A tower) and 4-stage cryogenic suspension (cryogenic payload).[6] Stages of

the cryogenic payload are platform, marionette, intermediate mass, and mirror from the top and there are corresponding recoil masses except for the platform. The intermediate mass has four sapphire blade springs and the mirror is suspended from them with four sapphire fibers. Since the sapphire fiber is too thick to make a loop, sapphire fibers are hooked on sapphire ears bonded on the side of the mirror.

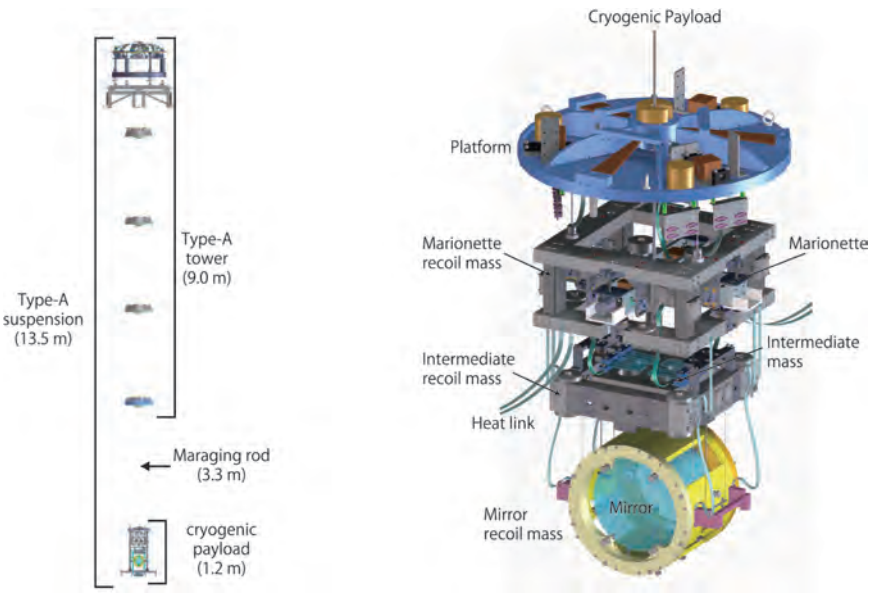



Fig. 1. Schematic view of the Type-A suspension and cryogenic payload.

To actuate angular and translational motions of mirrors, there are 6 coil-magnet actuators on both marionette stage and intermediate mass stage, and 4 on test mass stage[9]. To sense the local motion of the mirror, three kinds of local sensors are installed: optical levers(OLs) for mirror and marionette, length-sensing optical lever for mirror and reflective-type photosensors (PSs) for intermediate mass and marionette stages. These actuators and sensors are utilized for mirror alignment, local damping control, and interferometer operation.

### 2.2. *Sapphire mirror*

Hydroxide catalysis bonding (HCB) method is used for bonding the ear to the mirror because it is strong enough to support the mass at cryogenic temperature and not low-Q bonding to introduce additional thermal noise.[10,11] To mitigate a coupling between longitudinal and pitch motion, the position of the sapphire ears should be 48 mm below the center of gravity of sapphire mirrors.[12] It is also important to bond the ears to the center of the mirror along the optical axis to reduce coupling

between vertical and pitch motion. Coordinates of mirror shapes and ear position are shown in Fig. 2.

After bonding the sapphire ears to the mirrors, we measure the position of ears. Table 1 represents the measured result of the practical KAGRA mirrors.

After the confirmation of the ear positions, mirror stores for one month to promote chemical reaction of HCB. Then, tests of the ear strength are performed for one week. All the mirror for KAGRA were already passed these prcesses and deliverred at Kamioka-mine.

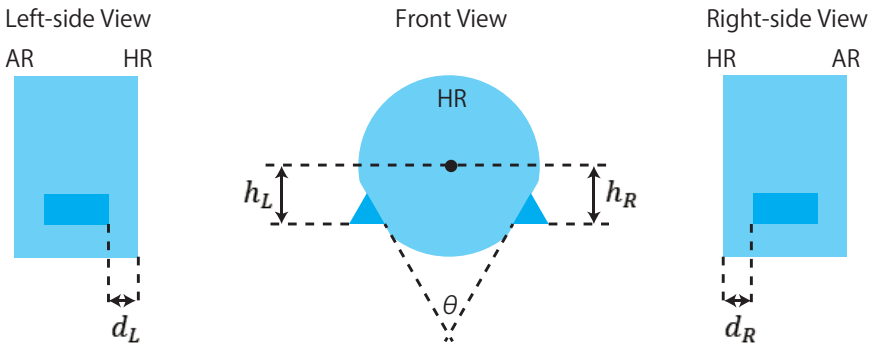



Fig. 2. Coordinates of mirror shapes and ear position.

Table 1. Summary of the mirror shapes and ear positions of KAGRA sapphire mirrors. Precision of the measurement is 50 $\mu$m for length and 0.1 degrees for angle.

| Mirror | Diameter (mm) | Thickness (mm) | $d_R$ (mm) | $d_L$ (mm) | $h_R$ (mm) | $h_L$ (mm) | $\theta$ (degrees) |
|---|---|---|---|---|---|---|---|
| ITMX | 220.34 | 150.12 | 35.03 | 35.02 | 48.27 | 48.38 | 36.0 |
| ITMY | 220.30 | 150.32 | 35.29 | 35.25 | 48.48 | 48.37 | 36.0 |
| ETMX | 220.28 | 149.96 | 35.09 | 35.05 | 48.03 | 48.08 | 36.0 |
| ETMY | 220.34 | 149.57 | 34.83 | 34.82 | 48.25 | 48.10 | 36.1 |

## 3. Installation of cryogenic payloads

Onece the mirrors are deliverred to the Kamioka-mine, they are suspended by cryogenic payloads and installed to KAGRA cryostats. Since the height of whole Type-A suspension is too large to install at onece, Type-A towers and cryogenic payloads are installed individually and connected with 3.3-meter maraging-steel rod after each installation (Fig. 3).

All the electrical cables on cryogenic payloads are connected from the top of Type-A tower and fixed along each stage of Type-A suspension to avoid vibration

injection through the cables. On the other hand, heat conductors for cooling is directly connected to a marionette recoil mass to cool the suspension effectively. Therefore, the material of the heat conductors is extremely pure aluminum that can balance high thermal conductivity with low stiffness to reduce vibration due to heat connection.

The first KAGRA cryogenic payload was suspended in November, 2017 and the last one was suspended in Feburary, 2019.

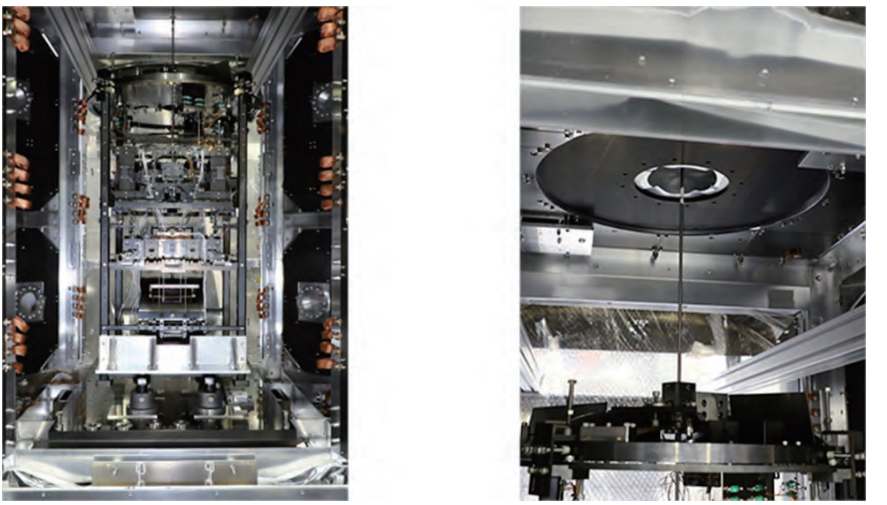


Fig. 3. Photos of a cryogenic payloads and its suspension rod from room temperature suspension. Left photo shows a cryigenic payloads installed inside a cryostat. Right photo shows a 3.3m maraging-steel rod.

## 4. Cooling

A cryogenic payload of the ETMY was cooled during the bKAGRA phase1 and the sapphire mirror reached 18 K. It takes almost one month to be a steady state, which is almost as we expected. Details are discussed in a paper.[6]

## 5. Angular motion

Mirror angular motion should be small sufficiently when we operate an interferometric gravitational wave detector. Active damping control is therefore implimented to the cryogenic payload during the bKAGRA phase1 operation. Figure 4 shows typical angular motion of the sapphire mirror with and without damping control of the cryogenic payload.

Yaw motion of the suspension could be reduced from the RMS of 4.4 $\mu$rad to that of 1.2 $\mu$rad. However, pitch motion of the suspension increased from the RMS

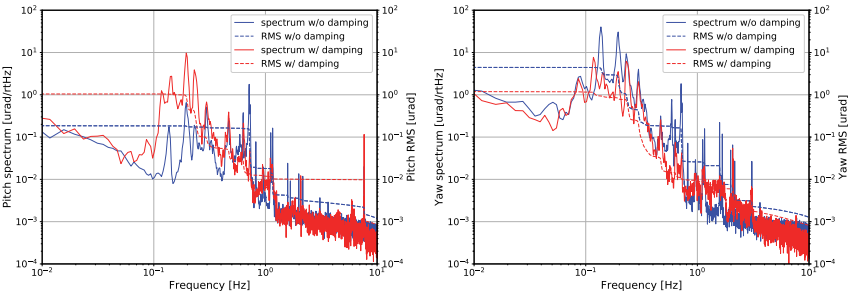



Fig. 4. Typical angular motion with and without damping control of cryogenic payload. Left figure shows pitch motion and right one shows yaw motion of the sapphire mirror. Red and blue lines in each graph show angular motion with and without damping, respectively. Dashed line represents RMS of mirror angular motion calculated by integration of the spectra from high frequency.

of 0.18 $\mu$rad to that of 1.0 $\mu$rad due to the actuator coupling between yaw and pitch motion. This large coupling is due to the rubbing of the electrical cables, such as actuators, thermometers, and sensors, between room temperature suspension and the platform.

Residual RMS of the angular motion is limited by several peaks around 0.2 Hz. These peaks are also due to the rubbing. This rubbing issue was solved after bKAGRA phase1 operation by tying cables to the suspension wires. So, our current suspension has no peak around there.

## 6. Summary and Future

All KAGRA sapphire mirrors were ready and already suspended by the Type-A suspension. Initial cryogenic suspension had some issures, especially on the angular motion of the mirror. However, these problems have been alredy solved and current cryogenic payload performance is quite stable.

All suspensions have been already installed, and all vacuum chambers were evacuated in April, 2019. At the begining of May, the final sapphire mirror started to be cooled, and the facility construction of KAGRA was completed. So, we are ready to start full-fledged commissioning. LIGO and VIRGO Observation run 3 has already started from April, 2019, and we are promoting interferometer comissioning to join the O3.

## Acknowledgments

This work was supported by MEXT, JSPS Leading-edge Research Infrastructure Program, JSPS Grant-in-Aid for Specially Promoted Research 26000005, JSPS Grant-in-Aid for Scientific Research on Innovative Areas 2905: JP17H06358,

JP17H06361 and JP17H06364, JSPS Core-to-Core Program A. Advanced Research Networks, JSPS Grant-in-Aid for Scientific Research (S) 17H06133, the joint research program of the Institute for Cosmic Ray Research, University of Tokyo, National Research Foundation (NRF) and Computing Infrastructure Project of KISTI-GSDC in Korea, the LIGO project, and the Virgo project.

## References

1. Abbott, B. P. *et al*, Phys. Rev. Lett. **116**, .061102 (2016)
2. The LIGO Scientific Collaboration, the Virgo Collaboration, arXiv:1811.12907 (2018)
3. Kentaro Somiya for KAGRA Collaboration, Class. Quantum Grav. **29**, 124007 (2012)
4. Unnikrishnan C. S., Int. J. Mod. Phys. D. **22**, 1341010 (2013)
5. T. Akutsu et al, Prog. Theor. Exp. Phys. **2018**, 013F01 (2018)
6. KAGRA Collaboration, arXiv:1901.03569 (2019)
7. R. Adhikari et al, LIGO Report No. T1400226 (2017)
8. M Punturo et al, Class. Quantum Grav. **27**, 194002 (2010)
9. Yuta Michimura *et al.*, 2017 Class. Quantum Grav. **34**, 225001
10. R Douglas et al, Class. Quantum Grav. 31 045001 (2014)
11. K. Haughian et al, Phys. Rev. D 94, 082003 (2016)
12. Kazuhiro Yamamoto, JGW-T1402773-v1 (internal document)

# Dynamical properties of binary stars hosting planets in the Galactic Center

Nazanin Davari* and Roberto Capuzzo-Dolcetta

*Department of Physics, University of Sapienza, Rome, Piazzale Aldo Moro 5/00185, Italy*
**E-mail: nazanin.davari@uniroma1.it*

We present some preliminary results of our work[1] about the close encounter of binary stars hosting planets on S-type orbits with the Sgr A* supermassive black hole in the center of our Galaxy.

*Keywords*: Galaxy: center; methods: numerical; planetary systems; stellar dynamics.

## 1. Introduction

Tidal breakup of binary stars passing close to a supermassive black hole (SMBH) in the center of galaxies may lead to the capture of one star around the SMBH (S-star) and the ejection of its companion as a hypervelocity star (HVS) (Hills mechanism).[2]

HVSs are fast-moving B-type main sequence stars observed in the Galactic halo at distances 50-120 kpc from the Galactic Center.[3] Some of them moving at extremely high velocities, i.e. they are not gravitationally bound to the Galaxy. The first observation of HVSs was in 2005,[17] a 3 $M_{\odot}$ main-sequence star, leaving the Galaxy with a heliocentric radial velocity of $853 \pm 12\ kms^{-1}$. HVSs that are originated from the Galactic Center can be used to constrain the Milky Way (MW) dark matter halo mass distribution since their orbits are completely determined by the MW potential.[6] A 3-4 $M_{\odot}$ star could be accelerated to such high velocities due to a close encounter with a relativistic gravitational potential.[4] So far, various mechanisms, as a result of close encounters with the relativistic potential well, have been proposed to study the ejection of HVSs. Yu & Tremaine[5] suggested a three-body interaction between a single star and a binary black hole (BBH). In this scenario, Sgr A* is assumed to be one component of a BBH. HVSs may also have been produced as a result of the interaction between stars within 0.1 pc of the Sgr A* black hole (BH) and a cluster of stellar-mass BHs that have segregated to that region.[7] Furthermore, the close interaction between a massive, orbitally decayed, globular cluster and the SMBH can give rise to ejection of some stars in the cluster as HVSs.[8]

Over the last decade monitoring the central arcsecond ($< 0.05$ pc) of the MW by two groups of researchers, [a] shows the presence of both population of early-type and late-type stars (about 40 bright stars),[14] the so-called S-stars[b] Unlike HVSs, S-stars are the most tightly bound stars which revolve around the SMBH residing in

[a] A group centered in the Max Planck Institute for Extraterrestrial Physics and a group at UCLA, California.
[b] "S" stands for "(infrared) source".[9]

the centre of the Galaxy (e.g. Schödel et al. 2002).[10] Refs. 11, 12, 13, 14 acquired orbital parameters of the S-stars using high-resolution near-infrared observations. Their results show that such stars would have semimajor axes in the range $\sim 0.005$ to $\sim 0.05$ pc, masses in the range 3-20 $M_{\odot}$. Similarly to HVSs, these are also classified as main-sequence stars, mostly of spectral type B. That is to say, they might be the former companion of the HVS in Hills mechanism.

On another side, it is known that stars borrow planets around them, also when they are in binaries. Therefore, it seems interesting to study the interaction of binary stars hosting planets with the SMBH in the MW center. In this work, we investigate the orbital properties of the plunging binary stars and their fate together with that of their planets after close interactions with the SMBH in the Galactic Center (GC).

## 2. Computational Method

We performed a huge set of simulations using a regularized $N$-body algorithm, the `AR-CHAIN` integrator,[15] which includes post-Newtonian corrections up to order 3.5 and properly modified (`ARGdf` code)[19] to account for an analytic external potential and its dynamical friction. Assuming spherical symmetry for the inner galactic region, we considered as mass distribution model the sum of a Dehnen's and a Plummer's distribution.

For our simulations, we considered both the case of a non-spinning (Schwarzschild) and a spinning (Kerr) with the dimensionless spin parameter $\chi = 0.76$ SMBH with mass $M_{\bullet} = 4 \times 10^6 M_{\odot}$ initially placed in the origin of the reference frame. We assumed binary stars of 3 $M_{\odot}$ each revolving around each other on initial circular orbit with (initial) separation in the range $a_* = 0.1 - 0.5$ AU.[16] The mass value chosen is comparable to the mass of the first HVS observed.[17] Each star has one planet initially at $a_p = 0.02$ AU from its host star with mass $m_p = 10^{-3} M\odot$ (Jupiter-like planet). The center of mass of the binary star is assumed initially at 2000 AU away from the SMBH. We give the system a transverse (respect to the line joining the binary and the SMBH) initial velocity of 66.5 km $s^{-1}$. We run simulations at varying the inclination of the binary orbital plane respect to that of the motion of its center of mass; 0°, 90° and 180°. We vary also the initial orbital phase angle values of the binary ($\phi$) in the $0° - 360°$ range at steps of 15°.[16]

Table 1. Set of initial conditions for our runs.

| inclination (degrees) | $a_\star$ (AU) | $a_P$ (AU) | $m_\star$ ($M_\odot$) | $m_P$ ($M_\odot$) | $\phi$ (degrees) |
|---|---|---|---|---|---|
| 0 | 0.1 – 0.5 | 0.02 | 3 | 0.001 | 0-360 |
| 90 | 0.1 – 0.5 | 0.02 | 3 | 0.001 | 0-360 |
| 180 | 0.1 – 0.5 | 0.02 | 3 | 0.001 | 0-360 |

## 3. Preliminary Results

When a binary star hosting planets, with our given parameters, is disrupted by the SMBH, stars and planets may eject as hypervelocity stars or hypervelocity planets (HVPs). In some cases, stars and/or planets may remain bound to the SMBH in highly eccentric orbits or are swallowed by it. In Figures 1 and 2 we sketch two possible examples for the fate of binary+planets system after a close interaction with a spinning SMBH.

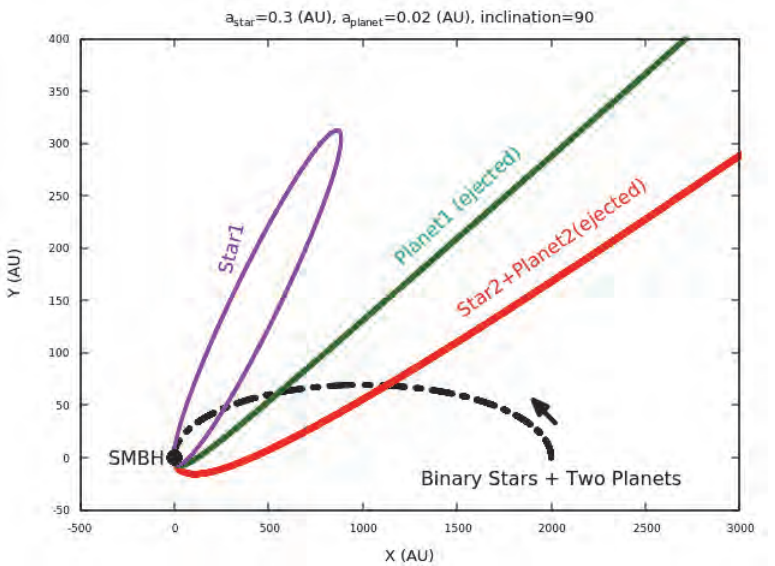



Fig. 1. HVS with planet.

The main aim of this work is to check conditions upon which binary stars borrowing planets around them can give rise to HVSs still keeping bound their planets even after the close interaction with the SMBH in the GC.

In Fig. 1, the initial semi-major axis of the binary star motion is 0.3 AU and the orbital plane inclination is 90°. After the first encounter, there is a break-up of the binary. One star is ejected keeping its planet bound (red line) and the former companion star starts revolving around the SMBH (magenta line) while its planet is ejected as HVP (green line). Fig. 2 refers to an initial separation of 0.1 AU for the two stars in the binary; again, the inclination of the plane is 90°. In this case, after the first encounter with the SMBH, the stars swap their planets, and one of the stars escapes the GC as HVS with the planet orbiting around it (green line). The companion star loses the planet after three passages around the SMBH (red line shows the three revolutions around the SMBH before they get separated); its planet ejects as HVP (blue line).

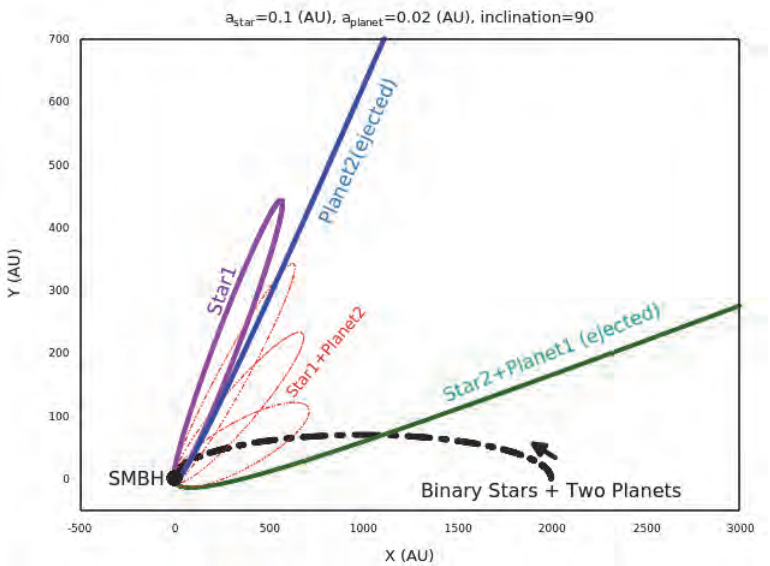



Fig. 2. Exchange of planets.

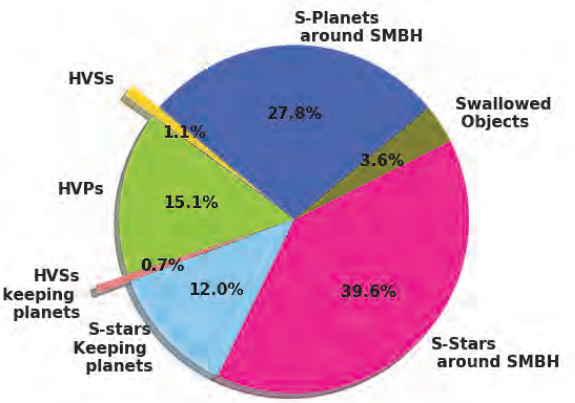



Fig. 3. Inclination 0°.

The following pie charts (Figs. 3 and 4) quantify the likelihood of different outcomes for the co-rotating and counter-rotating cases, after the binary star is broken up by the SMBH in our whole set of simulations. An intriguing aspect to examine, deepening the work by Fragione & Ginsburg,[18] will be checking the detection chance of such HVPs around their hosting stars, with transit techniques.

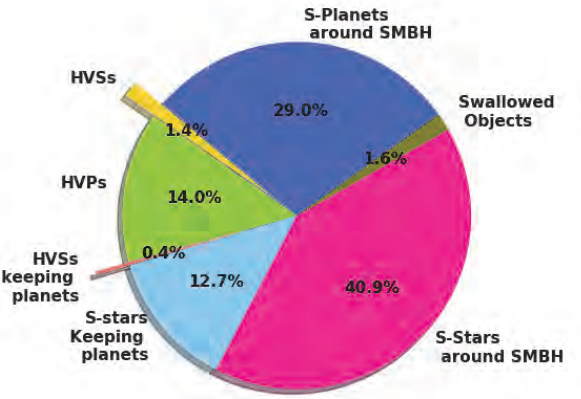



Fig. 4. Inclination 180°.

## 4. Forthcoming Research

An immediate and natural future development of this work is enlarging the study to binary and triple systems hosting multi-planet systems. This would allow an estimate of the number of high-velocity wandering planets, other than providing an estimate of the background contribution of multiple small bumps of gravitational wave emission.

## References

1. R. Capuzzo-Dolcetta, N. Davari, in preparation (2019).
2. J. G. Hills, Nature, 331, 687 (1988).
3. W. R. Brown, M. J. Geller & S. J. Kenyon, ApJ, 787, 89 (2014).
4. A. Sesana, F. Haardt & P. Madau, MNRAS, 379, L45 (2007).
5. Q. Yu & S. Tremaine, ApJ, 599, 1129 (2003).
6. G. Fragione & A. Loeb, NewA, 55, 32 (2017).
7. R. M. O'Leary & A. Loeb, MNRAS, 383, 86 (2008).
8. R. Capuzzo-Dolcetta & G. Fragione, MNRAS, 454, 2677 (2015).
9. A. Eckart & R. Genzel, Nature, 383, 415 (1996).
10. R. Schödel, T. Ott, R. Genzel, et al. Nature, 419, 694 (2002).
11. A. M. Ghez, G. Duchêne, K. Matthews, et al. ApJL, 586, L127 (2003).
12. F. Eisenhauer, R. Genzel, T. Alexander, et al. ApJ, 628, 246 (2005).
13. S. Gillessen, F. Eisenhauer, S. Trippe, et al. ApJ, 692, 1075 (2009).
14. S. Gillessen, P. M. Plewa, F. Eisenhauer, et al. ApJ, 837, 30 (2017).
15. S. Mikkola & D. Merritt, AJ , 135, 2398 (2008).

16. I. Ginsburg & A. Loeb, MNRAS, 368, 221G (2006).
17. W. R. Brown, M. J. Geller, S. J. Kenyon & M. J. Kurtz, ApJL, 622, L33 (2005).
18. G. Fragione & I. Ginsburg, MNRAS, 466, 1805 (2017).
19. M. Arca-Sedda & R. Capuzzo-Dolcetta, MNRAS, 483, 152 (2019).

# MOCCA survey database I. BHs in star clusters

Mirek Giersz*

*Nicolaus Copernicus Astronomical Center,*
*Polish Academy of Sciences, Warsaw, Poland*
** E-mail: mig@camk.edu.pl*

Abbas Askar

*Nicolaus Copernicus Astronomical Center,*
*Polish Academy of Sciences, Warsaw, Poland*
*and*
*Lund Observatory, Department of Astronomy, and Theoretical Physics,*
*Lund University, Lund, Sweden*

Jakub Klencki

*Institute of Mathematics, Astrophysics and Particle Physics,*
*Radboud University Nijmegen, Nijmegen, The Netherlands*

Jakub Morawski

*Astronomical Observatory,*
*Warsaw University, Warsaw, Poland*

We briefly describe and discuss the set-up of the project Mocca Survey Database I. The database contains more than 2000 Monte Carlo models of evolution of real star cluster performed with the Mocca code. Then, we very briefly discuss results of analysis of the database regarding the following projects: formation of intermediate mass black holes, abrupt cluster dissolution harboring black hole subsystems, retention fraction of black hole - black hole mergers, and tidal disruption events with intermediate mass black holes.

*Keywords*: Methods: numerical - globular clusters: general - stars: black holes.

## 1. Introduction

Recent high resolution observations of globular clusters (GC) provide a very detailed picture of their physical status and show complex phenomena connected with multiple stellar populations, binary evolution, black holes (BH) and the Galactic tidal field. Despite such great observational progress there are many theoretical uncertainties connected with the origins of GCs and their primordial properties. To bridge the gap between present-day observed star cluster properties and their properties at the time of cluster formation, we need to discriminate between different theories and models by means of numerical simulations of GC evolution. The best suited codes for such a task are N-body and Monte Carlo codes. In this paper we will describe results of Monte Carlo simulations done with the MOnte Carlo Cluster simulAtor - Mocca.[11,17] (and references therein) All those simulations were collected in a Mocca Survey Database I, which then was analyzed from the point of view of properties and evolution of different kinds of BH populations.

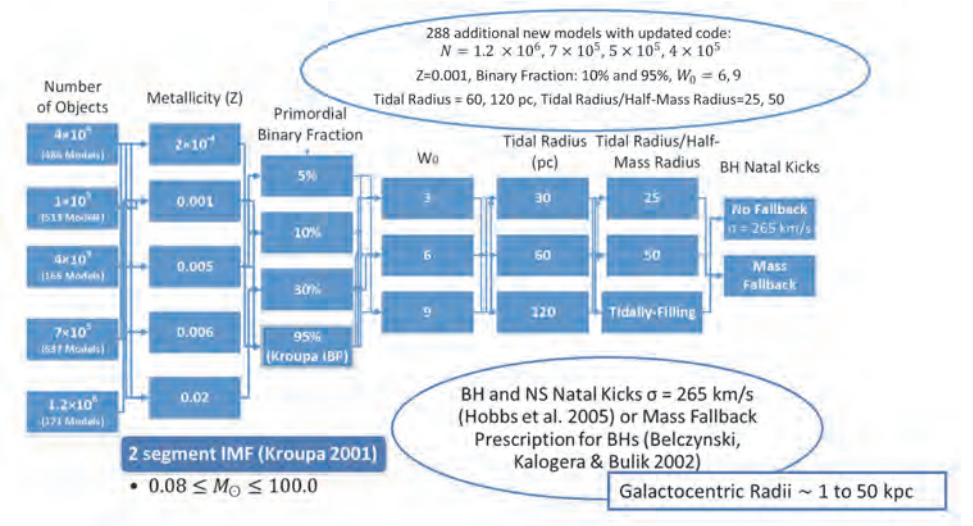



Fig. 1. The initial set-up of the MOCCA simulations of the real star clusters stored in the MOCCA-SURVEY DATABASE I. The model parameters and relevant references are listed in the Figure. In the models the pair-instability supernovae, pulsation pair-instability supernovae, electron capture supernovae and accretion induced supernovae were not taken into account. BHs formed in the simulations have masses smaller than about $15 - 20\ M_\odot$, depending on metallicity.

## 2. MOCCA and MOCCA Survey Database I

The MOCCA code is a version of the Monte Carlo codes and can be considered as so-called 'kitchen sink' code, which is able to follow most physical processes important during star cluster dynamical evolution. The MOCCA code treats the relaxation process using the method described by Henon,[13] that was significantly improved by Stodółkiewicz,[25,26] and more recently by Giersz and his collaborators.[10–12,17] (and reference therein). For stellar and binary evolution Jarrod Hurley's BSE code is used,[15,16] and for the scattering experiment John Fregeau's Fewbody code.[7] The realistic description of an escape process in a tidally limited cluster is done on the basis of the Fukushige & Heggie theory.[8] The MOCCA code provides as many details as N-BODY codes. It can follow evolution and movement of particular objects. The MOCCA code is extremely fast. It needs about a day to complete an evolution of a real size globular cluster. So, instead of just one N-body model, hundreds or thousands models can be computed with different initial conditions. The MOCCA code is ideal either for dynamical models of a particular cluster or for large surveys.

The MOCCA SURVEY DATABASE I[2] contains about 2000 models of GCs with different initial masses, structural and orbital parameters. The brief description can be found on Fig. 1 and more details is given in Askar et al., Table I.[2]

As it was pointed out by Askar et al.,[2] it can be assumed that the MOCCA SURVEY DATABASE I cluster models are more or less representative of the Milky Way GC population.

## 3. Intermediate Mass BH Formation in GCs

In the literature, so far, there were four possible groups of scenarios proposed for intermediate mass BH (IMBH) formation in GCs: Direct collapse of very massive Population III stars proposed by Madau and Rees,[20] runaway merging of very massive MS stars in dense young star clusters, first discussed by Portegies Zwart et al.,[23] accretion of the residual gas on stellar mass BHs formed from the first generation stars recently proposed by Leigh et al.[19] and the scenario based on results of Mocca simulations of evolution of dense stellar systems proposed by Giersz et al.[12] In this scenario an IMBH is formed because of buildup of BH mass solely due to mergers in dynamical interactions and mass transfers in binaries. This scenario is a much expanded and refined scenarios proposed ealier by Miller & Hamilton[21] and by Leigh et al.[19]

In the Mocca Survey Database I there were about 460 IMBHs formed out of 2000 models, and they formed even in very low N models, as small as consisting of 40000 objects. As it is described in detail in Ref. 12 there are generally two regimes of IMBH formation: 1) very fast an IMBH mass buildup (FAST) starting from the very beginning of the cluster evolution, which requires very large initial central densities. In a time of about one Gyr, an IMBH mass grows to up a few tens of thousands $M_{\odot}$, and 2) slow IMBH mass buildup (SLOW) starting later on in the cluster evolution, usually, around the core collapse. It requires rather moderate initial central densities, and masses of IMBHs are quite moderate, from a few hundred up to a few thousand $M_{\odot}$. The process of IMBH formation is highly stochastic. The larger the initial cluster concentration, the earlier, faster and with higher probability an IMBH will form.

Here is the detailed, slightly updated comparable to Ref. 12, description of the FAST and SLOW scenarios:

(1) SLOW and FAST formation scenarios;

   (a) SLOW scenario - either a single BH is left after the early phase of SN explosions (SNe), or a single BH is formed via mergers or collisions during dynamical interactions, usually around the core collapse time. The central density has to be greater than about $10^5\ M_{\odot}/pc^3$;
   (b) FAST scenario - several dozen/hundreds BHs remain in the system after the early phase of SNe, and form a dense central subsystem. The central density must be extremely high, greater than $10^8\ M_{\odot}/pc^3$, for an IMBH to form. Alternatively, all BHs are quickly and efficiently removed from the system via dynamical interactions. If at least one remains, then the SLOW scenario is followed;

(2) Initial mass buildup of IMBH progenitors:

   (a) If the cluster density is large enough, the collisions between main sequence (MS) stars lead to formation of very massive MS stars, hundreds of $M_{\odot}$. If

such a star collides with a BH then a very massive BH (already an IMBH) is formed.
   (b) If cluster density it is not large enough MS does not collide efficiently and stellar mass BH will be formed, because of stellar evolution and strong stellar winds.

(3) BHs are the most massive objects in GC, so they quickly form a binary via a three-body interaction;
(4) Further BH binary evolution is due to dynamical interactions with other binaries and stars, or because of gravitational wave radiation (GW);

   (a) orbit tightening leading to mass transfer from MS/Red Giant/Asymptotic Giant Branch companions;
   (b) exchanges and collisions, leaving the binary intact;
   (c) total collisions during dynamical interactions or GW mergers - in this case, the binary is destroyed and only a BH is left;

(5) Newly created single BH quickly forms a new binary via another three-body interaction, which is further free to undergo subsequent dynamical interactions with other single and binary stars, and the process repeats. In this way, the BH mass steadily increases.

Total collisions are the most important ingredient of an IMBH formation in the SLOW scenario. In this way BH mass can steadily increase and BH binaries are not kicked out from the system due to dynamical three- or four-body interactions. That is the reason why IMBH formation is so stochastic in the SLOW scenario.

As it can be seen in Fig 2, there are clearly visible two channels of dynamical interactions leading to formation IMBH seeds: runaway MS star mergers and BH formation because of collisions with a BH, and runaway mergers and BH formation because of pure stellar evolution. The larger the cluster density the larger the maximum mass of runaway merged MS star and consequently the larger the BH mass (collision product of stellar mass BH and very massive MS star). Less dense models form lower mass BHs, which later in the course of evolution will substantially grow. Very massive BHs are preferentially formed for very dense models at the very beginning of the cluster evolution. We should here caution readers, the above picture relies on very uncertain properties and evolution of extremely massive MS stars formed in runaway collisions and on the amount of accreted mass onto stellar mass BH during collision with extremely massive MS star. Such collision is more similar to the common envelope phase than to direct collision of similar mass BHs and a MS stars, in which probably only small amount of mass can be accreted onto BH. We checked, that in the case of only 10% of accreted mass an IMBH is still formed, so the FAST scenario seems to work in real physical systems.

We would like to stress that the presented scenarios for IMBH formation in GCs, in particular the SLOW scenario, do not require any specific conditions, unlike

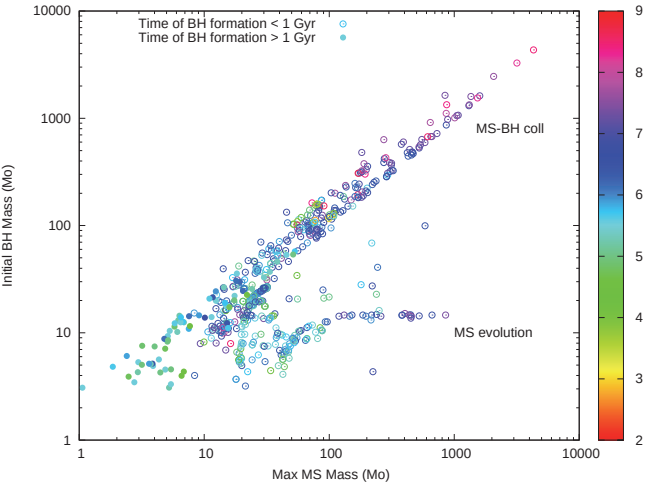



Fig. 2. Mass of just formed BH as a function of mass of MS star just before a supernova explosion. Open circles - formation of BH before 1 Gyr. Filled circles - formation of BH after 1 Gyr. The color bar on the right side shows the initial cluster central densities in $M_{\odot}/pc^3$.

other scenarios proposed in the literature. IMBH formation occurs solely via binary dynamical interactions and mass transfer in binaries. The FAST scenario is more probable to occur in galactic nuclei, or in extremely dense star clusters.

## 4. BH-BH Merger Gravitational Wave Radiation Kicks

We used Mocca Survey Database I to estimate GW kick retention fraction of BH-BH merger products. We found about 4500 such mergers. The amplitude of the GW recoil kick velocity depends on the spin magnitudes, degree of misalignment between BH spins and the binary angular momentum and on the BH mass ratio. The kick velocities were calculated according to Ref. 3 and final spins according to Ref. 24. We checked many different assumptions about the BH spins: random, constant equal to 0.5, or a function of metallicity and initial stellar mass.[6] There are two classes of mergers: mergers in 'primordial' binaries (binaries which keep the same stars during the whole evolution), which have only small spin misalignments, and 'dynamical' mergers in binaries which were involved in strong interactions, e.g. exchanges or disruptions, for which spin directions were distributed randomly. The evolution of each BH was tracked. Each product of BH-BH merger, also a BH, was assigned with a recoil kick velocity. Merger products with velocities greater than the cluster escape velocity were removed from any further interactions. Those merger products that remained in the cluster were assigned with a modified spin value, computed according to Ref. 24, and could contribute to next generations of

BH-BH mergers. 'Dynamical' mergers were mainly mergers with IMBHs and have small mass ratio, so nearly 0.7 of them were retained in the system. For 'primordial' binaries the retention fraction seems to saturate at about 0.2. This is connected with the fact that most 'primordial' BH-BH binaries consist of low mass BHs, which according to Belczynski et al.[6] have large spins. On average, we should expect that about 0.3 mergers will be retained in the system. This result does not strongly depend on the assumed initial BH spin distributions.

The retention fraction is a strong function of time. In the case of 'dynamical' mergers, it is changing from the initial values of 0.5-0.6, when the IMBH mass is still relatively small, and steadily increasing up to 0.9 in the later stages of the cluster evolution, when the mass ratios in IMBH-BH mergers drop down to 0.01 - 0.001 values. The retention fraction for 'primordial' binaries substantially increases after about 1 Gyr due to decrease of the mass ratio. It seems that first binaries with mass ratio close to 1 are merged and later on binaries with smaller and smaller mass ratios start to merge. The retention fraction for models with BH subsystems is very small and equal to about 0.1. This is because such models have relatively low concentration and low escape velocity. There is a clear correlation, the larger the cluster concentration the larger the retention fraction. Interested readers, we refer to the paper by Morawski et al.[22] where they can find all details connected with this work.

We would like to stress that our results confirm analytic and semi-analytic results obtained earlier by many authors, but for the first time real simulation data was used to obtain the BH merger retention fraction and its dependence on the different cluster evolutionary scenarios and global cluster parameters. This approach is not fully self consistent, but it is the first step to fully integrate BH merger kicks in dynamical evolution of GCs.

## 5. Tidal Disruption Events with IMBHs

Formation of IMBHs in the MOCCA models and their subsequent mass buildup have to be connected with tidal disruption of star intruders, so called tidal disruption events (TDEs). We analyzed the Mocca Survey Database I looking for disruptions of white dwarfs (WD), MS stars or other luminous stars. We found 344755 WD-IMBH type events, 750753 MS-IMBH type events and 42934 other-IMBH type events. The work is still in progress and we would like only to summarize briefly a very preliminary results. Most TDEs are formed in massive GCs with relatively small galactocentric distances. To form an IMBH in the FAST scenario a very large central density is needed. Such density can be achieved easily in massive GCs with small tidal radii (small galactocentric distances). Due to dynamical friction, some of such GCs will migrate towards to galactic center and merge with the nuclear cluster (NC). From that point, our models of GC evolution are no longer applicable. Interestingly, any TDE events associated with IMBHs hosted by those

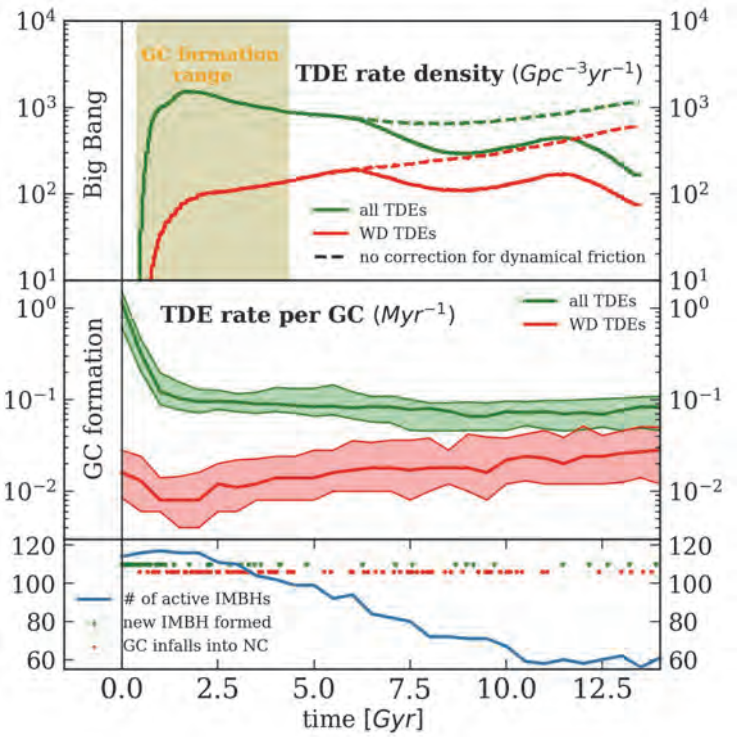



Fig. 3. The TDE rates as a function of the GC age. Red lines - rates for WDs. Green lines - rates for WDs and MS stars. Dashed line - for all models, solid line - only for models which do not migrate to the galactic center. Top panel - TDE density rate ($Gpc^{-3}yr^{-1}$). Middle panel - TDE rate per GC ($Myr^{-1}$). Bottom panel - number of active GCs. Red dots - time when GC merge with nuclear star cluster, green dots - time when new IMBH is formed.

GCs would be observable in the galactic nuclei. In any case, in order to calculate the IMBH TDE rate in actual GCs we exclude from our computations those GC models that have fallen into the NC. In the Fig. 3 we show the TDE rates computed assuming that all TDEs outside the galactic center can be observed. Present day TDE rate density is smaller by factor of about 3–4 when the evolution of GCs in the galactic environment (in a Milky-Way type galaxy) is taken into account. Also the number of "active" GCs in which TDEs are happening is much smaller now than was, showing strong influence of galactic environment on the observed TDEs. Interestingly, IMBHs formed in the SLOW scenario are responsible for only a small fraction of possible TDEs. The obtained TDE rates are comparable to other theoretical estimates.

In the future work we are planning to populate galaxies in the local universe with MOCCA GCs models and estimate the local TDE rates.

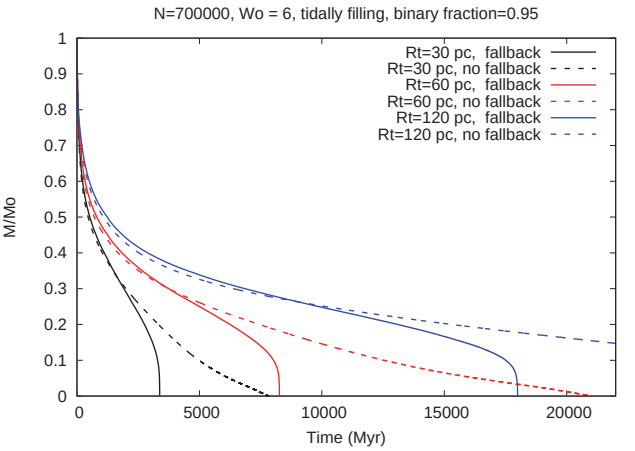



Fig. 4. Evolution of the fraction of cluster bound mass as a function of time for tidally filling Mocca cluster models with 700000 objects (stars and binaries), $W_0 = 6$ and binary fraction equal to 0.95, for different tidal radii and SNe natal kicks mass fallback set to ON or OFF.

## 6. BH Subsystem and a Third Cluster Dissolution Mechanism

We used the Mocca Survey Database I to investigate the dissolution process for dynamically evolving star clusters embedded in an external tidal field, with focus on the presence and evolution of a stellar-mass BH subsystem. We argue that the presence of a BH subsystem can lead to the dissolution of tidally filling star clusters and this can be regarded as a third type of cluster dissolution mechanism, in addition to well known mechanisms connected with strong mass loss due to stellar evolution and mass loss connected with the relaxation process. As it can be seen in Fig. 4 for models with mass fallback ON (high BH retention fraction), the third process is characterized by abrupt cluster dissolution connected with the loss of dynamical equilibrium. The abrupt dissolution is powered by the strong energy generation from a massive stellar-mass BH subsystem accompanied by tidal stripping. We argue that such a mechanism is universal and should also work for tidally under-filling clusters with top-heavy IMF. Observationally, star clusters which undergo dissolution powered by the third mechanism would look as 'dark clusters' i.e. composed of stellar mass BH surrounded by expanding halo of luminous stars,[4] and they should be different from 'dark clusters' harbouring an IMBH as discussed by Ref. 1. An additional observational consequence of an operation of the third dissolution mechanism should be larger than expected abundance of free floating BHs in the Galactic halo. Interested readers, we refer to the paper by Giersz et al.[9] where they can find all details connected with this work.

**Acknowledgments**

MG was partially supported by the Polish National Science Center (NCN) through the grant UMO-2016/23/B/ST9/02732. AA is currently supported by the Carl Tryggers Foundation for Scientific Research through the grant CTS 17:113 and was partially supported by NCN, Poland, through the grants UMO-2016/23/B/ST9/02732. AH was supported by Polish National Science Center grant 2016/20/S/ST9/00162. This work benefited from support by the International Space Science Institute (ISSI), Bern, Switzerland, through its International Team programme ref. no. 393 The Evolution of Rich Stellar Populations & BH Binaries (2017-18).

**References**

1. A. Askar, P. Bianchini, R. de Vita, M. Giersz, A. Hypki, and S. Kamann. Mocca-survey database i: Is ngc 6535 a dark star cluster harbouring an imbh? *MNRAS*, 464:3090–3100, 2017.
2. A. Askar, M. Szkudlarek, D. Gondek-Rosińska, M. Giersz, and T. Bulik. Mocca-survey database - i. coalescing binary black holes originating from globular clusters. *MNRAS*, 464:L36–L40, 2017.
3. J. G. Baker, W. D. Boggs, J. Centrella, B. J. Kelly, S. T. McWilliams, M. C. Miller, and J. R. van Meter. Modeling kicks from the merger of generic black hole binaries. *ApJ*, 682:L29, 2008.
4. S. Banerjee and P. Kroupa. A new type of compact stellar population: Dark star clusters. *ApJ*, 741:L12, 2011.
5. K. Belczynski, V. Kalogera, and T. Bulik. A comprehensive study of binary compact objects as gravitational wave sources: Evolutionary channels, rates, and physical properties. *ApJ*, 572:407–431, 2002.
6. K. Belczynski, J. Klencki, G. Meynet, C. L. Fryer, D. A. Brown, M. Chruslinska, W. Gladysz, R. O'Shaughnessy, T. Bulik, E. Berti, D. E. Holz, D. Gerosa, M. Giersz, S. Ekstrom, C. Georgy, A. Askar, D. Wysocki, and J.-P. Lasota. The origin of low spin of black holes in ligo/virgo mergers. *arXiv-1706.07053*, 2017.
7. J. M. Fregeau, P. Cheung, S. F. Portegies Zwart, and F. A. Rasio. Stellar collisions during binary-binary and binary-single star interactions. *MNRAS*, 352:1–19, 2004.
8. T. Fukushige and D. C Heggie. The time-scale of escape from star clusters. *MNRAS*, 318:753–761, 2000.
9. M. Giersz, A. Askar, L. Wang, A. Hypki, A. Leveque, and R. Spurzem. Mocca-survey database i: Dissolution of tidally filling star clusters harbouring bh subsystems. *MNRAS*, submitted, 2019.
10. M. Giersz, D. C. Heggie, and J. R. Hurley. Monte carlo simulations of star clusters - iv. calibration of the monte carlo code and comparison with observations for the open cluster m67. *MNRAS*, 388:429–443, 2008.

11. M. Giersz, D. C. Heggie, J. R. Hurley, and A. Hypki. Mocca code for star cluster simulations - ii. comparison with n-body simulations. *MNRAS*, 431:2184–2199, 2013.
12. M. Giersz, N. Leigh, A. Hypki, N. Lützgendorf, and A. Askar. Mocca code for star cluster simulations - iv. a new scenario for intermediate mass black hole formation in globular clusters. *MNRAS*, 454:3150–3165, 2015.
13. M. H. Hénon. The monte carlo method (papers appear in the proceedings of iau colloquium no. 10 gravitational n-body problem (ed. by myron lecar), r. reidel publ. co., dordrecht-holland.). *Ap&SS*, 14:151–167, 1971.
14. G. Hobbs, D. R. Lorimer, A. G. Lyne, and M. Kramer. A statistical study of 233 pulsar proper motions. *MNRAS*, 360:974–992, 2005.
15. J. R. Hurley, O. R. Pols, and C. A. Tout. Comprehensive analytic formulae for stellar evolution as a function of mass and metallicity. *MNRAS*, 315:543–569, 2000.
16. J. R. Hurley, C. A. Tout, and O. R. Pols. Evolution of binary stars and the effect of tides on binary populations. *MNRAS*, 329:897–928, 2002.
17. A. Hypki and M. Giersz. Mocca code for star cluster simulations - i. blue stragglers, first results. *MNRAS*, 429:1221–1243, 2013.
18. P. Kroupa. On the variation of the initial mass function. *MNRAS*, 322:231–246, 2001.
19. N. W. C. Leigh, T. Böker, T. J. Maccarone, and H. B. Perets. Gas depletion in primordial globular clusters due to accretion on to stellar-mass black hole. *MNRAS*, 429:2997–3006, 2013.
20. P. Madau and M. J. Rees. Massive black holes as population iii remnants. *ApJ*, 551:L27–L30, 2001.
21. M. C. Miller and D. P. Hamilton. Production of intermediate-mass black holes in globular clusterss. *MNRAS*, 330:232–240, 2002.
22. J. Morawski, M. Giersz, A. Askar, and K. Belczynski. Mocca-survey database i: Assessing gw kick retention fractions for bh-bh mergers in globular clusters. *MNRAS*, 481:2168–2179, 2018.
23. S. F. Portegies Zwart, H. Baumgardt, P. Hut, J. Makino, and S. L. W. McMillan. Formation of massive black holes through runaway collisions in dense young star clusters. *Nature*, 428:724–726, 2004.
24. L. Rezzolla, E. Barausse, E. N. Dorband, D. Pollney, C. Reisswig, J. Seiler, and S. Husa. Final spin from the coalescence of two black holes. *Phys. Rev. D*, 78:44, 2008.
25. J. S. Stodółkiewicz. Dynamical evolution of globular clusters. i. *Acta Astron*, 32:63–91, 1982.
26. J. S. Stodółkiewicz. Dynamical evolution of globular clusters. ii - binaries method. *Acta Astron*, 36:19–41, 1986.

# Microlensing events in the Galactic bulge

María Gabriela Navarro*,[1,2,3], Dante Minniti[1,3,5], Roberto Capuzzo-Dolcetta[2], Rodrigo Contreras Ramos[3,4] and Joyce Pullen[3]

[1] *Departamento de Ciencias Físicas, Facultad de Ciencias Exactas, Universidad Andres Bello, Av. Fernandez Concha 700, Las Condes, Santiago, Chile*
* *E-mail: gabriela.navarro@roma1.infn.it*

[2] *Dipartimento di Fisica, Università degli Studi di Roma "La Sapienza", P.le Aldo Moro, 2, I00185 Rome, Italy*

[3] *Millennium Institute of Astrophysics, Av. Vicuna Mackenna 4860, 782–0436, Santiago, Chile*

[4] *Instituto de Astrofisica, Pontificia Universidad Católica de Chile, Av. Vicuna Mackenna 4860, 782–0436 Macul, Santiago, Chile*

[5] *Vatican Observatory, V00120 Vatican City State, Italy*

For the first time we detected microlensing events at zero latitude in the Galactic bulge using the VISTA Variables in the Vía Láctea Survey (VVV) data.[9] We have discovered a total sample of $N = 630$ events within an area covering 20.7 sq. deg.[11,12] Using the near-IR color magnitude diagram we selected $N = 291$ red clump sources, allowing us to analyse the longitude dependence of microlensing across the central region of the Galactic plane. We thoroughly accounted for the photometric and sampling efficiency. The spatial distribution is homogeneous, with the number of events smoothly increasing toward the Galactic center. We find a slight asymmetry, with a larger number of events toward negative longitudes than positive longitudes, that is possibly related with the inclination of the bar along the line of sight. We also examined the timescale distribution which shows a mean on $17.4 \pm 1.0$ days for the whole sample, and $20.7 \pm 1.0$ for the Red Clump subsample

*Keywords*: Gravitational lensing: microlensing — Galaxy: bulge — Galaxy: structure.

## 1. Introduction

Gravitational microlensing is a geometrical effect related to the apparent increase in the brightness of a background source by an object (lens) located sufficiently close to the line of sight.[13] This effect is useful to detect objects independently of their intrinsic brightness, with a wide range of masses from planets to black holes.

The main surveys dedicated to detect microlensing events in the galactic bulge are the Massive Astrophysical Compact Halo Objects (MACHO),[2] the Optical Gravitational Lensing Experiment (OGLE),[16] the Microlensing Observations in Astrophysics (MOA),[4] the Expérience pour la Recherche d'Objets Sombres (EROS),[3] the Disk Unseen Objects (DUO),[1] the Wise Observatory[14] and the Korean Microlensing Telescope Network (KMTNet).[7,8]

Although a large number of events have been discovered, it is not possible to study the plane of the Milky Way ($|b| < 2°$) with these optical surveys due to the

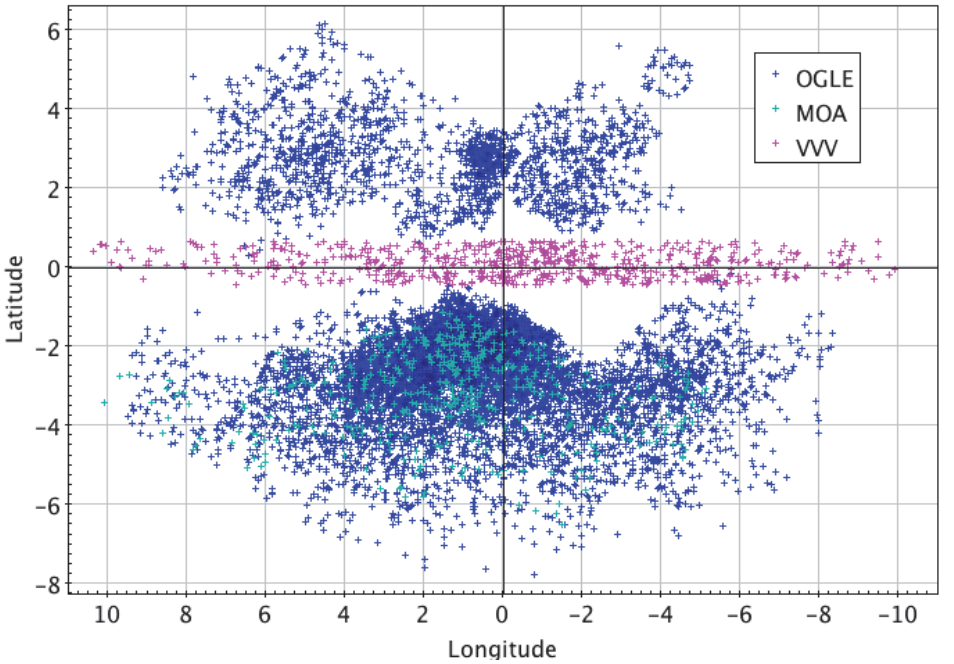



Fig. 1. Spatial distribution of the VVV microlensing events presented in this work (magenta crosses) around the Galactic center. The blue crosses are the microlensing events from the OGLE Early Warning System (EWS, http://ogle.astrouw.edu.pl) and the cyan crosses are the microlensing events discovered by MOA. All of the events happened between 2010 and 2015. Figure from Ref. 12 reproduced with permission.

high extinction and crowding. However, the Galactic plane is especially interesting because, as the density of stars increases, we expect to find microlensing events in large quantities.[19]

The only way to study the plane of the Galaxy is through observations in infrared bands. Motivated by this, we carried out the first systematic search of microlensing events at low latitudes using the *VISTA Variables in the Vía Láctea Survey* (VVV).[9] We analysed 14 VVV tiles (from b327 to b340) covering a large area within $-10.00° \leq l \leq 10.44°$ and $-0.46° \leq b \leq 0.65°$. The 3 innermost tiles (b332, b333 and b334) were published in Ref. 11. The other 11 tiles are presented in Ref. 12.

## 2. The VVV data

The *VISTA Variables in the Vía Láctea Survey* (VVV)[9] is an ESO public survey scanning the Milky Way bulge and adjacent section of the southern mid-plane in the near-infrared. Using the Visible and Infrared Survey Telescope for Astronomy (VISTA) 4-meter telescope located at ESO Cerro Paranal Observatory in Chile.

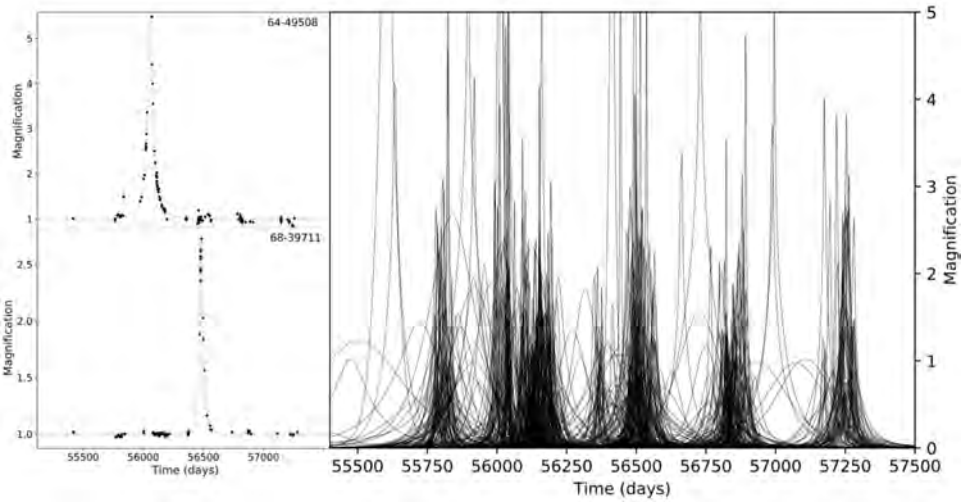



Fig. 2. Left: Examples of two typical microlensing event light curves. Right: Fits to all 630 microlensing events discovered between 2010 and 2015.

The area studied comprises $63 \times 10^6$ light curves for individual point sources and a multi-epoch campaign in the $K_s$-band ($K_s < 17.5$mag) with $\sim 90$ epochs spanning six seasons of observations (2010–2015). The PSF photometry was carried out using DAOPHOT II/ALLSTAR package.[5] The search procedure was done using the standard microlensing model, i.e., considering sources and lenses as point-like objects.

## 3. VVV Microlensing events

Initially we found 182 microlensing events in the three innermost tiles (b332, b333 and b334), with an excess in the number of events in the central tile and a relatively large number of long timescale events (with $t_E > 100$ days). We extended the sample in order to study in greater detail the behaviour along the Galactic longitude. The final sample consists of 630 microlensing events within $-10.00° \leq l \leq 10.44°$ and $-0.46° \leq b \leq 0.65°$. The complete sample is shown in Fig. 1 along with the OGLE and MOA events discovered during the same period. Fig. 2 shows all of the microlensing fits as well as two typical light curves. The animation that shows the microlensing events detected during the operation of the VVV survey is available online.[a] In the upper panel an image of the studied area is shown and the flash corresponding to each microlensing event detected in their respective positions. The intensity of the flash is associated with the amplitude. In the lower panel this figure

[a] https://www.youtube.com/watch?v=Uv08EAKVSQo

is reproduced as a function of time. The different colors correspond to the tile of each event and therefore are related to its position.

The events show a homogeneous distribution, smoothly increasing in numbers towards the Galactic centre, as predicted by different models due to the high stellar density. Considering that we expect to find a smaller amount of microlensing events due to the extreme extinction and crowding,[17] the increase in the number of events towards the center of the galaxy is expected to be even more peaked. Additionally we found an asymmetry in the density distribution with a higher number of events ($\sim 60\%$) at negative Galactic longitudes, this can be explained by the inclination of the bar along the line of sight.

Red Clump (RC) sources are low-mass stars that are burning Helium in the core. The study of the sources located in the RC is more reliable because they are more localised in the Color-Magnitude Diagram (CMD) and we can assume that they belong to the bulge. Consequently, using the near-IR Color-Magnitude Diagram we selected the Red Clump sources to analyze the longitude dependence of microlensing across the central region of the Galactic plane. Fig. 3 shows the CMD of the 14 VVV tiles along with the final sample of events. The red circles are the events located in the RC, corresponding to 46% of the sample ($N = 291$ events).

In order to study the timescale distribution it is crucial to implement an efficiency analysis. The photometric efficiency was performed using artificial star simulations for the RC stars.[17] The incompleteness is severe as we approach to the

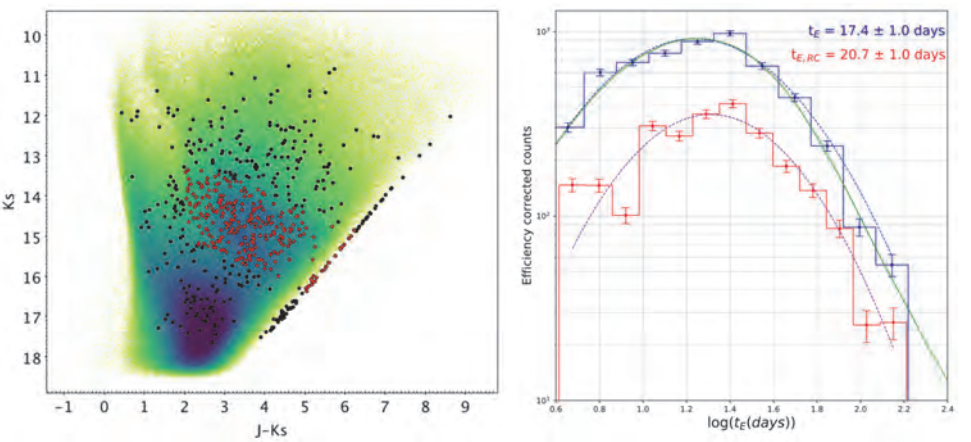



Fig. 3. Left: $K_s$ vs. $J$-$K_s$ color magnitude diagram for the 14 VVV tiles. The black circles are the microlensing sources, and the red circles are the red clump sources. Right: Efficiency corrected timescale distributions. The blue histogram is the distribution of the complete sample. The red histogram corresponds to the RC sources. Dashed lines show the best fit model of each histogram. The green shows the model proposed by Wegg et al.[18] normalised to the peak of our distribution. Figure from Ref. 12.

Galactic center, reaching values close to 40%. The sampling efficiency is cadence and timescale dependent, and was computed using Monte Carlo simulations for each fixed representative timescales with random impact parameters and times of maximum magnification. Due to the different cadence among the 14 tiles studied we performed the efficiency for each tile separately. With this analysis we conclude that the VVV Survey is limited by the sampling efficiency and is more efficient discovering long timescale events than short timescale events due to the large time coverage but low and irregular cadence.

The efficiency corrected timescale distribution is shown in Fig. 3 for the complete sample and for the RC sources. The mean timescale is $17.4 \pm 1.0$ days for the complete sample and $20.7 \pm 1.0$ days for the RC sources, in agreement with previous results. We also examined the timescale distribution for different longitudes, obtaining a clear decrease in the mean timescales as we approach to zero longitudes. A similar trend was previously reported by Wyrzykowski et al. 2015[19] but at higher latitudes using OGLE data.

## 4. Summary

We detected $N = 630$ microlensing events in the plane of the Galaxy using the VVV data. We performed for the first time a longitude analysis across the central Galactic plane at $b = 0$ deg. The spatial distribution shows a peak in the center as expected, although it is asymmetric, with more events towards negative longitudes.

The efficiency analysis demonstrates that the VVV survey is successful in detecting long/intermediate timescale events in highly reddened areas. Short timescale events are harder to find due to the low cadence of the survey (nightly at best). Our results can be complemented with other near-IR surveys to study the event rate, such as the UKIRT microlensing survey,[15] which is currently mapping the inner bulge with a higher cadence.

We computed the timescale distribution corrected by efficiency for the complete sample and RC sources. The distribution shows a peak consistent with previous measurements at higher latitudes in the bulge.

This study has many applications, from the study of the general behaviour of the complete microlensing population, to the study of specific events such as the events showing parallax effect, binary sources or lenses, as well as the very long timescale events which favor massive lenses (like black holes[10]). Furthermore, the analysis at very low latitudes is useful to optimise the observational campaign for the Wide Field Infrared Survey Telescope (WFIRST)[6] and to complement other recent near IR surveys.[15]

In this context, the study of a larger area is warranted to increase the statistics and in turn cover common areas with OGLE and MOA in order to complete a panoramic coverage of the event rates in the inner regions of the Milky Way.

## References


1. Alard, C., Guibert, J., Bienayme, O., et al. 1995, Messenger 80, 31.
2. Alcock, C., Akerlof, C. W., Allsman, R. A., et al. 1993, Nature, 365, 621.
3. Aubourg, E., Bareyre, P., Brhin, S., et al. 1993, Nature, 365, 623.
4. Bond, I. A., Abe, F., Dodd, R. J., et al. 2001, MNRAS, 327, 868.
5. Contreras Ramos, R., Zoccali, M., Rojas, F., et al. 2017, A&A, 608, 140.
6. Green, J., Schechter, P., Baltay, C., et al. 2012, arXiv:1208.4012.
7. Kim, S.-L., Park, B.-G., Lee, C.-U., et al. 2010, Proc. SPIE, 7733, 77733.
8. Kim, D.-J., Kim, H.-W., Hwang, K.-H., 2017, arXiv:1703.06883.
9. Minniti D., Lucas, P. W., Emerson, J. et al. 2010, New Astron., 15, 433.
10. Minniti, D., Contreras Ramos, R., Alonso-García, J., et al. 2015, ApJ, 810, L20.
11. Navarro, M. G., Minniti, D. & Contreras Ramos, R. 2017, ApJ, 851, L13.
12. Navarro, M. G., Minniti, D. & Contreras Ramos, R. 2018, ApJ, 865, L5.
13. Paczyński, B. 1986, ApJ, 304, 1.
14. Shvartzvald, Y., & Maoz, D. 2012, MNRAS, 419, 3631.
15. Shvartzvald, Y., Bryden, G., Gould, A., et al. 2017, AJ, 153 61.
16. Udalski, A., Szymanski, M., Kaluzny, J., et al. 1993, AcA, 43, 289.
17. Valenti, E., Zoccali, M., Gonzalez, O. A., et al. 2016, A&A, 587L, 6V.
18. Wegg, C., Gerhard, O., & Portail, M., 2016, ApJ, 843, L5.
19. Wyrzykowski Ł., Rynkiewicz A. E., Skowron J., et al., 2015, ApJS, 216, 12.

# Stellar black hole binary mergers in open clusters

S. Rastello[1,2,*], M. Arca-Sedda[3] and R. Capuzzo-Dolcetta[1]

[1] *Dep. of Physics, Sapienza, Universitá di Roma, P.le A. Moro 5, 00185, Roma, Italy*
[2] *Dip. di Fisica e Astronomia 'G. Galilei', Universitá di Padova, Vicolo dell'Osservatorio 3, 35122, Padova, Italy*
[3] *Astronomisches Rechen-Institut, Mönchhofstraße 12-14, 69120 Heidelberg, Germany*

[*] *E-mail: sara.rastello@uniroma1.it*

We study the evolution of a massive primordial hard black hole binary (BHB) in small- and intermediate-size isolated star clusters, modelled as proxies of galactic open clusters (OCs), by means of direct $N$-body simulations. Some of our models show a significant hardening of the BHB in a relatively short time. Some of them merge within the cluster. The perturbation of stars around BHB systems is key to induce their coalescence. Under our assumptions, we estimate a BHB merger rate of $R_{\rm mrg} \sim 2$ yr$^{-1}$ Gpc$^{-3}$. In some cases the BHB triggers tidal disruption events which, however, are not linked to the GW emission.

*Keywords*: Galaxy: open clusters and associations: general – stars: black holes stars: kinematics and dynamics – gravitational waves.

## 1. Introduction

The Laser Interferometer Gravitational-Wave Observatory (LIGO) and Virgo have detected five sources of GWs[1–4] corresponding to a system of BHB.

BHBs, can form either (i) in the field in isolation and via stellar evolution of a binary of two extended stars[5] (ii) or via dynamical interactions in a dense stellar system[6,7] Here we focus on the evolution of primordial stellar BHBs in low-mass star clusters. In such models the impulsive effect produced by the stochastic, large, fluctuations of the field felt by stars, whose amplitude over the mean field is of order $\sqrt{N}/N$, can significantly affect the BHB dynamical evolution. This effect is also reflected in the two-body relaxation time scale which is short in such systems if compared for example to globular clusters. In addition, the relatively low number of stars of OCs gives the possibility to integrate the system upon a long evolution time without needing a rescaling of the results.

We shortly resume as follows the methods and results of the recent paper by Rastello et al. (2019).[8]

## 2. Method and Models

To study the BHB evolution inside its parent cluster, we used `NBODY7`,[9] a direct summation $N$-body code that integrates in a reliable way the motion of stars in stellar systems, with a careful treatment of strong gravitational encounters, taking also into account stellar evolution. We created four sets of simulations representing OC models at varying initial number of stars (Table 1, column 1). Each cluster is modelled assuming a Kroupa IMF and a Plummer density profile at virial equilibrium with a core radius $r_c = 1$ pc, and solar metallicity ($Z_\odot$). The OCs have

masses in the range 300 $M_\odot$ − 3000 $M_\odot$ (Table 1, column 2). For simplicity, we do not take into account primordial binaries. Then we assumed that each cluster host a primordial hard massive BHB composed by two BH of mass $M_{BH} = 30\ M_\odot$ each) that is initially placed in the OC centre. The initial BHB semi major axis is 0.01 pc and we considered two initial eccentricities, $e = 0$ and $e = 0.5$ (Table 1, columns 3 and 4). To give statistical significance to the results we generated 150 different realizations of every model, which are denoted with names A00, A05, B00, B05, C00, C05, D00 and D05, where the letter refers to increasing $N$ and the digits to the initial BHB orbital eccentricity. All models were evolved up to 3 Gyr, which is about 3 times the simulated OC relaxation time.

Table 1. Model main parameters.

| $N_{cl}$ | $M_{cl}$ ($M_\odot$) | $a$(pc) | e | name |
|---|---|---|---|---|
| 512 | $3.2 \times 10^2$ | 0.01 | 0.0 | A00 |
| | | | 0.5 | A05 |
| 1024 | $7.1 \times 10^2$ | 0.01 | 0.0 | B00 |
| | | | 0.5 | B05 |
| 2048 | $1.4 \times 10^3$ | 0.01 | 0.0 | C00 |
| | | | 0.5 | C05 |
| 4096 | $2.7 \times 10^3$ | 0.01 | 0.0 | D00 |
| | | | 0.5 | D05 |

## 3. Dynamics of the black hole binary

### 3.1. *General evolution*

Our simulations show the existence of three main evolutionary scenarios for the fate of the BHB because of the gravitational interactions (Table 2): (i) the binary shrinks, becoming harder; (ii) the BHB gains energy, increasing its semi-major axis and, finally, (iii) the BHB can be disrupted.

We can see that, typically, about 90% of all binaries shrink their semi-major axis as they evolve (Fig. 1), as one can expect upon the so-called *Heggie's law* (softer binaries get softer while hard binaries get harder). In few of the studied cases, typically those of very low dense clusters (model A and B), the BHB becomes wider. Because of the initial choice of the BHB semi-major axis, gravitational encounters with other stars rarely *ionise* it, although we observe a few such events, typically below 7%.[8] This ionisations happen in the interval between $\sim$ 5Myr and $\sim$ 100Myr, usually driven by the encounter of the BHB with a massive star ($\gtrsim 10\ M_\odot$).[8]

## 4. Sources of gravitational waves

### 4.1. *Relativistic binaries*

The fraction of BHB coalescences in our simulations is $\sim$ 3%. In particular, in low density clusters, model A and B, the percentage of mergers is quite low,

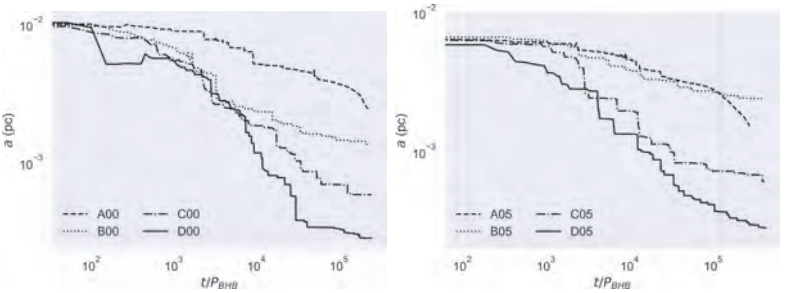



Fig. 1. Semi-major axis evolution in four of our simulations for initially circular and eccentric binaries.

Table 2. Percentage of BHBs (i) getting harder (col. 2); (ii) getting wider (col. 3) or (iii) break up (col. 4).

| Model | % Harder | % Wider | % Break up |
|---|---|---|---|
| A00 | 89.1 | 7.9 | 2.9 |
| A05 | 97.1 | 2.1 | 0.7 |
| B00 | 92.5 | 2.7 | 4.8 |
| B05 | 94.0 | 2.0 | 4.0 |
| C00 | 93.6 | 0 | 6.4 |
| C05 | 96.5 | 0 | 3.5 |
| D00 | 94.2 | 0 | 5.8 |
| D05 | 97.1 | 0 | 2.8 |

around 0.7%. As the mass (and the density) of the cluster increases, the number of relativistic mergers found is larger: 2.1% in model C00, 4.3% in C05 7.1% in D00 and 5.7% in D05. The majority of mergers occur in a time range between 5 Myr and 1.5 Gyr. As shown in Fig. 2, it is remarkable that the pericentre distances drop down to $7-8$ orders of magnitude with respect to the initial value. The eccentricities fluctuate significantly, episodically reaching values very close to unity. We found that about 50% of the mergers are mediated by a three body encounter with a perturber star which is thus a fundamental ingredient for BHB coalescence in low dense star clusters. An example of such mechanism is discussed as follows.

#### 4.1.1. *A detailed example of a merger event*

In order to check with accuracy the process of BHB coalescence upon perturbation, we followed the evolution of one of the allegedly merging BHB by mean of the few-body integrator `ARGdf`.[7] Based on the `ARCHAIN` code,[10] `ARGdf` includes a treatment of dynamical friction effect in the algorithmic regularization scheme, which models at high precision strong gravitational encounters also in a post-Newtonian scheme with terms up to the 2.5 order. We chose, at random, one of our simulations of the D00 sample to set initial conditions for the high precision evolution of a "pre merger" BHB considering its interaction with the closest 50 neighbours, number that we checked as sufficient to give accurate predictions at this regard.

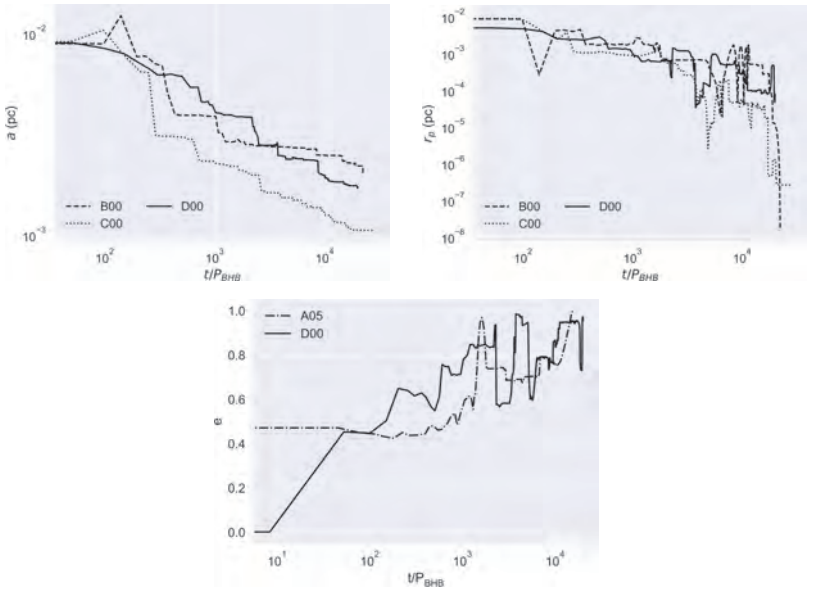



Fig. 2. Time evolution of the BHB semi major axis, pericentric distance, and eccentricity (from left to right, respectively) of some representative coalescence cases.

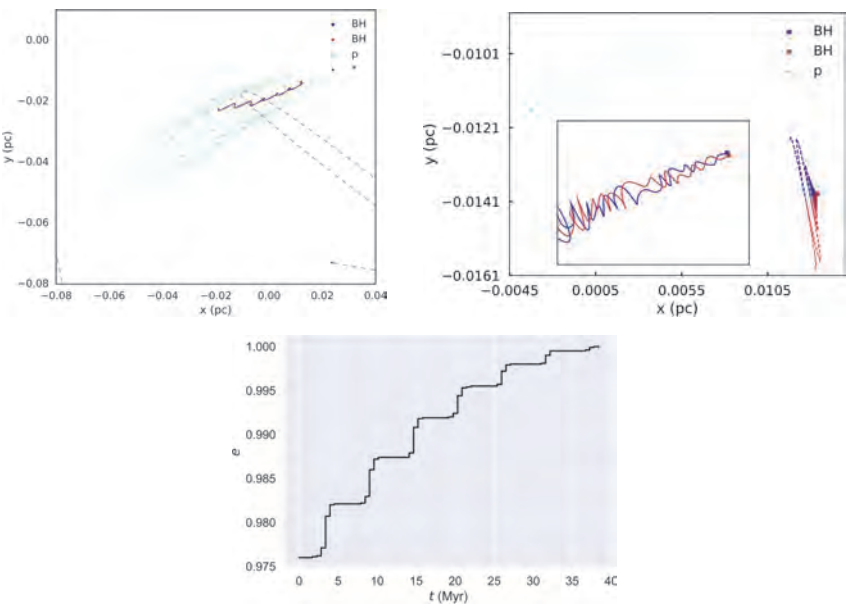



Fig. 3. *Left panel*: triple formation, *right panel*: zoom of the three body interaction, *bottom panel*: BHB eccentricity evolution.

This integration is a clear example of the relevance of dynamical interactions with other stars. Fig. 3 (*left panel*) is a snapshot of the BHB evolution and the formation of a triple system with a perturber star.[a] The BHB shrinks by interacting with such perturber, of mass 3.4 $M_\odot$, which is on a retrograde respect to the inner binary and with an inclination of 105° inducing an eccentric Kozai-Lidov mechanism. We note also a flyby star of mass 0.5 $M_\odot$ which interacts with the triple system (BHB & perturber). In Fig. 3 (*bottom panel*) we display the step-like increase of the BHB eccentricity, which is marked by the repeated interactions with the outer star. At any revolution of the perturber around the BHB we observe a step increasing of the eccentricity. On the contrary, this flyby is not sufficient to make a significant perturbation on the eccentricity evolution. Fig. 3 (*right panel*) sketches the evolution of the BHB latest orbits before the coalescence event. The plot in the rectangle is a zoom of the final part of the BHB trajectory (at its right side), spanning a length $\sim 10^{-7}$ pc. Therefore, in this particular case the triple system built up is the main ingredient that drives the BHB coalescence. A similar result is derived by Ref. 11 for low dense star clusters.

### 4.2. *Gravitational Waves*

In Fig. 4 we show the amplitude vs frequency of emitted gravitational waves for the case described in. 4.1.1 Using the last orbital parameters of the binary which correspond to the last integration made with `ARGdf`, we derive a coalescence time $T_{\rm mrg} \cong 7$ yrs. We have set the luminosity distance to that of the first source detected by LIGO,[1] which corresponds to a redshift of about $z = 0.11$. As described by the work of Chen & Amaro-Seoane (2017),[12] only circular sources are audible by LISA, which is "deaf" to eccentric binaries of stellar-mass black holes that emit their maximum power at frequencies farther away from LISA. Hence, this particular source would only enter the Advanced LIGO detection band.

### 4.3. *Merger Rate*

We estimate the merger rate following the prescription in Rastello et al., 2019[8], Sec. 4.5, Eq. (3) and Eq. (4). Assuming the total number of OCs in our Galaxy and the number of Milky way-like galaxies within redshift z = 1, we obtain a merger rate $\mathcal{R}_{\rm mrg} \approx 2{\rm Gpc}^{-3}\,{\rm yr}^{-1}$.

This estimate is however derived under the most favourable conditions, and thus represents the highest merger rate expected from low-mass OCs. Note that the BHB merger rate inferred from the first LIGO observations (GW150914) is in the range 2 - 600 $\rm Gpc^{-3}\ yr^{-1}$.[13] Our BHB merger rate is consistent with those found in Refs. 11, 14, 15 for BHB mergers in Young Massive Star Clusters. Although BHB mergers originating in open clusters-like systems might be less numerous than those

[a] An animation of the triple orbit and the eccentricity evolution is available at `https://youtu.be/zk8waNtubLk`.

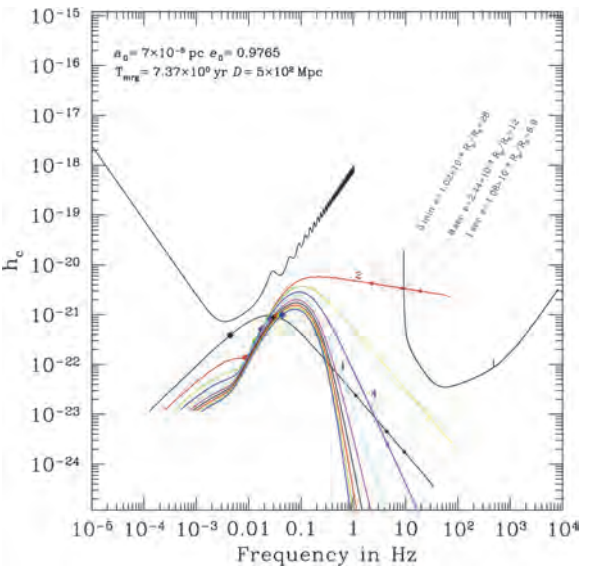



Fig. 4. Characteristic amplitude $h_c$ of the first most important harmonics for the model of described in Sec. 4.1.1 at a luminosity distance of $D = 500$ Mpc.

produced in massive star clusters, they would add a comparable amount to the BHB merger rate in the Universe because of their larger abundance.[11,15,16]

### 4.4. *Tidal Disruption Events*

As a serendipity outcome in our simulations, we found that in a few cases, the BHB disruption is mediated by a star, which binds to one of the two BHB former components.[8] The newly formed BH-star pair is characterized by a high eccentricity ($e > 0.9$) and a pericentre sufficiently small to rip the star apart and give rise to a tidal disruption event (TDE). Such events are however not linked to the BHB coalescence. In our models, TDEs involve either main sequence stars (MS), stars in the core He burning phase (HB) or in the early asymptotic giant branch (AGB) phase. Indeed, a component swap occurs in 28.5% of the cases, with the new companion star being swallowed by the heavier BH. Our findings suggest that X-ray or UV emission from OCs can be the signature of the presence of BHs with masses as high as $20 - 30$ M_$\odot$. Following the indications in Ref. 8 we found a TDE rate of $\Gamma_{\rm TDE} = 0.3 - 3.07 \times 10^{-6}{\rm yr}^{-1}$ per MW-like galaxies in the local Universe.

## 5. Conclusions

We studied the evolution of a primordial massive black hole binary in low density star clusters, and we derived that: (i) in $\sim$ 95% of the simulations performed, the BHB hardens (its semi-major axis reduces by 2 to 4 orders of magnitude) due to the repeated scatterings with flyby stars, while its eccentricity increases

significantly. This process takes place on a relatively short time-scale, $\sim$ 1 Gyr; (ii) in $\sim$ 1.2% of the cases, instead, the perturbations induced by massive stars that occasionally approach the BHB make it wider; (iii) in the remaining $\sim$ 4.8% cases, the interactions with OC stars are sufficiently strong to break up the BHB.

In $\sim$ 3% of the models, the star-BHB interactions are sufficiently effective to drive the BHB coalescence within a Hubble time. We find that a crucial ingredient to induce the BHB to merge is the interaction with a perturbing star, whose individual action considerably shortens the merger time. In our simulations, we see mergers take place in a time ranging from 5 Myr to 2.9 Gyr. In a few cases, the merging binaries emit GWs from the $10^{-3}$ to the 10 Hz frequency band. This suggests that merging BHBs in OCs can, potentially, be seen both by LISA, $\sim$ 200 yr before the merger, and LIGO, during the last phase preceding the merger.

Extrapolating our results to the typical population of OCs in MW-like galaxies in the local Universe, we found that the most optimistic merger rate for this type of BHB mergers in low-mass stellar systems is $\mathcal{R}_{\rm mrg} \sim 2$ yr$^{-1}$ Gpc$^{-3}$, a value compatible with the rate expected for galactic nuclei, but smaller than that inferred for globular clusters and young massive clusters.

Finally, we found that tidal disruption events in OCs would occur at a rate $\Gamma_{\rm TDE} = 3.08 \times 10^{-6}$ yr$^{-1}$ per MW-like galaxies in the local Universe.[8]

## Acknowledgments

SR acknowledges Sapienza, Universitá di Roma, which funded the research project "Black holes and Star clusters over mass and spatial scale" via the grant AR11715C7F89F177. SR is thankful to Sverre Aarseth of the Institute of Astronomy, Cambridge, for his helpful comments and suggestions during the development of this work. MAS acknowledges the Sonderforschungsbereich SFB 881 "The Milky Way System" (subproject Z2) of the German Research Foundation (DFG) for the financial support provided. MAS also acknowledges financial support from the Alexander von Humboldt Foundation and the Federal Ministry for Education and Research in the framework of the research project "The evolution of black holes from stellar to galactic scales".

## References

1. B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari and et al., Observation of Gravitational Waves from a Binary Black Hole Merger, *Physical Review Letters* **116**, p. 061102 (February 2016).
2. B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari and et al., GW151226: Observation of Gravitational Waves from a 22-Solar-Mass Binary Black Hole Coalescence, *Physical Review Letters* **116**, p. 241103 (June 2016).

3. B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya and et al., GW170608: Observation of a 19 Solar-mass Binary Black Hole Coalescence, *ApJ* **851**, p. L35 (December 2017).
4. B. P. Abbott, R. Abbott, T. D. Abbott, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari, V. B. Adya and et al., GW170814: A Three-Detector Observation of Gravitational Waves from a Binary Black Hole Coalescence, *Physical Review Letters* **119**, p. 141101 (October 2017).
5. K. Belczynski, D. E. Holz, T. Bulik and R. O'Shaughnessy, The first gravitational-wave source from the isolated evolution of two stars in the 40-100 solar mass range, *Nature* **534**, 512 (June 2016).
6. M. Mapelli, Massive black hole binaries from runaway collisions: the impact of metallicity, *MNRAS* **459**, 3432 (July 2016).
7. M. Arca-Sedda and R. Capuzzo-Dolcetta, The MEGaN project II. Gravitational waves from intermediate-mass and binary black holes around a supermassive black hole, *MNRAS* **483**, 152 (February 2019).
8. S. Rastello, P. Amaro-Seoane, M. Arca-Sedda, R. Capuzzo-Dolcetta, G. Fragione and I. Tosta e Melo, Stellar black hole binary mergers in open clusters, *MNRAS* **483**, 1233 (February 2019).
9. S. J. Aarseth, Mergers and ejections of black holes in globular clusters, *MNRAS* **422**, 841 (May 2012).
10. S. Mikkola and K. Tanikawa, Algorithmic regularization of the few-body problem, *MNRAS* **310**, 745 (December 1999).
11. S. Banerjee, Stellar-mass black holes in young massive and open stellar clusters and their role in gravitational-wave generation III: dissecting black hole dynamics, *MNRAS* **481**, 5123 (December 2018).
12. X. Chen and P. Amaro-Seoane, Revealing the Formation of Stellar-mass Black Hole Binaries: The Need for Deci-Hertz Gravitational-wave Observatories, *ApJ* **842**, p. L2 (June 2017).
13. B. P. Abbott, R. Abbott, T. D. Abbott, M. R. Abernathy, F. Acernese, K. Ackley, C. Adams, T. Adams, P. Addesso, R. X. Adhikari and et al., The Rate of Binary Black Hole Mergers Inferred from Advanced LIGO Observations Surrounding GW150914, *ApJ* **833**, p. L1 (December 2016).
14. S. Banerjee, Stellar-mass black holes in young massive and open stellar clusters and their role in gravitational-wave generation, *MNRAS* **467**, 524 (May 2017).
15. S. Banerjee, Stellar-mass black holes in young massive and open stellar clusters and their role in gravitational-wave generation − II, *MNRAS* **473**, 909 (January 2018).
16. U. N. Di Carlo, N. Giacobbo, M. Mapelli, M. Pasquato, M. Spera, L. Wang and F. Haardt, Merging black holes in young star clusters, *arXiv e-prints* (January 2019).

# The secular evolution of the Milky Way nuclear star cluster

S. Rastello[1,2,*], M. Arca-Sedda[3], R. Capuzzo-Dolcetta[1] and R. Spurzem[3,4,5]

[1] *Dep. of Physics, Sapienza, Universitá di Roma, P.le A. Moro 5, 00185, Roma, Italy*
[2] *Dip. di Fisica e Astronomia 'G. Galilei', Universitá di Padova, Vicolo dell'Osservatorio 3, 35122, Padova, Italy*
[3] *Astronomisches Rechen-Institut, Mönchhofstraße 12-14, 69120 Heidelberg, Germany*
[4] *National Astronomical Observatories and Key Laboratory of Computational Astrophysics, Chinese Academy of Sciences, 20A Datun Rd., Chaoyang District, 100012, Beijing, China*
[5] *Kavli Institute for Astronomy and Astrophysics at Peking University, 5 Yiheyuan Rd., Haidian District, 100871, Beijing, China*

[*] *E-mail: sara.rastello@uniroma1.it*

Nuclear Stellar Clusters are so dense stellar systems (containing up to $10^7$ stars in few parsec radius) that their dynamics cannot be followed with the high precision, direct summation available $N$-body codes over a long integration time. Here we present the main idea of our new computational approach to study the dynamics of the MW NSC exploiting the facilities of the high precision hybrid parallelized code NBODY6++GPU. This strategy will allow us to study the evolution of the NSC over its 2-body relaxation time with an acceptable numerical effort by mean of a massive computational platform.

*Keywords*: Stars: kinematics and dynamics; galaxies: black holes; galaxies: star clusters: general; methods: numerical.

## 1. Introduction

The Milky Way (MW) hosts a Super Massive Black Hole (SMBH) in its center, of mass of about $M_{SMBH} \sim 4 \times 10^6$ $M_\odot$ that is surrounded by a very dense stellar system: the Nuclear Stellar Cluster (hereafter MWNSC). The NSC is very massive ($10^6 - 10^7$ $M_\odot$) and luminous ($4.1 \times 10^7$ $L_\odot$) with a half-light radius of $\sim 4$ pc.[1,2] The density of the NSC is of the order of $\sim 10^7$ $M_\odot$ $pc^{-3}$.

A peculiarity of all NSCs known is that they always contain an old stellar population (age >1 Gyr), and most of them show also the presence of a young stellar population with ages below 100 Myr. This feature is probably connected with the NSC origin. According to the so called *in situ model*,[3] NSCs consist mostly of stars formed (locally) from radial gas inflow into the galactic center. On the other hand the *dry-merging scenario*[4,5] explains the origin of NSCs by massive clusters that migrated to the galactic center via dynamical friction and, further, merge to form a dense nucleus. Even if these models are still debated, there are many observational evidences which favor the dry merging model in case of dE galaxies and also in the case of the MW NSC.[6]

The MW NSC is the one we know better because we are able to observe with high resolution resolving many of its stars, so it constitutes a wonderful "laboratory" to study many astrophysical processes in the innermost Galactic region, within the SMBH influence radius, such as Tidal Disruption Events (TDEs) and Extreme Mass Ratio Inspirals (EMRIs).

### 1.1. *The Computational Challenge*

Despite the large quantity of works regarding the formation of NSCs, in literature there is a lack of material concerning their subsequent evolution.[7,8] This is mainly due to numerical problems because NSCs are so dense systems that their dynamics cannot be followed with high-precision, direct summation, available $N$-body codes.

The easiest way to model dense star clusters with a central massive black hole is the use of Fokker-Planck or Monte Carlo approaches, but these methods are severely approximated.[9] Actually, to achieve high accuracy and reliable results, it would me much better resorting on direct $N$-body simulations. Thanks to the recent improvements of the available $N$-body codes, it was possible to perform direct $N$-body simulation up to one million of stars[10] as done by Wang et al., (2016)[11] for globular clusters, or Panamarev et al. (2018)[12] for the Galactic centre, both using the code `NBODY6++GPU`.[10] Nevertheless, $N$-body simulations of star clusters remains very time consuming. Considering individual time-stepping, the computational cost per crossing time depends on the number of steps per particle, which varies with different time-step control criteria as well as on the integration method and star cluster properties.[10,11] Thus, an efficient parallelization of a direct $N$-body code is compulsory when dealing with large particle numbers. One of the best suitable codes for studying a very crowded star cluster as the MW NSC is `NBODY6++GPU`.[10] Such code is an MPI parallelization of the standard `NBODY6` code apt to an accurate integration of the movement of many bodies. It is optimal for collisional systems, where long times of integration and high accuracy are both required. `NBODY6++GPU` is designed to work across multiple nodes and, on a supercomputer, it allows to perform $N$-body simulations up to one million of stars. Our simulation tests and our future production simulations run on the super-computing cluster JUWELS (developed by Forschungszentrum Jülich Supercomputing Centre, Germany) qualified as the best German computer in the TOP500 list of the fastest supercomputers in the world, with a 12.6 Petaflops/s peak. It is composed by 2500 nodes equipped with Intel Xeon CPUs ($2 \times 24$ cores) and 48 nodes hosting 4 NVIDIA Tesla V100 GPUs each.

In the following section we present our planned implementation of the idea of one of the authors of this paper (RCD) as numerical strategy aiming to cope with the problem of the study of the evolution of a NSC towards relaxation with a portable numerical method and avoiding a mere re-scaling of short simulation results.

## 2. Numerical Strategy

In spite of its speed and efficiency, the `NBODY6++GPU` cannot deal with a NSC that hosts more than $\approx 10^7$ stars over many crossing times. So, some clever numerical strategy is needed which, at the same time, leads to a long term simulation run and a sufficiently high precision. A scheme of our "numerical approach" is shown in Fig. 1.

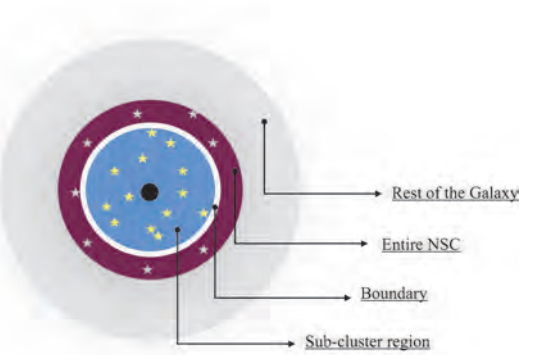



Fig. 1. Scheme of our approach (image is not to scale). The SMBH (central black circle) is surrounded by a light-blue region which corresponds to an NSC "sub-cluster" region, delimited by a (white) boundary. The magenta shell identifies the entire physical NSC. The (external) grey region refers to the external galactic components (nuclear bulge, bulge, halo and disk).

In the assumption of a "local" spherical symmetry in the NSC[a] star distribution, for computational convenience we consider a "sub-cluster" region of the NSC (light-blue region in Fig. 1) which contains approximately $N \sim 10^6$ stars (indicated in yellow in Fig. 1) which corresponds to the maximum number of stars that we can simulate with `NBODY6++GPU`. The boundary of the sub-cluster region is indicated by a white shell (Fig. 1) of radius $r_b$. The sub-cluster region is initially sampled from a Plummer density profile of solar metallicity stars whose masses follow the Kroupa IMF. The number of primordial binaries is assumed to be 5% of the total number of nuclear cluster stars. Stellar evolution is considered as implemented in the standard `SSE` and `BSE` tools.[13]

The entire NSC occupies the whole magenta sphere in Fig. 1. The dynamical role of stars which cross in or out the NSC region cannot be neglected to study the evolution of the entire NSC. Therefore to account for their contribute we developed a method described as follows:

- The boundary and the sub-cluster region are not physical limits and the stars can freely evolve across these regions. We keep track of the orbits of the stars that exit the sub-cluster region (moving through the magenta-ring) and re-enter either into the sub-cluster region or escape forever from it. This task is done by building up a proper *boundary condition* as described below.

[a]Since we are interested in the secular dynamical evolution of the MW NSC towards relaxation, the initial conditions for the $N$-body simulation are constructed assuming the dry-merger formation of the MW NSC.

We exclude or continue to consider in our computation stars that, on the base of their mechanical energy, are so energetic that they likely cross with no return within the $r_b$ boundary or, being less energetic, would actually re enter the boundary. These latter, indeed, will likely spend a time $\Delta t$ out of the sub-cluster region before re enter it. So, they are dynamically important for the global evolution of the NSC and need to be properly taken into account.

The "re-injection" of any, *i-th*, of these stars is done by a suitable estimate of the *re-enter* time $t_{re,i}$. Actually, in the hypothesis that outside the sub-cluster region the potential is spherically symmetric, the *i-th* object individual energy per unit mass is

$$E_i \;=\; \frac{1}{2}\,\dot{r_i}^2 \;+\; \frac{1}{2}\frac{L_i^2}{r_i^2} \;+\; U(r_i) \tag{1}$$

where $L_i$ is the specific angular momentum (per mass), $r_i$ is the norm of the position vector of the $i$-th-star, $U(r_i)$ is the gravitational potential.

An estimate of $t_{re,i}$ is obtained by computing

$$\Delta t_i \equiv\; 2\int_{r_b}^{r_{apo}} \frac{dr}{\sqrt{2\,E_i - \frac{L_i^2}{r_i^2} \;- 2\,U(r_i)}}, \tag{2}$$

where $r_{apo}$ is the star apocenter distance off the sub-cluster region.

As known, the apocenter distance is solution for $r_i$ of the equation:

$$E_i \;-\; \frac{1}{2}\frac{L_i^2}{r_i^2} \;-\; U(r_i) \;=\; 0. \tag{3}$$

After determination of $t_{re,i}$, we need to assign to the $i$-th-star a position and a velocity at the crossing of the boundary. We randomly derive the new position and the new velocity of the stars when it re-enters the boundary surface. In particular, in the acceptable assumption that the $i$-th-star feels out of the $r_b$ boundary only a smooth gravitational potential, we have that the norm of its velocity when it re enters the boundary is identical to the one when it exited. The orientation of the velocity vector at re-enter is done by extracting the velocity radial component ($v_{r,i}$) an inclination $-\pi/2 < \phi < \pi/2$ with respect to the plane tangent to the re-enter point on the boundary sphere.

- In our representation of the nuclear star cluster we need also to take into account stars that cross the boundary region coming from outside. If we have information on the distribution function $f(\mathbf{r}, \mathbf{v})$ of the stellar system at the boundary we can use it to sample velocity and position of such stars. Such stars, whose masses are distributed according to the chosen IMF, are indeed "inserted" within the sub-cluster region assigning them a randomly chosen initial position on the boundary.

Since the MWNSC is not an isolated system but it is embedded in Galactic environment we need to take care also of the other Galactic components. In the case of the MW center we face the interesting co-evolution of the NSC and the central SMBH, whose presence strongly influences the dynamical evolution of the entire MWNSC. According to our strategy the central SMBH is modelled as a fixed point-mass (Kepler) potential, $\mathrm{U}_{BH} = -GM_{BH}/r$. The mass, $M_{BH}$, of the SMBH is assumed 10% of the initial total mass of the NSC. Moreover, the central black hole is allowed to grow its mass via star capture within the accretion radius of. This is crucial when aiming to the study of events like tidal disruption of stars onto the central black hole.

The other, smooth, Galactic components are modelled as a Dehnen, a Miyamoto & Nagai and a logarithmic density profile, respectively, for the bulge, the disc and the halo.

In order to develop this numerical strategy we have modified the code `NBODY6++GPU` by implementing routines to ad an external potential for central black hole and unresolved stellar component and to let escaping stars return. Results of ongoing numerical simulations will be published elsewhere.

## Acknowledgments

SR acknowledges Sapienza, Universitá di Roma, which funded the research project "Black holes and Star clusters over mass and spatial scale" via the grant AR11715C7F89F177. SR acknowledges the Astronomisches Rechen Institut of Heidelberg (ARI), Germany. MAS acknowledges the Sonderforschungsbereich SFB 881 "The Milky Way System (subproject Z2) of the German Research Foundation (DFG) for the financial support provided. MAS also acknowledges financial support from the Alexander von Humboldt Foundation and the Federal Ministry for Education and Research in the framework of the research project "The evolution of black holes from stellar to galactic scales". The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project (ID: HHD28) by providing computing time on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC). The author acknowledge support by Volkswagen Trilateral Project 90411, "Dynamical Mechanisms of Accretion in Galactic Nuclei".

## References

1. A. Feldmeier, N. Neumayer, A. Seth, R. Schödel, N. Lützgendorf, P. T. de Zeeuw, M. Kissler-Patig, S. Nishiyama and C. J. Walcher, Large scale kinematics and dynamical modelling of the Milky Way nuclear star cluster, *A&A* **570**, p. A2 (October 2014).
2. R. Schödel, A. Feldmeier, D. Kunneriath, S. Stolovy, N. Neumayer, P. Amaro-Seoane and S. Nishiyama, Surface brightness profile of the Milky Way's nuclear star cluster, *A&A* **566**, p. A47 (June 2014).

3. K. Bekki, The Formation of Stellar Galactic Nuclei through Dissipative Gas Dynamics, *PASA* **24**, 77 (July 2007).
4. S. D. Tremaine, The formation of the nuclei of galaxies. II - The local group, *ApJ* **203**, 345 (January 1976).
5. R. Capuzzo-Dolcetta, The Evolution of the Globular Cluster System in a Triaxial Galaxy: Can a Galactic Nucleus Form by Globular Cluster Capture?, *ApJ* **415**, p. 616 (October 1993).
6. D. Minniti, R. Contreras Ramos, M. Zoccali, M. Rejkuba, O. A. Gonzalez, E. Valenti and F. Gran, Discovery of RR Lyrae Stars in the Nuclear Bulge of the Milky Way, *ApJ* **830**, p. L14 (October 2016).
7. O. Y. Gnedin, J. P. Ostriker and S. Tremaine, Co-evolution of Galactic Nuclei and Globular Cluster Systems, *ApJ* **785**, p. 71 (April 2014).
8. M. Arca-Sedda and R. Capuzzo-Dolcetta, The globular cluster migratory origin of nuclear star clusters, *MNRAS* **444**, 3738 (November 2014).
9. C. L. Rodriguez, M. Morscher, L. Wang, S. Chatterjee, F. A. Rasio and R. Spurzem, Million-body star cluster simulations: comparisons between Monte Carlo and direct N-body, *MNRAS* **463**, 2109 (December 2016).
10. L. Wang, R. Spurzem, S. Aarseth, K. Nitadori, P. Berczik, M. B. N. Kouwenhoven and T. Naab, NBODY6++GPU: ready for the gravitational million-body problem, *MNRAS* **450**, 4070 (July 2015).
11. L. Wang, R. Spurzem, S. Aarseth, M. Giersz, A. Askar, P. Berczik, T. Naab, R. Schadow and M. B. N. Kouwenhoven, The DRAGON simulations: globular cluster evolution with a million stars, *MNRAS* **458**, 1450 (May 2016).
12. T. Panamarev, A. Just, R. Spurzem, P. Berczik, L. Wang and M. Arca Sedda, Direct N-body simulation of the Galactic centre, *MNRAS* **484**, 3279 (April 2019).
13. J. R. Hurley, C. A. Tout and O. R. Pols, Evolution of binary stars and the effect of tides on binary populations, *MNRAS* **329**, 897 (February 2002).

# The future Milky Way and Andromeda galaxy merger

Riccardo Schiavi* and Roberto Capuzzo-Dolcetta

*Department of Physics, Università La Sapienza, Roma, 00185, Italy*
**E-mail: riccardo.schiavi@uniroma1.it*

Manuel Arca-Sedda

*ARI, University of Heidelberg, Heidelberg, 69117, Germany*

According to our current knowledge about physical and dynamical properties of the Milky Way-M31 system, it seems likely that these two galaxies will collide and eventually merge in a time very sensitive to initial conditions. Using the HiGPUs code[1], we have performed several numerical simulations to study the dynamics of the system, trying to define the role of indeterminacy in the present day observed relative velocities of the two galaxies and the time of the merger. At the same time, we have followed the dynamics of the two massive black holes sitting in the galactic centers, to check (within the space and time resolution limits of our simulation) their relative motion upon the completion of the galaxies merger process.

*Keywords*: Galaxy merger; supermassive black holes; Milky Way-Andromeda collision.

## 1. Purposes of the work

In this work we aim to follow the evolution of the "binary" system composed by our own galaxy, the Milky Way (MW), and the nearby Andromeda galaxy (M31). Observational data are currently not sufficient to strictly constrain the dynamics of this interaction and, therefore, a wide range of scenarios for their future evolution are possible. In any case, almost all of them imply an eventual merger between the two galaxies in a time which is very dependent upon parameters of the simulation.

In our study we considered a set of initial conditions spanning a range of values of the tangential component of the relative velocity of the MW and M31. This allowed us to quantify how these differences affect the interaction. A particular subtopic of relevant interest is the motion of the two massive black holes (SMBHs) known to be present in the two galaxies central regions, with especial regard to the last phase of the galactic merging process, when the two SMBHs are expected to form a close binary system. The results we obtained so far for the evolution of their orbits during the galaxies interaction, are interesting but the present simulation resolution is not enough to represent the very final stage, down to what is called the "final parsec problem".

## 2. Methods

For these simulations we have used the HiGPUs code[1], which performs direct $N$-body calculations exploiting a full parallelization on hybrid platforms (CPUs and Graphic Processing Units). So far, we have used a number of particles of order

$3.2 \times 10^4$, aiming to extend resolution by reaching, at least, one million of particles. Initial conditions are obtained with the code "GSAM" (Galaxy SAMpler), written by Arca-Sedda et al. (2015), and able to generate star clusters models with different density distributions. In our case the two galaxies are modelled according to a King density profile.

We have performed three different simulations at varying the transverse velocity of Andromeda as seen by the MW, keeping unchanged the other parameters.

In Table 1 all the initial parameters, common to the three simulations, are shown.

Table 1. Common initial conditions in the three simulations.

| | **Milky Way** | **Andromeda** |
|---|---|---|
| Mass ($M_\odot$) | $9.0 \times 10^{11}$ | $1.5 \times 10^{12}$ |
| $W_0$ | 6 | 6 |
| King radius $r_c$ (kpc) | 1.535 | 3.385 |
| Relative radial velocity (km/s) | $-120$ | |
| Initial separation (kpc) | 780 | |

## 3. Preliminary results

Modelling the two galaxies orbital evolution, we could investigate how the merger time changes at varying the relative initial transverse velocity ($V_{t0}$), as summarized in Table 2. Our results show that the great uncertainty on this kinematical parameter makes the time of merger undetermined for a factor $\sim 80$. Actually, a wide range of merger times are found in the literature: in some cases[4] a very low transverse velocity was reported ($V_{t0} \approx 17 km/s$), whereas in others[5] a value about 10 times larger, $V_{t0} \approx 164 km/s$. According to the most recent estimations[6] a reliable value seems $V_{t0} \approx 57 km/s$, but the level of uncertainty is still very large. For these reasons, our results are only indicative, as upper and lower limits. Moreover, in these present simulations the intergalactic medium (IGM) has not been considered, which, on its side, should act to reduce the time for the merger of the two galaxies.

Table 2. Time of the merger for the three initial tangential velocities of M31.

| $\mathbf{V_{t0}}$ **(km/s)** | **Time of the merger (Gyr)** |
|---|---|
| 20 | 6.7 |
| 50 | 108.8 |
| 80 | 544.6 |

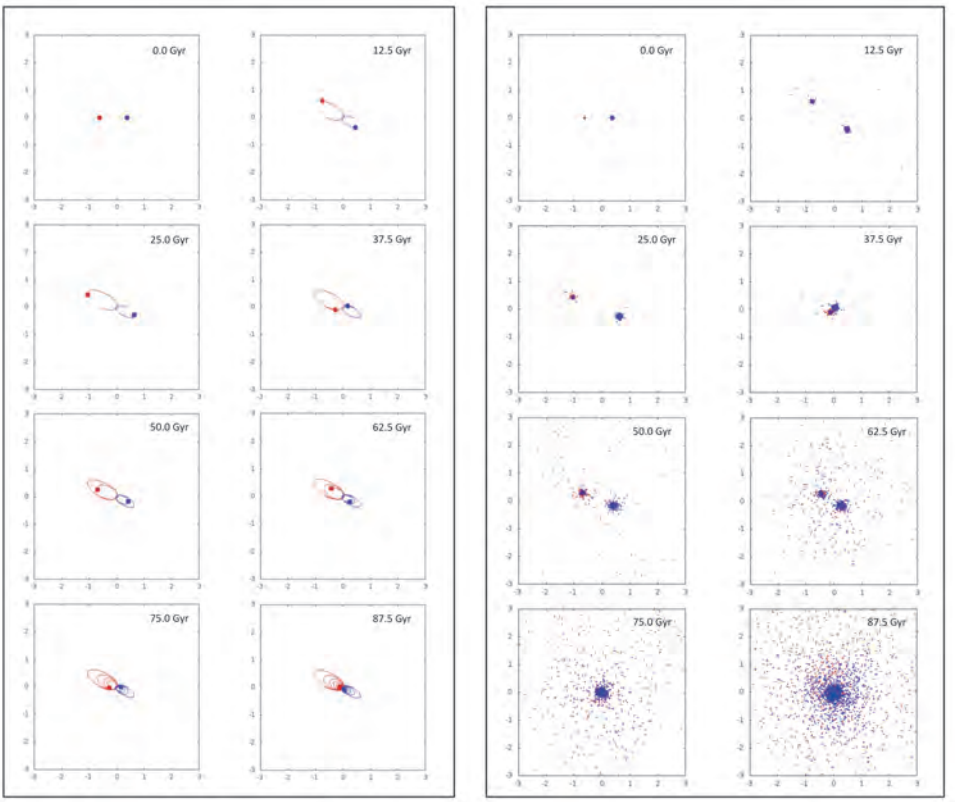



Fig. 1. Snapshots at various times of the configuration of the two BHs (left columns) and their host galaxies (right columns) in the case of $V_{t0} = 50 km/s$. Red colour refers to the Milky Way, blue to Andromeda. The unit of distance is 780 kpc. The time interval between each snapshot is 12.5 Gyr.

Figure 1 plots some snapshots of the merging process in the case with $V_{t0} =$ 50 km/s: in the left panel we show the relative motion of only the two central SMBHs while, on the right, the two interacting galaxies are shown.

We have plotted in Fig. 2 the time evolution of the distance between the two SMBHs in the three cases of $V_{t0}$ examined. It should be noticed that, after the merger of their host galaxies, the two SMBHs keep orbiting around each other at almost constant distance (roughly circular orbit). The reason why the SMBHs binary does not experience a further shrink is due to the loss of efficiency of gravitational encounters, because, due to unsufficient number of simulation particles, there are too few "stars" enclosed within the scale the two SMBHs are orbiting. This means that to reproduce satisfactorily the inner dynamics of the SMBHs binary we need at least a factor 10 larger $N$.

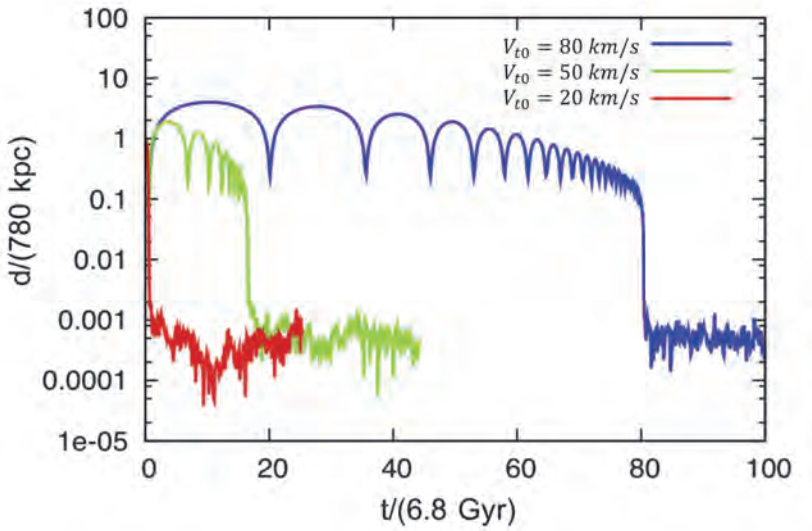



Fig. 2. Distance between the two BHs as function of time, for three different initial tangential velocities.

Anyway, an interesting result is that the distance at which the binary stalls (around 300 pc) seems to be independent of the initial tangential velocity, because it likely depends only on $N$ and not on the other initial conditions.

## 4. Tentative Conclusions

This short note reports some of our current work on the dynamics of the future merging process of our Galaxy and Andromeda. Since the tangential velocity of Andromeda is not precisely constrained by observations, we decided to perform three simulations varying this parameter. Depending of this, we have found that the time of the merger results great. This is partly due to that in these preliminary simulations we did not consider the role of the intergalactic medium, which, according to some previous simulations[3], can speed up significantly the merger process, via dynamical friction.

The relative motion of the two central SMBHs shows a regular behaviour on a large spatial scale, but we should increase the resolution to extend reliably the study also to the innermost region of the new cluster formed after the merger. Actually, we are planning other simulations with larger values of $N$, extending the range of values for $V_{t0}$, and with the inclusion of the intergalactic medium dynamical friction effect.

## References

1. Capuzzo-Dolcetta, R. et al., Journ. of Comp. Phys., v. 236, p. 580-593 (2013).
2. Binney J., Tremaine S., 1987, Galactic dynamics. Princeton, NJ, Princeton University Press, 1987.
3. Cox, T. J. and Loeb, A., MNRAS, 386, 461–474 (2008).
4. Van der Marel, R. P. et al., ApJ , 753, 9 (2012).
5. Salomon, J.-B. et al., MNRAS, 456, 4432 (2016).
6. Van der Marel, R. P. et al., ApJ, 872, 24 (2019).

# Ranking galaxies within a gravitational-wave sky localization

Francesco Brighenti*[1,2], Giuseppe Greco[1,2], Gianluca Maria Guidi[1,2], Francesco Piergiovanni[1,2], Frederique Marion[3], Benoit Mours[3], Damir Buskulic[3] and Florian Aubin[3]

[1,2] *Dipartimento di Scienze Pure E Applicate, Università degli Studi di Urbino "Carlo Bo", Urbino, Italy/INFN-Firenze*

* *E-mail: francesco.brighenti@uniurb.it*

[3] *Laboratoire d'Annecy de Physique des Particules (LAPP), Univ. Grenoble Alpes, Université Savoie Mont Blanc, CNRS/IN2P3, F-74941 Annecy, France*

The recent detection of a binary neutron star merger by the LIGO and Virgo collaborations (LVC) and its corresponding electromagnetic counterpart from several astronomer teams marks the birth of multi-messenger astronomy [1]. Due to the size of the sky localization from a GW signal only, which can span tens to thousands of square degrees, there are significant benefits to rank the galaxies inside these large sky areas to maximize the probability of counterpart detection. Here we present a new procedure to query the galaxy catalogs, rank the galaxies and eventually define a prior for time allocation and scheduling algorithms.

## 1. Introduction

Multi-messenger astronomy demonstrated its scientific potential from the very first event observed via gravitational waves and via electromagnetic (EM) radiation, GW170817. The success of the observational campaign following GW170817 originates also from some lucky coincidences, beginning with the fact that this has been the closest event to Earth and with the highest signal to noise ratio (SNR) so far detected by the gravitational waves detectors. This helped significantly in localising of the source, as the skymap produced shortly after limited the area containing the source with 90% probability to just 28 deg$^2$, instead of the usual hundreds or thousands deg$^2$.

This typical large value of the 90% area of the skymap is due to the method used to identify candidates gravitational events using the coincident SNR to assign a value to the total SNR: the current strategy adopted by the pipelines used by LVC for online detection of a gravitational wave signal is to consider whether all the online detectors find a trigger with high SNR from the same template in each detector in a time window which is compatible with a gravitational wave traveling time between each other. While using the coincident SNR limits the computational costs and allows a quick identification of candidates events, it comes at the cost of neglecting some informations that could reduce the area of the 90% probability region.

Assuming that the triggers are due to a binary neutron stars gravitational wave signal coming from a certain galaxy, one expects to measure not only trigger times compatible with the traveling time of a gravitational wave between the detectors, but also coherence in phases and amplitudes of the signals recorded in each detector. The coherent SNR includes exactly these informations to assign a value to the total

SNR measured by the detectors network. Clearly, a full-sky search with coherent SNR is computationally more expensive than a full-sky search with coincident SNR, and therefore the latter is still currently preferred by the online detection pipelines.

Our idea is to use the coherent SNR only for a reduced number of points in the sky, identified by the galaxies inside the region at the 90% c.l. of the skymap produced by BAYESTAR, and subsequently rank the galaxies according to their value of coherent SNR. The computational costs are further reduced if we avoid to compute the matched filtering output for a template bank and use instead the SNR time series already produced by an online detection pipeline, such as the Multi-Band Template Analysis (MBTA) pipeline[2].

In the following we report preliminary results from a "proof of concept" analysis performed with the LaL code lalapps_cohPTF_inspiral.c[3], that worked for us to test the viability of our approach.

## 2. Coincident and coherent SNR

Following the standard matched filtering theory, given the output of the detector $i$, $s(t) = n(t) + h(t)$, we assume that the noise $n(t)$ is Gaussian and stationary and with power spectral density $S_n(f)$ defined as $\langle \tilde{n}(f)\tilde{n}^*(f')\rangle = \frac{1}{2}\delta(f-f')S_n(f)$. This is used to define the single detector inner product between two time series $a(t_0)$ and $b(t_0+t)$,

$$(a|b)(t) = 4\Re \int_0^\infty df\, \frac{\tilde{a}(f)\tilde{b}^*(f)}{S_n(f)} e^{-i2\pi ft}\,. \tag{1}$$

The odds of a signal $h$ being present in the data can be evaluated via the likelihood $\Lambda_h(t) = \frac{P(s|h)}{P(s|0)} = \frac{e^{-(s-h|s-h)/2}}{e^{-(s|s)/2}}$, and therefore the log-likelihood for the single interferometer takes the form

$$\ln \Lambda_h(t) = (s|h)(t) - \frac{1}{2}(h|h)(t)\,. \tag{2}$$

This allows us to define the SNR time series for the waveform $h$ as

$$\rho_h^2(t) = 2\ln \Lambda_h(t) \tag{3}$$

Consider now to have two detectors, $i$ and $j$. What a typical online detection pipeline does in order to identify interesting candidates is to compute $\rho_h(t)$ for each detector output $(s^i(t)\,, s^j(t))$ and for every waveform template within a template bank. If the time series $(\rho_h^i(t), \rho_h^j(t))$ of the same template $h$ exceed a fixed threshold $\bar{\rho}$ for times $(t_i\,, t_j)$ such that $|t_i - t_j| \leq d_{ij}/c$, where $d_{ij}$ is the distance between the detectors, then the pipeline denotes it as a candidate event with coincident SNR given by

$$\rho_{coinc} = \sqrt{[\rho_h^i(t_i)]^2 + [\rho_h^j(t_j)]^2}\,. \tag{4}$$

The coherent SNR can be defined from a straightforward generalisation to a multi-detector likelihood given by $\ln \Lambda_h(t) = (\mathbf{s}|\mathbf{h}) - \frac{1}{2}(\mathbf{h}|\mathbf{h})$, where we denote the

scalar product for multiple detectors $i$ as $(\mathbf{a}|\mathbf{b}) := \sum_i (a^i|b^i)$. Taking advantage of some peculiarities of the compact binaries coalescence signals we are interested into,[a] we can obtain a simple expression for the coherent SNR as

$$\rho^2_{coh}(t) = 2\max(\ln \Lambda_h(t)) = (\mathbf{s}|\mathbf{h}_\mu)\mathcal{M}^{\mu\nu}(\mathbf{s}|\mathbf{h}_\nu)\,, \tag{5}$$

where $h^i_\mu$ are given by

$$\begin{aligned} h^i_1(t) &= F^i_+(\theta^i,\,\phi^i,\,\chi^i)h_0(t)\,,\\ h^i_2(t) &= F^i_\times(\theta^i,\,\phi^i,\,\chi^i)h_0(t)\,,\\ h^i_3(t) &= F^i_+(\theta^i,\,\phi^i,\,\chi^i)h_{\frac{\pi}{2}}(t)\,,\\ h^i_4(t) &= F^i_\times(\theta^i,\,\phi^i,\,\chi^i)h_{\frac{\pi}{2}}(t)\,, \end{aligned} \tag{6}$$

$F^i_{+,\times}$ being the detector response functions to the plus (+) and cross (×) polarization components of the waveform.[b] The matrix $\mathcal{M}^{\mu\nu}$ depends only on the sensitivities and response functions of the detectors.

The computation of the matched filtering terms $(\mathbf{s}|\mathbf{h}_\nu)$ are responsible for most of the computational cost of the coherent SNR, but are precisely what MBTA computes for its independent task of online detection. Computing coherent SNR using MBTA output will simply consists in summing these time series, properly time-shifted depending on the galaxy within the skymap for which we want to estimate the coherent SNR.

## 3. Methods

We select about 90 gravitational wave sky localizations from Ref. 4 in which a triple detection is considered. The 90% confidence level for each probability skymap is build using the MOC (Multi Order Coverage map) method based on HEALPix sky tessellation[5]. MOC is a multi-scale mapping based on HEALPix sky tessellation. It is essentially a simple way to map irregular and complex sky regions into hierarchically grouped predefined cells. Each MOC cell is defined by two numbers: the hierarchy level (HEALPIX ORDER) and the pixel index (HEALPIX NPIX). The NUNIQ scheme defines an algorithm for packing an (ORDER, NPIX) pair into a single integer for compactness[6]. We compute the MOC region at a given probability level and subsequently, we query databases for retrieving objects whose position falls within this MOC map at 90% confidence level. The GLADE catalog is used in the analysis[7].

We compute the sky position of the maximum probability pixel defined by the BAYESTAR pipeline with the healpy package[8].

The values of component masses of the template with highest coincident SNR are passed to a script that prepares a single-template bank that is used for the

[a] The interested reader can find a detailed derivation in Ref. 9.

[b] These are given in terms of $h_0$ and $h_{\frac{\pi}{2}}$ by $h_+(t) = \mathcal{A}^1 h_0(t) + \mathcal{A}^3 h_{\frac{\pi}{2}}(t)$, $h_\times(t) = \mathcal{A}^2 h_0(t) + \mathcal{A}^4 h_{\frac{\pi}{2}}(t)$.

analysis. The template bank and the GPS times related to the each event are input of the function that computes the coherent SNR, lalapps_cohPTF_inspiral, looping over the selected galaxies from the GLADE catalog cutting at 1-sigma distance reported in the image's header[3].

## 4. Preliminary results

In a small fraction of our datasample (10 cases) the injection positions meet with the highest coherent-SNR values. This fact has motivated our investigation to set up a EM follow-up strategy. Two different strategies are planned to optimise a follow-up activity in a few minutes after the first rapid sky localization using the coherent SNR measured for the galaxies within the 90% c. l. The skymap tiling is performed by the GWsky tool[10].

### 4.1. *Strategy for telescope with a large Field of View*

We measure the sky distances between (i) the position of the injection and the position of the galaxy with the highest coherent SNR; Dist (inj - MAXsnr), with median = 6.6. (ii) the position of the injection and the maximum probability pixel; Dist (inj - MAXpix), with median = 9.7. The two distances are compared in Fig. 1. The distance Dist (MAXsnr - inj) is narrower. If the trend will be confirmed, the injection could be promptly imaged in a one/few tile(s) by centering the FoV telescope at the position of the maximum coherent SNR.

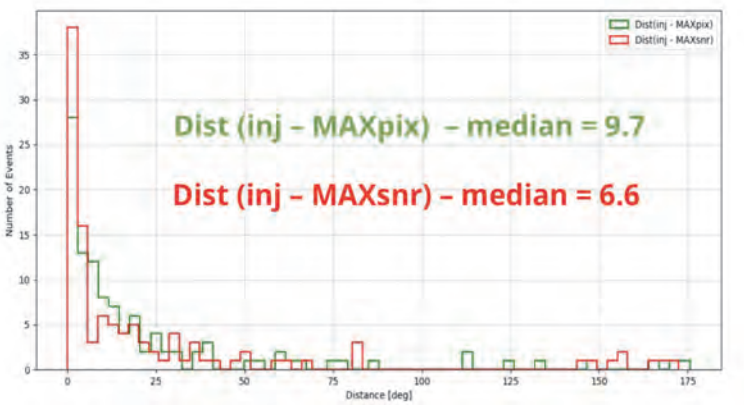



Fig. 1. Distributions of the angular distances between the true location of the injection and the pixel of maximum probability (green) or the galaxy with greatest coherent SNR (red). There is a mild indication of a better performance of our method.

### 4.2. *Strategy for telescope with a small Field of View*

We rank the coherent SNR from the highest value to the lowest value. The injection position in the rank is determined for testing. In 48 cases over 84 under examination, the galaxy target using the coherent SNR information captures the injection in a better position than galaxies localized in the confidence levels gradually growing. In Fig. 2, we show a case in which the injection is ranked at the position 3 using the the coherent SNR scenario while in the contour plots approach it is localized at the 41% c.l. in which 285 galaxies are distributed.

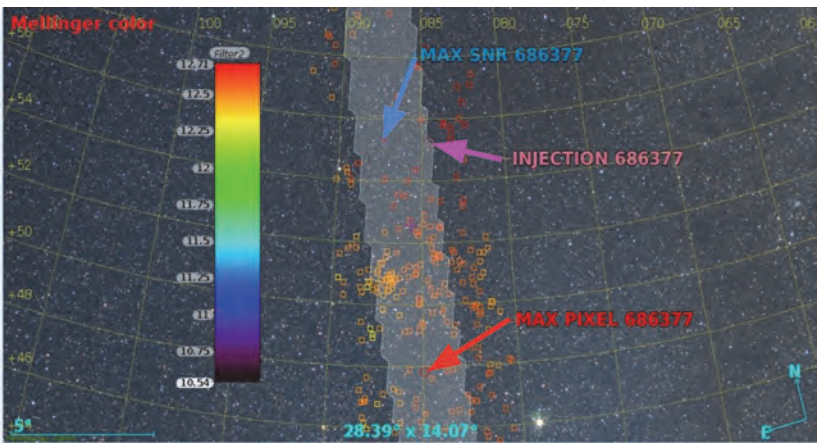



Fig. 2. A particular case in which the injection is ranked at the position 3 using the the coherent SNR scenario while in the contour plots approach it is localized at the 41% c.l. in which 285 galaxies are distributed.

## 5. Conclusion

We are testing our method to ensure its validity for a very large sample and to verify if it could significantly decrease the allocated time for an EM follow-up activity (or in which special cases, if there are). At this stage, we note a positive trend when the coherent SNR approach (measured for each galaxy in 90% c. l.) is applied to tile the skymap compared to the traditional contour- plot method provided in a rapid GW sky localization.

We are also preparing a separate code that implements this general algorithm on MBTA. Taking advantage of the matched filtering already performed by MBTA during the online search we expect to be able to obtain the rank of galaxies within few minutes from the trigger.

## References


1. LIGO and Virgo Collaboration, *Phys. Rev. Lett.* **119**, 161101 (2017).
2. T. Adams, D. Buskulic, V. Germain, G. M. Guidi, F. Marion, M. Montani, B. Mours, F. Piergiovanni, G. Wang, Class. Quant. Grav. **33** n.17, 175012 (2016).
3. lalapps_cohPTF_inspiral.c https://github.com/lscsoft/lalsuite/blob/5a47239a877032e93b1ca34445640360d6c3c990/lalapps/src/ring/coh_PTF_inspiral.c
4. L. Singer et al. `arXiv:1603.07333v4 [astro-ph.HE]`.
5. G. Greco, *Handling gravitational-wave sky maps with Multi-Order Coverage*, http://nbviewer.jupyter.org/gist/ggreco77/d43e5a1141b99f918672e70adc05864d.
6. MOC-HEALPix Multi-Order Coverage map, http://www.ivoa.net/documents/MOC/.
7. GLADE catalog on Vizier (Dalya+, 2016), http://vizier.u-strasburg.fr/viz-bin/VizieR?-source=VII%2F75.
8. Healpy https://healpy.readthedocs.io/en/latest/.
9. I. Harry, S. Fairhurst, *Phys. Rev.* **D 83**, 084002 (2011).
10. GWsky-tiling the skymap in FoV, https://github.com/ggreco77/GWsky.

# Data analysis techniques to search for the stochastic gravitational-wave background

Giancarlo Cella*

*INFN sez. Pisa,*
*Largo B. Pontecorvo 3, 56127 Pisa, Italy*
**E-mail: giancarlo.cella@pi.infn.it*

A stochastic background of gravitational waves can be described basically as the superposition of several unrelated events, of cosmological or astrophysical origin. This is a generic definition: depending on specific cases the background can show a rich statistical structure, both in the space and time domain, and in principle several physical information of great interest can be extracted from it. I will discuss the features of this kind of signal, giving some details about how detection and parameter estimation procedures can be designed to obtain optimal performances.

*Keywords*: Gravitational waves; Data Analysis, Stochastic Background.

## 1. Introduction

A stochastic background of gravitational waves can be seen as a gravitational wave field which evolves from an initially random configuration, or equivalently as the result of a superposition of signals generated by many uncorrelated and unresolved sources.

Expected stochastic backgrounds can be of cosmological origin. They can be generated by several physical mechanisms in the early evolution of the Universe, such as inflation, phase transitions, cosmic strings. These are of great importance as they can carry information about early Universe[1].

A background can be generated also by astrophysical sources[2], such as for example compact binaries coalescences, supernovae, rotating neutron stars.

The stochastic background couples to a given detector which gives as an output a signal $s_i^{\mathcal{A}}$, modeled by a stochastic process. This is really the sum of a contribution coming from the background $h_i^{\mathcal{A}}$ and another given by the detector's noise $n_i^{\mathcal{A}}$, namely

$$s_i^{\mathcal{A}} = h_i^{\mathcal{A}} + n_i^{\mathcal{A}} \tag{1}$$

Here the capital index labels the detectors, and the lowercase one enumerate the data.

## 2. Detection

If we know the statistical properties of the stochastic processes $h_i^{\mathcal{A}}$ we can evaluate the possibility of detecting a stochastic background[3]. We suppose that the noises $n_i^{\mathcal{A}}$ are Gaussian stochastic processes. We do not make at this stage particular assumptions about the signals $h_i^{\mathcal{A}}$: they are described by some probability density $p_h(h_i^{\mathcal{A}}|\underline{\alpha})$ which depends by a set of parameters $\underline{\alpha}$. Then the probability density for

observed data can be written as

$$p_s(s^{\mathcal{A}}|\underline{\alpha}) = \mathcal{N} \int p_h(h^{\mathcal{A}}|\underline{\alpha}) \exp\left\{-\frac{1}{2}[\mathbb{C}_n^{-1}]_{ij}^{\mathcal{A}\mathcal{B}}(s_i^{\mathcal{A}} - h_i^{\mathcal{A}})(s_j^{\mathcal{B}} - h_j^{\mathcal{B}})\right\} \mathcal{D}h_k^{\mathcal{A}} \quad (2)$$

where $\mathbb{C}_h$ is the covariance array of the background. Using the Bayes theorem this can be converted to the probability density for the model parameters $\underline{\alpha}$

$$p_s(\underline{\alpha}|s^{\mathcal{A}}) = \frac{p_s(s^{\mathcal{A}}|\underline{\alpha})p(\underline{\alpha})}{\int p_s(s^{\mathcal{A}}|\underline{\alpha})p(\underline{\alpha})d\underline{\alpha}} \quad (3)$$

where $p(\underline{\alpha})$ is the prior for these. This can be used to infer the properties of the background from the data. For example marginalizing over non interesting parameters we can get the joint probability distribution between the one we are interested to. It is also, at least in principle, the starting point to design a detection test such as

$$\frac{\int p_s(s^{\mathcal{A}}|\underline{\alpha}_1; \mathcal{H}_1)p(\underline{\alpha}_1)d\underline{\alpha}_1}{\int p_s(s^{\mathcal{A}}|\underline{\alpha}; \mathcal{H}_0)p(\underline{\alpha}_0)d\underline{\alpha}_0} > \lambda \quad (4)$$

where $\mathcal{H}_1$ and $\mathcal{H}_0$ are the hypothesis of presence and absence of a background. This approach is however not commonly used because it leads to quantities which are difficult to manage: a simpler strategy will be discussed in the following.

### 2.1. *The Gaussian case*

We give some details for a Gaussian stochastic background, which is the simpler to analyze. In this case

$$p_h(h^{\mathcal{A}}|\mathcal{H}_1) = \mathcal{N}' \exp\left\{-\frac{1}{2}[\mathbb{C}_h^{-1}]_{ij}^{\mathcal{A}\mathcal{B}} h_i^{\mathcal{A}} h_j^{\mathcal{B}}\right\} \quad (5)$$

where $\mathbb{C}_h$ is the spectral covariance array of the stochastic background signal. The integral (2) can be computed explicitly in this case. However here we want to follow a different but almost equivalent approach. Given a network of detectors, we want to find a statistic of the form

$$\hat{Y} = \int \hat{\mathbb{Q}}^{\mathcal{A}\mathcal{B}}(\omega)\mathbb{S}^{\mathcal{A}\mathcal{B}}(\omega)d\omega, \qquad \mathbb{S}^{\mathcal{A}\mathcal{B}}(\omega) = s^{\mathcal{A}*}(\omega)s^{\mathcal{B}}(\omega) \quad (6)$$

with the best signal to noise ratio. We work in the frequency domain assuming a stationary background and we set $\hat{\mathbb{Q}}^{\mathcal{A}\mathcal{B}} = 0$ when $\mathcal{B} = \mathcal{A}$. The signal to noise ratio is given (when $\mathbb{C}_h \ll \mathbb{C}_n$) by

$$\frac{<\hat{Y}>}{\sqrt{<\hat{Y}^2> - <\hat{Y}>^2}} \simeq \frac{T\sum_{\mathcal{A}\neq\mathcal{B}} \int \hat{\mathbb{Q}}^{\mathcal{A}\mathcal{B}}(\omega)\mathbb{C}_h^{\mathcal{A}\mathcal{B}}(\omega)d\omega}{\sqrt{2T\sum_{\mathcal{A}\neq\mathcal{B}} \int |\hat{\mathbb{Q}}^{\mathcal{A}\mathcal{B}}(\omega)|^2 \mathbb{C}_n^{\mathcal{A}\mathcal{A}}(\omega)\mathbb{C}_n^{\mathcal{B}\mathcal{B}}(\omega)d\omega}} \quad (7)$$

where $T$ is the observation time. This is maximized by choosing

$$\mathbb{Q}^{\mathcal{A}\mathcal{B}}(\omega) \propto \frac{\mathbb{C}_h^{\mathcal{B}\mathcal{A}}(\omega)}{\mathbb{C}_n^{\mathcal{A}\mathcal{A}}(\omega)\mathbb{C}_n^{\mathcal{B}\mathcal{B}}(\omega)} \quad (8)$$

The statistic of $\hat{Y}$ is a Gaussian one, and we can compute the false alarm probability for a given threshold $\lambda$ as

$$P_{FA}(\lambda) = \int_{\lambda}^{+\infty} \mathcal{N}(x; \mu_0, \sigma_0^2) dx = \frac{1}{2}\mathrm{erfc}\left(\frac{\lambda - \mu_0}{\sqrt{2}\sigma_0}\right) \tag{9}$$

The detection probability can be obtained in the same way, by substituting the mean $\mu_0 = 0$ and the variance $\sigma_0$ of $\hat{Y}$ evaluated in absence of the signal with the analogous quantities $\mu_1$ and $\sigma_1 \simeq \sigma_0$ evaluated when the signal is present. The detection probability for a given false alarm can now be written as

$$P_D \simeq \frac{1}{2}\mathrm{erfc}\left(\mathrm{erfc}^{-1}\left(2P_{FA}\right) - \frac{1}{2}\sqrt{T\sum_{A\neq B}\int \frac{|\mathbb{C}_h^{\mathcal{AB}}(\omega)|^2}{\mathbb{C}_n^{\mathcal{AA}}(\omega)\mathbb{C}_n^{\mathcal{BB}}(\omega)} d\omega}\right) \tag{10}$$

Note that $\mathbb{C}_n^{\mathcal{AA}}$ is the noise power spectrum in the detector $\mathcal{A}$. The detection probability increases with the square root of the measurement time, and is reduced when there is a loss of coherence between the detectors. This coherence is usually described by the overlap reduction function

$$\gamma^{\mathcal{AB}}(\omega) \equiv \mathbb{C}_h^{\mathcal{AB}}(\omega) / \sqrt{\mathbb{C}_h^{\mathcal{AA}}(\omega)\mathbb{C}_h^{\mathcal{BB}}(\omega)} \tag{11}$$

which depends on the separation and orientation of the two detectors.

## 3. Models

We can write the observed background as a superposition of modes in some specific basis, for example

$$h_{ij}(\vec{x}, t) = \sum_P \int df \int d\Omega_{\hat{n}} A^P(\hat{n}, f) \epsilon_{ij}^P(\hat{n}) e^{i\frac{2\pi f}{c}(\hat{n}\cdot\vec{x} - ct)} \tag{12}$$

where the amplitudes $A^P(\hat{n}, f)$ are modeled as stochastic variables. Here the sum is over the polarizations $P$ and the integrals over frequencies and directions of propagation. In this way the background is described by a (multivariate) probability distribution $p\left(A^P(\hat{n}, f)\right)$ or equivalently by the amplitude's moments (we can safely assumed the first order one to be zero)

$$\left\langle A^{P_1}(\hat{n}_1, f_1) \cdots A^{P_k}(\hat{n}_k, f_k) \right\rangle \qquad k = 2, \cdots \tag{13}$$

The coupling between the background and a detector is given by

$$h^{\mathcal{A}}(t) = D_{ij}^{\mathcal{A}}(t) h_{ij}(\vec{x}^{\mathcal{A}}(t), t) \tag{14}$$

Here $D_{ij}^{\mathcal{A}}(t)$ is the detector tensor, which is time dependent because the orientation of the $\mathcal{A}$–th detector can change with time owing to the earth's rotation. The gravitational wave strain field is evaluated at the position $\vec{x}^{\mathcal{A}}$ of the detector. Using Eq. (12) we get

$$h^{\mathcal{A}}(t) = \sum_P \int df \int d\Omega_{\hat{n}} A^P(\hat{n}, f) D_{ij}^{\mathcal{A}}(t) \epsilon_{ij}^P(\hat{n}) e^{i\frac{2\pi f}{c}\hat{n}\cdot\vec{x}^{\mathcal{A}}(t)} e^{i2\pi f t} \tag{15}$$

so we can associate a set of stochastic processes to the set of detectors.

If the amplitudes $A^P(\hat{n}, f)$ follow a (multivariate) Gaussian distribution, then the processes $h^{\mathcal{A}}$ are completely described by

$$\begin{aligned}\mathbb{C}_h^{\mathcal{AB}}(t,t') &= \left\langle h^{\mathcal{A}}(t) h^{\mathcal{B}}(t') \right\rangle \\ &= \sum_{P,Q} \int df \int d\Omega_{\hat{n}} \int df' \int d\Omega_{\hat{m}} \left\langle A^P(\hat{n}, f) A^Q(\hat{m}, f')^* \right\rangle \\ &\quad \times D_{ij}^{\mathcal{A}}(t) D_{kl}^{\mathcal{B}}(t')^* \epsilon_{ij}^P(\hat{n}) \epsilon_{kl}^Q(\hat{m})^* e^{i\frac{2\pi}{c}\left(f\hat{n}\cdot\vec{x}^{\mathcal{A}}(t) - f'\hat{m}\cdot\vec{x}^{\mathcal{B}}(t')\right)} e^{i2\pi(ft - f't')}\end{aligned} \tag{16}$$

On the typical time scale of observations, we can assume that the statistical properties of the background do not change. Under this assumption of stationarity we can write

$$\left\langle A^P(\hat{n}, f) A^Q(\hat{m}, f')^* \right\rangle \equiv \Lambda^{PQ}(\hat{n}, \hat{m}, f) \delta(f - f') \tag{17}$$

and we obtain

$$\begin{aligned}\mathbb{C}_h^{\mathcal{AB}}(t,t') &= \sum_{P,Q} \int df \int d\Omega_{\hat{n}} \int d\Omega_{\hat{m}} \Lambda^{PQ}(\hat{n}, \hat{m}, f) \\ &\quad \times D_{ij}^{\mathcal{A}}(t) D_{kl}^{\mathcal{B}}(t')^* \epsilon_{ij}^P(\hat{n}) \epsilon_{kl}^Q(\hat{m})^* e^{i\frac{2\pi f}{c}\left(\hat{n}\cdot\vec{x}^{\mathcal{A}}(t) - \hat{m}\cdot\vec{x}^{\mathcal{B}}(t')\right)} e^{i2\pi f(t-t')}\end{aligned} \tag{18}$$

Note that this does not mean that the stochastic processes $h^{\mathcal{A}}$ are stationary. The reason is that a non stationarity is induced by the motion of the detectors, which change their positions and their orientations owing to the earth's rotation and revolution around the sun. If we neglect revolution, which gives an effect which is usually negligible, we can write

$$\vec{r}^{\mathcal{A}}(t) = \mathbf{R}(t) \vec{r}^{\mathcal{A}} \tag{19}$$

where $\mathbf{R}$ is a rotation matrix and

$$\begin{aligned}\mathbb{C}_h^{\mathcal{AB}}(t,t') &= \sum_{P,Q} \int df \int d\Omega_{\hat{n}} \int d\Omega_{\hat{m}} \Lambda^{PQ}(\mathbf{R}(t)\hat{n}, \mathbf{R}(t')\hat{m}, f) \\ &\quad \times D_{ij}^{\mathcal{A}} D_{kl}^{\mathcal{B}*} \epsilon_{ij}^P(\hat{n}) \epsilon_{kl}^Q(\hat{m})^* e^{i\frac{2\pi f}{c}\left(\hat{n}\cdot\vec{x}^{\mathcal{A}} - \hat{m}\cdot\vec{x}^{\mathcal{B}}\right)} e^{i2\pi f(t-t')}\end{aligned} \tag{20}$$

The simpler situation is when the background is isotropic and there are no correlations between different directions in the sky. In this case $\Lambda^{PQ}$ does not depend on the directions and is insensitive to the rotation of the detectors. By using the previous expressions to connect $\mathbb{C}^{\mathcal{AB}}$ to physical quantities we obtain

$$\Omega_{gw}(f) = \frac{1}{\rho_c} \frac{d\rho_{gw}}{d\log f} = \frac{4\pi^2 f^3}{3H_0^2} \mathbb{C}_h^{\mathcal{AA}}(f) \tag{21}$$

where $\rho_{gw}$ is the energy density of the background and $\rho_c$ the present value of the critical energy density for closing the universe. In the same way it is possible to evaluate the overlap reduction function.

### 3.1. *Non standard polarizations*

General relativity predict only two different polarizations for gravitational waves. But a general metric theory can have up to six independent polarizations.

In principle it is possible to disentangle these polarizations[4] by using the signals of a network of detectors[5]. The idea is that each polarization has a peculiar overlap reduction function, so it is possible to introduce the statistics

$$Y^{\mathcal{AB}}(f) = \frac{20\pi^2}{3H_0^2 T} f^3 s^{\mathcal{A}}(f)^* s^{\mathcal{B}}(f) \tag{22}$$

whose expectation values are connected to the energy densities of the different polarizations by the overlap functions

$$\langle Y^{\mathcal{AB}}(f) \rangle = \gamma_P^{\mathcal{AB}}(f) \Omega_{gw}^P(f) \tag{23}$$

The different contributions are obtained by inverting this relation.

### 3.2. *Anisotropic backgrounds*

It is possible to investigate an anisotropic stochastic background[6] by exploiting the time dependence in Eq. (20). The basic idea[8] is that contribution to the measured signal are modulated in phase and amplitude in a way which depends on the detector and on the direction. If we introduce on each detector a demodulation factor which is appropriate for a given direction, when we evaluate the correlation we wash out other contributions. In this way it is possible to reconstruct a map of the background intensity in the sky, with an angular resolution which scales as $\lambda_{gw} d^{-1}$, where $d$ is the spatial separation between the detectors.

This can be seen as a solution of an inverse problem, and it comes out that at least three detectors are needed to remove the degeneracies.

### 3.3. *Non Gaussian backgrounds*

As we said in the introduction we can look at a stochastic background as a result of the superposition of many uncorrelated events. This is especially true for an astrophysical stochastic background[7], where the events can be for example coalescences of binary compact objects.

A key parameter that characterize the background is the duty cycle for a given upper cut-off in the red shift parameter $z$

$$D(z) = \int_0^z (1+z')\tau \times \frac{d\mathcal{R}}{dz'} dz' \tag{24}$$

where the first factor is the observed duration of the event and the second the observed rate of events. When $D \ll 1$ the events does not overlap and can be detected separately. When $D \gg 1$ we are in the opposite situation, there is a very large overlap and the signal can be described by a Gaussian stochastic background.

The intermediate regime $D \sim 1$ is a peculiar one. It is not possible to resolve each events owing to the overlap, but the typical number of overlapped signals is not too large and the stochastic background is not a Gaussian one.

Several approaches have been elaborated in order to deal with this, details can be found in the literature[9–13].

## 4. Conclusions and perspectives

During the last years several kind of analysis have been applied to the data, attempting to detect a stochastic background of gravitational waves.

No detections have been obtained until now, but interesting upper limit exists[14]. There are some promising candidates[15,16] that hopefully will be detected when the final sensitivity of current advanced detectors will be reached.

In the next generation of detectors, with an improvement of one order of magnitude in sensitivity, it will be possible to test for non trivial models, to detect sub–threshold events as a background, and maybe to start a synergy with other kind of detectors using multimessenger strategies.

## References

1. C. Caprini, J. Phys. Conf. Ser. **610** (2015) no. 1, 012004.
2. T. Regimbau, Res. Astron. Astrophys. **11** (2011) 369.
3. J. D. Romano and N. J. Cornish, Living Rev. Rel. **20** (2017) no. 1, 2.
4. B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) Phys. Rev. Lett. 120, 201102.
5. A. Nishizawa, A. Taruya, K. Hayama, S. Kawamura and M. Sakagami, Phys. Rev. **D** 79, 082002.
6. B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) Phys. Rev. Lett. 118, 121102 – Published 24 March 2017.
7. D. Meacher, E. Thrane and T. Regimbau, Phys. Rev. D **89** (2014) no. 8, 084063.
8. N.J. Cornish, Class. Quant. Grav. 18, 20, 2001.
9. S. Drasco and É. É. Flanagan, Phys. Rev. **D** 67, 082003.
10. E. Thrane, Phys. Rev. **D**87, 043009 (2013).
11. N. Seto, Phys. Rev. D 80, 043003 (2009).
12. N. Seto, Astrophys. J. Lett. 683, L95 (2008).
13. L. Martellini and T. Regimbau, Phys. Rev. **D** 89, 124009.
14. B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) Phys. Rev. Lett. 118, 121101.
15. B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) Phys. Rev. Lett. 116, 131102.
16. B.P. Abbott et al. (LIGO Scientific Collaboration and Virgo Collaboration) Phys. Rev. Lett. 120, 091101.

# Direction of arrival estimation for transient GW sources via time-frequency representations

Paolo Addesso, Maurizio Longo and Vincenzo Matta

*DIEM, University of Salerno, Fiscano (SA), 84084, Italy*
*E-mail: {paddesso, longo, vmatta}@unisa.it*
*www.unisa.it*

Elena Mejuto Villa, Vincenzo Pierro, Innocenzo M. Pinto and Luigi Troiano

*DING, University of Sannio, Benevento, 82100, Italy*
*E-mail: {mejutovilla, pierro, pinto, troiano}@unisannio.it*
*www.unisannio.it*

The localization of Gravitational Wave (GW) sources, that is crucial in identifying their physical nature via the joint use of GW interferometers and other messengers (electromagnetic, neutrino, etc.), is mainly based on the observed delays between pairs of interferometers. Time-Frequency (TF) representations can be effectively used for GW detection and parameter estimation. In particular, it is possible to estimate the arrival time delay between the GW signals detected by two interferometers by suitably aligning the related TF maps. In this work we compare different TF representatons and alignment techniques, by using numerical simulation based on recent public-domain GW data.

*Keywords*: Gravitational Waves, Time-Frequency Analysis, Wigner-Ville distribution, Constant Q Transform, Triangular Norms, Phase Correlation, Delay Estimation.

## 1. Introduction

Direction of arrival (DOA) estimation is a crucial aspect of gravitational wave (GW) and multi-messenger astronomy. The DOA can be retrieved with high accuracy from the observed arrival-time delays $d$ between the GW signals detected by different pairs of interferometers in a network of three or more non-colocated detectors.

Time-Frequency (TF) maps are tools of choice used for unmodeled transient GW signals affected by *nonstationary* disturbances of instrumental and/or environmental origin. The Constant Q-Transform (CQT) [1], in several flavours (including Q, Omega and Omicron pipelines), is currently used in LIGO-Virgo Collaboration (LVC) data analysis. Among available alternatives, the Wigner-Ville (WV) distribution features the best localization/resolution properties [2], but contains intermodulation artifacts, due to its bilinear nature. The Radial-Gaussian Kernel (RGK) *smoothed* WV [3] removes these artifacts, that hinder the visual readability of the WV distribution, with minimal resolution loss. The main purpose of this paper is to evaluate the effectiveness of the TF maps in estimating the GW signals DOA. Due to the image-like structure of the TF maps, it is possible to use alignment techniques borrowed from image processing. In particular, we will test the *Phase Correlation* (PhC) method [4] and the *Łukasiewicz Triangular Norm* (LTN) method [5], both implemented with *sub-pixel* accuracy, by using numerical simulations.

## 2. TF Maps for Delay Estimation: CQT and WV

The CQT of a signal $x(t)$ for a given value of the (so-called *quality factor*) parameter $Q$ is defined as follows:

$$X_{CQ}(t, f, Q) = \int_{-\infty}^{+\infty} x(\tau) w(\tau - t, f, Q) e^{-i2\pi f \tau} d\tau. \tag{1}$$

The best performance (the finest time-frequency resolution) is achieved when $w(\tau - t, f, Q)$ is a *Gaussian window*, but this latter has infinite support. Therefore, in GW data analysis pipelines the so-called *bisquare window* is used instead. For increasing values of $Q$, the CQT has lower temporal resolution. This effect is evident also in Fig. 1, displaying two CQT maps of the GW150914 signal as received by the LIGO-Hanford interferometer[6]. Lower values of $Q$ are therefore better suited for time delay $d$ estimation.

The WV distribution of a signal $x(t)$ is defined as follows

$$W_x(t, f) = \int_{-\infty}^{+\infty} \tilde{x}\left(t + \frac{\tau}{2}\right) \tilde{x}^*\left(t - \frac{\tau}{2}\right) e^{-i2\pi f \tau} d\tau, \tag{2}$$

where $\tilde{x}(t) = x(t) + i\,\mathcal{H}[x(t)]$ is the the analytic mate of $x(t)$, $\mathcal{H}[\cdot]$ denoting the Hilbert transform operator.

The WV transform is a *quadratic* representation and is affected by artifacts due to intermodulation (e.g. see Fig. 1). A possibility to solve this issue is to perform a *smoothing* procedure[3]. This amounts to computing the *Ambiguity Function* (AF), i.e. the 2D-Fourier transform of WV distribution[2]

$$A(\xi, \tau) = \mathcal{F}[X_W(t, f)], \tag{3}$$

where $\mathcal{F}[\cdot]$ denotes the $(t, f) \to (\xi, \tau)$ Fourier operator, and using a suitable kernel $K(\xi, \tau)$ to produce a new TF representation, given by

$$X_S(t, f) = \mathcal{F}^{-1}[A(\xi, \tau) K(\xi, \tau)], \tag{4}$$

where $\mathcal{F}^{-1}[\cdot]$ denotes the $(\xi, \tau) \to (t, f)$ inverse Fourier operator. In particular a radially-Gaussian smoothing kernel can be used, that in the polar coordinates $(\rho, \theta)$, of the AF domain can be written:

$$K(\rho, \theta) = \exp\left\{-\frac{\rho^2}{2\sigma^2(\theta)}\right\}. \tag{5}$$

Note that the AF symmetry property $A(-\xi, -\tau) = A(\xi, \tau)$ implies $K(\rho, \theta + \pi) = K(\rho, \theta)$. Therefore we may determine $\sigma(\theta)$ to maximize the energy content of the kernel-weighted AF subject to a volume constraint, namely:

$$\max_{\sigma(\theta)} \int |A(\rho, \theta) K(\rho, \theta)|^2 \rho d\rho d\theta : \int |K(\rho, \theta)|^2 \rho d\rho d\theta \leq \alpha. \tag{6}$$

The smoothed version of the WV distribution is thus tailored to the actual (energy) distribution in the AF plane. In Fig. 1 we show the WV-based TF maps of the GW150914 signal, as received by the LIGO-Hanford detector[6].

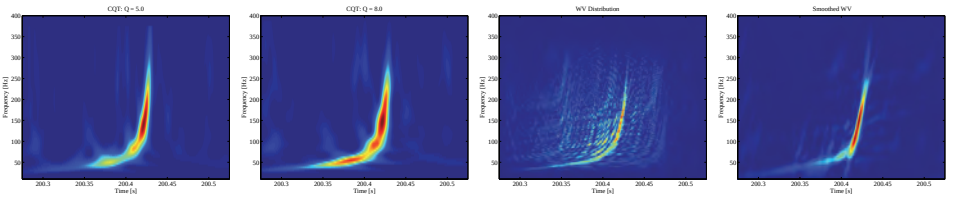



Fig. 1. TF maps for GW150914 (LIGO-Hanford)[6]: the two leftmost images are CQT with $Q = 5$ and $Q = 8$; the last rightmost ones are the plain and smoothed WV.

## 3. Delay Estimation

Consider two detectors that (disregarding the different directional gains) receive signals $x(t)$ and $y(t) = x(t-d)$ respectively, yielding the TF representations $|X(t, f)|$ and $|Y(t - d, f)|$. In practice, each TF map is evaluated at discrete times and frequencies, $(t_n, f_m) = (n/f_s, m/f_s)$, where $f_s$ is the sampling rate and $n$ and $m$ are integers such that $n \in [0, N - 1]$ and $m \in [0, M - 1]$, yielding a discrete 2D image $\mathbf{X} \in \mathbb{R}^{N\times M}$, whose elements are $X(n, m) = |X(t_n, f_m)|$.

In this subsection we will present two techniques to estimate the delay $d$ by aligning the two discretized TF representations $\mathbf{X}$ and $\mathbf{Y}$ at subpixel level, so as to minimize the effect of TF discretization on the time delay estimation.

### 3.1. *Phase Correlation*

The first alignment technique is based on the well known *Phase Correlation* (PhC) method, using a subpixel accuracy version implemented in a computational efficient way[4]. The 2D Discrete Fourier Transform (DFT) is denoted as $\mathcal{F}_{a\times b}[\cdot] : (n, m) \to (k, l)$, where $a \times b$ is the matrix dimension. Then, we compute the normalized cross-spectrum

$$\mathbf{S} = \circ \frac{\mathcal{F}_{N\times M}[\mathbf{Y}] \circ \mathcal{F}_{N\times M}[\mathbf{X}]^*}{|\mathcal{F}_{N\times M}[\mathbf{X}] \circ \mathcal{F}_{N\times M}[\mathbf{X}]^*|} \circ, \tag{7}$$

where $\mathbf{S} \in \mathbb{R}^{N\times M}$, and $\circ$ , $\circ - \circ$ denote element-wise (Hadamard) product and quotient, respectively. In order to implement a subpixel time shift estimator, $\mathbf{S}$ should be embedded in a suitably larger matrix of zeros $\mathbf{0} \in \mathbb{R}^{\kappa N\times \kappa M}$, where $\kappa > 1$ is an integer. Therefore, the inverse DFT will be performed on a larger dimension to compute the cross-correlation matrix via

$$\mathbf{R} = \mathcal{F}^{-1}_{\kappa N\times \kappa M}[\mathbf{S}] \in \mathbb{R}^{\kappa N\times \kappa M}. \tag{8}$$

Finally, we seek the cross-correlation maximum location, i.e.

$$(u, v) = \arg\max_{(n,m)} \mathbf{R}, \tag{9}$$

yielding the time delay estimator (for $\kappa N$ even, as usual)

$$\widehat{d}_{PhC} = \begin{cases} \dfrac{u}{\kappa f_s}, & u \in [0, \kappa N/2 - 1], \\ \dfrac{u - \kappa N}{\kappa f_s}, & u \in [\kappa N/2, \kappa N - 1]. \end{cases} \tag{10}$$

### 3.2. *Łukasiewicz Triangular Norm*

The second alignment technique considered in this work is based on *Łukasiewicz Triangular Norm* (LTN), selected for its simplicity and appealing properties[7], and defined as

$$T_L(x_0, y_0) = \max\{0, x_0 + y_0 - 1\}, \forall x_0, y_0 \in [0, 1]. \tag{11}$$

In order to implement a subpixel estimator, we interpolate the original TF map. To this end, we use the *bicubic interpolator*[8], denoted as $\mathcal{B}_{a\times b}[\cdot]$, where $a \times b$ is the dimension of the output matrix. Therefore, disregarding border effects, it is possible to define the estimator:

$$\widehat{d}_{LTN} = \frac{1}{\kappa f_s} \arg\max_u \sum_{k=1}^{\kappa N} \sum_{j=1}^{M} T_L \left( X_b(k, m), Y_b(k + u, m) \right), \tag{12}$$

where $X_b(k, m)$ and $Y_b(k + u, m)$ are elements of the interpolated matrices $\mathbf{X_b} = \mathcal{B}_{\kappa N\times M}[\mathbf{X}]$ and $\mathbf{Y_b} = \mathcal{B}_{\kappa N\times M}[\mathbf{Y}]$, respectively.

Note that it is possible to speed up the algorithm (otherwise rather slow) by finding a raw estimate of $d$ without performing the interpolation step, and then refining it via the subpixel version in Eq. (12) by considering only values of $d$ close to the raw estimate.

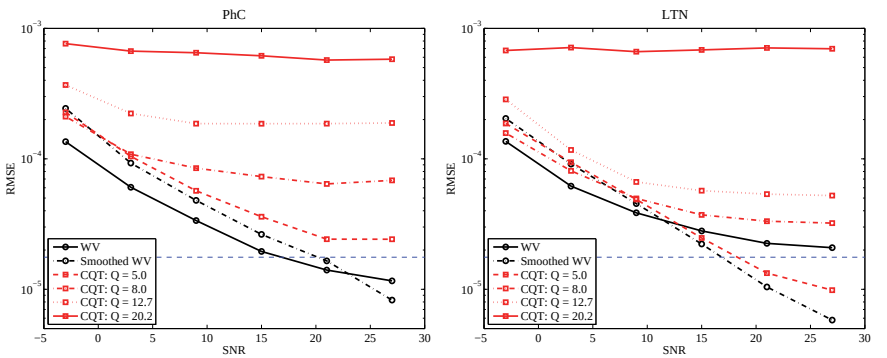



Fig. 2. RMSE for different time delay estimators, namely $\widehat{d}_{PhC}$ (left plot) and $\widehat{d}_{LTN}$ (right plot). The blue dashed lines refer to the quantization error $= 1/(f_s\sqrt{12})$. Other relevant parameters are the number of Monte Carlo trials ($N_{MC} = 100$) and the coefficient $\kappa = 100$ in Eq. (10) and Eq. (12).

## 4. Numerical Results

In this section we discuss the performance of the above TF-based time delay estimators via numerical simulations. We generate two chirping signals $s_x(t)$ and $s_y(t)$ differing by a time-delay $d$, using the *IMRPhenom* model[9], whose relevant parameters are the companion masses $M_1$ and $M_2$, and the *effective spin* $\chi_{\text{eff}}$. We set these parameters so as to fit the GW150914 signal [10]. Moreover, we set the observation interval $T = 1/8$ s, so that, for $f_s = 16384$ Hz, both discretized signals $s_x(n)$ and $s_y(n)$ consist of $N = 2048$ samples. Then, we add independent realizations of Gaussian white noise to the signal to obtain two noisy signals $x(n) = s_x(n) + w_x(n)$ and $y(n) = s_y(n) + w_y(n)$, with the same *Signal-to-Noise Ratio* (SNR), defined as $SNR = \sum_{n=0}^{N-1} s^2(n) / \sum_{n=0}^{N-1} w^2(n)$.

Finally, we compute the TF-maps of $x(n)$ and $y(n)$ and estimate the time delay $d$ using the techniques presented in the previous section. We gauge the estimation performance via standard Monte Carlo counting procedures, computing the Root Mean Square Error (RMSE), defined as

$$RMSE = \sqrt{\frac{1}{N_{MC}} \sum_{i=1}^{N_{MC}} (\widehat{d}_i - d)^2}, \tag{13}$$

where $N_{MC}$ is the number of Monte Carlo trials (here $N_{MC} = 100$) and $\widehat{d}_i$ is the time delay estimate for the $i$-th trial.

The results are shown in Fig. 2. We can see that, as expected, the sharpness of the TF representation is key for both PhC and LTN estimators. Indeed, the CQT performs poorly for high values of the Q factors (corresponding to lower time resolution), while the RMSE is significantly lower for low Q values. For low and medium SNR values, the uniform sharpness of the WV distribution has a clear advantage, but for high SNR its intermodulation artifacts cause performance degradation. This doesn't happen in the *smoothed* WV that, for high SNR, clearly outperforms the other TF representations considered. The combination *smoothed* WV-LTN outperforms PhC for high SNR values; on the other hand, for low and medium SNR values the combination WV-PhC yields the best performance.

## 5. Conclusion

In this paper we compared the performance of different Time-Frequency representations and alignment techniques for estimating the time delay between the GW signals received by two interferometers. The WV-based representations, as expected, outperform the CQT, due to their uniformly higher resolution. Moreover, the LTN is more effective than PhC for delay estimation, for high SNR values. even if its computational burden is higher. Future works will analyze more sophisticated TF maps capitalizing on *sparse* representations[11,12], and refined alignment techniques to further improve the DOA estimation.

**Acknowledgments**

This research has made use of data, software and/or web tools obtained from the Gravitational Wave Open Science Center (https://www.gw-openscience.org), a service of LIGO Laboratory, the LIGO Scientific Collaboration and the Virgo Collaboration. LIGO is funded by the U.S. National Science Foundation. Virgo is funded by the French Centre National de Recherche Scientifique (CNRS), the Italian Istituto Nazionale della Fisica Nucleare (INFN) and the Dutch Nikhef, with contributions by Polish and Hungarian institutes.

**References**

1. J. C. Brown, Calculation of a constant q spectral transform, *The Journal of the Acoustical Society of America* **89**, 425 (1991).
2. L. Cohen, Time-frequency distributions-a review, *Proceedings of the IEEE* **77**, 941 (Jul 1989).
3. R. G. Baraniuk and D. L. Jones, Signal-dependent time-frequency analysis using a radially gaussian kernel, *Signal Processing* **32**, 263 (1993).
4. M. Guizar-Sicairos, S. T. Thurman and J. R. Fienup, Efficient subpixel image registration algorithms, *Opt. Lett.* **33**, 156 (Jan 2008).
5. T. Dong, T. Deng, J. Dai, W. Xie, J. Yang and Q. Fu, A fuzzy reasoning method for multi-views image registration, *JCP* **8**, 2689 (2013).
6. M. Vallisneri, J. Kanner, R. Williams, A. Weinstein and B. Stephens, The ligo open science center, in *Proceedings of the 10th LISA Symposium, University of Florida, Gainesville, May 18-23, 2014; also arxiv:1410.4839*, 2014.
7. L. Troiano, L. J. Rodríguez-Muñiz, P. Marinaro and I. Díaz, Statistical analysis of parametric t-norms, *Inf. Sci.* **257**, 138 (February 2014).
8. R. Keys, Cubic convolution interpolation for digital image processing, *IEEE Transactions on Acoustics, Speech, and Signal Processing* **29**, 1153 (Dec 1981).
9. P. Schmidt, F. Ohme and M. Hannam, Towards models of gravitational waveforms from generic binaries: Ii. modelling precession effects with a single effective precession parameter, *Phys. Rev. D* **91**, p. 024043 (Jan 2015).
10. B. P. Abbott *et al.*, Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, p. 061102 (Feb 2016).
11. P. Flandrin and P. Borgnat, Time-frequency energy distributions meet compressed sensing, *IEEE Transactions on Signal Processing* **58**, 2974 (June 2010).
12. P. Addesso, M. Longo, S. Marano, V. Matta, M. Principe and I. M. Pinto, Compressed coding for time-frequency gravitational wave data analysis, *LIGO document P1200170* (2012).

# Polarization test of gravitational waves from compact binary coalescences

Hiroki Takeda*, Yuta Michimura, Kentaro Komori and Masaki Ando

*Department of Physics, University of Tokyo, Bunkyo, Tokyo 113-0033, Japan*
**E-mail: takeda@granite.phys.s.u-tokyo.ac.jp*

Atsushi Nishizawa

*Kobayashi-Maskawa Institute for the Origin of Particles and the Universe,Nagoya University, Nagoya, Aichi 464-8602, Japan*

Koji Nagano

*KAGRA Observatory, Institute for Cosmic Ray Research, University of Tokyo, Kashiwa, Chiba, 277-8582, Japan*

Kazuhiro Hayama

*Department of Applied Physics, Fukuoka University, Nanakuma, Fukuoka 814-0180, Japan*

The most general gravitational wave is composed of six polarization modes in metric theories of gravity. Polarization modes of gravitational waves can be used for the gravitational-wave tests of general relativity because the properties of polarization modes depend on the specific theory of gravity. We study the separability of the polarizations and the degeneracies between binary and polarization parameters for the inspiral gravitational waves from the compact binary coalescences.

*Keywords*: Gravitational waves, polarization, compact binary, modified gravity.

## 1. Introduction

The detection of gravitational waves from compact binary coalescences by Advanced LIGO and Advanced Virgo provides experimental approach to test general relativity [1-3]. In general metric theory of gravity, there are four non-tensorial polarization modes of gravitational waves in addition to two tensor modes allowed in general relativity [4]. Moreover, the number and the properties of the polarization modes reflect the nature of the theory of gravity. Thus, polarization test of gravitational waves make it possible to test general relativity. We focus on the polarization test of the inspiral gravitational waves from compact binary coalescences. We need to consider the source parameters, which are correlated with each other and determine the frequency evolution, to separate the polarizations from compact binary coalescences unlike some other waveforms. Polarization search by GW170814 had been conducted by the substitution of the antenna pattern functions [5]. However, the observational results of the polarizations may be affected by the inclination-angle dependence or the existence of other polarization modes. Thus, we study the separability of the polarizations and the degeneracies between binary and polarization parameters for the inspiral gravitational waves from the compact binary coalescences [6].

## 2. Angular dependence of a gravitational-wave waveform in modified gravity

In general relativity, the detector signal of I-th detector can be expressed as,

$$h_I = \frac{2}{5}\mathcal{G}_{T,I} h_{\rm GR}, \tag{1}$$

where $h_{\rm GR}$ is the gravitational-wave waveform in general relativity and $\mathcal{G}_{T,I}$ is the geometrical factor for the tensor modes defined by

$$\begin{aligned}\mathcal{G}_{T,I} := \frac{5}{2}\{&(1+\cos^2\iota)F_{+,I}(\boldsymbol{\theta_s},\boldsymbol{\theta_e}) \\ &+ 2i\cos\iota F_{\times,I}(\boldsymbol{\theta_s},\boldsymbol{\theta_e})\}e^{i\phi_{D,I}(\theta_s,\phi_s,\theta_e,\phi_e)}.\end{aligned} \tag{2}$$

Here $\iota$ is the inclination angle, $\boldsymbol{\theta_s} := (\theta_s, \phi_s, \psi_p)$ is the source direction angle parameters $(\theta_s, \phi_s)$ and polarization angle $\psi_p$, $\boldsymbol{\theta_e} := (\theta_e, \phi_e, \psi)$ is the detector location and orientation angle parameters, and $\phi_{D,I}$ is the Doppler phase. We calculated geometrical factors for non-tensorial modes having different inclination-angle dependence from those of tensor modes by quadrupole formula,

$$\mathcal{G}_{V_x,I} := \sqrt{\frac{525}{56}}\sin 2\iota F_{V_x,I}(\boldsymbol{\theta_s},\boldsymbol{\theta_e})e^{i\phi_{D,I}(\theta_s,\phi_s,\theta_e,\phi_e)}, \tag{3}$$

$$\mathcal{G}_{V_y,I} := \sqrt{\frac{15}{2}}\sin\iota F_{V_y,I}(\boldsymbol{\theta_s},\boldsymbol{\theta_e})e^{i\phi_{D,I}(\theta_s,\phi_s,\theta_e,\phi_e)}, \tag{4}$$

$$\mathcal{G}_{S_2,I} := \sqrt{\frac{225}{8}}\sin^2\iota F_{b,I}(\boldsymbol{\theta_s},\boldsymbol{\theta_e})e^{i\phi_{D,I}(\theta_s,\phi_s,\theta_e,\phi_e)}. \tag{5}$$

The inclination-angle dependence for the scalar dipole radiation is proportional to $\sin\iota$ in modified gravity theories with a scalar degree of freedom [7],

$$\mathcal{G}_{S_1,I} := \sqrt{\frac{45}{2}}\sin\iota F_{b,I}(\boldsymbol{\theta_s},\boldsymbol{\theta_e})e^{i\phi_{D,I}(\theta_s,\phi_s,\theta_e,\phi_e)}. \tag{6}$$

## 3. Polarization model

We adopt polarization models using the above geometrical factors and adding additional polarization amplitude parameters $A$s. The followings are examples of our polarization models.

Model TS1 is a tensor-scalar dipole model in which a scalar mode having the inclination-angle dependence of dipole radiation is added.

$$h_I = \{\mathcal{G}_{T,I} + A_{S_1}\mathcal{G}_{S_1,I}\}h_{\rm GR}. \tag{7}$$

Model TS2 is a tensor-scalar quadrupole model in which a scalar mode having the inclination-angle dependence of quadrupole radiation is added.

$$h_I = \{\mathcal{G}_{T,I} + A_{S_2}\mathcal{G}_{S_2,I}\}h_{\rm GR}. \tag{8}$$

Model TV is a tensor vector model in which the combination of vector x and vector y mode is added.

$$h_I = \{\mathcal{G}_{T,I} + A_{V_x}\mathcal{G}_{V_x,I} + A_{V_y}\mathcal{G}_{V_y,I}\}h_{\rm GR}. \tag{9}$$

## 4. Results

We estimate model parameters for 500 BBHs or 500 BNSs whose angular parameters are uniformly random, with a detector network such as three detectors aLIGO-AdV(HLV) or four detectors aLIGO-AdV-KAGRA(HLVK) in each polarization model by Fisher analysis. We use the inspiral waveform up to 3 post-Newtonian order in amplitude and 3.5 post-Newtonian order in phase as $h_{\rm GR}$, and consider 11 model parameters in general relativity and additional polarization amplitude parameters. Table 1 is the results of the medians of parameter estimation errors and correlation coefficients. We say that the polarization modes would be separable when the errors of the additional polarization amplitude parameters are less than unity because their fiducial values are set to be unity.

Table 1. Medians of parameter estimation errors and correlation coefficients. Here, $d_L$ is the luminosity distance of the source and $\Delta\Omega_s$ is the sky localization error. Masses of BBH and BNS are $10M_\odot - 10M_\odot$ and $1.4M_\odot - 1.4M_\odot$, respectively. Only correlation coefficients larger than 10% are shown.

| | parameter | BBH(HLV) | BBH(HLVK) | BNS(HLV) | BNS(HLVK) |
|---|---|---|---|---|---|
| | SNR | 33.3 | 40.2 | 36.4 | 44.3 |
| Model TS1 | $\Delta \ln d_L$ | 0.678 | 0.179 | 0.359 | 0.134 |
| | $\Delta\Omega_s[\mathrm{deg}^2]$ | 4.74 | 0.912 | 0.919 | 0.250 |
| | $\Delta A_{S1}$ | 1.16 | 0.284 | 0.606 | 0.197 |
| | $C(A_{S1}, \log d_L)$ | 0.998 | 0.989 | 0.996 | 0.984 |
| | $C(A_{S1}, \cos\iota)$ | -0.553 | -0.500 | -0.231 | -0.159 |
| Model TS2 | $\Delta \ln d_L$ | 0.676 | 0.182 | 0.358 | 0.134 |
| | $\Delta\Omega_s[\mathrm{deg}^2]$ | 4.74 | 0.913 | 0.862 | 0.246 |
| | $\Delta A_{S2}$ | 1.51 | 0.385 | 0.765 | 0.256 |
| | $C(A_{S2}, \log d_L)$ | 0.997 | 0.989 | 0.996 | 0.984 |
| | $C(A_{S2}, \cos\iota)$ | -0.609 | -0.564 | -0.246 | -0.189 |
| Model TV | $\Delta \ln d_L$ | 1.98 | 0.310 | 1.22 | 0.193 |
| | $\Delta\Omega_s[\mathrm{deg}^2]$ | 5.68 | 0.795 | 0.813 | 0.187 |
| | $\Delta A_{V_x}$ | 2.55 | 0.420 | 1.37 | 0.241 |
| | $\Delta A_{V_y}$ | 3.91 | 0.513 | 2.12 | 0.298 |
| | $C(A_{V_y}, \log d_L)$ | 0.999 | 0.993 | 0.998 | 0.991 |
| | $C(A_{V_y}, \cos\iota)$ | -0.846 | -0.335 | -0.307 | -0.207 |
| | $C(A_{V_x}, A_{V_y})$ | 0.987 | 0.814 | 0.948 | 0.624 |

In all models, the additional polarization amplitude parameters are strongly correlated with the amplitude parameters such as the luminosity distance and the inclination angle. In the model TS1 and TS2 having three polarizations, $A_S$ is determined even by three detectors HLV for BNS. However, it is difficult to separate the additional polarization mode by three detectors HLV for BBH due to the short

duration of the signal although the number of the detectors is equal to the polarization modes. In the model TV, the errors of the amplitude parameters are larger than unity with HLV for both BBH and BNS so that four detectors are always necessary to determine two additional polarizations.

Thus, we found that at least the same number of detectors is necessary to separate the polarization modes and obtain the polarization information of gravitational waves in principle. However, even when the number of detectors is equal to the number of the polarization modes, the modes would be inseparable in some cases, depending on the correlation among the amplitude parameters. Therefore, there are two conditions for the separation of polarization modes; (i) the same number of detectors or more as the number of polarization modes and (ii) significant signal-to-noise ratio (SNR) and the long duration of the signal.

## 5. Discussions

The second condition for the separation of polarizations (ii) indicates that the duration of the signal or frequency band of the detector has essential role in polarization test. Here, we use the networks composed of the second generation gravitational-wave detectors. However, a single next generation gravitational-wave detector such as the Einstein Telescope and the Cosmic Explorer could test several polarization modes because they have great sensitivity at lower frequency so that they could be effectively treated as a detector network including a set of detectors along its trajectory due to the Earth's rotation.

## Acknowledgments

H. T. and K. K. acknowledge financial support received from the Advanced Leading Graduate Course for Photon Science (ALPS) program at the University of Tokyo. H.T., A.N., K.N., K.K. and Y.M. are supported by JSPS KAKENHI Grant No. 18J21016., JSPS KAKENHI Grant No. JP17H06358., JSPS KAKENHI Grant No. 17J01176,JSPS KAKENHI Grant No. 16J01010 and JSPS Grant-in-Aid for Scientic Research(B) No. 18H01224, respectively.

## References

1. B. P. Abbott *et al.*, Physical Review Letters **116**, 221101 (2016).
2. B. P. Abbott *et al.*, Physical Review Letters **119**, 161101 (2017).
3. B. P. Abbott *et al.*, arXiv:astro-ph/1811.12907 (2018).
4. D. M. Eardley *et al.*, Physical Review Letters **30**, 884 (1973).
5. B. P. Abbott *et al.*, Physical Review Letters **119**, 141101 (2017).
6. H. Takeda *et al.*, Physical Review D **98**, 022008 (2018).
7. K. Chatziioannou, N. Yunes, and N. Cornish, Physical Review D **86**, 022004 (2012).

# Follow-up efforts of multi-messenger events and observational strategies with the MAGIC telescopes

Michele Peresano* on behalf of the MAGIC Collaboration

*Universit di Udine, and INFN Trieste, I-33100 Udine, Italy*
** E-mail: peresano.michele@gmail.com*

The MAGIC Collaboration operates two Imaging Atmospheric Cherenkov Telescopes (IACTs) which investigate the gamma-ray sky from 50 GeV to 50 TeV. In recent years, the ever-increasing development of new measurement instruments and the new synergies between research teams world-wide have contributed to the beginning of the multi-messenger era. The performance characteristics of the MAGIC telescopes make them well suitable for multi-messenger studies: numerous follow-ups of gravitational and neutrino alerts have been performed, along with the historic detection of the first neutrino-blazar association TXS 0506+56. In this context, since recent years it is also possible for external scientists to contribute to the MAGIC Collaboration observations and scientific output.

*Keywords*: Multi-messenger; neutrino; gravitational waves; MAGIC.

## 1. Introduction

The era of multi-messenger astronomy began formally already in the middle of the 20th century, with the study of cosmic rays and neutrinos from the Sun. It then progressed a few decades later progressed with the neutrino detection of SN1987A. In 2017 it acquired a more widespread and public interest. The discovery of gravitational waves (GW) from the neutron star merger GW170817 triggered a multi-messenger observational campaign. Such a multi-collaboration effort found another application also in the case of the first neutrino-blazar association of TXS 0506+056.

In the case of GWs, interest has been fueled by the possibility of detecting electromagnetic counterparts, required to constrain nature and physics of sources emissions - GW are not affected by gamma-ray attenuation during intergalactic propagation at Very-High-Energies (VHE). This happened only after the first pure GW signals reached the LIGO detectors (first binary black hole mergers - GW150914[1] and GW151226[2]).

Neutrinos have been studied for a longer time, since they represent a neutrally charged messenger which is also weakly interacting with matter. These properties made neutrinos the perfect tool for studies aiming to pinpoint sources responsible for Ultra-High-Energy (UHE) cosmic-rays acceleration. The interplay between neutrinos and electromagnetic radiation - in particular VHE photons - is critical in order to link the presence of neutrinos in a source's total emission to the production of cosmic rays at VHE and UHE.

Multi-messenger astronomy represents also a milestone in a multi-collaborative effort between many observatories and experiments world-wide. The MAGIC telescopes have been part of this community for many years, putting effort in both neutrino and GWs follow-ups.

## 2. The MAGIC telescopes and the IACT technique

The MAGIC collaboration is composed by more than 200 members from more than 10 countries, and operates a pair of Imaging Atmospheric Cherenkov Telescopes (IACTs) on the island of La Palma (Canary Islands, Spain). These instruments aim to detect and study astrophysical sources responsible for the emission of Very-High Energy (VHE) radiation.

Such phenomena cannot be properly investigated with sufficient statistics and precision by space-borne instruments, given the prohibitive energy domain of the emitted particles. In particular, due to their limited sizes, any material comprising scientific payloads on board of satellites will suffer from limited radiation lengths and collection areas at increasing energies. This effect results in a progressive decrease in the triggering rates of events with energies over hundreds of GeV.

The Imaging Cherenkov Atmospheric technique (IACT) aims to solve this issue by exploiting the atmosphere as part of the measuring process. In this way it is possible to gain more powerful stopping power and greater effective area at higher energies.

Instruments like MAGIC detect VHE primary particles which impact the Earth's atmosphere by observing the product of such impacts: cascade showers. Any primary particle (gamma rays, electrons or hadrons) will produce an associated shower, which develops through the atmospheric medium due to successive diffusion between secondary particles and nuclei. In the case of IACT facilities, the signal is constituted often by primary gamma rays, whereas the remaining particle populations are to be considered background.

Instead of detecting directly the secondary products of each shower, the IACT technique aims to detect the Cherenkov radiation emitted as a result of the propagation in the atmosphere of such particles. This radiation is collected by each IACT telescope in a camera, after the shower's image is projected on a reflecting surface. Each camera is composed of more than 1000 pixels, each represented by a photo-multiplier (PMT) which converts light signals into electrical impulses. Starting from this input, each event's image is properly cleaned, removing background features which will depend on the observing conditions. Subsequently, a statistical analysis is performed in order to retrieve, energy, direction and nature of the primary particle (gamma ray or hadron), given the cascade shower it produced. All the analysis chain is dependent on specific MonteCarlo simulations covering both the instrument's performance and a model of the atmosphere at the observing site.

Modern IACT facilities are composed of more than one of such instruments, providing increased precision in the estimation of the shower's directions which in turn results in competitive angular resolutions and collection areas at VHE.

The current setup of the MAGIC telescopes is the result of an upgrade performed in 2011-2012[3], which produced an identical configuration between the two telescopes, operating in stereoscopic mode since 2009. Multi-messenger studies have in particular few critical requirements, which are covered at VHE by the current

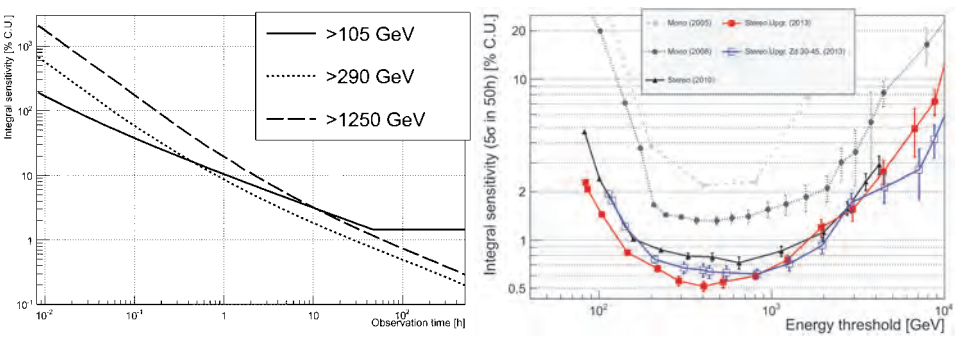
MAGIC performance:

- fast repositioning,
- low energy thresholds,
- high sensitivities within hour timescales.

In Table 1 we emphasise the MAGIC performance features more relevant for multi-messenger studies.

Table 1. Summary of the main features of the MAGIC telescopes and performance relevant for multi-messenger studies.

| Hardware features | Performance |
|---|---|
| 3.5° Field Of View | Angular resolution of 0.06° @ 1TeV - 0.1° at 100 GeV |
| Low voltage photo-multipliers | moderate / strong moonlight operations |
| $\sim 480m^2$ total reflective surface | Energy resolution of 15% at 1TeV and 23% at 100 GeV |
| Lowest structural weight | Fast repositioning up to 7 deg/s |
| Standard and "Sum" trigger systems | Energy range from $\lesssim$50 GeV to $\sim$50 TeV |




(a) (b)

Fig. 1. See Ref. 4 for details. (a) Dependence of the integral sensitivity of the MAGIC telescopes against observation time for different energy thresholds. (b) Evolution of the integral sensitivity of the MAGIC telescopes as a function of energy threshold for different configurations of the system.

Such characteristics result in a sensitivity of $\sim$ 10% of the Crab Nebula flux in 1 hour above 100 GeV ($\sim$ 0.66% Crab, $5\sigma$ in 50 hours above 220 GeV) as shown in Fig. 1(a) and Fig. 1(b).

Particularly relevant for multi-messenger studies have been the continuous improvements of the MAGIC telescopes in increasing their duty cycle. This has been done not only by progressively reducing technical and operational delays each year, but also by developing new hardware and software resources in order to allow observations during higher Night-Sky Background conditions, such as moderate and strong moonlight.

The effect that this kind of observing conditions have on registered data results mainly in a higher energy threshold, due to the higher cleaning levels required to extract the shower signal from a stronger NSB. Depending on moonlight intensity, the MAGIC telescopes can compensate by reducing the high voltage (HV) in the PMTs (which in turn reduces their gain by $\sim$1.7 times) and by inserting an additional UV-pass filter in front of both cameras.

The overall performance in such conditions has been recently described in Ref. 5 and results in a sensitivity which is still competitive, even if affected by unavoidable degradations dependent on the single setups:

- $\lesssim 10\%$ for nominal HV,
- $\lesssim 15-30\%$ for nominal HV,
- $\lesssim 60-80\%$ when using UV-pass filters.

## 3. Multi-messenger astronomy with the MAGIC telescopes

### 3.0.1. *Neutrinos*

Since 2012 MAGIC is part of the Gamma-Ray Follow-up program[6]. Only during the last observation cycle in 2017 the MAGIC telescopes have invested more than 30 hours on mainly two different types of neutrino follow-up searches from IceCube astrophysical neutrinos:

- a persistent source searches of $\nu_\mu$ tracks (HESE-37, HESE-38 and a multi-PeV track[7]),
- follow-ups of HESE/EHE real-time alerts (HESE-160427 and HESE/EHE-160731A).

A visualisation of both population of alerts observed by the MAGIC telescopes is shown in Fig. 2. No significant VHE gamma-ray signal has been detected by MAGIC for such events, which have been described in Ref. 8.

A major historical milestone in multi-messenger astrophysics was achieved by MAGIC being the first VHE instrument to detect gamma rays from a direction consistent with a HE neutrino. The blazar TXS 0506+056 became the first known source to emit both messengers.

Given the performance of MAGIC over neutrino follow-ups, it has been also possible to preliminary constrain the flux of astrophysical neutrino sources at at $\lesssim 0.1\%$ of the diffuse flux of neutrinos (3 years of Icecube data) - such result is comparable to the recent limits from IceCube in the Northern Sky[9].

### 3.0.2. *Gravitational waves*

MAGIC is has been also part of MoU with LIGO/Virgo, aimed to identify and follow gravitational wave candidate events since 2014. This kind of follow-ups are still limited by the difficult search for the possible EM counterpart, due to the

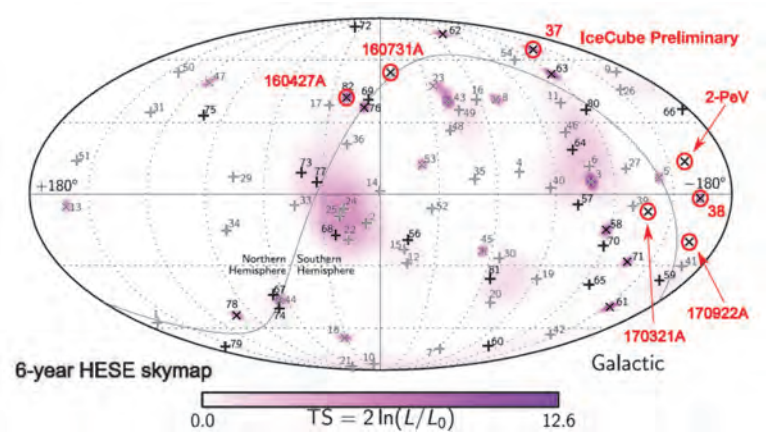



Fig. 2. Arrival directions of the HESE/EHE real-time alerts and $\nu_\mu$ tracks searches performed by the MAGIC collaboration, superimposed over a 6-years point-source clustering made by Icecube. Credits: Berti A., adapted from Ref. 10.

limited field of view of IACTs and the extent of the localisation areas of the available GW facilities (which reached a sensible improvement with the addition of the Virgo detector to the LIGO ones).

An important milestone for the Collaboration and the IACT community has been set during the Observation-1 (O1) cycle of LIGO, when the MAGIC telescopes performed the first ever IACT ToO observation of GW151226.

The case of GW151226 has been of particular importance for MAGIC, IACT and GW communities because it represented the first experiment of multi-messenger communication between these two kinds of facilities. Delays due to e.g. the interplay between the different types of alerts (GCN and MAGIC notices) and duty cycle (bad weather in particular), contributed to the development of early strategies to select LIGO sky regions within maximum source probability confidence level, in particular:

- choice of the visibility region area depending on a chosen probability threshold (50%, 90%),
- additional observations of other EM partners,
- overlap with existing catalogs[11].

## 4. Latest developments and future perspectives

These multi-messenger efforts have encouraged the members and working groups of the MAGIC Collaboration to develop and adapt new ideas and apply them to software, hardware and strategies:

- Off-axis upper limit calculation on flux over sky-maps for multi-messenger related alerts,

- GW follow-up observational strategies with selection of potential host galaxies from catalogs and application of emission models to estimate detectability,
- update of the MAGIC Automatic Alert system to improve prompt response,[12]
- plans to apply "Sum-trigger" to future automatic follow-ups in order to reach a lower energy threshold,
- Very Large Zenith Angle observations will expand the available observational limits while shifting sensitivity to higher energies[13].

In 2016 the MAGIC Collaboration started a program aimed at fostering scientific cooperation with external scientists. Each year this program is renewed with usual deadline around October. Interested individuals or teams are invited to visit the web page `https://magic.mpp.mpg.de/outsiders/magicop` and make contact as soon as possible.

The MAGIC telescopes continue to monitor the VHE sky with continuous developments on both hardware and software, keeping up with the latest discoveries.

**Acknowledgements**

The author would like to thank the organizers and the session's Chairs for the invitation.

**References**

1. Abbott B.P. et al. (2016), Phys. Rev. Lett. 116, 061102.
2. Abbott B.P. et al. (2016), Phys. Rev. Lett. 116, 241103.
3. Aleksi, J. et al. (2015) Astroparticle Physics, 72, 6175.
4. Aleksi, J. et al. (2015) Astroparticle Physics, 72, 7694.
5. Ahnen, M. L. et al. (2017) Astroparticle Physics, 94, 2941.
6. The IceCube, MAGIC and VERITAS collaborations (2016), JINST, 11, P11009.
7. Schoenen, S. and Raedel, L. on behalf of the IceCube Collaboration, Astronomer's Telegram (ATel) #7856.
8. Santander, M. et al., PoS (ICRC 2017) 618.
9. Aartsen M. G. et al., Astrophys. J., 835 (2017) no. 2, 151.
10. Kopper, C. et al., PoS (ICRC 2015) 1081.
11. Stamerra A. et al. on behalf of the MAGIC collaboration, GCN CIRCULAR #18776.
12. Berti et al. for the MAGIC collaboration, Marcell Grossman 15 (these proceedings).
13. Mirzoyan, R., et al. (2018) NIMA RICH10 conference proceedings in press (arxiv:1903.04989).

# The unprecedented VHE $\gamma$-ray outburst of PKS 1510-089 in May 2016

Tomislav Terzić*[a], Michael Zacharias[b], Julian Sitarek[c], Manuel Meyer[d], Dijana Dominis Prester[a], Felix Jankowsky[e], Elina Lindfors[f], Mahmoud Mohamed[e], David Sanchez[g], for the H.E.S.S. and MAGIC Collaborations

[a] *Croatian MAGIC Consortium, Rudjer Boskovic Institute, University of Rijeka, University of Split, University of Zagreb and University of Osijek, Croatia*
[b] *Centre for Space Research, North-West University, 2520 Potchefstroom, South Africa*
[c] *University of Lodz, 90236 Lodz, Poland*
[d] *KIPAC, Stanford University, Stanford, USA*
[e] *Landessternwarte, Universität Heidelberg, Königstuhl, 69117 Heidelberg, Germany*
[f] *Finnish MAGIC Consortium, Tuorla Observatory, University of Turku and Astronomy Division, University of Oulu, Oulu, Finland*
[g] *Laboratoire d'Annecy-le-Vieux de Physique des Particules, Universite Savoie Mont-Blanc, CNRS/IN2P3, 74941 Annecy-le-Vieux, France*
* *University of Rijeka, Departent of Physics, Radmile Matejčić 2, Rijeka, HR-51000, Crotia, E-mail: tterzic@phy.uniri.hr*

PKS 1510-089 ($z = 0.361$), one of only a handful of flat spectrum radio quasars detected in the very high energy (VHE, $E > 100$ GeV) $\gamma$-rays, is known for its flux variability and complex multiwavelength behaviour. VHE observations by H.E.S.S. and MAGIC in May 2016 detected an unprecedented flare, both in intensity and in the shortness of its variability timescale. The flare lasted less than 48 hours, during which time the flux reached about 80 per cent of the Crab Nebula flux above 200 GeV. In addition, the intranight variability of this source was detected for the first time. Simultaneous observations in high energy (HE, $E > 100$ MeV) $\gamma$-rays performed with Fermi-LAT and optical $R$-band performed with ATOM show behaviour not consistent with simple simultaneous brightening in all bands. While a significant hardening of the spectrum is visible in HE, the flux increased only moderately. A simultaneous rise in daily-averaged optical flux was seen in the $R$-band. However, the intranight $R$-band flux evolution shows two prominent peaks, while only one is visible in the VHE range. These intriguing features of the flare will be presented in detail. We will also discuss possible explanations for the observed emission.

*Keywords*: Gamma-rays; multiwavelength; blazars; PKS 1510-089.

## 1. Introduction

PSK 1510-089 is a flat spectrum radio quasar located at RA=$15^{\rm h}12^{\rm m}52.2^{\rm s}$, Dec=$-09°06'21.6''$ (J2000), with the redshift of $z = 0.361$. It is a highly variable source of electromagnetic radiation in all energy bands, exhibiting a complex MWL behaviour.[1–4] It was first detected in the very high energy (VHE, $E > 100$ GeV) $\gamma$-ray band with High Energy Stereoscopic System (H.E.S.S., see Section 2.1) in 2009[5] Major Atmospheric Gamma-Ray Imaging Cherenkov (MAGIC, see Section 2.1) detected VHE $\gamma$-ray signal from PSK 1510-089 in 2012.[6] The first detection of variability in the VHE band occurred during an outburst in 2015.[7,8] The flux was variable on a daily time-scale. An exceptional VHE $\gamma$-ray outburst of PKS 1510-089 in May 2016 was observed both with H.E.S.S. and MAGIC, revealing intra-night variability for the first time, on which we report here. We also present results in the high energy (HE, $E > 100$ MeV) $\gamma$-rays and optical band.

## 2. Data analysis and results

### 2.1. *VHE γ-ray band*

H.E.S.S. is an Imaging Atmospheric Cherenkov Telescope (IACT) array located in the Khomas Highland in Namibia at an altitude of about 1800 m a.s.l. It consists of four 12-m telescopes (CT1-4) and a 28-m telescope (CT5) in the centre of the array. The energy threshold of the system under optimal conditions is $\sim 50$ GeV. However, because of technical problems in the observation period, only CT2 − 4 data were used in the present analysis, resulting in higher energy threshold of $\sim 200$ GeV. H.E.S.S. observed PSK 1510-089 from MJD 57535 to 57545. A total of 31 runs (about 28 min per run) passed the standard quality selection[9] resulting in a total live time of 13.6 h. The data were analysed with the Model analysis chain using loose cuts.[10] The results were cross-checked and verified using the independent reconstruction and analysis chain ImPACT.[11]

The H.E.S.S. light curve for the entire observation period (MJD 57535 − 57545) with nightly averaged flux above 200 GeV is shown with red points in panel (a) of Fig. 1. The signal was significantly detected only on the nights of MJD 57537/8 and 57538/9. The average flux above 200 GeV on the night of the highest flux (MJD 57538) was $(14.3\pm0.6)\times10^{-11}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, corresponding to 56% of the flux from Crab Nebula above the same energy threshold (0.56 C.U.).[12] This was $\sim 15$ times higher than what H.E.S.S. measured in 2015.[8] We fitted the run-by-run based light curve with a constant. The hypothesis of constant flux was rejected at the level of $> 10\,\sigma$. On the night of the highest flux the source was observed for 4 runs, totalling in 1.8 h of data. The light curve with flux above 200 GeV is shown with red points in panel (a) of Fig. 2. Again the constant flux is ruled out on a run-by-run basis with $5.4\,\sigma$, confirming the first ever detection of VHE intra-night variability in PSK 1510-089. The flux above 200 GeV reached $(20 \pm 1) \times 10^{-11}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$, equivalent to 0.8 C.U. The spectrum is best fitted with a power-law (PL) $F(E) = N_0 \times (E/E_0)^{-\Gamma}$ folded with exponential function representing $\gamma$-ray flux attenuation due to the extragalactic background light (EBL).[13] Spectral parameters are as follows: normalisation flux $N_0 = (19.0 \pm 0.8) \times 10^{-10}\,\mathrm{TeV}^{-1}\,\mathrm{cm}^{-2}\,s^{-1}$, spectral index $\Gamma = 2.9 \pm 0.2$ and decorrelation energy $E_0 = 268$ GeV.[12] The intrinsic spectrum is shown in Fig. 3 with red points.

MAGIC is a stereoscopic system located in Canary Island of La Palma, at the height of 2200 m a.s.l.[14] It consists of two identical 17-m IACTs. The standard trigger threshold of the MAGIC telescopes for low zenith angle observations is $\sim 50$ GeV.[15] However, from the MAGIC location, PSK 1510-089 is observable at zenith angles above 38°, resulting in higher trigger threshold of $\sim 90$ GeV for the present analysis. Nevertheless, we integrated flux above 200 GeV, in order to have both H.E.S.S. and MAGIC light curves in the same energy range. MAGIC observed PSK 1510-089 on MJD 57535/6 and for five consecutive nights between MJD 57538/9 and 57542/3. A total of 7.5 h of data were collected for the whole observation period, while 2.7 h of data were collected on the night of flare. For

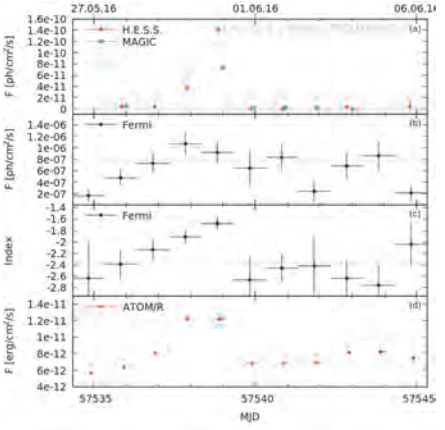



Fig. 1. PSK 1510-089 light curves for the MJD 57535 − 57545 period. **(a)** Nightly-averaged flux above 200 GeV measured with H.E.S.S. (red) and MAGIC (green). **(b)** Flux above 100 MeV averaged over 24 h and centred on VHE observations as measured by *Fermi*-LAT. **(c)** *Fermi*-LAT spectral index above 100 MeV obtained from integration over 24 h assuming PL spectral shape. **(d)** Nightly-averaged optical light curve in *R*-band from ATOM. Plot adopted from Ref. 12.

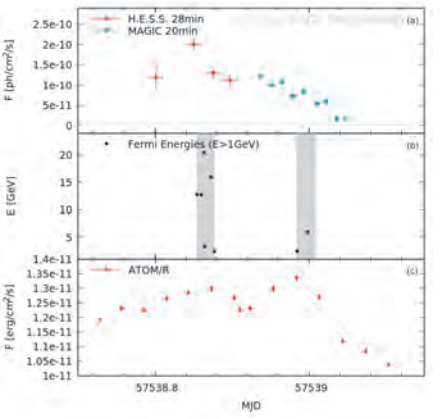



Fig. 2. PSK 1510-089 light curves for the night of the flare (MJD 57538/9). **(a)** Flux measured with H.E.S.S. (red) and MAGIC (green) above 200 GeV. **(b)** *Fermi*-LAT detected photons with $E > 1$ GeV. Grey bands represent LAT visibility windows of PSK 1510-089. **(c)** Optical light curve in *R*-band from ATOM. Each point represents an individual exposure of $\sim$ 8 min. Plot adopted from Ref. 12.

a short period during the observation on MJD 57538/9, partial cloudiness caused variable rates in the MAGIC data acquisition. Those data were excluded based on pyrometer measurements,[16] leaving the final data sample of 2.53 h. The data were analysed using the the MAGIC analysis and reconstruction software (MARS).[15,17]

The MAGIC light curve for the entire observation period with nightly averaged flux above 200 GeV is shown with green points in panel (a) of Fig. 1. The signal was significantly detected only on the night of MJD 57538/9, when the average flux above 200 GeV was $(7.36 \pm 0.40) \times 10^{-11}\,\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$,[12] equivalent to 0.32 C.U. This was $> 5$ times higher than upper limits of the flux on the rest of the nights from this data set, $\sim 4$ times the flux MAGIC measured in 2015,[7] and as much as $\sim 20$ times the flux MAGIC measured in 2012.[6] Intra-night variability on MJD 57538/9 was also detected in MAGIC light curve, shown with green points in panel (a) of Fig. 2. Each point represents one 20-min run. A fit with a constant is ruled out with $> 10\,\sigma$. The flux changed from $\sim 0.5$ C.U. in the first to $\sim 0.075$ C.U. in the last run. The observed spectrum on the night of the flare was reconstructed between 60 and 700 GeV. Again it is best fitted with a PL folded with the EBL attenuation.[13] Spectral parameters are: $N_0 = (34.7 \pm 1.5) \times 10^{-10}\,\mathrm{TeV}^{-1}\,\mathrm{cm}^{-2}\,s^{-1}$, $\Gamma = 3.37 \pm 0.09$

and $E_0 = 175\,\mathrm{GeV}$.[12] The intrinsic spectral energy distribution (SED) is shown in Fig. 3 with green points.

All uncertainties quoted are statistical only, for both observatories, while the systematic uncertainty on the energy scale is 15%.

We calculated variability time-scale from $t_{var} = \frac{1}{2}(F_i + F_{i+1})(t_{i+1} - t_i)/|F_{i+1} - F_i|$.[18] Smallest $t_{var}$ is $18 \pm 5\,\mathrm{min}$ between $7^{th}$ and $8^{th}$ MAGIC points (see Fig. 2). Based on causality argument we set upper limit on the size of the VHE $\gamma$-ray emission region $R \leq c\delta t_{var}/(1+z) = 1.2 \times 10^{15}(\delta/50)\,\mathrm{cm}$, where $\delta$ is relativistic Doppler factor. Adopting $\theta_j \sim 0.2/\Gamma$ for the jet opening angle,[19] where $\Gamma$ is the relativistic Lorentz factor (here we use $\Gamma = \delta$), and assuming the $\gamma$-ray emission region fills the entire cross section of the jet, we estimated the distance of the emission region from the black hole (BH) to be $d \sim R/\theta_j \sim 3 \times 10^{17}(\delta/50)^2\,\mathrm{cm}$. It should be noted that both $\theta_j \sim 0.2/\Gamma$ and $\delta \sim 50$ are extreme values. Larger opening angle, or smaller Doppler factor would put the emission region closer to the BH. On the other hand, abandoning the assumption of filling the entire cross section of the jet, would allow the emission region to be located further down the jet.

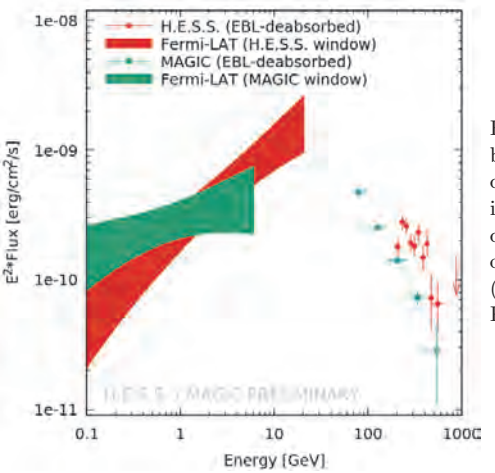



Fig. 3. PSK 1510-089 SED in HE and VHE band for the night of the flare. The H.E.S.S. data are given in red and the MAGIC ones in green points. The *Fermi*-LAT confidence regions are obtained by integrating over precise H.E.S.S. (red) and MAGIC (green) observation windows respectively. Plot adopted from Ref. 12.

### 2.2. *Multiwavelength picture*

We combined our data with observations in HE with *Fermi*-LAT and optical *R*-band observations with ATOM.

LAT is sensitive in the $0.02 - 300\,\mathrm{GeV}$ interval.[20] In an all-sky-survey mode, it scans the entire sky in 3 hours, while any point in sky is observed continuously for 30 min. We analysed publicly available Pass 8 SOURCE class events with energy $> 100\,\mathrm{MeV}$ in the region within 15° from the PSK 1510-089 position. We applied a zenith angle cut of $< 90°$. The analysis was performed with the ScienceTools

software package version v10r0p5 using the P8R2_SOURCE_V6 instrument response function and the GLL_IEM_v06 and ISO_P8R2_SOURCE_V6_v06 models for the Galactic and isotropic diffuse emission,[21] respectively. We produced a light curve for the whole period of VHE observations, shown in panel (b) in Fig. 1, where each point represents flux integrated over 24 h, centred on VHE observation. For this purpose, we assumed a PL spectral model with parameters free to vary between bins. The flux variability is apparent, however smaller than in the VHE band. Indeed the change is slower and the flux level significantly lower than the historical maximum[2] in this band.

Inspecting the spectral index on 24-hour scale, a hardening of the HE spectra with the increase of the VHE flux can be seen, reaching $1.7 \pm 0.1$, and then dropping to average value of $2.368 \pm 0.004$ after the flare (see Fig. 1c). On the night of the flare, LAT detected 8 photons with $E > 1\,\mathrm{GeV}$ (Fig. 2b); 6 of them in the H.E.S.S. observation window (4 with energies in the $10 - 25\,\mathrm{GeV}$ interval), when the VHE flux was higher, compared to the MAGIC observation window when only two photons were detected, both with energies below 10 GeV. Combined HE – VHE SED is shown in Fig. 3. The LAT confidence regions were obtained by integrating over precise VHE observation windows and calculated up to the highest detected energies for the respective period. The spectral indices were $\Gamma = 1.4 \pm 0.2$ and $\Gamma = 1.7 \pm 0.2$ in the H.E.S.S. and MAGIC windows, respectively. We calculated the difference in the spectral indices between HE and VHE spectra ($\Delta\Gamma$) and logarithm of ratio between extrapolated HE and measured VHE fluxes ($\tau = \ln(F_{extra}/F_{obs})$) for both observation windows. We obtained $\Delta\Gamma = 1.5 \pm 0.3$ and $\Delta\Gamma = 1.7 \pm 0.2$, and $\tau = 3.9 \pm 1.4$ and $\tau = 5.4 \pm 0.9$ for H.E.S.S. and MAGIC windows, respectively. The SED peak was located in the $10 - 60\,\mathrm{GeV}$ interval. Possible causes of the spectral break are absorption of $\gamma$-rays in the broad line region (BLR), but also intrinsic effects such as Klein-Nishina regime, break in electron spectrum etc. Our calculations put the emission zone (see Section 2.1) just outside of the BLR if $R_{BLR} = 2.6 \times 10^{17}\,\mathrm{cm}$[6] is adopted.

ATOM is a 75 cm optical telescope located on the H.E.S.S. site,[22] monitoring H.E.S.S. $\gamma$-ray sources. The magnitudes of each flux point were derived with differential photometry using five comparison stars in the same field of view. The resulting fluxes were corrected for Galactic extinction. Nightly averaged light curve (Fig. 1d) shows a similar flux evolution as in VHE band. However, a comparison on a finer time-scale for the flare night (Fig. 2c) shows double-peaked structure, in contrast to the VHE light curve shape.

## 3. Summary

H.E.S.S. and MAGIC observed an exceptionally strong and fast flare from PSK 1510-089 in VHE, detecting for the first time intra-night variability in this band. Although we saw no obvious counterparts in lower energies, a significant hardening of the spectrum was detected in the HE. A HE–VHE spectral break could

indicate a BLR absorption, however the shortest variability time-scale of $18 \pm 5$ min put the $\gamma$-ray emission region just outside of the BLR.

**Acknowledgments**

H.E.S.S. gratefully acknowledges financial support from the agencies and organisations listed at `https://www.mpi-hd.mpg.de/hfm/HESS/pages/publications/auxiliary/HESS-Acknowledgements-2019.html`.

MAGIC gratefully acknowledges financial support from the agencies and organisations listed at `https://magic.mpp.mpg.de/ack_201805/`.

**References**

1. A.M. Brown, *MNRAS* **431**, 824 (2013).
2. S. Saito, L. Stawarz, Y.T. Tanaka, T. Takahashi, G. Madejski, F. D'Ammando, *ApJL* **766**, L11 (2013).
3. S. Saito, L. Stawarz, Y.T. Tanaka, T. Takahashi, M. Sikora, R. Moderski, *ApJ* **809**, 171 (2015).
4. P. Kushwaha, S. Chandra, R. Misra, S. Sahayanathan, K.P. Singh, K.S. Baliyan, *ApJL* **822**, L13 (2016).
5. H.E.S.S. collaboration, Abramowski A., et al., *A&A* **554**, A107 (2013).
6. J. Aleksić, et al. (MAGIC Collaboration), *A&A* **569**, A46 (2014).
7. M.L. Ahnen, et al. (MAGIC collaboration), *A&A* **603**, A29 (2017).
8. M. Zacharias, M. Böttcher, N. Chakraborty, et al., for the H.E.S.S. collaboration, *Proc. of the GAMMA16* `arXiv:1611.02098` (2016).
9. F.A. Aharonian, et al. (HESS collaboration), *A&A* **457**, 899 (2006).
10. M. de Naurois & L. Rolland, *Astropar. Phys.* **32**, 231 (2009).
11. R.D. Parsons & J.A. Hinton, *Astropar. Phys.* **56**, 26 (2014).
12. M. Zacharias, et al. (HESS and MAGIC collaborations), *Proc of 35th ICRC, Busan, Korea*, 655 (2017).
13. A. Franceschini, G. Rodighiero, M. Vaccari, *A&A* **487**, 837 (2008).
14. J. Aleksić, et al. (MAGIC collaboration), *Astropar. Phys.* **72**, 61 (2016).
15. J. Aleksić, et al. (MAGIC collaboration), *Astropar. Phys.* **72**, 76 (2016).
16. M. Will, *EPJWC* **144**, 01002 (2017).
17. R. Zanin R., E. Carmona, J. Sitarek, et al., *Proc of 33rd ICRC, Rio de Janeiro, Brazil*, Id. 773 (2013).
18. Y.H. Zhang, A. Celotti, A. Treves, L. Chiappetti, et al., *ApJ* **527**, 719 (1999).
19. E. Clausen-Brown E., T. Savolainen, A.B. Pushkarev, Y.Y. Kovalev, J.A. Zensus, *A&A* **558**, A144 (2013).
20. W.B. Atwood, A.A. Abdo, M. Ackermann, et al., *ApJ* **697**, 1071 (2009).
21. F. Acero, M. Ackermann, M. Ajello, et al., *ApJS* **223**, 26 (2016).
22. M. Hauser, C. Möllenhoff, G. Pühlhofer, S.J. Wagner, H.-J. Hagen, M. Knoll, *AN* **325**, 659 (2004).

# On the potential of KM3 Neutrino Telescopes and Cherenkov Telescope Arrays for the detection of extended sources

Silvia Celli*

*Gran Sasso Science Institute, Viale F. Crispi 7, 67100, L'Aquila, Italy*
** E-mail: silvia.celli@gssi.infn.it*

The detection of a cosmic neutrino flux by the IceCube telescope triggers the search for the astrophysical accelerators responsible for it. Among them, Galactic sources are expected to contribute at some level: thanks to its location in the Northern hemisphere, KM3NeT is optimally suited to constrain their contribution with a clean event sample of upgoing muon tracks. Therefore, it is timely to investigate the discovery potential of KM3NeT to extended sources of very-high-energy neutrinos. The study presented is based on a comparative analysis of the sensitivity of KM3NeT and CTA. The methods are then applied to two interesting Galactic gamma-ray sources: the brightest TeV supernova remnant, RX J1713.7-3946, and the Galactic Center Ridge.

*Keywords*: Neutrino telescopes; Imaging atmospheric Cherenkov telescopes; Extended sources: RX J1713.7-3946, Galactic Center Ridge.

## 1. Introduction

The identification of the sources contributing to the diffuse neutrino flux detected by IceCube (Ref. 1), as well as the discovery of discrete sources of VHE neutrinos, are the major objective of neutrino astronomy for the coming years. In the feasible future, the upgraded IceCube and the planned KM3NeT will serve as the major tools of neutrino astronomy. The production mechanisms of very-high-energy (VHE) neutrinos are connected to the hadronic interactions of ultra-relativistic protons with the ambient gas and radiation. These processes are accompanied by the production at comparable rates of VHE gamma rays. The ground-based gamma-ray detectors, in particular the current arrays of Imaging Atmospheric Cherenkov Telescopes (IACTs) provide lower flux sensitivities for point-like sources around 1 TeV, compared to the sensitivities of present neutrino detectors, as IceCube, and forthcoming ones, as KM3NeT. With the exception of compact objects and sources located at cosmological distances, TeV gamma-ray fluxes should be taken as a robust criterion regarding the expectations of discovery of discrete VHE neutrino sources. Given the difference in the TeV flux sensitivities of IACT arrays and KM3-scale neutrino detectors, the gamma-ray fluxes are especially constraining for point-like sources. For mildly-extended sources with an angular size $\sim 1^{\circ}$, which is one order of magnitude larger than the point spread function (PSF) of IACTs but still comparable to the PSF of VHE neutrino detectors, the gamma-ray flux sensitivity degrades, while the flux sensitivity of neutrino detectors does not change significantly. Here, this question is studied based on the comparative analysis of the sensitivities of CTA and KM3NeT for extended sources. In Sec. 2 the procedure defined to compute the instrument sensitivity is defined, considering different sizes of the sources and analyzing the different impact they have on the sensitivity of these instruments.

As an application of this study, the case of two Galactic objects is considered, for which the gamma-ray and neutrino connection has been widely discussed in literature: the young supernova remnant (SNR) RX J1713-3946 is presented in Sec. 3, while the region of the Galactic Center Ridge is investigated in Sec. 4. Finally, conclusions are derived in Sec. 5.

## 2. Sensitivity studies to extended sources

A common procedure for gamma-ray and neutrino telescopes is here introduced for the computation of the sensitivity curves: the same analytical approach is applied, thus the calculated sensitivities of two detectors can be compared. Since the sensitivity curves have to be compared point by point, the same energy binning is used for both the gamma-ray and the neutrino sensitivities. Three bins per logarithmic decade are defined, so that the energy resolution of both instruments is covered in each bin. The minimum detectable flux is defined as the flux that gives in each energy bin: (i) a minimum number of signal events, $N_s^{min}$, (ii) a minimum significance level of background rejection, $N_\sigma^{min}$, and (iii) a minimum signal excess over the background uncertainty level. Thus, the instrument sensitivity is fixed by the one condition among the three listed above which dominates over the other two. The number of signal events, $N_s$, is obtained by folding an $E^{-2}$ power-law spectrum with the instrument response: these are reported in Fig. 1 for the CTA Southern Array configuration (4 large size telescopes, 24 medium size telescopes and 72 small size telescopes) and to the six building block KM3NeT in the observation channel of up-going muons.

In the following, $N_s^{min} = 10$ is set for CTA and $N_s^{min} = 1$ for KM3NeT. The significance level of the detection is expressed by the standard deviation $\sigma$, defined as $\sigma = N_s/\sqrt{N_b}$, where $N_b$ is the number of background events in the energy bin. The threshold on the minimum number of $\sigma$ is set to $N_\sigma^{min} = 5$ for CTA and to $N_\sigma^{min} = 3$ for KM3NeT. The values of $N_s^{min}$ and $N_\sigma^{min}$ are smaller in the case of neutrino telescopes in order to investigate the limits of the source detection capability. Concerning the condition on the background uncertainty, in the case of CTA a 1% systematic uncertainty on the modeling of the background is assumed and a signal of at least five times this background accuracy level is required, i.e. $N_s/N_b \geq 0.05$, following the approach adopted by CTA in Ref. 3. In the case of neutrinos, instead, the approach adopted by KM3NeT in Ref. 4 is followed and a 25% background systematic uncertainty is assumed, mainly related to uncertainty in the theoretical modeling of the atmospheric neutrino background. In the neutrino case, a signal of at least three times higher than the background accuracy is assumed: this converts into requiring a signal to background ratio of at least $N_s/N_b \geq 0.75$. An observation time of 50 hours is assumed for the CTA (a pointing of the source is required), while 10 years are assumed for KM3NeT.

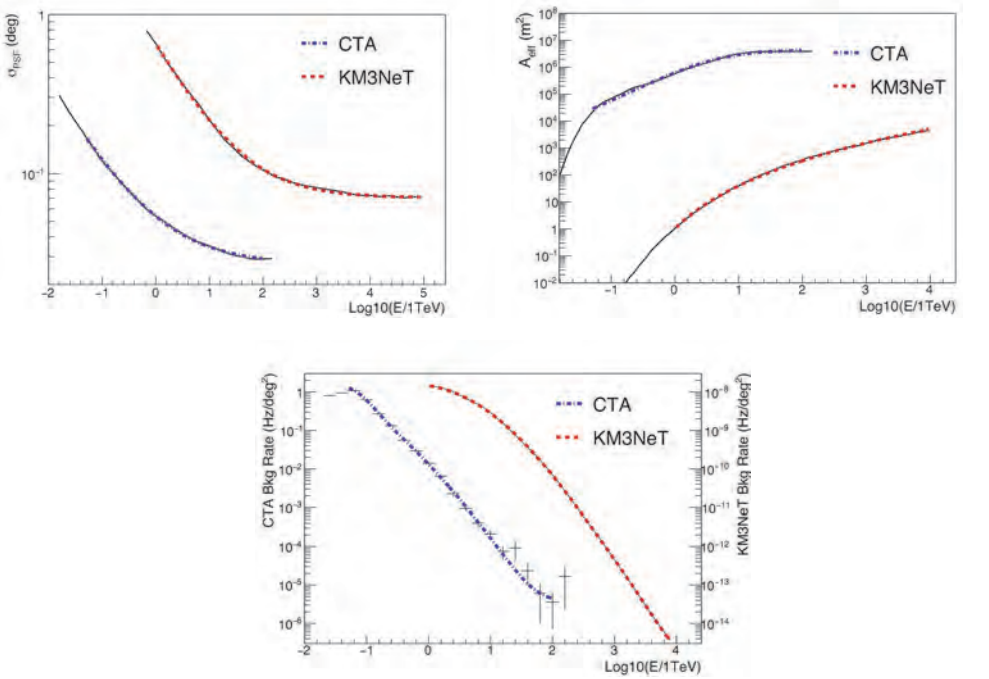
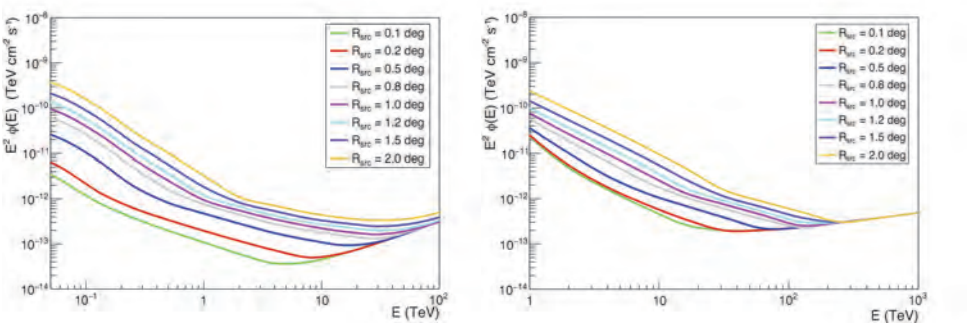



Fig. 1. Performances of the CTA and KM3NeT telescopes: angular resolution (top-left), effective area (top-right) and background rates per unit of solid angle (bottom). For both the telescopes, black lines correspond to the publicly available instrument responses, respectively for CTA Southern array (see Ref. 5) and KM3NeT (see Ref. 4). Dotted curves are the best fits, valid in the energy range $E \in [0.05 - 100]$ TeV for CTA and above 1 TeV for KM3NeT. The KM3NeT angular resolution is for $\nu_\mu$ charged current events and the muon neutrino effective area (six building blocks) corresponds to triggered events with a zenith angle greater than 80°, averaged over both $\nu_\mu$ and $\bar{\nu}_\mu$. Further details in Ref. 2.

The radial dimension of the source strongly affects the sensitivity of detection. Here eight different source sizes are considered, i.e. $R_{src} = [0.1, 0.2, 0.5, 0.8, 1.0, 1.2, 1.5, 2.0]$ deg. In the sensitivity computation, the angular resolution $\sigma_{PSF}$ affects the actual size of the observed region of interest (ROI). The radius of the ROI is defined as $R_{ROI} = \sqrt{\sigma^2_{PSF} + R^2_{src}}$. Spherical sources placed at the center of the field of view are considered, covering a solid angle $\Omega = \pi R^2_{ROI}$ for the background computation. The resulting sensitivity curves are shown in Fig. 2(a) for the gamma-ray telescope and in Fig. 2(b) for the neutrino telescope: note that these curves correspond to the differential sensitivities, thus the per bin sensitivity allows not only the identification of a source but also its spectroscopic analysis. Fig. 2 demonstrate that the deterioration of the sensitivity with source size shows an energy dependence for both instruments. In principle, a simple re-scaling of the point-source sensitivity according the actual extension of the source (i.e. through an energy-dependent scaling-factor proportional to $R_{ROI}/\sigma_{PSF}$), would predict a

stronger deterioration of the sensitivity for extended objects at energies at which the angular resolution is smaller. Thus, since the angular resolution is improved with energy, one would expect stronger effect at higher energies. Nevertheless, Fig. 2 does not show such a tendency for both telescopes. The reason is that at very high energies the detection of the signal proceeds at the very low background rate, thus the detection condition is determined by the signal statistics rather than by the background, i.e. by condition 1) listed in Sec. 2. Indeed, it is seen from Fig. 2 that the sensitivities become almost independent of the source extension at few tens of TeV for both gamma rays and neutrinos.




(a) (b)

Fig. 2. Minimum detectable flux computed according to the procedure described in Sec. 2 in the case of extended sources for: (a) 50 hour observation with CTA and (b) 10 year exposure with KM3NeT.

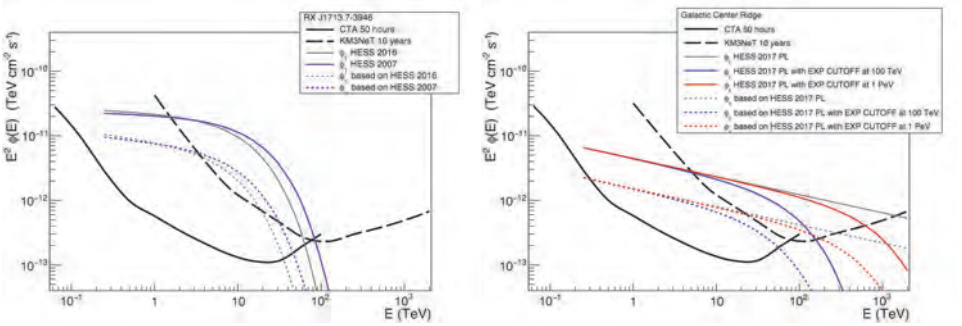
## 3. The case of RX J1713.7-3946

The case of such SNR is of great interest for neutrino telescopes, given that it is one of the brightest sources in the TeV sky. Moreover, its location in the sky makes it observable with up-going events at the latitude of KM3NeT for 70% of the time. The recent data from the H.E.S.S. Collaboration (see Ref. 6) suggest a spectrum in the form of a power-law with an exponential suppression as:

$$\frac{dN_\gamma}{dE}(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha} \exp\left[-\frac{E}{E_{\rm cut}}\right] \tag{1}$$

with $E_0 = 1$ TeV, $\phi_0 = 2.3 \times 10^{-11}$ TeV$^{-1}$ cm$^{-2}$ s$^{-1}$, $\alpha = 2.06$ and $E_{\rm cut} = 12.9$ TeV for the best-fit model. Note that the best-fit flux of the source, previously published by the H.E.S.S. Collaboration in Ref. 7, with the parameters $\phi_0 = 2.13 \times 10^{-11}$ TeV$^{-1}$ cm$^{-2}$ s$^{-1}$, $\alpha = 2.04$ and $E_{\rm cut} = 17.9$ TeV, predicts a noticeably higher flux of neutrinos at the most relevant energies for their detection, namely

$E \geq 10$ TeV. This is visible in Fig. 3(a), where the expected neutrino fluxes from both H.E.S.S. measurements are shown together with the flux sensitivities of the two instruments for a source with a radius of 0.6 deg (the $5\sigma$ minimum detectable flux is shown for CTA and the $3\sigma$ one for KM3NeT). High quality spectroscopic measurements of gamma rays at the highest energies are still missing, due to the limited sensitivity of current instruments. Apparently, this uncertainty will be substantially diminished by CTA for a rather short observation time. Remarkably, even for the lowest predicted neutrino flux, a statistically significant detection of the latter by KM3NeT seems realistic within a time scale of 10 years.




(a) (b)

Fig. 3. Minimum detectable flux computed according to the procedure described in Sec. 2 for CTA and KM3NeT to: (a) the extended SNR RX J1713.7-3946 (spherical source with radius of 0.6 deg) and (b) the diffuse emission from the Galactic Center Ridge (rectangular box with longitudinal size of 2.0 deg and latitudinal size of 0.6 deg). Solid lines refer to gamma-ray fluxes, dashed curves to muon neutrino fluxes, computed in the hypothesis that the corresponding gamma-ray flux results entirely from hadronic *pp* interactions (see Ref. 8).

## 4. The case of the Galactic Center Ridge

Another promising object from the point of view of detection of multi-TeV neutrinos is the Galactic Center Ridge. The observations with the H.E.S.S. telescope revealed the presence of a diffuse component in the emission from the 200 pc region of the Galactic Center (see Refs. 9, 10). The hard energy spectrum of the emission, as reported in Ref. 10, extends above 10 TeV without any indication of a spectral break or a cut-off. This condition makes such source another candidate to be detected by KMeNeT (see Ref. 11). The most recent H.E.S.S. measurements of the Galactic Center Ridge in Ref. 12 point toward an unbroken power-law spectrum of the form:

$$\frac{dN_\gamma}{dE}(E) = \phi_0 \left(\frac{E}{E_0}\right)^{-\alpha} \tag{2}$$

with $E_0 = 1$ TeV, $\phi_0 = 1.2 \times 10^{-8}$ TeV$^{-1}$ cm$^{-2}$ sr$^{-1}$ s$^{-1}$, $\alpha = 2.28$. This spectrum corresponds to the region $|l| \leq 1.0$ deg, $|b| \leq 0.3$ deg. The estimation of the

sensitivities is thus done for a rectangular box with longitudinal size of 2.0 deg and latitudinal size of 0.6 deg. The results are shown in Fig. 3(b), where three spectra of neutrinos are calculated: the unattenuated power law that closely follows the gamma-ray measurements and two more spectra obtained assuming an exponential cut-off in gamma rays at 100 TeV and 1 PeV. One can see that only in the case of location of the gamma-ray cut-off beyond 100 TeV, the gamma-ray data could guarantee a statistically significant detection of the counterpart neutrinos by KM3NeT. Gamma-ray data above 100 TeV are too weak to be detected by the H.E.S.S. telescopes, even after a decade of continuous monitoring of this region. The exploration of this energy domain requires more powerful gamma-ray instruments, such as CTA. This can be seen in Fig. 3(b). However, it is worth to note that gamma rays, even in the case of their effective production above 100 TeV, would hardly escape the Ridge because of the absorption at interactions with the enhanced far infrared radiation fields in the central 100 pc region. Thus it is likely that the neutrinos remain the only messengers of information about the cosmic-ray protons with energies larger than 1 PeV. This opens a unique opportunity for KM3NeT to provide a major contribution to the exploration of the inner Galactic Center region.

## 5. Conclusions

This analysis presented here shows that, assuming a source emitting a gamma-ray $E^{-2}$ differential energy spectrum through a fully hadronic mechanism, a minimum gamma-ray flux of $E^2\phi_\gamma(10\ \mathrm{TeV}) > 1 \times 10^{-12}$ TeV cm$^{-2}$ s$^{-1}$ is necessary in order to investigate with a $3\sigma$ significance its neutrino counterpart on a time scale of 10 years with KM3NeT, if the source has an angular size of $R_{\mathrm{src}} = 0.1$ deg. In the extreme case of a source with a radial extent of $R_{\mathrm{src}} = 2.0$ deg, only sources brighter than $E^2\phi_\gamma(10\ \mathrm{TeV}) > 2 \times 10^{-11}$ TeV cm$^{-2}$ s$^{-1}$ will be in the reach of neutrino telescopes. In particular, RX J1713.7-3946 and the Galactic Center Ridge remain potential sources for the upcoming generation of neutrino detectors.

## References


1. M.G. Aartsen *et al.*, *Science* **342**, 1242856 (2013).
2. L. Ambrogi, S. Celli and F.A. Aharonian, *Astrop. Phys.*, **100** 69 (2018).
3. K. Bernlöhr *et al.*, *Astrop. Phys.* **43**, 171 (2013).
4. S. Adrian-Martinez *et al.*, *J. of Phys. G*, **43** 8 (2016).
5. `https://www.cta-observatory.org/science/cta-performance/` (2017).
6. H. Abdalla *et al.*, *Astron. Astrophys.* **612**, A6 (2018).
7. F.A. Aharonian *et al.*, *Astron. Astrophys.* **464**, 235 (2007).
8. F.L. Villante and F. Vissani, *Phys. Rev. D* **78**, 103007 (2008).
9. F.A. Aharonian *et al.*, *Nature* **439**, 695 (2006).
10. A. Abramowski *et al.*, *Nature* **531**, 476 (2016).
11. S. Celli, A. Palladino and F. Vissani, *Eur. Phys. J. C* **77**, 2 (2016).
12. H. Abdalla *et al.*, *Astron. Astrophys.* **612**, A9 (2018).

# Relic supernova neutrino spectrum and the nuclear equation of state

G. J. Mathews

*Center for Astrophysics, Department of Physics, University of Notre Dame*
*Notre Dame, IN 46556, USA*
**E-mail: gmathews@nd.edu*
*https://physics.nd.edu/people/faculty/grant-j-mathews/*

Jun Hidaka

*Mechanical Engineering Department, Meisei University*
*Tokyo 191-8506, Japan*
*National Astronomical Observatory of Japan*
*Tokyo, 181-8588, Japan*
*jun.hidaka@meisei-u.ac.jp*

Toshitaka Kajino

*National Astronomical Observatory of Japan*
*Tokyo, 181-8588, Japan*
*International Research Center for Big-Bang Cosmology and Element Genesis, and School of Physics, Beihang University, Beijing 100083, Peoples Republic of China*
*Graduate School of Science, The University of Tokyo, Tokyo, 113-0033, Japan*
*kajino@buaa.edu.cn, kajino@nao.ac.jp*

We explore the sensitivity of the relic supernova neutrino background from core-collapse supernova explosions to the nuclear equation of state (EoS). This sensitivity arises largely from the contribution to neutrino emission from failed supernovae (fSNe). We consider a variety of astrophysical scenarios, including different progenitor masses, different cosmological star formation rates, starbursts, quiescent star formation, and a metallicity dependence of the initial mass function. We find that the EoS signature remains robust under a variety of conditions. We demonstrate the viability of future neutrino detectors to distinguish the nuclear EoS via the relic supernova neutrino spectrum.

*Keywords*: Diffuse radiation - neutrinos - stars: formation - stars: massive - supernovae: general - stars: supergiants.

## 1. Introduction

We have studied[1–3] the contribution to the relic supernova neutrino (RSN) spectrum from different types of SNe. These previous works have shown that there is a clear EoS dependence in the spectrum. In particular in Ref. 3 we clarified the sensitivity to various astrophysical scenarios in order to quantify the EoS dependence of the RSN spectrum. For example, any change in the occurrence of core-collapse supernovae (CCSNe) vs. failed supernovae (fSNe) can affect the expected spectrum. Therefore, we have applied different criteria for a successful SN explosion and analyzed their effects on the RSN spectrum.

Moreover, the observational estimation of the cosmic star formation rate (SFR) is an involved procedure based upon different sources, e.g., UV light from galaxies and the far infrared (FIR) luminosity[4]. It also depends on modeling for the interpretation of the observational data, e.g., the fit to the spectral energy distribution

(SED).[5] Therefore, it is valuable to consider different SFR estimates and study their effects on the RSN spectrum.

The star formation rate also depends upon the choice of the stellar initial mass function (IMF). The IMF is a fundamental quantity not only for SFR estimation but also for the occurrence of each type of SN. A universal IMF, typically Salpeter-like IMF, is commonly assumed in almost every aspect of the astronomical derivation. Theoretical studies have attempted to derive this universality[6,7]. Nevertheless, a convincing theory for the universal IMF has not yet been established. Another uncertainty surrounding the IMF is its metallicity dependence. In general cooler molecular clouds should more easily collapse to form stars. Therefore, metal-poor molecular clouds, which may be common at high redshift, could tend to form more massive stellar objects. This leads to the concept of a top-heavy IMF in the early universe, i.e., a $z$-dependent IMF.

In Refs. 1, 2, 3 we have investigated the EoS dependence of the RSN spectrum by applying a variety of scenarios. These works show that the EoS dependence is sufficiently robust that the relic supernova neutrino spectrum can be used to gain insight into the nuclear EoS for core-collapse supernova explosions almost independently of the many uncertainties in the astrophysical scenarios.

One must know which progenitors become which type of SN because each SN emits a different spectrum of neutrinos [cf. Ref. 1]. Many numerical simulations have been performed which produce supernova explosions similar to SN1987A. Recent progress in multi-dimensional supernova simulations has led to successful explosions[8–11]. There is a common feature in the relationship between the structure of the progenitor and successful explosions[12–17]. The progenitors with greater compactness are less able to explode because they experience more accretion of material before the shock passes through the core. Indeed, successful/failed SNe occur with a non-monotonic dependence on progenitor mass[12,14,16,17]. Previously, most of studies of relic supernova neutrinos assumed definitive mass ranges within which a certain type of SN can occur. The relic neutrino spectrum must be re-evaluated by taking into account the sporadic dependence of the SN explosion on the progenitor mass.

In Ref. 3 we followed the work of Ref. 14 and adopted the fraction $f_{\rm fSN} = 0.26$ for progenitors in the mass range of $10 - 40{\rm M}_\odot$ to include the dependence of the SN explosion on progenitor in this mass range. We then take $f_{\rm fSN} = 1.00$ for ${\rm M} > 40{\rm M}_\odot$. We use this SN-progenitor relationship as the fiducial case throughout this work. For comparison, we also apply $f_{\rm fSN} = 1.00$ for ${\rm M} > 18{\rm M}_\odot$, which has been adopted in previous work with respect to the red supergiant problem[2].

Several different SFR formulae have been proposed to fit the observational data. The data mostly are interpreted from the UV luminosity of galaxies at different redshifts (e.g., Refs. 20, 21). This derivation is complicated because of the possible extinction due to the dust surrounding galaxies[22]. Also, the stellar composition at different ages must be taken into consideration so that the color of each galaxy is

matched to the observations.[5] Typically a Salpeter-like IMF is assumed for both the deduction of the SFR from the UV light and the color matching. Moreover, not only the UV source but also the IR emission is used for the SFR estimate.[23,24]

The relic supernova neutrino spectrum depends strongly upon the SFR through the rate of occurrence for each type of SN. In Ref. 3 the SFR fitting formula proposed by Ref. 4 was adopted. It has been argued[25] that the observationally deduced supernova rate tends to be smaller than the theoretical value. This is the so-called supernova rate problem[1,25]. This depends upon the choice of SFR. The SFR adopted in this work, gives a more conservative relic SN neutrino estimation and is independent of the supernova rate problem.[4]

As mentioned above, most evaluations of the SFR are based upon the UV luminosity, which is mainly emitted by OB stars. The dust surrounding galaxies, however, always complicates the evaluation of the UV. There is a claim[26] that the SFR based upon UV light is an underestimate especially for starburst galaxies. Recently, a new method based upon SED modeling has been proposed[26] in order to take into account the star formation embedded in dense molecular clouds. This study has shown that the SFR for $z > 3.5$ is higher by a factor of 2 to 3 than the estimate from UV light. We also apply this SFR to see how it affects the relic SN neutrino spectrum compared to the fiducial case.

The IMF is commonly assumed to be a universal function that is independent of the physical conditions of the star formation site. One of the popular functional forms is the Salpeter-A IMF (Sal-A).[27] This is a two-segment power-law given by $\phi_0(M) = M^{-\zeta} = M^{-(1+\Gamma)}$, with $\zeta = 2.35 \pm 0.2$ ($\Gamma = 1.35 \pm 0.2$) for stars with $M \geq 0.5 M_\odot$ and $\zeta = 1.5$ ($\Gamma = 0.5$) for stars with $0.1 < M < 0.5 M_\odot$. This is the adopted IMF for the present work. Most of the popular IMFs including the Sal-A are observationally deduced for stars in the Milky Way. There are also theoretical derivations of the power-law index for the IMF. For example, a power law IMF is explained by the turbulent motion of molecular clouds in a magnetic field[7]. There is, however, no definitive theory for the IMF, and the universality of the IMF remains an open question.

The relationship between metallicity and the IMF was also considered.[5,28–30] The IMF in metal poor environments molecular clouds may not sufficiently cool to collapse gravitationally on small scales. We explored[3] the IMF-metallicity relationship by adopting a cosmological metallicity evolution. If we consider a variable IMF, it is necessary to make a correction to the SFR, which is IMF dependent.

We follow the standard method to evaluate the detection rate of relic SN neutrinos as described in Refs. 1, 2, 3, 31, 33. For each type of SN we adopt $z_{max} = 5$ as the redshift at which star formation begins and $R_{SN}$ is the cosmic supernova rate. $dN_\nu(E'_\nu)/dE'_\nu$ is the neutrino spectrum emitted at the source, and the energy $E'_\nu = (1+z)E_\nu$ is the energy at emission. $E_\nu$ is the redshifted energy observed in the detector. A standard cosmology with $H_0 = 70$ km s$^{-1}$ Mpc$^{-1}$, $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$[25,34] was adopted and a Fermi-Dirac distribution was assumed

for all neutrino species from each type of SN. Two models of fSNe, one with a soft EoS[35] (LS-EoS, $K = 180$) and a stiff EoS[36] (Shen-EoS) were considered from the numerical models of Ref. 37.

The supernova rate $R_{SN}$ is depends on the cosmic SFR, and $\phi_0(M)$ is the IMF. Here, $M_{SN}$ denotes the range of progenitor masses that lead to a particular type of SN. Hence, it is obvious that the neutrino flux depends not only upon the cosmic SFR and the IMF, but also upon the occurrence of each type of SN. Then, the detector event rate is

$$\frac{dN_{\rm event}}{dE_{e^+}} = N_{\rm target}\frac{1}{c}\frac{dN_\nu}{dE_\nu}\sigma(E_\nu)\frac{dE_\nu}{dE_{e^+}}, \tag{1}$$

where $N_{target}$ is the number of target particles in the detector. $\sigma(E_\nu)$ is the cross section for neutrino interactions within the detector. In the calculations, we assume a water Čerenkov detector so that the neutrino absorption is dominated by the $\bar{\nu}_e + p \rightarrow e^+ + n$ reaction with $E_\nu = E_{e^+} + 1.3$ MeV.

The effects of neutrino oscillations on the final neutrino signals have also been considered.[1–3] Two possible cases were considered i.e. either a normal or inverted mass hierarchy associated with complete non-adiabatic mixing (Case I) and an inverted mass hierarchy associated with complete adiabatic mixing (Case II) through the MSW high-density resonance. We denote the case without neutrino oscillations as Case III. These schemes were applied to each different astronomical scenario.

## 2. Results

We numerically treated both Type-II SNe and SNeIb,c as luminous core-collapse SNe in the same manner and denote them simply as CCSNe. We then showed the results within different astrophysical scenarios. For each case we considered the neutrino oscillations (Cases I & II) in the inverted and normal hierarchy and cases without neutrino oscillations (Case III). Comparisons among different scenarios for Case I, II, and III are made by looking at the locations of peak and tail of spectrum.

Figure 1 shows the the esseential result of these studies. It shows the differences between the peak and the tail of the observable observed neutrino detector response for all uncertainties included in the no-oscillations studies (Case III). This figure reveals the robustness of the EoS dependence by indicating the characteristic locations with the rectangular symbols as described above. The EoS dependence appears as a robust separation between the predicted RSN spectra for a stiff vs. a soft EoS. Other cases show a similar pattern.[3]

In summary, the fiducial case in this study, we have adopted the SFR parametrized by Ref. 4 along with the sporadic CCSN/fSN occurrence in the mass range of $10-40M_\odot$ based upon the numerical studies of Ref. 14. We then estimated the RSN spectrum. In addition, we have considered a wide variety of astrophysical scenarios and investigated the EoS dependence of the RSN spectrum for each case with and without the occurrence of neutrino oscillations. We find that the EoS dependence of the RSN spectrum manifests prominently in the high energy tail in

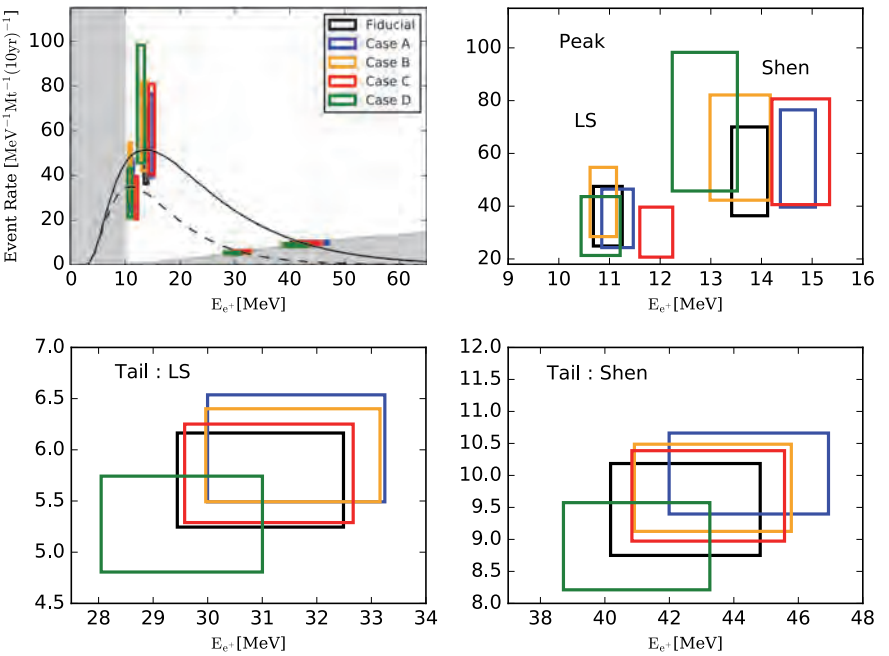



Fig. 1. Comparison of Case III (without neutrino oscillation) for all scenarios.. The figure includes only rectangles for the location of peak and tail which characterize the spectrum for each cases. In the left upper panel, the spectrum of the fiducial case without the error band are included for reference. The solid line is for the Shen EoS and the dashed line for the LS EoS. The right upper panel shows the detail of the peak locations for both the Shen and LS EoSs. Similarly the details of tails are shown in the left bottom (for the LS EoS) and right bottom panel (for the Shen EoS).

any scenario considered. The robustness of this EoS dependence can be understood in that the RSN spectrum is determined mostly by the SFR for $z < 2$. For example, in the case which adopts an enhanced cosmic SFR for $z > 3$, it is clearly seen that the RSN exhibits an EoS dependence similar to that of the fiducial cosmic SFR.

## Acknowledgments

Work at the University of Notre Dame (GJM) supported by the U.S. Department of Energy under Nuclear Theory Grant DE-FG02-95-ER40934. This work was also supported in part by Grants-in-Aid for Scientific Research of JSPS (15H03665, 17K05459) of the Ministry of Education, Culture, Sports, Science and Technology of Japan.

## References

1. Mathews, G. J., Hidaka, J., Kajino, T., & Suzuki, J. 2014, ApJ, 790, 115.
2. Hidaka, J., Kajino, T., & Mathews, G. J. 2016, ApJ, 827, 85.

3. Hidaka, J., Kajino, T., & Mathews, G. J. 2019, ApJ, 874, 11.
4. Madau, P., & Dickinson, M. 2014, ARA&A, 52, 415.
5. Conroy, C., & van Dokkum, P. G. 2012, ApJ, 760, 71.
6. Padoan, P., & Nordlund, A. 2002, ApJ, 576, 870.
7. Padoan, P., Nordlund, A., & Jones, B. J. T. 1997, MNRAS, 288, 145.
8. Melson, T., Janka, H.-T., Bollig, R., et al. 2015a, ApJ, 808, L42.
9. Melson, T., Janka, H.-T., & Marek, A. 2015b, ApJ, 801, L24.
10. Lentz, E. J., Bruenn, S. W., Hix, W. R., et al. 2015, ApJ, 807, L31.
11. B. Müller, B. 2015, MNRAS, 453, 287.
12. Ugliano, M., Janka, H.-T., Marek, A., & Arcones, A., ApJ, 757, 69 (2012).
13. Horiuchi, S., Nakamura, K., Takiwaki, T., Kotake, K., & Tanaka, M. 2014, MNRAS, 445, L99.
14. Pejcha, O., & Thompson, T. A. 2015, ApJ, 801, 90.
15. Nakamura, K., Takiwaki, T., Kuroda, T., & Ko- take, K. 2015, PASJ, 67, 107.
16. Sukhbold, T., Ertl, T., Woosley, S. E., Brown, J. M., & Janka, H.-T. 2016, ApJ, 821, 38.
17. Ertl, T., Janka, H.-T., Woosley, S. E., Sukhbold, T., & Ugliano, M. 2016, ApJ, 818, 124.
18. Horiuchi, S., Beacom, J. F., & Dwek, E. 2009, Phys. Rev. D, 79, 083013.
19. Keehn, J. G., & Lunardini, C. 2012, Phys. Rev. D, 85, 043011.
20. Cucciati, O., Tresse, L., Ilbert, O., et al. 2012, A&A, 539, A31.
21. R. J. Bouwens, G. D. Illingworth, P. A. Oesch, et al., ApJ, 752, L5 (2012).
22. Buat, V., Noll, S., Burgarella, D., et al. 2012, A&A, 545, A141.
23. Magnelli, B., Elbaz, D., Chary, R. R., et al. 2011, A&A, 528, A35.
24. Gruppioni, C., Pozzi, F., Rodighiero, G., et al. 2013, MNRAS, 432, 23.
25. Horiuchi, S., Beacom, J. F., Kochanek, C. S., et al. 2011, ApJ, 738, 154.
26. Rowan-Robinson, M., Oliver, S., Wang, L., et al. 2016, MNRAS, 461, 1100.
27. I. K. Baldry, & K. Glazebrook, ApJ, 593, 258 (2003).
28. Hoversten, E. A., & Glazebrook, K. 2008, ApJ, 675, 163.
29. Meurer, G. R., Wong, O. I., Kim, J. H., et al. 2009, ApJ, 695, 765.
30. Gunawardhana, M. L. P., Hopkins, A. M., Sharp, R. G., et al. 2011, MNRAS, 415, 1647.
31. Strigari, L. E., Beacom, J. F., Walker, T. P., & Zhang, P. 2005, J. Cosmology Astropart. Phys., 4, 017.
32. Strumia, A., & Vissani, F. 2003, Physics Letters B, 564, 42.
33. Yüksel, H., & Beacom, J. F., Phys. Rev. D, 76, 083007 (2007).
34. Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, ApJS, 208, 19.
35. Lattimer, J. M., & Swesty, D. F. 1991, Nuclear Physics A, 535, 331.
36. Shen, H., Toki, H., Oyamatsu, K., & Sumiyoshi, K. 1998, Nuclear Physics A, 637, 435.
37. Sumiyoshi, K., Yamada, S., & Suzuki, H. 2008, ApJ, 688, 1176.

# Investigating gamma-ray bursts by joining Insight-HXMT and other gamma-ray spacecraft

C. Guidorzi*, R. Martone, M. Marongiu, F. Frontera, P. Rosati and E. Virgilli

*Dept. Physics and Earth Science, University of Ferrara, via Saragat 1, Ferrara, I-44122, Italy*
**E-mail: guidorzi@fe.infn.it*

L. Amati, M. Orlandini and J. Stephen

*INAF–OAS Bologna, Via Gobetti 93/3, Bologna, I-40129, Italy*

C. Giuri

*DESY, Platanenallee 6, Zeuthen, D-15738, Germany*

S.-N. Zhang, S. Xiong

*Institute of High Energy Physics, Chinese Academy of Sciences, Beijing, 100049, China*

The Chinese mission Insight–HXMT allows for a high sensitivity study of the X-ray sky in 1250 keV, thanks to a smart combination of collimated detectors operating in different energy bands and providing a source location accuracy of 1 arcmin for a 20-$\sigma$ source. In addition to observing Galactic sources, the HXMT High Energy (HE) instrument also operates in the so-called GRB mode, that makes it possible to detect and characterise GRB prompt emission from 200 keV to 3 MeV with an effective area as high as 2000 cm$^2$, thus filling the sensitivity gap of presently flying main GRB detectors in this energy range. We report the results of intensive simulations of the expected performances of HXMT for GRB science, showing how, especially in combination with, e.g., Swift/BAT and Fermi/GBM, the HXMT/HE will provide a significant improvement in the characterisation of temporal and spectral properties of the prompt emission of long GRBs, as well as an improved detection rate and better spectral characterisation of short GRBs.

*Keyword*: Gamma-ray bursts.

## 1. Introduction

The Insight Hard X-ray Modulation Telescope (HXMT) was launched June 15, 2017 in a 43°-inclination orbit and is devoted to X-ray astrophysics. The payload consists of three collimated instruments covering a total energy band of 1–250 keV[1]. The High–Energy instrument (HE), made of 18 phoswich detection units NaI(Tl) + CsI(NaI), operates in the 20–250 keV and has a total geometric area of $\sim$5000 cm$^2$, time resolution of 25$\mu$s and energy resolution of 19% at 60 keV. All the information for each photon is recorded. The detection system and the design of the HE units are very similar to those of the BeppoSAX/PDS instrument[2], thus making it an excellent instrument for the study of hard X-ray sources as well as a sensitive full-sky gamma-ray burst (GRB) monitor. In the normal operation mode, the energy band is 20–250 keV for NaI detectors and 40–600 keV for CsI detectors. The HE electronics and on-board data handling system are designed in such a way

to identify NaI and CsI events, and to measure the energy of these events along with their occurrence time. In the normal mode, the CsI detectors above 100 keV exhibit an effective area ranging from few hundreds $cm^2$ for front incident radiation, to ∼1500–2000 $cm^2$, for radiation coming from the rear of the instrument. In the GRB operation mode, activated during Earth occultation of pointed sources, the HV of the HE instrument units is lowered to reduce the gain by a factor of ∼5, thus changing the energy band of the NaI detectors from 20–250 keV to 100–1250 keV and that of the CsI from 40–600 keV to 200–3000 keV (right panel of Fig. 1). In

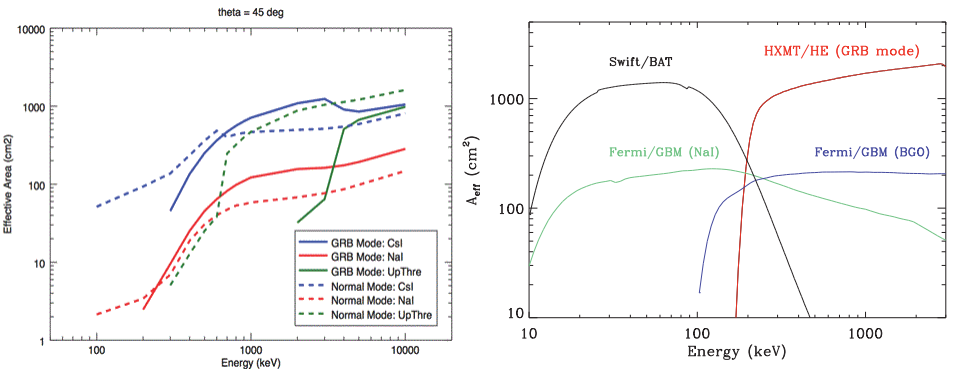



Fig. 1. *Left*: effective area of the NaI and CsI detectors of the HXMT/HE instrument in both normal and GRB modes, for an offset angle $\theta = 45°$ with respect to the detector axis. For each mode green lines show the efficiency of the high-energy threshold as a function of photon energy. The response function at different directions were determined through extensive Monte Carlo simulations (GEANT4) of the whole satellite. *Right*: effective area as a function of energy for Swift/BAT (on–axis), Fermi/GBM (sum of two NaI detectors and BGO, for a random direction in the sky) and HXMT/HE (GRB mode, $\theta = 135°$).

this paper we focus on the capabilities of HXMT/HE to detect and characterise the prompt emission of GRBs. In fact, the HE instrument is also operated in a GRB mode that allows to derive with high accuracy the temporal and spectral properties of the detected GRBs in the energy range from 200 keV to 3 MeV, thanks to an average effective area of ∼1500–2000 $cm^2$. This way, HXMT complements the sensitivities of other important and presently flying GRB detectors, like, e.g., Swift/BAT (15–150 keV) and Fermi/GBM (8 keV–30 MeV but with small effective area), thus providing a relevant contribution to the understanding of the physics of the prompt emission, one of the main open issues in the GRB field (e.g., see Ref. 3 for a review), and to the use of spectrum–intensity correlations for cosmology[4] (left panel of Fig. 1). For instance, the spectral peak energy, $E_{\rm p}$, corresponding to the peak of the $\nu F_\nu$ spectrum, is a key parameter for GRB prompt emission models. In addition, $E_{\rm p}$ strongly correlates with the GRB radiated energy during the prompt emission[5,6], during the early X-ray afterglow[7], and with the time variability and properties of the power density spectrum of the gamma-ray time profile[8]. With

reference to the "Band function" [9], the high-energy spectral index $\beta$ is also very important for testing and discriminating among various models, thus providing clues to the physics and geometry of the ultrarelativistic jet that is thought to act in a GRB. Swift/BAT measures mostly power-law spectra, thus providing estimates of $E_{\rm p}$ only for a few GRBs and almost no estimates of $\beta$, whilst most Fermi/GBM spectra can be fitted with a cut-off power-law, with no clues on $\beta$.

## 2. GRB spectroscopy

We quantified the improvements expected by joining HXMT/HE with other missions through simulations with the XSPEC package (v.12) in terms of the accuracy in determining $E_{\rm p}$ and $\beta$. The response matrix of HXMT/HE in GRB mode was obtained through intensive Monte Carlo simulations of the instrument and of the whole satellite, combined with extensive on-ground calibrations [1]. We used the response matrix for an offset incident angle of $\theta = 135°$, which is the rear side direction. The background spectrum was derived by taking into account the expected in-flight background components, including CXB, primary and albedo protons, electrons and positrons, gamma-rays albedo and SAA proton activated fluorescence by instrument and satellite structures [10]. Figure 2 (*left*) shows the Swift/BAT (measured)

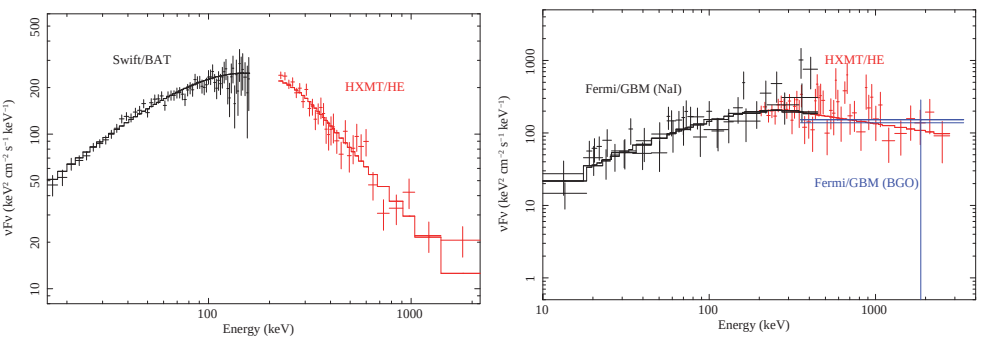



Fig. 2. Left: $\nu F_\nu$ spectrum of GRB 161218A as measured by Swift/BAT and expected with HXMT/HE (GRB mode) by assuming the spectral parameters measured by Konus–WIND [11]. Right: simulated Fermi/GBM and HXMT/HE (GRB mode) spectra of a medium–weak GRB (10–1000 keV fluence of $5 \times 10^{-6}$ erg cm$^{-2}$) with a typical Band spectral shape with $\alpha = -1$, $\beta = -2.3$ and $E_{\rm p} = 300$ keV and 25-s duration. Channels have been grouped to ensure $> 2\sigma$ significance.

and HXMT/HE (simulated) joint spectrum of the medium intensity GRB 161218A. While BAT alone would measure a simple power-law spectrum, the HE data are essential to constrain both $E_{\rm p}$ and $\beta$ by measuring the whole spectral curvature. A comparison between Fermi/GBM and HXMT/HE spectral capabilities is also shown (*right*): we simulated the spectrum of a medium–weak GRB (10–1000 keV fluence of $5{\times}10^{-6}$ erg cm$^{-2}$) with a typical Band spectral shape: $\alpha = -1$, $\beta = -2.3$ and $E_{\rm p} = 300$ keV. In particular, we simulated the spectra of the two NaI detectors and

the BGO detector most illuminated by the considered GRB, as usual with real data. Although the nominal energy band of the GBM extends up to 20–30 MeV, its effective area is poor with respect to that of HXMT/HE at energies higher than ∼300 keV (Fig. 1). Thus, also in this case HXMT/HE data are essential to constrain the spectral shape and determine $E_{\rm p}$ and $\beta$. We then simulated a set of Fermi/GBM (using real background data) and HXMT/HE (GRB mode) spectra by assuming the above typical spectral shape and a 10–1000 keV fluence in the range $10^{-4}$–$10^{-7}$ erg cm$^{-2}$. The results are shown in Fig. 3 in terms of fractional $1\sigma$ uncertainty in the estimates of $E_{\rm p}$ and $\beta$ as a function of GRB fluence. Clearly, HXMT/HE

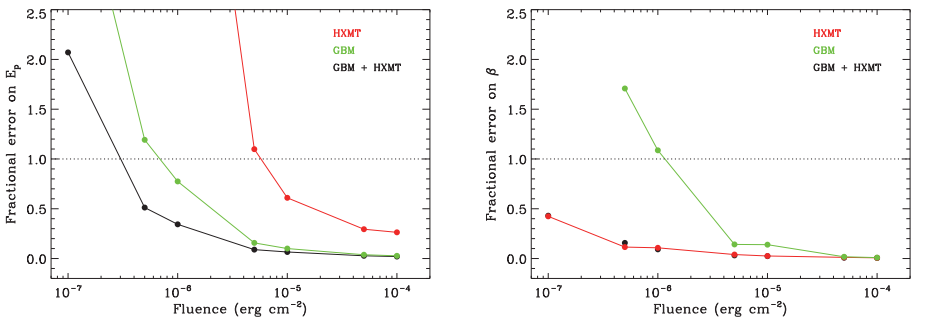



Fig. 3. Fractional error on spectral parameters $E_{\rm p}$ (*left*) and $\beta$ (*right*) as a function of fluence (10–1000 keV) for a GRB with $\alpha = -1$, $\beta = -2.3$ and $E_{\rm p} = 300$ keV, and duration 25 s, as would be measured with Fermi/GBM alone, HXMT/HE alone and a joint spectral analysis of spectra from the two instruments.

alone constrains $E_{\rm p}$ well only for the brightest GRBs, whereas $\beta$ is measured well down to the faintest ones. Conversely, the Fermi/GBM alone constrains $E_{\rm p}$ down to medium-bright GRBs. Yet, only the joint analysis of both instruments enables a good estimate of $E_{\rm p}$ down to fluences as low as $5{\times}10^{-7}$ erg cm$^{-2}$. This will significantly increase the number of GRBs that can be used for testing prompt emission models and for the study of the $E_{\rm p}$–intensity correlations. For progressively lower GRB fluences and higher values of $E_{\rm p}$ the HXMT spectral capabilities become increasingly relevant, especially for short GRBs, whose average $E_{\rm p}$ is higher than that of long ones[12].

## 3. GRB timing

The degree of variability exhibited by GRB light curves is long known to be related to luminosity[13,14] and spectral properties[8]. Among the various characterising temporal observables, such as the average[15] and the individual power density spectrum[8,16], noteworthy is the so-called spectral (or "time") lag, i.e. the delay between the time profiles of a given GRB at two different energy channels with softer photons lagging behind harder ones. The spectral lag is important for several aspects: (i) it anti-correlates with luminosity for long GRBs[17,18] as well as for GRB X-ray

flares[19]; (ii) it discriminates between long and short GRBs, for which typical values are several ten or hundred ms vs. just a few ms, respectively[20]; (iii) an exhaustive self-consistent explanation is still missing, although two main interpretations have been put forward: (a) a spectral hard-to-soft evolution within the fluid comoving frame, or (b) a kinematic/geometric effect due to the Doppler factor connected with the bulk Lorentz factor of the jet, or the combination of both. The spectral lag is usually calculated through the cross-correlation function (CCF) of the two profiles. As such, its accuracy strongly depends on the number of counts, especially in the harder channel, which is more significantly hampered by low-count statistics: indeed, above $> 300$ keV the lag calculation is an almost uncharted territory, apart from rare, exceptionally bright events. This limitation is even more severe for short GRBs due to the lower fluence and harder photons than for long GRBs. Thanks to its large effective area above a few hundreds keV, HXMT/HE holds a great promise for the spectral lag calculation even for less fluent GRBs, such as short ones. We assessed the potential of joining HXMT/HE to Fermi/GBM through a number of simulations of short GRB profiles, assuming the response function for an off axis angle of $\theta = 135°$ in GRB mode. In particular, we simulated a fast-rise exponential decay (FRED) with an instantaneous spectrum described by a Band function with typical values for short GRBs: $\alpha = -0.5$, $\beta = -2.3$, and $E_{\rm p}$ monotonically decreasing from the initial value of 1 MeV according to Ref. 21. Figure 4 shows

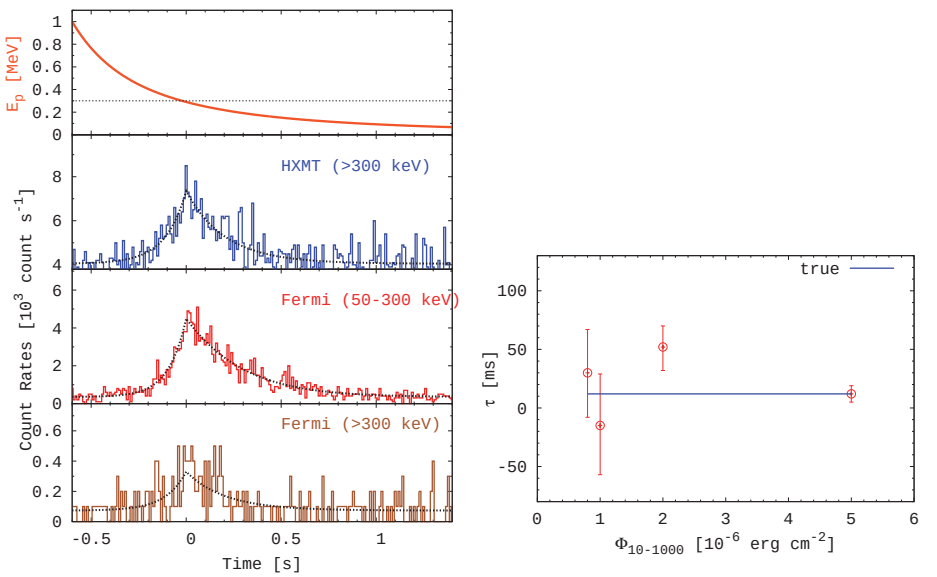



Fig. 4. *Left*: simulated profile of a short GRB ($T_{90} = 0.84$ s; 10–1000-keV fluence of $5 \times 10^{-6}$ erg cm$^{-2}$) with hard-to-soft evolution due to a decreasing $E_{\rm p}$ (top). The HXMT/HE profile above 300 keV is shown in comparison with the Fermi/GBM soft (50–300 keV) and hard ($> 300$ keV) bands. In the hard channel the better statistical quality of HXMT/HE is evident. *Right*: spectral lag between HXMT/HE ($> 300$ keV) and Fermi/GBM (50–300 keV) profiles as a function of the 10–1000-keV fluence of the short GRB, compared with the true value (solid line).

the different simulated profiles for HXMT/HE and for Fermi/GBM (*left*), where the statistical quality of the latter is clearly lower than that of the former. We then calculated the lag via CCF between the soft energy channel of Fermi/GBM and that of HXMT/HE for a number of identically shaped short GRBs within increasing $\gamma$–ray fluences and compared the results with the true value (measured without statistical noise): the result (right panel) shows the unprecedented good accuracy obtained at any given fluence in the considered range. This will allow for a better characterisation of the spectral lag of a larger number of short GRBs, whose relevance cannot be overstated given their association with the mergers of binary neutron stars revealed through the detection of gravitational waves, as was first the case for GRB170817A/GW170817[22].

## Acknowledgments

We acknowledge financial contribution from the agreement ASI-INAF n.2017-14-H.0.

## References

1. S. Zhang, F. J. Lu, S.N. Zhang, T.P. Li, *Procs. SPIE*, **9144**, 914421 (2014).
2. F. Frontera, E. Costa, D. Dal Fiume, et al., *A&AS*, **122**, 357 (1997).
3. P. Kumar and B. Zhang, *Physics Reports*, **561**, 1 (2015).
4. L. Amati and M. Della Valle, *Int. J. of Mod. Phys. D*, **22**, 30028 (2013).
5. L. Amati, F. Frontera, M. Tavani, et al., *A&A*, **390**, 81 (2002).
6. R.-J. Lu, J.-J. Wei, E.-W. Liang et al., *ApJ*, **756**, 112 (2012).
7. M. G. Bernardini, R. Margutti, E. Zaninoni, et al., *MNRAS*, **425**, 1199 (2012).
8. S. Dichiara, C. Guidorzi, L. Amati, et al., *A&A*, **589**, A97 (2016).
9. D. Band, J. Matteson, J. Ford et al., *ApJ*, **413**, 281 (1993).
10. F. Xie, J. Zhang, L.-M. Song et al., *Ap&SS*, **360**, 47 (2015).
11. D. Fredericks, S. Golenetskii, R. Aptekar et al. *GCN*, **20292** (2016).
12. D.S. Svinkin, D.D. Frederiks, R.L. Aptekar et al., *ApJS*, **224**, 10 (2016).
13. D.E. Reichart, D.Q. Lamb, E.E. Fenimore et al., *ApJ*, **552**, 57 (2001).
14. C. Guidorzi, F. Frontera, E. Montanari et al., *MNRAS*, **363**, 315 (2005).
15. S. Dichiara, C. Guidorzi, L. Amati et al., *MNRAS*, **431**, 3608 (2013).
16. C. Guidorzi, S. Dichiara and L. Amati, *A&A*, **589**, A98 (2016).
17. J.P. Norris, G.F. Marani and J.T. Bonnell, *ApJ*, **534**, 248 (2000).
18. T.N Ukwatta, M. Stamatikos, K.S. Dugha et al., *ApJ*, **711**, 1073 (2011).
19. R. Margutti, C. Guidorzi, G. Chincarini et al., *MNRAS*, **406**, 2149 (2010).
20. M.G. Bernardini, G. Ghirlanda, S. Campana et al., *MNRAS*, **446**, 1129 (2015).
21. D. Kocevski and E. Liang, *ApJ*, **594**, 385 (2003).
22. LIGO Scientific Collaboration and Virgo Collaboration, *ApJ*, **848**, L13 (2017).

# An updated calculation of the high energy diffuse gamma and neutrino flux from the galactic disk

Francesco L. Villante[1,2], Maddalena Cataldo[1,2], Giulia Pagliaroli[2,3] and Vittoria Vecchiotti[2,3]

[1] *Department of Physical and Chemical Science (DSFC), University of L'Aquila, L'Aquila I-67100, Italy*

[2] *Laboratori Nazionali del Gran Sasso (LNGS), Istituto Nazionale Fisica Nucleare (INFN), Assergi (AQ) I-67100, Italy*

[3] *Gran Sasso Science Institute (GSSI), L'Aquila I-67100, Italy*

We provide updated expectations for the diffuse gamma and neutrino emission produced in the TeV domain by the interaction of cosmic rays with the gas contained in the galactic disk. Motivated by recent analyses of the Fermi-LAT data, we consider different assumptions for the cosmic ray space and energy distribution in the Galaxy, including the possibility that cosmic ray energy spectrum depends on the galactocentric distance.

*Keyword*: High energy photons and neutrinos.

## 1. Introduction

Cosmic rays (CR) that propagate in different regions of the Galaxy interact with the gas contained in the galactic disk through hadronic processes and produce a diffuse flux of high energy (HE) gammas and neutrinos. These particles propagate to Earth along straight lines providing us with information on the CR space and energy distributions and, thus, in turn, on the CR transport in the galactic magnetic field. In this work, we present updated predictions for the diffuse galactic gamma-ray and neutrino flux produced by this mechanism in the TeV domain. Our results are obtained by using the phenomenological approach introduced in Ref. 1 that allows us to implement different assumptions for the CR space and energy distribution in a simple and direct way, including the possibility of a position dependent CR spectral index, recently emerged from analyses of the Fermi-LAT data[2–5].

## 2. The diffuse gamma and neutrino fluxes

The diffuse gamma (at energies $E_\gamma \leq 20\,\text{TeV}$ where gamma ray absorption in the Galactic radiation field is expected to be negligible) and neutrino fluxes produced by the interaction of CR with the interstellar medium in the galactic plane can be written as:

$$\varphi_{i,\text{diff}}(E_i,\hat{n}_i) = A_i \left[ \int_{E_i}^{\infty} dE \; \frac{\sigma(E)}{E} \; F_i\left(\frac{E_i}{E}, E\right) \right.$$
$$\left. \int_0^{\infty} dl \, \varphi_{\text{CR}}(E, \mathbf{r}_\odot + l\,\hat{n}_i)\, n_{\text{H}}(\mathbf{r}_\odot + l\,\hat{n}_i) \right], \qquad (1)$$

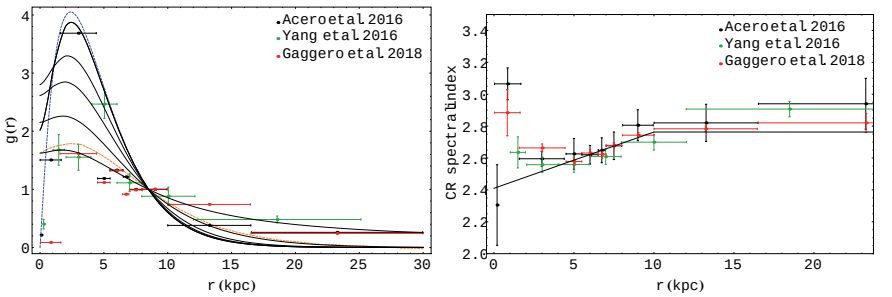



Fig. 1. Left Panel: The black lines show the functions $g(\mathbf{r})$ obtained from Eq. (5) for different smearing radii $R$. Going from top to bottom at $r \simeq 2\,\mathrm{kpc}$, the different lines correspond to $R = 1, 3, 5, 10, \infty$ kpc, respectively. The data points[3–5] show the CR density at $E \simeq 20\,\mathrm{GeV}$ (normalized to one at the Sun position) and the $\gamma$-ray emissivity above $E_\gamma = 1\,\mathrm{GeV}$, as a function of the galactocentric distance obtained from Fermi-LAT data. Right Panel: The black line gives the spectral index of CR at $E_{\mathrm{CR}} = 20\,\mathrm{GeV}$ adopted in this calculation. The data points[3–5] show the CR spectral index as a function of the galactocentric distance obtained from Fermi-LAT data.

where $i = \nu, \gamma$ stands for neutrinos and gamma respectively, $E_i$ and $\hat{n}_i$ indicate the energy and arrival direction of the considered particles and $r_\odot = 8.5$ kpc is the position of the Sun. The total inelastic cross section in nucleon-nucleon collision, $\sigma(E)$, is given by:

$$\sigma(E) = 34.3 + 1.88\ln(E/1\mathrm{TeV}) + 0.25\ln(E/1\mathrm{TeV})^2 \ \ \mathrm{mb},$$

where $E$ is the nucleon energy, while the spectra $F_i\,(E_i/E, E)$ of produced secondary particles are described (with 20% accuracy) by the analytic formulas given in Ref. 6. The constant $A_i$ is equal to 1 for photons and 1/3 for neutrinos, since the one flavour neutrino flux $\varphi_{\nu,\mathrm{diff}}$ is obtained by summing over the production rates of $\nu_e$ and $\nu_\mu$ in the sources, i.e.:

$$F_\nu\,(E_i/E, E) \equiv F_{\nu_\mu}\,(E_\nu/E, E) + F_{\nu_e}\,(E_\nu/E, E)\ , \tag{2}$$

and then assuming flavour equipartition at Earth, as it is expected with good accuracy due to neutrino mixing, see e.g. Ref. 7. The gas density $n_{\mathrm{H}}(\mathbf{r})$ is described as in Ref. 8; the heavy element contribution is included by assuming that the total mass of the interstellar gas is a factor 1.42 larger that the mass of hydrogen, as it is expected if the solar system composition is considered representative for the entire galactic disk.

Following Ref. 1, the CR nucleon flux is given as a function of position and nucleon energy by:

$$\varphi_{\mathrm{CR}}(E, \mathbf{r}) = \varphi_{\mathrm{CR},\odot}(E)\, g(\mathbf{r})\, h(E, \mathbf{r}) \tag{3}$$

where $\varphi_{\mathrm{CR},\odot}(E)$ represents the local flux,[a] $g(r)$ is an adimensional function (normalized to one at the Sun position $\mathbf{r}_\odot$) introduced to describe the spatial distribution

[a] In this work, we use the symbol $\varphi$ for angle-differential fluxes and $\Phi$ for angle-integrated fluxes.

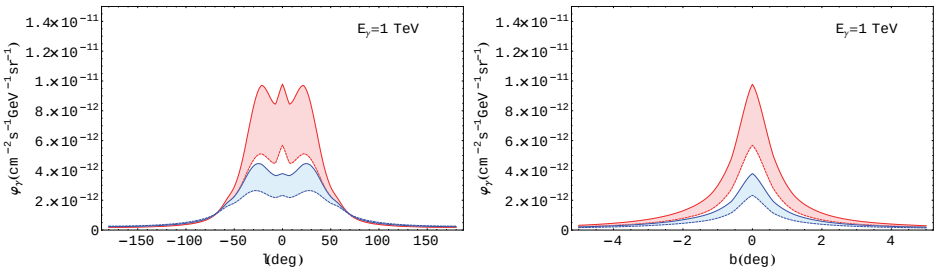



Fig. 2. The diffuse HE gamma-ray flux at $E_\gamma = 1\,\mathrm{TeV}$ as a function of the Galactic longitude $l$ (for $b = 0$) obtained by assuming that the CR spectrum is position-independent (blue lines) and by implementing CR spectral hardening in the inner Galaxy (red lines) as described by Eqs. (4), (8). Solid and dashed lines (in each group) are obtained by assuming that the CR spatial distribution follows that of SNR with smearing radius $R = 1\,\mathrm{kpc}$ and $R = \infty$, respectively.

of CR while the function $h(E, \mathbf{r})$, given by:

$$h(E, \mathbf{r}) = \left(\frac{E}{\overline{E}}\right)^{\Delta(\mathbf{r})} \tag{4}$$

with $\overline{E} = 20\,\mathrm{GeV}$ and $\Delta(\mathbf{r}_\odot) = 0$, introduces a position-dependent variation $\Delta(\mathbf{r})$ of the CR spectral index.

We describe the local CR nucleon flux $\varphi_{\mathrm{CR},\odot}(E)$ according to the data driven parameterization given in Ref. 9 that relies as little as possible on theoretical assumptions. In the energy range of interest for the present analysis, that roughly corresponds to $E \simeq 10 \times E_\gamma \sim 10\,\mathrm{TeV}$ when we consider photons with energy $E_\gamma \sim 1$ TeV, the total nucleon flux is lower than the broken-power law parameterization[10] adopted in our previous calculation[1] by $\sim 12\,\%$. For neutrinos, we are interested at larger CR energies $E \simeq 20 \times E_\nu \sim 2\,\mathrm{PeV}$, as can be understood by considering that the astrophysical signal emerges from background components at $E_\nu \sim 100$ TeV in IceCube data. At these energies, the determination of the local flux is affected by relevant uncertainties and the total nucleon flux adopted in this work is reduced by a factor $\sim 2$ with respect to that given in Ref. 10.

The function $g(\mathbf{r})$ is determined by the distribution of the CR sources $f_\mathrm{S}(\mathbf{r})$, that is assumed to follow the Supernova Remnants (SNR) number density parameterization given by Ref. 12, and by the propagation of CR in the Galactic magnetic field. Following Ref. 11, we define the function $g(\mathbf{r})$ as:

$$g(\mathbf{r}) = \frac{1}{\mathcal{N}} \int d^3x \; f_\mathrm{S}(\mathbf{r} - \mathbf{x}) \; \frac{\mathcal{F}(|\mathbf{x}|/R)}{|\mathbf{x}|} \tag{5}$$

where $\mathcal{N}$ is a normalization constant given by:

$$\mathcal{N} = \int d^3x \; f_\mathrm{S}(\mathbf{r}_\odot - \mathbf{x}) \; \frac{\mathcal{F}(|\mathbf{x}|/R)}{|\mathbf{x}|} \tag{6}$$

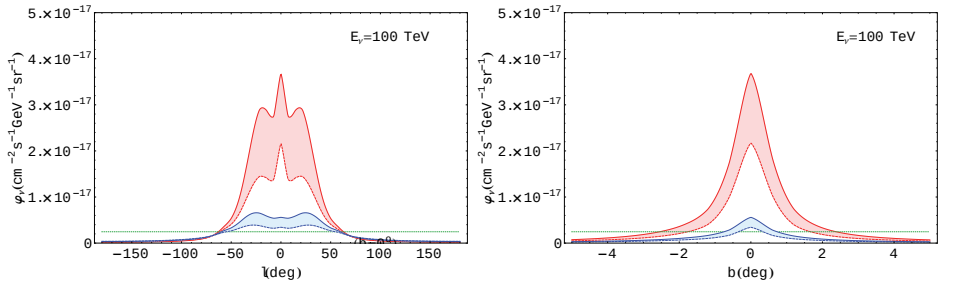



Fig. 3. The diffuse HE neutrino flux at $E_\nu = 100\,\mathrm{TeV}$ as a function of the Galactic longitude $l$ (for $b = 0$) obtained by assuming that the CR spectrum is position-independent (blue lines) and by implementing CR spectral hardening in the inner Galaxy (red lines) as described by Eqs. (4), (8). Solid and dashed lines (in each group) are obtained by assuming that the CR spatial distribution follows that of SNR with smearing radius $R = 1\,\mathrm{kpc}$ and $R = \infty$, respectively.

while the function $\mathcal{F}(\nu)$ is defined as:

$$\mathcal{F}(\nu) \equiv \int_\nu^\infty d\gamma \, \frac{1}{\sqrt{2\pi}} \exp\left(-\gamma^2/2\right) \tag{7}$$

This kind of behaviour can be motivated as the solution of 3D isotropic diffusion equation with constant diffusion coefficient and stationary CR injection. In this context, the smearing parameter $R$ represents the diffusion length $R = \sqrt{2\,D\,t_\mathrm{G}}$, where $D$ is the diffusion coefficient and $t_\mathrm{G}$ is the integration time.

The functions $g(\mathbf{r})$ calculated for different smearing radii $R$ are shown as a function of the galactocentric distance by the black lines in the left panel of Fig. 1 where they are compared with the CR density at $E \simeq 20\,\mathrm{GeV}$ and the $\gamma$-ray emissivity integrated above $E_\gamma = 1\,\mathrm{GeV}$ (which is a proxy of the CR flux) obtained from Fermi-LAT data in Refs. 3, 4, 5. While we believe that the data show a clear trend as a function of $r$, we think that they still do not allow to discard any of the proposed curves. We thus consider the two extreme assumptions $R = 1\,\mathrm{kpc}$ and $R = \infty$, corresponding to the thick black lines in Fig. 1, in order to encompass a large range of possibilities and to provide a conservative estimate of the uncertainty in the gamma-ray flux connected with different descriptions of the CR spatial distribution in the Galaxy.[b]

The possibility of CR spectral hardening in the inner Galaxy is implemented by modelling the function $\Delta(\mathbf{r})$ in Eq. (4) as:

$$\Delta(r, z) = \Delta_0 \left(1 - \frac{r}{r_\odot}\right) \tag{8}$$

with $r_\odot = 8.5\,\mathrm{kpc}$, in galactic cylindrical coordinates. This choice is equivalent to what is obtained in Ref. 2 by considering a phenomenological CR propagation model

[b]Since the assumed smearing length is larger than the thickness of the galactic disk, we neglect in both cases the variation of the CR flux along the galactic latitudinal axis. In other words, we assume $g(r, z) \simeq g(r, 0)$ in the disk, where we used galactic cylindrical coordinates.

characterized by radially dependent transport properties. The numerical parameter $\Delta_0$ that physically corresponds to the difference between the CR spectral index at the Sun position, $\alpha_\odot \simeq 2.7$ at $E = 20\,\mathrm{GeV}$, and its value close to the galactic center, is taken as $\Delta_0 = 0.3$ since this assumption allows us to reproduce the trend with $r$ observed in Refs. 3, 4, 5 for $r \leq 10\,\mathrm{kpc}$, as it is shown in the right panel of Fig. 1.[c] In more external regions, the evidence for a variation of CR energy distribution is much weaker and we assume that $\Delta(r)$ is constant, as shown by the black solid line.

## 3. Results

The diffuse HE gamma-ray flux at $E_\gamma = 1\,\mathrm{TeV}$ is plotted in Fig. 2 as function of the Galactic longitude (at the fixed latitude $b = 0$, left panel) and latitude (at the fixed longitude $l = 0$, right panel). Blue lines are obtained by assuming that the CR spectrum is independent from the position in the Galaxy (we refer to this in the following as the "standard" scenario) while red lines implement CR spectral hardening in the inner Galaxy. Solid and dashed lines in (each group) are obtained by assuming that the CR spatial distribution is described by the function $g(r)$ given in Eq. (5) with smearing radius equal to $R = 1\,\mathrm{kpc}$ and $R = \infty$, respectively.

The angle-integrated gamma-ray flux is equal to $\Phi_\gamma = (7.0 - 8.0) \cdot 10^{-13}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{GeV}^{-1}$ at 1 TeV in the standard scenario with upper and lower bounds corresponding to $R = 1\,\mathrm{kpc}$ and $R = \infty$, respectively. The inclusion of CR hardening increases the integrated flux by $1.2 - 1.3$, the exact enhancement factor being dependent on the assumed CR spatial distribution. Even if the effect on the total flux is relatively small, CR hardening may be responsible for a much more significant increase of the gamma-ray flux in the central region $-60^\circ \leq l \leq 60^\circ$. We see indeed from Fig. 2 that the enhancement factor can be as large as $\sim 2$ in the direction of the Galactic center. This factor is larger than the uncertainty due to CR spatial distribution that can be conservatively estimated from the width of the coloured bands in the figure. Additional uncertainty sources are the normalization of the local CR flux at $E \sim 10\,E_\gamma \sim 10\,\mathrm{TeV}$, the distribution of gas in the Galaxy, the hadronic interaction cross section, etc. All these are expected to produce a total error smaller than the difference between red and blue lines in the figure, suggesting that effects of CR hardening can be probed by TeV scale gamma-ray observations of the galactic central region. A detailed comparison of the obtained predictions with the observational data in the TeV domain by Argo-YBJ, HESS, Milagro and HAWC experiments is performed in Ref. 11.

Finally, the diffuse neutrino flux at $E_\nu = 100\,\mathrm{TeV}$ is given as a function of the galactic longitude and latitude in Fig. 3 where we also show with a green line the isotropic astrophysical neutrino flux that is obtained from the six years HESE event rate observed by IceCube[13]. Our new calculations provide $\sim 40\%$ smaller neutrino

[c] Refs. 4, 5 report the spectral index $\alpha_\gamma$ of gamma emission associated to $\pi_0$ decay. This is converted into the spectral index $\alpha$ of the parent CR by $\alpha_\gamma \simeq \alpha + 0.1$.

fluxes than what obtained in our previuos work[1], as a combined consequence of the revised description of the local CR flux (which accounts for a factor $\sim 2$ reduction) and the inclusion of heavy element contribution to the galactic gas distribution (which is responsibile for a $\sim 40\%$ enhancement). The angle-integrated neutrino flux is equal to $\Phi_\nu = (1.0-1.2)\cdot 10^{-18}\ \mathrm{cm}^{-2}\ \mathrm{s}^{-1}\ \mathrm{GeV}^{-1}$ at 100 TeV in the standard scenario and it is equal to $\sim 3-4\%$ of the isotropic astrophysical flux observed by IceCube. The inclusion of CR hardening increases the integrated flux by $1.5-1.8$, the exact enhancement factor being dependent on the assumed CR spatial distribution. The effect is larger in the direction of the inner Galaxy where CR spectral hardening increases the predicted neutrino flux up to a factor $\sim 6$, producing a visibly different angular distribution with respect to the standard scenario.

## Acknowledgments

This work was partially supported by research grant number 2017W4HA7S "NAT-NET: Neutrino and Astroparticle Theory Network" under the program PRIN 2017 funded by Italian Ministero dell'Istruzione, dell'Università e della Ricerca (MIUR).

## References

1. G. Pagliaroli, C. Evoli and F. L. Villante, JCAP **1611** (2016) no. 11, 004.
2. D. Gaggero, D. Grasso, A. Marinelli, A. Urbano and M. Valli, Astrophys. J. **815** (2015) no. 2, L25.
3. F. Acero *et al.* [Fermi-LAT Collaboration], Astrophys. J. Suppl. **223** (2016) no. 2, 26.
4. R. Yang, F. Aharonian and C. Evoli, Phys. Rev. D **93** (2016) no. 12, 123007.
5. M. Pothast, D. Gaggero, E. Storm and C. Weniger, JCAP **1810** (2018) no. 10, 045.
6. S. R. Kelner and F. A. Aharonian, Phys. Rev. D **78** (2008) 034013 Erratum: [Phys. Rev. D **82** (2010) 099901].
7. A. Palladino, G. Pagliaroli, F. L. Villante and F. Vissani, Phys. Rev. Lett. **114**, no. 17, 171101 (2015) doi:10.1103/PhysRevLett.114.171101 [arXiv:1502.02923 [astro-ph.HE]].
8. http://galprop.stanford.edu/.
9. H. P. Dembinski, R. Engel, A. Fedynitch, T. Gaisser, F. Riehn and T. Stanev, PoS ICRC **2017** (2018) 533.
10. M. Ahlers, Y. Bai, V. Barger and R. Lu, Phys. Rev. D **93**, no. 1, 013009 (2016).
11. M. Cataldo, G. Pagliaroli, V. Vecchiotti and F. L. Villante, arXiv:1904.03894 [astro-ph.HE].
12. D. A. Green, Mon. Not. Roy. Astron. Soc. **454** (2015) no.2, 1517.
13. C. Kopper [IceCube Collaboration], PoS ICRC **2017**, 981 (2018).

# Gamma-ray emission from the near black hole environment in AGN

Katsoulakos, Grigorios* and Rieger, Frank M.

*Max-Planck-Institut für Kernphysik, P.O. Box 103980, 69029 Heidelberg, Germany*
*ZAH, ITA Universität Heidelberg, Philosophenweg 12, 69120 Heidelberg*
** E-mail: gkats@mpi-hd.mpg.de*

"Gap"-type particle acceleration in the vicinity of supermassive black holes (BHs), accompanied by curvature and Inverse Compton radiation, could in principle lead to variable gamma-ray emission that may be detectable with current instruments. We shortly comment on the occurrence of magnetospheric gaps at the jet base, and the realisation of different potentials. The detection of rapid variability becomes most instructive by imposing a constraint on possible gap sizes, thereby limiting extractable gap powers and allowing to assess the plausibility of a magnetospheric origin. The relevance of this is discussed for the radio galaxies M87 and IC310. The detection of magnetospheric gamma-ray emission generally allows for a sensitive probe of the near black hole region and is thus of prime interest for advancing our understanding of the (astro)physics of extreme environments.

*Keywords*: Acceleration of particles — galaxies: individual (M87, IC310) — gamma rays.

## 1. Introduction

With the experimental progress during the last few years the high energy processes occurring in the vicinity of a supermassive black holes (BHs) have gained particular attention.[2,4–9,13,14] The occurrence of strong, unscreened electromagnetic field components around rotating BHs could in principle facilitate an efficient particle acceleration to very high energies (Ref. 11 for a review). This process is naturally accompanied by $\gamma$-ray production via curvature emission (the radiation of charged particles moving along curved magnetic fields) and inverse Compton upscattering of ambient (accretion disk) soft photons. Provided that suitable conditions are met, the close BH environment could enable high (Blandford-Znajek-type) power extraction and account for rapid gamma-ray variability on timescales of $\sim r_g/c = 1.4(M_{BH}/10^9 M_\odot)$ hr and shorter (Ref. 2). A characteristic feature in this context is the occurrence of charge-deficient regions (the so-called "gaps") close to the BH.

## 2. "Gaps" around supermassive black holes

Magnetospheric emission models usually rely on efficient "gap"-type particle acceleration. According to Ohm's law, $\mathbf{J} = \sigma(\mathbf{E} + \mathbf{V}/c \times \mathbf{B})$, a deficiency of electric charges (i.e., low conductivity $\sigma$) within the BH magnetosphere can lead to the formation of regions with an unscreened parallel electric field component, i.e. where $\mathbf{E} \cdot \mathbf{B} \neq 0$. Thus, magnetospheric particles moving along magnetic fields into such charge-empty regions ("gap") can be strongly accelerated to high energies by these parallel electric field components.

The formation of "gaps" is closely associated with mechanisms for the charge density supply into the magnetospheric environment. The conditions under which charges are injected in the magnetosphere are still not fully understood. It seems likely, however, that the presence of a hot, radiatively inefficient accretion flow (RIAF) could facilitate particle injection via photon annihilation ($\gamma+\gamma \rightarrow e^+ + e^-$) of MeV disk photons (Bremsstrahlung peak). As it turns out, below a certain accretion rate the resulting charge density is not sufficient to screen the electric field, so that magnetospheric particle acceleration becomes a plausible process (Ref. 6).

## 3. The structure of magnetospheric "gaps"

It is usually believed that the magnetospheric structure of astrophysical BH can be reasonably approximated (at least in a time-averaged sense) by an electromagnetic force-free solution. Such a configuration is known to facilitates an efficient electromagnetic extraction of the rotational energy of the BH (Ref. 3). For this to be possible, however, some quasi-steady electric currents that pervade the magnetospheres as sources of the magnetic field and that are carried by charged particles flowing through them, need to be sustained despite the presence of inflows and outflows. If we suppose, to the contrary, that under-dense ($\rho_e < \rho_{GJ}$) regions ("gaps") are formed in which a parallel electric field $E_\parallel$ is established, this does not necessarily invalidate over-all force-freeness.

Vacuum "gaps" could occur around at least two locations, the null surface (NS) and the stagnation surface (SS) (Refs. 5, 7, 8). Figure 1 provides an illustration of their locations for an exemplary (split monopole) BH magnetosphere (split monopole). The former surface (NS) designates the region where the generalised Goldreich-Julian charge density $\rho_{GJ}$ vanishes, changing sign across it (red solid line). This happens close to the location where the field line rotation frequency $\Omega_F$ equals the Lense-Thirring angular frequency $\omega(r)$, i.e., usually on radial scales $r \sim r_g = GM/c^2$. In order for the BH magnetosphere to be force-free (vanishing $E_{||}$) the real charge density $\rho_e$ should correspond to $\rho_{GJ}$. As $\rho_{GJ}$ changes signs across the null surface, $\rho_e$ is required to have opposite signs on opposite sites, hence a parallel electric field component could easily arises around this surface (Refs. 5, 7). The stagnation surface (SS), on the other hand, naturally occurs in an MHD outflow driven by a rotating (Kerr) BH and designates the surface that separates plasma motion inwards (inflows) due to the gravitational field from plasma motion outwards (outflows) above it (green solid line). This surface is in general non-spherical and located inside the (outer) light cylinder, typically on radial scales of some $r_g$. To allow for a general (time-averaged) force-free MHD description, plasma would need to be continuously replenished to maintain a charge density $\rho_e \geq \rho_{GJ}$.

## 4. Different realisations of the "gap" potential

In its simplest (one-dimensional, non-relativistic) form the "gap" electric field $E_\parallel$ and potential $\Phi_e$ along a path length $s$ in the presence of a non-zero charge density

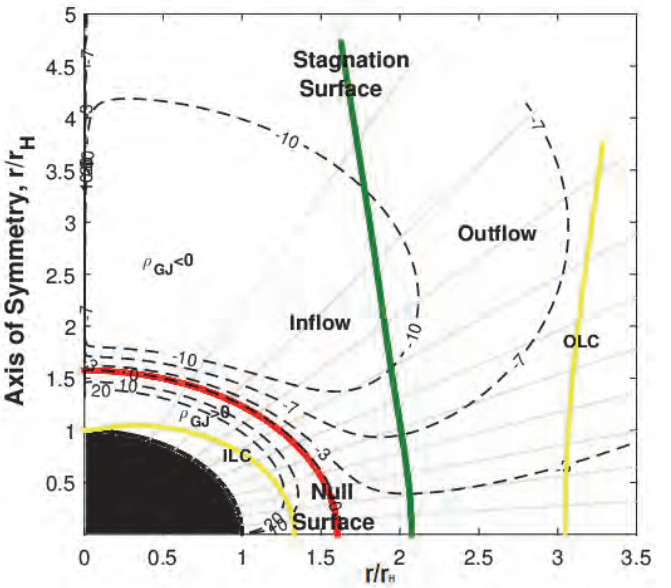



Fig. 1. Structure of the BH magnetosphere for a split-monopole and field line rotation frequency $\Omega_F = c/4r_g$. Gray solid lines represent field lines and dashed black lines denote contour levels of Goldreich-Julian density. Between the inner and outer light cylinder (yellow lines) the null (red) and stagnation (green) surfaces are located. Across the null surface the charge density changes sign, while the stagnation surface separates matter inflow and outflow, respectively.

$\rho_e$ can be determined from Gauss law

$$\frac{dE_{||}}{ds} = 4\pi(\rho_e - \rho_{GJ}) \quad \text{with} \quad \frac{d\Phi_e}{ds} = -E_{||}\,, \tag{1}$$

so that the relevant voltage drop becomes $\Delta V_{gap} = \Phi_e(s = h) - \Phi_e(s = 0)$, where $h$ denotes the characteristic gap height. It is easily seen that a "gap" emerges only if the real charge density $\rho_e$ diverges from the expected $\rho_{GJ}$.

Different solutions are obtained however, dependent on what type of boundary conditions (e.g., highly or weakly under-dense) are considered to be realised, as we show in Ref. 8. In general, the resultant voltage drop can be written as $\Delta V_{gap} = \Phi_o(h/r_g)^\nu$, with $\Phi_o \approx \Omega_F r_g^2 B_H/c$, where $B_H$ corresponds to the strength of the normal magnetic field component threading the horizon, and the coefficient $\nu$ is dependent on the chosen "gap" scenario:

(a) if $\rho_e << \rho_{GJ}$, for example, then

$$\Delta V_{gap} = \Phi_o \left(\frac{h}{r_g}\right)^2, \tag{2}$$

(b) while for $\rho_e \sim \rho_{GJ}$ one instead finds

$$\Delta V_{gap} = \Phi_o \left(\frac{h}{r_g}\right)^3, \tag{3}$$

i.e., a scaling $\propto h^3$ where the power index is now increased by one.

## 5. VHE gamma-ray emission from magnetospheric "gaps"

If the potential is sufficiently strong as expected around supermassive BHs, a single particle entering the gap can be accelerated to very high energies (up to Lorentz factors $\gamma \sim 10^{10}$), emitting curvature and inverse Compton gamma-ray photons on its way.[8] In the presence of an ambient photon field, a pair cascade ($\gamma + \gamma \rightarrow e^+ + e^-$) is triggered, that can lead to vacuum breakdown and the formation of highly conducting pair plasma, i.e., to enough charges ($\rho_e \geq \rho_{GJ}$) to annihilate the parallel electric field and ensure gap closure $\mathbf{E} \cdot \mathbf{B} = 0$.[6] As a consequence, the magnetic field lines can be considered as nearly orthogonal to the $\mathbf{E}$ field, with just enough $\mathbf{E} \cdot \mathbf{B}$ remaining to produce sparks of $e^+e^-$-pairs and keep the magnetosphere filled with plasma. An electromagnetic force-free solution could then provide a reasonable approximation to the time-averaged structure of such a magnetosphere.[3]

## 6. Associated "gap" luminosities

Given the anticipated potential strengths around supermassive BHs a significant amount of the non-thermal emission of magnetospheric "gaps" is expected to occur in the high and very high energy $\gamma$-ray domain (Refs. 11, 6, 5, 9). The different "gap" potential noted above, however, lead to different expectations for the maximum "gap" luminosity $L_{gap} \approx n_e V_{gap} dE_e/dt$ (with characteristic gap volume $V_{gap} \propto r_g^2 h$) for a gap of height $h$. As we show in (Ref. 8) the maximum extractable "gap" power is in general proportional to the classical Blandford-Znajek jet power, $L_{BZ} \propto r_g^2 B_H^2 \propto \dot{m} M_{BH}$ (with $B_H \propto \dot{m}^{1/2}$, $\dot{m}$ being the accretion rate), and a sensitive function of the gap height $h$, i.e.,

$$L_{gap} \sim \eta_\beta L_{BZ} \left(\frac{h}{r_g}\right)^\beta , \tag{4}$$

where the index $\beta \geq 1$ depends on the respective "gap"-setup: for case (a) noted above (Eq. (2)) one has $\beta = 2$ with $\eta_2 = 1$, while for case (b) $\beta = 4$ with $\eta_4 = 1/6$.

## 7. The phenomenological relevance of magnetospheric "gaps"

The anticipated occurrence of magnetospheric gamma-ray emission in AGN in principle allows for a fundamental probe of their near BH environment including accretion physics and jet formation. For this to be feasible, however, several constraints need to be met. On the one hand, the escape and detectability of VHE gamma-rays up to a few TeV at least requires the presence of a radiatively inefficient accretion flow (RIAF) close to the BH as otherwise severe $\gamma\gamma$-absorption in the disk photon field will occur (Ref. 8). On the other hand, in classical blazar-type sources (with jets viewed at small inclination) the observed VHE gamma-ray emission is expected to be dominated by the strongly Doppler-boosted emission of their jets. Hence

only in sources where the jet is sufficiently misaligned, such that Doppler effects are modest, might magnetospheric emission become detectable. This has made misaligned and under-luminous AGN, in particular nearby radio galaxies, to the most promising targets (see e.g. Ref. 10 for review and further references). The occur-

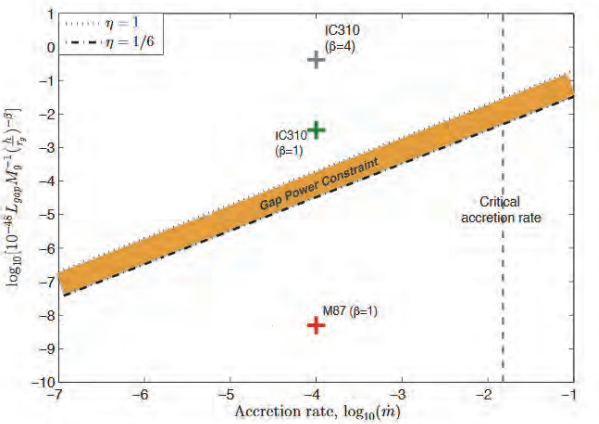



Fig. 2. Maximum possible $\gamma$-ray power (orange band) of a magnetospheric "gap" as a function of accretion rate $\dot{m}$, cf. Eq. (4). Different assumptions for the gap potential, Eqs. (2) and (3), lead to differences in extractable powers. The required transparency of the accretion environment to VHE photons introduces an upper limit $\dot{m}_c \leq 0.01$ on possible rates. If the observed VHE variability is used to constrain the gap height $h$, the extractable power is in principle sufficient to account for the gamma-ray emission seen from M87, yet under-predicts the VHE emission seen from IC 310.

rence of magnetospheric emission in these sources could contribute to the spectral hardening at GeV energies seen in some sources[12] and possibly account for rapid VHE variability on horizon crossing-times and shorter.[2,5] As the extractable power depends on the gap size $h$, Eq. (4), the plausibility of a magnetospheric origin can be investigated for sources in which significant VHE variability is detected.

In the radio galaxy M87 (distance $d = 16.7$ Mpc), for example, day-scale VHE variability has been detected during several TeV high states (with $L_{VHE} \simeq 10^{41} erg/s$). There are also indications that the TeV emission is accompanied by delayed radio core flux enhancements supporting the conclusion that the VHE emission may originate at the jet base very close to the BH (Ref. 1). One typically expects magnetospheric "gaps" to possess maximum heights of $h \lesssim r_g$ (if efficient pair production takes place, screening may well occur earlier, see Ref. 6). M87 has a BH mass of $M_{BH} = (2-6) \times 10^9 M_\odot$, corresponding to light crossing times of $\Delta t_H = r_g/c \simeq (3-9)$ hr. The VHE variability observed to date does not impose a stronger constraint on $h$. To ensure transparency to VHE photons the accretion flow would need to be radiatively inefficient (as indeed expected in M87) and this on average constrains inner accretion rates to satisfy $\dot{m}_c \leq 0.01$, see e.g. Ref. 8 for details. When this is put in context, extractable gap powers are such as to allow accommodation of the $\gamma$-ray emission seen from M87, see Figure 2.

The situation is different for the Perseus Cluster radio galaxy IC 310 (at $d = 80$ Mpc), believed to host a BH of $M_{BH} = 3 \times 10^8 M_\odot$. The minute-scale VHE variability seen during a strong VHE flare in Nov. 2012 (with isotropic $L_{VHE} = 2 \times 10^{44}\ erg/s$) would imply a gap height $h \leq c\Delta t = 0.2 r_g$ (Ref. 2). When this is employed in Eq. (4) along with the ADAF constraint $\dot{m}_c \leq 0.01$, extractable powers tend to become too small to account for the observed VHE emission (see Figure 2), thus disfavouring conventional magnetospheric scenarios for its origin.

## 8. Conclusions

Non-thermal magnetospheric processes could in principle lead to a non-negligible contribution at gamma-ray energies that could vary on horizon crossing times. As this contribution is non-boosted, potential targets would need to be close enough and possess jets sufficiently misaligned for this emission to become detectable by current instruments. The extractable gap power is a sensitive function of the gap height ($h/r_g$). Detection of rapid variability on timescales smaller than $r_g/c$ thus becomes most constraining and could allow to probe different gap descriptions. When put in context of recent observations, this suggests that the gamma-ray emission seen from M87 may have a magnetospheric origin, while such a scenario appears rather disfavoured in the case of IC 310.

## Acknowledgments

Support by the International Max Planck Research School for Astronomy and Cosmic Physics at the University of Heidelberg (IMPRS-HD) is greatly acknowledged.

## References

1. V. A. Acciari, E. Aliu, T. Arlen, et al., *Science*, **325**, 444, (2009).
2. J. Aleksić, S. Ansoldi, L. A. Antonelli, et al., *Science*, **346**, 1080, (2014).
3. R. D. Blandford, R. L. Znajek, *MNRAS*, **179**, 433 (1977).
4. A. E. Broderick, A. Tchekhovskoy, *ApJ*, **809**, 97, (2015).
5. K. Hirotani, H.-Y. Pu, *ApJ*, **815**, 50, (2016).
6. A. Levinson, F.M. Rieger, *ApJ*, **730**, 123, (2011).
7. A. Levinson, N. Segev, *PhRvD*, **96**, 123006, (2017).
8. G. Katsoulakos, F.M. Rieger, *ApJ*, **852**, 112, (2018).
9. K. Ptitsyna, A. Neronov, *A&A*, **593**, 8 (2016).
10. F. M. Rieger, A. Levinson, *Galaxies*, **6**, 116 (2018).
11. F.M. Rieger, *Int. J. Mod. Phys. D*, **20**, 1547, (2011).
12. N. Sahakyan, R. Yang, F. A. Aharonian, F. M. Rieger, *ApJL*, **770**, 6, (2013).
13. A. Levinson, B. Cerutt, *A&A*, **616**, A184, (2018).
14. A.Y. Chen, Y. Yuan, H. Yang, *ApJ*, **863**, L31, (2018).

# Searching for high-energy neutrino emitter active galactic nuclei

E. Kun

*Institute of Physics, University of Szeged,*
*H-6720 Szeged, Dóm tér 9., Hungary*
*E-mail: kun@titan.physx.u-szeged.hu*

P. L. Biermann

*Max-Planck-Institute for Radioastronomy,*
*D-53121 Bonn, Auf dem Hügel 69, Germany*
*E-mail: plbiermann@mpifr-bonn.mpg.de*

L. Á. Gergely

*Institute of Physics, University of Szeged,*
*H-6720 Szeged, Dóm tér 9., Hungary*
*E-mail: gergely@physx.u-szeged.hu*

We employ two radio surveys to find young energetic active galactic nuclei (AGN) with flat spectrum, and search for their counterpart in the Second Planck Catalogue of Compact Sources (PCCS2) at 353 GHz. We find 21 and 6 such AGN in the S4 and S5 Strong Source Surveys, respectively. Although none of them are within the error circle about the sky position of the track-type high-energy neutrinos already detected by the IceCube South Pole Neutrino Observatory, their flat spectrum and high-frequency detection imply they might emit neutrinos to be observed in the near future. Extending the list of these sources with the 114 flat spectrum AGN appearing with PCCS2-counterpart we already found in the Parkes Catalogue, we conclude that most of them have flux density between $400 \div 1400$ mJy at 5 GHz. This range is a good selection criterion for the future campaigns at 5 GHz aiming to find the possible sources of high-energy neutrinos.

*Keywords*: Active galactic nuclei; high-energy neutrino; catalogues.

## 1. Introduction

In the last years extensive studies have been carried out aiming at identifying the origin of the cosmic high-energy (HE) neutrinos detected by the IceCube South Pole Neutrino Observatory[1–5]. Active galactic nuclei (AGN), powered by supermassive black holes (SMBHs), are promising candidates, as they are able to produce the dominant isotropic neutrino background between $10^4 \div 10^{10}$ GeV[6].

IceCube is located at the Amundsen-Scott South Pole Station, and the observatory uses one cubic kilometre of the Antarctic ice as detector matter. The detector strings penetrate up to 2500 m deep underground, and are sensitive to the Cherenkovradiation of electrons and muons emerging from the interaction of electron and muon neutrinos with the ice. Electrons induce shower-type, while muons induce track-type event topologies. The arrival direction of the parent neutrino is constrained more precisely for the track-type events (within 2°).

Track-type HE neutrinos pinpoint their specific sources, the blazars, a class of radio-loud AGN pointing their jet towards the Earth. We first demonstrated[7] that in the case of the track-type neutrino ID5 and blazar PKS 0723-008 by cross-

correlating the Parkes Catalogue[8] and the Second Planck Catalogue of Compact Sources (PCCS2)[9] at 353 GHz with the arrival direction of track-type HE neutrino events detected by IceCube[2,10,11].

Recently the IceCube Collaboration with partner observatories on Earth and in space has reported on the multimessenger observations of the blazar TXS 0506+056[12]. The HE neutrino event IC-170922A detected at 22/9/2018 was in temporal and positional coincidence with a prominent $\gamma$-ray flare shown by this source. Subseqently we reported on the radio brightening of the neutrino emitting blazar TXS 0506+056[13], supporting its identification as source of the HE neutrino IC-170922A. The radio brightening of TXS 0506+056 preceding the detection of a HE neutrino is similar to the one we reported for PKS 0723-008 and IceCube event ID5[7], suggesting something similar happened with these blazars. The high energy of these neutrinos imply violent events as progenitors, such as the coalescence of compact objects, as we already suggested[7,13,14]. The published candidates for track-type IceCube neutrinos, with the neutrino ID-s in brackets: PKS 0723-008[7] (ID5), PKS B1206-202[7] (ID8), PKS B2300-254[7] (ID18), PKS B2224+006[7] (ID44), TXS 0506+056[12] (IC-170922A), GB6 J1040+0617[15] (IC-141209A), all six of them are flat spectrum sources at GHz-frequencies.

Here we investigate two other radio catalogues, the S4 Strong Source Survey[16] and the S5 Strong Source Survey[17], to identify further possible sources of track-type HE neutrinos. We report on the results here, also on the integrated flux density of the flat spectrum AGN from these surveys, as well as from the Parkes Catalogue[8], that have counterparts in the PCCS2[9] at 353 GHz.

## 2. Flat-spectrum AGN with Planck-counterpart at 353 GHz

An enormous shock-wave generated in the coalescence of two compact objects accelerates both the light and the heavy particles in their immediate environment. The former process leads to enhanced synchrotron emission and flat spectrum up to high radio frequencies ($\sim$ THz) due to the relativistic electrons, while the latter process leads to the emission of ultra-high energy cosmic rays (UHECRs) and HE neutrinos due to the relativistic protons and nuclei. The gravitational wave (GW) signal of a merger is accompanied by huge flares emerging across all regimes of the EM spectrum, as well as by significant excess in HE particles. Due to their much larger mass the coalescence of two SMBHs produces much more energetic GW signal as compared to the merger of astrophysical black holes or neutron stars. Consequently a more violent shock-wave propagates through the immediate environment of the SMBH merger, capable to induce the highest energy particles, and therefore we focused on AGN.

We searched for flat-spectrum AGN in the S4 Strong Source Survey[16] (flux limit 500 mJy, sky coverage $+35^0 < \delta < +70^0$, where $\delta$ is the decliantion), in the S5 Strong Source Survey[17] (flux limit 250 mJy, sky coverage $+70^0 < \delta < +90^0$), and in the Parkes Catalogue[8] (flux limit 250 mJy, sky coverage all the sky south of

declination $+27^0$, excluding the Galactic Plane and the Magellanic Cloud regions). The spectral index $\alpha$ was defined with the convention $S \sim \nu^{+\alpha}$ between 2.7 GHz and 5 GHz. We cross-correlated the coordinates of flat spectrum AGN from these catalogues with the coordinates from the Second Planck Catalogue of Compact Sources (PCCS2)[9]. We used data from the PCCS2 taken at 353 GHz, which is the highest frequency measurements at which the contamination from the Galactic plane can be safely corrected. The error on the sky position of the point sources at this frequency is 8 arcmin. Then we cross-correlated the coordinates of flat spectrum AGN having PCCS2 counterpart with the coordinates of the 15 track-type HE neutrinos detected by the IceCube, considering the errors on the positions as well. The detection data of these neutrino events are summarized in Table 1.

Table 1. Data of the 15 track-type HE neutrino events detected by the IceCube[2,10,11]. ID: identification, $E_{dep}$: deposited energy, $RA$: right-ascension, $\delta$: declination, $\theta$: median angular error.

| ID | $E_{dep}$ $(TeV)$ | $RA$ $(^0)$ | $\delta$ $(^0)$ | $\theta$ $(^0)$ |
|---|---|---|---|---|
| 3 | $78.7^{+10.8}_{-8.7}$ | 127.9 | -31.2 | $\lesssim 1.4$ |
| 5 | $71.4^{+9.0}_{-9.0}$ | 110.6 | -0.4 | $\lesssim 1.2$ |
| 8 | $32.6^{+10.3}_{-11.1}$ | 182.4 | -21.2 | $\lesssim 1.3$ |
| 13 | $253^{+26}_{-22}$ | 67.9 | 40.3 | $\lesssim 1.2$ |
| 18 | $31.5^{+4.6}_{-3.3}$ | 345.6 | -24.8 | $\lesssim 1.3$ |
| 23 | $82.2^{+8.6}_{-8.4}$ | 208.7 | -13.2 | $\lesssim 1.9$ |
| 28 | $46.1^{+5.7}_{-4.4}$ | 164.8 | -71.5 | $\lesssim 1.3$ |
| 37 | $30.8^{+3.3}_{-3.5}$ | 167.3 | 20.7 | $\lesssim 1.2$ |
| 38 | $200.5^{+16.4}_{-16.4}$ | 93.34 | 13.98 | $\lesssim 1.2$ |
| 43 | $46.5^{+5.9}_{-4.5}$ | 206.63 | -21.98 | $\lesssim 1.3$ |
| 44 | $84.6^{+7.4}_{-7.9}$ | 336.71 | 0.04 | $\lesssim 1.2$ |
| 45 | $429.9^{+57.4}_{-49.1}$ | 218.96 | -86.25 | $\lesssim 1.2$ |
| 47 | $74.3^{+8.3}_{-7.2}$ | 209.36 | 67.38 | $\lesssim 1.2$ |
| 53 | $27.6^{+2.6}_{-2.2}$ | 239.02 | -37.73 | $\lesssim 1.2$ |
| 55 | $2600^{+300}_{-300}$ | 110.34 | 11.48 | $\lesssim 1.0$ |

## 3. Discussion

We found 21/6/114 flat spectrum AGN with PCCS2-counterpart at 353 GHz in the S4-S5 Surveys and in the Parkes Catalogue, respectively. The map of the flat spectrum AGN with PCCS2-counterparts and the track-type HE neutrino events are presented in Figure 1 in galactic coordinates. The number of these sources in 200 mJy-bins is presented in Figure 2, separately at 2.7 GHz and at 5 GHz. Most of these AGN have flux density between $400 \div 1400$ mJy at 5 GHz. This range might be a good selection criterion for future campaigns at 5 GHz, aiming to find the possible sources of HE neutrinos.

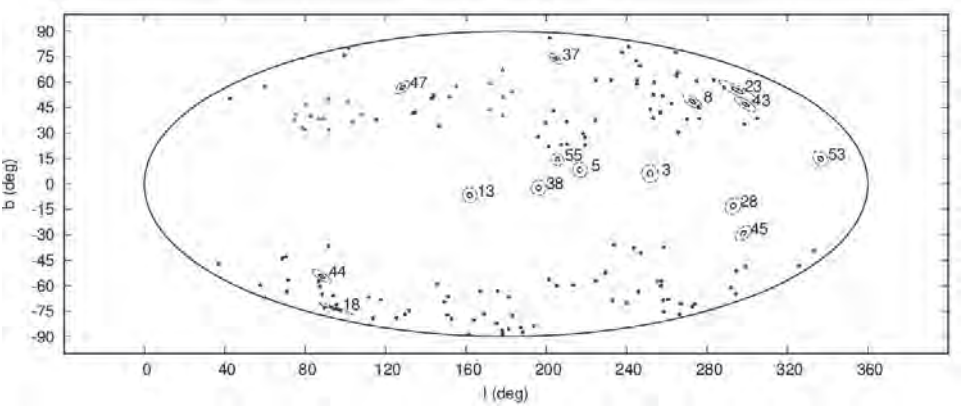



Fig. 1. Map of flat spectrum AGN with Planck-counterpart at 353 GHz in galactic coordinates, employing Mollweide projection. The error circle of the track-type neutrino events (denoted by their ID-number) is presented by continuous black line and three times it's size by a dashed line. The event ID5 is marked by a red cross.

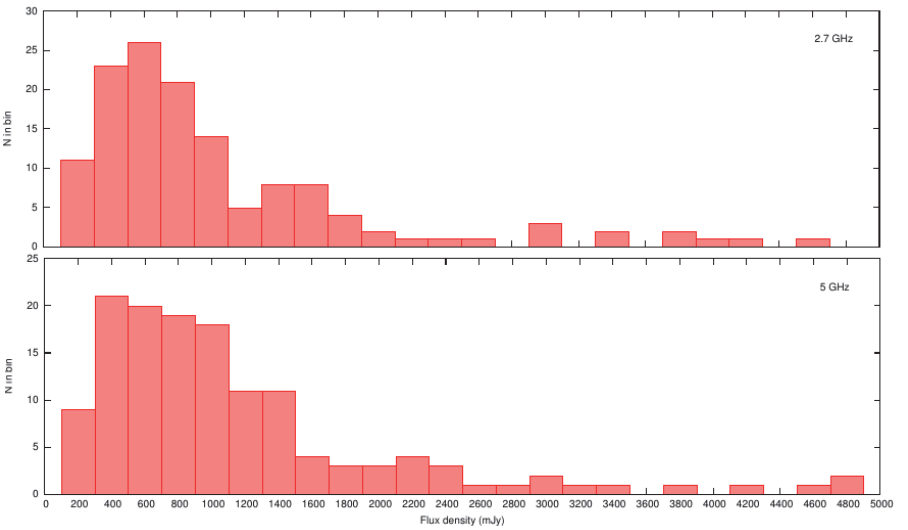



Fig. 2. The number of flat spectrum AGN with Planck-counterpart at 353 GHz, binned by their flux density measured at 2.7 GHz (upper panel) and at 5 GHz (bottom panel). There are 6, respectively 4 AGN that are brighter than 5000 mJy at 2.7 GHz, respectively 5 GHz.

We proposed a scenario of binary SMBH evolution, which naturally explains the observed HE neutrino emission leading to the emission of GWs through a sequence induced by the merger[7,13,14]. We expect that there are three main phases of the emission of HE particles in this scenario. The first one is the process of

spin-flip, when the jets sweep through a large cone (see M82 where this may have been detected,[18,19]). The second one is after the spin-flip, when a new jet is boring into the environment, leading to more injection of the seed particles to create HE nuclei, gammas and neutrinos. The third one is probably in the instant of the coalescence of SMBHs, when a giant shock-wave may be generated by low-frequency GWs (lfGWs) to accelerate particles to high energies, leading to a final burst of HE nuclei, gammas and neutrinos. Only the HE neutrinos are detectable across the universe. Considering TXS 0506+056, the post-merger SMBH binary scenario provides a consistent physical picture explaining the already occurred multimessenger observations (radio, $\gamma$-ray, and HE neutrino).

The next-generation IceCube detector, the future ten-cubic-kilometre detector IceCube-Gen2, is a very promising development regarding this work. This instrument is aimed to detect very high-energy astrophysical neutrinos in the PeV to EeV range, yielding hundreds of neutrinos across all flavours at energies above 100 TeV. This will enable detailed spectral studies, significant point source detections, important for quantifying the HE particle emission of the individual candidate-sources. The IceCube-Gen2 spectra together with the EM-observations will teach as how to find AGN harbouring SMBHs, that will merge within a very short time emitting GWs.

A great challenge will be the direct detection of these lfGWs (of order of microHz to milliHz) by the future Laser Interferometer Space Antenna (LISA) mission. The ultimate proof of the above scenario will be this direct detection of lfGW-signal of the candidate sources.

## Acknowledgments

The authors acknowledge the support of the Hungarian National Research, Development and Innovation Office (NKFIH) in the form of the grant 123996.

## References

1. IceCube Collaboration, Probing the origin of cosmic rays with extremely high energy neutrinos using the IceCube Observatory, *PRD* **88**, p. 112008 (December 2013).
2. IceCube Collaboration, Observation of High-Energy Astrophysical Neutrinos in Three Years of IceCube Data, *Physical Review Letters* **113**, p. 101101 (September 2014).
3. ANTARES Collaboration, ANTARES constrains a blazar origin of two IceCube PeV neutrino events, *A&A* **576**, p. L8 (April 2015).
4. IceCube Collaboration, All-sky Search for Time-integrated Neutrino Emission from Astrophysical Sources with 7 yr of IceCube Data, *ApJ* **835**, p. 151 (February 2017).
5. IceCube Collaboration, Search for Astrophysical Sources of Neutrinos Using Cascade Events in IceCube, *ApJ* **846**, p. 136 (September 2017).

6. F. W. Stecker, C. Done, M. H. Salamon and P. Sommers, High-energy neutrinos from active galactic nuclei, *Physical Review Letters* **66**, 2697 (May 1991).
7. E. Kun, P. L. Biermann and L. Á. Gergely, A flat-spectrum candidate for a track-type high-energy neutrino emission event, the case of blazar PKS 0723-008, *MNRAS Lett.* **466**, L34 (March 2017).
8. A. Wright and R. Otrupcek, Parkes Catalog, 1990, Australia telescope national facility., in *PKS Catalog (1990)*, 1990.
9. Planck Collaboration, Planck 2015 results. XXVI. The Second Planck Catalogue of Compact Sources, *A&A* **594**, p. A26 (September 2016).
10. IceCube Collaboration, The IceCube Neutrino Observatory - Contributions to ICRC 2015 Part II: Atmospheric and Astrophysical Diffuse Neutrino Searches of All Flavors, *ArXiv e-prints* (October 2015).
11. S. Schoenen and L. Raedel, Detection of a multi-PeV neutrino-induced muon event from the Northern sky with IceCube, *The Astronomer's Telegram* **7856** (July 2015).
12. IceCube Collaboration, Fermi-LAT, MAGIC, AGILE, ASAS-SN, HAWC, H.E.S.S., INTEGRAL, Kanata, Kiso, Kapteyn, Liverpool Telescope, Subaru, Swift/NuSTAR, VERITAS and VLA/17B-403, Multimessenger observations of a flaring blazar coincident with high-energy neutrino IceCube-170922A, *Science* **361**, p. 147 (July 2018).
13. E. Kun, P. L. Biermann and L. Á. Gergely, Very long baseline interferometry radio structure and radio brightening of the high-energy neutrino emitting blazar TXS 0506+056, *MNRAS Letters* **483**, L42 (February 2019).
14. E. Kun, P. Biermann, S. Britzen and L. Gergely, On the High-Energy Neutrino Emission from Active Galactic Nuclei, *Universe* **4**, p. 24 (February 2018).
15. Fermi-LAT collaboration, :, S. Garrappa, S. Buson, A. Franckowiak, ASAS-SN, :, B. J. Shappee, J. F. Beacom, S. Dong and et al., Investigation of two Fermi-LAT gamma-ray blazars coincident with high-energy neutrinos detected by IceCube, *arXiv e-prints* (January 2019).
16. I. I. K. Pauliny-Toth, A. Witzel, E. Preuss, H. Kühr, K. I. Kellermann, E. B. Fomalont and M. M. Davis, The 5 GHz strong source surveys. IV - Survey of the area between declination 35 and 70 degrees and summary of source counts, spectra and optical identifications, *AJ* **83**, 451 (May 1978).
17. H. Kühr, I. I. K. Pauliny-Toth, A. Witzel and J. Schmidt, The 5-GHz strong source surveys. V - Survey of the area between declinations 70 and 90 deg, *AJ* **86**, 854 (June 1981).
18. P. P. Kronberg, P. Biermann and F. R. Schwab, The nucleus of M82 at radio and X-ray bands - Discovery of a new radio population of supernova candidates, *ApJ* **291**, 693 (April 1985).
19. P. L. Biermann, J. Becker Tjus, W. de Boer, L. I. Caramete, A. Chieffi, R. Diehl, I. Gebauer, L. Á. Gergely, E. Haug, P. P. Kronberg, E. Kun, A. Meli, B. B. Nath and T. Stanev, Supernova explosions of massive stars and cosmic rays, *Advances in Space Research* **62**, 2773 (November 2018).

# Hunting for transients with the SVOM mission

Diego Götz*

*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France*
**E-mail: diego.gotz@cea.fr*

Jianyan Wei

*National Astronomical Observatories/Chinese Academy of Science 20A Datun Road, Beijing, 100012, China*

Bertrand Cordier

*AIM, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, F-91191 Gif-sur-Yvette, France*

on behalf of the SVOM collaboration

The Space Variable astronomical Object Monitor (SVOM) is a mission dedicated to the detection and characterization of Gamma-Ray Bursts and other high-energy astrophysical transients. SVOM is jointly developed by the Chinese Academy of Science (CAS), the Chinese National Space Administration (CNSA) and the French space agency (CNES) with a launch date foreseen at the end of 2021. SVOM will enable to observe new sources in a large wavelength domain, from the near infrared to gamma rays, thanks to its unique combination of space and ground based instruments. The space borne instruments include two wide field of view monitors, ECLAIRs and GRM, operating in the hard X-ray to gamma-ray energy band, and two narrow field telescopes, MXT and VT, operating in the X-ray and visible domain. On the ground three dedicated robotic telescopes, F-GFT-Colibri, C-GFT, and GWACs will provide complementary coverage in the near infrared and visible bands. SVOM alerts will be distributed publicly to the scientific community in order to enhance the scientific return of the mission.

*Keywords*: Gamma-Ray Bursts; Space Instrumentation.

## 1. Introduction

The Space Variable astronomical Object Monitor (SVOM) mission is being jointly developed by the Chinese Academy of Science (CAS), the Chinese National Space Administration (CNSA) and the French space agency (CNES), and its launch date is foreseen for the end of 2021. The science goals are shared between a Core Programme (CP), a General Programme (GP), and a Target of Opportunity (ToO) Programme. The goals of the CP include the detection and characterization of Gamma-Ray Bursts (GRBs) with a particular emphasis on the detection of high redshift ($z > 5$) events and their use as probes of the early Universe, the broad band coverage of the GRB prompt emission, and the building of a uniform afterglow data sample from the X-rays to the near infrared. The CP will make use of 25% of the SVOM observation time. During the rest of the time SVOM will be operated as an observatory performing pre-planned (GP, 60% of the SVOM time) and un-anticipated ToO observations (15% of the time). This sharing applies to the

nominal lifetime of the mission (i.e. the first three years on orbit) and will evolve during the extend mission including a higher fraction of ToO time (up to 40% of the SVOM observation time). The reason for it is that SVOM is also suited to play a role in the upcoming era of multi-messenger astronomy, being able to provide follow-up of triggers related to gravitational wave, neutrino, very high energy gamma-ray or synoptical astronomy.

In order to favour the follow-up of SVOM GRBs by ground based observatories, SVOM will be mainly pointing to the anti-solar direction. In addition, in order to maximize the GRB detection sensitivity, at least during the nominal mission lifetime, the SVOM pointing has been defined as to avoid the Galactic plane (which in turn facilitates the follow-up from ground) and other bright X-ray sources (such as Sco X-1). Thanks to this pointing strategy we expect that about 75% of the SVOM GRBs can be followed promptly from ground, increasing the chances of measuring their redshift through spectroscopy.

For a detailed view of the SVOM science objectives we refer the reader to the SVOM white paper [1]. In this paper we will mainly review the SVOM space payload and the associated ground segment.

## 2. The SVOM Space Segment

SVOM will be launched from Xichang in China by a LM-2C rocket and injected into a circular low Earth orbit with an inclination of about 30° at an altitude of ∼625 km. It will carry two wide field monitors and two narrow field telescopes. The wide field monitors are ECLAIRs, a coded mask telescope operating in the 4–250 keV energy range, provided by CNES in collaboration with French laboratories (IRAP, APC, CEA-Irfu), and the GRM, a set of three non-imaging gamma-ray spectrometers, sensitive in the 15 keV–5 MeV energy range, provided by IHEP Beijing. The narrow field telescopes are the MXT, sensitive in the 0.2–10 keV energy range provided by CNES in collaboration with CEA-Irfu, LAL, MPE, and the University of Leicester, and the Visible Telescope (VT), provided by the NAOC in collaboration with XIOPM.

### 2.1. *ECLAIRs*

ECLAIRs is composed by a focal plane of 80×80 CdTe pixels of 4×4×1 $mm^3$ size. A 54×54 $cm^2$ mask with an open fraction of 40% and a pixel size of about 1.1 cm is placed 46 cm above the detection plane. The system provides a ∼2 sr field of view and a point spread function of 52 arc min (FWHM). The ECLAIRs data stream (monitored on time scales from 10 ms to about 20 minutes on 4 energy bands and in 9 detector zones) is analysed in real time on board the satellite by the UGTS electronic box in order to search for new sources appearing in the field on view. Once a new source is detected its coordinates are computed and transmitted to the satellite in order to send the information to ground based observatories and to slew

to place the ECLAIRs error box (<12 arc min) into the field of view of the narrow filed telescopes for more accurate localization and physical characterization in the X-rays and in the visible domain. ECLAIRs will be able to detect all kind of GRBs, see Fig. 1, but thanks to its low energy threshold of only 4 keV it will be more sensitive than previous experiments to X-ray rich GRBs which could represent the population of high distance GRBs due to cosmological energy shift[2]. ECLAIRs is expected to detect about 60 GRBs per year.

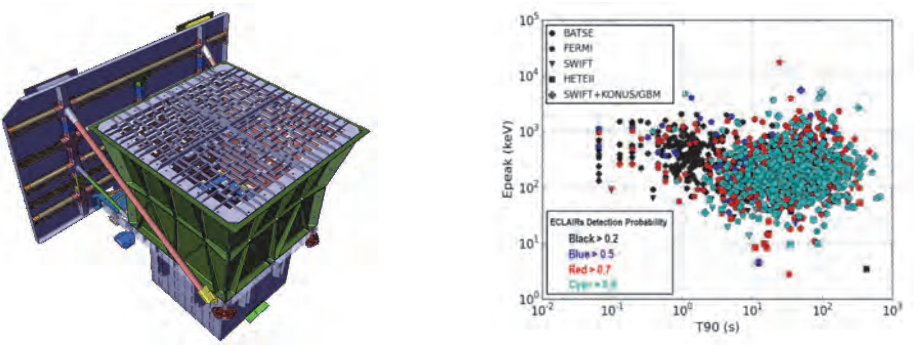



Fig. 1. Left: the ECLAIRs coded mask telescope. Right: the probability of detection by ECLAIRs (color code) of a sample of GRBs observed by previous experiments, plotted as a function of the GRB duration and spectral peak energy.

### 2.2. *GRM*

The Gamma-Ray Monitor (GRM) system is composed by three Gamma-Ray Detectors (GRD) and a particle monitor. Each GRD is made of a NaI(Tl) scintillating crystal, a photomultiplier and its readout electronics, and has a geometrical area of 200 $cm^2$, thickness of 1.5 cm, FoV of $\pm 60$ degrees with respect to its symmetry axis, dead time $<8$ $\mu$s, temporal resolution $<20$ $\mu$s and energy resolution of 16% at 60 keV. A plastic scintillator in front of the NaI(Tl) crystal is used to distinguish low energy electrons from gamma-rays. The 3 GRD-modules point at different directions in order to cover, globally, a field of view which is larger than the ECLAIRs one, see Fig. 2. This will allow to detect a larger number of GRBs ($\sim$90 per year), although most of them will be poorly localized (error box $> 10°$). For the GRBs detected both in the field of view of ECLAIRs and in the one of one or more GRD a combined spectral analysis will be possible from a few keV to a few MeV, allowing a complete physical description of the SVOM GRBs, as shown in the right panel of Fig. 2.

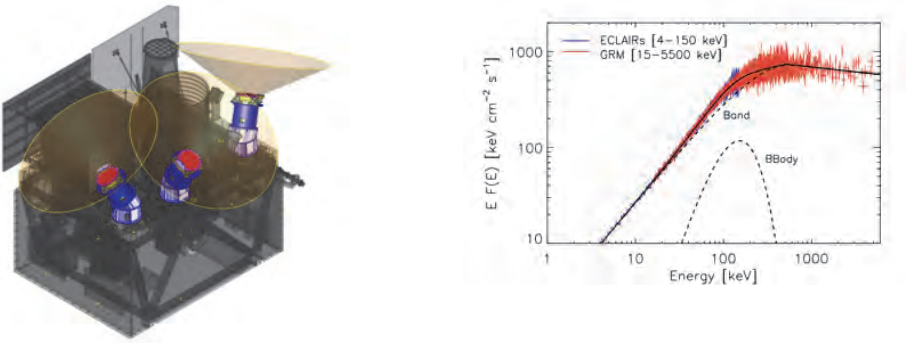



Fig. 2. Left: the GRDs and their individual field of views. Right: a simulation showing a simultaneous spectral fit of GRB 100724B. GRB 100724B was observed by Fermi/GBM and a thermal component (BBody in the figure) has been reported: this component can be detected in the simulated data with the same statistical significance with respect to the GBM data. For more details, see the SVOM white paper [1].

### 2.3. *MXT*

The Microchannel X-ray Telescope (MXT) [3] is an X-ray focussing telescope with a field of view of 57×57 arc min, whose innovative optical concept is based on the so-called *Lobster Eye* optics [4]. This optical concept is realized making use of square pore lead glass micro pore optics (MPOs) plates with 40 $\mu$m pore side and 1-2 mm thickness (with inner Ir coating to enhance the reflectivity), produced by Photonis. 25 plates of 40 mm side are used to build the complete MXT optics, which is coupled to a camera sensitive in the 0.2–10 keV energy range, based on a pnCCD [5] composed of 256×256 pixels of 75 $\mu$m side. The pnCCD and is read-out rate every 100 ms and actively cooled to –65°C. A filter wheel allows to put a calibration source or additional optical/UV filters in front of the detector when needed.

The MPO Point Spread Function (PSF) is composed by a central spot and two cross arms: about 50% of the incident X-rays are reflected twice and focused in the central PSF spot, X-rays reflected just once and focused in both PSF arms (2×22%), and the rest produces a diffuse patch. In the *Lobster Eye* geometry the vignetting is very low, reaching 10–15% at the edge of the FoV. Simulations indicate that the PSF of such system could reach 4.5 arcmin FWHM at 1.5 keV (central peak). In reality, due to defects in the alignment of the channels, distortions of the channel walls, and misalignment of the plates, such system may finally provide a ∼10 arc min FWHM PSF. The GRB afterglow position is computed on-board in near real-time by the MXT Data Processing Unit. The expected MXT effective area is about 23 cm$^2$ at 1 keV for the central spot, and expected GRB afterglow localization performance, obtained by folding the entire Swift/XRT afterglow dataset through the MXT response, shows that MXT is well adapted to study GRB afterglows, see Fig. 3. Indeed, 50% of the bursts will be localized to better than

1 arc min (90% c.l. radius; statistical uncertainties only) within 5 min from the MXT stabilization time.

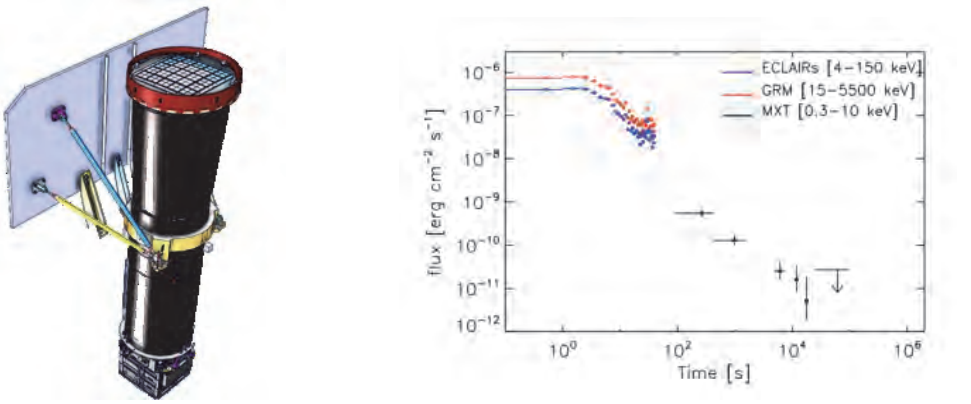



Fig. 3. Left: the MXT telescope. Right: Simulation of the prompt and afterglow emission of GRB 091020 in SVOM instruments. In MXT the afterglow is detected up to $\sim 10^4$ s. For more details, see the SVOM white paper [1].

### 2.4. *VT*

The Visible Telescope (VT) main purpose is to detect and observe the optical afterglows of GRBs localized by ECLAIRs. It is a Ritchey-Chretien telescope with a 40 cm diameter and its limiting magnitude is about $M_V$=22.5 for a 300 s integration time. The VT is designed to maximize the detection efficiency of GRB's optical afterglows. A dichroic beam splitter divides the light into two channels, in which the GRB afterglow is observed simultaneously, a blue channel with a wavelength from 0.4 to 0.65 $\mu$m and a red channel from 0.65 to 1 $\mu$m. Each channel is equipped with a 2K×2K CCD detector. The Quantum Efficiency (QE) of the red-channel CCD is over 50% at 0.9 $\mu$m, which gives VT the capability of detecting GRBs with redshifts larger than 6.5. The VT FoV is about 26×26 arc min$^2$, covering the ECLAIRs error box in most cases.

## 3. The SVOM Ground Segment

Three optical/infrared instruments compose the ground-based SVOM system. The first unit of the Ground-based Wide Angle Cameras (GWACs), currently located Xinglong (China), and is already operational. Another unit may be installed later at a different site. The GWACs located in China consist of 40 cameras of 180 mm diameter, covering $\sim$6000 deg$^2$ in total in the wavelength range 500-850 nm. The GWAC system can reach a magnitude limit of V=16 in 10 s. In addition to the GWACs, whose main goal is to try to catch the prompt optical emission of SVOM

GRBs, two robotic 1-m class telescopes are being included in the SVOM system. The Chinese Ground Follow-up Telescope (C-GFT) has a 1.2 m diameter primary mirror, is located at the Jilin observatory, has a field of view of 21×21 arc min$^2$ and is sensitive in the 400-950 nm wavelength range. The French Ground Follow-up Telescope (F-GFT/Colibri) will have a 1.3 m diameter primary mirror and will be installed in San Pedro Martir (Mexico). It offers a field of view of 26×26 arc min$^2$ and operates simultaneously in 3 bands from 400 to 1700 nm.

The SVOM ground segment includes a set of VHF antennas (about 40), which will be deployed on the projected track of the satellite in order to collect in real time the alerts produced by the platform. These VHF ground stations will be connected to the French Science Centre (FSC), which will immediately dispatch the alerts produced by all the SVOM instruments to the scientific community. We expect that 65% of the SVOM alert messages are delivered within less than 30 s, and that 95% of them reach the observers within less than 20 minutes.

## 4. Conclusions

SVOM will play a key role in the upcoming era of time domain and multi-messenger astronomy. At the time of SVOM launch, new large sky area time domain oriented instruments, like the LSST in the optical or the SKA at radio wavelengths, will deliver an unprecedented number of well localized ToOs that may be worth following up with the space-borne and ground-based SVOM instruments. Gravitational wave detectors (advanced LIGO, VIRGO, LIGO-India, KAGRA, etc.) will provide more alerts than now with improved localisation accuracies. Neutrino detectors IceCube-Gen2, KM3NeT (extension of ANTARES), etc. will also trigger a significant number of alerts. In this respect, the SVOM/MXT grasp appears to be well adapted to the large error boxes delivered by gravitational wave and neutrino detectors.

## Acknowledgments

We acknowledge the financial support of the UnivEarthS Labex program at Sorbonne Paris Cité (ANR-10-LABX-0023 and ANR-11-IDEX-0005-02).

## References

1. J. Wei et al., *Scientific prospects of the SVOM mission* (2016) arXiv:1610.06892.
2. O. Godet et al., *The x-/gamma-ray camera ECLAIRs for the gamma-ray burst mission SVOM* (2014) SPIE 9144, 24.
3. D. Götz et al., *The Microchannel X-ray Telescope Status* (2016) SPIE 9905, 4.
4. J.R.P. Angel, *Lobster eyes as X-ray telescopes* (1979) ApJ, 233, 364.
5. N. Meidinger et al., *Next generation of pnCCDs for X-ray spectroscopy and imaging* (2006) Nuclear Instruments and Methods in Physics Research A, 568, 141.

# Mercury perihelion advance in perspective (1906–1915)

Christian Bracco

*Syrte, Observatoire de Paris, Université PSL, CNRS, Sorbonne Université, LNE, Paris, 75014, France*
*E-mail: Christian.Bracco@obspm.fr*

*CRHI, Université Côte d'Azur, Nice, 06204, France*

Jean-Pierre Provost

*720 route de la Turbie*
*Eze, 06360, France*
*E-mail: provost@oca.eu*

Thanks to the introduction of a general Lagrangian, we derive a simple expression for Mercury perihelion advance (MPA) which can be applied to many historical examples, such as Poincaré's 1906 application of his relativistic action, Einstein's probable test of his 1912 metric, Einstein-Besso's 1914 application of the *Entwurf*... as well as to modern presentations. We argue that Einstein's failure to derive in his Zurich Notebook the right equations of general relativity has a physical rather than mathematical origin, which amounts to an incomplete application of the 1895 Lorentz Transformations to his 1911 formulation of the Equivalence Principle. MPA could have been already in his hands in 1912.

*Keywords*: Mercury; Lagrangian; Lorentz; Poincaré; Einstein.

## 1. Introduction: Between history and teaching

In this presentation, we discuss Mercury perihelion advance (MPA) in relation to relativity both from an historical and a pedagogical point of view [1]. Implementing efficiently historical issues in the teaching of physics (instead of referring to them either anecdotally, critically or emphatically) is in our opinion an important goal both for the history and the teaching of science. Of course, this necessitates to formulate them in a modern perspective, which implies that one acknowledges that historical issues are often close to modern ones, a point of view generally shared by mathematicians in the tradition of Henri Poincaré, but not by physicists who are rather inclined to adopt the Bachelardian epistemological point of view of scientific revolutions (or a Kuhnian one in terms of change of paradigm). In return, such an approach may shed a new light on these issues. As an example, which concerns MPA, we explained in [2] how Lorentz 1895 transformations $x' = x - Vt$, $t' = t - Vx$ (with $c = 1$) (thereafter 1895 LT) can be used to introduce Special Relativity (SR) early in teaching and discuss relativistic physics without calling for the exact 1904 LT;[a] then applying these LT to Einstein's approach of gravitation through the

[a]Other examples of the interest of historical issues for present teaching are our discussions of Kepler's equant model [3] and Newton's geometrical derivation of Kepler laws [4].

equivalence principle (EP) in 1907 and 1911, we realized that Einstein's "bumpy road" to General Relativity (GR) mainly originated from an incomplete use of them (see section 3 below).

As well documented [5], the observation of Mercury's orbital perihelion (Le Verrier 1859, Newcomb 1882 and 1895, etc.) has left a residue of about 43" per century unexplained by perturbations due to other planets within the framework of Newton's theory of gravitation, or by attempts to modify the gravitational force.[b] Less known is that the calculation of this residue, proposed as an exercise in some text books, from the Lagrangian (with $c = 1$)

$$L_{HP} = -\sqrt{1 - v^2} + \alpha u \quad (\alpha = GM, u = 1/r), \tag{1}$$

has been taken as a test of relativistic mechanics as soon as 1906 by Poincaré in his Sorbonne lectures on Mercury perihelion[c] [6]. History of the MPA has remembered that SR with the above Lagrangian only accounts for $1/6^{\text{th}}$ of the correct value, forgetting to consider

$$L_{SR} = -\sqrt{1 - v^2} + \alpha\gamma u, \quad (\gamma = (1 - v^2)^{-1/2}), \tag{2}$$

which introduces the physical idea of "inertia of energy". It has essentially memorized the final explanation of MPA by Einstein in November 1915 [8] from the metric

$$ds^2 = (1 - 2\alpha u)dt^2 - (1 + 2\alpha u)dr^2 - r^2 d\Omega^2, \tag{3}$$

(based on the solution of $R_{\mu\nu} = 0$ at first order) as one of his most famous discoveries. History generally ignores Einstein-Besso calculation[d] of MPA from

$$ds^2 = \left(1 - 2\alpha u + \frac{3}{2}(\alpha u)^2\right) dt^2 - d\vec{r}^2, \tag{4}$$

based on Einstein-Grossman's *Entwurf* [10] and the hypothesis that gravitational energy (the quadratic term in Eq. (4)) be also source of gravitation, or Einstein 1912 [11] likely one from

$$ds^2 = (1 - 2\alpha u)dt^2 - d\vec{r}^2, \tag{5}$$

after his incomplete application of the EP (see section 3), or Einstein-Fokker 1914 possible one from

$$ds^2 = e^{-2\alpha u}(dt^2 - d\vec{r}^2) \tag{6}$$

[b] We recall that in Newton's theory, an inverse square law for the acceleration $-\alpha\hat{r}/r^2$ accounts for the elliptical trajectory $r = (l^2/\alpha)(1 + e\cos K\varphi)^{-1}$ ($l$ area constant or angular momentum per unit mass), which for small eccentricity $e \ll 1$ corresponds to Kepler's equant model.

[c] We recall that in 1905 Poincaré deduced the relativistic free Lagrangian (he took already $c = 1$!) from the Lorentz contraction and justified it by the invariance of action under the Lorentz group [7]. In 1906, he compared his result for MPA with those derived from Abraham's and Langevin's Lagrangians (based on different hypotheses concerning the shape of a moving electron).

[d] A noticeable exception is [9].

(equivalent to Nordström scalar theory).[e] Historical discussions also generally put aside the question of why Einstein 1915 metric is equivalent for MPA to Schwarzschild's exact solution

$$ds^2 = (1-2\alpha u)dt^2 - (1-2\alpha u)^{-1}dr^2 - r^2 d\Omega^2 \tag{7}$$

or why it is not equivalent to

$$ds^2 = (1-\alpha u)^2 dt^2 - (1+\alpha u)^2 dr^2 - r^2 d\Omega^2 \tag{8}$$

although like Eq. (3) it also fully takes into account the EP. In section 2, we present a simple formalism which allows to address all these proposals and to test the different physical ideas which have anticipated GR.[f]

## 2. A general Lagrangian approach to Mercury perihelion advance

Quite remarkably, the eight above metrics, with the identification $ds^2 = L^2 dt^2$ and the Lagrangians given by Eqs. (1) and (2), correspond to a squared Lagrangian of the type (for plane trajectories)

$$L^2 = A(\alpha u) - B(\alpha u)\dot{r}^2 - C(\alpha u) r^2 \dot{\varphi}^2, \tag{9}$$

$A$, $B$ and$C$ being functions of $\alpha u$, with $A(0) = B(0) = C(0) = 1$ and the dot meaning a derivation with respect to time. These functions are e.g. respectively $(1-\alpha u)^2$, $(1-\alpha u)$ and $(1-\alpha u)$ for Eq. (1) and $(1-2\alpha u)$, $(1+2\alpha u)$ and 1 for Eq. (3). Finding the equation of the trajectory from $L$, with the use of angular momentum $l$ and energy $E$ conservations, is then a standard exercise of Lagrangian calculus (see [1] for details). It leads to (prime denoting a derivation with respect to $\varphi$):

$$A = E^2 - l^2(au^2 + bu'^2), \quad \left(a = \frac{A}{C},\ b = \frac{AB}{C^2}\right). \tag{10}$$

Because of Mercury's small orbital eccentricity, one can set $u = u_0 + \delta u$, where $\delta u$ is a small deviation with respect to a circular trajectory centered at origin. Then introducing the a-dimensional distance parameter $\varepsilon = \alpha u_0 \simeq \alpha^2/l^2 \ll 1$ (*cf.* note 2) and writing $a(\varepsilon) = 1 + \varepsilon a'(0)$ (idem for $b(\varepsilon)$ and $A'(\varepsilon)$), the perturbation calculus at order $\varepsilon$ yields:

$$\delta u'' + K^2 \delta u = 0; \quad K^2 = 1 + \varepsilon\left(3a'(0) - b'(0) + \frac{A''(0)}{2}\right). \tag{11}$$

[e]The calculations from Eqs. (5) and (6) are merely historical assumptions based on Einstein's acknowledgement that he thought about MPA since 1907. Einstein has also probably tested the relativistic equations of motion in a gravitational field which he puts forward at the top end of his extensive review paper of November 1914.

[f]In present teaching [12], this formalism can also be applied to the PPN metric $ds^2 = (1-2\alpha u + 2\beta\alpha^2 u^2)dt^2 - (1+2\gamma\alpha u)d\vec{r}^{\,2}$ or to the isotropic form $ds^2 = (1-\alpha u)^2(1+\alpha u)^{-2}dt^2 - (1+\frac{\alpha}{2}u)^4 d\vec{r}^{\,2}$ of the Schwarzschild metric.

So the trajectory reads $u = u_0(1 + e\cos K\varphi)$, and the perihelion advance per revolution is:

$$\Delta\varphi = \pi(1 - K^2) = \pi\varepsilon\left(-3a'(0) + b'(0) - \frac{A''(0)}{2}\right). \tag{12}$$

In particular, applications to Eqs. (1) ($\Delta\varphi = \pi\varepsilon$), (2) ($\Delta\varphi = 3\pi\varepsilon$), (4) ($\Delta\varphi = 4\pi\varepsilon$) and (8) ($\Delta\varphi = 5\pi\varepsilon$) allow to appreciate how much[g] the associated physical hypotheses mentionned in section 1 bring the solution close to the exact one obtained from Eq. (7) as well as from Eq. (3) ($\Delta\varphi = 6\pi\varepsilon$). One may also note that a quadratic term in A (*cf.* Eqs. (1), (2), (4), (6) and (8)) interpreted in section 1 as "gravitation due to gravitational energy" systematically contributes negatively to $\Delta\varphi$[8].[h] Finally, we emphasize that the EP often discussed in the context of GR is clearly insufficient since it gives information on $A$, $B$ and $C$ only up to first order in $\alpha u$ and says nothing for $A''(0)$. In conclusion, whereas the EP can account for the gravitational redshift and the deviation of light, one cannot escape $R_{\mu\nu} = 0$ in order to explain MPA; so MPA is a nice introduction to the teaching of GR.[i]

## 3. How far was Einstein from explaining Mercury perihelion advance in 1912?

Following Einstein, it has been claimed that him and Grossmann (in the 1912 Zurich Notebook [13]) where close to the equations of GR (at least in vacuum) since in particular the Ricci tensor $R_{\mu\nu}$ was clearly for them the geometrical counterpart of the energy momentum tensor. The problem was that Einstein 1912 metric deduced from his application of the EP did not satisfy $R_{\mu\nu} = 0$ even at first order in the gravitational potential. The metric was right for time coordinates but false for spatial ones. Up to our knowledge, the question has not been previously raised as to whether 1895 LT $x' = x - Vt$, $t' = t - Vx\,(c = 1)$, which come into play when one conceives, as Einstein, an accelerated system $\Sigma$ moving along the $x$-axis as a succession of inertial frames $R_t = R$, $R_{t+\delta t} = R'$ with relative velocities $V = a\delta t$, imply or not a non trivial $x$-metric for $\Sigma$. We now briefly show that the answer is affirmative.[j]

Let us indeed consider a rod of apparent length $l$ in uniform motion at velocity $v$ in $R_t$, defined by the events $0 < x - vt < l$. It is straightforward to verify that it is defined in $R_{t+\delta t}$ by $0 < x' - v't' < l'$ with in particular $l' = l(1 + Vv)$.[k] This calculation is the same as that of Poincaré in 1900 for a light wave packet ($v = 1$),

[g] $\Delta\varphi = 2\pi\varepsilon$ can be obtained from the non-Lagrangian equation $d\vec{p}/dt = -\alpha\gamma\hat{r}/r^2$ [1].
[h] $A''(0)$ is the only contribution for the scalar theory (Eq. (6)).
[i] Instructive for teaching are also the PPN metric result $\Delta\varphi = \pi\varepsilon\,(4 - 2\beta + 2\gamma)$, and the result issued from the isotropic metric $\Delta\varphi = \pi\varepsilon\,(12 - 4 - 2)$ which shows that $6\pi\varepsilon$ may be obtained in very different ways according to the choice of coordinates.
[j] See [1–2] for more issues and details.
[k] The same analysis for a moving clock (*i.e.* writing $\Delta x = v\Delta t$ with prime coordinates) gives the inverse relation $\Delta t' = \Delta t\,(1 - vV)$.

and the fact that the result is identical to the Doppler effect $\lambda' = \lambda(1+V)$ is simply a consequence of the invariance of the phase already considered by Lorentz in 1895. So, a light wave packet is nothing else but a rod moving at $v = 1$! In 1911, Einstein applied, as is well known, the EP to a light signal sent from the altitude $h$ to the ground and showed, with the help of the Doppler effect, that the received signal has a frequency $\nu_0 = \nu_h(1+gh)$ [14]. He then deduced that the proper times of identical clocks at rest at $z = 0$ and $z = h$ were related by $\tau_h = \tau_0(1+gh)$ and he extended this relation to $\tau = t(1 - GM/r)$ for clocks in the gravitational potential of a spherical mass $M$, thanks to the EP. Had he realized that $\lambda_0 = \lambda_h(1-gh)$ implied for the proper lengths the inverse relation $\Delta l_h = \Delta l_0(1+gh)$, he would probably have been led to the metric given by Eq. (3) instead of Eq. (4). Then he would have explained the MPA and be confident in Grossmann's hypothesis $R_{\mu\nu} = 0$ outside matter as soon as 1912.

In, conclusion, whereas Einstein has found in 1907 and in 1911 subtle ways to apply the Equivalence Principle and relativity at first order to the issue of time, he completely missed the opportunity to address an analogous issue for lengths. He turned to other considerations which have contributed to his "bumpy road" to General Relativity.

## References


1. J.-P. Provost and C. Bracco, Lorentz' 1895 transformations, Einstein equivalence principle and the perihelion shift of Mercury, *Eur. J. Phys.***37**/6, (2018).
2. J.-P. Provost and C. Bracco 2016, The 1895 Lorentz transformations: historical issues and present teaching, *Eur. J. Phys.* **37**/4 (*Highlights*), (2016).
3. C. Bracco and J.-P. Provost, Had Mars not existed: Kepler's equant model and its physical consequences, *Eur. J. Phys.* **30,** 1085 (2009).
4. J.-P. Provost and C. Bracco, A simple derivation of Kepler's law without solving differential equations, *Eur. J. Phys.* **30,** 581 (2009).
5. N. T. Roseveare, *Mercury's perihelion from Le Verrier to Einstein* (Oxford, Clarendon Press, 1982).
6. H. Poincaré, Les limites de la loi de Newton, H. Vergne and M. Chopinet (eds.) *Bull. Astron.* **17,** 121 (1953).
7. C. Bracco and J.-P. Provost, De l'électromagnétisme à la mécanique: le rôle de l'action dans le Mémoire de Poincaré de 1905, *Rev. Hist. Sci.* **61,** 457 (2009). `http://www.cairn-int.info/article.php?ID_ARTICLE=E_RHS_622_0457`.
8. A. Einstein, Explanation of the Perihelion Motion of Mercury from the General Theory of Relativity, (1915) in *Collected Papers of Albert Einstein* (*CPAE*) *Vol.* 6 *doc. 24* (Princeton: Princeton University Press). `http://einsteinpapers.press.princeton.edu/`.
9. M. Janssen, What did Einstein know and when did He know it? A Besso Memo dated August 1913 in *The Genesis of General Relativity* (Boston Studies in the Philosophy of Science, vol. 1, 2007).

10. A. Einstein and M. Grossmann, Outline of a generalized theory of relativity and of a theory of gravitation, (1913) in *CPAE Vol. 4 doc. 13*.
11. A. Einstein, The speed of light and the statics of the gravitational field, (1912) in *CPAE Vol. 4 doc. 3*.
12. H. C. Ohanian and R. Ruffini, *Gravitation and Space Time* (Cambridge, Cambridge University Press, 2013).
13. M. Janssen, J. D. Norton, J. Renn, T. Sauer and J. Stachel, Einstein's Zürich notebook, introduction and sources in *The Genesis of General Relativity*.
14. A. Einstein, On the influence of gravitation on the propagation of light, (1911) in *CPAE Vol. 3 doc. 23*.

# May 29, 1919 total solar eclipse: Brazilian, British and North-American expeditions to Sobral, Brazil

Luís C. B. Crispino*

*Faculdade de Física, Universidade Federal do Pará, 66075-110, Belém, Pará, Brazil*
**E-mail: crispino@ufpa.br*
*www.ufpa.br*

I comment on the three (Brazilian, British and North-American) expeditions organized to observe in Sobral, Brazil, the May 29, 1919 total solar eclipse.

*Keywords*: Total Solar Eclipse; 1919; Expeditions; General Relativity.

Among the most famous eclipses in History of Science, stands out the May 29, 1919 total solar eclipse, during which observations led to the first experimental confirmation of General Relativity.[1]

The photographs which confirmed Einstein's Theory, in 1919, were taken by two expeditions organized by the Royal Greenwich Observatory: one sent to Príncipe Island, in Africa (team composed by Arthur Stanley Eddington and Edwin Turner Cottingham); and the other sent to the city of Sobral, in Brazil (team composed by Andrew Claude de la Cherois Crommelin and Charles Rundle Davidson).[2]

Less known is the fact that, apart from the Cambridge expedition, two other scientific commissions (cf. Fig. 1) were sent to Sobral in 1919.[3] One was the Carnegie commission, composed by Daniel Maynard Wise and Andrew Thomson;[4] and the other was the Brazilian National Observatory commission, leaded by Charles Henry Morize.[5–7] The Carnegie expedition had the purpose to investigate Earth's magnetism and atmospheric electricity, while the Brazilian Observatory team main goal was to investigate the solar corona.

The British and the North-American teams installed their equipment at the racecourse of the Jockey Club at Sobral (cf. Fig. 2), while the Brazilian team set their installation in Patrocinio square (cf. Fig. 3).

Both the British and the Brazilian teams obtained very nice pictures of the eclipsed Sun (cf. Figs. 4 and 5). In those pictures, an impressive solar prominence appears. The measurements of the relative positions of the stars appearing in the British photographic glass plates led to the confirmation of Einstein's theory.

The results obtained by the British expeditions sent to Africa and Brazil were published in Ref. 1. The results of the North-American expedition were published in Refs. 8, 9, 10, 11; and the results of the Brazilian expedition were published in Ref. 12.

Fig. 1. Observers of the May 29, 1919 Total Solar Eclipse, in Sobral, Brazil. From left to right: Luiz Rodrigues, Theophilo Henry Lee, Daniel Maynard Wise, Henrique Charles Morize, Charles Rundle Davidson, Andrew Claude de la Cherois Crommelin, Allyrio Hugueney de Mattos, Andrew Thomson, Domingos Fernandes da Costa, Lelio Itapuambyra Gama, Antonio C. Lima, and Primo Flores. Courtesy of the National Observatory, Rio de Janeiro, Brazil.


Fig. 2. Racecourse of the Jockey Club at Sobral, where the British and the North-American teams installed their equipment. Courtesy of Carnegie Institution, Department of Terrestrial Magnetism, Washington, USA.

Fig. 3. Patrocínio square in Sobral, where the Brazilian team installed their equipment. Courtesy of the National Observatory, Rio de Janeiro, Brazil.

**Acknowledgments**

I thank Marcelo Costa de Lima for helpful discussions and Luiz Carlos dos Santos Leite for visiting the Cambridge University Library in the search for documents related to the 1919 British expedition to Sobral. I am also very grateful to: (i) Shaun J. Hardy and Mary Ferranti, from Carnegie Institution, Department of Terrestrial Magnetism, Washington, USA; (ii) Adam Perkins and Emma Saunders, from Cambridge University Library, United Kingdom; (iii) Emily Beech and Louise Devoy, from National Maritime Museum, Greenwich, London, United Kingdom; (iv) Diógenes Farias Gomes, from *Vale do Acaraú* University, Sobral, Brazil; (v) Emerson Ferreira de Almeida, from *Museu do Eclipse*, Sobral, Brazil; (vi) Everaldo Pereira Frade, Luci Meri Guimarães, and Maria Celina Soares de Mello e Silva, from *Museu de Astronomia e Ciências Afins* (MAST), Rio de Janeiro, Brazil; (vii) Carlos Henrique Veiga, Jailson Souza de Alcaniz, João Carlos Costa

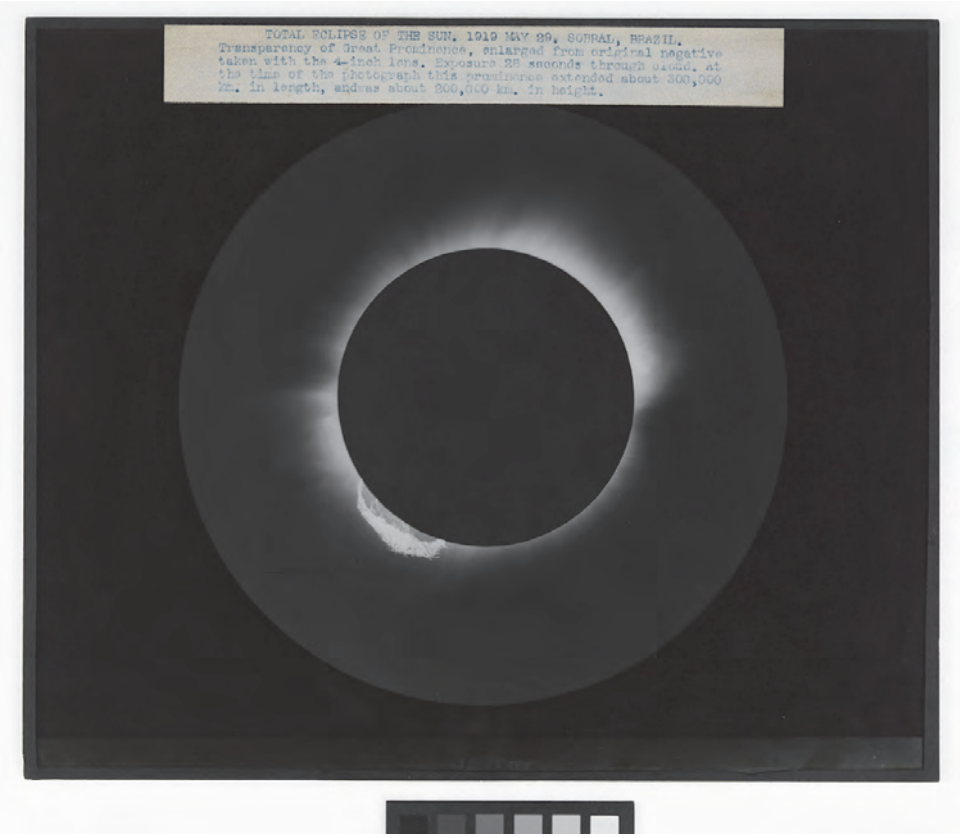



Fig. 4. Image of the eclipsed Sun taken by the British expedition at Sobral, on May 29, 1919. Courtesy of National Maritime Museum, Greenwich, London, United Kingdom.

dos Anjos, Katia Teixeira dos Santos, and Matheus Santos de Oliveira, from *Observatório Nacional* (ON), Rio de Janeiro, Brazil; and (viii) *Fundação Biblioteca Nacional*, Rio de Janeiro, Brazil. I thank also to Christian Gérard Wise for email correspondence and for sharing some pictures and documents of his grandfather, Daniel Maynard Wise. I acknowledge the partial financial support from *Conselho Nacional de Desenvolvimento Científico e Tecnológico* (CNPq) and *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* (CAPES) – Finance Code 001, from Brazil.


Fig. 5. Image of the eclipsed Sun taken by the Brazilian expedition at Sobral, on May 29, 1919. Courtesy of the National Observatory, Rio de Janeiro, Brazil.

## References


1. F. Dyson, A. Eddington and C. Davidson, *A determination of the deflection of light by the Sun's gravitational field, from observations made at the Total Eclipse of May 29, 1919*, Philosophical Transactions of the Royal Society of London, **220** (1920), pp. 291-333.
2. A. C. D. Crommelin, *The Eclipse Expedition to Sobral*, The Observatory, London, **42** (1919), pp. 368-371.
3. L. C. B. Crispino, *Expeditions for the observation in Sobral, Brazil, of the May 29, 1919 total solar eclipse*, International Journal of Modern Physics D, **27** (2018), 1843004.
4. L. C. B. Crispino and M. C. de Lima, *Expedição norte-americana e iconografia inédita de Sobral em 1919* (in Portuguese), *Revista Brasileira de Ensino de Física*, **40** (2018), e1601.
5. L. C. B. Crispino and M. C. de Lima, *Amazonia introduced to General Relativity: The May 29, 1919, Solar Eclipse from a North-Brazilian Point of View*, Physics in Perspective, **18** (2016), pp. 379-394.
6. M. C. de Lima and L. C. B. Crispino, *Crommelin's and Davidson's visit to Amazonia and the 1919 total solar eclipse*, International Journal of Modern Physics D, **25** (2016), 1641002.
7. L. C. B. Crispino and M. C. de Lima, *A teoria da relatividade de Einstein apresentada para a Amazônia* (in Portuguese), *Revista Brasileira de Ensino de Física*, **38** (2016), e4203.
8. L. A. Bauer *et al.* (editors), *Terrestrial Magnetism and Atmospheric Electricity*, volume XXV. The Johns Hopkins Press, Baltimore, Maryland, 1920.
9. S. J. Mauchly and A. Thomson, *Results of Atmospheric-Electric Observations made at Sobral, Brazil, during the Total Solar Eclipse of May 29, 1919*, Terrestrial Magnetism and Atmospheric Electricity, **25** (1920), pp. 41-48.
10. L. A. Bauer, J. A. Fleming, H. W. Fisk and W. J. Peters, *Land Magnetic Observations 1914-1920. Researches of the Department of Terrestrial Magnetism*, volume IV. Carnegie Institution of Washington, Washigton D.C., 1921.
11. L. A. Bauer, *Results and Analysis of Magnetic Observations during the Solar Eclipse of May 29, 1919*, Terrestrial Magnetism and Atmospheric Electricity, **25** (1920), pp. 81-98.
12. H. Morize, *Resultados obtidos pela Commissão Brasileira do eclipse de 29 de Maio de 1919* (in Portuguese), Revista de Sciencias, **4** (1920), pp. 65-81.

# The large numbers hypothesis in cosmology

Eve-Aline Dubois* and André Füzfa and Dominique Lambert

*Namur Institute for Complex Systems and Espace philosophique de Namur*
*University of Namur - Belgium*
**E-mail: eve-aline.dubois@unamur.be*

In 1937, P.A.M. Dirac suggested the idea that the dimensionless constants of physics must be in relation with the epoch (age of the universe expressed in atomic units). From this hypothesis, known as Large Numbers Hypothesis or Dirac's Principle, he built a cosmological model in 1938 and abandoned it.

Following this principle, P. Jordan developed a series of articles, translated by us, based on the conservation of the dimensionless numbers coincidence. He suggested a model of matter creation to counterbalance the expansion of the universe.

Surprisingly, in the seventies, Dirac came back to his Large Numbers Hypothesis and published a new cosmological model, based on a description of the universe using two metrics.

We intend to review and present the historical development of the Large Numbers Hypothesis and its consequences in cosmology through the works of these two famous authors.

*Keywords*: Cosmology - Large number hypothesis - History of science.

## 1. Dirac's Principle

In a short letter to the editor in *Nature*[1], P.A.M. Dirac, following Eddington's work on dimensionless numbers[2], noticed a coincidence between the constants used in cosmology and enunciated his Large Number Hypothesis. Indeed, it could be observed that the ratio between the Coulombian and the Newtonian gravitational forces between an electron and a proton is about $10^{39}$; the ratio between the masses of the universe and of a proton is about $10^{78}$. These two large numbers need different types of explanations because they are not physically linked. But, if you add the coincidence that the age of the universe, according to the contemporary cosmological models, expressed in atomic units, so-called the epoch, is $10^{39}$; it seems logical to put the two previous large numbers in relation with the epoch. It is what Dirac did: "*This suggests that the above-mentioned large numbers are to be regarded not as constants, but as simple functions of our present epoch, expressed in atomic units.*" [1]

This principle has two direct consequences. First, the number of protons and electrons has to increase like the square of the epoch, conserving the null electrical charge of the universe. Secondly, the gravitational constant can not be constant anymore and must decrease with time.

Dirac concluded his letter by a brief paragraph about cosmological applications of his principle, which he studied in a later article, as it will be shown in the next section.

## 2. Dirac's Cosmology of 1938

In 1938, Dirac published a paper in which he suggested a cosmological model based on the Large Numbers Hypothesis[3]. He rewrote his principle as "*Any two of the very large dimensionless numbers occurring in Nature are connected by a simple mathematical relation, in which the coefficients are of the order of magnitude unity.*" [3]

With this hypothesis, he tackled one of the main problems of cosmology, the determination of the form of $f(t)$, similar to the current scale factor, giving the recession law of galaxies, since any cosmological model must explain Hubble's observations. Doing so, he arrived at the possibility of creation or annihilation of protons and neutrons assuming that the effect will be so faint that it could not be detected in laboratory. However, Dirac noted that "*However, such a spontaneous creation or annihilation of matter is so difficult to fit in with our present theoretical ideas in physics as not to be worth considering, unless a definite need for it should appear, which has not happened so far, since we can build up a quite consistent theory of cosmology without it.*" [3]

Dirac also studied the curvature of the slice of three-dimensional surfaces given for each value of the epoch, or $t$-space. The curvature cannot be positive, because, in this case, the mass of the universe is a very large number and will be constant, thanks to the assumption of mass conservation. This is in contradiction with his fundamental principle so it should be ruled out. The case of a negative curvature can also be excluded: working in a sphere of radius equal to the radius of curvature of the $t$-space, the mass contained in this sphere will not evolve with time which contradicts Dirac's principle. Dirac concluded that "*We are thus left with the case of zero-curvature, or flat t-space, as the only one consistent with our fundamental principle and with conservation of mass.*" [3]

The article finished with this summary: "*It is proposed that all the very large dimensionless numbers which can be constructed from the important natural constants of cosmology and atomic theory are connected by simple mathematical relations involving coefficients of the order of magnitude unity. The main consequences of this assumption are investigated and it is found that a satisfactory theory of cosmology can be built up from it.*" [3]

## 3. Jordan's work

From 1937, Pascual Jordan developed a parallel work based on Eddington's study of dimensionless numbers[2] and Dirac's idea that very large numbers could be expressed in relation with the epoch. Jordan's work has been published in a series of articles[4, 5, 6]. We worked on our own translation of them.

Like Eddington, Jordan hoped to find the way to unify quantum mechanics and genral relativity by finding the relation between their two characteristic constants $\hbar$ and $c$. Following Dirac's reasoning, Jordan reached the conclusion that the gravitational constant cannot be constant with respect to the time and that matter must be created.

To have a continuous and spontaneous matter creation process, Jordan considered the possible creation of stars. These stars must have the good radius and mass ratio to counterbalance their mass energy with their own gravitational energy. So that, according to him, the energy cost of this creation is null. Jordan found an argument in favour of his theory of star creation in the observation of younger and older stars.

## 4. Jordan and Hoyle

The history has very often ignored Jordan's German pre-World War II model. In 1948, two articles, due to Hermann Bondi and Thomas Gold for the first[7] and to Fred Hoyle for the second[8], founded the Steady State Theory. Max Born seemed to see some similarities between Jordan's work and Hoyle's model[8]. Therefore, he invited Pascual Jordan to publish in English in the prestigious review *Nature*[9].

However, Hoyle's and Jordan's models are really different. If both of them referred to Dirac's work, they did not develop it in the same way. Jordan worked with the dimensionless constants and their variations when Hoyle modified Einstein's equations to describe a universe with a constant density of matter. And, to create matter, the former considered spontaneous appearance of stars while the latter suggested creation of hydrogen atoms.

That is why Jordan finished his comparison between their models with: "*Several decisive ideas of Hoyle's are in full harmony with my own theory [...] But there are also considerable differences between Hoyle's theory and my own.*"[9]

## 5. Dirac's Cosmology of 1973

Surprisingly, Dirac used a communication at the Pontifical Academy of Science on evolutionary cosmology[10] to come back to his cosmological model with a matter creation process. He published two other articles[11][12] on this subject.

In this series of papers, Dirac studied two ways to create matter: "*A: Matter is created uniformly throughout space, and hence mainly in intergalactic space. B: Matter is created where it already exists, in proportion to the amount existing.*"[10] Thereafter, he called them *additive and multiplicative creation*[12].

According to his Large Numbers Hypothesis, the gravitational constant must vary. To reconcile this idea with the successful Einstein's theory of gravitation, Dirac suggested the use of two metrics: Einstein's one $ds_E$ and $ds_A$, measured by atomic apparatus. From that, he built two cosmological models waiting observations to come to make the distinction between the two, as Shapiro's time delay experiment.

In the conclusion, Dirac wrote: "*The foregoing work is all founded on the Large Numbers Hypothesis, in which I have great confidence.*"[12]

## Conclusion

The present paper described the historical development of cosmological models based on Large Numbers Hypothesis and reviewed the work of two renowned physicists who built cosmological models on this hypothesis. If this hypothesis is now considered as mere numerology and close to pseudo-science, it is interesting to study its past applications in physics. For a review on controversies about the Big Bang theory and the Steady State theory, we refer, among others, to Helge Kragh's work[13]. Our work gives us the opportunity to illustrate the fact that the Steady State theory was not the only one in competition with the Big Bang Theory and, moreover, not the only one to suggest a process of continuous creation of matter.

## Acknowledgement

The authors would like to thank D. Bertrand for his precious help in the translation of Jordan's German works.

## References

1. Paul A.M. Dirac. The cosmological constants. *Nature*, 139:323, February 1937.
2. Arthur Eddington. *Relativity theory of protons and electrons*. Cambridge University Press, 1936.
3. Paul A.M. Dirac. A new basis for cosmology. *Proceedings of the Royal Astronomical Society of London*, 165(921):199–208, 1938.
4. Pascual Jordan. Die physikalischen Welkonstanten. *Die Naturwissenschaften*, (32):513–517, 1937.
5. Pascual Jordan. Zur empirischen Kosmologie. *Die Naturwissenschaften*, 26:417–421, 1938.
6. Pascual Jordan. Bemerkungen zur Kosmologie. *Annalen der Physik*, 5-36:64–70, 1939.
7. Herman Bondi and Thomas Gold. The steady-state theory of the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 108:252–270, 1948.
8. Fred Hoyle. A new model for the expanding universe. *Monthly Notices of the Royal Astronomical Society*, 108:372–382, 1948.
9. Pascual Jordan. Formation of the stars and development of the universe. *Nature*, 164:637–640, october 1949.
10. Paul A.M. Dirac. Evolutionary cosmology. *Commentarii Pontificia Academia Scientiarum*, 46-II:1–16, 1973.
11. Paul A.M. Dirac. Long range forces and broken symmetries. *Proceeding of the Royal Society*, 333:403–418, 1973.
12. Paul A.M. Dirac. Cosmological models and the Large Numbers hypothesis. *Proceeding of the Royal Society*, 338:439–446, 1974.
13. Helge Kragh. *Cosmology and controversy*. Princeton University Press, 1999. first edition in 1996.

# Mapping gravity with hierarchical networks

M. Espinosa-Aldama*

*Universidad Autonoma Metropolitana, Unidad Cuajimalpa, Vasco de Quiroga 4871, Contadero, 05370, Ciudad de Mexico, Mexico*

In order to map and clarify the field of gravitation and dozens of alternative theories to General Relativity developed during the past century, I propose several visualizations of networks that illustrate the complexity of theoretical relations, hierarchy of concepts, historical evolution and authorship relations. Conceptual lattices are data driven, and formal concepts are analyzed using *Concept Explorer* and visualized in *D3*.

The interactive web display allows exploration of the network elements that contain information about the theories and their attributes. It also highlights super-concepts and sub-concepts, serving as a tool for the study of physical theories, their models and foundations. A tree type visualization, called the *Gravity Apple Tree*, relates common terms, authors and publications in time series for 12 branches of gravitation. Portraits of authors and links to more than 100 seminal articles are provided. Crucial experiments and observations, located in the time-line space, contrast with the theoretical branches. A citation network based on the seminal articles shows another view to the development of this theoretical field. Visualizations are available for exploration at (http://remo.cua.uam.mx/vis/Exploratorium/).

*Keywords*: Networks, Visualizations, Alternative Theories of Gravity, Concept Lattices, Citation Network, Foundations of Space-Time Theories.

## 1. Introduction

Every now and then a new theory of gravity is proposed. Keeping track of hundreds of proposals and communicating a synthesis to the general public very often results in long dissertations that loose the objective of giving a wide view. One can easily get lost in such attempts. An elegant way of presenting a thorough panorama of the state of the art has been the graphic approach. Maps of knowledge place in order references and milestones that allow us to navigate in a sea of data, such as publication dates, authors and significant terms. Visualizations may emphasize some structures of science, such as networks, clusters and tendencies. Multiple examples can be found in the *Atlas of Science*.[1] The employed methodology involves multidisciplinary skills, such as data mining, data analysis, programming, design, visualization, notions of structuralism in philosophy of science, network theory, lattice theory and some understanding of the discussed subject, in this case, physics and mathematics. Such skills and approaches are usually not found together in the scope of philosophers of science, nor physicists, even communicators. I hope this interdisciplinary approach, inserted in what is called *Science of science*, may contribute with a refreshing perspective to the field of gravitation and serve as a tool for education and philosophical discussion. Here I present three kinds of hierarchical networks for gravitation theories: the *Gravity Apple Tree*, that illustrates figuratively the development in time of 12 theoretical branches, anchored by publication dates of seminal articles from 1905 to 2016; concept lattices that illustrate the

complex structure of theories in a synchronic way, using a formal concept analysis (FCA) of renowned classifications of theories of gravity; and a citation network from 200 seminal articles arranged within 18 clusters. These visualizations have interactive properties that can be explored in a web, coded with *Data Driven Documents* (*D3*), *java script*, *css* and *html*.[a]

## 2. The Gravity Apple Tree

*The Gravity Apple Tree* was born with the intent of giving some order to an overwhelming number of alternative proposals in a synthetic, graphic manner.[2] The tree type network connects authors and publication dates of seminal articles anchored by the year of publication (green colored), classifying theories (in red) in 12 branches. A curved time-line serves as the space where crucial experiments and observations feature. The lower part of the trunk symbolizes some 200 years of newtonian gravitation. In the center of the branches, main principles and applications drive their growth.

Metric theories, that is, theories where bodies follow geodesics of space-time in free fall, as defined in Ref. 3, occupy the lower right side of the tree. The upper right side is occupied by MoNDian theories and f(R) theories, following the classification in Ref. 4. The left side is occupied by non-metric theories that suppose torsion, conformal geometries, physical fields, extra dimensions and other exotic proposals (see Refs. 5–9 for more on the history of alternative theories of gravity). Even though it represents a vast plurality of proposals, it does not include many others such as Numerical Gravity, Weyl-Dirac's theory, Plevansky's Complex Gravity, Franklin's Relativist Rotation or the Exponential Gravity of Milne.

The *Gravity Apple Tree* helps us pinpoint interesting events, such as when theoretical branches grow or stop growing in correlation with the experimental developments. It also orders the data in a mnemonic manner, which enables to go further into the mapping enterprise. Its 200 selected seminal articles compose a data base of which we can extract their cites and references (see Section 4). It is accessible to the public in a PREZI format with links to *doi* addresses of seminal articles and includes portraits of main authors and a programmed tour through the tree.[b]

## 3. Concept Lattices of Gravitation Theories

To represent the conceptual structure of gravitation theories we'll turn to the set-semantic model based view of philosophy of science, and focus on the set of *attributes* (axioms) of theories. Here, authors, dates and famous theorems do not feature, but the foundations from which we can deduce a theory. In *Foundations of Space-time Theories*,[11] Michael Friedman represents a theory $T$ by a model set comprised by geometric objects $GO$ in a manifold $M$ of 4 dimensions and a set of relations $R$. The

[a] http://remo.cua.uam.mx/vis/Exploratorium/.
[b] https://prezi.com/rdkivznlhgga/the-gravity-apple-tree/.

structuralist view of Bunge et al.,[12] is one of the most detailed approaches among philosophers of science. It takes $T$ (or more precisely the theoretical element $ET$) to be composed by a formal nucleus $K$, a set of Intentional Aplications $IA$, a Scientific Community $CC$, and of historical data $h$. $ET =< K, IA, CC, h >$ Here, we are focusing only on $K$, which is formed by classes of models $M$, $Mp$, $Mpp$ and Links between models $L$ such as the Lorentz transformations. $K = \langle D, Tip, Cond, R, E, L\rangle$ Classes of models are formed by combinations of domains $D_i$, with certain tipifications $Tip_j$, under certain conditions $Cond_n$, that follow certain relations $R_m$ and where we can distinguish certain distinguished elements $E$.

*Previous partial models*: $M_{pp} = \langle D_i, Tip_j, Cond_n\rangle$ refere to the *non-theoretical* terms, such as *Protophysics*,[13,14] where pure geometrical space-time theories lie, with respect to gravitation theories.

*Partial models*: $M_p = \langle M_{pp}, D_{i+l}, Cond_q\rangle$ include the elements of $M_{pp}$ plus some new domains expressed under *T-theroetical* terms and conditions.

*Actual models*: $M = \langle M_{pp}, M_p, R_m, E_q\rangle$ includes all of the above, plus some relations (field equations and equations of motion), as well as the distinguished elements such as the gravitational constant or the speed of light.

In this study, $ET$ will only contain the formal nucleus $K$, with out the links between models: $ET = \langle D, Tip, Cond, R, E\rangle$. As several theories are included, partial and previous models are expected to appear in the lattices.

### 3.1. *Methodology: Formal Concept Analysis*

After the selection of different classifications or theoretical reconstructions, a binary data set called *context* $(G, M, I)$ relates a set of objects $G$ (theories or models) with a set of attributes $M$ (axioms) of all objects. Using Formal Concept Analysis (FCA) — a theory and analysis methodology developed by Rudolf Willy[10] in the 1970s, based on lattice and set theory — we can extract a set of concepts and order them in a hierarchical fashion through a coverture relation, composing a concept lattice.[15,16] That is, a network where nodes represent concepts of a context, and any two nodes must have an infimum and a supremum node. These networks are read in a vertical way. A formal concept is a pair of sets *(B,A)*, where $B \subseteq G$ and $A \subseteq M$, such that $B'' = B$ and $A = A''$ where $\ll''\gg$ is a closure operator. An extra title row corresponds to the classes of attributes, which enables extra encoding. It also occupies multiple algorithms to draw and analyze lattices. The free software *Concept Explorer* (conexp 1-3), written in *java* may do the work, providing the corresponding lattices, but it requires to enter the context by hand and does not provide good importing or exporting options, such as *json* files (Java Script Object Notation). In order to solve this rough patch, and to include in the analysis other categories such as classes of attributes, the FCA is done through a code in *Wolfram Mathematica* that imports the *csv* file, uses *conexp* as a library from which we extract the necessary algorithms that provide information corresponding to the

context, nodes, attributes, objects, tags and node levels among others.[c] The *json* file is then occupied in a web that uses *D3*[d] to visualize the lattice. Visualization in *D3* allows interactivity in the web, as nodes can be moved by the user, double clicks on nodes highlights upper and lower connected nodes, and single clicks shows additional information related to attributes and theories. For this, extra *csv* files are needed, as well as the *MathJax* library.[e]

Finite ordered sets as the ones used here have a supreme and a minimum element for they are called lattices. The supreme node, generally called *Meet*, represents the concept that is possessed by all of the theories and should be presented at the top of the diagram. In our case, all theories are based on a four dimension manifold. The minimum node, also called *Join*, or *null node* that posses all of the attributes, but generally does not belong to any object is located at the bottom of the diagram. Diagrams should be as symmetric as possible, keeping nodes in their hierarchical level. Lattices are much more informative when they are presented in such a way and for small contexts. For big contexts, diagrams may appear messy or confusing for it is convenient to present them with interactive properties that allow the highlighting of certain parts in order to allow readability.

The *Space-Time Theories Eploratorium* shows six interactive concept lattices, besides other visualizations on gravitation. A lattice for physical geometry, that relates the Leibnitzian, Galilean, Newtonian, Affine, Metric Affine, Riemannian, Lorentzian and Minkowski space-times; and a lattice for classical physical theories that includes Newton's gravitation, Newton-Cartan Gravitation, Special Relativity and General Relativity were both based on Fiedman's classification of space-time theories.[11,17] The lattice for f(R) theories is based on Ref. 18, while the lattice for non-local theories is based on Refs. 19–22. In *Theory and Experiment in Gravitational Physics*[3] we can find Clifford Will's cookbook to obtain the post newtonian parameters of any metric, Lagrangian based theory. Also, several tables that provide classes of attributes such as variables (*Domains*), field equations and motion equations (*Relations*), constrictions and cosmological parameters, as well as special functions and constants (*distinguished Elements*). From this cookbook, we can extract two data bases for 14 viable theories: one for the Lagrangian formulation and a second one for the PPN formulation.

The concept lattices clearly show their theoretical foundations and structure: a Manifold, a connection, a metric tensor, as well as their conditions, typifications and relations. Previous partial models, such as a Lorentzian space-time for GR rests in higher levels, while specific models of a theory, such as SR for GR would rest under the object-concept node. So foundations rest above the object-concepts, while specific models of T under the object-concept.

[c]Similarly, this can also be done with the free software Jupyter.
[d]D3: a java script library developed by Mike Bostock. Explore d3js.org and https://bl.ocks.org/mbostock.
[e]Explore the resulting visualizations at: http://remo.cua.uam.mx/vis/Exploratorium/.

The *Space-Time Theories Exploratorium* shows also a *holon* of theories, that is, a big lattice obtained by the addition of all contexts explored so far. Such addition requires carefulness, as some attributes are not considered in the original theoretical element *ET*, but are necessary in a wider context and may need specific conditions too. That is the case of the affine connection, common to all connections, that may posses torsion or non metricity, with the condition that it is zero for all points of the manifold for some theories. As exploring these lattices, one can easily point out the different notions for metric theories (those that follow geodesics, those that depend only on curvature with no torsion, etc.); point those that follow Einstein's equation, those that are local or non-local, those that have a flat metric, or the attributes of some specific theory, among other features.

## 4. Citation Network

Order between citations and references of seminal articles provides also insight into the path and tendencies of research. Papers provide cited references that tell us where their ideas come from, and citations that tell us about their resonance. The *Exploratorium* includes a citation network for the 200 seminal articles of the *Gravity Apple Tree* in order to represent the popularity of some ideas and their distances. One can pinpoint seminal articles and locate thematic clusters for more than eighty thousand citations and references. Data was downloaded from the Science Citation Index (SCI) of Thomson Reuters; the electronic archive for pre-publications of scientific articles and automatized distribution server ArXiv.org, operated by the University of Cornell; the Astrophysics Data System (ADS), a portal and digital library of NASA that keeps three bibliographic data bases with more than 11.3 million registrations of bibliographic data on physics, astronomy and astrophysics. Larger nodes correspond to seminal articles as their in-degree is sized compared. It is a force directed layout visualized with Gephi 0.9.2, using Sci2 tool.[f] Process is explained in Ref. 23.

## 5. Insights and Future Work

I have presented several examples of hierarchical networks for gravitation theories: a temporal network, concept lattices and a citation network, expecting to bring some insight on the plurality, composition and evolution of the field of gravitation. These networks can be found in an interactive display in the web, facilitating exploration. It is clear that there is a hierarchy that dominates the structure and development of scientific theories, and that such orders can be well represented in directed networks, and that complex networks are easily explored in interactive displays. These are very rich representations full of valuable information and interactive links. Design

[f]Sci2 Team. (2009). Science of Science (Sci2) Tool. Indiana University and SciTech Strategies, https://sci2.cns.iu.edu.

in several immersing levels, color coding, and extra information displays are important to visualize the distribution of attributes and not to overcrowd the image. One can color code theories according to their viability or experimental adequacy, and distinguish classes of attributes and their level of appearance. Aesthetic designs such as the tree-like shape may likely be hanged as pleasant posters while displaying a synthesis full of information. Meanwhile, interactive displays may turn an abstract and hard matter into a ludic, mnemonic experience. The concept lattices show topologic and geometric considerations standing at the top levels, following a downward construction of theories, leaving physical considerations in the lower levels. Classes of theories and classes of models, such as potential and actual models may be shown explicitly. As expected, the criteria used in most classifications is variable. We can find different formalisms (Lagrangian, PPN, non-Lagrangian), formulations (holonomic and anholonomic), direction of reconstruction (upwards and downwards), classes of attributes, global and local considerations, relativist or absolutist views, etc. So combining or adding up contexts becomes quite a delicate enterprise, but a clarifying one.

This work in progress attempts to include non-metric theories, MoNDian theories and M theories in order to be more inclusive, and finding a common classification that allows to join all conceptual networks. Other classifications being currently analyzed and visualized are included in Refs. 4, 5, 17, 24. All these complementary graphic representations should give us a fair idea of the paths taken by scientists as they explore the obscurities of gravitation and the general mathematical suppositions of the geometric approach to gravitation. Future work on the citation network will include the temporal component that will order the publications in terms of their timestamp and the reference network, as well as extending the seminal data base.

**Acknowledgements**

This work is part of my PhD Thesis on Studies of Science at Universidad Autonoma Metropolitana-Cuajimalpa, with the direction of Mario Casanueva, Diego Mendez (UAM-C) and Sergio Mendoza (IA- UNAM). Special thanks to Arturo Espinosa and Irving Morales (C3-UNAM) for the assistance on programming, as well as all the feedback and validation received at the Marcell Grossmann Meeting (MG 2018).

**References**

1. K. Borner, *Atlas of Science: Visualizing what We Know* (MIT Press, 2010).
2. M. E. Aldama, The gravity apple tree, *Journal of Physics: Conference Series* **600**, p. 012050 (2015).
3. C. Will, *Theory and Experiment in Gravitational Physics* (Cambridge University Press, 1993).

4. B. Famaey and S. S. McGaugh, Modified newtonian dynamics (mond): Observational phenomenology and relativistic extensions, *Living Reviews in Relativity* **15** (2012).
5. M. Blagojevic and F. W. Hehl, *Gauge Theories of Gravitation: A Reader with Commentaries*Classification of Gauge Theories of Gravity, Classification of Gauge Theories of Gravity (Imperial College Press, 2013).
6. J. Moffat, *Reinventing Gravity: A Physicist Goes Beyond Einstein* (Dundurn, 2008).
7. J. Magueijo, *Faster Than the Speed of Light: The Story of a Scientific Speculation* (Perseus Book Group, 2003).
8. K. Thorne, *Black Holes and Time Warps: Einstein's Outrageous Legacy* Commonwealth Fund Book Program, Commonwealth Fund Book Program (W.W. Norton, 1994).
9. R. Sanders, *The Dark Matter Problem: A Historical Perspective* (Cambridge University Press, 2010).
10. B. Ganter, G. Stumme and R. Wille, *Formal Concept Analysis: Foundations and Applications*Lecture Notes in Artificial Intelligence, Lecture Notes in Artificial Intelligence (Springer, 2005).
11. M. Friedman, *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science* (Princeton University Press, 1983).
12. C. Balzer, W.and Moulines and J. D. Sneed, *Una arquitectonica para la ciencia : el programa estructuralista* (Universidad Nacional de Quilmes, 2012).
13. M. Bunge, *Foundations of physics*Springer tracts in natural philosophy, Springer tracts in natural philosophy (Springer-Verlag, 1967).
14. C.-U. Moulines, *A Study in Protophysics*, in *Scientific Philosophy Today: Essays in Honor of Mario Bunge*, eds. J. Agassi and R. S. Cohen (Springer Netherlands, Dordrecht, 1982), Dordrecht, pp. 207–224.
15. B. Davey and H. Priestley, *Introduction to Lattices and Order*Cambridge mathematical text books, Cambridge mathematical text books (Cambridge University Press, 2002).
16. R. Belohlavek, Introduction to formal concept analysis (2008).
17. N. Dewar, Maxwell gravitation, *Philosophy of Science* **85**, 249 (2018).
18. T. P. Sotiriou and V. Faraoni, f(r) theories of gravity, *Reviews of Modern Physics* **82**, 451 (January 2010).
19. S. Mendoza, Extending cosmology: the metric approach, *ArXiv e-prints* (August 2012).
20. D. A. Carranza, S. Mendoza and L. A. Torres, A cosmological dust model with extended f($\chi$) gravity, *European Physical Journal C* **73**, p. 2282 (January 2013).
21. E. Barrientos, F. S. N. Lobo, S. Mendoza, G. J. Olmo and D. Rubiera-Garcia, Metric-affine f (R ,T) theories of gravity and their applications, *Physical Review D* **97**, p. 104041 (May 2018).

22. E. Barrientos and S. Mendoza, MOND as the weak field limit of an extended metric theory of gravity with a matter-curvature coupling, *Physical Review D* **98**, p. 084033 (Oct 2018).
23. K. Börner and D. E. Polley, *Visual Insights: A Practical Guide to Making Sense of Data* (The MIT Press, 2014).
24. K. S. Thorne, D. L. Lee and A. P. Lightman, Foundations for a theory of gravitation theories, *Phys. Rev. D* **7**, 3563 (Jun 1973).

# The Struble–Einstein correspondence

Marcus C. Werner

*Center for Gravitational Physics, Yukawa Institute for Theoretical Physics,*
*Hakubi Center for Advanced Research,*
*Kyoto University,*
*Kitashirakawa Oiwakecho Sakyoku, Kyoto 606-8502, Japan*
*E-mail: werner@yukawa.kyoto-u.ac.jp*

This article presents hitherto unpublished correspondence of Struble in 1947 with Menger, Chandrasekhar, and eventually Einstein, about a possible observational test supporting Einstein's special relativity against Ritz's emission theory using binary stars. This 'Struble effect,' an acceleration Doppler effect in emission theory, appears to have been overlooked, and the historical context, including de Sitter's binary star test of special relativity, is also discussed.

*Keywords*: History of relativity; special relativity; emission theory.

## 1. Introduction

The nature of light propagation is central to the history of relativity theory. An early rival of Einstein's special relativity was emission theory, which was also consistent with the negative result of the Michelson-Morley experiment. In 1947, Raimond Struble (1924–2013), then a student at the University of Notre Dame and later a professor of mathematics at North Carolina State University in Raleigh, corresponded with Chandrasekhar at Yerkes Observatory and Einstein at the Institute for Advanced Study about an observational test of emission theory using binary stars. While de Sitter, among others, had shown earlier that binary star observations do, indeed, support special relativity against emission theory, Struble pointed out a new Doppler effect due to the binary's orbital acceleration, which appeared to have been overlooked.

These letters and related documents, hereinafter denoted *Struble-Einstein correspondence* (SEC), are held privately and are listed in the Appendix. The purpose of this article is twofold: on the one hand, to present this unpublished correspondence with Einstein and, on the other hand, to draw attention to Struble's effect and place it in its historical context of binary star tests of emission theory.

## 2. Historical background

### 2.1. *Emission theory*

After Einstein's paper on special relativity[1] was published in 1905, emission theory was considered as a viable alternative to understand the negative result[2] of the Michelson-Morley interferometer experiment without the need to adopt special relativistic kinematics. This theory was developed in particular by Walter Ritz (1878–1909), Einstein's junior by one year at the Zurich Polytechnic Institute (presently the ETH), who proposed such a framework[3] in 1908 (see, e.g., Ref. 4 for historical

aspects): light is emitted at a constant speed $c$ in any direction $\mathbf{u}$, $|\mathbf{u}| = 1$, relative to a source moving at velocity $\mathbf{v}$ relative to an observer. Thus, the velocity of light relative to the observer becomes

$$\mathbf{c}_r = c\mathbf{u} + \mathbf{v}\,. \tag{1}$$

Just as Einstein's theory, this is consistent with Michelson-Morley because light has a constant speed relative to the interferometer in this case (without being *universally* constant), so there is no diurnal variation of the light travel time.

To test the emission theory with a similar interferometer experiment, one would either require suitably moving mirrors, as was indeed considered by Michelson[5] himself in 1913, or a light source moving fast with respect to the interferometer, such as a celestial light source. This point was made by Tolman in 1912:

> "Hence if the Ritz theory should be true, using the sun as source of light we should find on rotating the apparatus a shift in the [interference] fringes of the same magnitude as originally predicted for the Michelson-Morley apparatus where a terrestrial source was used. If the Einstein theory should be true, we should find no shift in the fringes using any source of light." (Ref. 6)

A negative result using star light was finally achieved by Tomaschek[7] in 1924, providing strong evidence against emission theory and corroborating special relativity. Meanwhile, other experimental tests of emission theory were sought, and it is here that binary stars enter the picture.

### 2.2. *Binary star tests*

Consider an idealized Newtonian binary star system consisting of a small component $S$ orbiting a large component with speed $v$ in a circular orbit of radius $a$, whose centre is effectively the barycentre of the system. We assume also that the observer $E$ is coplanar with the orbit at large distance $d$, $d \gg a$, and that light is emitted isotropically by $S$ at the emission time $t_0$. Denoting the constant orbital angular velocity of $S$ by $\omega$, we have $v = a\omega$. This setup is illustrated in Fig. 1.

Suppose, now, that emission theory holds. Then at emission time $t_0$,

$$c_r(t_0) = c - v_r(t_0) \simeq c - v\cos\theta(t_0) = c - a\omega\cos\omega t_0\,, \tag{2}$$


Fig. 1. Schematic binary star system as considered by Struble in SEC 3.

from (1). This speed is constant from $S$ to $E$, which are separated by the distance $d(t_0) \simeq d + a \sin \omega t_0$, so a light signal emitted by $S$ at $t_0$ is received by $E$ at the observation time

$$\begin{aligned} t &= t_0 + \frac{d(t_0)}{c_r(t_0)} \simeq t_0 + \frac{d + a \sin \omega t_0}{c - a\omega \cos \omega t_0} \\ &= \frac{d}{c} + t_0 + \frac{a}{c} \sin \omega t_0 + \frac{ad\omega}{c^2} \cos \omega t_0 + \mathcal{O}\left(\frac{a^2\omega^2}{c^2}\right) \end{aligned} \tag{3}$$

at leading order, with $v \ll c$. Hence, the observation time $t$ is a linear function of the emission time $t_0$ plus a small oscillation, as would be the case in Einstein's theory with $c_r(t_0) = c = \text{const}$. However, note that $d\omega/c$ can be large, so the fourth term in (3) need not be negligible, and this gives rise to characterstic optical effects of emission theory: if $d\omega/c$ is sufficiently large, the map $t \mapsto t_0$ is no longer one-to-one, and a given observation time corresponds to multiple emission times. In other words, $E$ would observe $S$ at the same time in multiple points of its orbit. In his *Introduction to the Theory of Relativity*, Bergmann referred to this phenomenon of multiple star images as "ghost stars", whose absence provided strong evidence against the emission theory of Ritz and for the constancy of the speed of light in the sense of Einstein's special relativity theory:

> "In some cases, we should observe the same component of the double star system simultaneously at different places, and these 'ghost stars' would disappear and reappear in the course of their periodic motions. [...] However, no trace of any such effect has ever been observed. This is sufficiently conclusive to rule out further consideration of this hypothesis." (Ref. 8, pp. 19–20)

Daniel Comstock (1883–1970), who visited J. J. Thomson at Cambridge during 1906–1907 before returning to MIT where he later became a co-founder of *Technicolor*[9], seems to have been the first to seriously consider astronomical tests of emission theory, in particular its implications for binary stars. At a meeting of the American Physical Society in October 1909 at Princeton, he noted that

> "The assumption that the velocity of light depends on that of the source has, so far as the author is aware, never been properly examined. This is strange, but is probably explainable as a natural result of the complete trust which has been put for years in the conception of an ether." (Ref. 10)

In 1910, Comstock devised a criterion[11] to check whether the orbit of an observed binary obeys Keplerian motion, thus revealing anomalies such as a changing speed of light according to emission theory. Apparently unaware of Comstock's work, de Sitter also proposed[12,13] to use binary star orbits to constrain emission theory in February 1913. This seems to have had more impact, and generated considerable

interest within the physics community during that year, e.g. with papers by Guthnick[14] and Freundlich[15]. In his reply[16], de Sitter agreed to a parametric extension generalizing Eq. (1),

$$\mathbf{c}_r = c\mathbf{u} + \kappa\mathbf{v}\,,$$

where $\kappa = 1$ in Ritz's original theory and $\kappa = 0$ in special relativity. The task, then, was to constrain $\kappa$, and by considering the system $\beta$ Aurigae, he found $\kappa < 0.002$ and hence good agreement with special relativity. Einstein himself was, naturally, very interested in these developments. In May 1913, he wrote to Ehrenfest,

> "The matter concerning binary stars is very nice, provided that the [spectral] line movements are really measured accurately enough to check the Keplerian motion to some extent." (Ref. 17, doc. 441, p. 523)[a]

At this time, Einstein was working on the generalization of his theory of relativity to include gravity. Having predicted a gravitational deflection of light in 1911 using an early version of general relativity[18], he was in touch with Freundlich who had proposed an observational test of gravitational lensing[19] (cf. also Ref. 20) in January 1913. In a letter to Freundlich in August 1913, mainly about this issue of gravitational lensing, Einstein also expressed the paramount importance of the binary star tests for relativity theory:

> "I am also very curious about the results of your investigations concerning the binary stars. If the speed of light depends on the speed of the light source even only in the slightest, then my entire theory of relativity, including the theory of gravity, is wrong." (Ref. 17, doc. 472, p. 555)[b]

In addition to orbital properties, Struble thought that binary systems can provide other observational evidence for special relativity, as we shall discuss now.

## 3. Struble's effect

### 3.1. *Motivation and results*

As a student, in the wake of the atomic bombs ending the Second World War, Struble became interested in the underlying physics, the mass defect $\Delta E = \Delta m c^2$ and its origin in special relativity:

[a] Own translation. Original: "Die Sache mit den Doppelsternen ist sehr hübsch, vorausgesetzt, dass die Linienwanderungen wirklich exakt genug gemessen sind, um einigermassen die Keplersche Bewegung nachzuprüfen."

[b] Own translation. Original: "Sehr neugierig bin ich auch auf die Ergebnisse Ihrer Untersuchungen über die Doppelsterne. Wenn die Lichtgeschwindigkeit auch nur im Geringsten von der Geschwindigkeit der Lichtquelle abhängt, dann ist meine ganze Relativitätstheorie inklusive Gravitationstheorie falsch."

> "Since the constancy [of the speed of light] seemed to me unbelievable, I started thinking about other possible evidence and came upon the double-star phenomena [...]." (SEC 11)

Struble's adviser was initially the Viennese mathematician Karl Menger (1902–1985), a pioneer of fractal and probabilistic metric geometry, who had emigated to the USA joining the University of Notre Dame in 1937, where he started the mathematics PhD program and hosted Gödel. Even after moving to the Illinois Institue of Technology in 1946, he stayed in touch with Struble, whom he had encouraged to work on the binary star problem:

> "I asked Mr. Struble to study the geometrical aspects of DeSitter's theory and in pursuing this study he discovered another phenomenon which would be observable if the principle [of the constancy of the speed of light] were wrong, namely, a change in the wave length of the light emitted by the companion. This change would be entirely different from the Doppler effect, and quantitatively much larger than DeSitter's phenomena." (SEC 4)

The argument (SEC 3) can be summarized as follows. Consider again a binary star system in emission theory as discussed in Sec. 2.2, with speed of light $c_r$ relative to the observer $E$ given by Eq. (2). Now the observed frequency of light $\nu$ is related to the frequency $\nu_0$ at the emitter $S$ by

$$\nu = \nu_0 \frac{c_r}{c} \simeq \nu_0 \left(1 - \frac{v}{c} \cos\theta\right).$$

In other words, there is a standard Doppler effect in frequency which is linear in the orbital speed $v$, but it should be noted carefully that the wavelength of light remains unchanged, $\lambda = \lambda_0$, as pointed out by Tolman[21] in 1910. However, Struble observed that the changing of $c_r$ with $\theta$ or emission time $t_0$, that is, the orbital acceleration, gives rise to an actual change in wavelength which can be very large. This is the "change [...] entirely different from the Doppler effect" which Menger referred to in the quotation above.

To see this, note that during the time $\Delta t_0 = \lambda_0/c$ needed to emit a monochromatic wave of length $\lambda_0$ at $S$, the speed of light relative to $E$ has changed by $\Delta c_r \simeq \Delta t_0 a\omega^2 \sin\omega t_0$, from Eq. (2). So after a travel time of approximately $d/c$ to reach $E$, the wavelength has changed by

$$\Delta\lambda = \lambda - \lambda_0 \simeq -\Delta c_r \frac{d}{c} \simeq -\lambda_0 \frac{ad\omega^2}{c^2} \sin\theta \tag{4}$$

at leading order.[c] As before in Eq. (3), the term containing $d$ may give rise to a large observational effect.

In February 1947, Struble expounded his argument in a letter to Chandrasekhar at Yerkes Observatory, who noted in his reply, shown in Fig. 2, that

[c]This is corrected from Struble's derivation (SEC 3), which contains ambiguities in the angle and sign.

> "[...] it does seem that no one has thought of the effect of acceleration on the velocity of light on classical lines." (SEC 1)

However, he cautioned that group velocity rather than phase velocity should be considered, and that

> "[...] one's interest is somewhat dimmed by the consideration that the effect is not present anyway." (SEC 1)

Special relativity was well established and tested by that time, and the latter remark is also echoed by Tate (SEC 2), the editor of *Physical Review*, where Struble had tried to publish his work as a *Letter to the Editor*. Six days after this rejection, Menger tried intercede on behalf of his former student (SEC 4), and received a reply from the assistant editor Hill (SEC 6). While acknowledging that there are theoretical issues with accelerated light sources even in classical electrodynamics,[d] Hill agreed with Tate that special relativity was already sufficiently corroborated and therefore questioned that

> "[...] at the present time Ritz' theory requires a final coup de grâce, [...] [having,] like numerous other trial theories of its period, [...] earned a decent oblivion." (SEC 6)

He also pointed out that it remains unclear in Struble's derivation how the spectral energy distribution changes, a remark that reflects Chandrasekhar's concern about the group velocity.

### 3.2. *Correspondence with Einstein*

Having thus received encouragement as well as rejection, Struble decided to approach Einstein himself as arbiter "for final judgement" (SEC 7) regarding the validity of his derivation. In October 1947, Struble sent a letter to Einstein at the Institute for Advanced Study, emphasizing his new effect:

> "Many spectroscopic binary systems would exhibit these ghost stars due to their geometric characteristics, but as of now I have been unable to find any visible binary systems which would. Mostly because of this I investigated geometrical consequences of Ritz' hypothesis which concerned these visible double stars, in an attempt to agument [*sic*] de Sitter's repudiation." (SEC 7)

The accompanying computation (SEC 8) recapitulated the one sketched above (SEC 3), although Struble expressed his result now in terms of frequency rather than wavelength.

[d] This remark presumably refers to the Abraham-Lorentz backreaction force which is proportional to the *third* time derivative of the trajectory, thus causing problems with inital values.

In his reply of November 1947 (SEC 9), shown in Fig. 3, Einstein agreed that Struble's argument "is essentially right", although he preferred his own derivation. Using again the notation of Sec. 2.2, Einstein obtained an approximation for the observation time as a function of emission time,

$$t \simeq t_0 + \frac{ad\omega}{c^2}\cos\omega t_0\,,$$

as in Eq. (3), ignoring the constant $d/c$ and neglecting the other terms. Infinitesimal intervals are therefore related by

$$\mathrm{d}t \simeq \mathrm{d}t_0\left(1 - \frac{ad\omega^2}{c^2}\sin\omega t_0\right). \tag{5}$$

Assuming that the phase $\varphi$ of the light wave obeys a simple harmonic oscillation, Einstein noted that the frequency is

$$\nu^2 = -\frac{1}{\varphi}\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2}\,,$$

and the chain rule implies[e]

$$\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2} = \frac{\mathrm{d}^2\varphi}{\mathrm{d}t_0^2}\left(\frac{\mathrm{d}t_0}{\mathrm{d}t}\right)^2 + \frac{\mathrm{d}\varphi}{\mathrm{d}t_0}\frac{\mathrm{d}^2t_0}{\mathrm{d}t^2}\,,$$

so that, neglecting the last term and using equation (5),

$$\nu \simeq \nu_0\frac{\mathrm{d}t_0}{\mathrm{d}t} \simeq \frac{\nu_0}{1 - \frac{ad\omega^2}{c^2}\sin\omega t_0}\,,$$

which, in the given approximation, is in agreement with Struble's result. Einstein also supported his interpretation, stating that

> "The effect is, of course, in most cases of observation much greater than the ordinary Doppler-effect and without doubt incompatible with the experimental facts" (SEC 9),

but, in closing his letter, reminded Struble that careful checking of the literature is in order since

> "I should be quite astonished if De Sitter would not have made this little calculation. In any case, you should carefully look into his paper before publishing something about it." (SEC 9)

[e]The letter has $\frac{\mathrm{d}^2\varphi}{\mathrm{d}t^2} = \frac{\mathrm{d}^2\varphi}{\mathrm{d}t_0^2}\frac{\mathrm{d}t_0}{\mathrm{d}t} + \frac{\mathrm{d}\varphi}{\mathrm{d}t_0}\frac{\mathrm{d}^2t_0}{\mathrm{d}t^2}$, but the mistake in the first term of the right-hand side does not propagate to the following computations. Despite the approximations made, equality signs are used throughout the letter (SEC 9).

### 3.3. *Aftermath and assessment*

Einstein's letter made it clear that the question of priority for Struble's result had to be settled before any further attempt at publication should proceed. In January 1948, Menger advised Struble to obtain the original papers regarding the binary star effects of emission theory (SEC 10), in particular the 1913 articles in German by de Sitter[12] and Freundlich[15]. As discussed in Sec. 2.2, these are concerned primarily with apparent distortions of the binary orbits – indeed, Struble's effect is treated neither there nor in the other related papers, by Comstock[10,11], de Sitter[13,16], Guthnick[14] and Tolman[6,21].

A heuristic argument which comes close to Struble's is found in Bergmann's *Introduction to the Theory of Relativity*: using the notation of Sec. 2.2 as before, the light travel time to the observer is approximately $d/c$ so that

$$\frac{\Delta t}{\Delta c} \simeq -\frac{d}{c^2}, \quad \text{and since } \Delta c = v, \quad \Delta t \simeq -\frac{vd}{c^2}\,.$$

The amplitude of emission theory effects will then be determined by the ratio $\Delta t/T$, where $T := 2\pi/\omega$ is the orbital period of the binary (Ref. 8, p. 20). Hence,

$$\Delta t\ \omega \simeq -\frac{ad\omega^2}{c^2}\,,$$

as in Eq. (4) for the amplitude of Struble's effect on wavelength.[f] Still, there is no explicit discussion of Struble's effect and how it differs from the ordinary Doppler effect here.

Hence, it does appear that Struble's effect had not been noticed before or, at any rate, was not at all well known, in accordance with the assessment by Chandrasekhar in SEC 1. It is also interesting to note that the effect does not appear in the 1965 paper by Fox[22], which is a detailed critique of all available evidence against emission theories. On the other hand, it must be conceded that Struble's argument is purely kinematical as it stands, and a final assessment of its physical merit should include a deeper investigation of dynamical aspects as well, along the lines suggested by Chandrasekhar (SEC 1) and Hill (SEC 6) mentioned before.

In the end, Struble pursued his interests in mathematics and its applications, and never resumed his attempt to publish this work at the interface of astronomy and theoretical physics:

> "[...] the main reason I couldn't follow through was that I would need to use Einstein's superior version. Though his letter granted me precidence

[f] In fact, it is likely that Struble knew Bergmann's treatment, since the first edition had been published in 1942 and it had become a popular textbook on relativity theory: there had been four printings by 1948, according to the colophon of the Prentice-Hall, New York NY, 1948, edition. Moreover, Struble used the rather imaginative expression "ghost stars" in his letter to Einstein (SEC 7), which is also found in Bergmann (Ref. 8, p. 19) but not in the original papers on binary star effects of emission theory mentioned above.

[*sic*] on the idea, I couldn't publish his modifications as mine. I probably should have discussed it with him at the time." (SEC 11)

## 4. Final remarks

In addition to its scientific interest, the Struble–Einstein Correspondence gives detailed insight into individual motivations and practical aspects of the scientific process, as described in this article. Einstein's reply to the student Struble is also a testament to his often generous attitude towards researchers outside his immediate circle of professional colleagues.[g]

Finally, it may be mentioned that the *Albert Einstein Archives* preserve a copy of Struble's correspondence (SEC 7 in Ref. 24, SEC 8 in Ref. 25) as well as the autograph draft of Einstein's letter written in German[26], although the English version that Einstein eventually sent to Struble (SEC 9) appears to be lacking.

## Acknowledgments

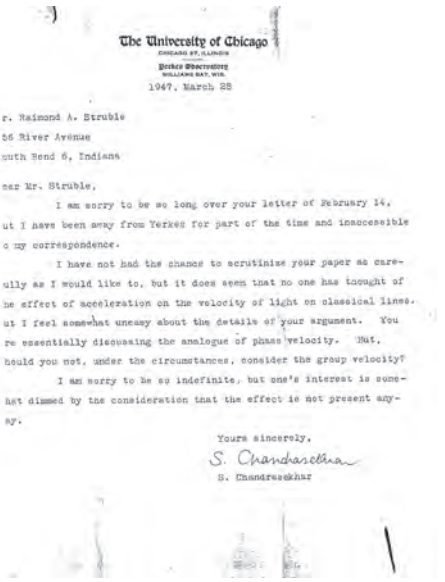
MCW would like to thank Masataka Fukugita, Domenico Giulini and Dennis Lehmkuhl for helpful discussions.

## Appendix

The documents of the Struble-Einstein Correspondence are held by Struble's estate in Raleigh, NC, USA, and consist of the following.

SEC 1. S. Chandrasekhar to R. A. Struble, typed and signed letter, 1 page, 1947 March 28.
SEC 2. J. T. Tate to R. A. Struble, typed and signed letter, 1 page, 1947 July 23.
SEC 3. [R. A. Struble:] An observable consequence of the ballistic hypothesis, typed and annotated document, 1 page (incomplete), no date.
SEC 4. K. Menger to J. T. Tate, typed letter, 2 pages, 1947 July 29.
SEC 5. K. Menger [to R. A. Struble], signed autograph letter, 1 page, 1947 August.
SEC 6. E. L. Hill to K. Menger, typed, signed and annotated letter, 2 pages and copy, 1947 August 11.
SEC 7. R. A. Struble to A. Einstein, typed letter, 1 page, 1947 October 15.
SEC 8. R. A. Struble: Consider a double star system [...], typed document, 2 pages, [1947 October 15].
SEC 9. A. Einstein to R. A. Struble, typed, signed and annotated letter, 2 pages, 1947 November 5.
SEC 10. [K. Menger] to R. A. Struble, autograph letter, 1 page, 1948 January.
SEC 11. R. A. Struble, autograph letter, 1 page, 1996 May 5.

[g]Another example of this is his interaction with Mandl in the early history of gravitational lensing, cf. Ref. 23.



The University of Chicago
CHICAGO 37, ILLINOIS
Yerkes Observatory
WILLIAMS BAY, WIS.

1947. March 28

r. Raimond A. Struble
56 River Avenue
outh Bend 6, Indiana

ear Mr. Struble,

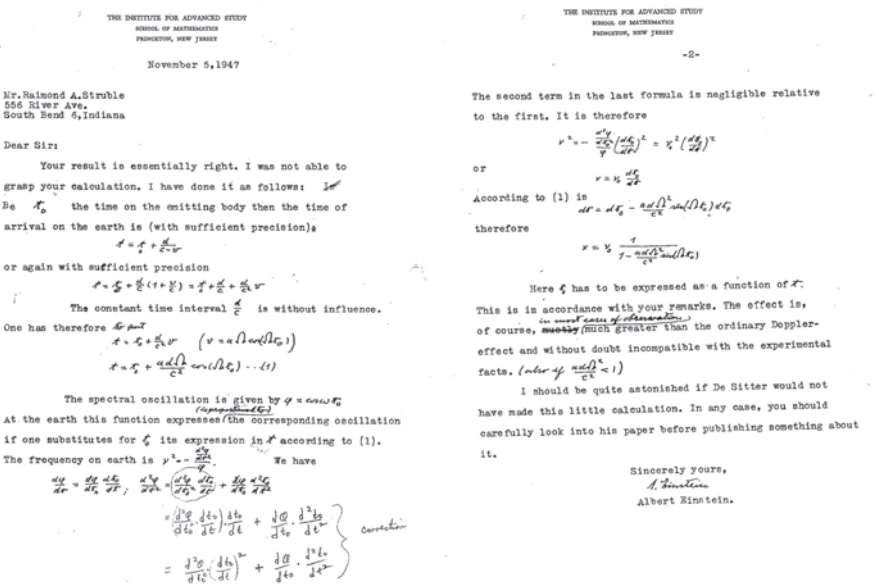
I am sorry to be so long over your letter of February 14, ut I have been away from Yerkes for part of the time and inaccessible o my correspondence.

I have not had the chance to scrutinize your paper as carefully as I would like to, but it does seem that no one has thought of he effect of acceleration on the velocity of light on classical lines. ut I feel somewhat uneasy about the details of your argument. You re essentially discussing the analogue of phase velocity. But, hould you not, under the circumstances, consider the group velocity?

I am sorry to be so indefinite, but one's interest is somehat dimmed by the consideration that the effect is not present anyay.

Yours sincerely,
S. Chandrasekhar
S. Chandrasekhar

Fig. 2. Letter by Chandrasekhar to Struble (SEC 1).



THE INSTITUTE FOR ADVANCED STUDY
SCHOOL OF MATHEMATICS
PRINCETON, NEW JERSEY

November 5,1947

Mr.Raimond A.Struble
556 River Ave.
South Bend 6,Indiana

Dear Sir:

Your result is essentially right. I was not able to grasp your calculation. I have done it as follows:

Be $t_0$ the time on the emitting body then the time of arrival on the earth is (with sufficient precision):

$$t = t_0 + \frac{R}{c - v}$$

or again with sufficient precision

$$t = t_0 + \frac{R}{c}\left(1 + \frac{v}{c}\right) = t_0 + \frac{R}{c} + \frac{R}{c^2}v$$

The constant time interval $\frac{R}{c}$ is without influence. One has therefore

$$t = t_0 + \frac{R}{c^2}v \qquad \left(v = a\Omega\cos(\Omega t_0)\right)$$

$$t = t_0 + \frac{a\Omega}{c^2}\cos(\Omega t_0) \quad \cdots (1)$$

The spectral oscillation is given by $\varphi = \cos\nu t_0$ (approximately). At the earth this function expresses the corresponding oscillation if one substitutes for $t_0$ its expression in $t$ according to (1). The frequency on earth is $\nu = -\frac{d\psi}{dt}$. We have

$$\frac{d\varphi}{dt} = \frac{d\varphi}{dt_0}\cdot\frac{dt_0}{dt}; \quad \frac{d^2\varphi}{dt^2} = \frac{d}{dt_0}\left(\frac{d\varphi}{dt_0}\right)\frac{dt_0}{dt} + \frac{d\varphi}{dt_0}\cdot\frac{d^2t_0}{dt^2}$$

$$= \frac{d^2\varphi}{dt_0^2}\left(\frac{dt_0}{dt}\right)^2 + \frac{d\varphi}{dt_0}\cdot\frac{d^2t_0}{dt^2}$$

correction

THE INSTITUTE FOR ADVANCED STUDY
SCHOOL OF MATHEMATICS
PRINCETON, NEW JERSEY

-2-

The second term in the last formula is negligible relative to the first. It is therefore

$$\nu^2 = -\frac{\frac{d^2\varphi}{dt_0^2}\left(\frac{dt_0}{dt}\right)^2}{\frac{d\varphi}{dt_0}\left(\frac{dt_0}{dt}\right)^2} = \nu_0^2\left(\frac{dt_0}{dt}\right)^2$$

or

$$\nu = \nu_0\frac{dt_0}{dt}$$

According to (1) is

$$dt = dt_0 - \frac{a\Omega^2}{c^2}\sin(\Omega t_0)\,dt_0$$

therefore

$$\nu = \nu_0\frac{1}{1 - \frac{a\Omega^2}{c^2}\sin(\Omega t_0)}$$

Here $t_0$ has to be expressed as a function of $t$. This is in accordance with your remarks. The effect is, in most cases of observation, of course, much greater than the ordinary Doppler-effect and without doubt incompatible with the experimental facts. (since $\frac{a\Omega^2}{c^2} < 1$)

I should be quite astonished if De Sitter would not have made this little calculation. In any case, you should carefully look into his paper before publishing something about it.

Sincerely yours,
A. Einstein
Albert Einstein.

Fig. 3. Letter by Einstein to Struble (SEC 9).

## References


1. A. Einstein, *Ann. Phys. (ser. 4)* **17 (322)**, 891 (1905).
2. A. A. Michelson and E. W. Morley, *Am. J. Sci. (ser. 3)* **34**, 333 (1887).
3. W. Ritz, *Ann. de Chim. et de Phys. (ser. 8)* **13**, 145 (1908).
4. A. A. Martínez, *Phys. perspect.* **6**, 4 (2004).
5. A. A. Michelson, *Astrophys. J.* **37**, 190 (1913).
6. R. C. Tolman, *Phys. Rev. (ser. 1)* **35**, 136 (1912).
7. R. Tomaschek, *Ann. Phys. (ser. 4)* **73 (378)**, 105 (1924).
8. P. G. Bergmann, *Introduction to the Theory of Relativity* (Prentice Hall, New York NY, 1948).
9. J. M. Cattell and D. R. Brimhall (eds.), *American Men of Science*, 3rd edition (The Science Press, Garrison NY, 1921), p. 140 [biography]. *Physics Today* **23**, 92 (1970) [obituary].
10. D. F. Comstock, *Phys. Rev. (ser. 1)* **30**, 267 (1910).
11. D. F. Comstock, *Astrophys. J.* **31**, 364 (1910).
12. W. De Sitter, *Physik. Zeitschr.* **14**, 429 (1913).
13. W. De Sitter, *Proc. R. Netherlands Acad. Arts Sci.* **15 II**, 1297 (1913).
14. P. Guthnick, *Astron. Nachr.* **195**, 265 (1913).
15. E. Freundlich, *Physik. Zeitschr.* **14**, 835 (1913).
16. W. De Sitter, *Proc. R. Netherlands Acad. Arts Sci.* **16 I**, 395 (1913).
17. M. J. Klein, A. J. Kox and R. Schulmann R. (eds.), *The Collected Papers of Albert Einstein, vol. 5* (Princeton University Press, Princeton NJ, 1993).
18. A. Einstein, *Ann. Phys. (ser. 4)* **35 (340)**, 898 (1911).
19. E. Freundlich, *Astron. Nachr.* **193**, 369 (1913).
20. T. Sauer, *Arch. Hist. Exact Sci.* **62**, 1 (2008).
21. R. C. Tolman, *Phys. Rev. (ser. 1)* **31**, 26 (1910).
22. J. G. Fox, *Am. J. Phys.* **33**, 1 (1965).
23. J. Renn and T. Sauer, in: H. W. Duerbeck and W. R. Dick (eds.), *Einstein's Kosmos. Untersuchungen zur Geschichte der Kosmologie, Relativitätstheorie und zu Einsteins Wirken und Nachwirken* (Harri Deutsch, Frankfurt am Main, 2005), pp. 210–239.
24. R. A. Struble to A. Einstein, 1947 October 15, in *The Albert Einstein Archives*, Hebrew University of Jerusalem, 27-320.
25. R. A. Struble, 1947 October 15, in *The Albert Einstein Archives*, Hebrew University of Jerusalem, 27-321.
26. A. Einstein A. to R. A. Struble, 1947 October, in *The Albert Einstein Archives*, Hebrew University of Jerusalem, 27-322.

# Angelo Secchi and beyond: From the beginning of stellar spectroscopy to the expansion of the universe

Matteo Galaverni

*Specola Vaticana,*
*V-00120 Vatican City, Vatican City;*
*INAF/OAS Bologna, Osservatorio di Astrofisica e Scienza dello Spazio,*
*Area della ricerca CNR-INAF, via Gobetti 101, I-40129 Bologna, Italy*
*E-mail: matteo.galaverni@gmail.com*

The spectral classification introduced by Father Angelo Secchi, SJ (1818–1878) opened a new way to study the stars and the universe. We summarize the main achievements of his work and show how the following developments lead to the foundation of modern astrophysics.

*Keywords*: Angelo Secchi; Astrophysics; Spectroscopy; History of Astronomy.

## 1. Beyond astrometry and celestial mechanics

At the beginning of the Nineteenth Century, astronomy was concerned with astrometry and celestial mechanics (calculation of the orbits of planets, compilation of catalogues of position and brightnesses of stars and other celestial objects). In 1835 the French philosopher Auguste Comte (1798–1857) was so confident about the impossibility of obtaining experimental data on the physics of stars that he wrote:

> "Every research in relation to stars not reducible in the end to simple visual observations is perforce barred to us [...] we could never study, by any means, either their chemical composition or their mineral structure [...] our positive knowledge concerning the stars is necessarily restricted just to their geometrical and mechanical phenomena, being it entirely impossible to undertake any physical or chemical research." [1]

Already at the beginning of the Nineteenth Century the German physicist Joseph von Fraunhofer (1787–1826) began to study in detail the myriad of dark lines that appear when sunlight is refracted by a prism. Fraunhofer classified more than 570 dark lines in the solar spectrum. By the middle of the 19th Century Gustav Kirchhoff (1824–1877) explained the presence of those thin dark lines through the presence of particular chemical elements in the atmosphere of the Sun, they absorb the light at a fixed wavelength. It was the beginning of the chemical analysis of the universe through spectroscopy (see Fig. 1).

These new studies provided the necessary motivations to overcome classical astronomy and foster the birth of modern Astrophysics. The observations of the spectrum of the Sun encouraged the study of the spectra of other stars by Giovanni Battista Donati (1836–1873) in Florence, William Huggins (1824–1910) in London, Lewis Morris Rutherfurd (1816–1892) in New York and Angelo Secchi in Rome.

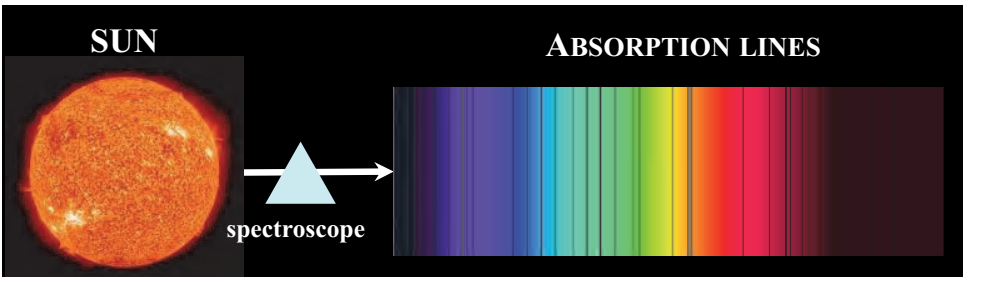



Fig. 1. Spectrum of the Sun. The outermost layers of the Sun's atmosphere absorb the radiation at fixed wavelengths, so are producing dark absorption lines in the continuous spectrum. The spectral lines became precious indicators, like a fingerprint, of the chemical elements present in the atmosphere of the observed star. Moreover, the width of the lines gives information on the abundances and physical conditions (temperature, pressure, magnetic fields) of the various chemical elements in the atmosphere of the star. Using the spectroscope it is possible to study in detail the dark lines sequence: the light coming from the star passes through a slit and then through a prism (or a diffraction grating) that separates the light into various wavelengths.

## 2. The Secchi classification of stellar spectra

The work of Father Angelo Secchi of the Society of Jesus (born Reggio Emilia 1818, died Rome 1878) contributed in a decisive way to the birth of modern Astrophysics. When he was 15 years old he entered the Jesuit order and began his formation. He studied extensively not only theology and philosophy, but also science and, in particular, physics. When in 1850 he was appointed as director of the observatory of the Roman College, he continued to look at astronomy from the point of view of a physicist. He used spectroscopy as a tool to answer a new question: not *where* are the stars (positional astronomy), but *what* are the stars, what is their composition:

> "when the field of astronomical research seemed to be dried-up, and we could only glean where others had richly reaped, here a new discovery opens up an immense horizon, which will reveal us the physical nature of the stars and show us the kind of the matter that compose them. This is the spectroscopy and its applications made by Kirchhoff and Bunsen." [2]

In 1852 he got permission to build a new observatory on the roof of the church of St. Ignatius. He had two main instruments: a Cauchoix refractor telescope (with a 16.9 cm aperture) and — after 1854 — also a Merz refractor telescope (with a 24.4 cm aperture). He also obtained an objective prism, usable on both telescopes, for spectroscopic observations of stars.

Thanks to this advanced instrumentation he started a systematical project on stellar classification based on spectra. In 1863 he published a first study, based on the observations of the spectrum of 35 of the brightest stars: *Sugli spettri prismatici della luce de'corpi celesti* in *Bullettino meteorologico dell'Osservatorio del Collegio Romano*, 31 July 1863. The following ten years were dedicated to improving this

first study, and in 1877 he published the essay *Le Stelle. Saggio di astronomia siderale*, based on the observation of more than four thousand spectra. This last classification included four stellar types based on absorption spectra, plus a fifth type containing prominent emission lines.

We present here a brief description of Secchi's classification following his treatise *Le Soleil* (1st ed: 1870; 2nd ed. vol. I: 1875, vol. II: 1877):

- The first type (see first spectrum of Fig. 2) is characterized by bluish-white stars, like Sirus: "all these stars are commonly called white, although they are slightly blue. They offer a spectrum formed by the ordinary set of the 7 colors, interrupted by four large black lines: the first in the red, the other in the blue-green, the last two in the violet" [3] (predominant hydrogen lines).
- The second group (second spectrum of Fig. 2) includes white-yellowish colored stars like the Sun: "the spectrum of these stars is perfectly similar to that of our Sun, that is, formed of very fine and very tight black lines which occupy the same position as those of the solar spectrum." [3]
- In the third type (third spectrum of Fig. 2) Secchi grouped the stars with color close to "red or orange" as $\alpha$ of Orion; "the spectrum of the third kind in very extraordinary including a double system of nebulous bands and dark lines." [3]
- "The fourth type is even more extraordinary" [3] (fourth spectrum of Fig. 2 these few stars are not necessarily colder than those of the third type, but only chemically different (faint reddish stars, predominant carbon lines).
- Lastly he also added a fifth type (see Fig. 3) characterized by the presence of pronounced emission lines in their spectra: "direct spectrum of hydrogen." [3]

Thanks to his pioneering work, Secchi is rightly considered among the founders of modern astrophysics and the father of the classification of stars based on their spectra. Finally, we would like to emphasize that all the spectra shown in Secchi's books are not photos, but they are all hand-drawn work.

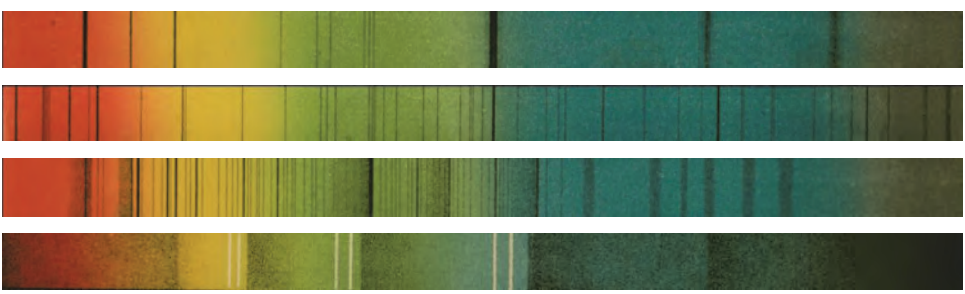


Fig. 2. Absorption spectra of a star of type I (Sirius),type II (Sun), type III (Orion), type IV (152 Schjellerup Catalogue also known as "La Superba" - Y CVn). [4]

Fig. 3. Typical emission spectrum of nebulae (type V).[4]

## 3. After Secchi

Fr. Secchi was fully aware of the new potentialities of spectroscopy: "there remain many things for us to learn: for nature is inexhaustible in its wonders; [when] we think we are arrived at the end, we find that it is only the beginning."[3,5]

### 3.1. *The Harvard classification and the Hertzsprung–Russell diagram*

The work started by Secchi was continued by numerous astronomers. In 1872 Henry Draper (1837–1882) realized the first photographs of a stellar spectrum. At the end of the Nineteenth Century the astronomer Edward Charles Pickering (1846–1919) started a large project focused on stellar classification using photographic techniques. Draper's widow decided to finance this project and in 1890 Pickering published a first catalogue based on the observation of 10.351 stars: *Draper Catalogue of Stellar Spectra*. The work continued in the subsequent years and thanks to the devotion to the project of Annie Jump Cannon (1863–1941) and collaborators it was possible to publish the *Henry Draper Catalog* containing the magnitude, position and stellar classification of 225.300 stars. Later the astronomers Ejnar Hertzsprung (1873–1967) and Henry Norris Russell (1877–1957), starting from the data contained in this catalogue, combined in a diagram the information on the temperature (color) and absolute luminosity of different stars. The H-R diagram is a key step towards understanding stellar evolution. Spectral Classification remains, even today, an important tool to identify peculiar and astrophysically interesting stars.[6]

### 3.2. *Proper motion and expansion of the universe*

Secchi was fully aware of another important information contained in the spectra: "It is the opinion of many talented spectroscopists that spectral measurements can give us a very delicate means to know also the absolute motions in the direction of the visual ray."[4] If the light source is in motion towards the observer, the wavelength will be shortened compared to the wave emitted by a source at rest (*blueshift*), vice versa if the source is moving away from the observer, the wavelength will increase (*redshift*).

Unfortunately, the instruments used by Secchi were not accurate enough to detect this modification of the wavelengths. Only at the end of the Nineteenth Century that it became possible to obtain the first reliable measurements of radial velocities when photographic spectra became available. In the 1920s Edwin Hubble

(1889–1953) and Georges Lemaître (1894–1966) derived the famous law of the expansion of the universe (Hubble-Lemaître law) using this technique.

This discovery marked the passage from a static universe to the model of an universe evolving in time. The expansion of the universe is, together with the Cosmic Microwave Background (CMB) and primordial nucleosynthesis, one of the main pillars of the Big Bang cosmological model.

## 4. The bicentenary of the birth of Angelo Secchi

Angelo Secchi gave an important contribution with his pioneering work to the birth of "New Astronomy", the modern Astrophysics. Therefore many initiatives were organized to remember the bicentenary of his birth.

The Italian Ministry of Cultural Heritage and Activities and Tourism established a National Committee for the celebrations of the bicentenary of the birth of Angelo Secchi. This committee was coordinated by the National Academy of the Sciences and the Italian National Institute for Astrophysics.

Secchi was remembered during the International Astronomical Union XXXth General Assembly in Vienna (August 20–31, 2018).[7]

In Italy conferences and workshops were held both in Rome ("The Legacy of Angelo Secchi SJ 200 years after his birth", September 3–5, 2018; "Angelo Secchi (1818–1878) SJ: the scientist and his time", September 6, 2018) and in his hometown Reggio Emilia ("Father Angelo Secchi, the person, the works, astrophysics", May 11–12, 2018). The Fifteenth Marcel Grossmann Meeting dedicated a parallel session to "Angelo Secchi and Astrophysics" (July 2, 2018). The Italian Astronomical Society (Società Astronomica Italiana) — counting Secchi among its founders when it was still called "Società delgi Spettroscopisti Italiani" — remembered Secchi during its LXII congress (Teramo, May 2–5, 2018). The City Museums of Reggio Emilia organized an exhibition dedicated to him: "All the colors of stars. Father Angelo Secchi and the birth of astrophysics" (October 20, 2018 – February 3, 2019). Also a virtual exhibition "Between Heaven and Earth"[8] was prepared by the National Committee for the celebrations of the bicentenary. A completely new biography "Decoding the Stars" was written by Ileana Chinnici and will be soon published.[9]

We conclude with the words of Secchi himself: "How many other wonders should not be found in the immensity of that space, which we cannot fathom? By taking advantage of art and science, the astronomer reveals to us more and more the greatness of God, and makes us exclaim with the Prophet: *May your works be magnified, O Lord; you did all of them in wisdom* (Psalm 104)."[10]

## Acknowledgments

The author is grateful to Christopher J. Corbally, SJ for valuable comments on this work. Thanks to Vatican Observatory (Castel Gandolfo and Tucson) for hospitality during the development of this project.

## References


1. A. Compte, *Cours de philosophie positive,* 2 vol., Bachelier, Parigi 1835.
2. A. Secchi, *Le scoperte spettroscopiche in ordine alla ricerca della nature dei corpi celesti,* Tipografia delle Belle Arti, Roma 1865.
3. A. Secchi, *Il Sole,* Tipografia della Pia Casa di Patronato, Firenze 1884.
4. A. Secchi, *Le Stelle,* Fratelli Dumolard, Milano 1877.
5. M. F. McCarthy, *Fr. Secchi and Stellar Spectra,* Popular Astronomy **58** (1959) 153.
6. R. O. Gray and C. J. Corbally, *Stellar Spectral Classification,* Princeton University Press, Princeton 2009.
7. C. J. Corbally and R. O. Gray, *The Spectral Classification of Stars over the last 200, 100, 75 years, and in the future,* in "Under One Sky - The IAU Centenary Symposium Proceedings IAU Symposium No. 349, 2018", eds. D. Valls-Gabaud, J. Hearnshaw and C. Sterken.
8. See: https://tracieloeterra.bicentenarioangelosecchi.it.
9. I. Chinnici, *Decoding the Stars: A Biography of Angelo Secchi, Jesuit and Scientist,* Brill, Leiden 2019.
10. A. Secchi, *Fisica terrestre. Coll'aggiunta di due discorsi sopra la grandezza del creato,* Loescher, Torino 1879.

# Angelo Secchi and Gnomonics

Bacchini, Sara[1]; Sigismondi, Costantino
[1] *Sapienza University of Rome, Italy;* [2] *ICRA/Sapienza*
[2] *E-mail: sigismondi@icra.it*

The equation of time determines the difference between the mean solar noon and the true one. Graphically it is represented with an analemma, drawn also in the sundials made by Secchi (e.g. the ones of Grottaferrata and Augusta). The relationship between the day of the year and the corresponding equation of time, from Ptolemy to Clavius, is presented, including some medieval cases. Finally the Sundial of Boville Ernica made in 1865 by Secchi for public utility, as in his vision of science, is presented to wish a new restauration.

## 1. Theoretical exposition

The length of each day through the course of the year is variable, it can be longer or shorter than 24 hours of many minutes, about half an hour at most, and this offset is defined as the equation of time. Let us first introduce some basic concepts.

We has to distinguish two different systems of time based respectively on the apparent solar time and the mean one. Obviously a natural unit of time is the rotation of the Earth. It represents the time in which the Sun returns to the local meridian, but it is variable due to two features of our solar system, which we analyze beyond. It is useful introduce also a motion of the Sun without irregularities, namely a uniform motion on the ecliptical plane. On this time is defined the so-called mean solar time system. A comparison of a sundials and a mechanical clock shows that the times are different and the deviation between them can be described by the equation of time. Historically, the term equation was used to tabulate the difference between an observed valued and the expected one, so the equation of time is a quantity that should be added or subtracted to the hour of the passage at the meridian to obtain the mean local noon. So, the precise definition of the equation of time is:

$$E_d = h(S) - h(M) \tag{1}$$

where $h(S)$ is the true solar time in term of hour angle and $h(M)$ is the hour angle of the equatorial mean time.

The features of our solar system are at the basis of the irregularities in the apparent motion of the Sun and the most important contributes are due to the variation of the angular velocity, according to the second Kepler's law, and the obliquity of the ecliptic. The second Kepler's law state: the area swept out per time interval is constant, thus the Sun appears to move faster (relative to the background stars) at perihelion and slower at aphelion. A Sun slower corresponds to an apparent solar day shorter, instead, a Sun faster to a longer solar day. Let us consider Kepler's equation, which allows to connect the position occupied by the planet along the keplerian orbit with the time elapsed from the moment of passage from perihelion

$$E = M + e \sin E \tag{2}$$

where M is the mean anomaly, namely the angle covered by Earth on a circular orbit with constant speed, and E the eccentric anomaly, that is the angle between the position of real Earth projected on the circular orbit and the major axis. For a better comprehension see Fig. 1.

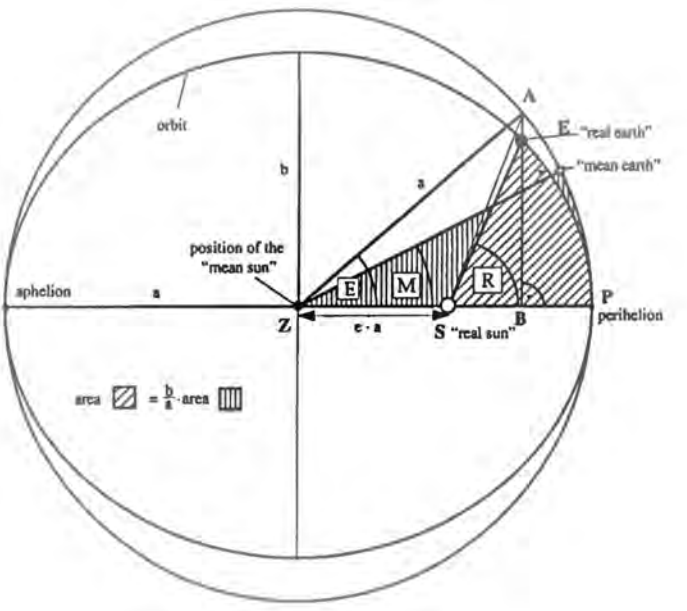



Fig. 1. Angles M and E at a specific time point.

It isn't possible to solve the equation (2) for E in closed form. Therefore, the position of the Earth can only approximated as function of time. More useful is the expansion in series of Eq. (2) in term of R, that is the real angle spanned from the real Earth along the ecliptic orbit, or the true anomaly. In fact if we make the first derivative of the expression of R respect to the time, we obtain the angular velocity and its corrections. The relation is

$$\omega(t) = \frac{2\pi}{T}[1 + 2e\cos M + ...] \tag{3}$$

The deviations amount to about double of eccentricity of the orbit.

The true solar day depends on the declination of the Sun, that is the elevation from the celestial equator. Let us ignore the variation of angular velocity and consider a circular orbit. A uniform circular motion is

$$\theta = 30^\circ(t - t_0) \tag{4}$$

where $t$ is the time passed from vernal equinox in term of months (January has a value from 0 to 1, February from 1 to 2 an do on), and $t_0 = 2,69$ the value which

correspond to 21 March. On the celestial sphere, a circle with inclination $e = 23,5^\circ$ appears as an ellipse, which can be obtained compressing all the coordinates of the circle of a factor $a/b$. In our case $b = 90^\circ$ and $a = e$. So if the coordinate of minor axis has the parametric equation $y = 90^\circ \sin\theta$, for the ellipse becomes the solar declination we are searching. Thus

$$\delta = e \sin\theta \tag{5}$$

So the solar day for a given latitude $\lambda$ and time $t$ is

$$T = 12 + \frac{2}{15} \arcsin\Big[\tan\lambda \cdot \tan\Big(e \sin(30^\circ(t - t_0))\Big)\Big] \tag{6}$$

Obviously this equation is valid for a uniform motion of the Sun. If we consider also the effect of the second Kepler's law, $T$ is affected by a small variation which tends to cumulate during a year. This discrepancy is due to the changing of the declination between two culminations. The projection of an arch of ecliptic value a on the equator gets larger of a factor $1/\cos\delta$, because the relation is $a\cos\alpha/\cos\delta$.

The mean ecliptical arch crossed by the real Sun in a day is about $1^\circ$. Instead the projection of this arch on the celestial equator is

$$\Delta\alpha = 1^\circ \frac{\cos\Big[\arctan\Big(\tan e \cos(30^\circ(t - t_0) + EC)\Big)\Big]}{\cos\Big[e \sin(30^\circ(t - t_0) + EC)\Big]} \tag{7}$$

If we make the product of $\Delta\alpha$ and $v$, an adimensional value of the angular velocity, defined as

$$v = \bar{v}[1 + 2\Delta\cos(30^\circ t)] \tag{8}$$

taking $\bar{v} \simeq 1$ and $\Delta \simeq 0,03$, the semiamplitude of the annual variation of angular velocity, we obtain an expression with dimensions of an angle. Thus the equation of time is the difference

$$ET = v\Delta\alpha - 1^\circ \tag{9}$$

Let us plot the cumulative effect of the discrepancy in time in Fig. 2.

## 2. Historical cases

The ancient name of the equation of time was *a'dilal-'ayynmbi- laydihha*, which can be translated into latin as *aequatio dierum cum noctibus suis*, that is correction of the day and their nights. The difference between apparent solar time and mean time was recognized by astronomers since antiquity, but prior to the invention of accurate mechanical clocks in the mid-17th century, sundials were the only reliable timepieces, and apparent solar time was the generally accepted standard. Ptolemy's Almagest is primarily concerned with the Sun's anomaly, and he tabulated the equation of time in his Handy Tables. Ptolemy discusses the correction needed to convert the meridian crossing of the Sun to mean solar time and takes into

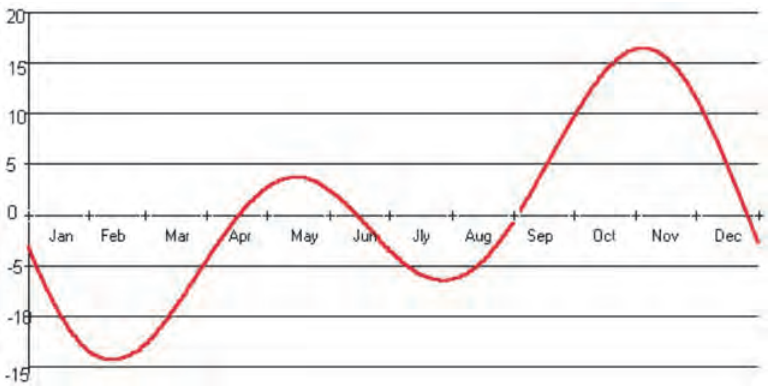



Fig. 2. Graph of true Sun's delay in the passage at the meridian.

consideration the nonuniform motion of the Sun along the ecliptic and the meridian correction for the Sun's ecliptic longitude. Some other medieval islamic astronomers made improvements to the solar tables and published tables of the equation of time. If we study some values taken from this tables, we can see that they were calculated with an high precision. Some of the most important tables we have to quote are of *Al-Battānī*, Arabic astronomer lived in IX century, John of Murs (1321), Levi ben Gerson (1288–1344) and Giovanni Bianchini (1410–1469).

## 3. The analemma

The analemma is a graph which shows the position of the Sun in the sky during a year in the same place and at the same time. A rigorous analemma can be obtained taking a photograph everyday in the same place for a year, but this is obviously impossibile because of cloudy day. The most important feature of the analemma is his form, namely an eight with a lobe bigger than the other (see Fig. 3). Due to the eccentricity of the orbit of Earth and the obliquity of the poloidal axis the form of the analemma can vary with the time and the place. If the Earth's orbit was

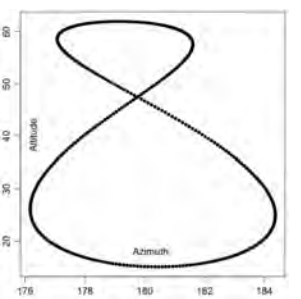



Fig. 3. The analemma, based on observed values of the azimuth and altitude of the Sun at 12.00 GMT at the Royal Observatory, Greenwich for 2006 (see Ref. 2). Note the unequal axes.

circular, two lobes would be equal, instead, if the polar axis hadn't an obliquity, the analemma would be a linear segment.

In some meridians we can find the representation of analemma for each hour of the day, except for the noon. In the meridian showed in photo (Fig. 4) is an example of that. The photo is taken in 7/23 in 2003, for that date the equation of

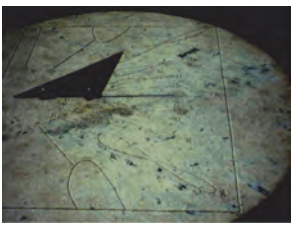


Fig. 4. Meridian with analemma. Photo taken in July 2003, see Ref. 6.


Fig. 5. Sundial of Angelo Secchi in Boville Ernica (FR) as in April 2019 and when restaured in 1996; the gnomon is parallel to the polar axis, the analemma and the equation of time vanished.[7] This sundial was restaured in 1950 by the astronomer Giuseppe Armellini, buried in the nearby Church's crypt, and again by Nicola Severino in 1996.[8] Now it deserves another restauration as it is evidenced with the rectangle. In the left low side its status in 1996.

time was −6 m 24 s. For this longitude the local noon is so at 14:50 UT. Thus, the place is about 42°30′ from Greenwich. In fact the meridian is located in Rio de Janeiro.

## References


1. D.W. Huges, B.D. Yallop, C.Y. Hohenkerk, The equation of time, 1989MNRAS.238.152 9H (1989).
2. P. Lynch, the equation of time and the analemma, Irish Math. Soc. Bulletin Number 69, Summer 2012, 47–56 ISSN 0791-5578 (2012).
3. M. Muller, The equation of time - Problem in astronomy (1993).
4. B. Van Dalen, On Ptolemy's Table for the Equation of Time, Centaurus 1994: Vol. 37: pp. 97-153 (1994).
5. S. Pietroni, Equation of time, historical examples (2018).
6. C.Sigismondi, L'equazione del tempo, arXiv:1211.2755.
7. M. L. Tuscano, Secchi e la Gnomnica, in Angelo Secchi, Quater Foligno (2012).
8. N. Severino, Meridiane Ciociare, ilmiolibro.it (2011).

# The Sodium D-Lines in Rome from Angelo Secchi to Alessandro Cacciani

Bordoni, Luigi[1] and Sigismondi, Costantino*

[1] *Sapienza University of Rome;* * *ICRA/Sapienza and ITIS G. Ferraris, Roma, Italy*
* *E-mail: sigismondi@icra.it www.icra.it/gerbertus*

Optical spectroscopy developped so much in XIX century that many results required the quantum physics to be fully explained from selection rules to quantum probability. Angelo Secchi designed or modified some instruments like the Hoffman prisms to reach a great resolution, namely better than 1 Angstrom around the Sodium D-Lines. Alessandro Cacciani, who worked in Monte Mario Observatory and Sapienza University of Rome, invented and developped the MOF Magneto-Optical Filter by tuning the magnetic field in a vaporized sodium cell. The MOF is being used in the two MOT telescopes in Anctartica, and has been used in US and EU telescope reaching an extraordinary velocity resolution in the dopplerograms of the Sun. Two inventors in comparison, at a century of temporal distance, with Mount Mario Observatory ideally inherited the Collegio Romano story, because had the Merz telescope of Secchi until the fire of July 15, 1958.

*Keywords*: History of Science, Sodium D-lines, Angelo Secchi, Alessandro Cacciani.

## 1. Father Secchi and Spectral lines

A spectral line is a dark or bright line in an otherwise uniform and continuous spectrum, resulting from emission or absorption of light in a narrow frequency range, compared with the nearby frequencies. Spectral lines are often used to identify atoms and molecules. The spectral lines of a specific element or molecule at rest in a laboratory always occur at the same wavelengths. **For this reason, we are able to identify which element or molecule is causing the spectral lines**[1].

Fig. 1. *Absorption lines are seen as black lines on a coloured background.* © *Swinburne University of Technology.*

Absorption lines appear as dark bands, often superimposed on a coloured continuum, and are the result of specific wavelengths being absorbed along the line-of-sight, and they are produced when photons from a hot, broad spectrum source pass through a cold material.

In 1814 the German Physics Joseph Fraunnhofer repeated the Newton's experiment of breaking the light of the Sun into his colors through a prism[2]. This time he greatly enlarged the resulting colored spectrum, discovering that he was interrupted by more than **600 dark lines**: the spectral lines (today we know about a million!).

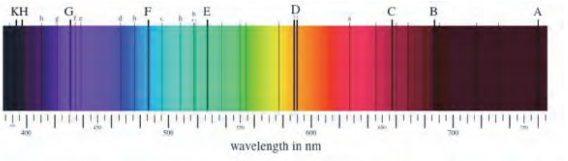



Fig. 2. *Fraunnhofer lines labelled from A to K.* © *2001–2017 The Regents of the University of California. See also*[25,26] *for identifying all the lines.*

Together with Fraunhofer, **Angelo Secchi** was the pioneer of the discovery of essence material of the stars : the stars are not simple point-like objects that move in the cosmos and shine with their own light, but they are big objects made of variegated chemical composition matter that can be investigated quantitatively.

Angelo Secchi **was the first to connect the spectroscope (a prism) to the telescope**. In this way he made an important classification of the stars, which was based on their spectral analysis. This is considered the most significant part of the astronomical work of Father Secchi, enough to make him considered the father of modern astrophysics.

Between 1863 and 1867, he examined more than 4000 stars, trying to identify their resemblances[3]; in this way he cataloged more than 500 stars, subdividing them into four main classes, according to their color and the type of spectral lines. Father Secchi was the first to put in evidence the fact that the color of the stars is related to their spectral lines and that the temperature greatly influences that spectrum. Thus identified the presence of hydrogen, a pioneering discovery of undoubted importance: today we know that the matter of the Sun is made up of 71% of hydrogen, 26% of helium, while the remaining 3% is due to other elements of the scale periodic, heavier.

## 2. Angelo Secchi and prism

In his book Il Sole[4] [pg. 220] Angelo Secchi describes the spectroscope and the prism (the so called triangular glass). A prism is any object that separates white light into the colors of the rainbow (red, orange, yellow, green, blue, indigo, and violet). It works by refracting light, and breaking it according to its wavelength.

The direct vision spectroscope, conceived by the Florentine astronomer Giovanni Battista Amici, consists of **a series of glass prisms** of different refrangibility, glued in succession, through which the beam of light passes to be analyzed[6]. In 1862, Secchi had the oppurtunity to test the performance of a similar instrument used by the French astronomer Jules Janssen. Secchi then ordered the Parisian optician Hoffman a five prism spectroscope which he used in 1863 for his first observations of stellar spectra. Subsequently, Secchi commissioned the Merz company another direct vision spectroscope, of better quality, with which in 1867 he was able to realize his spectral classification of the stars.

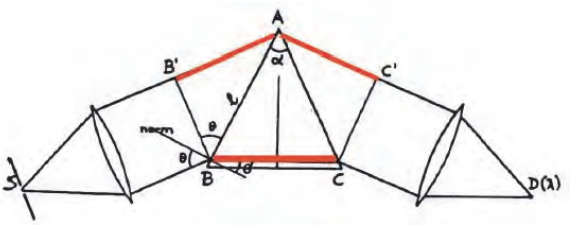



Fig. 3. *How it works a prism: the light coming from a source S is transformed into a parallel beam by an achromatic objective (called collimator), crosses the dispersing prism (ABC) and the resulting spectrum is focused by another achromatic objective (or telescope if observed with the eye) on the detector D.*[5]


Fig. 4. *A) Objective prism, Merz, Munchen, 1864, glass, brass, diameter 162 mm, angle of refraction 12 degrees. INAF museum - Osservatorio Astronomico di Roma; B) TRA CIELO E TERRA Làvventura scientifica di Angelo Secchi a cura di Ileana Chinnici e Mauro Gargano Comitato Nazionale per il Bicentenario della Nascita di Angelo Secchi* © *2018 All Rights Reserved.*

In fact Secchi used the **prism placing it in front of the objectives of the Cauchoix and Merz telescopes** of the Collegio Romano's Observatory and in his book[4] [pg. 222] he wrote that, by adding many prisms, even to small telescope, it is possible to obain a large spectrum: spectral lines become more separate and distinguishable from each other. Using powerful instuments, it is often possible to resolve lines which seem simple at first glance. Sodium D-Line is resolved very easily, but with a powerful spectroscope, it is possible to recognize that in reality it is fivefold, because between the two main lines D1 and D2, there are 3 others at least, not counting those that are outside. You can compare these observations with those of multiple stars: in both cases the result depends on the power of the instruments used. For these researches it is necessary to use 9 prisms and more; and these prisms are made of a very dispersive substance, such as heavy flint or coal sulfide.


Fig. 5. *Sodium D-Lines published in Secchis̀ book Il Sole pg. 223.*

## 3. The Sodium D-lines

Sodium is the seventh most common element in the Universe[7]. On Earth, sodium compounds such as salt give sea water its salty taste and give the white colour of salt pans in deserts. In animal life, sodium is known to regulate heart activity and metabolism. Sodium is also used in technology, e.g. in the sodium-vapour street lights, where it produces yellow-orange light.

A strong doublet at 589.0 nm and 589.6 nm is observed in the Sodium emission spectrum, identifying the D-lines as due to Sodium in the sun[8].

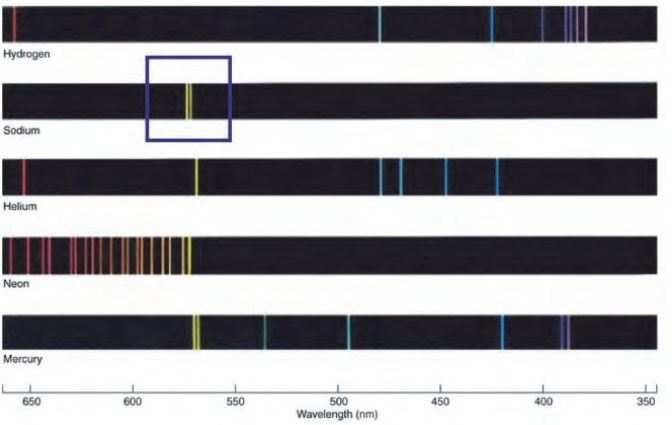



Fig. 6. *Spectral lines observed in Solar specturm.*[8]

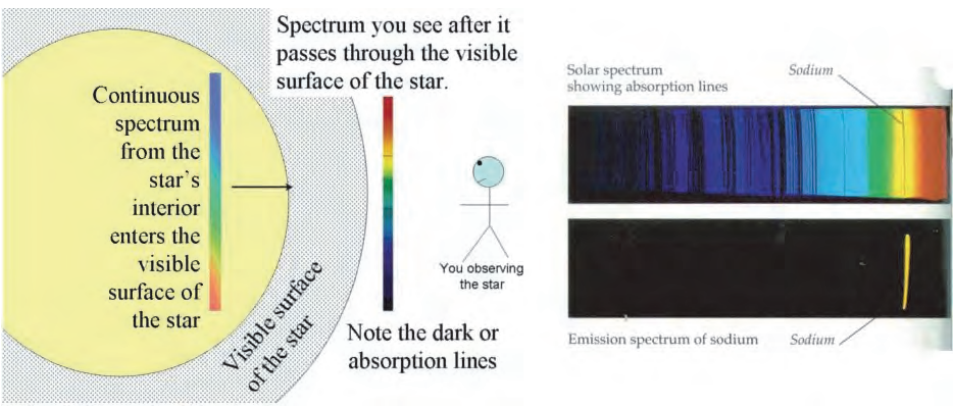



Fig. 7. *Scheme of Solar Spectrum from an observer on Earth.*[8]

## 4. Sodium D-lines and Astrophysical research

The importance of sodium is evident in many astrophysical research. By using a narrow band filter, astronomers can image all sorts of objects for excited Sodium[8]:

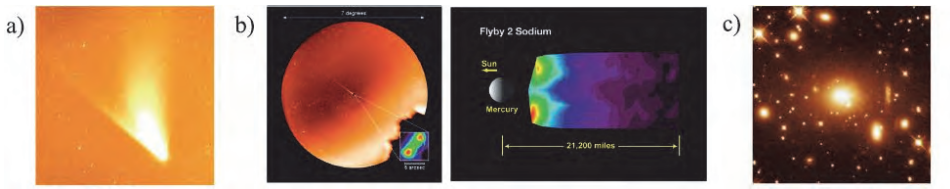



Fig. 8. *a) Sodium D-line image of the tail of the Hale-Bopp comet. b) Sodium D-line images of the comet tail of the planet Mercury. c) Supercluster of Galaxies BAS11.*

**The Sodium D-Lines are also used to measure Doppler shifts and the size of the universe** (Hubble's constant)[8].

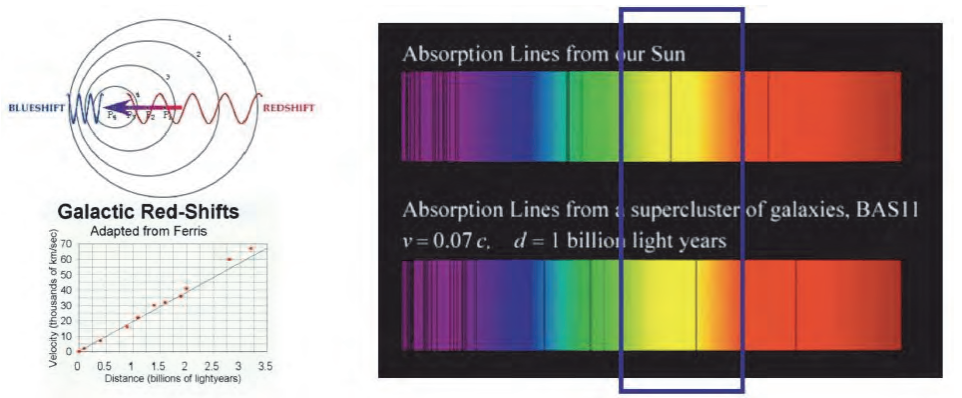



Fig. 9. *Galactic Red-Shifts measured through Sodium D-Lines.*

Sodium is very important in the Physics of Exoplanets. The study of extraterrestrial atmospheres is an active field of research[9]. Several planets outside the Solar System (exoplanets) have been observed to have atmospheres. At the present time, most atmosphere detections are of hot Jupiters or hot Neptunes that orbit very close to their star and thus have heated and extended atmospheres. **The first observed extrasolar planetary atmosphere** was made in 2001[10]. **Sodium in the atmosphere** of the planet HD 209458 b was detected during a set of four transits of the planet across its star. HD 209458 b, also given the nickname Osiris, is an exoplanet that orbits the solar analog HD 209458 in the constellation Pegasus, some 159 light-years from the Solar System.

In a recent article published in Nature An absolute sodium abundance for a cloud-free 'hot Saturn' exoplanet (May 2018)[11], Scientists have detected an

Fig. 10. *Comparison of "hot Jupiter" exoplanets (artist concept). From top left to lower right: WASP-12b, WASP-6b, WASP-31b, WASP-39b, HD 189733b, HAT-P-12b, WASP-17b, WASP-19b, HAT-P-1b and HD 209458b.*[9]

exoplanet atmosphere that is free of clouds, marking a pivotal breakthrough in the quest for greater understanding of the planets beyond our solar system. Just as an individual's fingerprints are unique, atoms and molecules have a unique spectral characteristic that can be used to detect their presence in celestial objects. The spectrum of WASP-96b shows **the complete fingerprint of sodium, which can only be observed for an atmosphere free of clouds**.

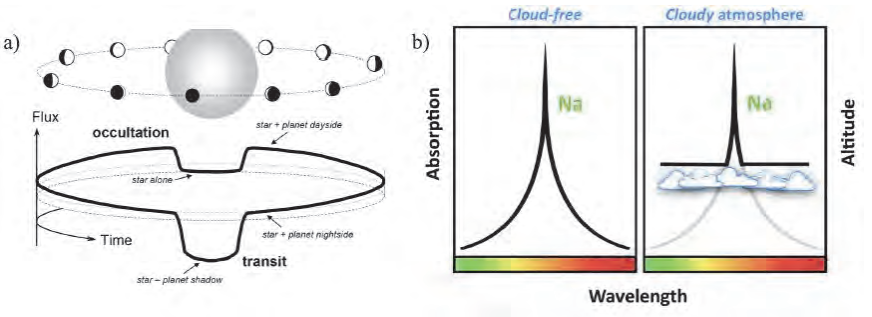



Fig. 11. *a) Exoplanets in orbits close to the line of sight for us on Earth periodically pass in front (transit) and behind (secondary eclipse) of their host stars. Transits and eclipses are a powerful indirect way to study the composition of exoplanet atmospheres*[12]*. b) Sodium fingerprint in an exoplanet spectrum. Shown is the absorption due to sodium at each wavelength. More absorption means that we are looking higher up in the atmosphere,and the vertical axis therefore a measure of altitude in the atmosphere of the planet. An atmosphere free of clouds produces an intact sodium fingerprint (left panel). A cloud deck blocks part of the sodium in the atmosphere, partially removing its spectral signature (right panel). Image credit: N. Nikolov/E. de Mooij.*

Using Europeès 8.2m Very Large Telescope in Chile, the team studied the atmosphere of WASP-96b when the planet passed in front of (transited) its host-star (see figure). This enabled the team to see the starlight shining through the planetary atmosphere, and so determine its composition. It has long been predicted that **sodium exists in the atmospheres of hot gas-giant exoplanets**, and in a cloud-free atmosphere it would produce spectra that are similar in shape to

the profile of a camping tent. Until now, sodium was revealed either as a very narrow peak or found to be completely missing. This is because the characteristic tent-shaped profile can only be produced deep in the atmosphere of the planet and for most planets clouds appear to get in the way. The sodium signature seen in WASP-96b suggests an atmosphere free of clouds. The observation allowed the team to measure how **abundant sodium is in the atmosphere of the planet, finding levels similar to those found in our own Solar System**.

Whether we are talking about exoplanets or our Earth, the role of the atmosphere is crucial in the study of the chemical composition of the sun. In fact the atmosphere of our Earth absorbs the light of the solar photosphere determining the so-called **Telluric lines**[13]. They are those of molecular oxygen and water, which could not exist in the solar atmosphere, much warmer. Furthermore, the atmosphere does not uniformly absorb solar radiation and could contribute to the 10% distortion from solar data.

In his book Il Sole[4] [pg. 235] Angelo Secchi writes *The gases that make up our atmosphere have a considerable absorption force and can therefore, by making some light rays disappear, give rise to some lines that are observed in the spectrum. Since this atmospheric absorption is greater than the layer crossed by light is thicker, the spectrum must vary with the position of the Sun: the lines must be less numerous when it is closer to the zenith, more abundant when it is near the horizon.* Angelo Secchi was referring to what we call today the **Air Mass**. The Air Mass (AM) is the path length which light takes through the atmosphere normalized to the shortest possible path length (that is, when the sun is directly overhead)[14]. The Air Mass quantifies the reduction in the power of light as it passes through the atmosphere and is absorbed by air and dust. The Air Mass is defined as:

$$AM = \frac{1}{\cos(\theta)} \tag{1}$$

where $\theta$ is the angle from the vertical (zenith angle)

The standard spectrum outside the Earth's atmosphere is called AM0, because at no stage does the light pass through the atmosphere.

Angelo Secchi writes in his book[4] [pg. 322] that **water vapor** absorbs solar radiation 100 times greater than atmospheric air.

The measures taken by Angelo Secchi for the Solar Spectrum and consequently for the Sodium D lines, were of course influenced by the Earth's atmosphere. This fact can be seen from the shape of the spectral lines. Many processes tend to broaden these lines so that the lines develop a characteristic shape or profile. When light penetrates the earth atmosphere, the AM0 spectrum is changed due to the following main mechanisms[16]:

- Absorption due to molecules and particles in the earth atmosphere.

- Scattering at molecules and particles in the atmosphere.

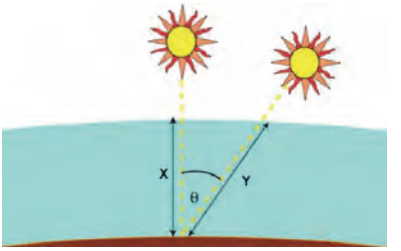



Fig. 12. *The air mass represents the proportion of atmosphere that the light must pass through before striking the Earth relative to its overhead path length, and is equal to Y/X. Copyright 2019 PVeducation.*

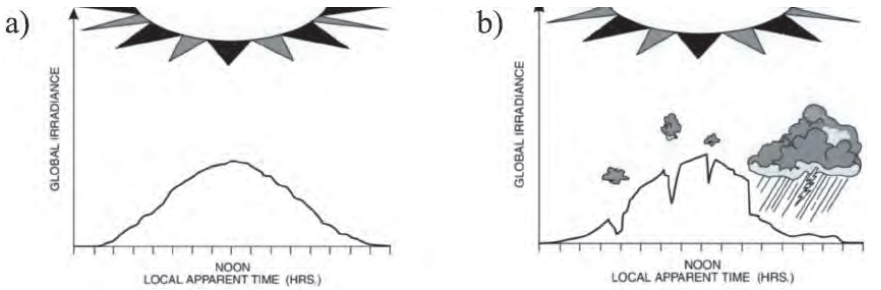



Fig. 13. *a) Diurnal variations of global solar radiative flux on a sunny day. b) Diurnal variations of global solar radiative flux on a cloudy day* ©*2019 Newport Corporation. All rights reserved.* [15]

- Reflection from the earth surface.

- The air mass the light goes through, depending on the elevation of the sun over the horizon.

The Fig. 14 shows the AM0 spectrum as the black line [17] and the terrestrial spectrum without the blue sky (red line) [18], and the transmittance of those atmospheric gases that are the main responsible for absorption [18].

In Fig. 15 we see how the Solar spectrum varies at different AM.

The first experiments made by Father Secchi were done in Rome using a prism with an angle of 10 degrees and 16 cm diameter, in front of a telescope 22 cm achromatic (lost in 1958 in the Monte Mario Observatory fire) [19].

Chromatic resolving power of a prism is given by:

$$R = \frac{\lambda}{\Delta\lambda} = \frac{BCdn(\lambda)}{d\lambda} \tag{2}$$

where $n$ is the refractive index of the prism for the wavelength $\lambda$ and it depends from $\lambda$, $\Delta\lambda$ is the limit of resolution, the difference in wavelength between two lines of equal intensity that can be distinguishedand BC is the maximum thickness of

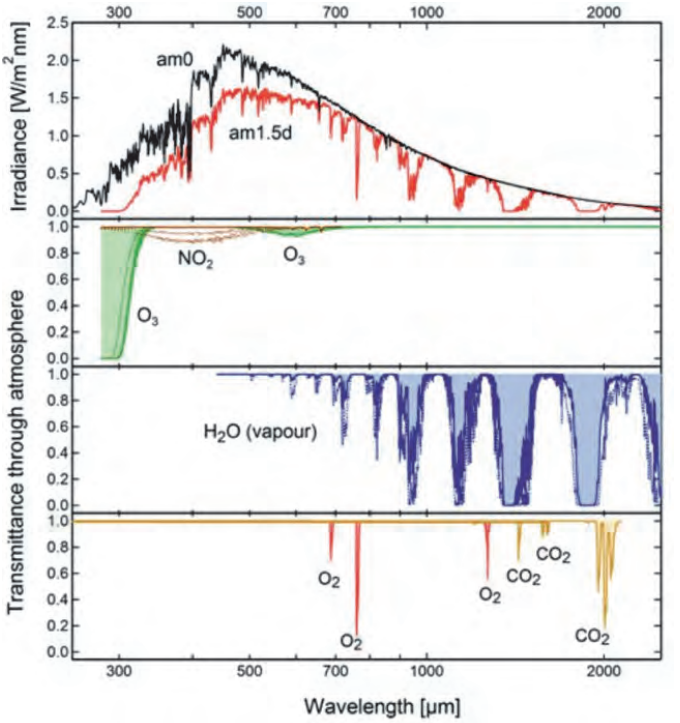



Fig. 14. *Sources of absorption in the AM1.5d spectrum.*[11]

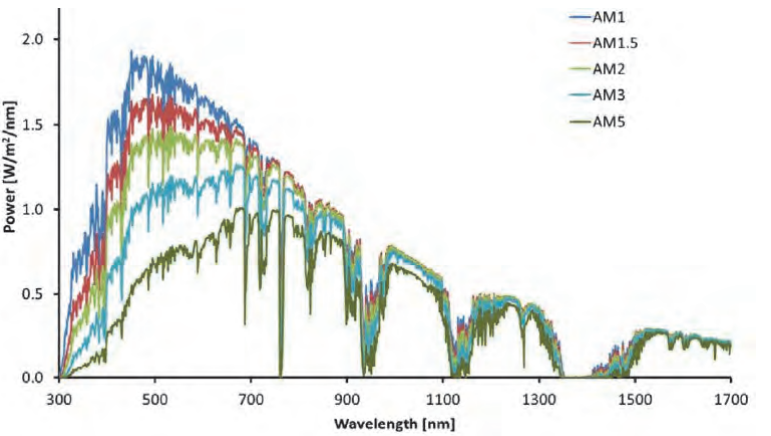



Fig. 15. *The direct incidence solar spectrum at different air mass values.*[11]

the prism traversed by light rays. The quantities $d(n)/d(\lambda)$ and BC often are called the dispersion and baselength of the prism, respectively.

Considering a different value of Air Mass during the day, supposing in the morning an Air Mass = 5, and the usage of seven or more prisms in series, as reported in Angelo Secchiś book[4] [pg. 222], the resolution obtained by Father Secchi was a very good result.

In fact for $n$=1.60 - 1.89 (flint glass), $\lambda(D1) = 589.418$ nm, $\lambda(D2)$ 588.821 nm, $\Delta\lambda = 0.597nm$, R = 986, which is the resolution we can obtain for resolving the the Sodium D-Lines doublet. But Angelo Secchi obtained more: using 9 prisms he was able to see what after will be called the Zeeman effect, discovered 30 years after (1898), more than 30 years after Angelo Secchi's obeservations.

**The Zeeman effect** is the splitting of a spectral line into several components in the presence of a static magnetic field. **Zeeman's original data was exactly on Sodium D lines**.

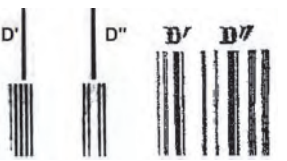



Fig. 16. *Zeeman effect on Sodium D-lines compared with Sodium D-Lines published in Secchiś book Il Sole pg. 223.*[9]

We can deduct that **Angelo Secchi was able to see the effects of the solar magnetic field**, which generates the sunspots. Father Secchi dedicate many pages in his book[4] to these aruguments (sunsposts pg. 8–20, solar magnetic field 335–343).

## 5. Alessandro Cacciani and Sodium Filter

Professor Alessandro Cacciani (1938–2007) has been known world wide for the invention and development of the Magneto-Optical Filter[20]. Hundreds publications present its applications from solar to stellar helioseismology. Astronomer at the Observatory of Roma (1964–1982), he continued his research at the University of Rome La Sapienza where directed the Solar Laboratory and taught Spectroscopy. The Magneto-Optical Filter (MOF) is an instrument that gives a really narrow bandwidth, high transmission (almost 50%) and perfect stability. It can work only in a small range of wavelengths, well defined, like **Sodium (Na)** and K doublets, and can be thought as an high resolution spectrograph within those wavelengths.

It was developed by Alessandro Cacciani at the end of the 1960s Its main use is for the study of the Sun; during the last years it has been adopted also for the analysis of Jupiter oscillations and in Anctartica for long-duration solar observations.

A complete MOF instrumentation is made of two separated unit: the MOF itself and the WS (Wing Selector) that, together, provide accurate Doppler and magnetic measurements. The first unit is composed by two crossed polarizers (P1, P2) and a

metallic vapour (Sodium) between them in a longitudinal magnetic field B between 2000 and 4000 Gauss. A schematic diagram of the filter is shown in the following figure.

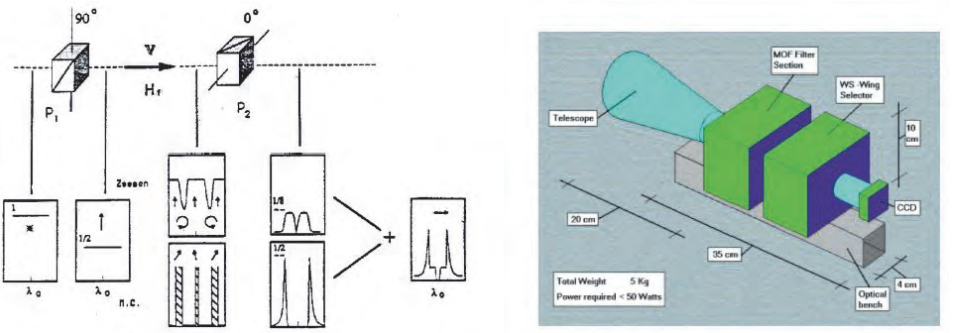



Fig. 17. *Schematic diagram of the MOF filter (top) and spectral behaviour along the optical path. Following the Zeeman rules, the vapour immersed in a magnetic field absorbs two circularly polarized components,leaving the residual light also circularly polarized (in opposite direction). This wavelengths cannot be stopped by the crossed polarizers and are transmitted as a couple narrow bands that can be close at will (depending on the magnetic field strength).*

The working principle of the MOF is based on two concurrent effects, namely the Zeeman effect in absorption and a sort of Faraday rotation close to the lines wings called Macaluso-Corbino effect. Both change the polarization in and around the resonance lines, leading to total transmission profile shown below where the experimental transmissions profile at different temperature and magnetic field values is shown[21].

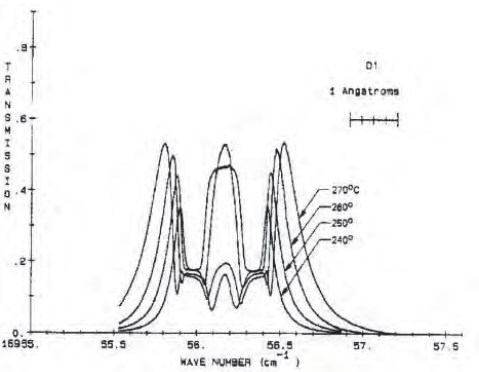



Fig. 18. *transmission profile at different temperature and magnetic field.*

Professor Cacciani develepoed the Sodium Filter even thanks to the "solar tower" present at the Rome Astronomical Observatory. **The invention allowed to "see" the solar surface**. Through the processing of digital images, taken at the

Monte Wilson Astronomical Observatory, it was possible to show to the world the filmed footage of the "flames" movement of the surface of the Sun, lasting 8 hours, summarized in 10 minutes.

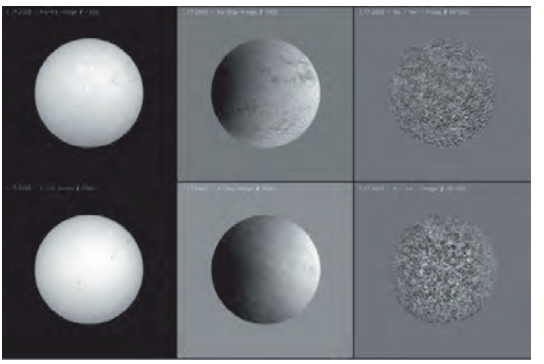


Fig. 19. *Intensity (left), Velocity (middle) and rotation suppressed Velocity (right) sample images in the Sodium Na line (above) and in the K line (below).*

It is possible to find many articles in astrophisical journals which refers to **Professor Caccianiš filter MOF**[22, 23, 24]. Helioseismolgy is a relatively new field of Solar Physics, whereby the internal structure and dynamics of the sun can be studied. This is a field of research that is being actively pursued both from the ground and from the space. Its aim is to derive information about the interior of the sun measuring the waves that travel inside the star and become resonant. From the frequency spectrum of such a resonant waves it is possible to compute the radial profile of relevant physical parameters (temperature, density and chemical composition) as well as its dynamics (internal rotation as a function of radius, large scale convective motions that mix the thermonuclear core etc). To monitor solar oscillations, a series of dopplergrams must be taken over a very long time[24]. At the 60 Foot Solar Tower (Univerisy of Southern California), **dopplergrams are obtained by using a magneto-optical-filter (MOF)**.

**A red/blue pair of sodium images** (see below) is taken once every minute with a separation of 5 seconds between the red and blue images. As a result, one dopplergram is computed once a minute for several hours during the day. The two dopplergrams can be subtracted to obtain a velocity image map of the line-of-sight components of the rising and sinking regions of the sun. A two minute separation is chosen since 2 minutes is roughly half of a solar oscillation. The images below show doppler velocity of the sun at a two minute separation and at a three minute separation, to the left and right respectively. Light regions are rising and dark regions are sinking.

So **a doppler image reveals those individual areas of the sun that are either rising or sinking**. In the images, rotation can be seen. The left hand side of the 2 images is rotating toward the Earth.

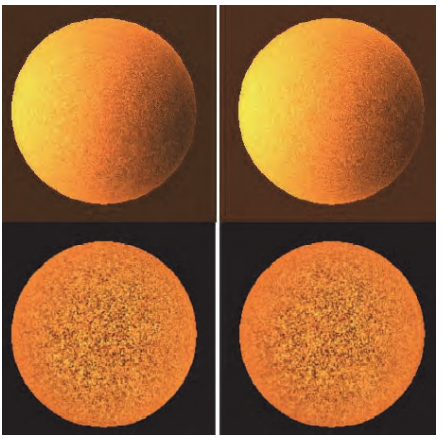


Fig. 20. *Dopplergrams.*

## References


1. http://astronomy.swin.edu.au/cosmos/S/Spectral+Line.
2. Capuzzo Dolcetta R. Alla Scoperta Dell'Astronomo Angelo Secchi, presentazione in occasione del bicentenario presso il Collegio Romano, Gennaio 2019.
3. Rota E. In memory of Father Angelo Secchi, Jesuit and astronomer, from Reggio Emilia, Quaderni di Storia della Fisica n. 21, 2019 - DOI 10.1393/qsf/i2019-10058-6.
4. Secchi A. Il Sole - esposizione delle principali scoperte moderne sopra la sua struttura, la sua influenza nell'Universo e le sue relazioni con gli altri corpi celesti 1884.
5. De Bernardis P. Corso di Laboratorio di Astrofisica, Laura Magistrale in Astronomia e Astrofisica, Sapienza University of Rome.
6. A. Altamore, M. Faccini, https://tracieloeterra.bicentenarioangelosecchi.it/gli strumenti/prisma-obiettivo/.
7. https://wasp-planets.net/tag/sodium/.
8. University of California, Irvine Department of Chemistry Natural Sciences II Irvine, CA 92697 https://www.chem.uci.edu/ unicorn/249/Handouts/ RWFSodium.pdf © 2001-2017.
9. Charbonneau, David; et al. (2002). "Detection of an Extrasolar Planet Atmosphere". The Astrophysical Journal. 568 (1): 377–384. arXiv:astro-ph/0111544. Bibcode:2002ApJ...568..377C. doi:10.1086/338770.
10. http://exoplanets.co/exoplanets-tutorial/extrasolar-planet-hd-209458-b.html.
11. N. Nikolov, et al., An absolute sodium abundance for a cloud-free hot Saturn exoplanet, May 2018.
12. G. Tinetti, Fisica degli Esopianeti, Cattedra Fermi 2019.
13. C. Sigismondi, Dalla Trasmittanza del vetro nell'Infrarosso allo Spettroscopio a Reticolo alla temperatura del Sole, 2019.

14. https://www.pveducation.org/pvcdrom/properties-of-sunlight/air-mass.
15. https://www.newport.com/t/introduction-to-solar-radiation.
16. Altermatt Lecture: The Solar Spectrum The influences of Earth's atmosphere on the solar spectrum; Solar constant and zero air mass solar spectral irradiance tables, http://www.astm.org, Standard E490-00a, The American Society for Testing and Materials, 2000.
17. C. Gueymard, The suns total and spectral irradiance for solar energy applications and solar radiation models, Solar Energy, pp. 423–453, 2004.
18. C. Gueymard, Parameterized transmittance model for direct beam and circumsolar spectra irradiance, Solar Energy 71(5), pp. 325–346, 2001.
19. R. Nesci Dispense Laboratorio Astrofisica, Sapienza University of Rome, 2008.
20. http://www.icra.it/solar/MOF/welcome.html.
21. A. Cacciani D. Ricci P. Rosati E. J. Rhodes E. SmithS. Tomczyk R. K. Ulrich, Solar magnetic fields measurements with a magneto-optical filter, Il Nuovo Cimento C January 1990, Volume 13, Issue 1, pp. 125–130.
22. D.Ricci, P.Rosati Helioseismology: Observational methods and requirements Article in Il Nuovo Cimento B105 (8): 1009–1020 August 1990.
23. David H. Hathaway, E. Rhodes, A. Cacciani, Sylvain G. Korzennik "The supergranulation spectrum" from book Conference on Challenges to Theories of the Structure of Moderate mass Stars (pp.163-169) January 2006.
24. University of Southern California, The 60th Foot Solar Tower http://physics1.usc.edu/solar/.
25. bass2000.obspm.fr/solar_spect.php.
26. http://web.usm.my/chem/pastyear/files/SpectrumSeries_Concept.pdf.

# Dark sky in Rome now and at the time of Angelo Secchi

Sigismondi, Costantino*

*ICRA/Sapienza and ITIS G. Ferraris, Roma, Italy
*E-mail: sigismondi@icra.it www.icra.it/gerbertus

Rome as every modern big city is heavily affected by the light pollution and the observation of the celestial objects is hampered by a luminous background. While a comparison with a rural place is made by observing the Orion nebula at the same heigth above the horizon and in the same meteorological conditions, an historical consideration is made by using the observations of years 1867–68 by Angelo Secchi of the same nebula, and upon the use of the meridian line of santa Maria degli Angeli for measuring the position of the Moon, and the instant of the full Moon used for the computation of the Easter's sunday in Catholic Church. The observation of the Moon on the meridian line nowadays on the meridian line is obstacoled by the external lights of Termini Station and Repubblica Square.

*Keywords*: History of Science, Light Pollution, Urban and Rural site for Astronomy, Pinhole Astronomy.

## 1. Introduction: The luminosity of the night sky in 1700 in Rome

The meridian line of santa Maria degli Angeli was build with the intention of measuring the variation of the obliquity of Earth's orbit for the forthcoming centuries in 1702. Another evident service of this line was to measure the tropical year, in order to check the Calendar's reformation of Gregory XIII. Finally on the meridian line are mentioned the two "Terminus Paschae" between 22 March and 25 April. Being Easter related to the full moon after or equal to 22 March (being 21 March the so called Easter equinox), the observation of the full Moon had to be possible on that meridian line. We tried to observe the Moon on the meridian line during the late evening, and we found it extremely difficult because of the illumination of the church from the external light coming from the great roman windows. The church in 1700s was completely darkened during the noon time for the observations of the Sun and the stars in daylight; during the night the absence of artificial light made unnecessary the use of tents for covering the great windows. So the astronomer Francesco Bianchini (1662–1729) who built this meridian line upon the will of pope Clement XI, did not write expecial notes on lunar observations, probably because they were rather simpler with respect to the observation of the stars in full daylight, as it was done in this meridian line.

## 2. Stars in daytime at Santa Maria degli Angeli

It has been written that Bianchini used very long telescopes in the basilica for observing the stars in daytime. It it true that Bianchini itself used such telescopes of 50 meters of focal lenght and single objective lens (made by Campani and purchased by the Cardinal Melchiorre De Polignac) for observing and mapping Venus from the Palatine Hill, but I exclude he did use the same instruments in the Basilica.

Not because of lack of space, but because they weren't necessary. A star does not show details like the planets, and to him only the position of the star was required, in order to compare the timing of the Solar transit with the one of the fixed star. He used for this purpose the coordinates of the stars reported in the atlas of Philippe de la Hire (1701) made in Paris observatory. The passage of a star on the meridian line can be done by defocusing the star with respect of a crosshair and averaging the first contact of the defocused disk and the second contact, exactly as it was done for the solar image on the floor of the basilica. So the draw of a small telescope used on the meridian line by Bianchini explains fully this procedure.

## 3. Moon in night time

Nothing is left about lunar observation on the floor of the basilica of S. Maria degli Angeli by Bianchini. But evidently he could use the telescope used for the stars in daytime, in the same exact way: he opened the window of about 40 cm wide x 60 cm high above the pinhole and could observe the Moon at the telescope. Our question now is: was it possible to observe the Moon in the same way of the Sun, projected on the floor? Was it useful for calibrating the Easter Computus of the spring full Moons?

## 4. Observed Full Moon and Ecclesiastic Moon

The Computus of Easter was based essentially on the Methonic cycle of 19 years. After this period of time the full Moon occurs in the same day of the year. 19 years contain always 5 bissextile years, unless we are in the surrounding of secular years not multiple of 400. So the problem of ecclesiastic Moon was only a mathematical problem for the astronomers who worked at the reformation upon the will of pope Gregory XIII, Christopher Clavius and Egnazio Danti. They used the mathematical solution of Luigi Giglio. From a physical point of view the solution of Methonic cycle is very stable, and there is the possibility that the lunar eclipses repeat over a Methonic series, as well as it occurs for the Saros cycles, for about thousand years. A lunar eclipse is the certitude that the Moon is full. It can occur that the days of such series changes of 1 or 2 days from a given eclipse, rightly for the aforementioned problem of secular years not multiple of 400. For example the year 2019 we had the lunar eclipse of 21 January, which follows the one of 21 January 2000, and the one of 21 January 1981. But there was also a Saros series dealing with the eclipse of 9 January 1982 and the one of 21 January 2000. So the eclipse of 21 January 2000 belong both to a Saros series and to a Methonic series. The Saros series of lunar eclipse are usually longer than thousand years, while the Methonic ones are slightly shorter. The calculated full Moon therefore is prevalent over the observed one in the ecclesiastic computus. The validation of this computus has been already done by the ancient astronomers.

## 5. Optics and energetics

The original circular pinhole was 20 mm wide and it projected the image of the Sun between 21 and 50 meters, and when the full Moon is the source, it is about 15 magnitudes fainter than the Sun. The windows of the church of S. Maria degli Angeli let enter nowadays a luminous background much higher than the sky background already much brighter than natural one. We can reconstruct some parameters by simplifying the situation as if the meridian line would be in an open space. The same considerations of this paragraph can be done by using a flat mirror and a pinhole, and projecting through the pinhole the lunar light over a screen at a given distance. For reproducing the original pinhole of st. Maria degli Angeli meridian line it is necessary to use a 2 cm wide circular pinhole and the screen from 21 to 50 m. To scale to shorter dimensions like 10 m, representative of a similar situation in S. Maria degli Angeli during spring, i.e. 30 meters of focal lenght for the spring Moon (due to its height above the horizon near the celestial equator... 27.5 m) we should use a smaller pinhole. Focal lenght in scale 10 m; focal length in truth 30 m. Energy per unit area is $E/30^2$ and has to remain constant scaling to 10 m, so the input energy has to diminish to $E'/10^2 = E/30^2$; $E' = E/9$, therefore the pinhole has to be 3 times smaller, i.e. 7.3 mm. To see the Moon projected by a 7 mm pinhole at 10 m of distance it is necessary to have the screen in full darkness. The magnitude of such pinhole lunar image on a 10% white reflecting paper screen or 20% reflecting white marble (the one of the meridian line) screen is 3 or 2.5 magnitudes dimmer than the ideal image at 10 m, which is 10 cm wide and it should contain all the energy fallen upon 7mm. The ratio between these areas is $(7.3/100)^2$ and the scaling in magnitude is 2.5 its logarithm (according to Pogson's law) i.e. 5.7. Finally the full Moon of magnitude $-13.6$ becomes $-13$ for the height above the horizon (atmospheric loss for transparency) and $-7.4$ for the projections and around $-4$ for the reflection efficiency. Now the problem of the sky background is that is integrated energy attains magnitude $-4$ and more! If we consider the brightness of the image of total magnitude $-4$ in term of magnitudes per arcsec$^2$ we obtain for the area of the Moon on the floor, of about 160 mm of radius, 8.26 magnitudes per mm$^2$ or 3.26 magnitudes per cm$^2$. Considering that this diffused luminosity for a single square centimeter is of the same order of the integrated magnitude of the Orion nebula, if the sky background does not allow to see the Orion nebula to the naked eye... it is difficult to see such projected Moon as well. A photo of the Moon on the floor of the Basilica required a long pose to be caught.

## 6. Sky background levels

In a rural site, the value of +22 magnitude/squared arcsecond is attained. The integrated energy of such sky over 2 pi of solid angle, or $2.7 * 10^{11}$ arcsec$^2$ is 11.4 magnitudes more than the single pixel of a single square arcsec of 22 mag, i.e. 10.5 magnitude. Evaluating the integrated energy equal to $-4$ means to have a single

pixel about 15 magnitudes more bright than the ideal rural site. This is not the case of the city of Rome, and probably of none of the urban sites worldwide, but in our case we have to remember that many of the lights around the basilica enters through the windows directly in the church letting the ambient rather luminous instead of dark.

## 7. Comparison with the Orion nebula observations

Rome was ending to be an *astronomical rural place* when in 1851 an English Company for gas illumination was charged to start the night illumination by the pope Pius XI. Father Angelo Secchi was the director of Collegio Romano Observatory at that time and published in 1868 a memoir on the observations and spectra of the Orion Nebula M42. The details he could see with a 24 cm refracting telescope with f=4.1m are astonishing for the astronomers of nowadays, they are indeed invisible for modern observers with equivalent and bigger telescopes without narrow bandpass filters. And even with such filters the rapid spread of bright LED illuminations are growing and polluting also the wavelengths which were left free of pollution by the gas lines. Using Newtonian telescopes f/4 of 25 cm of diameter (of the Urania firm, made in Rome and a brand new one made in China) only the greenish color of the nebula is visible (see Fig. 5), but not the details showed by Secchi. Rome passed from being a rural place (astronomically speaking) in 1851 to a present urban place. The angular dimensions of M42 observed through a 18 mm spotting scope in Rome and in Vigna di Valle (on Bracciano lake 60 km North from Rome) show the loss in terms of background magnitude. The observations of M42 have been made on April 5 at 20:42 UTC in Vigna di Valle $h = 9^o$ i.e. 4 airmasses and on April 7 at 19:10 UTC in Roma, Monteverde SW from the center of Rome $h = 23^o$

Fig. 1. Sequence of photos of the Moon on the meridian line of S. Maria degli Angeli. Pietro Oliva, 19 december 2007, from left to right. The luminous background determines the illumination of the rest of the meridian line.

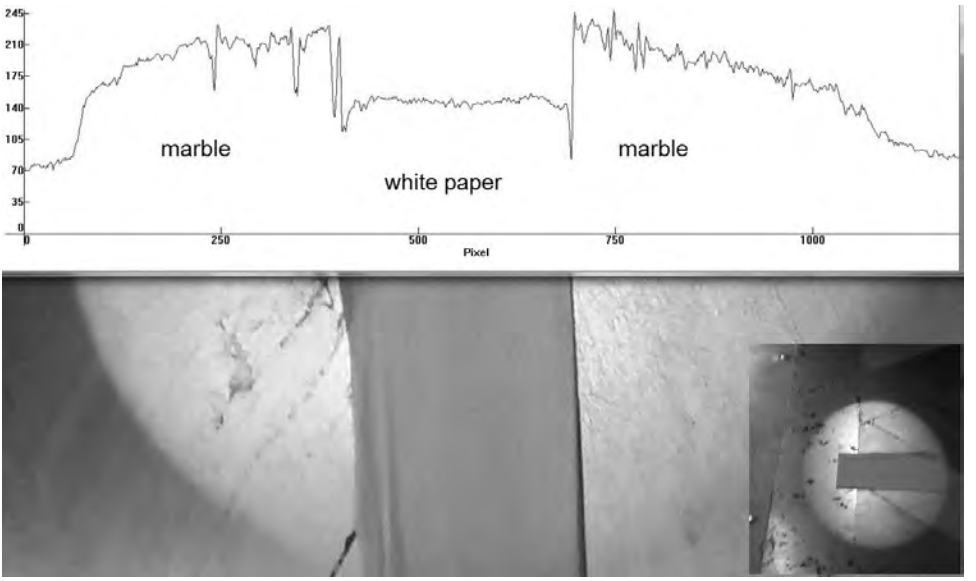



Fig. 2. The solar image at s. Maria degli Angeli shows the different reflective index of marbles and white paper. The histogram is obtained with IRIS software. While the white paper shows about 140 counts, the marbles get 245 counts on the same photo and in the most luminous regions, where the rays are reflected at the same incidence angle.

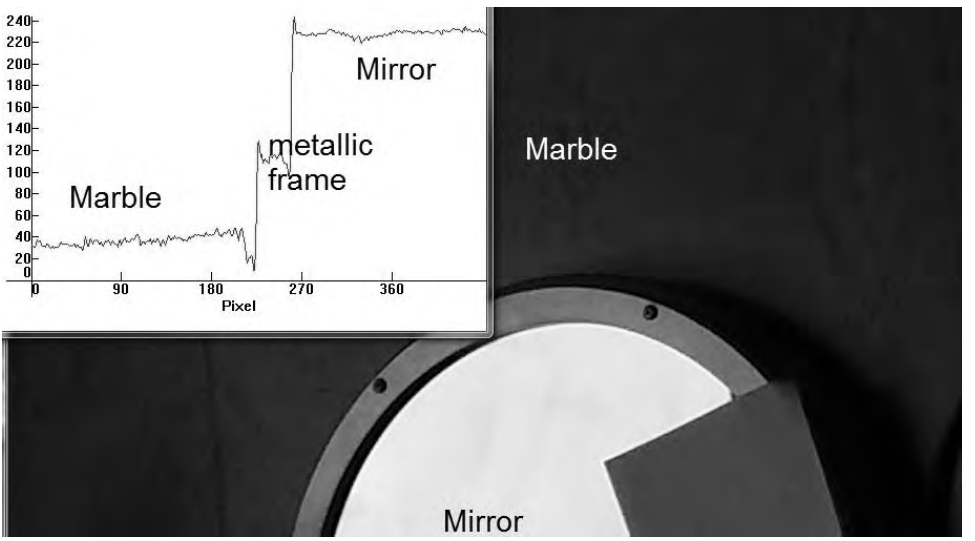



Fig. 3. The reflection of the marble compared with a mirror with 99.999% of reflectance, a glass solar filter of the Science Museum of Virginia. The marble reflects 40 units while the mirror 230. The reflectance of this marble is about 17%. We assumed 20% for the one of S. Maria degli Angeli.

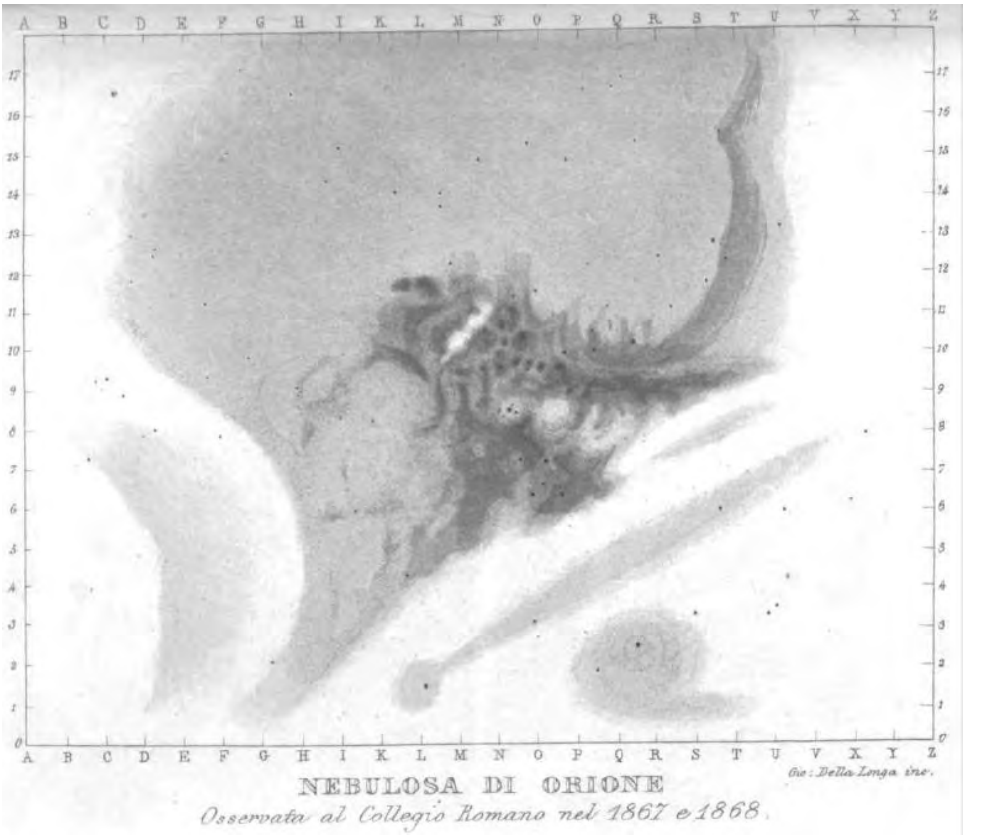



Fig. 4. The Orion nebula as seen by father Angelo Secchi in 1867–68.

i.e. 2.5 airmasses, and 20:27 UTC in Roma, Prati NW of the center of Rome, at $h = 10^o$. These observations showed the following results: in the rural place at same height of Rome the angular dimension of M42 is twice. The dimension of M42 is 1.5 smaller in Rome the rural place when compared the observations at $9^o$ and at $23^o$ of altitude. The passage from a rural place to Rome SW is the equivalent of 1.5 airmasses of absorption. Being 0.235 magnitudes for each airmass in Rome, the lower limit of the loss due to the ligth pollution is at least 0.35 magnitudes.

## 8. Conclusions

We studied two situations dealing with light pollution: the classical observation of a deep sky object like the Orion Nebula M42 in Rome now and in 1867–68 by father Angelo Secchi at Collegio Romano (center of Rome) and the observation of

Fig. 5. The Orion nebula photographed by Daniele Nocco and Emanuele D'Aleo in 2019 with Canon EOS 600d ISO 3200 exposition 10s, tube of 1.5 m, from Rome, Casaletto.

the Moon projected at 30 meters of distance on the meridian line of Santa Maria degli Angeli through a pinhole of 20 mm by Francesco Bianchini in 1702 and the similar experiment at 10 m through a 7mm pinhole. The difficulty to observe the projected Moon is due to the integrated magnitude of the sky of $-4$, 15 magnitude brighter than the rural site, this value included the direct lights entering through the windows in the church during the night. The Orion nebula observed from Rome and from a rural place, Vigna di Valle, showed a reduction of perceived angular dimensions from Rome of about 50% which would correspond to 0.75 magnitudes. A similar lower limit of 0.35 magnitudes is obtained by comparing an observation made in Rome at $h = 23^o$ with the one at $h = 9^o$ in Vigna di Valle. The site of Rome, when looking offcenter is not too bad with respect to a rural environment like Vigna di Valle located 60 Km from the city, even if many details have been lost when comparing with the observations made by Father Secchi on the Orion nebula.

## References

1. Sigismondi, C., Lo Gnomone Clementino, Roma (2009).
2. Secchi, A., Sulla Grande Nebulosa di $\Theta'$ Orione, Firenze, Stamperia Reale (1868).
3. Buil, C., IRIS astronomical software (2010).
4. UNESCO, et al., http://www.darkskiesawareness.org/defense-declaration.php (2007).

# The Secchi-Rosa law on the solar diameter

Sigismondi, Costantino*

*ICRA/Sapienza and ITIS G. Ferraris, Roma, Italy
*E-mail: sigismondi@icra.it www.icra.it/gerbertus

As Secchi-Rosa law has been named the observed variation of the solar diameter with the solar cycle. It was considered a consequence of the solar activity on the outer layer of the stellar structure, and there was a discussion on the irradiation effect due to the Earth's atmosphere. A comparison with the penumbra effect measured on the meridian line of Santa Maria degli Angeli and its debate since Cassini and Manfredi is made, and the definition of the solar limb as inflexion point are discussed. A relationship between position on the meridian line and extension of the penumbra is proposed from the analyzed data.

*Keywords*: History of Science, Solar diameter, Angelo Secchi, Rosa.

## 1. Secchi-Rosa law as reported by Giorgio Abetti (1937)

The solar physicits giorgio Abetti in his book "il Sole" (1937, pp. 77–80) reported the Secchi-Rosa effect. The original text in Italian is here fully reported and after commented.

### La forma del Sole e la sua supposta variabilità

Già nel 1750 Bradley eseguiva misure sistematiche del diametro solare a varie latitudini e, a parte il problema di una sua eventuale diminuzione secolare, ben presto si incominciò a discutere sulla sfericità del globo solare e della variabilità del suo o dei suoi diametri, forse in relazione con le varie fasi di attività solare. Le misure del diametro sono state cominciate, e si fanno tuttora, con i cerchi meridiani, determinando a mezzo dei tempi dei passaggi in meridiano dei lembi ovest e est, il diametro orizzontale e, dalla differenza di declinazione fra i lembi Nord e sud, quello verticale. Il diametro viene inoltre determinato a mezzo degli eliometri o con la fotografia. Il P. Angelo Secchi e il P. Paolo Rosa, dell'Osservatorio del Collegio Romano, hanno discusso il problema delle variazioni del diametro solare dalle osservazioni allora esistenti, ed il secondo di questi, in un lavoro dettagliato, concludeva che in quelle epoche in cui il numero delle macchie e delle protuberanze è minore, il diametro equatoriale è maggiore. Similmente R. Wolf, dalla discussione delle osservazioni eseguite da Hilfiker col cerchio meridiano, trova, in corrispondenza del grande massimo del 1870, un risultato corrispondente a quella che egli chiama legge Secchi-Rosa. Anche all'Osservatorio del Campidoglio, dal 1877 si continuano ad eseguire sistematiche osservazioni con il passaggio al meridiano del diametro solare le quali dimostrano la sua variabilità. Ma è difficile stabilire in qual grado essa sia propria del Sole o dovuta ad altre cause. Intanto il problema è probabilmente complicato dal che i due diametri, polare ed equatoriale, presentano una lieve differenza, pur essa variabile, in relazione con il ciclo di attività. Secondo le ricerche di Lane Poor,

su osservazioni fatte da varie investigatori con diversi metodi, il rapporto fra i diametri polare ed equatoriale sarebbe periodicamente variabile. La lunghezza del periodo resta incerta, ma sembra essere approssimativemente quella del ciclo delle macchie. L'amplitudine di questa variazione è circa 0.2", la differenza fra i massimi valori positivi e negativi essendo di circa 0.5". A Greenwich, dove si osservano regolarmente col cerchio meridiano i diametri del Sole, dalla discussione del decennio 1915–1925 risulta, che il diametro verticale è maggiore di quello orizzontale di 0.3", differenza che, almeno in parte, può essere dovuta al differente metodo di osservazione, il quale, come già detto, per il diametro orizzontale consiste nel misurare i tempi del passaggio al meridiano, con un micrometro autoregistratore, per quello verticale nel misurare con i cerchi la differenza di distanza zenitale fra il lembo Nord e quello Sud. Un esame dei diametri osservati, in realzione col periodo delle macchie, ha condotto a risultati incerti; invece, ordinando per ogni mese durante il decennio, per ambedue i diametri, le differenze fra i diametri calcolati nel *Nautical Almanac* e quelli osservati, risultano delle variazioni massime di un secondo d'arco, essendo il diametro maggiore in estate e minore in inverno. Discutendo le varie cause, che possono produrre un simile effetto, si giunge alla conclusione che esso sia dovuta alla **irradiazione** e alla sua variazione nel corso dell'anno. Infatti, se si fa l'ipotesi che l'irradiazione sia variabile al variare dell'assorbimento atmosferico, secondo la diversa altezza del Sole sull'orizzonte, ed anche per variazioni di contrasto fra il disco luminoso del Sole, e lo sfondo del cielo, Cullen, il quale ha discusso le osservazioni di Grenwich, trova che le variazioni del diametro sono proporzionale alla secante della distanza zenitale del Sole, raggiungendo appunto un secondo d'arco. D'altra parte l'effetto totale della irradiazione sul diametro del Sole, che è in media di 1920", sembra sia di 3". Se ne deduce che le variazioni annue del diametro solare, dovute alla irradiazione, sono piuttosto notevoli, e quindi non deve far meraviglia che risulti assai difficile distinguere, da questa variazione apparente, una eventuale variazione effettiva del globo solare. La questione è lungi dall'essere risolta: sembra che sia necessario escogitare dei metodi di misura più precisi, ed eseguire le osservazioni in luoghi dove le variazioni dovute a cause terrestri siano le minori possibili. Che possa esistere una differenza fra il diametro polare e quello equatoriale del Sole, e che questa sia variabile, potrebbe venire confermato dal fatto che sembrano esistere delle variazioni nell'inviluppo esterno del Sole, chiamato "cromosfera", e dal fatto che hanno luogo periociamente notevolissime variazione nell'altro inviluppo più esterno, cioè la "corona solare" a noi visibile soltanto durante le eclissi totali. Era già stato notato da Respighi che la cromosfera ha altezza variabile alle diverse latitudini; dalle regolari osservazioni eseguite in Arcetri durante l'ultimo ciclo sembra resultare, che l'altezza della cromosfera è maggiori ai poli che all'equatore alle epoche di minima attività solare, ed ha invece approssimativamente la stessa altezza attorno a tutto il bordo durante le epoche di massima attività. La corona solare poi presenta, come vedremo, delle configurazioni ben definite con espansione della materia coronale ora all'equatore ed ora ai poli, in corrispondenza ai massimi e minimi dell'attività solare.

## 2. Comments on the text of Abetti and penumbra problem in pinhole gnomon

The text reports a full account of the debate on the Secchi-Rosa law, as named by R. Wolf, according to wich the solar diameter would vary along the solar cycle, with the maximum occurring at minimum solar activity, with typical amplitudes of 0.5". The reports made by Cullen on the measurements held at the Greenwich Observatory evidence also a monthly variation of the solar diameter, both vertical and horizontal, lead to the introduction of the "irradiation" effect, responsible of the different perceptions of the solar diameter because of the contrast of the solar disk with the sky. This irradiation effect would amount up to 3" for the whole mean apparent diameter of 1920". The contrast with the sky background is due to the amount of humidity in the air and to the Mie scattering of the atmosphere; this produces as well a variation on the measured penumbra in the pinhole meridian line as the Clementine Gnomon in Santa Maria degli Angeli in Rome, an effect already discussed by Cassini, Montanari and Bianchini in the XVII century. The phrases reported by Catamo and Lucarini (2012) who rediscovered the inscription on the penumbra algorithm on the "equinox meters" of the Clementine Gnomon (1702) explains the fact: Bianchini prendeva sempre in considerazione, per le sue rilevazioni, l'ellisse luminosa diminuita dell'*anello* di minore luminosità (la penombra). La larghezza di questo anello era stimata, con un criterio all'epoca condiviso da Gian Domenico Cassini e da Eustachio Manfredi, altro importante astronomo coevo, pari al semidiametro del foro gnomonico, quindi nel caso di Santa Maria degli Angeli, pari a 50 particelle, cioè un centimetro. Altri hanno ritenuto che l'anello è di misura variabile in relazione a diverse circostanze, tra cui quelle meteorologiche, e può raggiungere e anche superare, nel caso di un foro come quello di Santa Maria degli Angeli, la larghezza di due centimetri.

## 3. Penumbra measurements at the Clementine Gnomon

During the meridian transit of 21 August 2019 the phenomenon lasted 141.11± 0.58 s at a declination of +12d 07m 29.4s, corresponding to an angular extent of 2069" ± 9". The real diameter of the Sun was 1897"; the penumbral ring was 86.1" wide. The pinhole is presently 25.4 mm wide and at the distance of 57.201 centesimal parts on the meridian line corresponds to an angle of 223.4", therefore the penumbra was the 38.6% of the pinhole in these conditions of clear sky (with increased humidity with respect to 18 august) and illumination in the church. The timing was obtained by video inspection. Ephemerides from calsky.com. Air humidity from meteoblu.com archive: for 21 August at noon time it was 55%.

From a pure geometrical consideration the penumbra extension should be a ring of a thickness equal to the diameter of the pinhole, but the contrast inside the church is different if we consider the winter image, with respect to the summer image, which is more contrasted, falling in a darker place of the basilica. See Fig. 1 for the case of 18 August 2019 with penumbra ring as 41.1% of the pinhole diameter.


Fig. 1. *The Sun approaching the meridian line of the Clementine Gnomon on August 18, 2019. Penumbra 41% of pinhole diameter and air humidity 60%.*

The transit lasted 142.46± 0.65 s at a declination of +13d 06m 26.6s, corresponding to an angular extent of 2081" ± 10". The real diameter of the Sun was 1896"; the penumbral ring was 92.7" wide. The pinhole is presently 25.4 mm wide and at the distance of 54.948 centesimal parts on the meridian line corresponds to an angle of 225.6", therefore the penumbra was the 41% of the pinhole in these conditions of sky clear and illumination in the church. The timing was obtained by video inspection. Ephemerides from calsky.com. The humidity of the air was 60%. Moreover the sky conditions has its influence on the penumbra: a blue sky is less luminous than a white one, and the contrast of the solar image is larger; a wider penumbra is to expect, while in the case of the white sky the contrast is less and the sky is more luminous, so its light penetrates into the church through the large windows lowering the contrast of the observerd image with respect to the floor. In the first years of use of the meridian line the basilica's windows were closed eliminating the problem of the contrast inside the church, leaving only to the sky contrast the cause of the "irradiation" effect.

## 4. Inflexion point of the ligth curve

It has been evindenced that for the limb darkening function the horizontal and the vertical luminosity profiles of the Sun have their maximum at the center, and decrease with a typical slope toward the limbs, at which the luminosity if 16% of the center. This curve has an inflexion point slightly before the end. This inflexion point is considered stable, even if the conditions of background contrast vary. This concept is applied either in modern solar observations with appropriate detectors (digital) either in total eclipse images obtained with different density filters; the inflexion point has not to be cut out with too high density filter.

Fig. 2. *The Sun on the XX August MDCCII inscription, on 21 August 2019. Its penumbra is 38.6% of the pinhole width, instead of 100% geometrical. Air humidity 55%.*

## 5. Conclusions

The problem of the definition of the solar diameter in the XVIII and XIX centuries was heavily influenced by the luminosity of the sky background. Either in the pinhole meridian line of S. Maria degli Angeli in Rome or in S. Petronio in Bologna, for quoting the main instruments devoted to the measurements of the

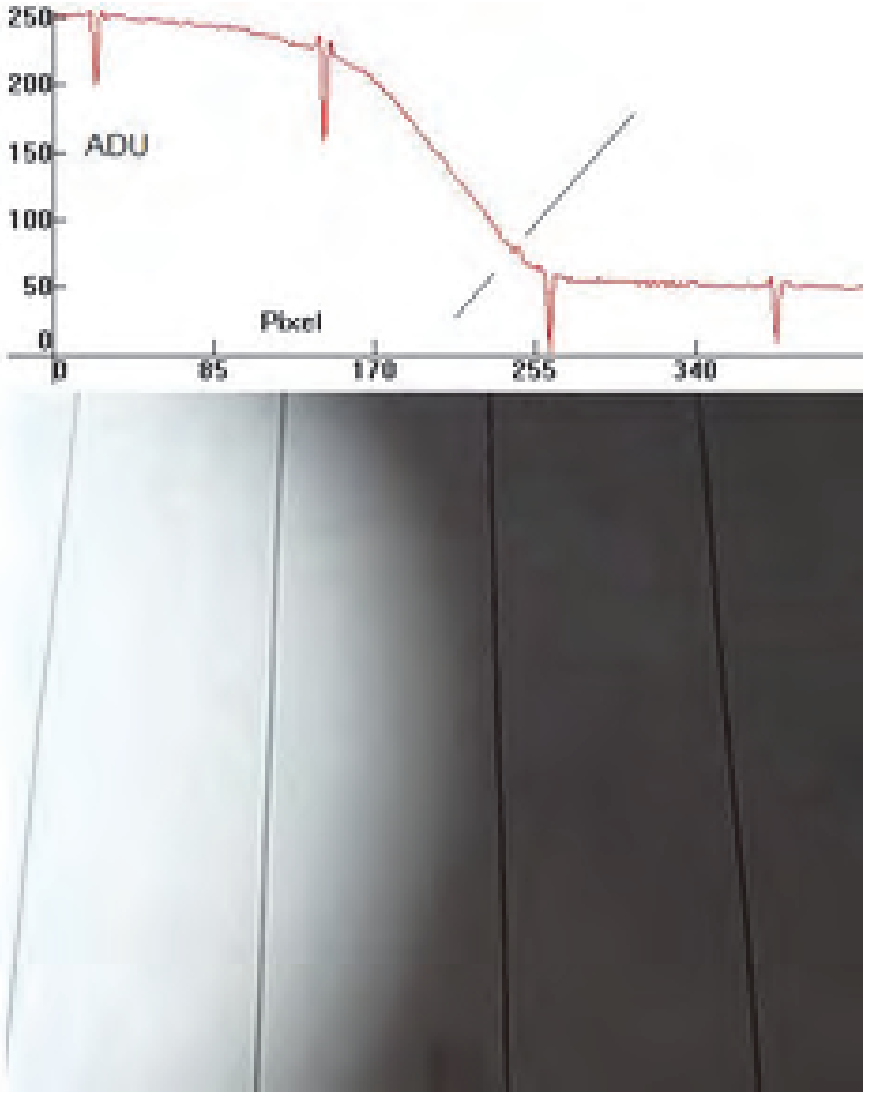



Fig. 3. *The inflexion point of the Limb Darkening Function of the Sun combined with the 25.4 mm Pinhole transfer function at the floor of Santa Maria degli Angeli, image of 13 July 2019.*

solar position and of the solar diameter, the dependance on the weather conditions is rather important. Two measurements devoted exactly to this purpose made on 18 and 21 August 2019 at the meridian line of the Basilica of Santa Maria degli Angeli proved a variation of the penumbra extent from 41% to 38.6% of the pinhole diameter, instead of being 100% as the geometry implies. The illuminination inside the Basilica was rather the same, being the Sun in a position only slightly different (less than 30 cm); but the conditions of humidity and color of the sky, which

was clear in both days, changed from 60% to 55%. This produced a reduction in the observed penumbra from 41%to 38.6%, suggesting a relationship between percentual humidity and percentual Penumbra. After these first measurements, other video have been inspected: 7 Dec 2018, 12 Jan 2019, 1 May 2019 and 13 Jul 2019. Other relationships between Penumbra and Temperature or air Pressure have been investigated; none of them are convincing, excepted the one with the position. The conclusion is that the Penumbra depends strongly on the contrast conditions inside the Basilica, more than from sky conditions. To observe the influences of the sky conditions it should be necessary obscure the church like Francesco Bianchini did in XVIII century. The assumption made by Francesco Bianchini that the Penumbra was half of the pinhole, say 50% of it, has been verified as optimal with respect to our videos with time resolution down to 0.26 s, and angular resolution down to 3 arcsec.

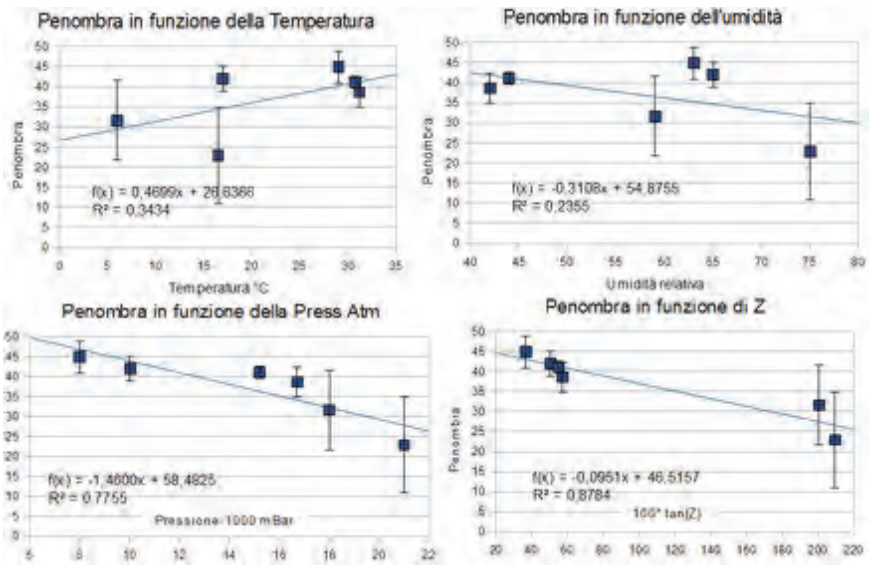



Fig. 4. *The relationships between Penumbra and air humidity (at 2 m), air temperature (at 2 m), air pressure (at sea level) and position on the meridian line; the best correlation is the last one, with z, zenital angle.*

The seasonal effect invoked by Secchi and Abetti to explane the seasonal variation of the observed solar diameter could be of the same type devised here, through the measurements made in Santa Maria degli Angeli and here presented and discussed. The inflexion points of the intensity profile along a solar diameter define in a more stable way the limbs of the Sun, but this was possible only in XX century with the advent of photography and digitalization. For visual inspection the inflexion point is missed, because the Penumbra varies from about 30% to 45% from Winter to Summer, but it would provide an evaluation very similar to the

50% chosen by Bianchini, who could work with solar images produced in the fully darkened basilica, by using tents outside the large windows.

**Acknowledgements**

To Irene Salmaso, Osservatorio Astrofisico di Asiago and University of Padova, for her bibliographic research.

**References**

1. G. Abetti, Il Sole, pp. 77–80 (1937).
2. M. Catamo and C. Lucarini, Il Cielo in Basilica, Roma (2012).
3. E. Manfredi, De Gnomone Bononiensi, Bologna (1736) pp. 8–9.
4. J. L. Heilbron, The Sun in the Church, Harvard (1999) pp. 109–112.
5. L. Ximenes, Del Vecchio e nuovo Gnomone Fiorentino, Firenze (1757). Scanned on https://books.google.it/books/?id=PM7v4pPvZeAC.
6. P. Rosa, Studii intorno al diametro solare, Roma, Befani (1873). Scanned on http://books.google.com/books?vid=IBCR:BC000032664.
7. F. Bianchini, De Nummo et Gnomone Clementino, Roma (1703). Scanned on http://books.google.com/books?vid=IBNR:CR000685435
8. C. Sigismondi, Meridian Transit of 18 August 2019. https://www.youtube.com/watch?v=eRzIQB2X5xE.
9. C. Sigismondi, Meridian Transit of 21 August 2019. https://youtu.be/q9cy9Btw3c.
10. C. Sigismondi, The Sun on the XX August MDCCII inscription https://youtu.be/bKYpOuKIrhI.
11. https://www.meteoblue.com/it/tempo/historyclimate/weatherarchive/roma_italia_3169070.
12. A. Raponi, C. Sigismondi, K. Guhl, et al. Solar Physics, 278, 269 (2011).
13. https://www.timeanddate.com/weather/italy/rome/historic?month=7&year=2019.
14. C. Sigismondi, All Meridian Transit in S. Maria degli Angeli (2018–2019). https://docs.google.com/document/d/1i63OT0jYgSvvfsD-hXAG_U2VinAlkE_ZzwN7LeduHDE/edit?usp=sharing.

# Agro-pastoral astronomical algorithms for seasonal feasts

Giannini, Francesco (1) and Sigismondi, Costantino*

*(1) FEMI srl, Chieti, Italy*

**ICRA/Sapienza University of Rome, E-mail: sigismondi@icra.it www.icra.it/solar*

The research started from the birthdate of Angelo Secchi: June, 28, that is the vigil of St. Peter and Paul in Roman Catholic Church. In this congress Matteo Galaverni assessed the baptismal names of Angelo Secchi, clarifying that the pointed letter "P." stands for Father, in Italian "Padre". So there is no connection between the names of Angelo Secchi, with the Prince of the Apostoles, but the question about the storicity of the feast of the two Apostoles dead both in Rome under the prosecution of Nero, remains. An archaic algorithm to fix relevant days of the year is still connected with some important dates of the present globalized World. These are ±40 days from solstices and equinoxes. The action of the Catholic Church to Christianize some of these feasts is outlined, but the conclusion is that the 29 of June is not connected with such algorithm and procedure, having reliably and historical root.

*Keywords*: Angelo Secchi, Astronomy, Gnomonics, Meridian line, Santa Maria degli Angeli, San Pietro, Equinoxes, Solstices, Zodiacal Signs, Play Cards.

## 1. The birthday of Secchi: A case study

Angelo Secchi was born in 1818 on June 28, the vigil of St. Peter and Paul. The "P." found in the internet before his name was interpreted not only by wikipedia, but also by D. Menzel in the book "Our Sun" and reliably by many other authors as his first name, but Matteo Galaverni demostrated that it was simply "Padre" i.e. Father (a Jesuit one in his case). The starting question for us was concerning this feast and its date: was it original or the result of an adaptation made by the Catholic Church to substitue some important dates coming from the ancient agro-pastoral culture.

## 2. Agro-pastoral Algorithms for feast and annual rithms

The transhumance, the periodical migration of the shepherds with their flocks, between the mountains of Abruzzo and the flat regions of the Apulia in Italy, required a knowledge of the annual solar rithms. To the equinoxes and the solstices were associated 40 days before and after some relevant yearly events for this society: occasions of gathering and celebrating together some feasts. The reckoning of such days was done by counting, and 40 was an apotropaic number obtained by the sum of the fingers of hand, feet of the man and of the woman. This is one reason why the Neapolitan play cards are 40.

The present feasts so individuated are 1st May (labor day) 31 October (All Saints Eve, Hallowe'en) 2 February (Candlemas) 15 August (August holiday, Assumption of Virgin Mary to the Heaven) 2 August (Santa Maria degli Angeli) 11 November (St. Martin, end of the ancient fiscal year). The 11 February is Our Lady of Lourdes, but it is only after 1858, when the Virgin Mary appeared in Lourdes at

the base of French Pyrenees. The 11–12 of May also are no relevant feast, even if we found the death date of two popes Sylvester II and Sergius IV respectively in May, 12 1003 and 1012, whose inscriptions are preserved in the Cathedral of Rome, St. John the Lateran; Sylvester II was an astronomer.

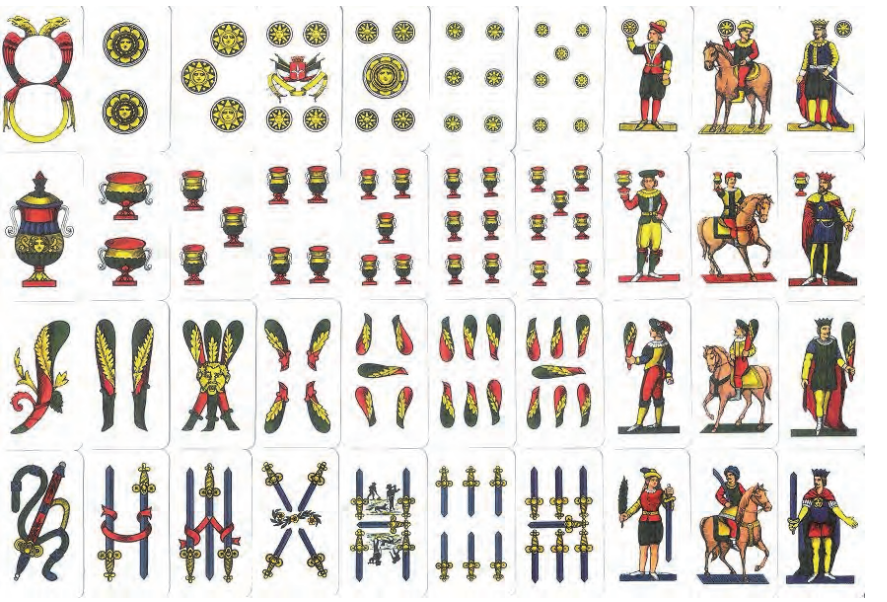


Fig. 1. The 40 Neapolitan Cards.

There are 4 card types (seeds) related to the seasons: cups for wine (autumn), money for summer, sword for spring (the time for was champaings), club for winter (needed for fire). The meaning of Kings, Knight and Soldier for each card type are related to the beginning days of the seasons: equinoxes and solstices. The numbers from 1 to 7 are the weekdays.

## 3. French cards and lunar months

Alternative interpretations are in the French cards with lunar months: 13 cards for each seed, as the 13 lunar months in the year. 52 cards and 1 jolly.

## 4. Zodiacal signs and annual rithms

Conventionally the Zodiacal Signs are referred to the geocentric annual orbit of the Sun around the Earth: since it is inclined of about 23° with respect to the celestial

equator we have the variation of the insulation and consequently the 4 seasons, corresponding to separation of 90° in this presumably circular orbit. The change in rising and setting points of the Sun is evident along the year for a give point of view: the extreme positions correspond to the solstices, and the mean position correspond to the equinoxes. The zodiacal signs represents a further division in 3 parts for each sector, slightly larger than a lunar month. But assuming precisely sectors of 30° implies the capability to compute the times of the year when these sectors are transpassed. And these dates and times are not constant from an year to the following because of the length of the tropical year approximately 365.2422 days, and because the obliquity of the Earth's orbit is changing from about 21° to 24° in 41000 years.

The Zodiacal signs have been used to describe the annual orbit of the Sun, and after Ptolemy's Almagest they were used for marking the great meridian lines built in the churches, namely St. Petronius in Bologna and Santa Maria degli Angeli in Rome.

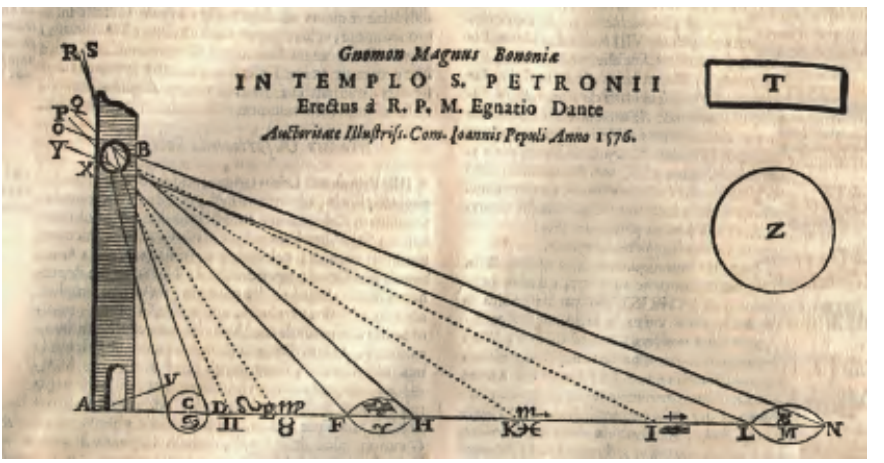



Fig. 2. The scheme of the Zodiacal Signs in St. Petronius's meridian line in Bologna, made by Egnazio Danti in 1577.

It happens that when the dates of the Zodiacal signs are rounded to an integer value, some incongruences appear. This is the case of the meridian line in St. Peter's square in Vatican, realized by Peter Maccaranius in 1817.

The dates are not compatible with the second law of Kepler and apsides on 4 July and 4 January. It is immediate by looking at them: the ingress in Acquarius should be on 20 January instead of 21.


Fig. 3. The dates of meridional Zodiacal Signs in St Peter's meridian line.


Fig. 4. The dates of septentrional Zodiacal Signs in St Peter's meridian line.


Fig. 5. The winter solstice in St Peter's meridian line.

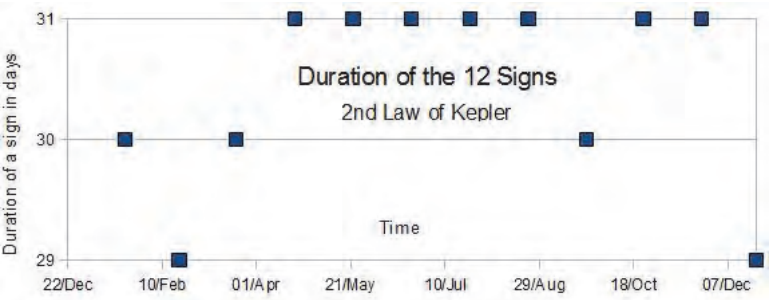



Fig. 6. The irregular duration of the signs in St Peter's meridian line.

## References

1. Morpurgo, Enrico, Gli Orologi, 7, Fratelli Fabbri Editori, Milano (1966).
2. Danti, Egnazio, Tractatio Usi Gnomoni Magni, Bologna (1577).

# Meridian Service in Rome at S. Maria degli Angeli with Francesco Bianchini and at Collegio Romano with Angelo Secchi

Baschetti, Beatrice (1) and Sigismondi, Costantino*

*(1) Sapienza University of Rome * ICRA/Sapienza University and ITIS G. Ferraris, Rome, Italy*

** E-mail: sigismondi@icra.it www.icra.it/solar*

The father of Italian geodesy is Francesco Bianchini (1662–1729), who built the meridian line in the Basilica of S. Maria degli Angeli in Rome with the purpose of measuring accurately the variation of the obliquity of Earth's orbit and the tropical year duration. He compared the observations of eclipses made in Rome and in Bologna, at the meridian line (1655) made by Giandomenico Cassini and found that the meridian of the Pontifical State was from Rome to Rimini. While Bianchini published many details of the meridian line in 1703, the presence of two decorations near the Summer solstice position has remained unexplained until 2018. Only one of them receives the image of the Sun nowadays and allows immediately to evaluate the secular shift of the solstice's position. The position of the red marble strip under this decoration is the materialization of the solstitial center of the solar image in 1702. For 150 years the noon signal was given by that meridian line; Angelo Secchi was able to automatize the procedure of the ball-drop, as in Greenwich observatory. A signal was spread to the city, through the fall of a sphere visible from far on top of the Flamsteed house, to give the instant of the local meridian transit. This was established in 1833 and automated in 1852 by George B. Airy. Secchi realized a similar device upon the roof of St. Ignatius, where his Observatory was located. This ball-drop gave the signal to a cannon located at Castel Sant'Angelo, the cannon was moved later on Gianicolo hill. Nowadays the tradition of the cannon continues, without the intervention of astronomers. Secchi measured carefully the meridian of Rome, and paved the way to the modern geodesy and the fundamental meridian of Italy at 12° 27' 08" from Greenwich on top of Monte Mario.

*Keywords*: Angelo Secchi, Astronomy, meridian line, Santa Maria degli Angeli.

## 1. The story of the meridian line in Santa Maria degli Angeli

The church of S. Maria degli Angeli is built inside a part of the old roman Baths of Diocletian in Rome, the present entrance is from Piazza della Repubblica. Michelangelo Buonarroti first adapted the old building to a church in 1561–64, further interventions were then made by Luigi Vanvitelli in 1749–50.

The meridian line of S. Maria degli Angeli was built by Francesco Bianchini between 1701 and 1702, upon the will of pope Clement XI who commissioned the work. This device was built for multiple purposes, one of which was measuring accurately the variation of the obliquity of Earth's orbit and the tropical year duration. However, it also had religious purposes as it was used to verify the date for Easter.

Bianchini took inspiration from other great meridian lines like the one built by Giandomenico Cassini in the church of San Petronio, Bologna, in 1655. Bianchini compared the lunar eclipse's observations in Rome and Bologna and found out

that the meridian of the Pontifical State was from Rome to Rimini. Bianchini's meridian line built in Santa Maria degli Angeli materialized this meridian, which crosses Rome at 12° 29' 51" from Greenwich.

Santa Maria degli Angeli's meridian line was inaugurated on 6th October 1702. It is made of bronze and is enclosed in white marble surrounded by a line of yellow marble. Originally the pavement in the church was made of earthenware tiles which made the meridian line stand out from all the rest. All around the line the twelve zodiac signs are represented.

The light of the Sun comes from a circular pinhole, the gnomon, situated in the middle of pope Clemens XI's emblem.

The image of the Sun traverses the meridian line at local noon everyday and during the year it crosses through the whole line. The image of the Sun projected on the pavement goes from being almost a circle during summer solstices to being very oblong at winter solstice. The presence of two decorations near the Summer solstice position has remained unexplained until 2018. During summer solstice the sun passes at the closest point from the base of the gnomon and is expected to meet the two decorations with the papal emblem. However, only one of them receives the image of the Sun nowadays, therefore this can be used to evaluate the secular shift of the solstice's position. The position of the red marble strip under this decoration is the materialization of the solstitial center of the solar image in 1702.

## 2. Luigi Vanvitelli and the alterations on the meridian line

Between 1749 and 1750 the church underwent some architectural changes, made by Luigi Vanvitelli and the meridian line was altered. From a manuscript of the Jesuit astronomer Ruggiero Boscovich found in Italian National Library Vittorio Emanuele II in Rome in 2009, we get a deeper insight about the work of Vanvitelli and how it affected the meridian line. The document, whose aim was to suggest a solution to fix the line after the work of Vanvitelli, clearly illustrates the changes that were made. Before Boscovich's inspection, the monks of the church had already noticed that the cornice built under the gnomon prevented the light of the Sun from reaching the meridian line. They therefore operated a cut in it that let the light coming from the gnomon pass. Boscovich focused on other unfixed alterations in the meridian line, like the precise determination of the point where the perpendicular line from the gnomon meets the meridian line. This fundamental part should had been indicated on a slab on the pavement but Vanvitelli had it eliminated. Moreover, the base of one of the column added by Vanvitelli during the works was built on a part of the meridian line and obstructed it. Boscovich recommended for that part to be levigated and a movable piece was built and substituted to the part that was on the meridian line. So nowadays what we see is, in many parts, not the original work of Bianchini due to the alterations made during Vanvitelli's works.


Fig. 1. Santa Maria degli Angeli's meridian line with the modern epigraph at its beginning.

## 3. Meridian services and the procedure of the ball drop in Greenwich

In the past, many people relied on sundials to tell the time, however time slightly changes basing on the location. The difference in time, in England, would be up to 30 minutes depending on the longitude of the place. This led to the introduction of the Standard Time which was based on the Prime Meridian at Greenwich. Having a device to signal the exact time of noon was therefore very important, an instrument to measure it was by using a meridian line. Everyone would then check their clocks and adjust them basing on a certain signal. In Greenwich, a bright red ball on the top of Flamsteed House is used as a signal: each day the ball rises up its mast and is dropped exactly at 13:00. In the past this served as a signal to ships and to the whole population and corresponded to the instant of the meridian transit at the location. This procedure was established in 1833 and automated in 1852 by Airy.

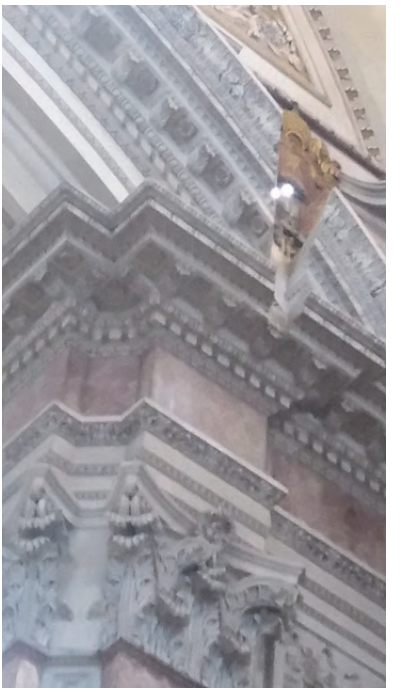


Fig. 2. The cut in Vanvitelli's cornice made in order to let the light from the gnomon reach the meridian line. The photo is on April 6, 2019 at 13:19:53. The two luminous points are the pinhole and, lower, the light reflected by the carved marble of the frieze.

## 4. Rome's meridian services

In Rome, for 150 years the signal was given by the meridian line at Santa Maria degli Angeli. In fact, as an epigraph inside the church reads, the meridian line served as a mean to indicate the time util 1846, when a cannon shot from Castel Sant'Angelo was established to mark midday, after receiving the visual signal from St. Ignatius roof. This part of the information is lacking in the modern epigraph visible in figure 1 on the pilaster at the beginning of the meridian line, leaving in the reader the wrong impression that a cannon could measure the local noon given by the solar transit on the local meridian.

A similar device to that of Greenwich was then used to give the signal to a cannon at the top of Castel Sant'Angelo. This device was created by Angelo Secchi and inspired by that of Greenwich, the aim of it was almost the same. However, the reason of such device was also to be found in the recent reform that concerned the day hours. The Pontificial State had been using a way of counting the hours in which the sunset coincided with the end of the day, that is the hour 24:00 was just after the sunset. Precisely, the sunset was set at 23:30. However, that meant a different time everyday for midday. This was the so called Italian way of counting

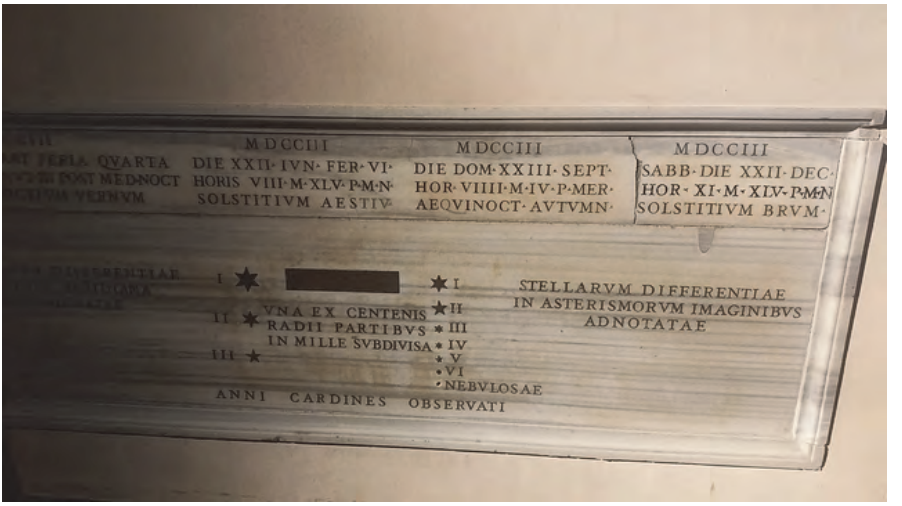



Fig. 3. The epigraph of the results of the observations made in the year 1703 with the meridian line. The photo is of 23 Jan 2019 at 13:26, when the sunlight hits the epigraph directly from the great window of SSW. The spring equinox was at HOR IX MIN III after local midnight on 21 March 1703.

hours. The sundials of Clavius at Collegio Romano, are still showing this use of measuring time. Using a meridian line, it is very easy to determinate the moment of midday, however with the method in use the hour 12:00 did not correspond to the moment the Sun was at its highest. Tables were required, basing on which people could get to know what time it was at the moment of the peak.

Pius IX changed this by adopting a method of measuring the time of the day which was used in France and in many other European countries. With this method the midday coincided with the hour 12:00.

Angelo Secchi's observatory in the Roman College was in charge of communicating the time of midday through a visible signal. As in Greenwich, a ball located on rooftop of St. Ignatius Church, gave the signal to the artillerymen that fired a cannon from Castel Sant'Angelo. Only at the beginning of the 20th century the cannon has been moved to the top of Gianicolo hill. There was now a reliable and unique signal to indicate the midday throughout the city.

Secchi also measured carefully the meridian of Rome, the fundamental meridian of Italy at 12° 27' 08" from Greenwich. The meridian passes through St. Peter's church in Rome and is physically indicated by a tower on Monte Mario hill. Santa Maria degli Angeli's meridian line is not along this meridian but about 8 seconds of time East of it. The measures of the meridian were part of a project with the aim of measuring the central European meridian, this was wanted mostly by Prussian government but many countries gave their contribute to it. Angelo Secchi was the person designated by the pontificial state to be in charge of the commission taking part in the measurements in the state's territory. With this work Secchi not only

demonstrated once again his great competence and capability but also paved the way to the modern geodesy.

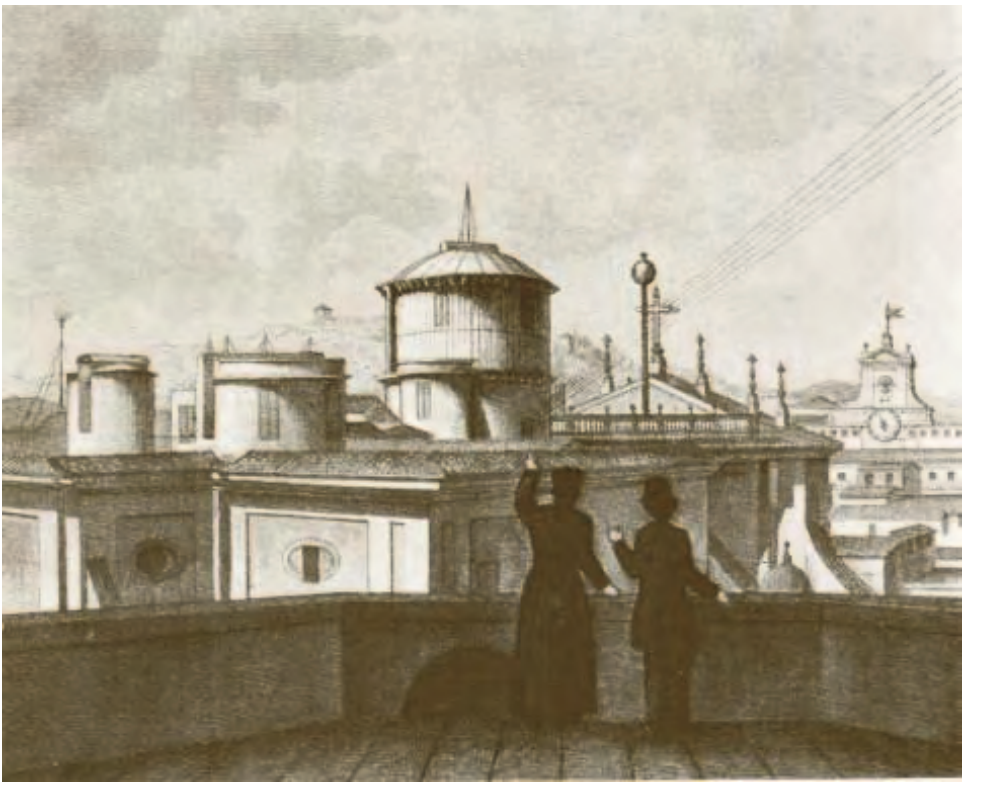


Fig. 4. The top of St. Ignatius church and the ball.

## References


1. Sigismondi, C. Lo Gnomone Clementino, Gerbertus 7 2012.
2. Catamo, M. Giornale di Astronomia, vol. 36 (3) 2010, pp. 41–45.
3. Sigismondi, C., C. Sterken and S. Pietroni, The rediscovery of the obliquity meter in the meridian line of St. Maria degli Angeli in Rome, Vienna, IAU XXX GA 2018.
4. Sigismondi, C., Christopher Clavius astronomer and mathematician, arXiv1203. 0476S 2012.
5. Lay, R., Il nuovo osservatorio del Collegio Romano, in Angelo Secchi Quater, Foligno 2012.
6. Aebischer, T., Le misure geodetiche, in Angelo Secchi, Quater, Foligno 2012.

# Angelo Secchi, the tradition of Gnomonics at the Collegio Romano, and the Equation of Time during the centuries

Pietroni, Silvia (1) and Sigismondi, Costantino*

*(1) Sapienza University of Rome * ICRA/Sapienza University and ITIS G. Ferraris, Rome, Italy*

** E-mail: sigismondi@icra.it www.icra.it/solar*

Angelo Secchi, following the tradition started with father Christopher Clavius (1535–1612) who wrote a mathematical treatise (1581) on Gnomonics, realized sundials and quadrants. The equation of time determines the difference between the mean solar noon and the transit of the Sun at the meridian line. Its behavior depends on two facts: the obliquity of the ecliptic and its eccentricity. The first is determinant for the double sinus curve during a year and the second for the asymmetry of the approach to the solstices, which determines some cares when the parabolic fit is adopted. It confirms the difficulty of measuring the solstices with respect to the equinoxes already evidenced in Ptolemy's Almagest.

*Keywords*: Angelo Secchi, Astronomy, Gnomonics, Christopher Clavius, meridian line, Santa Maria degli Angeli.

## 1. The book on Gnomonics of Clavius and the Sundials at Collegio Romano

Fig. 1. The Clavius sundial at the Collegio Romano between the Calandrelli tower (right) and the top of the façade (left): the Italic hours, with Sunset at 23:30 and Ave Maria half hour later at 24 h.

## 2. Sundials at the time of Father Secchi

Secchi built some sundials in Grottaferrata and in Sicily, but Gnomonics at his time was a subject of study of some extent. The priest Angelo Giuseppe Sarto, born in

1835, built two sundials in his parish at Tombolo and Fontaniva in Veneto. He later become pope Pius X (1903–1914). When he was pope had to answer about the equation of time to his former parishioners, because in the passage between the local time to the European one, TMEC, every sundial and also his own showed immediately the equation of time to everyone.

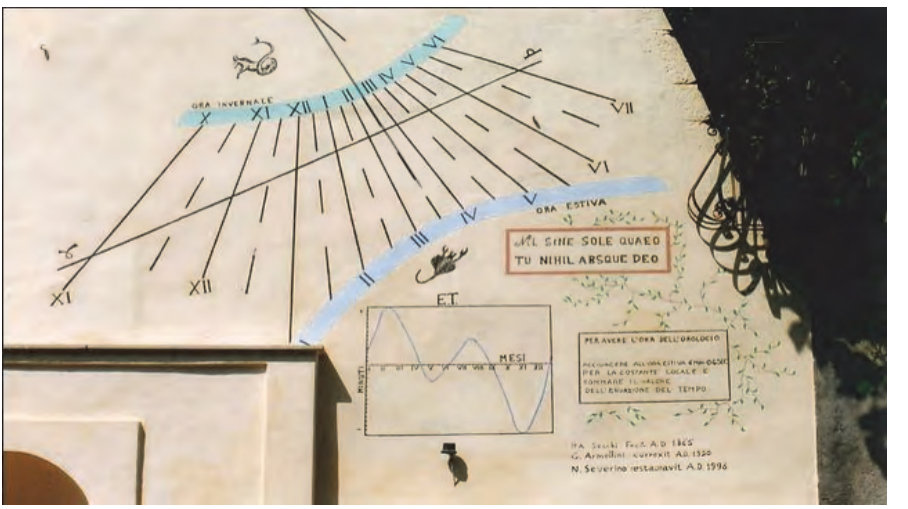



Fig. 2. The equation of time obtained with approximation.

## 3. Cause of the equation of time and first representations

Se puoi Jean des Murs e gli altri se no solo Tolomeo in Almagesto.

## 4. The bisection of the eccentricity from Ptolemy to Kepler

Kepler introduced the elliptical orbits, but it happened, already with Cassini data, that the ptolemaic eccentricity was exactly the double of the keplerian one, by definition of eccentricity of an ellipse. Therefore the introduction of an eccentricity, well before Ptolemy, after Eudoxus with Apollonius that in 3rd century BC who explained the equivalence with epicycles, already was an early evidence that the orbit of the Sun or of the Earth was showing different velocities in time and the existence of two opposite apsides. The perigee and the apogee become perihelion and aphelion, but the date remained the same. Arab astronomers understood their westward rotation, and between 15 and 17 centuries the astronomers tried to measure these modifications to the perfect two-body problem with increasing accuracy.

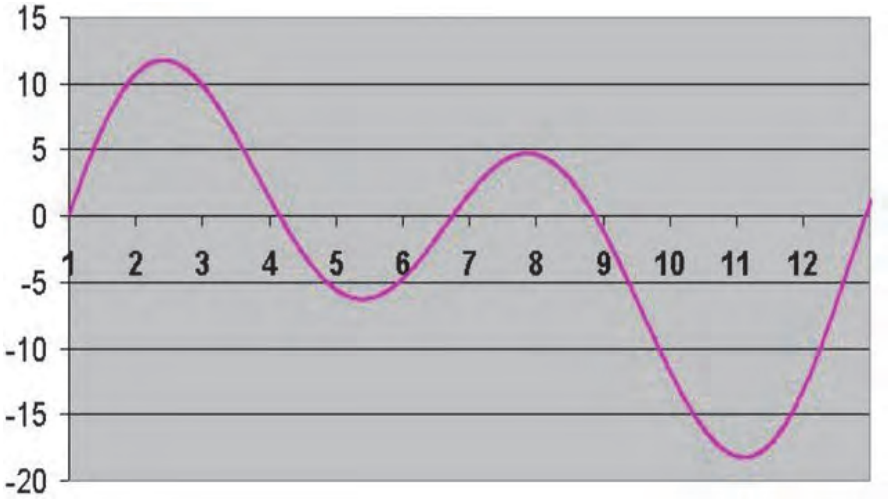



Fig. 3. The equation of time obtained with approximation.

The presence of the axis of apsides now occurring on 3–4 January and 3–4 July determines an asymmetry in the solar approach to the solstices: slower before the perigee on January with a faster farthening, conversely the summer solstice of June 21 is approached with a faster speed than later. This fact produces an estimate of the solstice instant which is sistematically earlier than the true instant as we increase the timespan between symmetric dates with the meridian Sun at the same heigth. Example the cardinal Nicholas Cusano estimated the winter solstice averaging the two dates in which his pinhole allowed the sunlight entering in his room at Andraz Castle. Now it happens between 2 February and 18 November at the 19 of December is the result of this estimate, because the Sun reach the same height of 18 November slightly faster after the perigee which now is Jan 4. At the time of Cusano the perigee was in January 1st 1450 and dividing by 2 the timespan the solstice estimate yielded probably 13 of december (the tradition of St. Lucy in Italian sounds "santa Lucia il giorno più corto che ci sia" says that st Lucy feast on Dec. 13 is the shortest day. Cusano already was able to measure that the solstice occurred well earlier than 21 december. The difference between the legth of the days near the winter or the summer solstice was not measurable accurately without chronometers. Cusano and other scholars already knew that the solstice was a minimum in the meridian height of the Sun, then its estimate by symmetry was one rapid solution to obtain that day, and the proverb tells us that the people was aware of the true solstice at the beginning of the Renaissance.

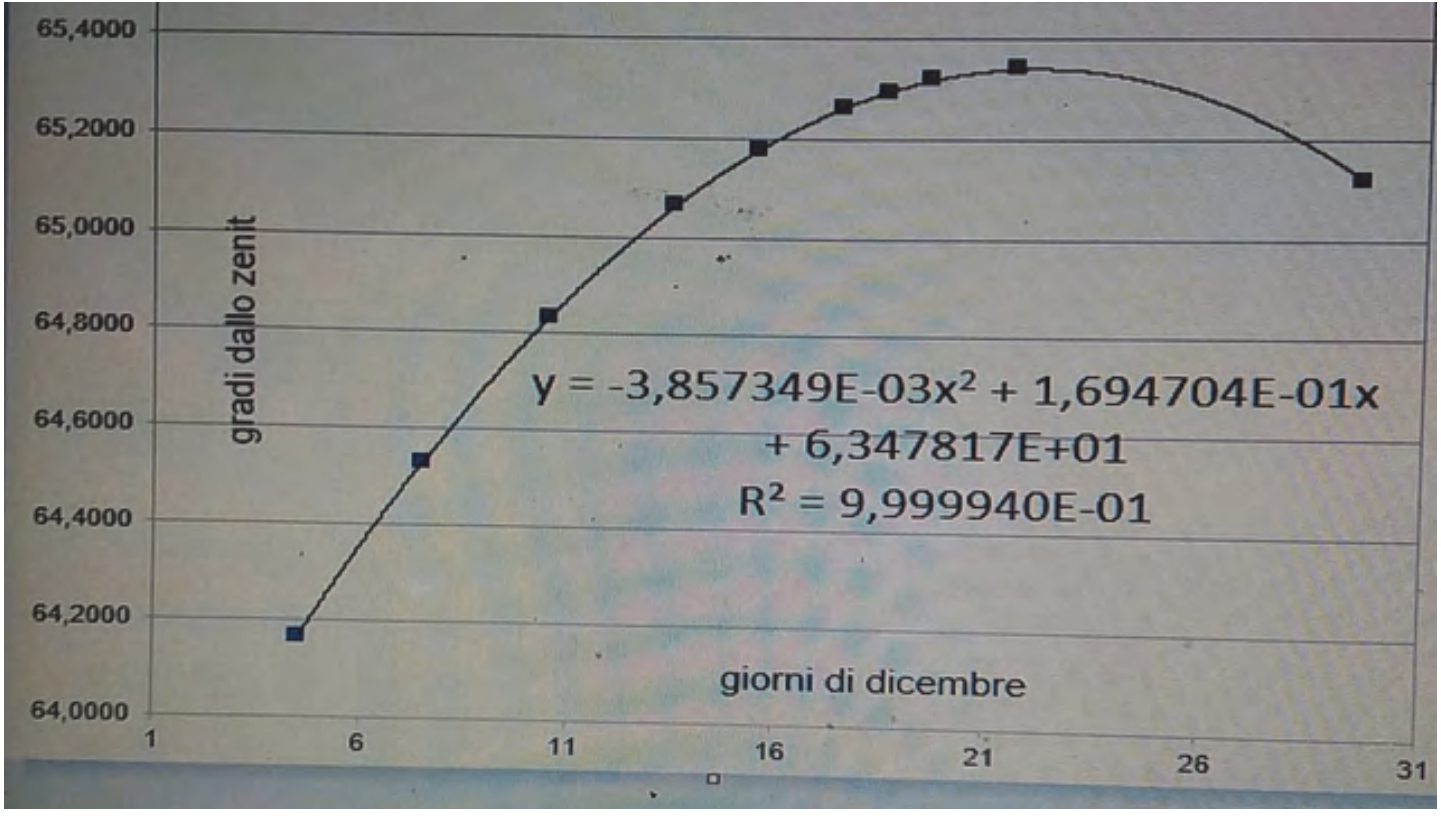



Fig. 4. The winter solstice measured in S. Maria degli Angeli in 2018–19 until December 29, with the asymmetry not yet evident.

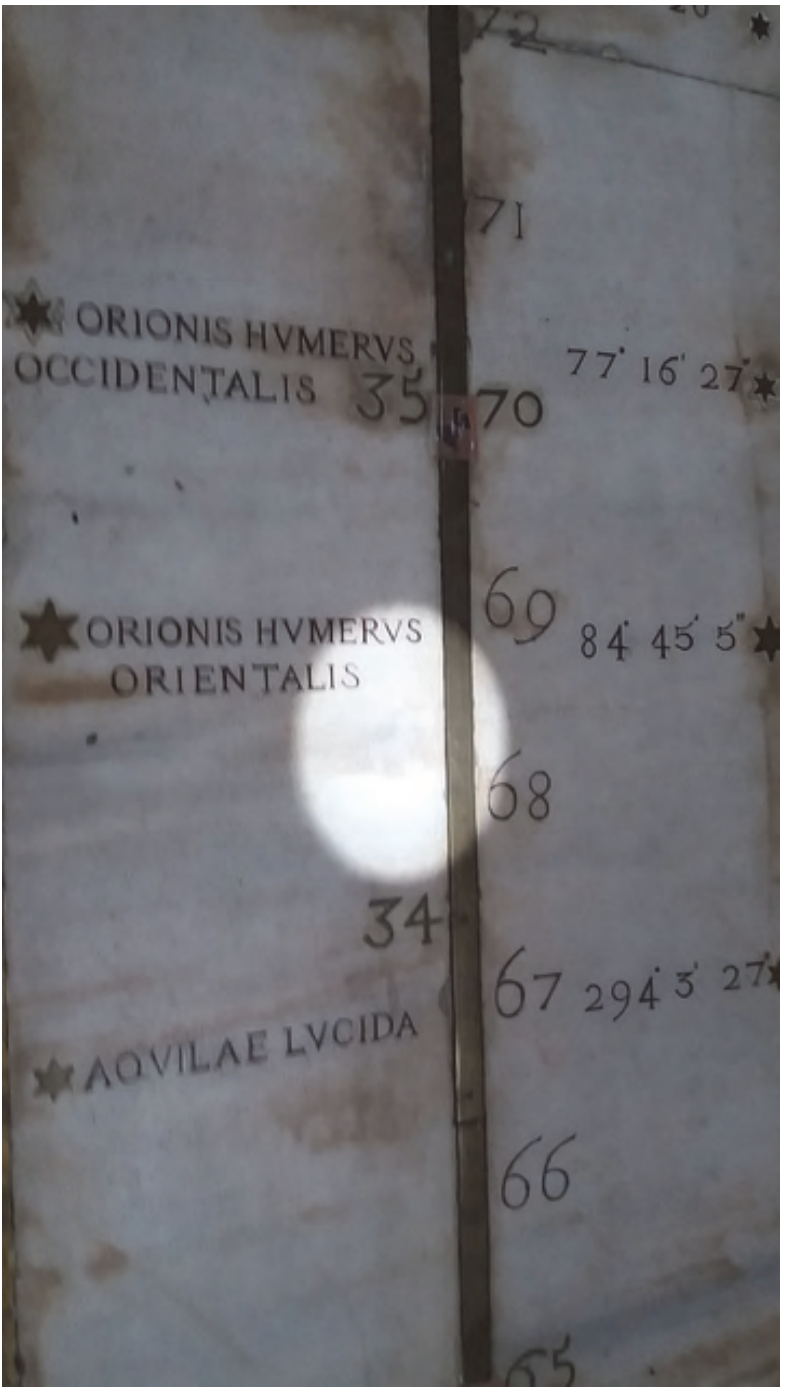



Fig. 5. The position of Betelgeuse, Humerus Orientilis Orionis, in 1701 atlas of Philippe de la Hire. The Southern solar limb was on it near April 10; nowadays the Sun is much southern with respect to Betelgeuse, for the precession. Betelgeuse is in the opposite direction of local motion of the Celestial North Pole from 1701 to present.

## References

1. Sigismondi, C. Lo Gnomone Clementino, Gerbertus 7 2012.
2. De Donà, G. Il Foro Gnomnomico di Cusano al Castello di Andraz, Gerbertus 4 2011.
3. Sigismondi, C., C. Sterken and S. Pietroni, The rediscovery of the obliquity meter in the meridian line of St. Maria degli Angeli in Rome, Vienna, IAU XXX GA 2018.
4. Sigismondi,C., Christopher Clavius astronomer and mathematician, arXiv1203. 0476S 2012.
5. Tuscano, M. T., Secchi e la Gnomonica, in Angelo Secchi Quater, Foligno 2012.

# Angelo Secchi: Un inquadramento storico-scientifico

Namara, Alexandre

This paper is rewritten from an original paper of 1976. It is in Italian and gives a complete framing of the scientific activity of Angelo Secchi in comparison with his contemporaries. Alexandre Namara published this contribution in a famous book "les secrets de l'Astronomie" published in Geneva in 1976 and translated into 4 languages.

*Keywords*: Angelo Secchi, Astronomy, Astrophysics.

## 1. I primi incontri con la neonata astrofisica

Padre Angelo Secchi (1818–1878) esiliato dall'Italia con i Gesuiti, fu inviato all'osservatorio astronomico di Georgetown. Tornato in patria, venne nominato direttore dell'osservatorio del Collegio Romano, che dotò di nuovi e più adatti strumenti, come egli stesso racconta: "In questi ultimi anni è sorto un ramo di studio non guari prima considerato. Cioè lo studio fisico degli astri. Se ne togliamo i lavori degli Herschel, i quali non mai ebbero un osservatorio regolare propriamente detto, questi studi non furono coltivati che sporadicamente e da pochissimi amatori. Anzi gli Herschel stessi sparirono dalla scena dopo l'immortale spedizione di Sir John al Capo di Buona Speranza. Talché l'astronomia fisica era quasi un terreno abbandonato all'epoca in cui fummo chiamati a dirigere l'Osservatorio del Collegio Romano, onde ci risolvemmo di dedicare ad essa le nostre fatiche... Questo concetto fu quello che regolò tutti i nostri studi posteriori e la scelta stessa degli strumenti per il nuovo osservatorio. Questo studio fu allora creduto una superfetazione e ci fu perfino chi disse che al Collegio Romano non si coltivava la scienza astronomica, ma la fisica, e vi è stato chi ha rifiutato il titolo di astronomo, quasi che Galileo e gli Herschel, la cui vita fu spesa in tali studi, non fossero astronomi. Ma il tempo ha fatto giustizia, e senza vanità possiamo dire che ora, sulle nostre pedate, sorgono altrove osservatori esclusivamente fisici per lo studio dei corpi celesti, come a Oxford, a Berlino, a Parigi stessa, a Calcutta e in altri siti. Questa fisica degli astri, allora bambina, si è svolta nell'intervallo di 25 anni dacché lavora l'osservatorio e questo ha tenuto un qualche posto nel suo avanzamento".

## 2. Gli strumenti e gli studi di Secchi a Roma

Gli studi di padre Secchi erano cominciati già quando poteva disporre solo di strumenti inadeguati. Si era interessato allora del Sole e, con una pila termoelettrica, ne aveva misurato l'intensità di radiazione in occasione dell'eclisse del 1851 e in seguito sull'immagine solare fornita dal telescopio. Era sua intenzione di valutare il rapporto fra la radiazione al centro del Sole e quella sul bordo. Sull'immagine proiettata dal rifrattore egli riuscì a calcolare le deviazioni galvanometriche provocate dalla pila, ottenendo la conferma che l'intensità di radiazione al centro del disco solare è quasi doppia che ai bordi. Padre Secchi si dedicò anche alle comete: ne scoprì una nel 1853 che aveva un nucleo multiplo e l'anno prima aveva rintracciato i due

resti della cometa di Biela, che si era frantumata durante il suo ritorno periodico del 1846. Avendo intravisto attraverso il nucleo di uno dei due pezzi una debole stella, aveva concluso che le comete sono costituite di materiale molto rarefatto. Ancora non si conosceva l'origine delle stelle cadenti, ma, dagli studi condotti, egli propendeva per una loro natura cosmica e per accertarsene eseguì delle regolari osservazioni annuali dei punti radianti.

Quando fu finalmente in possesso di strumenti più moderni e più precisi, si occupò di Saturno e analizzò it suo anello determinandone l'intensità luminosa delle varie parti, notò che l'anello interno è nebuloso e che la superficie non è piana e ne calcolò it diametro, lo schiacciamento e l'eccentricità. Osservò anche Giove, di cui ha lasciato molti disegni riguardanti le particolarità presenti sulla sua superficie, e rilevò la presenza nell'atmosfera di grandi perturbazioni. Fu in grado perfino di seguire le orbite dei pianeti Medicei e ne misurò i diametri, il periodo di rotazione assiale e la loro diversa capacità di riflettere la luce solare. Fu poi la volta di Marte, sul quale scoprì la presenza di due formazioni di colore rosso cupo e di carattere stabile, che egli chiama ⟪canali⟫. Si accorse inoltre che le calotte erano eccentriche rispetto ai poli e le considerò erroneamente nubi.

### 3. La Spettroscopia

Cominciò ben presto ad introdurre l'uso dello spettroscopio nell'analisi dei pianeti e in tal modo osservò che nello spettro di Giove e Saturno vi erano righe relative alla presenza di vapor acqueo, ma soprattutto che erano presenti righe caratteristiche nel rosso, che lo portarono a concludere che la loro atmosfera ⟪non era ancora purgata⟫ ed era costituita diversamente dalla nostra. Gli spettri di Nettuno e Urano apparivano discontinui, mentre da quello di Marte ricavava che la sua atmosfera era molto trasparente. Riguardo alla Luna egli eseguì dei disegni e delle fotografie. Nei disegni insiste su alcune caratteristiche dei crateri, in particolare su "Copernico". Le fotografie risalgono addirittura al 1859 e furono eseguite come esperimento di fotografia celeste. Padre Secchi si avvide in tal modo che rifrattori non erano adatti per la occhio umano, perché i loro obiettivi erano corretti per l'occhio umano, ma non adeguati alla sensibilità della lastra fotografica. Bisognava allora utilizzare a tale scopo i riflettori.

Venuto a conoscenza che G. B. Donati era riuscito ad ottenere lo spettro di 15 stelle, P. Secchi volle cimentarsi nell'impresa. Potendo utilizzare spettroscopi già perfezionati per questo uso, gli fu possibile distinguere vari tipi di spettri stellari. Ora, la sua intenzione era di studiare il maggior numero possibile di stelle per vedere se "come sono innumerevoli le stelle, fosse pure proporzionatamente varia la loro composizione". Dopo aver messo a punto uno strumento appositamente adattato allo scopo, il prisma-obiettivo di Merz, iniziò le sue osservazioni che lo portavano a concludere che "le stelle sono numerosissime, eppure i loro spettri si riducono a poche forme ben definite e distinte, che per brevità noi chiamiamo tipi. L'esame delle stelle ci ha occupati per parecchi anni; furono esaminate quasi tutte le principali e

moltissime altre, almeno 4000 in tutto, perché oltre la stella principale si esaminava tutto il suo contorno".

Le stelle furono suddivise a seconda delle fondamentali caratteristiche spettrali in tre tipi principali: stelle bianco-azzurre, gialle e rosse; inoltre ne esiste un quarto definito "assai bizzarro e vario. Potrebbe credersi a prima vista che esso risulti dal terzo sopprimendo alcune delle righe oscure"; il quinto tipo è formato dalle stelle che hanno spettri con righe brillanti di emissione. Questa classificazione ebbe molto successo e venne subito adottata, anche perché apparve chiaro che essa doveva corrispondere ai diversi stadi dell'evoluzione stellare.

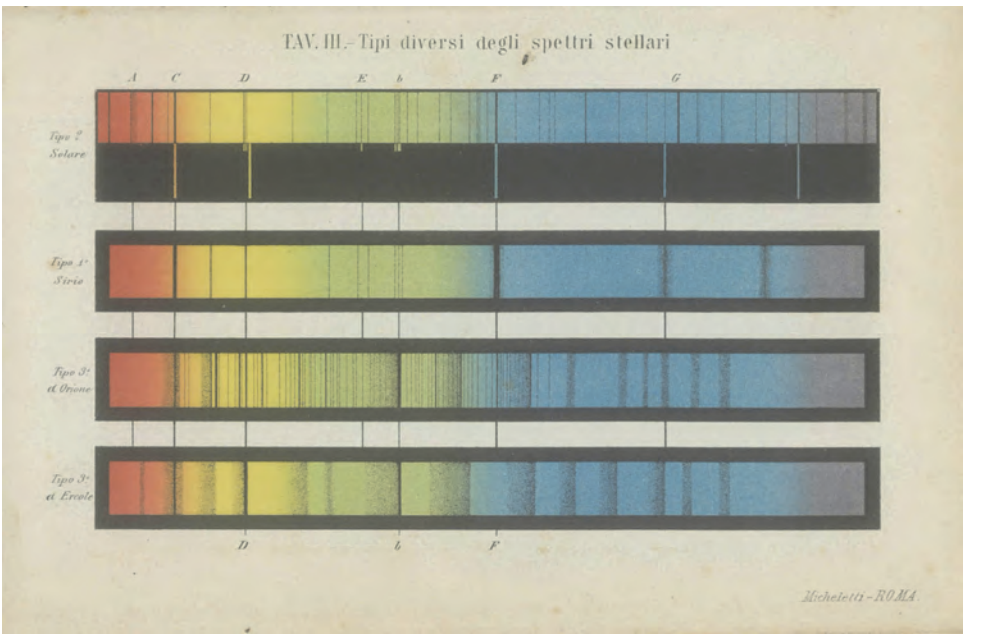



Fig. 1. Spettri di alcune stelle studiate da Secchi.

## 4. La fotografia di spettri

Intanto Sir William Huggins (1824–1910) analizzando vari spettri terrestri e celesti e ponendoli a confronto, osservò alcune analogie: entrambi tipi di spettri presentavano dei caratteristici due raggruppamenti di righe corrispondenti a determinati elementi come soprattutto l'idrogeno, il sodio e il magnesio. Le nebulose fino ad allora erano considerate degli ammassi di stelle troppo lontani per essere investigati anche coi più potenti telescopi mediante l'analisi spettroscopica poteva scoprire che alcune nebulose dovevano considerarsi piuttosto degli ammassi di gas. Infatti gli spettri di varie nebulose presentavano delle righe brillanti di emissione, caratteristiche dei

gas incandescenti. I suoi potenti spettroscopi gli avevano fra l'altro consentito di scoprire la presenza in certi spettri di un debole effetto Doppler. Dopo qualche anno Huggins potè determinare la velocità di allontanamento o di avvicinamento al Sole di varie stelle. Nel 1880, utilizzando lastre fotografiche "asciutte", Huggins fotografò sulla stessa lastra uno spettro terrestre e uno stellare, riuscendo così a determinare immediatamente la velocità della stella lungo la visuale. In seguito al perfezionamento di questo metodo, Vogel nel 1881 confrontò lo spettro dei due lembi, del Sole durante la rotazione e potè confermare l'esistenza dell'effetto Doppler. P. Secchi ebbe il merito di considerare il Sole un utile mezzo di conoscenza delle proprietà stellari, difatti esso è l'unica stella vicina a noi ed è la sola di cui possiamo osservare ampiamente la superficie. Le caratteristiche rilevabili dall'analisi spettroscopica si sarebbero potute generalizzare per la altre stelle. Egli stesso si occupò della natura delle macchie solari e si convinse che esse sono zone a temperatura inferiore rispetto alla fotosfera come sarebbe poi risultato dall'analisi dello spettro. Anche R. C. Carrington si occupò di questo problema e giunse a conclusioni sulla legge di rotazione del Sole a varie latitudini, l'esistenza e la direzione di correnti sistematiche, la distribuzione delle macchie sulla superficie solare.

Spetta invece a J. F. Janssen di aver fotografato per la prima volta la fotosfera, che si presenta ricoperta da piccoli granuli luminosi, per la maggior parte di forma ovale. Fotografò, inoltre, un'eclisse parziale di Sole, con la quale si poteva dimostrare l'inesistenza di un'atmosfera lunare: infatti i granuli che si trovavano presso il lembo lunare non subivano la minima distorsione. Tali granulazioni erano considerate da Padre Secchi come la punta delle colonne di gas surriscaldato, che uscivano dalla fotosfera. La presenza di macchie, penombre, granulazioni e facole faceva pensare che sul Sole esistesse una zona di grandi perturbazioni.

## 5. La natura delle macchie del Sole

A causa di tale zona sorse una polemica tra Padre Secchi e un membro dell'Accademia delle scienze di Francia, Hervé Faye. Quest'ultimo sosteneva che il Sole poteva essere paragonato a una macchina termica, che emanava calore dalla sua superficie. La sorgente di calore era la massa stessa del Sole, la sorgente del freddo lo spazio celeste, il condensatore era la fotosfera. Il meccanismo di funzionamento consisteva in correnti ascendenti e discendenti, le une portanti i vapori, le altre le sostanze raffreddate e solidificate. Le macchie solari non sarebbero che dei cicloni. Contro questa ipotesi P. Secchi obiettava, in primo luogo, che le macchie solo di rado presentano un movimento vorticoso, necessario per ammettere loro natura ciclonica. Infatti di trecento macchie da lui osservate nel corso di un anno solo otto rilevavano un movimento vorticoso arguibile dalla struttura a spirale. Fra l'altro, questa, quando è presente, permane solo uno o due giorni, mentre le macchie mostrano un carattere più duraturo. Egli quindi riteneva che la massa solare fosse costituita prevalentemente di gas, come facevano supporre le caratteristiche delle macchie, la cui struttura era paragonabile a quella delle nuvole o delle fiamme.

Un'ulteriore conferma della sua ipotesi venne dallo studio di quelle eruzioni gassose che spesso si osservano nei nuclei delle macchie. La fotosfera sarebbe allora un fluido elastico analogo ai nostri gas e i movimenti a spirale possono effettivamente essere ricondotti a fenomeni simili a cicloni terrestri. I vortici sono più frequenti nel periodo di formazione delle macchie, ma in seguito scompaiono e al loro posto subentrano correnti regolari convergenti verso il nucleo delle macchie.

## 6. Le foto del Sole

Nel 1857 fu costruito il primo "fotoeliografo" con lo scopo di fotografare regolarmente il Sole ogni giorno. Si ottenne così una preziosa documentazione dell'attività variabile del Sole durante l'anno. Le fotografie stereoscopiche delle macchie e delle facole evidenziarono che queste si trovano nella parte alta della fotosfera, mentre le macchie sembrano dei buchi o vortici sprofondanti nella parte più bassa di essa. Padre Secchi ne potè determinare allora, la "parallasse di profondità" cioè gli spostamenti apparenti determinati dalla depressione della superficie solare durante la rotazione. Egli potè proseguire le ricerche di Angstrom e Kirchoff sullo spettro solare e inoltre notò che l'aspetto delle righe di Fraunhofer (più o meno diffuse) dipendeva dallo stato fisico di pressione e di temperatura delle zone d 'atmosfera in cui avviene l'assorbimento.

## 7. Le eclissi del Sole

Nel 1860 vi fu un'eclisse totale di Sole e Padre Secchi ne approfittò per studiare che natura avessero quelle protuberanze rosse, cioè quelle grandi eruzioni di idrogeno che erano già state notate dagli antichi nel corso delle eclissi. Dalle fotografie che egli eseguì sull'immagine diretta del disco solare occultato dalla Luna, e da quelle che realizzò Warren de la Rue con l'eliografo di Kew risultò che le protuberanze effettivamente esistono e non sono giochi di luce; che appartengono alla superficie solare; che anche la corona è reale e che le sue dimensioni sono meno estese ai poli che all'equatore, e raggiungono un massimo a 450 di latitudine. Janssen, invece, osservò l'eclisse di Sole del 1868 in India, per osservare le protuberanze allo spettroscopio. Queste presentavano uno spettro di emissione molto brillante ed egli pensò che quindi potevano essere osservate anche normalmente, senza eclisse. Adottò, a tale scopo, un metodo che sarebbe risultato utilissimo. Egli puntò la fessura dello spettroscopio radialmente, in, parte sul disco solare, in parte al di fuori, e osservò una parte del lembo. Notò così che la riga C di Fraunhofer presentava un prolungamento, una sottile riga rossa e brillante che perdurava nonostante gli spostamenti della fessura su altre regioni, ma che variava notevolmente in altezza e splendore a seconda delle regioni perlustrate. C. A. Young, osservando una nuova eclisse totale, riuscì a dimostrare la costituzione della corona, la cui natura appariva gassosa. Dall'analisi del suo spettro si osservava la presenza di una riga verde di emissione che egli identificò con la riga cromosferica vicina, che ha analoga lunghezza d'onda.

La cromosfera, un attimo prima dell'oscuramento è molto brillante, ma scompare rapidamente subito dopo il principio della totalità quando diventa visibile la riga coronale molto più debole. à probabile che per la piccola dispersione allora usata, le due righe possano essere state confuse.

Contemporaneo di Young fu Sir J. Norman Lockyer, il quale mise a punto, quasi negli stessi anni di Janssen, la tecnica per studiare le protuberanze solari in piena luce. Con le sue ricerche scoprì che esse si originavano da un involucro gassoso che circonda il Sole e a cui diede nome di ⟪cromosfera⟫. Dallo studio degli spettri solari e soprattutto da quello della cromosfera, egli potè formulare una sua ipotesi sulla dissociazione degli elementi alle alte temperature presenti nelle atmosfere stellari. Secondo la sua teoria le righe spettrali H e K della regione del violetto, sono dovute alla dissociazione del calcio, ed esse aumentano di intensità man mano che aumenta la carica elettrica.

## 8. L'Elio e le perturbazioni magnetiche

Nello spettro delle protuberanze era presente una riga gialla brillante D3, che non era determinata da nessun elemento presente sulla Terra. essa fu quindi attribuita a un nuovo gas che prese il nome di ⟪elio⟫. In Italia, allievi e successori di Padre Secchi furono Pietro Tacchini e Annibale Riccò che si interessarono soprattutto alle relazioni fra fenomeni solari e fenomeni terrestri. Il primo si occupò di valutare l'influenza delle eruzioni solari sulle aurore polari nelle parti alte dell'atmosfera terrestre. Il Riccò eseguì molte osservazioni sulle macchie solari e sulle protuberanze e studiò le eventuali relazioni fra esse e le perturbazioni del magnetismo terrestre. Riuscì così a rilevare un ritardo nella perturbazione magnetica sulla Terra di circa 45 ore dopo il passaggio delle macchie solari per il meridiano centrale del Sole. In Russia nel 1894 A. Belopolsky scopriva variazioni periodiche nella velocità radiale di Delta Cephei, dimostrando che questa è una proprietà generale delle variabili a corto periodo.

## 9. I primi cataloghi stellari e i principali temi dell'Astronomia dell'ottocento

Intorno alla metà dell'ottocento incominciano ad apparire cataloghi stellari redatti con sempre maggiore accuratezza e precisione, e Bessel riesce a calcolare, con ottima approssimazione, la prima distanza stellare; sempre in questo Draper ottiene la prima fotografia Luna e Daguerre fotografa un'eclisse solare. Galle, nel 1846, scopre il pianeta Nettuno e, poco più tardi, Doppler riesce dare una spiegazione al fatto che le righe spettrali di una sorgente luminosa si spostano verso il violetto o verso il rosso a seconda che l'astro si avvicini od allontani dall'osservatore (fenomeno poi conosciuto col nome di effetto Doppler). Con lo studio dell'analisi spettrale, che consente di calcolare la temperatura e la composizione chimica di qualsiasi sorgente luminosa, ha inizio una nuova corrente astronomica, l'astrofisica che, come abbiamo visto, ebbe per primi ferventi sostenitori A. Secchi e W. Huggins. Verso la

fine del XIX secolo viene determinata, prima in modo approssimato, poi sempre più accuratamente, la parallasse solare, la cui importanza abbiamo già rilevato nei capitoli precedenti. Nel 1891 H. Deslandres e G. E. Hale eseguono particolari fotografie solari servendosi di un nuovo strumento: lo spettroeliografo. Intanto gli apparati strumentali vengono ancora maggiormente perfezionali seguendo il parallelo sviluppo della tecnica; allora incominciano ad essere determinabili quei piccoli spostamenti degli astri che prima non potevano essere notati. Così Chandler potè riscontrare alcuni piccoli mutamenti net valori delle latitudini di alcuni osservatori e, di conseguenza, scoprire il fenomeno della variazione di latitudine. Vengono frattanto pubblicati alcuni cataloghi stellari tra cui il Draper Catalogue, che comprende classificazione di oltre diciannovemila stelle, e quello di F. A. Argelander di 324 mila stelle dell'emisfero boreale.

Le scoperte di piccoli pianeti si susseguono freneticamente, potendosi ora gli astronomi servire del valido aiuto della fotografia. Schiaparelli, l'astronomo italiano più rappresentativo di questo periodo, compie all'osservatorio di Brera di Milano i suoi studi sui canali di Marte, che tanto hanno fatto discutere fino ai nostri tempi; egli scopre inoltre che il periodo di rotazione di Mercurio intorno al proprio asse è uguale a quello di rivoluzione intorno al Sole. J. C. Maxwell espone la sua teoria secondo cui gli anelli di Saturno sono composti di piccoli corpi o frammenti di corpi celesti, che ruotano intorno al pianeta. Intanto si scopre che la Luna è priva di atmosfera e si osserva la ⟪macchia rossa⟫ di Giove. Pickering, servendosi delle teorie espresse da Doppler, scopre un nuovo tipo di stelle doppie che, trovandosi ad una distanza apparente molto vicina, non potevano essere osservate visualmente: a queste venne dato il nome di ⟪binarie spettroscopiche⟫. Come era stato previsto da Bessel in linea teorica, si riesce ora ad osservare le compagne di Sirio e Procione.

## Reference

1. Alexandre Namara, Le avventure dell'astronomia, Ferni ed., Ginevra 1976.

# Father Angelo Secchi: A giant solar physicist of the 19th century

Koutchmy, Serge*

*emeritus scientist at the Paris Institut d'Astrophysique de Paris (CNRS) and the Sorbonne University
E-mail: koutchmy@iap.fr

Why Angelo Secchi has been a giant in solar physics? the following considerations will demonstrate this thesis. Both observationally, experimentally and theoretically Secchi was always at the cutting edge of his times in the realm of solar physics and astrophysics. His book Le Soleil has recently been reprinted by Forgotten Books (August 6, 2018) and Il Sole, Italian version, is being published by Gerbertus Academic Journal (2019). The experience of Secchi drove me during the PhD Thesis at Sorbonne University, through his book Le Soleil.

*Keywords*: Angelo Secchi, Solar Physics.

## 1. Introduction

It is difficult to pay a full or even a partial tribute to father A. Secchi S. J. so extended are his contributions to different fields of the Astrophysics of the 19th century and so highly appreciated has been his life almost entirely devoted to Science and Education. Let me just restrict myself to a memory that I vividly keep since the late 1960ies, when I discovered the double big volume of his wonderfully written in French books "Le Soleil" edited by the House "Gauthier- Villars" in Paris. I mean the 2d Edition issued in 1877 made of 8 so-called "books" of a total of more than 900 pages plus an atlas containing the 1st solar spectra (1st never published I guess). The reading of each volume that I performed several times at the Paris Observatory Library, impressed me so much that I seriously decided to continue and deepen my research on the Sun in general, instead of restricting on the analysis of total eclipse observations and their applications to the study of the expanding solar corona and structures which was the topic of my PhD at Sorbonne University.

## 2. Secchi's book and solar eclipses

The 1st volume indeed contains a very detailed account of the observations performed by several famous scientists and by father Angelo Secchi himself, at eclipses, starting with the 1842 eclipse. Secchi described with great details the chromosphere and prominences discovered at time of the totality and just before and after: it seems legitimate to claim that he discovered indeed the solar chromosphere and prominences, including its finest details that today are called spicules. It was amazing to read in 1970–71 how the Jesuit-Astronomer interpreted the observation of the white- light aureola seen around the Sun during the totality that now is called the solar corona, with the chromospheric red fringe at the feet. I cannot resist in showing in Figure 1, a copy of one of his drawing reproduced in the book IV, p. 350 of vol. I showing the eclipse observation of 1868 which reflects for the 1st time the

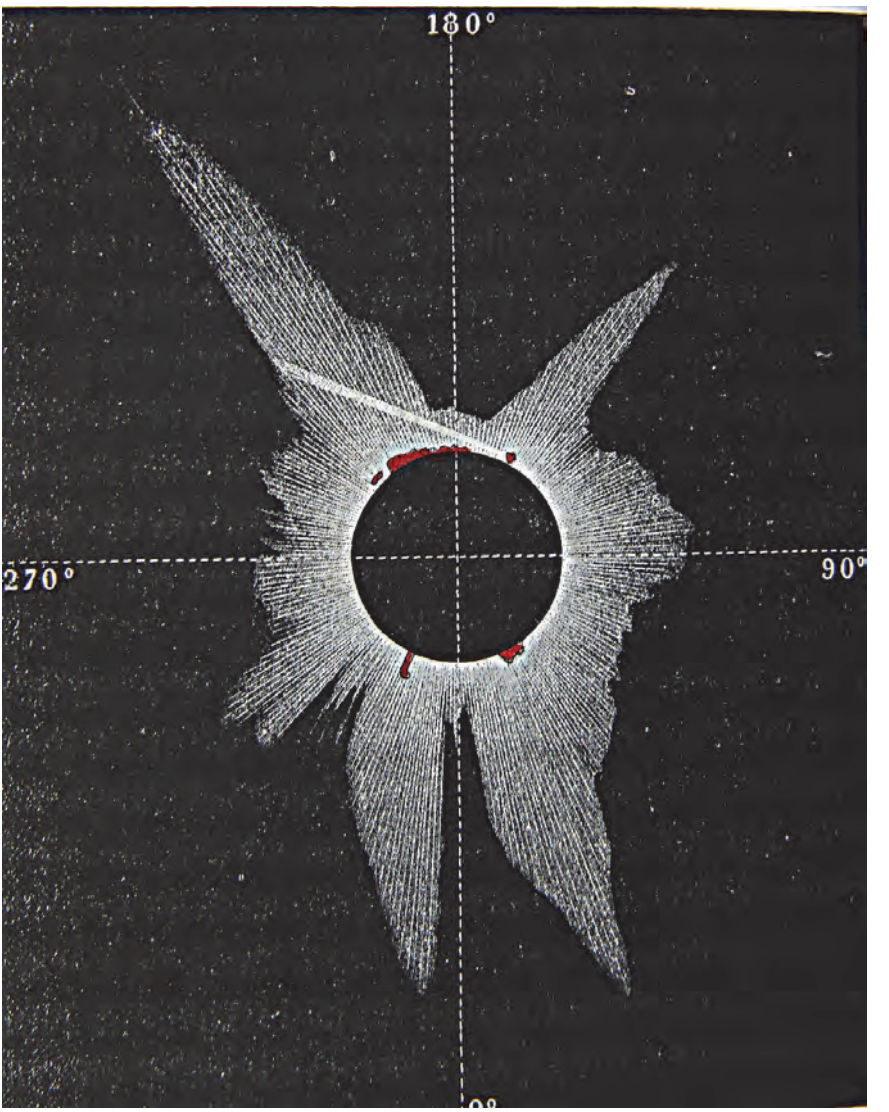



Fig. 1. Drawing of the 1868 solar total eclipse corona from the Angelo Secchi's book "Le Soleil" 1877, p. 350, vol. I. Note the prominences, in red, and large streamers shown for the 1st time. Because observations were done near a solar minimum of activity, the polar axis is horizontal (270 to 90° direction) and streamers extend about the equatorial plane.

behavior of the outer corona with what he called "aigrettes" and that today are called streamers. This was at a time when the Sun itself was still poorly understood (remember that the great Newton thought it is a giant planet and his follower W. Herschel considered until the beginning of the 19th century, the Sun was inhabited); the plasma high temperature properties of the corona, its magnetism and

its expansion towards the interplanetary medium was completely ignored or totally not understood. His very accurate and critical account of observations made by eclipse observers who started to use the photographic method to help in making a more objective drawing of their fugitive and transient observations. This was of great help for a young student, not only for analyzing the solar corona, but also from the epistemological point of view. Just by comparing the far reaching Secchi's suggestions with what was established in the late 1960s, following the numerous and enlightening works produced on the same subject by Grotrian, Lyot, Edlen, Alfven, Shklovsky, Van de Hulst, Parker et al., the delicate processes leading to the discovery of new properties in Physics and in Nature can be evaluated.

## 3. The mission Stereo and the solar corona

Angelo Secchi is today broadly recognized as the founder of the stellar classification based on spectroscopic data; in his honor the most famous experimental package of the international Space Mission "Stereo" hold his name. He was clearly influenced and even impressed by spectra in general and he seriously described solar spectra and spectroscopic methods in his books. Regarding eclipses, he was the 1st to extensively comment the coronal spectrum obtained by Rayet at the 1868 eclipse in Malacca, p. 395 of Vol. 1. Interestingly he reported the observation of a new strong line he called E, situated between the line F (identified with H beta line of HI that Secchi knew from the prominence spectra) and line D (much later identified as the D3 line of HeI). The line E of the Rayet spectrum is now known as being the brightest emission line of the corona due to Fe XIV, the so-called green line that the today literature is used to attribute to Young who described it a little bit later and paid attention to the absence of any identification using the known at that time prominence spectra, concluding on the coronal nature of the line. The Rayet spectrum was reproduced in the Secchi's book (his Fig. 142) and it is easy to notice that the green line E is very well apparent without any additional comment! Regarding prominences, Secchi also published a large number of observations in his treatise. This work was illustrated with magnificent color plates of the chromosphere and of the various types of prominences. Limb eruptive phenomena were extensively observed in the Halpha line of hydrohen with his 25 cm aperture refractor equipped with a broad slit with a large dispersion spectroscope thanks to 3 consecutive prims put inside. Several plates were later used to illustrate and comment new works of the genial observer, including the study of the famous very fine spicules that the chromosphere is made of. I notice that following the D. Menzel studies of the 1930s, small parts of the books were translated after the WWII at Sacramento Peak Observatory and both J. Beckers and especially R.B. Dunn developed the topic. An instrument called the chromopsheric telescope was built by R.B. Dunn to specially analyzed the Secchi's spicules and the chromospheric limb. Figure 2 shows one of the plate of the book illustrating how deep and extended was the analyzed limb features! Several of his plates could even today illustrate a book dealing with

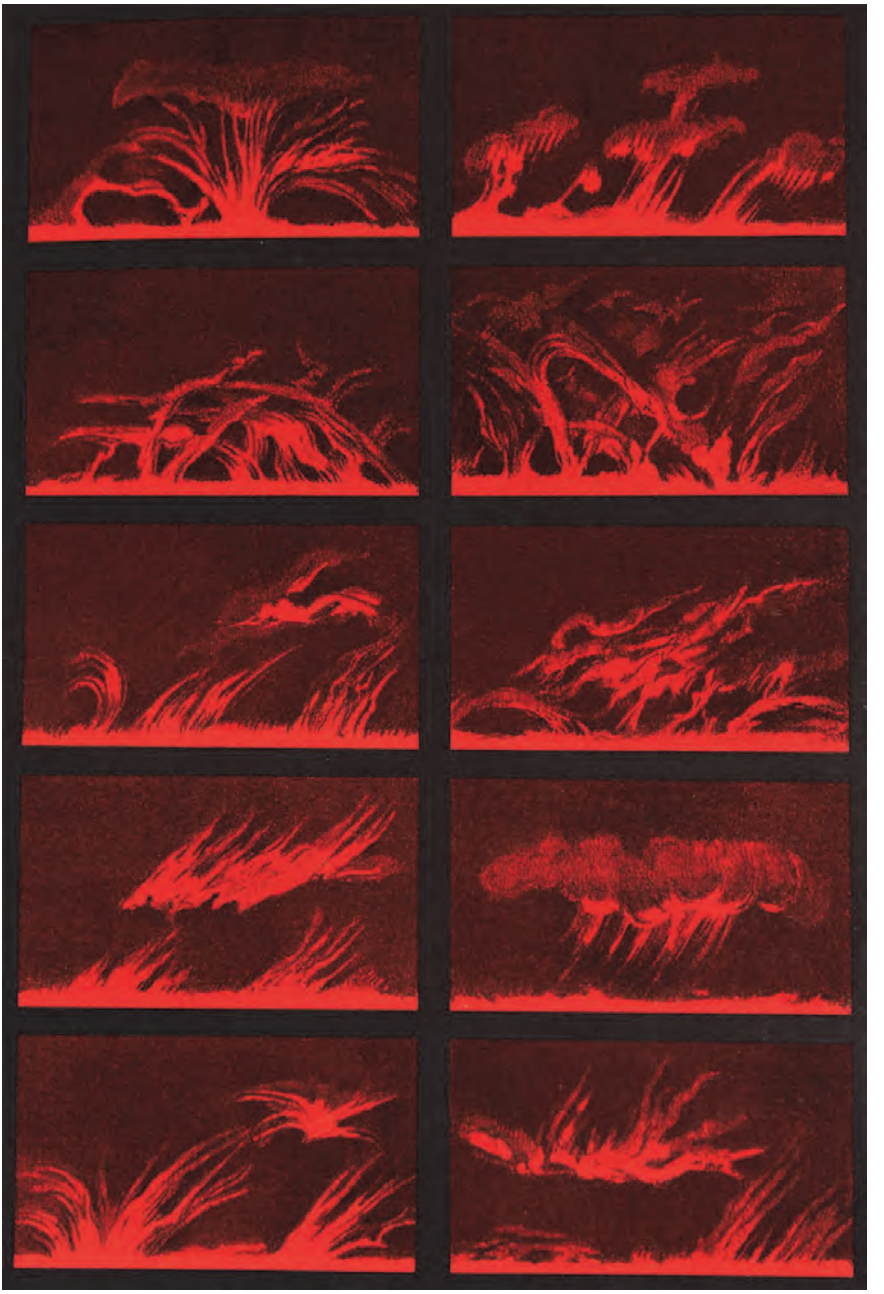


Fig. 2. Contrasted reproduction of one of the plate from the Secchi's book "Le Soleil" 1877 to give a sample of his gallery of drawings of prominences. Note the chromosphere at the bottom of each drawing, with spicules protruding.

the solar chromosphere which is not yet understood even with the extended today space-born experiments and numerical simulations.

Prominent scientists like the father Angelo Secchi are missing in order to progress not only in solar and helio-physics, but as well, to deepen our understanding of Mother Nature.

## References

1. Secchi, Angelo, Le Soleil, Paris (1877).
2. Secchi, Angelo, Il Sole, Firenze (1884).

# Modeling anisotropic magnetized compact objects

D. Alvear Terrero[a], V. Hernández Mederos[a], S. López Pérez[a,b],
D. Manreza Paret[b,c], A. Pérez Martínez[a,*] and G. Quintero Angulo[b]

[a] *Instituto de Cibernética, Matemática y Física,*
*Calle E esq 15 No. 309, Vedado, La Habana, 10400, Cuba.*

[b] *Facultad de Física, Universidad de la Habana,*
*San Lázaro y L, Vedado, La Habana, 10400, Cuba*

[c] *Instituto de Ciencias Nucleares, Universidad Nacional Autónoma de México,*
*Apartado Postal 70-543, CdMx 04510, Mexico.*

[*] *aurora@icimaf.cu*

Magnetic fields affect the structure of compact objects. The anisotropy produced by the magnetic field in the pressures suggests the necessity of using structure equations considering the axial symmetry of the magnetized system. In this work, we propose a model that generalizes the Tolman-Oppenheimer-Volkoff equations for the magnetized case and discuss some preliminary results for WDs equations of state. Our calculations are based on the $\gamma$-metrics where the parameter $\gamma$ relates the deformation with the anisotropy in the pressures.

*Keywords*: White dwarfs; magnetic fields; general relativity; anisotropic structure equations.

## 1. Introduction

The influence of a magnetic field is relevant in compact objects. Measurements of periods and spin down of soft-gamma repeaters (SGR) and X-ray luminosities of anomalous X-ray pulsars (AXP), support the idea of the existence of magnetars[1] which exhibit surface magnetic fields as large as $10^{14} - 10^{16}$ G[2]. However, the interior magnetic fields can not be observed directly, so one can only estimate their bounds with theoretical models based on macroscopic and microscopic analysis. The maximum magnetic fields estimated for magnetized white dwarfs are around $10^{13}$ G[4] and about $5 \times 10^{18}$ G for quark stars[3].

The action of a magnetic field in a fermion system is breaking the spherical symmetry, which gives rise to an anisotropy in the quantum-statistical average of the energy-momentum tensor. This produces a pressure splitting in two components, one along the magnetic field (the parallel pressure $P_{\parallel}$) and another in the transverse direction (the perpendicular pressure $P_{\perp}$), $T^{\mu}_{\nu} = \mathrm{diag}(\rho, -P_{\perp}, -P_{\perp}, -P_{\parallel})$, $\rho$ is the energy density[5]. Consequently, the system composed by fermions under the action of a uniform magnetic field has an anisotropic equation of state (EoS)[5]. This form of the energy-momentum tensor suggests that when modeling the structure of magnetized compact objects, one should consider axial symmetry.

Our first attempt addressing this issue was to consider a cylindrical coordinates $(t, r, \phi, z)$ metric to obtain Einstein field equations following the procedures of Trendafilova and Fulling[6]. In order to do so and for simplicity, we assumed that all the magnitudes in the model depend only on the radial coordinate $r$. This

approximation allowed us to obtain upper limits for the values of the magnetic field that compact stars can sustain, but compromised the computation of a total mass [3,4].

Then, in order to be able to determine the total masses, we went back to spherical coordinates. First, we would like to remark that anisotropies in the energy-momentum tensor are not new to spherical symmetry, which admits tensors of the form $T^{\mu}_{\nu} = \mathrm{diag}(\rho, -p_r, -p_t, -p_t)$, where $p_r$ is a radial pressure and $p_t$ is a tangential one [7,8]. However, this is not compatible with the anisotropies due to magnetic fields. Therefore, in this work, we take into account a generalized version of spherical metric, with a $\gamma$ parameter associated with the deformation of the stars. We propose an ansatz that relates it with the perpendicular and the parallel pressures [9]. This allows us to obtain structure equations that generalizes the Tolman-Oppenheimer-Volkoff (TOV) equations to the magnetized case, where the mass is computed as for spheroidal objects. In Section 2, we present the model and in Section 3 illustrate its solutions with EoS for magnetized white dwarfs.

## 2. $\gamma$-Metric and structure equations for magnetized compact objects

In Refs. 10–12 it is shown that a deformed compact object with axial symmetry can be described by the metric

$$ds^2 = -\left[1 - 2M(r)/r\right]^{\gamma} dt^2 + \left[1 - 2M(r)/r\right]^{-\gamma} dr^2 + r^2 \sin\theta d\phi^2 + r^2 d\theta^2, \qquad (1)$$

where $\gamma = z/r$ parametrizes the polar radius in terms of the equatorial one and $M(r)$ is the mass of the star. Note that the Schwarzschild metric is recovered for $\gamma = 1$.

As we are interested in the effects on the structure equations coming from the magnetic field and the related anisotropy and using smaller central pressures in TOV equations leads to smaller radii, we propose to interpret $\gamma$ as the ratio between the central pressures, parallel and transverse to the magnetic field respectively [9],

$$\gamma = \frac{P_{\parallel_0}}{P_{\perp_0}}. \qquad (2)$$

Then, since the parallel pressure has its maximum central value at $z = 0$ and goes to zero at the surface, we assume it depends just on the $z = \gamma r$ coordinate. The perpendicular pressure, on the contrary, is zero at $r = 0$, therefore depending on the radial coordinate.

Taking into account these considerations and using the mass of a spheroid to compute the star's mass, we obtain the following structure equations

$$\frac{dM}{dr} = r^2 \frac{(E_\parallel + E_\perp)}{2}\gamma, \tag{3a}$$

$$\frac{dP_\parallel}{dz} = \frac{1}{\gamma}\frac{dP_\parallel}{dr} = -\frac{(E_\parallel + P_\parallel)[\frac{r}{2} + r^3 P_\parallel - \frac{r}{2}(1-\frac{2M}{r})^\gamma]}{\gamma r^2 (1-\frac{2M}{r})^\gamma}, \tag{3b}$$

$$\frac{dP_\perp}{dr} = -\frac{(E_\perp + P_\perp)[\frac{r}{2} + r^3 P_\perp - \frac{r}{2}(1-\frac{2M}{r})^\gamma]}{r^2 (1-\frac{2M}{r})^\gamma}, \tag{3c}$$

which describe the mass variation and pressures with the spatial coordinates $r, z$ for an anisotropic magnetized compact object. The solutions of this system are computed similarly to how it is done usually for the TOV equations. Starting from a central point with $E_0 = E(r=0)$, $P_{\parallel_0} = P_\parallel(z=0)$ and $P_{\perp_0} = P_\perp(r=0)$ taken from the EoS, the structure equations are integrated until reaching the surface, where $P_\parallel(Z) = 0$ and $P_\perp(R) = 0$. In practice, this condition is established by the lower central pressure, which determines the value of the corresponding radius ($R$ if $P_{\perp_0}$ and $Z$ if $P_{\parallel_0}$), from where the other radius can be computed by means of $\gamma$ and the mass of the star is set to $M = M(R)$.

The difference of the integration process with respect to TOV solution lies in the manner of computing the energy density from the EoS at $r + dr$, since Eqs. (3) are updated taking into account $E = E_\parallel(P_\parallel(r))$ and $E = E_\perp(P_\perp(r))$ respectively, while the expression for the mass must include information from both, $E_\parallel$ and $E_\perp$, in order to consider the mass density anisotropy, so it is updated with its average value $E = (E_\parallel + E_\perp)/2$.

## 3. Results and important remarks

To validate the model we show some results using the equations of state given in Ref. 15 for magnetized white dwarfs for magnetized WDs with a carbon/oxygen composition at fixed values of the magnetic field, $10^{12}$ G, $10^{13}$ G and $10^{14}$ G, compared to the non-magnetized solution (Fig. 1).

Contrary to the case of Neutron Stars and Quark Stars[13], where the energy density and pressures are of the same order, WDs energy densities are three orders higher than the pressures. Therefore, the value delimiting when the Maxwell term becomes relevant is determined by the pressure, with $P_\perp^B$ comparable to the statistical pressures for magnetic fields higher than $1.78 \times 10^{11}$ G.[a] Anyhow, we tackle two situations with respect to the Maxwell contribution to the pressures and the energy density, $P_\perp^B = E^B = -P_\parallel^B = B^2/8\pi$. In the first case we neglect this term (upper panels of Fig. 1) while the second one is done considering it (lower panels).

[a]The magnetic energy density becomes relevant at $10^{14}$ G[14].

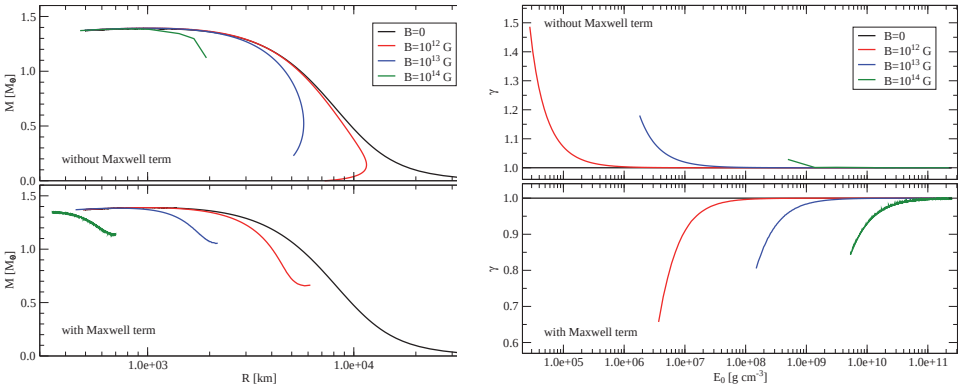



Fig. 1. Left: Mass versus the equatorial radius $R$ and the polar radius $Z$. Right: $\gamma$ parameter as a function of the central density of the magnetized white dwarfs.

The left graph on Fig. 1 displays the mass versus the equatorial radius. At the highest central densities and smaller radii, the masses reach values close to the Chandrasekhar limit [16,17], and both radius are equal and quite similar to the radius of the $B = 0$ case.

Analyzing the magnetized solution at lowest densities and biggest radii, we note on the right graph of Fig. 1 that for a certain value of mass and energy density, the case without (with) Maxwell term features $\gamma > 1$ ($\gamma < 1$), i.e. the polar radius is higher (lower) than the equatorial one. So, the corresponding star is a prolate (oblate) object, which gives us information about the anisotropy in the compact stars with those densities.

Also, it is important to remark that when $B = 0$, the model automatically yields $P_\perp = P_\parallel$ and $\gamma = 1$. This means that we recover the spherical TOV equations from Eqs. (3) and thus, the standard non-magnetized solution for the structure of compact objects.

Our results point out that the magnetic fields effects become relevant in the low density regime and can be practically neglected for high densities. Even more, such effects are more relevant to the deformation of the magnetized compact objects than to their masses, at least when we study magnetized white dwarfs. Also, note that once again, we do not obtain masses above the Chandrasekhar limit.

As $\gamma$-structure equations are general, they can be useful to study other types of magnetized compact objects.

## Acknowledgments

D.A.T, S.L.P, D.M.P, A.P.M and G.Q.A have been supported by the grant No. 500.03401 and V.H.M by the grant No. 500.03501, both of PNCB-MES, Cuba. D.M.P has been also supported by a DGAPA-UNAM fellowship. A.P.M thanks the LOC of Marcel Grossman meeting and Prof. Remo Ruffini and Jorge Rueda that made possible for her to attend this conference.

## References


1. Woods, P.M. & Thompson, C.: 2006. *Soft gamma repeaters and anomalous X-ray pulsars: magnetar candidates*, pages 547–586.
2. R. C. Duncan and C. Thompson, Astrophys. J. **392** (1992) L9. doi:10.1086/186413.
3. D. Manreza Paret, J. E. Horvath and A. Pérez Martínez, Res. Astron. Astrophys. **15**, no. 7, 975 (2015) doi:10.1088/1674-4527/15/7/005, arXiv:1407.2280 [astro-ph.HE].
4. D. M. Paret, J. E. Horvath and A. P. Martínez, Res. Astron. Astrophys. **15**, no. 10, 1735 (2015) doi:10.1088/1674-4527/15/10/010, arXiv:1501.04619 [astro-ph.HE].
5. M. Chaichian, S. S. Masood, C. Montonen, A. Perez Martinez and H. Perez Rojas, Phys. Rev. Lett. **84**, 5261 (2000) doi:10.1103/PhysRevLett.84.5261, arXiv:hep-ph/9911218.
6. C. S. Trendafilova and S. A. Fulling, Eur. J. Phys. **32**, 1663 (2011) doi:10.1088/0143-0807/32/6/020, arXiv:1101.4668 [gr-qc].
7. K. Dev and M. Gleiser, Gen. Rel. Grav. **34** (2002) 1793 doi:10.1023/A:1020707906543, arXiv:astro-ph/0012265.
8. T. Harko and M. K. Mak, Annalen Phys. **11**, 3 (2002) doi:10.1002/1521-3889(200201)11:1<3::AID-ANDP3>3.0.CO;2-L, arXiv:gr-qc/0302104.
9. D. Alvear. Terrero, V. Hernandez. Mederos, S. Lopez. Pérez, D. Manreza. Paret, A. Pérez Martínez and G. Quintero. Angulo, arXiv:1807.09943 [astro-ph.HE].
10. O. Zubairi and F. Weber, doi:10.1142/9789814699662_0002.
11. O. Zubairi, D. Wigley and F. Weber, Int. J. Mod. Phys. Conf. Ser. **45** (2017) 1760029. doi:10.1142/S2010194517600291.
12. L. Herrera, F. M. Paiva and N. O. Santos, J. Math. Phys. **40** (1999) 4064 doi:10.1063/1.532943, arXiv:gr-qc/9810079.
13. J. M. Lattimer and M. Prakash, Phys. Rept. **442** (2007) 109, doi:10.1016/j.physrep.2007.02.003, arXiv:astro-ph/0612440.
14. D. Alvear. Terrero, D. Manreza. Paret and A. Pérez Martínez, Int. J. Mod. Phys. D **27** (2017) no.02, 1850016. doi:10.1142/S0218271818500165.
15. D. Alvear Terrero, M. Castillo García, D. Manreza Paret, J. E. Horvath and A. Pérez Martínez (2015), Astron. Nachr., 336: 851-855. doi:10.1002/asna.201512236.
16. S. Chandrasekhar. ApJ, **74** (1931) 81. doi:10.1086/143324.
17. S. L. Shapiro y S. A. Teukolsky. Black holes, white dwarfs, and neutron stars: The physics of compact objects (1983).

# Constraint on the equation of state from the quasi-periodic oscillations in giant flare

Hajime Sotani[1,*], Kei Iida[2], and Kazuhiro Oyamatsu[3]

[1]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[2]*Kochi University, 2-5-1 Akebono-cho, Kochi 780-8520, Japan*
[3]*Aichi Shukutoku University, 2-9 Katahira, Nagakute, Aichi 480-1197, Japan*
[*]*E-mail: sotani@yukawa.kyoto-u.ac.jp*

We examine crustal torsional oscillations, newly taking into account the effect of the pasta structure. We find from eigenmode analyses for various models of the equation of state of uniform nuclear matter that the fundamental frequencies of such oscillations are almost independent of the incompressibility of symmetric nuclear matter $K_0$, but strongly depend on the slope parameter of the nuclear symmetry energy $L$. On the other hand, we also find that the frequencies of the 1st overtones depend strongly on not only $L$ but also $K_0$. By comparing the resultant frequencies to the quasi-periodic oscillations observed in the giant flares, we can constrain the values of $L$ and $K_0$. Furthermore, considering the constraints on $K_0$ obtained from the terrestrial nuclear experiments, we can successfully make a more severe constant on $L$.

*Keywords*: Equation of state, neutron star crust, shear oscillations.

## 1. Introduction

Neutron stars would be produced via supernova explosions at the last moment of massive stars. The density inside the star significantly exceed the normal nuclear density, the magnetic field inside/around the star can become very strong, and the gravitational field is also so strong. The neutron stars, thus, are good candidate to probe physics under such extreme conditions. In order to extract the interior information of stars, the asteroseismology is very powerful technique. In practice, it would be possible to probe the stellar mass, radius, equation of state (EOS) of dense matter, and theory of gravity in the strong-field regime via future observations of gravitational waves[1–3].

The possible oscillations of neutron stars are observed as quasi-periodic oscillations (QPOs) in the giant flares from the soft-gamma repeaters[4]. In order to theoretically explain these observed QPOs, several studies have been done in terms of crustal torsional oscillations[5–9] or magnetic oscillations[10,11]. The magnetic oscillations strongly depend on uncertain magnetic field structure, field strength inside the star, and uncertain EOS in core region if magnetic fields penetrate the core. On the other hand, the crustal torsional oscillations are confined into the crust region because the shear modulus inside the core should be zero. With such reasons, we focus on the crustal torsional oscillations in the present study by especially taking into account the effect of presence of cylindrical nuclei as well as the phase of spherical nuclei. Then, comparing the observed QPO frequencies to the frequencies of crustal torsional oscillations, we can constrain the nuclear saturation parameters, which are important parameters for describing the neutron star crust. More details of this study can be seen in Ref. 9.

## 2. Crust in equilibrium

To avoid the uncertainties in the EOS of core region, we construct the crust in equilibrium by integrating Tolman-Oppenheimer-Volkoff equations from the stellar surface inward to the bottom of the curst, assuming the stellar mass $M$ and radius $R$. The bulk energy per baryon of uniform nuclear matter at zero temperature for any EOS can be expanded as a function of baryon number density and neutron excess with five expansion parameters, i.e., the saturation density $n_0$, the saturation energy, incompressibility $K_0$ of symmetric nuclear matter, symmetry energy at $n_0$, and the slope parameter $L$. Among five parameters, $K_0$ and $L$ are more difficult to be constrained by terrestrial nuclear experiments, because such parameters are higher order coefficients with respect to the change in density from $n_0$. In the present study, we adopt the phenomenological EOS, where $K_0$ and $L$ are considered as free parameters and the other parameters are determined in such a way as to reproduce the empirical nuclear data[12,13]. The shear modulus is another important property describing the shear oscillations. Here, we adopt the standard shear modulus in the phase of spherical nuclei[14] together with that in the phase of cylindrical nuclei[15], where we also adopt the results by Ref. 16 as the effect of neutron superfluidity inside the phase of spherical nuclei, while the effect of it inside the phase of cylindrical nuclei is considered by introducing a parameter of $N_s/N_d$ with the superfluid neutron density $N_s$ and the dripped neutron density $N_d$. At last, to construct the crust in equilibrium, one has to choose $M$, $R$, $K_0$, $L$, and $N_s/N_d$ in the phase of cylindrical nuclei.

## 3. Torsional oscillations and comparison to the observations

In the left panel of Fig. 1, the calculated frequencies of fundamental torsional oscillations with various values of $\ell$ are shown as a function of $L$ together with the QPO frequencies observed in SGR 1806–20 for the neutron star model with

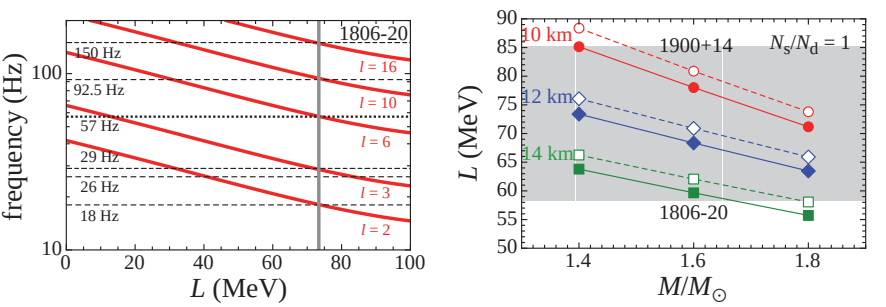



Fig. 1. Comparison of the low-laying QPOs observed in SGR 1806–20 with the fundamental frequencies of crustal torsional oscillations with various values of $\ell$ in the left panel[9]. For various stellar models, the optimal value of $L$ obtained by the comparison of QPOs in SGR 1806–20 and in 1900+14 with the fundamental torsional oscillations in the right panel[9].

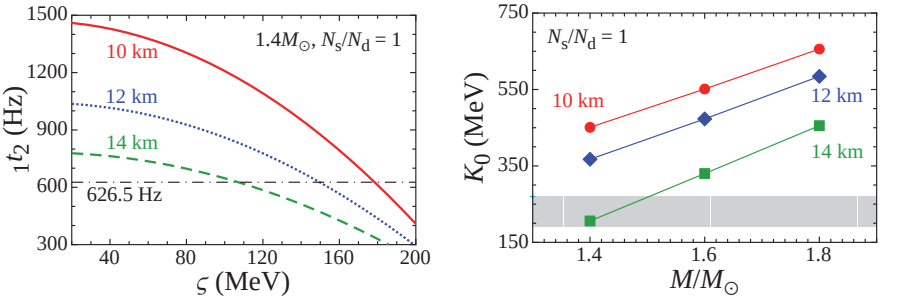



Fig. 2. The frequencies of 1st overtones for the $1.4M_\odot$ neutron star models with different stellar radii in the case of $N_s/N_d = 1$ in the phase of cylindrical nuclei are compared with the 626.5 Hz QPO in SGR 1806 (left panel), while the value of $K_0$ constrained by the combination of constraints on $L$ from the fundamental torsional oscillations and those on ς from the 1st overtones for the various stellar models (right panel)[9].

$M = 1.4M_\odot$ and $R = 12$ km. From this figure, one can see that the QPO frequencies except for the 26 Hz can be identified by the crustal torsional oscillations with $L \simeq 73.4$ MeV. We remark that the 26 Hz QPO may be identified by the torsional oscillations confined in the phase of bubble structure[8]. In a similar way, the optima value of $L$ for identifying the observed QPOs with the fundamental crustal torsional oscillations is shown in the right panel of Fig. 1 for various stellar models, where the filled and open marks denote the results for identifying the QPOs observed in SGR 1806–20 and 1900+14, respectively, and the painted region is allowed region of $L$.

On the other hand, the frequencies of overtones depend on the crust thickness, which depends on $L$ and $K_0$[17]. As a result the frequencies of overtones depend on $L$ and $K_0$. We successfully find that the frequencies of the 1st overtones can be expressed well with a combination of $L$ and $K_0$[9] as

$$\varsigma \equiv \left(K_0^4 L^5\right)^{1/9}. \tag{1}$$

In the left panel of Fig. 2, we show the 1st overtone frequencies for the $1.4M_\odot$ stellar model with different radius as a function of ς. By identifying the QPO frequencies observed in SGR 1806–20 with the 1st overtones, one can determine the optimal value of ς, with which one can determined the value of $K_0$ by combining the results obtained from the identification of the QPO frequencies with the fundamental oscillations as shown in Fig. 1. Such a value of $K_0$ is shown in the right panel of Fig. 2, where the value of $K_0$ constrained from terrestrial nuclear experiments also denote by the painted region[18]. By considering the both constraints from the observations of QPOs and terrestrial nuclear experiments, the preferable stellar model as SGR 1806–20 is relatively low mass and larger radius neutron star, such as $M = 1.2 - 1.4M_\odot$ with $R = 13$ km or $M = 1.4 - 1.5M_\odot$ with $R = 14$ km. With this constraint on the stellar model for SGR 1806–20, we further constrain the value of $L$ from the right panel of Fig. 1, i.e., $L \simeq 58 - 73$ MeV, which is consistent with the terrestrial nuclear experiments[19].

## 4. Conclusion

In this study, we successfully constrain the value of $L$ and $K_0$ by identifying the low-laying QPOs and higher QPO observed in giant flares with the fundamental oscillations and 1st overtones of crustal torsional oscillations. The constraint on $K_0$ obtained from terrestrial nuclear experiments makes a constraint on the stellar model and further constraint on $L$.

## Acknowledgments

This work was supported in part by Grant-in-Aid for Scientific Research (C) (No. 17K05458) provided by JSPS.

## References

1. N. Andersson and K. D. Kokkotas, Phys. Rev. Lett. **77**, 4134 (1996).
2. H. Sotani, K. Kohri, and T. Harada, Phys. Rev. D **69**, 084008 (2004).
3. H. Sotani and K. D. Kokkotas, Phys. Rev. D **70**, 084026 (2004).
4. A. L. Watts and T. E. Strohmayer, Adv. Space Res. **40**, 1446 (2006).
5. A. W. Steiner and A. L. Watts, Phys. Rev. Lett. **103**, 181101 (2009).
6. H. Sotani, K. Nakazato, K. Iida, and K. Oyamatsu, Phys. Rev. Lett. **108**, 201101 (2012).
7. H. Sotani, K. Nakazato, K. Iida, and K. Oyamatsu, Mon. Not. R. Astron. Soc. **428**, L21 (2013); **434**, 2060 (2013).
8. H. Sotani, K. Iida, and K. Oyamatsu, Mon. Not. R. Astron. Soc. **464**, 3101 (2017).
9. H. Sotani, K. Iida, and K. Oyamatsu, Mon. Not. R. Astron. Soc. **479**, 4735 (2018).
10. H. Sotani, K. D. Kokkotas, and N. Stergioulas, Mon. Not. R. Astron. Soc. **375**, 261 (2007); **385**, L5 (2008).
11. M. Gabler, P. Cerdá-Durán, N. Stergioulas, J. A. Font, and E. Müller, Mon. Not. R. Astron. Soc. **421**, 2054 (2012).
12. K. Oyamatsu and K. Iida, Prog. Theor. Phys. **109**, 631 (2003).
13. K. Oyamatsu and K. Iida, Phys. Rev. C **75**, 015801 (2007).
14. T. Strohmayer, H. M. Van Horn, S. Ogata, H. Iyetomi, and S. Ichimaru, Astrophys. J. **375**, 679 (1991).
15. C. J. Pethick and A. Y. Potekhin, Phys. Lett. B **427**, 7 (1998).
16. N. Chamel, Phys. Rev. C **85**, 035801 (2012).
17. H. Sotani, K. Iida, and K. Oyamatsu, Mon. Not. R. Astron. Soc. **470**, 4397 (2017).
18. E. Khan and J. Margueron, Phys. Rev. C **88**, 034319 (2013).
19. W. G. Newton et al., Eur. Phys. J. A **50**, 41 (2014).

# Contributions to neutron star's tidal deformability from the low density equation of state

A. M. Kalaitzis, T. F. Motta and A. W. Thomas

*CSSM and ARC Centre of Excellence for Particle Physics at the Terascale, Department of Physics, University of Adelaide SA 5005 Australia*

The low density contribution to the tidal deformability and moment of inertia of a neutron star are calculated via various well known equations of state. The contributions to the moment of inertia are directly calculated, whilst the tidal deformability's are constructed through comparing an equation of state with a fit with the low density region removed. With the recent measurement of GW170817[1,2] providing constraints on the tidal deformability, it is very important to understand what features of the equation of state have the biggest effect on it.

*Keywords*: Neutron Stars, Tidal Deformability, Low Density, Equation of State, Nuclear Matter, GW170817.

## 1. Introduction

Neutron stars (NS) being the densest known repositories of nuclear matter contain fascinatingly complex phase changes between their regions of finite and infinite nuclear matter, more commonly known as their layers, denoted crust and core[3–5]. The crust poses challenges with regards to modelling because of the variety of finite nuclei that may appear and complex phase structures that can occur[6].

On the other hand, after the recent GW170817 measurement[1,2], there has been a tremendous push to calculate the tidal deformability (TD) and the moment of inertia (MOI)[7–13] across a variety of equations of state (EOS). This is due to the bounds produced on the TD in Ref 1, 2 of $\Lambda_{1.4} = 190^{+390}_{-120}$, which has recently been analysed via the I-Love-Q relations in Ref. 14 to produce a constraint on the MOI of $\overline{I} = 11.10^{+3.64}_{-2.28}$.

In this paper we explore the proportional contribution that the low density region (LDR) of the EOS makes to the MOI and TD.

## 2. Theoretical Framework

We will be using the well known EOS shown in Fig. 1 acquired from the online data base in Ref. 15 which details the models in question. In order to calculate the NS profile, ordinarily one would match the model EOS description of the core to some low density EOS in order to account for the NS crust or model it directly[4,6]. We have fit to each EOS using a polytropic of the form given in equation 1 where $A, B, C$ and $D \in \mathbb{R}$. This will act as the EOS without the LDR modelling, with the original EOS data from Ref. 15 containing LDR modelling. As seen in the enlargement in Fig. 1

$$P(\epsilon) = A\epsilon^B + C\epsilon^D. \tag{1}$$

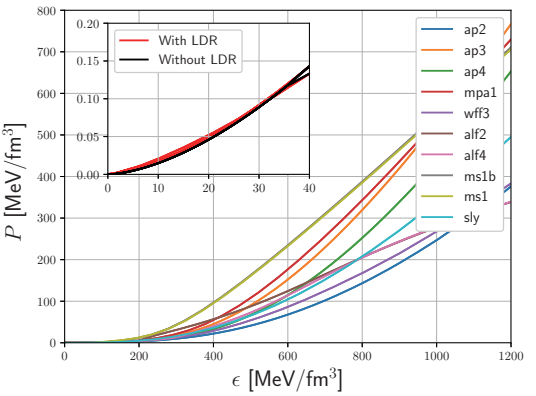



Fig. 1. EOS (pressure $P$ vs energy density $\epsilon$) for all the models[15], nucleon only. Two EOS for each EOS, one with the LDR and one produced from the polytropic fit. The enlargement shows the LDR with all EOS and their fits displayed (There is much overlap between the EOS for the LDR).

We will use these fits to analyse the percentile contributions to the TD via direct comparison of their values. This is done as the TD is defined mathematically as a boundary point and not something cumulative (e.g. integral)[16,17], so that the direct calculation of its LDR contribution is impossible. This is equivalent to modelling a core only NS, consisting of only infinite nuclear matter. The LDR contribution to the MOI may be calculated directly as it is defined as a cumulative integral[18,19].

The NS masses are obtained by numerically integrating the Tolman-Oppenheimer-Volkoff (TOV) equations[20]

$$\frac{\mathrm{d}P}{\mathrm{d}r} = -\frac{(P(r)+\epsilon(r))(m(r)+4\pi r^3 P(r))}{r(r-2m(r))}, \tag{2}$$

$$\frac{\mathrm{d}m}{\mathrm{d}r} = 4\pi r^2 \epsilon(r), \tag{3}$$

until $P(r = R) = 0$ ($R$ the radius of the NS). This produces the mass radius relations given in Fig. 2. We will define the variation of a parameter $\alpha$ as $\Delta\alpha/\alpha$, which is reported as the percentage difference of that parameter when compared to its LDR removed counterpart. This is calculated via (MOI variation $\Delta I/I$ is defined differently as seen in Refs. 12, 21)

$$\frac{\Delta\alpha}{\alpha} = \frac{(\alpha_{\text{with LDR}} - \alpha_{\text{Without LDR}})}{\alpha_{\text{with LDR}}}. \tag{4}$$

Note that this requires one parameter to be fixed. For example the compactness parameter $\beta = M/R$, we would fix the mass and look at the variations in the radius or visa versa. The MOI $I$ is calculated through the formula in Refs. 18, 19, with $\Delta I/I$ defined in Refs. 12, 21.

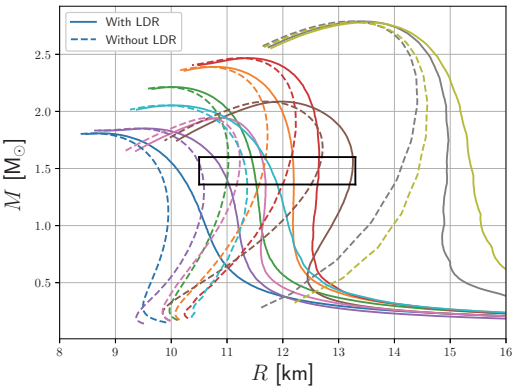



Fig. 2. Mass radius relation for EOS given in Fig. 1. The box represents the 90% confidence constrained mass region given from the GW170817 measurement[1,2]. Mass radius curves retain name and colour convention as in Fig. 1 with dashed versions being the fits.

The tidal Love number, $k_2^{\text{tidal}}$, TD, $\Lambda$, and binary tidal deformability, $\widetilde{\Lambda}$, equations are verbatim those seen in Refs. 16, 17, 11. For future reference, the TD and binary tidal deformability are calculated via

$$\Lambda = \frac{2}{3} k_2^{\text{tidal}} \beta^{-5}, \tag{5}$$

$$\tilde{\Lambda}^{\text{Tid}} = \frac{16}{13} \frac{(12q+1)\Lambda_1 + (12+q)q^4\Lambda_2}{(1+q)^5}, \tag{6}$$

where $q = m_2, m_1$ and $m_2 \leq m_1$. With these equations we can now calculate our desired quantities and the percentage contributions from the LDR.

## 3. Results and Discussion

Numerically solving the MOI equations[11,18,19,21] produces Fig. 3 which shows the proportion of crustal MOI $\Delta I$ over the total MOI $I$. It should be noted that for any realistic star[21] of solar mass

$$0.8 M_\odot \leq M \leq 1.92 M_\odot, \tag{7}$$

we find $|\Delta I/I| < 17.5\%$, with this percentage decreasing rapidly as the mass increases. In particular, a star of solar mass $1.4 M_\odot$ has $|\Delta I/I| < 7.5\%$.

In order to calculate the LDR contribution to the tidal deformability, we simply solve the equations in Refs. 16, 17 for each set of mass radius curves in Fig. 2 and then make use of equation 4 which produces Fig. 4. We find that $|\Delta\Lambda/\Lambda| < 15\%$ for all NS within equation 7's mass range. However, it is important to note that if we disregard the three EOS that lie outside the 90% confidence interval from Refs. 1, 2

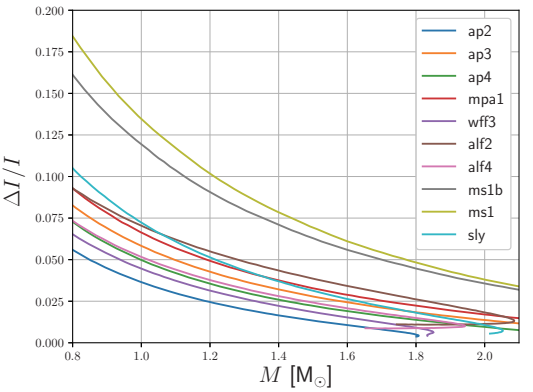



Fig. 3. Percentage of the MOI due to the LDR, scaled via the neutron star's total MOI. Only shown for the EOS with the LDR.

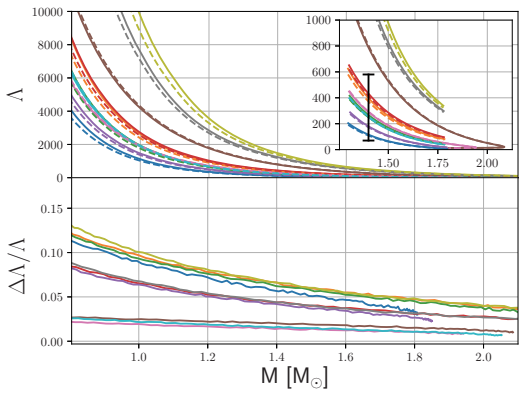



Fig. 4. Percentage of the TD due to the LDR, scaled via the neutron star's total TD. The enlargement shows a shorter range with the black bar representing the bound $\Lambda_{1.4} = 190^{+390}_{-120}$.

in Fig. 4, then our new bounds become

$$|\Delta I/I| < 9.0\%, \qquad |\Delta\Lambda/\Lambda| < 12.0\%, \tag{8}$$

$$|\Delta I/I|_{M=1.4M_\odot} < 3.5\%, \qquad |\Delta\Lambda/\Lambda|_{M=1.4M_\odot} < 7.0\%, \tag{9}$$

within mass range given in Eq. (7). For clarity the tidal Love numbers are displayed in Fig. 5

This happens in spite of the visibly large difference in the mass radius relations (by extension the compactness parameter $\beta = M/R$) and tidal Love numbers shown in Fig. 2 and Fig. 5. The variance in the TD remains relatively low due the large negative power from the $\beta$ term in Eq. (5).

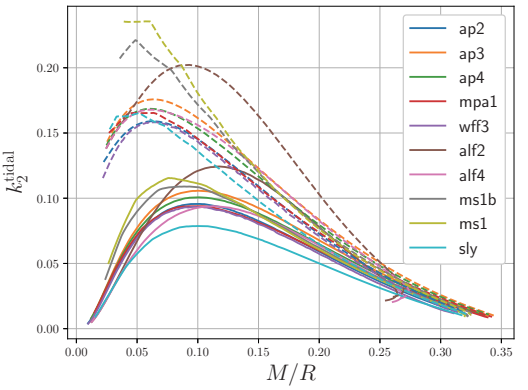



Fig. 5. Tidal Love numbers $k_2^{\text{tidal}}$ keeping same conventions as in Fig. 2 and Fig. 1

Due to the low variance in the TD, by extension the variance in the BTD is of a similar order of magnitude as it depends linearly on $\Lambda_1$ and $\Lambda_2$, with $m_1, m_2$ being its free parameters, i.e, these two parameters offer no more increase in the variance.

These results have been summarised as the following:

- The impact on macroscopic quantities (e.g. mass, radius etc.) from LDR decreases in significance as the mass increases.
- The TD and MOI are affected to just a small percentage via the inclusion of the low density EOS for all stars with mass $M \geq 0.8M_\odot$, and to a greater extent for stars with mass $M \geq 1.4M_\odot$. Percentages are shown in Eqs. (8) and (9).

This is not to say that the LDR's effects are uniformly negligible, as seen in Fig. 2 where the LDR contributes significantly to the radius. With the imminent measurements of the neutron star radii by NICER, this low density EOS modelling will prove incredibly significant in ascertaining correct predictions for the radius.

## 4. Conclusion

In summary we see that despite the LDR's proportionally significant contribution to the radius and percentage of the NS content, its contributions to the TD and MOI are small, especially when considered for a star of $1.4M_\odot$ or greater.

## Acknowledgements

This work was supported by the University of Adelaide and the Australian Research Council through grants DP150103101 and DP180100497 and the Masters (No Honours) academic scholarship.

**References**


1. B. P. Abbott *et al.*, Properties of the binary neutron star merger GW170817 (2018).
2. B. P. Abbott *et al.*, GW170817: Measurements of neutron star radii and equation of state, *Phys. Rev. Lett.* **121**, p. 161101 (2018).
3. F. Weber, Strange quark matter and compact stars, *Progress in Particle and Nuclear Physics* **54**, 193 (2005).
4. H. Sotani, K. Iida and K. Oyamatsu, Probing crustal structures from neutron star compactness, *Monthly Notices of the Royal Astronomical Society* **470**, 4397 (06 2017).
5. N. K. Glendenning, *Compact stars Nuclear physics, paricle physicsand general relativity* (Springer, 2000).
6. N. Chamel and P. Haensel, Physics of neutron star crusts, *Living Reviews in Relativity* **11**, p. 10 (Dec 2008).
7. J. B. Wei, A. Figura, G. F. Burgio, H. Chen and H. J. Schulze, Neutron star universal relations with microscopic equations of state (2018).
8. S. De, D. Finstad, J. M. Lattimer, D. A. Brown, E. Berger and C. M. Biwer, Tidal Deformabilities and Radii of Neutron Stars from the Observation of GW170817, *Phys. Rev. Lett.* **121**, p. 091102 (2018).
9. S. Han and A. W. Steiner, Tidal deformability with sharp phase transitions in (binary) neutron stars (2018).
10. Z. Carson, A. W. Steiner and K. Yagi, Constraining nuclear matter parameters with GW170817, *Phys. Rev.* **D99**, p. 043010 (2019).
11. T. Zhao and J. M. Lattimer, Tidal Deformabilities and Neutron Star Mergers, *Phys. Rev.* **D98**, p. 063020 (2018).
12. Z. Qian, R. Y. Xing and B. Y. Sun, Moments of inertia of neutron stars in relativistic mean field theory: the role of the isovector scalar channel (2018).
13. P. Landry and B. Kumar, Constraints on the moment of inertia of PSR j0737-3039a from GW170817, *The Astrophysical Journal* **868**, p. L22 (nov 2018).
14. P. Landry and B. Kumar, Constraints on the moment of inertia of PSR J0737-3039A from GW170817, *Astrophys. J.* **868**, p. L22 (2018).
15. Observations and models of compact stars, EOS available from `http://xtreme.as.arizona.edu/neutronstars/`.
16. S. Postnikov, M. Prakash and J. M. Lattimer, Tidal Love Numbers of Neutron and Self-Bound Quark Stars, *Phys. Rev.* **D82**, p. 024016 (2010).
17. K. S. Thorne and A. Campolattaro, Non-Radial Pulsation of General-Relativistic Stellar Models. I. Analytic Analysis for L $\geq=$ 2, *Astrophys. J.* **149**, p. 591 (September 1967).
18. J. B. Hartle, Slowly Rotating Relativistic Stars. I. Equations of Structure, *Astrophys. J.* **150**, p. 1005 (December 1967).
19. J. B. Hartle and K. S. Thorne, Slowly Rotating Relativistic Stars. II. Models for Neutron Stars and Supermassive Stars, *Astrophys. J.* **153**, p. 807 (September 1968).
20. J. R. Oppenheimer and G. M. Volkoff, On massive neutron cores, *Phys. Rev.* **55**, 374 (Feb 1939).
21. J. M. Lattimer and M. Prakash, Neutron Star Observations: Prognosis for Equation of State Constraints, *Phys. Rept.* **442**, 109 (2007).

# Hadron-quark phase transition and the QCD phase diagram

Clebson A. Graeff†, Constança Providência** and Débora P. Menezes*

† *Universidade Tecnológica Federal do Paraná, campus Pato Branco*
*Via do Conhecimento, Km 1 CEP 85503-390 Pato Branco – PR, Brazil*
*E-mail:* cgraeff@utfpr.edu.br
** *Departamento de Fsica, Universidade de Coimbra, 3004-516 Coimbra, Portugal*
* *Departamento de Física, Universidade Federal de Santa Catarina,*
*Florianópolis, SC, CP 476, CEP 88.040-900, Brazil*

Different extensions of the Nambu-Jona-Lasinio model, known to satisfy expected QCD chiral symmetry aspects, are used to investigate a possible hadron-quark phase transition at zero temperature and to build the corresponding binodal sections.

*Keywords*: NJL; QCD; Effective models; Binodals.

## 1. Introduction

Effective models remain a good source of information about regions of the QCD phase diagram inaccessible by terrestrial experiments or by LQCD methods[1], providing qualitative results and theoretical insights. The present work intends to help in advancing our knowledge towards some of the regions of the QCD phase diagram through this strategy. When this approach is applied to the study of the transition of hadronic matter to the deconfined quark matter, it is sugested that the QCD phase diagram shows a first order phase transition[2] at high chemical potentials and low temperatures, while from the LQCD perspective, the hadron-quark transition is believed to be a crossover at low chemical potentials and high temperatures. This seemingly contradictory picture can be reconciled by the existence of a critical end point in the phase transition curve, linking the LQCD crossover with the effective model first order transition. This idea is reinforced by experimental results[3] signaling to a first order phase transition and pointing out to the possible existence and location of the critical end point.

Considerations on the phase transition at zero temperature have already been done in many works[4–9], but we do believe the formalism we employ in the present work is more adequate, as the effective models employed here exhibit chiral symmetry in both hadronic and quark phases, which is demanded to take seriously the appearance of the quarkyonic phase[10]. The models used here are all included in the Nambu–Jona-Lasinio (NJL) model framework[11], in order to naturally describe the chiral characteristics of QCD matter.

In Refs. 6 and 9, the hadron-quark phase transition was investigated with the help of two different models, namely, the non-linear Walecka model (NLWM) for the hadronic phase and the MIT bag model for the quark phase. A formalism we understand as a more adequate one was used in Refs. 7, 12, 13 at zero temperature and in Ref. 8 for finite temperatures, all considering NJL-type models for the two

phases. To describe the hadron phase, the standard NJL model with vector interaction is extended to include a scalar-vector channel in order to render the model capable of saturation at low densities. We revisit the approach of Refs. 7 and 8, but applying an extended NJL model for the hadron phase that includes additional channels to achieve a better description of important nuclear bulk properties[14]. A similar extension of the NJL model for hadronic matter has been developed[15,16] with a different choice of interaction channels. Recently, this version was also applied to investigate the hadron-quark phase transition[17], but the quark phase was still described by the MIT bag model. Hence, we describe the hadronic matter with the extended NJL model from Ref. 14 and the quark matter with the NJL model in its SU(2) version in order to check for which parameters a phase transition is possible, considering both symmetric and asymmetric systems. Whenever possible, the binodal sections are obtained.

## 2. Binodals

The QCD phase-diagram is characterised by potentially multiple phases, whose phase separation boundaries are referred as *binodals*[18]. Over those boundaries, the phases from the regions of either side of the boundary can coexist. The binodals may be determined using the Gibbs conditions[6]:

$$\mu_B^Q = \mu_B^H, \qquad T^Q = T^H, \qquad P^Q = P^H, \tag{1}$$

where the indexes $H$ and $Q$ refer to the hadronic and quark phases. The chemical potentials $\mu_B^i$ are obtained from the chemical potentials of the particles of each phase[6].

At a certain fixed temperature ($T = 0$ in the present context), the phase coexistence condition may be obtained by plotting $P^i \times \mu_B^i$, $i = Q, H$, and looking for the

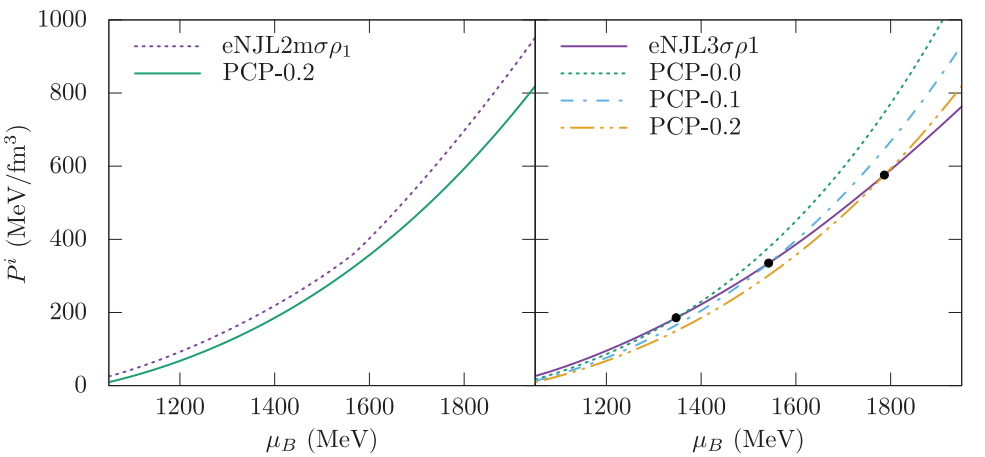



Fig. 1. Combinations of parameter sets for which hadron-quark phase transition is not allowed to happen (left), and combinations for which the transition is allowed to happen (right).

intersection of both curves. In Fig. 1 (left) we display a combination of parameterizations for which there are no intersections, implying that there are no transitions (i.e., the hadron phase is always more stable). In Fig. 1 (right), we compare the results for a hadron parameter set with three quark matter parameter sets, with varying strenght of the vector coupling, obtaining the transitions indicated by the dots.

The conditions of phase coexistence are also important in asymmetric matter and to obtain the binodal sections as a function of the system asymmetry, we use the prescription given in Ref. 4. The isospin chemical potentials are defined as

$$\mu_3^H = \mu_p - \mu_n, \qquad \mu_3^Q = \mu_u - \mu_d, \tag{2}$$

and enforced to be identical according to the Gibbs conditions. The asymmetry parameters of the hadron and quark phases are respectively

$$\alpha^H = (\rho_n - \rho_p)/(\rho_n + \rho_p), \qquad \alpha^Q = 3(\rho_d - \rho_u)(\rho_d + \rho_u), \tag{3}$$

in such a way that $0 \leq \alpha^H \leq 1$ (just nucleons) and $0 \leq \alpha^Q \leq 3$ (just quarks).

The *binodals* are obtained through the determination of hadron and quarks phase pressures for each value of $\mu_B$ and $\mu_3$ (the $\mu_3$ parameter directly controls the proton fraction of both phases): whenever the pressure difference is below 0.1 MeV, we assume that both phases coexist. This procedure leads to the results shown in Fig. 2. The pressures shown for $\alpha = 0$ in the right panel correspond to the intersections marked in Figure 1. Also from this figure, we can clearly see that the increase in strength of the vector coupling causes a substantial modification on the transition point, which reflects in the values of the pressure in the binodal sections.

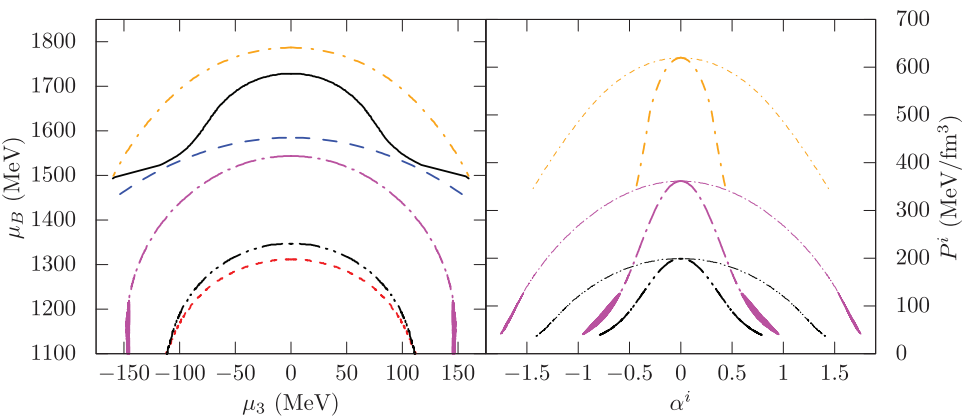



Fig. 2. Baryonic chemical potentials as a function of $\mu_3$ (left) and pressure as a function of asymmetry (right) at the coexistence point for: BuballaR-2 and eNJL2m$\sigma\rho$1 (black line), BuballaR-2 and eNJL3$\sigma\rho$1 (blue long-dashed line), PCP-0.0 and eNJL2m$\sigma\rho$1 (red short-dashed line), PCP-0.0 and eNJL3$\sigma\rho$1 (black double-dot dashed line), PCP-0.1 and eNJL3$\sigma\rho$1 (magenta long-dash dotted line), and PCP-0.2 and eNJL3$\sigma\rho$1 (orange short-dash dotted line).

## 3. Conclusions

We have revisited the study of hadron-quark phase transition at zero temperature with different extensions of the NJL model, which are more appropriate to describe systems where chiral symmetry is an important ingredient. We analysed possible phase transitions from a hadron phase described by an extended NJL model to a quark phase described by the SU(2) NJL with the inclusion of a vector interaction of arbitrary strenght. We have first considered symmetric matter and checked that not all parameterization combinations produce a system in which a phase-transition is favored. Another manifestation of the dependence of the results on the choice of parameters is the range of barionic chemical potentials for which the transition takes place, spanning from around 1300 MeV to around 1800 MeV, indicating a strong parameter dependence. We have next analysed assymetric systems and whenever possible, binodal sections were obtained. Both pressures and chemical potential increase drastically with the increase of the vector interaction strenght in the quark sector. As a next step on this analysis, we plan to expand our results to include finite temperature in the system and obtain the complete binodal sections.

## References


1. H. B. Meyer, PoS **LATTICE2015**, 014 (2016), arXiv:1512.06634 [hep-lat].
2. F. Karsch, Nuclear Physics A **590**, 367 (1995).
3. Y. Akiba *et al.*, ArXiv e-prints (2015), arXiv:1502.02730 [nucl-ex].
4. H. Müller, Nuclear Physics A **618**, 349 (1997).
5. M. Di Toro, *et al.*, Nuclear Physics A **775**, 102 (2006).
6. R. Cavagnoli, C. Providência, and D. P. Menezes, Phys. Rev. C **83**, 045201 (2011).
7. Y. Tsue, J. da Providência, C. Providência, and M. Yamamura, Progress of Theoretical Physics **123**, 1013 (2010).
8. T.-G. Lee, Y. Tsue, J. da Providência, C. Providência, and M. Yamamura, Progress of Theoretical and Experimental Physics **2013**, 013D02 (2013).
9. K. D. Marquez and D. P. Menezes, Journal of Cosmology of Astroparticle Physics **12**, 028 (2017), arXiv:1709.07040 [astro-ph.HE].
10. L. McLerran, PoS **CPOD2014**, 046 (2015).
11. Y. Nambu and G. Jona-Lasinio, Phys. Rev. **122**, 345 (1961).
12. C. Providência, J. M. Moreira, J. da Providência, and S. A. Moszkowski, AIP Conf. Proc. **660**, 231 (2003).
13. C. da Providência, J. da Providência, and S. A. Moszkowski, Int. J. Mod. Phys. **B17**, 5209 (2003).
14. H. Pais, D. P. Menezes, and C. Providência, Phys. Rev. C **93**, 065805 (2016).
15. T. H. Phat, N. T. Anh, and D. T. Tam, Phys. Rev. C **84**, 024321 (2011).
16. N. T. Anh and D. T. Tam, Phys. Rev. C **84**, 064326 (2011).
17. N. T. Anh, Communications in Physics **27**, 71 (2017).
18. H. Müller and B. D. Serot, Phys. Rev. C **52**, 2072 (1995).

# Dark mechanism for nucleation inside old neutron stars

M. Ángeles Pérez-García*

*Department of Fundamental Physics, University of Salamanca, Plaza de la Merced, s/n, E-37008, Spain*
** E-mail: mperezga@usal.es*

Neutron stars are objects where matter and fields can be tested to the extreme. In their interior, ordinary matter with a majority of nucleons ($N$) may coexist with a component of exotic matter that constitutes the so-called dark matter. It is believed that this latter type of matter constitutes nearly 80% of all mater in the currently accepted cosmological paradigm for our Universe. Popular candidates for such a kind of matter belong to extensions of the Standard Model of particle physics. In this contribution we discuss the impact of a component of self-annihilating fraction of dark matter ($\chi$) that may be present in the star. We focus on the microscopics of the energy deposit arising from the annihilation process $\chi\chi \to N\gamma$ in the dense core of a NS. We show that even if a tiny portion of matter inside the NS is dark matter it may trigger critical changes in the equation of state (EoS) of regular hadronic stars, provided its mass, $m_\chi$, and cross-section scattering off nucleons, $\sigma_{\chi N}$, resides in a suitable allowed region of the dark parameter phase space.

*Keywords*: Neutron star; dark matter; nuclear matter; phase transition.

## 1. Introduction

Neutron stars (NSs) are compact objects born in the aftermath of a supernova explosive event[1]. They are believed to be among the densest known objects in our Universe and where extreme conditions determine the region of the matter phase space that can be tested. i.e. cold and dense matter. Namely, in most of the volume of the quasi-spherical object with mass $M_*$ and radius $R_*$ densities are several times that of nuclear saturation density $n_0 \sim 0.16$ fm$^{-3}$. Solutions to the structure equations provided by the theory of General Relativity in the static approximation provide values of the compactness parameter $M_*/R_* \sim 0.1$ which is much larger than for the Sun or any other stellar body. NS typical values for $M_* \sim 1.5 - 2M_\odot$ and $R_* \sim 11 - 12$ km. Thus NSs can be considered to be efficient attractors of a massive $\chi$ dark matter (DM) component from a halo where they may be inmerse. That is the specific case under study in this contribution. So far this kind of elusive matter has been searched for using different techniques including direct searches with terrestrial targets where scattering events off nuclei would allow to measure the corresponding recoiling energies[2]. Another existing effort is that of the so-called indirect searches[3] where emission of secondaries (photons or additional Standard Model particles) may signal the presence of DM. This means for some popular models of DM a dark component that could self-annihilate into gamma rays or any other particle pairs as many reviews on the subject explain[4]. As there is so far no consensus about the fermionic or bosonic nature of the DM candidate $\chi$, we will assume an effective treatment where they are considered as a semiclassic, non relativistic matter component. The validity of our treatment is assured as

we consider a finite fraction of an accreted dark component that is able to exist thermalized in the interior of these dense stars. The combined effect of gravitational capture and multiple scattering suffered by an incoming DM particle traversing the NS volume can be summarized in an expression for the global capture rate[5]

$$C_\chi \simeq 6\times 10^{25}\left(\frac{M_*}{1.5M_\odot}\right)\left(\frac{R_*}{12\,\mathrm{km}}\right)\left(\frac{1\,\mathrm{GeV}}{m_\chi}\right)\left(\frac{\rho_\chi^{ambient}}{0.3\frac{\mathrm{GeV}}{\mathrm{cm}^3}}\right)\left(\frac{\sigma_{\chi N}}{\sigma_0}\right)\ \mathrm{s}^{-1}, \qquad (1)$$

where $\sigma_0 \sim 10^{-44}$ cm$^2$. The typical mean free path of a DM particle inside the core is approximated by $\lambda_\chi^{-1} \sim \sigma_{\chi N} n_N$ with $\sigma_{\chi N}$ the $\chi N$ scattering cross-section and $n_N$ the number density of nucleons inside the NS core. At least one scattering event is needed to capture the DM particle then $\lambda_\chi/R \lesssim 1$ in the NS core at a typical density $\sim 3n_0$ yields $\sigma_{\chi N} \gtrsim 10^{-45}$ cm$^2$. At current level of sentitivity this is allowed for heavy DM candidates with masses beyond the $m_\chi \sim 1$ TeV scale.

## 2. Role of Dark Matter and changes in the NS EoS.

From an accreted component of DM and once it is thermalized with the rest of matter we can assume that a spatial distribution arises under a gaussian form $n_\chi(r) = n_{0,\chi}\exp(\frac{\mathrm{r}}{\mathrm{r_{th}}})^2$ where $r_{\mathrm{th}} = \sqrt{\frac{3k_BT}{2\pi G\rho_N m_\chi}}$ stands for the thermal radius and $\rho_N \simeq m_N n_N$. $G$ and $k_B$ are the gravitational and Boltzmann constants, respectively. We consider that the inner core in the old NS has attained an internal temperature $T \sim 10^5$ K, low with respect to Fermi energies in the baryon sector.

It is in this context that DM is gathered forming a density enhancement in a central stellar region with volume $V_{\mathrm{th}} \sim r_{\mathrm{th}}^3$ and can undergo self-annihilation following a chain of possible reactions into photonic final states $\chi\chi \to N\gamma$. PYTHIA package[6] can simulate this using Monte Carlo techniques and it is obtained that the energy injection due to this kind of events yields $E_\gamma \sim 0.6 m_\chi$. We are interested in the study of the possible consequences of such a microscopic energy transfer at the nuclear and sub-nuclear (quark) level. As we know, a description of nuclear matter from an effective field theory picture based on nucleons (neutrons and protons) is adequate for a limited set of (low energy) thermodynamical conditions. A novel effect, such as the local nucleation of quark droplets could be in principle studied, provided stability conditions for their existence are fulfilled. This intriguing possibility is indeed not new in the literature as pioneering works[7] referred to it early in the 1970's and 1980's and are coined today as the Bodmer-Witten conjecture. However the actual procedure leading to the deconfined state is far from being understood. A stochastic procedure where NS central density (pressure) is increased due to spin-down or mass accretion has been invoked, as well as fluctuations in high temperature young protoneutron stars[8,9].

On a more technical ground, relativistic nucleation theory developed by Lifshitz-Kagan[10] can model phenomenological bubble nucleation by using a relativistic lagrangian that describes the formation of a fluctuation i.e. a spherical bubble of

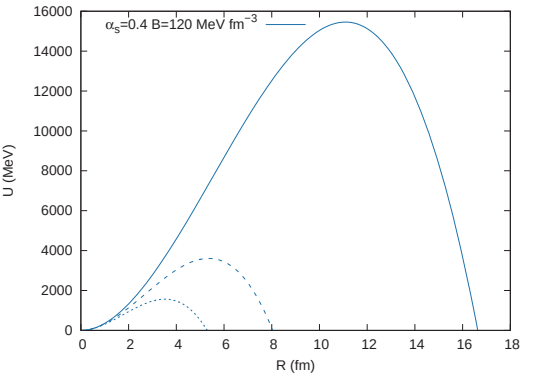



Fig. 1. Potential barrier energy as a function of radius. Solid, dashed and dotted lines depict pressures $P = P_0 + 10$ MeV, $P = P_0 + 20$ MeV, $P = P_0 + 30$ MeV, respectively for each case. We fix $\alpha_s = 0.4$, $B = 120$ MeV fm$^{-3}$, $\sigma = 30$ MeV fm$^{-2}$. $P_0 = 525$ MeV fm$^{-3}$.

mass $M$ and radius $R$,

$$\mathcal{L}(R, \dot{R}) = -M(R)\sqrt{1 - \dot{R}^2} + M(R) - U(R), \tag{2}$$

where $\dot{R} = dR/dt$ is the radial growth rate and $U(R)$ is an effective potential that confines the quark content. The latter depends on the thermodynamical conditions of the medium and it can be written in terms of the leading contributions approximately by

$$U(R) = \frac{4}{3}\pi R^3 n_{\rm Quark}(\mu_{\rm Quark} - \mu_{\rm Had}) + 4\pi\sigma R^2. \tag{3}$$

We label $\mu_{\rm Had}$ ( $\mu_{\rm Quark}$) as the chemical potentials of the metastable hadronic (stable quark) phases of matter at fixed pressure value, $P$. $n_{\rm Quark}$ is the number density of the stable phase and $\sigma$ is the surface tension. Contributions from volume as well as surface terms are the most energetically relevant in the nucleation process although some other corrections can be accounted for as well. This expression may result familiar to the reader as in terrestrial experiments for DM search based on bubble chambers such as PICO[11] another target is used but the spirit is still the same.

Back to our description of both high density phases, this is done in this contribution by a relativistic mean field (RMF) approach including neutrons, protons, electrons and mesonic fields $\sigma, \omega, \rho$ as in Ref. 12 while for the quark deconfined phase we choose the MIT bag model[15] with massive uds quarks for simplicity. The system fulfills electrical charge neutrality and baryon number conservation.

## 3. Results

Examples of U potential barriers can be seen in Fig. 1 for the case of beta-equilibrated matter at T=0 and three values of pressure. Solid, dashed and dotted lines depict $P = P_0 + 10$ MeV, $P = P_0 + 20$ MeV, $P = P_0 + 30$ MeV, respectively. To

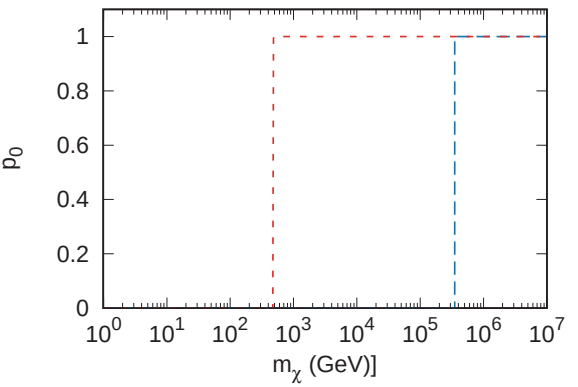



Fig. 2. Tunneling probability as a function of mass of the DM particle. We fix $\alpha_s = 0.4$, $B = 120$ MeV fm$^{-3}$ and $P = 535$ MeV fm$^{-3}$. Short dashed and long dashed lines depict cases with $\sigma = 30$ MeV fm$^{-2}$ and $\sigma = 10$ MeV fm$^{-2}$ respectively.

describe quark matter in the droplet we fix $\alpha_s = 0.4$, $B = 120$ MeV fm$^{-3}$, $\sigma = 30$ MeV fm$^{-2}$. The transition pressure of both equations of state $P(\varepsilon)$ for the hadronic and quark degrees of freedom is $P_0 = 525$ MeV fm$^{-3}$. The potential height and critical radius fulfilling $U(R_c) = 0$ decreases as pressure grows, for example for the $P = 535$ MeV fm$^{-3}$ case we obtain $R_c = 16.6$ fm. Note that the hadronic and electromagnetic cascades expected from the self-annihilating DM are of that size $X_\chi \lesssim R_c$ and can be considered to be contained in the droplet region.

The probability of nucleation is determined by the tunneling process for each thermodynamical case considered. In the WKB approximation[13] this is given by $p_0 = \exp\left(-\frac{\mathrm{A(E)}}{\hbar}\right)$ where the action $A(E)$ can be written as

$$A(E) = 2\int_{R_-}^{R_+} \sqrt{[2M(R) + E - U(R)]\,[U(R) - E]}\, dR, \tag{4}$$

being $E$ the energy of the particle (quark) confined in the region between the returning point $R_-$ and $R_+$. It is worth to remind that the treatment sketched here is valid for pressures above that of transition $P_0$. Nucleon confinement prevents from having any quark freedom in hadronic objects.

The determination of the transition point from a nuclear to quark EoS in an extended system is established by the Gibbs criteria relating the equality of baryonic chemical potentials at a given pressure $P_0$ where equilibrium holds. Namely,

$$\mu_b(P_0)\Big|_{\mathrm{Had}} = \mu_b(P_0)\Big|_{\mathrm{Quark}}, \tag{5}$$

where $\mu_b$ is the baryonic chemical potential at $T = 0$ , $\mu_b = \frac{\varepsilon + P}{n_b}$. It is assumed that it corresponds to a first order phase transition[14]. The probability of the formation of the quark droplet is related to the tunneling probability for the selected

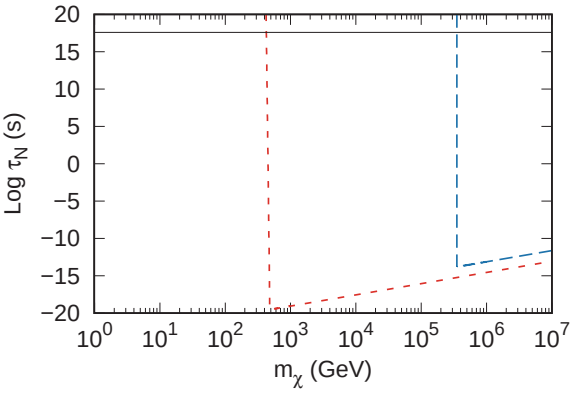



Fig. 3. Logarithm of the nucleation time as a function of mass of the DM particle. We fix $\alpha_s = 0.4$, $B = 120$ MeV fm$^{-3}$ and $P = 535$ MeV fm$^{-3}$. Short dashed and long dashed lines depict cases with $\sigma = 10$ MeV fm$^{-2}$ and $\sigma = 30$ MeV fm$^{-2}$ respectively. The black thin line shows the Universe lifetime value.

thermodynamical conditions. In Fig. 2 we show $p_0$ as a function of the DM particle mass. We fix $\alpha_s = 0.4$, $B = 120$ MeV fm$^{-3}$ and take $P = 535$ MeV fm$^{-3}$ (10 MeV above the transition density $P_0$). Short dashed (red) and long dashed (blue) lines depict cases with $\sigma = 10$ MeV fm$^{-2}$ and $\sigma = 30$ MeV fm$^{-2}$, respectively. We can see that the larger the $\sigma$ the more massive DM needs to be in order to trigger the quark droplet formation. The nucleation time can be obtained from the expression $\tau_N = \frac{\tau}{N_C} = \frac{1}{\nu_0 p_0 N_C}$ where $N_C$ is the number of nucleation centers available in the star and determined from the sites (bubbles with volume $\sim R_c^3$) where energy injection can happen $N_C = \frac{V_{\rm th}}{V_B} \sim (r_{\rm th}/R_c)^3$. $\nu_0 = 10^{-23}$ $s^{-1}$ is a frequency that is obtained from the fundamental energy configuration [12] and in this U height range can be approximately taken as constant. In Fig. 3 we show the logarithm (base 10) of the nucleation time as a function of the mass of the DM particle for the same sets of parameters as in Fig. 2. The critical radii for the $\sigma = 30$ MeV fm$^{-2}$ and $\sigma = 10$ MeV fm$^{-2}$ cases are $R_c = 16.6$ fm and $R_c = 5.5$ fm, respectively. The thin solid line signals the upper limit of the Universe lifetime $\tau_U \sim 4 \times 10^{17}$ s. The oldest NSs have a characteristic age of nearly $\sim 4.9$ Gyrs such as for PSR J04374715 [17]. We can see that the surface tension (still poorly known) highly influences the nucleation process. Light DM mass prevent from having nucleation ($\tau_N > \tau_U$) while it is surely predicted for higher DM masses where $\tau_N$ is tiny.

## 4. Discussion and conclusions

From the current constraints of the DM phase space and for $m_\chi \gtrsim 500$ GeV these results can be complementary to the ones arising from scans of the parameter space of typical SUSY models such as CMSSM, NUHM1, NUHM2, pMSSM10, see Fig. 26.1 in Ref. 16, which integrates constraints set by ATLAS Run 1 and from

SuperKamiokande. As observable consequence of the explained dark mechanism the nucleated stellar object changes its EoS. It is a matter of further study to determine the adjustment of the stellar equilibrium[18] but is is anticipated that could end in the emission of an energetic and very short gamma ray burst[19]. This could be measured with modern capabilities of satellite missions in the $\sim 0.1$ MeV energetic range such as Astrogam[20].

**Acknowledgments**

This work was performed with financial support from the project SA083P17 by Junta de Castilla y León. We also acknowledge CA16214 (PHAROS) and the Spanish Red Consolider MultiDark FPA2017-90566-REDC.

**References**

1. S. A. Teukolsky, S. L. Shapiro, *Black holes, white dwarfs, and neutron stars: the physics of compact objects*. New York City, Wiley (1983).
2. M. Schumann, arXiv:1903.03026v1 [astro-ph.CO].
3. J. Carr, G. Lamanna and J. Lavalle, Rept. Prog. Phys. **69**, 2475 (2006).
4. G. Bertone, D. Hooper, J. Silk, Phys.Rept. **405** (2005) 279.
5. I. Goldman, S. Nussinov, *Phys. Rev. D* **40** (1989) 3221.
6. T. Sjostrand, S. Mrenna and P. Skands, *JHEP05* **26** (2006).
7. A.R. Bodmer, *Phys. Rev. D* **4** (1971) 1601; E. Witten, *Phys. Rev. D* **30** (1984) 272.
8. T.Fischer, N. F. Bastian, M. R. Wu et al., *Nature Astronomy* **2** (2018) 980.
9. Z. Berezhiani, I. Bombaci, A. Drago, F. Frontera, A. Lavagno, *The Astrophysical Journal* **586** (2003) 1250.
10. I. M. Lifshitz, Yu. Kagan, *Zh. Eksp. Teor. Fiz.* **62** (1972) 385.
11. PICO Collab., C. Amole et al., *Phys. Rev. Lett.* **118** (2017) 251301.
12. K. Iida, K. Sato, *Phys. Rev. C* **58** (1998) 2538.
13. T. Nakamura, S. Takagi, *Mod. Phys. Lett. B* **9** (1995) 591.
14. I. Bombaci, D. Logoteta, P. K. Panda, C. Providencia, I. Vidana, *Phys. Lett. B* **680** (2009) 448.
15. A. Chodos, R. L. Jaffe, K. Johnson, Charles B. Thorn, V.F. Weisskopf, *Phys. Rev. D* **9** (1974) 3471.
16. M. Tanabashi et al. (Particle Data Group), *Phys. Rev. D* **98** (2018) 030001.
17. S. Johnston, D. R. Lorimer, P. A. Harrison, et al., *Nature* **361** (1993) 613.
18. M. A. Pérez García, J. Silk, *Physics Letters B* **711** 6.
19. M. A. Pérez García, F. Daigne, J. Silk, *The Astrophysical Journal* **768** (2013) 145.
20. A. de Angelis et al., *Journal of High Energy Astrophysics* **19** (2018) 1.

# The equation of state and cooling of hyperonic neutron stars

Laura Tolos*

*Institut für Theoretische Physik, Goethe Universität Frankfurt, Max von Laue Strasse 1, 60438 Frankfurt, Germany*
*Frankfurt Institute for Advanced Studies, Goethe Universität Frankfurt, Ruth-Moufang-Str. 1, 60438 Frankfurt am Main, Germany*
*Institute of Space Sciences (ICE, CSIC), Campus UAB, Carrer de Can Magrans, 08193, Barcelona, Spain*
*Institut d'Estudis Espacials de Catalunya (IEEC), 08034 Barcelona, Spain*
**E-mail: tolos@th.physik.uni-frankfurt.de*

Mario Centelles and Angels Ramos

*Departament de Física Quàntica i Astrofísica and Institut de Ciències del Cosmos (ICCUB), Facultat de Física, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain*

Rodrigo Negreiros

*Instituto de Física, Universidade Federal Fluminense, Av. Gal. Milton Tavares S/N, Niterói, Brazil*

Veronica Dexheimer

*Department of Physics, Kent State University, Kent OH 44242 USA*

We present two recent parametrizations of the equation of state (FSU2R and FSU2H models) that reproduce the properties of nuclear matter and finite nuclei, fulfill constraints on high-density matter stemming from heavy-ion collisions, produce $2M_\odot$ neutron stars, and generate neutron star radii below 13 km. Making use of these equations of state, cooling simulations for isolated neutron stars are performed. We find that two of the models studied, FSU2R (with nucleons) and, in particular, FSU2H (with nucleons and hyperons), show very good agreement with cooling observations, even without including nucleon pairing. This indicates that cooling observations are compatible with an equation of state that produces a soft nuclear symmetry energy and, thus, generates small neutron star radii. Nevertheless, both schemes produce cold isolated neutron stars with masses above $1.8M_\odot$.

*Keywords*: Neutron stars, cooling, mass-radius constraints, equation of state, hyperons.

## 1. Introduction

The thermal evolution (or cooling history) of neutron stars strongly depends on the equation of state (EoS) and the associated composition of these compact objects[1,2]. Several works have addressed the cooling history of neutron stars and they agree that it is imperative to establish whether fast cooling processes take place, since if the star cools down too fast, it will yield to a disagreement with most observed data. If the stellar proton fraction is high enough, the most important fast processes (the so-called direct URCA, DU) will take place, thus leading to a direct connection between the thermal behavior and the symmetry energy of nuclear matter.

In this paper we investigate the thermal evolution of neutrons stars using EoSs for the nucleonic and hyperonic inner core[3,4] that reconcile the $2M_\odot$ mass observations[5,6] with determinations of stellar radii below 13 km (see Ref. 7 for an overview), the latter being confirmed from analysis of the gravitational-wave emission of GW170817 (see Table II in Ref. 8 and references therein) detected by the LIGO and Virgo collaborations[9]. Moreover, the aforementioned microscopic models satisfy the properties of nuclear matter and finite nuclei[10,11] and constraints from heavy-ion collisions (HICs)[12–14].

In particular, the two models FSU2R and FSU2H, based on the nucleonic FSU2 model[11], differ on the onset of appearance of each hyperon, whereas the neutron star maximum masses calculated with these models show only a moderate dispersion of about $0.1M_\odot$[4]. In the present study, we focus on how the hyperons, as well as the symmetry energy of the microscopic model influence the cooling history of isolated neutron stars, considering also different nucleon pairing scenarios[15].

## 2. FSU2R and FSU2H models

Our models are based on two new parametrizations of the FSU2 RMF model[11]. We start by considering only nucleons. The FSU2R(nuc) model produces a soft symmetry energy and a soft pressure of neutron matter for densities $n \lesssim 2n_0$, as seen in Fig. 1 (left). This is done by increasing the $\Lambda_\omega$ coefficient of the mixed quartic isovector-vector interaction that modifies the density dependence of the nuclear symmetry energy, thus turning the EoS softer[3,4]. As a result, radii within the range of 11.5–13 km are obtained from FSU2R for neutron stars with masses between the maximum mass and $M = 1.4M_\odot$, as seen in Fig. 1 (right). FSU2R predicts $E_{\rm sym}(n_0) = 30.7$ MeV and $L = 46.9$ MeV[4], that differs from the FSU2 model with $E_{\rm sym}(n_0) = 37.6$ MeV and $L = 112.8$ MeV and is in agreement with the limits of recent determinations (see Fig. 1 of Ref. 4 for a summary and references therein). In the high-density sector of the EoS, the FSU2R and FSU2 EoSs are similar, and, hence, FSU2R also reproduces heavy neutron stars (see Fig. 1 (right)).

When hyperons are incorporated, a softening of the high-density EoS with hyperonic degrees of freedom takes place (compare the FSU2R(hyp) and FSU2R(nuc) EoSs in Fig. 1 (left)), so we obtain a reduction of the maximal neutron star mass below $2M_\odot$ in FSU2R(hyp). We may readjust the parameters of the nuclear model by stiffening the EoS of isospin-symmetric matter for densities above twice the saturation density, where hyperons set in. We do that by changing the quartic isoscalar-vector self-interaction (with reducing coupling $\zeta$), which stiffens the EoS at high densities[3,4]. The FSU2H allowing for hyperons, FSU2H(hyp), successfully fulfills the $2M_\odot$ mass limit with moderate radii for the star (see Fig. 1 (right)), while the base nuclear model still reproduces the properties of nuclear matter and nuclei, with $E_{\rm sym}(n_0) = 30.5$ MeV and $L = 44.5$ MeV[4]. The parameter set of FSU2R(nuc) and FSU2H(hyp) and some nuclear matter properties can be found in Refs. 4, 15.

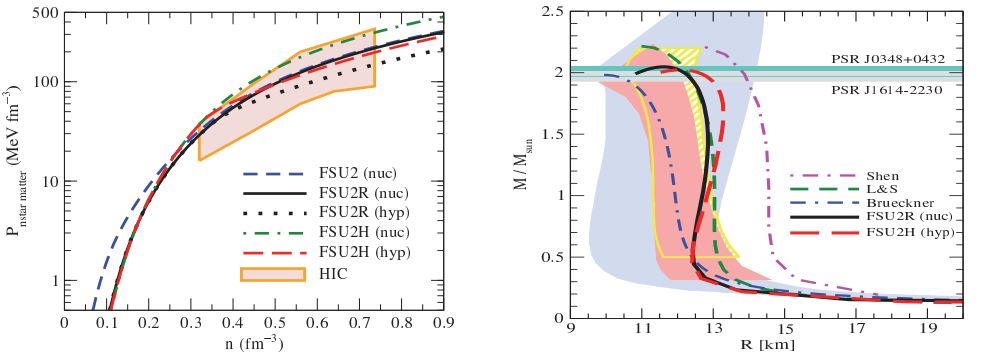



Fig. 1. Left: Pressure of $\beta$-stable neutron star matter as a function of baryon number density for different models[15], with the colored area being compatible with HiCs[12]. Right: Mass versus radius for neutron stars from the models FSU2R and FSU2H[4] together with Shen[16], L&S[17], Brueckner[18]. The thin horizontal bands indicate the $2M_{\odot}$ observations[5,6], the vertical blue band is the region constrained from chiral nuclear interaction[19], the vertical red band is the area derived from five quiescent low-mass X-ray binaries and five photospheric radius expansion X-ray bursters[20], and the vertical striped yellow band is the mass-radius constraint from cooling tails of type-I X-ray bursts in three low-mass X-ray binaries[21].

## 3. Cooling of neutron stars

We first consider the thermal evolution of neutron stars without taking into account any sort of pairing (neither in the core nor in the crust). This is done in Figs. 2 and 3 for the FSU2R and FSU2H models, with and without hyperons. The figures also display the observed surface temperature versus the age of a set of prominent neutron stars, including that of the remnant in Cas A. The nucleonic DU and hyperonic DU thresholds from each microscopic model are given in Table 1 for low-mass to high-mass neutron stars. When DU reactions are allowed in the model, that is, the DU threshold takes place for densities below the central density of the star, they lead to an enhanced cooling of the star, thus to a fast cooling.

Several conclusions can be drawn from the comparison between the different plots in Figs. 2 and 3 together with the analysis of Table 1:

- Low-mass stars of $1.4M_{\odot}$ when only nucleons are considered: the cooling pattern of a $1.4M_{\odot}$ star is slow for FSU2R(nuc) and FSU2H(nuc). This is mainly due to the density dependence of the symmetry energy around saturation and, hence, to the symmetry energy slope parameter ($L$). The smaller the value of $L$ is, the less protons are produced and, thus, the DU process appears at higher densities, making the cooling less efficient.
- High-mass stars of $1.8-2M_{\odot}$ when only nucleons are considered: the different behaviours exhibited by the cooling curves of FSU2R(nuc) (fast cooling) and FSU2H(nuc) (slow cooling) are correlated with the different values of the central densities in these stars, as seen in Table 1. The stiffer the EoS is, the lower central densities are and the slower the cooling is.

Table 1. The nucleonic DU thresholds for FSU2(nuc), FSU2R(nuc), FSU2R(hyp), FSU2H(nuc) and FSU2H(hyp) models, together with the hyperonic DU threshold for each of them. Also shown is the central density $n_c$ for three selected neutron star masses.[15]

| Models | $DU_{th}$ (fm$^{-3}$) | hyp $DU_{th}$ (fm$^{-3}$) | $1.4M_\odot$ $n_c$ (fm$^{-3}$) | $1.76M_\odot$ $n_c$ (fm$^{-3}$) | $2.0M_\odot$ $n_c$ (fm$^{-3}$) |
|---|---|---|---|---|---|
| FSU2 (nuc) | 0.21 | — | 0.35 | 0.47 | 0.64 |
| FSU2R (nuc) | 0.61 | — | 0.39 | 0.51 | 0.72 |
| FSU2H (nuc) | 0.61 | — | 0.34 | 0.39 | 0.45 |
| FSU2R (hyp) | 0.57 | 0.37 | 0.40 | 0.87 | — |
| FSU2H (hyp) | 0.52 | 0.34 | 0.34 | 0.44 | 0.71 |

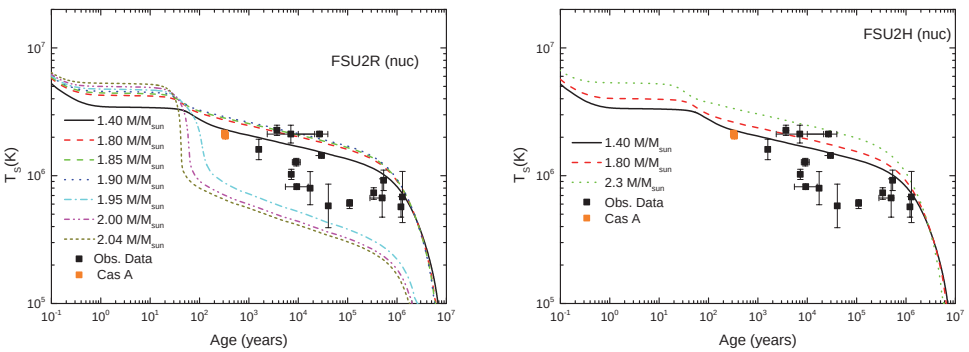



Fig. 2. Surface temperature as a function of the stellar age for different neutron star masses for FSU2R(nuc) (left) and FSU2H(nuc) (right). Also shown are different observed thermal data.[15]

- The inclusion of hyperons in medium- to heavy-mass stars speeds up the cooling (compare in Figs. 2 and 3 FSU2R(nuc) and FSU2R(hyp) or FSU2H(nuc) and FSU2H(hyp)). On the one hand, this is because the presence of hyperons (mostly $\Lambda$ particles) reduces the neutron fraction at a given baryon number density and, consequently, the DU restriction $\vec{k}_{Fn} = \vec{k}_{Fp} + \vec{k}_{Fe}$, can be fulfilled at a lower density. Here $k_{Fn}$, $k_{Fp}$ and $k_{Fe}$ are the Fermi momenta of the neutron, proton and electron, respectively. On the other hand, the appearance of hyperons softens the EoS, thus FSU2R(hyp) and FSU2H(hyp) produce stars with higher central densities than the nucleonic counterparts, so that they may overcome the DU threshold (see Table 1).

We now investigate the effect of neutron superfluidity and proton superconductivity on the cooling of neutron stars. We study the two most relevant cases, the FSU2R(nuc) and the FSU2H(hyp) models, which are the ones that best reproduce the observed data on cooling in Figs. 2 and 3. On the one hand, the introduction of a superfluidity (conductivity) gap in the energy spectrum of baryons reduces the neutrino reaction rates, leading to a sharp drop of neutrino emissivity after the matter temperature drops below the pairing critical temperature. On the

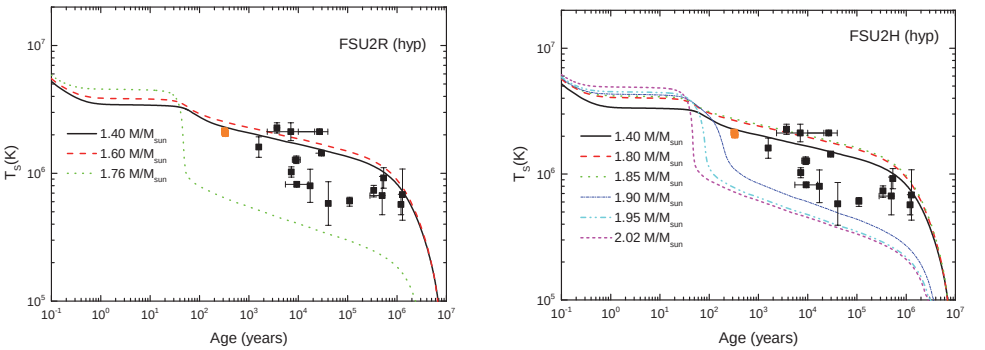



Fig. 3. Surface temperature as a function of the stellar age for different neutron star masses for FSU2R(hyp) (left) and FSU2H(hyp) (right). Also shown are different observed thermal data.[15]

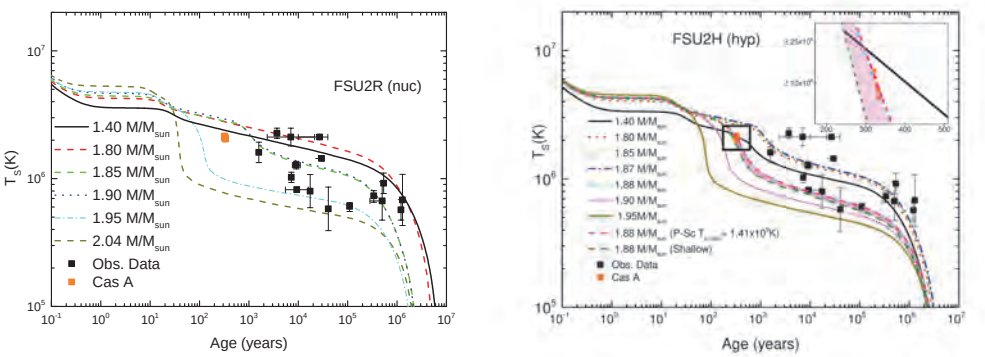



Fig. 4. Surface temperature as a function of the stellar age for different neutron star masses for FSU2R(nuc) (left) and FSU2H(hyp) (right) with medium proton pairing and neutron pairing.[15]

other hand, a new transient neutrino emission process appears, commonly known as pair breaking-formation (PBF) process, where two quasi-baryons with similar anti-parallel momenta annihilate into a neutrino pair.

The analysis of the plots in Fig. 4 indicates that the inclusion of medium proton pairing, in addition to the neutron pairing, improves the agreement of the cooling curves of the FSU2R(nuc) and FSU2H(hyp) models with data, specially for Cas A in the case of the FSU2H(hyp) model, as can be seen in the inset of Fig. 4 (right). Note that in a very recent publication[22], Posselt and Pavlov have revised the Cas A cooling data, concluding that the cooling of this object could be slower. In other words, we find that a shallow/medium proton superconductivity does not lead to an over-suppression of the DU processes, a fact that, combined with miscroscopic models with a soft nuclear symmetry energy, leads to an optimum agreement with observations. However, the calculations favour rather large stellar masses to explain the observed colder stars with surface temperatures $T \lesssim 10^6$ K.

**Acknowledgements**

R.N. acknowledges support from CAPES, CNPq and INCT-FNA Proc. No. 464898/2014-5, whereas L.T. from Grant No. FPA2016-81114-P from MINECO, Heisenberg Project Nr. 383452331 and PHAROS COST Action CA16214. V.D. acknowledges support from NSF under grant PHY-1748621, while M.C. and A.R. from Grant No. FIS2017-87534-P and MDM-2014-0369 (ICCUB) project from MINECO.

**References**

1. D. Page and S. Reddy, *Annual Review of Nuclear and Particle Science* **56**, 327 (2006).
2. F. Weber, R. Negreiros, P. Rosenfield and M. Stejner, *Progress in Particle and Nuclear Physics* **59**, 94 (2007).
3. L. Tolos, M. Centelles and A. Ramos, *Astrophys. J.* **834**, p. 3 (2017).
4. L. Tolos, M. Centelles and A. Ramos, *Publ. Astron. Soc. Austral.* **34**, p. e065 (2017).
5. P. Demorest, T. Pennucci, S. Ransom, M. Roberts and J. Hessels, *Nature* **467**, 1081 (2010).
6. J. Antoniadis *et al.*, *Science* **340**, p. 6131 (2013).
7. M. Fortin, J. L. Zdunik, P. Haensel and M. Bejger, *Astron. Astrophys.* **576**, p. A68 (2015).
8. G. Montana, L. Tolos, M. Hanauske and L. Rezzolla, arxiv: 1811.10929, 2018.
9. B. Abbott *et al.*, *Phys. Rev. Lett.* **119**, p. 161101 (2017).
10. M. B. Tsang *et al.*, *Phys. Rev.* **C86**, p. 015803 (2012).
11. W.-C. Chen and J. Piekarewicz, *Phys. Rev.* **C90**, p. 044305 (2014).
12. P. Danielewicz, R. Lacey and W. G. Lynch, *Science* **298**, 1592 (2002).
13. C. Fuchs, A. Faessler, E. Zabrodin and Y.-M. Zheng, *Phys. Rev. Lett.* **86**, 1974 (2001).
14. W. G. Lynch, M. B. Tsang, Y. Zhang, P. Danielewicz, M. Famiano, Z. Li and A. W. Steiner, *Prog. Part. Nucl. Phys.* **62**, p. 427 (2009).
15. R. Negreiros, L. Tolos, M. Centelles, A. Ramos and V. Dexheimer, *Astrophys. J.* **863**, p. 104 (2018).
16. H. Shen, H. Toki, K. Oyamatsu and K. Sumiyoshi, *Nucl. Phys.* **A637**, 435 (1998).
17. J. M. Lattimer and D. F. Swesty, *Nuclear Physics A* **535**, 331 (1991).
18. B. K. Sharma, M. Centelles, X. Viñas, M. Baldo and G. F. Burgio, *Astron. Astrophys.* **584**, p. A103 (2015).
19. K. Hebeler, J. M. Lattimer, C. J. Pethick and A. Schwenk, *Astrophys. J.* **773**, p. 11 (2013).
20. J. M. Lattimer and A. W. Steiner, *Eur. Phys. J.* **A50**, p. 40 (2014).
21. J. Nättilä, A. W. Steiner, J. J. E. Kajava, V. F. Suleimanov and J. Poutanen, *Astron. Astrophys.* **591**, p. A25 (2016).
22. B. Posselt and G. G. Pavlov, *Astrophys. J.* **864**, p. 135 (2018).

# BlackHoleCam — Testing general relativity with pulsars orbiting Sagittarius A*

Ralph P. Eatough*[1], Gregory Desvignes[1], Kuo Liu[1], Robert S. Wharton[1], Aristedis Noutsos[1], Pablo Torne[2,1], Ramesh Karuppusamy[1], Lijing Shao[3,1], Michael Kramer[1,4], Heino Falcke[5,1] and Luciano Rezzolla[6]

[1] *Max-Planck-Institut für Radioastronomie, Auf dem Hügel 69, Bonn, D-53121, Germany*
* *E-mail: reatough@mpifr-bonn.mpg.de*
*www.mpifr-bonn.mpg.de*

[2] *Instituto de Radioastronomía Milimétrica, IRAM, Avenida Divina Pastora 7, Local 20, E-18012, Granada, Spain*

[3] *Kavli Institute for Astronomy and Astrophysics, Peking University, Beijing 100871, China*

[4] *Jodrell Bank Centre for Astrophysics, The University of Manchester, Alan Turing Building, Manchester M13 9PL, UK*

[5] *Department of Astrophysics, Institute for Mathematics, Astrophysics and Particle Physics (IMAPP), Radboud University,P.O. Box 9010, 6500 GL Nijmegen, The Netherlands*

[6] *Institut für Theoretische Physik, Goethe-Universität Frankfurt, Max-von-Laue-Straße 1, D-60438 Frankfurt am Main, Germany*

BlackHoleCam is a project funded by a European Research Council *Synergy Grant* to build a complete astrophysical description of nearby supermassive black holes by using a combination of radio imaging, pulsar observations, stellar astrometry and general relativistic magneto-hydrodynamic models. BlackHoleCam* scientists are active partners of the Event Horizon Telescope Consortium†. In this talk I will discuss the use of pulsars orbiting Sagittarius A* for tests of General Relativity, the current difficulties in detecting such sources, recent results from the Galactic Centre magnetar PSR J1745−2900 and how BlackHoleCam aims to search for undiscovered pulsars in the Galactic Centre.

*Keywords*: Pulsars; Black Holes; Gravity Tests.

## 1. The science case for observing pulsars orbiting Sagittarius A*

For over 40 years observations double neutron star systems − the collapsed and degenerate remnants of massive stars − where one, or both, stars are active radio pulsars have demonstrated that they form exceptional natural "laboratories" for precision tests of theories of gravitation[1,2]. In this vein, a pulsar in a close orbit around the supermassive black hole at the centre of our Galaxy (Sagittarius A* − Sgr A* for short) would be at the apex of gravity experiments made possible using pulsars[3,4].

*www.blackholecam.org
†www.eventhorizontelescope.org

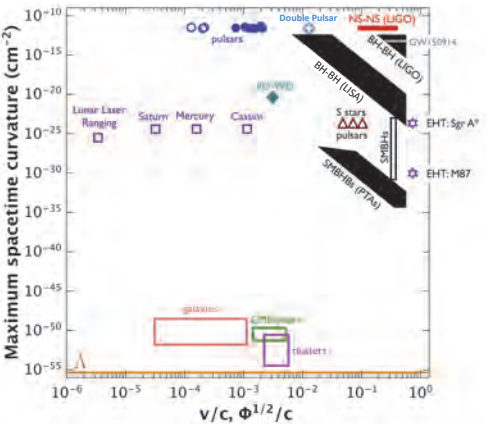



Fig. 1. Parameter space of observations and tests of gravity. On the x-axis, $v$ denotes the typical velocity of the system's components while $\Phi$ denotes the gravitational potential being probed by photons propagating in the corresponding spacetime. On the y-axis the maximum spacetime curvature (taken at the horizon for black holes) in the system is indicated as a measure of how much the system deviates from flat spacetime. Filled areas indicate gravitational wave tests, while hollow areas stand for quasi-stationary tests, including accretion onto compact objects. The rightmost hollow blue circle stands for the Shapiro delay test in the double pulsar. Figure and caption reproduced from Ref. 6 with the kind permission of Wex and Kramer.

Such a system will allow the fundamental predictions of black-hole properties in General Relativity (GR) to be tested; properties that Advanced LIGO, which measures the strongly dynamical regime of a merger, potentially cannot[5]. These include the *no-hair theorem* and the *cosmic censorship conjecture*[3,4]. For example, the latter is tested through measurements of frame dragging caused by the spinning black hole, which manifests itself as a contribution to the precession of the pulsar orbit. A Kerr black hole should exhibit a dimensionless spin parameter $\chi$ which is no larger than unity; a spin parameter greater would be in conflict with GR posing a direct contest to the theory. Fig. 1 shows how pulsar tests fit into the relativistic regime of other gravity tests past and present. In Fig. 2 (*l*eft panel) from Ref. 4 the expected signature of the black hole quadrupole moment in pulsar timing residuals for a putative pulsar closely orbiting Sgr A* is displayed. Ref. 7 have shown that by combining measurements of the black hole spin from stellar astrometry, pulsar timing and interferometric imaging of the black hole shadow, an unbiased and quantitative test of the no-hair theorem is possible (Fig. 2 *right* panel).

## 2. Galactic Centre pulsar searches

Over the last couple of decades a number of searches of the inner tens of parsecs have taken place[8–12]. In 2013, the Galactic Centre (GC) pulsar population[a] was

[a]Here we define the GC pulsar population as those with pulses detected in radio at a projected offset of less than 0.5° (∼73 pc) from Sgr A*.

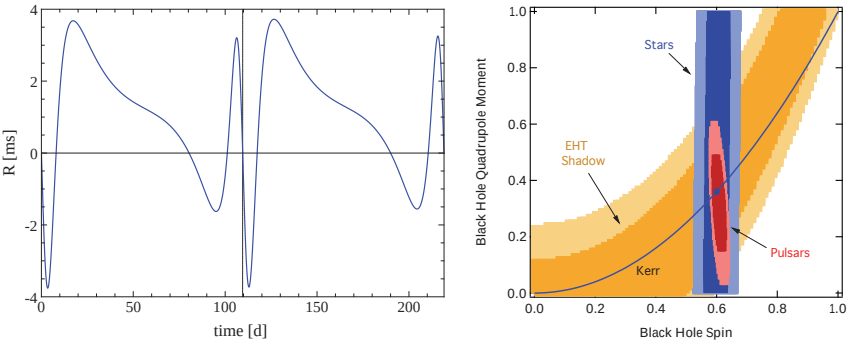



Fig. 2. (*left*) A simulation showing the effect of the quadrupole moment of Sgr A* "imprinted" upon pulsar timing residuals over two orbital cycles. Here an orbital period of 0.3 yr, an eccentricity of 0.5 and a dimensionless black hole spin parameter of unity were assumed. Figure reproduced from Ref. 4. (*right*) Comparison of the posterior likelihood of measuring the spin and quadrupole moment of Sgr A* using the orbits of two stars (blue), timing of three periapsis passages of a normal pulsar (red) and the shape of the black hole shadow from radio interferometric imaging (gold). The solid curve shows the expected relation between the two quantities for a Kerr metric. The filled circle marks the assumed spin and quadrupole moment. Figure reproduced from Ref. 7.

brought to a total of six with the detection of PSR J1745−2900, first through its gamma ray emission[13,14] and then finally in radio[15,16].

As radio telescope hardware and analysis techniques are improved, GC pulsar searches are carving out additional areas of parameter space neglected by previous searches. An example is the progression to higher than "normal" pulsar observing frequencies ($f_{\rm obs} \gtrsim 10\,{\rm GHz}$) where the deleterious effects of interstellar dispersion and scattering (effects that are at their highest in GC) are reduced. Unfortunately, improvements in sensitivity by choosing higher frequencies are limited by the steep spectrum of pulsar emission, where detection signal-to-noise ratio S/N is typically $\propto f_{\rm obs}^{\sim -1.7}$. This is one case where competing selection effects affect the performance of GC pulsar searches (Table 1). In general, higher observing frequencies and longer integration lengths $T_{\rm obs}$ (given the large distance of the GC) are more favourable, however this predisposes searches to flatter spectrum isolated pulsars. Binary pulsar searches can become computationally prohibitive above certain $T_{\rm obs}$ where even single parameter acceleration search computations $C_a$ scale with $T_{\rm obs}^3$.

## 3. BlackHoleCam pulsar work

BlackHoleCam[18], and members of the Event Horizon Telescope (EHT) Consortium, aim to tackle these observational constraints by using the largest and most sensitive telescopes operating at millimetre wavelengths which form elements of the EHT telescope. Because both EHT VLBI imaging and pulsar observations can utilise the same raw data product from each array element, EHT VLBI and pulsar observations can be commensal.

Table 1. Competing observational selection effects for GC radio pulsar searches. Pulse broadening effects, such as dispersion smearing $\tau_{\rm ch}$, scatter broadening $\tau_{\rm s}$ and acceleration broadening $\tau_a$, add in quadrature to the intrinsic pulse width $W$ making $W_{\rm eff}$, which scales with the detection signal-to-noise ratio as S/N $\propto \sqrt{(P - W_{\rm eff})/W_{\rm eff}}$ where $P$ is the spin period. ✓ indicates favourable observational configurations while ✗ are unfavourable.

| Observing frequency related | Higher $f_{\rm obs}$ | Lower $f_{\rm obs}$ |
|---|---|---|
| Intra-channel pulse dispersion smearing ($\tau_{\rm ch} \propto f_{\rm obs}^{-3}$) | ✓ | ✗ |
| Pulse scattering broadening ($\tau_{\rm s} \propto f_{\rm obs}^{-3.8}$) | ✓ | ✗ |
| Intense GC background continuum emission[17] (S/N $\propto f_{\rm obs}^{-0.8}$) | ✓ | ✗ |
| Pulsar spectra (S/N $\propto f_{\rm obs}^{\sim -1.7}$) | ✗ | ✓ |
| **Integration length related** | **Longer $T_{\rm obs}$** | **Shorter $T_{\rm obs}$** |
| GC distance (S/N $\propto \sqrt{T_{\rm obs}}$) | ✓ | ✗ |
| Pulse broadening from acceleration ($\tau_a \propto T_{\rm obs}^2$) | ✗ | ✓ |
| Pulse broadening from jerk ($\tau_j \propto T_{\rm obs}^3$) | ✗ | ✓ |
| Computational operations for acceleration search ($C_a \propto T_{\rm obs}^3$) | ✗ | ✓ |

Thus far, pulsar search efforts have concentrated on analysing data from the single most sensitive element of the EHT: "fully phased" ALMA. In the future we can envisage using a phased array of the largest components of the EHT to further increase sensitivity or to mitigate site specific interference contamination.

## Acknowledgments

The authors acknowledge financial support by the European Research Council (ERC) Synergy Grant "BlackHoleCam: Imaging the Event Horizon of Black Holes" (grant 610058).

## References

1. J. H. Taylor and J. M. Weisberg, *ApJ*, **253**, 908, (1982).
2. M. Kramer et al., *Science*, **314**, 97, (2006).
3. N. Wex and S. Kopeikin, *ApJ*, **513**, 388, (1999).
4. K. Liu et al., *ApJ*, **747**, 1, (2012).
5. E. Thrane, P. Lasky and Y. Levin, *ArXiv*, `arXiv:gr-qc/1706.05152`, (2017).
6. M. Kramer, *Astronomy & Geophysics*, **58**, 3.31, (2017).
7. D. Psaltis, M. Kramer and N. Wex, *ApJ*, **818**, 121, (2016).
8. S. Johnston et al., *MNRAS*, **373**, L6, (2006).
9. J. S. Deneva et al., *ApJ*, **702**, L177, (2009).
10. J. P. Macquart et al., *ApJ*, **715**, 939, (2010).
11. A. Siemion et al., *IAU Symposium*, **291**, 57, (2013).
12. R. P. Eatough et al., *IAU Symposium*, **291**, 382, (2013).
13. J. A. Kennea et al., *ApJ*, **770**, L24, (2013).
14. K. Mori et al., *ApJ*, **770**, L23, (2013).
15. R. P. Eatough et al., *Nature*, **501**, 391, (2013).
16. R. M. Shannon and S. Johnston, *MNRAS*, **435**, L29, (2013).
17. C. J. Law et al. *ApJS*, **177**, 255, (2008).
18. C. Goddi et al., *IJMPD*, **26**, 1730002-239, (2017).

# Deflection of light in equatorial plane of Kerr-Sen black hole

Rashmi Uniyal*

*Department of Physics, Government Degree College, Narendranagar, Tehri Garhwal, 249 175 Uttarakhand, India*
**E-mail: rashmiuniyal001@gmail.com; uniyal@associates.iucaa.in*

Hemwati Nandan

*Department of Physics, Gurukula Kangri Vishwavidyalaya, Haridwar 249 404, Uttarakhand, India*
*E-mail: hnandan@associates.iucaa.in*

Philippe Jetzer

*Physik-Institut, University of Zürich, Winterthurerstr 190 8057 Zürich, Switzerland*
*E-mail: jetzer@physik.uzh.ch*

We study the gravitational lensing by a Kerr-Sen black hole spacetime which arises a solution to the low-energy effective field theory for four-dimensional heterotic string theory. A closed form expression of the deflection angle of light rays lensed by the Kerr-Sen black hole is derived as a function of impact parameter, spin and charge in the equatorial plane of the black hole. It is observed that charge parameter behaves differently from the spin parameter for the photons in direct and retrograde orbits around the black hole. The deflection angle becomes larger in strong field limit with an increment in the value of charge parameter and it is observed that this effect is more perceptible in case of the direct orbits as compared to the retro orbits. The results obtained are also compared with the corresponding cases of well known Kerr black hole in general relativity (GR).

*Keywords*: Gravitational lensing; Kerr-Sen black hole; Heterotic string theory; Deflection angle.

## 1. Introduction

One of the consequences of Einstein's general relativity (GR) is that the light rays passing a massive body are deflected by the virtue of gravity and the resulting phenomenon is known as Gravitational lensing[2] (GL), as first observed by Eddington during the solar eclipse of 1919. The GL, theory and observations, is one of the most important areas in modern astronomy[3] and it also provides a clean and unique probe of the dark matter at all the distance scales since it is independent of the nature and physical state of the lensing mass[4].

The main objective of the present work is to study the GL by a Kerr-Sen BH in equatorial plane to have exact closed-form solutions for the deflection angle of light such that both the strong and weak field limits are satisfied[5–11].

In the next section, structure of KSBH spacetime is discussed alongwith the first integrals of geodesic equations for null geodesics. The critical parameters for obtaining the exact deflection angles are then calculated in section 3. The closed-form expression for the deflection angle as a function of impact parameter and BH spin is derived in section 4. Finally, the results are summarised in last section.

## 2. Kerr-Sen BH Spacetime

The Einstein metric for KSBH,

$$ds^2 = -\left(\frac{\Delta - a^2\sin^2\theta}{\Sigma}\right)dt^2 + \frac{\Sigma}{\Delta}dr^2 - \frac{4\mu a r\cosh^2\alpha\sin^2\theta}{\Sigma}dtd\phi + \Sigma d\theta^2 + \frac{\Xi\sin^2\theta}{\Sigma}d\phi^2, \tag{1}$$

where the metric functions are described as, $\Delta = r^2 - 2\mu r + a^2$, $\Sigma = r^2 + a^2\cos^2\theta + 2\mu r\sinh^2\alpha$, and $\Xi = \left(r^2 + 2\mu r\sinh^2\alpha + a^2\right)^2 - a^2\Delta\sin^2\theta$. The parameters $\mu$, $\alpha$ and $a$ are related to the physical mass $M$, charge $Q$ and angular momentum $J$ as follows,

$$M = \frac{\mu}{2}(1+\cosh 2\alpha), \qquad Q = \frac{\mu}{\sqrt{2}}\sinh^2 2\alpha, \qquad J = \frac{a\mu}{2}(1+\cosh 2\alpha)\,. \tag{2}$$

For a nonextremal BH, there exist two horizons, determined by $\Delta(r) = 0$ as,

$$r_\pm = M - \frac{Q^2}{2M} \pm \sqrt{\left(M - \frac{Q^2}{2M}\right)^2 - a^2}\,, \tag{3}$$

where $r_+$ and $r_-$ represent the outer and the inner horizons of the BH respectively.

### 2.1. *First Integrals of Geodesic Equations*

The first integral of radial and latitudinal coordinates (i.e. $r$ and $\phi$) for null geodesics[13] can be obtained as,

$$\dot{r}^2 \;=\; L^2\left(\frac{1}{b^2} + \frac{a^2}{b^2 r(r+x)} + \frac{2Ma^2}{b^2 r(r+x)^2} - \frac{4Ma}{b_s r(r+x)^2} - \frac{1}{r(r+x)} + \frac{2M}{r(r+x)^2}\right) \tag{4}$$

and

$$\dot{\phi} \;=\; \frac{L}{\Delta}\left(1 - \frac{2M}{r+x} + \frac{2Ma}{r+x}\frac{E}{L}\right), \tag{5}$$

where $x = Q^2/2M$ and $M = GM_\star/c^2$ (gravitational radius) with $M_\star$ defined as the physical mass of the BH. The time derivatives in 4 and 5 are taken with respect to the coordinate time variable $t$ defined as, $t = c\tau$ where $\tau$ represents the physical time with,

$$b_s \;=\; s\left|\frac{L}{E}\right| \equiv sb. \tag{6}$$

## 3. Critical Parameters

Using the change of variable as $u = 1/r$, the orbit equation can be obtained from Eq. (5) and Eq. (4) as,

$$\left(\frac{du}{d\phi}\right)^2 = \frac{\left(a^2u^2 + (1-xu)(1-2Mu)\right)^2}{\left(1-2Mu+2Mau/b_s\right)^2(1-xu)}\mathcal{B}(u), \tag{7}$$

where,

$$\mathcal{B}(u) = \frac{(1-xu)}{b^2} + \left(\frac{a^2}{b^2}-1\right)u^2 + \left(1-\frac{a}{b_s}\right)^2 2Mu^3. \tag{8}$$

We will further consider the case of one real negative root $u_1$ and two real distinct positive roots $u_2$ and $u_3$ given in terms of two intermediate constants $P$ and $Q$ that allow one to line up the order such as, $u_1 < u_2 < u_3$, $u_1 = \frac{P-2M-Q}{4Mr_0}$, $u_2 = \frac{1}{r_0}$, $u_3 = \frac{P-2M+Q}{4Mr_0}$.

By comparing the coefficients in $\mathcal{B}$(u) to those in the original polynomial in Eq. (8), one can first obtain the following relationship between $P$ and $\{a, b, s, r_0\}$,

$$P = r_0\frac{\left(1-\frac{a}{b_s}\right)}{\left(1+\frac{a}{b_s}\right)}. \tag{9}$$

It leads to the following relation between the critical parameters,

$$r_{sc} = 3M\frac{\left(1-\frac{a}{b_s}\right)}{\left(1+\frac{a}{b_s}\right)}. \tag{10}$$

$$r_0 = \frac{2\Theta}{\sqrt{3}}\sqrt{\left(b^2\left(1-\frac{a^2}{b^2}\right)+\frac{x^2}{3}\right)}. \tag{11}$$

where,

$$\Theta = \cos\left[\frac{1}{3}\arccos\left\{-\frac{\sqrt{3}}{18}\left(\frac{54Mb^2\left(1-\frac{a}{b_s}\right)^2 - 9b^2x\left(1-\frac{a^2}{b^2}\right) - 2x^3}{\left(b^2\left(1-\frac{a^2}{b^2}\right)+\frac{x^2}{3}\right)^{3/2}}\right)\right\}\right]. \tag{12}$$

The above relation among the distance of closest approach ($r_0$) and the invariant impact parameter is extremely important in Strong Deflection Limit (SDL) as well as Weak Deflection Limit (WDL) series expansions in terms of the invariant normalised quantity $b'$. The expression also reduces to the corresponding Schwarzschild limits when charge and spin parameters are zero.

In the strong deflection limit, $P = 3M$and the following expressions involving the critical quantities:

$$r_{sc} = 3M\frac{\left(1-\frac{a}{b_{sc}}\right)}{\left(1+\frac{a}{b_{sc}}\right)} \tag{13}$$

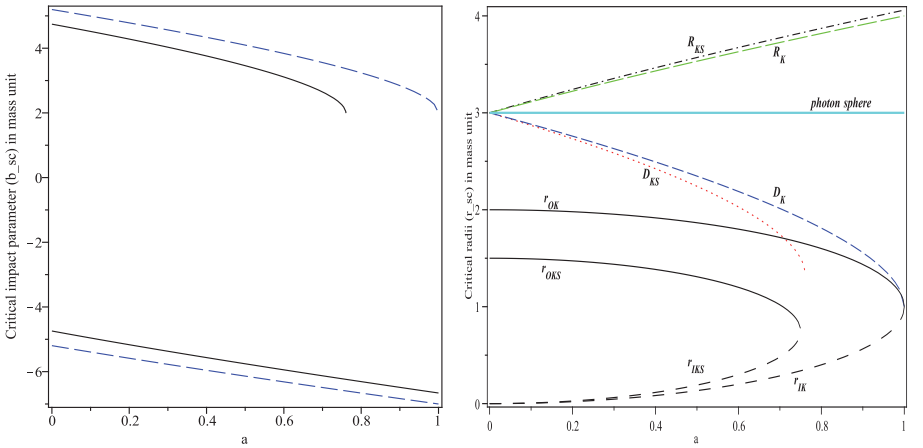



Fig. 1. ***Left Panel:*** Here solid and dashed lines represent the critical impact parameter (in mass unit) for KSBH and KBH respectively with $x = 0.5$ for KSBH; further upper portion of the plot corresponds to $b_{+c}$ solutions while the lower portion of the plot corresponds to the $b_{-c}$ solutions; ***Right Panle:*** Variation of critical radii in mass units with spin parameter $a$. Here $D_K$ and $D_{KS}$ depict the corresponding radii for direct orbits and $R_K$ and $R_{KS}$ depict the corresponding radii for retro orbits of KBH and KSBH respectively. Lower portion of the plot represents inner and outer horizons for KBH and KSBH.

$$b_{sc} = -a + 2s\sqrt{3M(3M-x)}\cos\left[\frac{1}{3}\arccos\left(\frac{-3sa}{3M-x}\sqrt{\frac{3M}{3M-x}}\right)\right]. \tag{14}$$

and

$$r_{sc} = x + \frac{2}{3}(3M-x)\left[1+\cos\left(\frac{2}{3}\arccos\left(\frac{-3sa}{3M-x}\sqrt{\frac{3M}{3M-x}}\right)\right)\right]. \tag{15}$$

## 4. Bending Angle of Light Rays

Now if one considers a light ray starting in an asymptotic region and approaching a BH, with $r_0$ as distance of its closest approach. The exact expression for the bending angle can easily be expressed in terms of elliptical integrals of third kind[5] as,

$$\alpha = -\pi + \frac{4}{1-\omega_s}\sqrt{\frac{r_0}{Q}}\Big\{\Omega_+\Big[\Pi(n_+,k) - \Pi(n_+,\psi,k)\Big] + \Omega_-\Big[\Pi(n_-,k) - \Pi(n_-,\psi,k)\Big]\Big\}. \tag{16}$$

where $\Pi(n_\pm,k)$ and $\Pi(n_\pm,\psi,k)$ are the complete and the incomplete elliptic integrals of the third kind respectively. The argument $k^2$ is defined through the elliptic integral as usual in the range $0 \le k^2 \le 1$. The other variables in the above expression are defined as, $\Omega_\pm = \frac{C_\pm}{u_\pm - u_1}$, $k^2 = \frac{Q-P+6M}{2Q}$, $\psi = \arcsin\sqrt{\frac{Q+2M-P}{Q+6M-P}}$ and $n_\pm = \frac{u_2-u_1}{u_\pm-u_1}$.

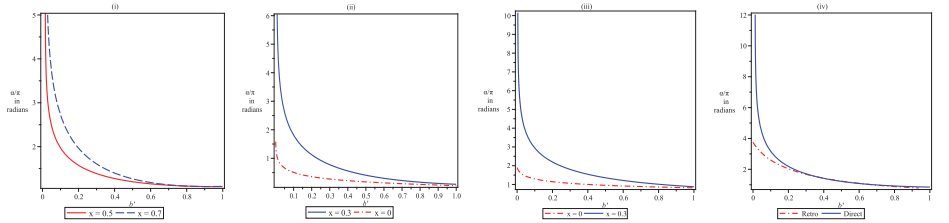



Fig. 2. Exact deflection angle as a function of normalised impact parameter with spin parameter value $a = 0.5$. Left section of the above plots where $b' \to 0$ corresponds to the strong deflection limit while the right section where $b' \to 1$ corresponds to weak deflection limit.

## 5. Summary and Conclusions

We have studied the GL for a KSBH in its equatorial plane and derived an exact expression for the bending angle of light. The effect of frame-dragging on the bending angle of photons in such cases has previously been discussed for KBH[5]. The additional charge parameter behaves similar to the spin parameter for direct orbiting photons but oppositely for retrograde orbiting photons as the bending angle increases in either case on increasing the numerical value of the charge parameter. Though this increment is still much larger for direct orbiting photons. This difference in the bending angle can clearly be visible through the shifts of the corresponding relativistic images. In order to study this shift in relativistic images, one needs to study the series expansion of the bending angle in weak as well as strong deflection limits. Hence, as a further work we will derive the series expansion of the above obtained bending angle formulas in both the strong and weak field limits. This will allow an easier comparison with similar results obtained for other BH types.

## Acknowledgments

The authors RU and HN would like to thank Science and Engineering Research Board (SERB), New Delhi for financial support through grant no. EMR/2017/000339. The authors RU and HN are also thankful to IUCAA, Pune (where a part of the work was completed) for support in form of academic visits under its Associateship programme. RU and HN would also like to thank Prof. Philippe Jetzer for the study visit to Physik-Institut, University of Zurich, Switzerland (where a part of this work was initially carried out) and for the hospitality during the stay there.

## References

1. Fourteenth Marcel Grossmann Meeting (2015), `http://www.icra.it/MG/mg14/`.
2. P. Schneider, J. Ehlers and E. E. Falco (1992), Gravitational Lensing, Springer Verlag, Berlin.

3. S. Refsdal and J. Surdej (1994), Gravitational lenses. Rep. Prog. Phys. 56 117; R. Narayan and M. Bartelmann (1996), Lectures on Gravitational Lensing. astro-ph 9606001; N. Straumann, Ph. Jetzer and J. Kaplan (1998), Topics on Gravitational Lensing. Napoli Series on Physics and Astrophysics (Bibliopolis, Naples).
4. P. Jetzer, http://www.physik.uzh.ch/groups/jetzer/notes/Gravitational-Lensing.pdf.
5. S.V. Iyer, E.C. Hansen, Phys. Rev. D80 (2009) 124023.
6. V. Bozza, Gen.Rel.Grav.42 (2010) 2269-2300.
7. N. S. Barlow, S. J. Weinstein, J. A. Faber, arXiv: 1701.05828 v1 [gr-qc].
8. N. S. Barlow, A. J. Schultz, S. J. Weinstein and D. A. Kofke (2012), J. Chem. Phys. 137 204102.
9. N. S. Barlow, A. J. Schultz, S. J. Weinstein and D. A. Kofke (2014), AIChE J. 60 3336-3349.
10. N. S. Barlow, A. J. Schultz, S. J. Weinstein and D. A. Kofke (2015), J. Chem. Phys. 143 071103:1.
11. N. S. Barlow, C. R. Stanton, N. Hill, S. G. Weinstein and G C A (2017), Q. J. Mech. Appl. Math. doi:10.1093/qjmam/hbw014.
12. A. Sen (1992), Phys. Rev. Lett. 69, 1006.
13. R. Uniyal, H. Nandan and K. D. Purohit (2018), Class. Quant. Grav. 35, 025003.

# Reconstructing the metric of the local Universe from number counts observations

Sergio Andrés Vallejo Peña[1,3,†] and Antonio Enea Romano[1,2,3]

[1] *Instituto de Fisica, Universidad de Antioquia, A.A.1226, Medellin, Colombia*
[2] *Theoretical Physics Department, CERN, CH-1211 Geneva 23, Switzerland*
[3] *ICRANet, Piazza della Repubblica 10, I–65122 Pescara*
[†] *E-mail: sergio.vallejo@udea.edu.co*

Number counts observations available with new surveys such as the Euclid mission will be an important source of information about the metric of the Universe. We compute the low red-shift expansion for the density contrast using an exact spherically symmetric solution in presence of a cosmological constant. At low red-shift the expansion is more precise than linear perturbation theory prediction. We then use the local expansion to reconstruct the metric from the monopole of the density contrast. We test the inversion method using numerical calculations and find a good agreement within the regime of validity of the red-shift expansion. The method could be applied to observational data to reconstruct the metric of the local Universe with a level of precision higher than the one achievable using perturbation theory.

*Keywords*: Number counts, density contrast, redshift, metric, local Universe.

## 1. Introduction

The standard cosmological model is based on the assumption that the Universe is homogeneous and isotropic on sufficiently larges scales, and is confirmed by different observations such as for example the cosmic microwave background (CMB) radiation[1] or of galaxy catalogues. However the presence of structure at smaller scales can affect local observations as it was shown in Ref. 2, and it is therefore important to understand its consequences. The effects of inhomogeneities on cosmological observables have been studied in different cases, such as dark energy, the luminosity distance[3–5] or the expansion scalar.[6] These effects are due to the fact that spatial inhomogeneities change the energy of photons, modifying the cosmological red-shift due to the Universe expansion. As a consequence some errors are produced in the estimation of parameters based on homogeneous cosmological models.

One important source of information about the Universe are galaxy catalogues since they allow to map the local density field. Since we can only measure the red-shift of astrophysical objects for which other distance measurement methods such as stellar parallax cannot be applied, it is important to take into account the effects of these inhomogeneities on the metric in order to compute self-consistently the density in red-shift space. This is particularly important when trying to determine the metric of the Universe.

## 2. Modeling the local Universe

In order to model the monopole component of the local structure we use the LTB solution[7–11]

$$ds^2 = -dt^2 + \frac{R'(t,r)^2}{1+2\,E(r)}dr^2 + R(t,r)^2 d\Omega^2\,, \tag{1}$$

where $E(r)$ is an arbitrary function of $r$, and $R'(t,r) = \partial_r R(t,r)$. The analytical solution of Einstein's equations can be derived[12,13] if we introduce a new function $k(r)$, and a new coordinate $\eta = \eta(t,r)$ given by

$$\frac{\partial\eta}{\partial t}|_r = \frac{r}{R} = \frac{1}{a}\,, \quad k(r) = -\frac{2E(r)}{r^2}\,. \tag{2}$$

The Einstein equations imply

$$\left(\frac{\partial a}{\partial \eta}\right)^2 = -k(r)a^2 + \frac{\rho_0(r)}{3}a + \frac{\Lambda}{3}a^4\,, \tag{3}$$

where $\rho_0(r)$ is an arbitrary function of $r$, and we adopt a system of units in which $c = 8\pi G = 1$. Without any loss of generality we adopt the coordinate system in which $\rho_0(r)$ is a constant, which is known as the FLRW gauge. The solution to the above equation can be written in the form[14]

$$a(\eta,r) = \frac{\rho_0}{k(r) + 3\wp(\frac{\eta}{2}; g_2(r), g_3(r))}\,. \tag{4}$$

where $\wp(x; g_2, g_3)$ is the Weierstrass elliptic function and

$$g_2(r) = \frac{4}{3}k(r)^2\,, \quad g_3(r) = \frac{4}{27}\left(2k(r)^3 - \Lambda\rho_0^2\right). \tag{5}$$

The relation between $t$ and $\eta$ can be found by integrating Eq. (2) and is given by[6]

$$t(\eta,r) = \frac{2\rho_0}{3\wp'\left(\wp^{-1}\left(-\frac{k(r)}{3}\right)\right)}\left[\ln\left(\frac{\sigma\left(\frac{\eta}{2} - \wp^{-1}\left(-\frac{k(r)}{3}\right)\right)}{\sigma\left(\frac{\eta}{2} + \wp^{-1}\left(-\frac{k(r)}{3}\right)\right)}\right) + \eta\zeta\left(\wp^{-1}\left(-\frac{k(r)}{3}\right)\right)\right].$$

The radial null geodesic equations take the form[15]

$$\frac{d\eta}{dz} = -\frac{\partial_r t(\eta,r) + G(\eta,r)}{(1+z)\partial_\eta G(\eta,r)}\,, \qquad \frac{dr}{dz} = \frac{a(\eta,r)}{(1+z)\partial_\eta G(\eta,r)}\,, \tag{6}$$

where

$$G(\eta,r) \equiv \frac{R'(t(\eta,r),r)}{\sqrt{1-k(r)r^2}}\,.$$

The density profile is given by

$$\rho(\eta,r) = \rho(t(\eta,r),r) = \frac{\rho_0}{a(\eta,r)^2 R'(t(\eta,r),r)}\,. \tag{7}$$

The background density of the universe is given by the sub-horizon volume average of $\rho$ on constant time slices

$$\overline{\rho}(t) = \frac{\int_V \rho(t,r)dV}{\int_V dV} = \frac{\int_0^{r_{Hor}(t)} \rho(t,r)\frac{R(t,r)^2 R'(t,r)}{\sqrt{1-k(r)r^2}}dr}{\int_0^{r_{Hor}(t)} \frac{R(t,r)^2 R'(t,r)}{\sqrt{1-k(r)r^2}}dr}\,, \tag{8}$$

where the upper limit of the integrals $r_{Hor}(t)$ is the comoving horizon as a function of time, and determines the region of space causally connected with the central observer at time $t$. We can then evaluate $\overline{\rho}(t)$ at the time $t(z)$ corresponding to a given redshift $z$, i.e. the time along null radial geodesics, and define the background value of $\rho$ at redshift $z$ as $\overline{\rho}(z) \equiv \overline{\rho}(t(z))$. We can then define the density contrast

$$\delta(z) = \frac{\rho(z)}{\overline{\rho}(z)} - 1\,. \tag{9}$$

If the size of the local inhomogeneity is sufficiently smaller than the volume over which the integral in Eq. (8) is performed then $\overline{\rho}$ will get most of its contribution from the asymptotically homogeneous region and the average density will be well approximated by the asymptotic density

$$\overline{\rho}(z) = 3(H_0^b)^2\Omega_M^b(1+z)^3\,; \quad H_0^b = \overline{H}(0)\,, \quad \Omega_M^b = \frac{\overline{\rho}(0)}{3(H_0^b)^2}\,, \tag{10}$$

where the upper-script $^b$ stands for background and $H$ is the expansion scalar $H(t,r)$[6].

## 3. Reconstruction of the local metric

In order to reconstruct the metric of the local universe from the density contrast we expand the curvature function $k(r)$ as

$$k(r) = k_0 + k_1 r + k_2 r^2 + \cdots\,. \tag{11}$$

We also expand the solution of the geodesic equations according to

$$r(z) = r_1 z + r_2 z^2 + r_3 z^3 + \cdots\,, \quad \eta(z) = \eta_0 + \eta_1 z + \eta_2 z^2 + \cdots\,. \tag{12}$$

We expand $t(\eta, r)$ as

$$t(\eta,r) = t_0(r) + a(\eta_0,r)(\eta-\eta_0) + \frac{1}{2}\partial_\eta a(\eta_0,r)(\eta-\eta_0)^2 + \cdots\,, \tag{13}$$

where $t_0(r) \equiv t(\eta_0, r)$. We now define the dimensionless parameters $K_n \equiv k_n(a_0H_0)^{-(n+2)}$ and applying the chain rule to $t_0'(0)$ and $t_0''(0)$ we find

$$t_0'(0) = \frac{\partial t_0(r)}{\partial k}\frac{\partial k}{\partial r}|_{r=0} = a_0\alpha K_1\,, \quad t_0''(0) = a_0(a_0H_0)(\beta K_1^2 + 2\alpha K_2)\,, \tag{14}$$

where $\alpha$ and $\beta$ are dimensionless parameters.

We will consider the case in which $k_0 = 0$, which is enough to understand qualitatively the effects of the inhomogeneity, since this term corresponds to the homogeneous component of the curvature function $k(r)$. Expanding the density contrast up to second order we find

$$\delta(z) = \delta_0 + \delta_1 z + \delta_2 z^2 \,, \tag{15}$$

$$\delta_0 = \left(\frac{H_0}{H_0^b}\right)^2 \left(\frac{\Omega_M}{\Omega_M^b} - 1\right) \,, \tag{16}$$

$$\delta_1 = \left(\frac{H_0}{H_0^b}\right)^2 \frac{4K_1\,(3\alpha\Omega_M + 1)}{3\Omega_M^b} \,, \tag{17}$$

$$\begin{aligned}\delta_2 = &-\left(\frac{H_0}{H_0^b}\right)^2 \frac{1}{36\Omega_\Lambda\Omega_M\Omega_M^b}\Bigg[18K_1\Omega_\Lambda\Omega_M^2\,(3\alpha\Omega_M + 2) \\ &+ K_1^2\Bigg\{\Omega_\Lambda\Big(-18\left(25\alpha^2 - 4\alpha + 5\beta\right)\Omega_M^2 + 81\alpha^2\Omega_M^3 - 300\alpha\Omega_M - 40\Big) \\ &+ 20\,(\Omega_M - \zeta_0)\Bigg\} - 60K_2\Omega_\Lambda\Omega_M\,(3\alpha\Omega_M + 1)\Bigg] \,.\end{aligned} \tag{18}$$

From these equations we can finally obtain

$$k(r) \approx K_1(a_0H_0)^3 r + K_2(a_0H_0)^4 r^2 \,, \tag{19}$$

$$K_1 = \left(\frac{H_0^b}{H_0}\right)^2 \frac{3\Omega_M^b\delta_1}{4\,(3\alpha\Omega_M + 1)} \,, \tag{20}$$

$$\begin{aligned}K_2 = &\left(\frac{H_0^b}{H_0}\right)^4 \frac{3\Omega_M^b}{320\Omega_\Lambda\Omega_M\,(3\alpha\Omega_M + 1)^3}\Bigg[8\Omega_\Lambda\Omega_M\,(3\alpha\Omega_M + 1)\Bigg\{9\alpha\delta_1\Omega_M^2 \\ &+ 6\,(4\alpha\delta_2 + \delta_1)\,\Omega_M + 8\delta_2\Bigg\} + \delta_1^2\Omega_M^b\left(\frac{H_0}{H_0^b}\right)^2\Bigg\{\Omega_\Lambda\Big(-18\left(25\alpha^2 - 4\alpha\right. \\ &\left. + 5\beta\right)\Omega_M^2 + 81\alpha^2\Omega_M^3 - 300\alpha\Omega_M - 40\Big) + 20\,(\Omega_M - \zeta_0)\Bigg\}\Bigg] \,.\end{aligned} \tag{21}$$

In the above equations we have introduced the parameters $a_0$, $H_0$, $\Omega_M$, $\Omega_\Lambda$, $T_0$ and $\zeta_0$ according to their corresponding definitions given in Refs. 14, 6.

## 4. Testing the accuracy of the method

In order to test the formulae we have derived we chose models defined by the spatial curvature function $k(r)$ according to

$$k(r) = \pm(a_0H_0)^2\frac{a_0H_0r}{5}\Big[1 - \tanh\left(2a_0H_0r\right)\Big] \,. \tag{22}$$

We also compare our formula for the density contrast with the linear perturbation theory prediction that around a flat homogeneous background[16]

$$\delta(z) \approx -3\delta H(z)(\Omega_M^b)^{-0.55} \,. \tag{23}$$

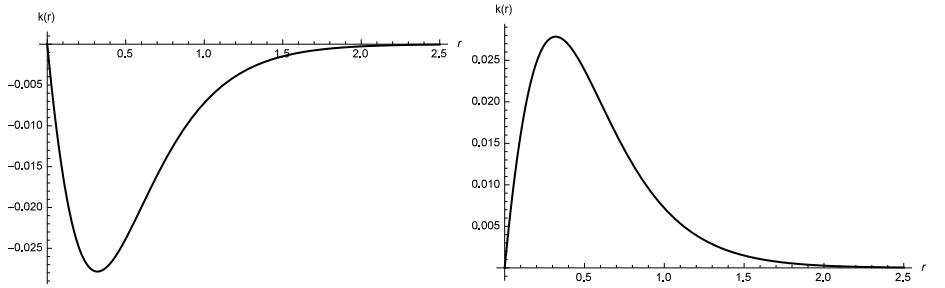



Fig. 1. The function $k(r)$ is plotted in units of $H_0^2$ as a function of the radial coordinate in units of $H_0^{-1}$.

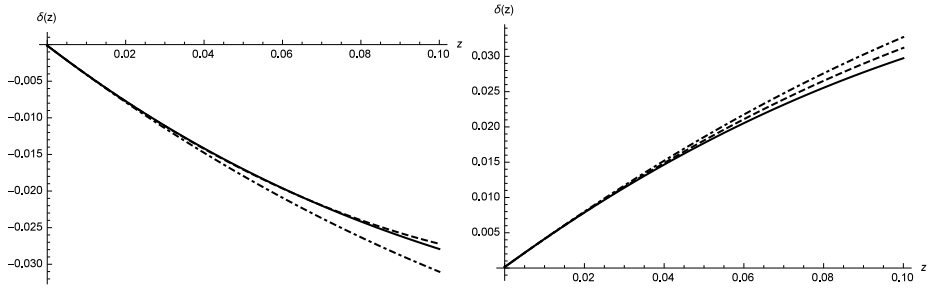



Fig. 2. The density contrast is plotted as a function of redshift. The left and right plots are for the inhomogeneities corresponding to Fig. 1. The solid lines correspond to the numerical solution, the dashed lines to the analytical formula we derived and the dot-dashed lines to the perturbation theory result.

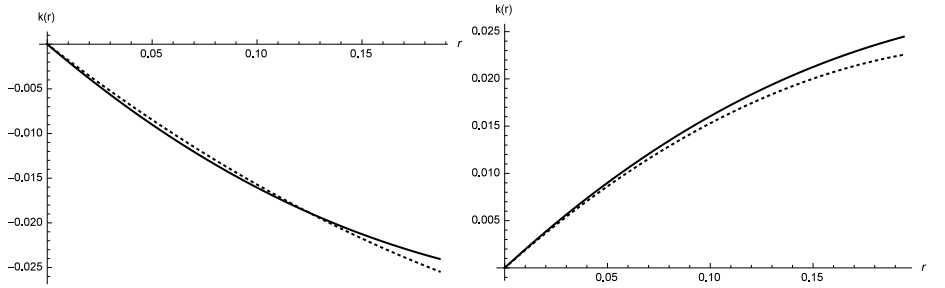



Fig. 3. The reconstructed metric function $k(r)$ is plotted in units of $H_0^2$ as a function of the radial coordinate in units of $H_0^{-1}$ for the inhomogeneities corresponding to Fig 1. The black solid line corresponds to the original $k(r)$ function and the black dotted line to the reconstructed one.

## 5. Conclusions

We have derived the low-redshift expansion for the monopole of the density contrast. At low red-shift the formula is in good agreement with numerical solutions and

is more accurate than the linear perturbation theory approximation. Using this formula we have then developed a new analytical inversion method to reconstruct the metric from the monopole of the density contrast. The inversion method could be applied to low red-shift observational data to determine the metric with a level of precision higher than the one achievable using perturbation theory.

For a full reconstruction of the metric beyond the monopole contribution other solutions of the Einstein equations could be used for the analytical approach, in order to accommodate more complex geometries. For a general numerical inversion able to reconstruct any type of metric more sophisticated methods in numerical relativity will be required.

## References


1. Planck Collaboration et al., *Astronomy & Astrophysics* **594**, A16 (2016), arXiv:1506.07135.
2. A. E. Romano and S. A. Vallejo, *Eur. Phys. Lett.* **109**, 39002 (2015), arXiv:1403.2034.
3. A. E. Romano, M. Sasaki and A. A. Starobinsky, *Eur. Phys. J. C* **72**, 2242 (2012), arXiv:1006.4735.
4. A. E. Romano and P. Chen, *Eur. Phys. J. C* **74**, 2780 (2014), arXiv:1207.5572.
5. A. E. Romano and P. Chen, *JCAP* **1110**, 016 (2011), arXiv:1104.0730.
6. A. E. Romano and S. A. Vallejo, *Eur. Phys. J. C* **76**, 216 (2016), arXiv:1502.07672.
7. G. Lemaître, *Annales de la Société Scientifique de Bruxelles* **53**, (1933).
8. G. Lemaître, *Gen. Rel. Grav.* **29**, 641 (1997).
9. G. Lemaître, *Mon. Not. Roy. Astron. Soc.* **91**, 490 (1931).
10. R. C. Tolman, *Proc. Nat. Acad. Sci.* **20**, 169 (1934).
11. H. Bondi, *Mon. Not. Roy. Astron. Soc.* **107**, 410 (1947).
12. A. Zecca, *Adv. Stud. Theor. Phys.* **7**, 1101 (2013).
13. D. Edwards, *Mon. Not. Roy. Astron. Soc.* **159**, 51 (1972).
14. A. E. Romano, *Int. J. Mod. Phys. D* **21**, 1250085 (2012), arXiv:1112.1777.
15. A. E. Romano and M. Sasaki, *Gen. Rel. Grav.* **44**, 353 (2012), arXiv:0905.3342.
16. E. L. Turner, R. Cen, and J. P. Ostriker, *Astrophys. J.* **103**, 1427 (1992).

# Gravitational lensing in area metric spacetimes

Marcus C. Werner

*Center for Gravitational Physics, Yukawa Institute for Theoretical Physics,*
*Hakubi Center for Advanced Research,*
*Kyoto University,*
*Kitashirakawa Oiwakecho Sakyoku, Kyoto 606-8502, Japan*
*E-mail: werner@yukawa.kyoto-u.ac.jp*

We consider light propagation as a probe of non-metricity in area metric spacetimes, and find a deviation from the standard Etherington relation for linearized area metric Schwarzschild. This is joint work with Frederic P. Schuller (Erlangen University).

*Keywords*: Gravitational lensing; Etherington relation; constructive gravity; area metrics.

## 1. Introduction

Since the 1919 eclipse expeditions, whose centennial will be celebrated in the year following this Marcel Grossmann Meeting, gravitational lensing has provided important precision tests of the Lorentzian metric spacetimes of general relativity. This raises the question of how modified theories with *non-metric* spacetime structures may be probed by gravitational lensing.

The answer to this question requires, of course, knowledge of the gravitational theory of such non-metric spacetime kinematics, and the standard approach is to start by stipulating some modification of the Einstein-Hilbert action as the gravitational dynamics. However, recent work in geometrodynamics (cf. Refs. 1, 2) has shown that predictive gravitational dynamics can, in fact, be *derived* from the underlying spacetime kinematics, such that the theory is predictive by construction. Hence, this new approach is called *constructive gravity*, and session AT5 of this meeting is dedicated to it.

Here, we describe the first concrete example of how constructive gravity can be employed to derive a prediction for gravitational lensing in a non-metric spacetime. Specifically, we consider area metric geometry for which a perturbative Schwarzschild-like solution has been obtained, and find a deviation from the standard metric Etherington distance duality relation. This proceedings paper is based on Ref. 3 and references therein.

## 2. Geometrical background

### 2.1. *General kinematics*

We begin by discussing general spacetime kinematics and the notion of predictivity, following Ref. 4. Consider a smooth 4-dimensional manifold $M$ endowed with smooth tensor fields $G$ and $F$, which we refer to as *geometry* and its *test matter*,

respectively, governed by a general linear field PDE,

$$\left[\sum_{d=1}^{k} D^{\bar{\mu}\nu_1\dots\nu_d}_{\bar{\lambda}}(G)\frac{\partial}{\partial x^{\nu_1}}\dots\frac{\partial}{\partial x^{\nu_d}}\right] F_{\bar{\mu}} = 0\,, \tag{1}$$

where $\nu_i \in \{0,\dots,3\}$ denote general spacetime coordinates, $\bar{\mu}$ multi-indices of test matter fields components, and $i \in \{1,\dots,d\},\ d \in \{1,\dots,k\}$ partial derivative order where $k$ is highest. Now, causal properties of Eq. (1) are governed by its principal polynomial, which emerges from the eikonal approximation: letting

$$F_{\bar{\mu}}(x,\epsilon) = e^{\frac{\mathrm{i}S(x)}{\epsilon}}\sum_{j=0}^{\infty} F_{\bar{\mu}j}(x)\epsilon^j\,, \quad \text{and} \quad \epsilon \to 0\,,$$

where $S$ is the eikonal (phase) function, then from (1),

$$e^{\frac{\mathrm{i}S(x)}{\epsilon}}\left(\frac{\mathrm{i}}{\epsilon}\right)^k\left[D^{\bar{\mu}\nu_1\dots\nu_k}_{\bar{\lambda}}(x)\frac{\partial S}{\partial x^{\nu_1}}\dots\frac{\partial S}{\partial x^{\nu_k}}\right]F_{\bar{\mu}0}(x) + \text{ lower terms in } \frac{1}{\epsilon} = 0\,,$$

whence the first term stemming from the highest derivative order remains in the limit, and for this equation to have non-trivial solutions, the determinant of the square bracketed term needs to vanish. This yields the principal polynomial of Eq. (1) $P : T^*M \to \mathbb{R}$, reducing repeated powers,

$$P(x,p) = w_G \det\left[D^{\bar{\mu}\nu_1\dots\nu_k}_{\bar{\lambda}}(x)p_{\nu_1}\dots p_{\nu_k}\right] = P^{\nu_1\dots\nu_{\deg P}}p_{\nu_1}\dots p_{\nu_{\deg P}}\,, \tag{2}$$

where $P^{\nu_1\dots\nu_{\deg P}}$ is the totally symmetric principal polynomial tensor, and $w_G$ some appropriate weight function to render $P$ a scalar. Thus, the general *null cone* (or, null dispersion relation) at $x \in M$ becomes

$$N_x = \{p \in T^*_x M \ :\ P(x,p) = 0\}\,. \tag{3}$$

We are interested in causal kinematics of the generalized spacetime $(M, G, F)$, which is determined by the Cauchy problem. Given (1) and initial data, the Cauchy problem is well-posed if there is a unique solution depending continuously on the initial data, so $P$ is necessarily hyperbolic.

So far, we have only considered covectors, or particle momenta. However, we also need the dual vectors for particle trajectories. Now it turns out that if $P$ is hyperbolic, then its dual polynomial $P^\sharp : TM \to \mathbb{R}$ exists, although hyperbolicity of $P$ does not imply hyperbolicity of $P^\sharp$. But for predictivity in the sense of distinguishing future and past, or the sign of particle energies, $P^\sharp$ needs to be hyperbolic as well. This is called *bihyperbolicity*, and is the geometric requirement for predictive kinematics.

**Example: standard Maxwell.** In $M = \mathbb{R}^4,\ x^\nu = (t, \mathbf{x})$, consider Maxwell's equations in vacuum, in suitable units. Then we have two constraint equations $\nabla \cdot \mathbf{E} = 0,\ \nabla \cdot \mathbf{B} = 0$, and two dynamical equations,

$$\frac{\partial \mathbf{E}}{\partial t} - \nabla \times \mathbf{B} = 0\,, \qquad \frac{\partial \mathbf{B}}{\partial t} + \nabla \times \mathbf{E} = 0\,.$$

Introducing $F_{\bar{\mu}} = (-\mathbf{E}, \mathbf{B})$, $\bar{\mu} \in \{1, \ldots, 6\}$, they can be recast as test matter field equations in the form of Eq. (1) to obtain

$$D^{\bar{\mu}\nu}_{\bar{\lambda}} \frac{\partial F_{\bar{\mu}}}{\partial x^{\nu}} = 0\,.$$

Now according to Eq. (2), the corresponding principal polynomial is computed from

$$\det[D^{\bar{\mu}\nu}_{\bar{\lambda}} p_\nu] = \det \begin{bmatrix} -p_0 & 0 & 0 & 0 & p_3 & -p_2 \\ 0 & -p_0 & 0 & -p_3 & 0 & p_1 \\ 0 & 0 & -p_0 & p_2 & -p_1 & 0 \\ 0 & p_3 & -p_2 & p_0 & 0 & 0 \\ -p_3 & 0 & p_1 & 0 & p_0 & 0 \\ p_2 & -p_1 & 0 & 0 & 0 & p_0 \end{bmatrix} = p_0^2(-p_0^2 + p_1^2 + p_2^2 + p_3^2)^2\,,$$

whence, after reducing repeated powers, we read off the principal polynomial,

$$P(x,p) = -p_0^2 + p_1^2 + p_2^2 + p_3^2 = P^{\mu\nu} p_\mu p_\nu = \eta^{\mu\nu} p_\mu p_\nu\,,$$

identifying the principal polynomial tensor as the inverse Minkowski metric.

### 2.2. *Area metric spacetimes*

We can now specialize the previous discussion to test matter defined by general linear electromagnetism, which has the Lagrangian

$$L = -\frac{1}{8} \mathcal{G}^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}\,, \tag{4}$$

where $F_{\mu\nu} = -F_{\nu\mu}$ is the electromagnetic tensor, and $\mathcal{G}$ is called constitutive tensor density, describing the properties of the optical medium. This is the structure defining pre-metric electromagnetism (see, e.g., Ref. 5), although in the standard Maxwell vacuum we consider $\mathcal{G}$ to be induced by a Lorentzian spacetime metric $g$. However, more generally, $\mathcal{G}$ is an area metric structure.

An *area metric* is a smooth 4th order tensor field $G$ with symmetries such that

$$G_{\mu\nu\rho\sigma} = G_{\rho\sigma\mu\nu}\,,\; G_{\mu\nu\rho\sigma} = -G_{\nu\mu\rho\sigma}\,,\; G_{\mu\nu\rho\sigma} = -G_{\mu\nu\sigma\rho}\,,$$

and we define an area metric induced by a metric $g$ as follows,

$$(G_g)_{\mu\nu\rho\sigma} = g_{\mu\rho} g_{\nu\sigma} - g_{\mu\sigma} g_{\nu\rho} - \sqrt{\det g}\; \epsilon_{\mu\nu\rho\sigma}\,.$$

The name stems from its area measuring property.[a] Now *vacuum* electromagnetism on an area metric background can be described like Eq. (4) with $F$ subject to

$$L = -\frac{1}{8} \mathcal{G}^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma} \quad \text{with} \quad \mathcal{G}^{\mu\nu\rho\sigma} = \omega_G G^{\mu\nu\rho\sigma}\,, \tag{5}$$

[a] To see this, consider two vectors $X, Y$ and a Riemannian metric $g$, then

$$G_g(X,Y,X,Y) = (X \cdot X)(Y \cdot Y) - (X \cdot Y)^2 = |X|^2 |Y|^2 \sin^2 \angle(X,Y) = |X \wedge Y|^2\,,$$

that is, the squared area spanned by $X, Y$.

where $\omega_G^{-1} = \frac{1}{4!}\epsilon_{\mu\nu\rho\sigma}G^{\mu\nu\rho\sigma}$. Thus, an area metric spacetime with predictive kinematics in the sense of Sec. 2.1 can be defined as the triple $(M, G, F)$ with a corresponding bihyperbolic principal polynomial $P$.

## 3. Gravitational lensing

### 3.1. *Light propagation*

Let us know proceed to study light propagation on such an area metric spacetime $(M, G, F)$. As discussed in general in Sec. 2.1, we can apply the eikonal approximation to the field equations of Eq. (5) to define geometrical optics. The null cone (3) in this case is, of course, found to be *quartic* ($\deg P = 4$),

$$P(x, p) = P(G)^{\alpha\beta\gamma\delta}p_\alpha p_\beta p_\gamma p_\delta = 0\,,$$

with the principal polynomial tensor constructed from the area metric $G$. It is clear, then, that light rays in such a spacetime are, in general, subject to *birefringence*.

However, since observationally any such effect would necessarily be very small, we shall from now on investigate a perturbation about an effective Minkowski spacetime. Then the (inverse) area metric is given by

$$G^{\mu\nu\rho\sigma} = \eta^{\mu\rho}\eta^{\nu\sigma} - \eta^{\mu\sigma}\eta^{\nu\rho} - \epsilon^{\mu\nu\rho\sigma} + H^{\mu\nu\rho\sigma}\,,$$

where $H$ is small. In this case, the principal polynomial tensor and its dual are found to be merely *quadratic*,

$$P^{\mu\nu} = \eta^{\mu\nu} + H^{\mu\nu}\,, \quad P^\sharp_{\mu\nu} = \eta_{\mu\nu} - H_{\mu\nu}\,, \tag{6}$$

where $H^{\mu\nu} = \frac{1}{2}H^{\alpha\mu\beta\nu}\eta_{\alpha\beta} - \frac{1}{4!}\epsilon_{\alpha\beta\gamma\delta}H^{\alpha\beta\gamma\delta}\eta^{\mu\nu}$. Hence, there is an *effective* Lorentzian metric $P^\sharp_{\mu\nu}$ for light rays, and standard ray optical arguments for light rays apply. This fact, however, does not eliminate effects of the non-metricity of this spacetimes entirely: one result important for our later discussion is that energy-momentum conservation (for the general case, cf. Ref. 6) implies that the photon current $N^\mu$ obeys the conservation law

$$(\omega_G N^\mu)_{,\mu} = 0\,, \tag{7}$$

which corresponds to the well-known result for metric spacetimes (see, e.g., Eq. (3.26) of Ref. 7), except that the volume measure here, $\omega_G$, is area metrical.

### 3.2. *Linearized gravity*

Having established our non-metric spacetime kinematics, we need gravitational dynamics next in order to obtain a concrete prediction for light propagation. But rather than stipulating a modification of general relativity, contructive gravity can be employed to derive it.

This geometrodynamical technique originates in Ref. 8, which treats the Lorentzian metric case, but has been extended recently to any tensorial background

(see Refs. 1 and 2). Briefly, the idea is that a bihyperbolic spacetime kinematics yields a generalized ADM split with lapse and shift, defining a hypersurface deformation algebra. Recognizing that kinematical deformation and dynamical evolution coincide, the constraint algebra with supermomentum and superhamiltonian can be determined, resulting eventually in a PDE system for the gravitational Lagrangian.

The gravitational dynamics for area metric kinematics can, so far, be derived perturbatively. The solution for a point mass $M$, that is, the linearized area metric Schwarzschild, is

$$\begin{aligned} G^{0a0b} \equiv& -\gamma^{ab} + H^{0a0b} = -\gamma^{ab} + (2A - \tfrac{1}{2}U + \tfrac{1}{2}V)\gamma^{ab}\,, \\ G^{0bcd} \equiv& \epsilon^{bcd} + H^{0bcd} = \epsilon^{bcd} + \left(\tfrac{3}{4}U - \tfrac{3}{4}V - A\right)\epsilon^{bcd}\,, \\ G^{abcd} \equiv& \gamma^{ac}\gamma^{bd} - \gamma^{ad}\gamma^{bc} + H^{abcd} = (1 + 2U - V)(\gamma^{ac}\gamma^{bd} - \gamma^{ad}\gamma^{bc})\,, \end{aligned}$$

with Euclidean metric $\gamma^{ab}$, Euclidean distance $r$ of the unperturbed background, Levi-Civita symbol $\epsilon^{abc}$, and scalar perturbations

$$A = -\frac{M}{8\pi r}\left(\kappa - \lambda\eta e^{-\mu r}\right)\,,\ U = -\frac{M}{4\pi r}\eta e^{-\mu r}\,,\ V = \frac{M}{4\pi r}\left(\kappa - \tau\eta e^{-\mu r}\right)\,,$$

where $\eta, \kappa, \lambda, \mu, \tau$ are constants.

### 3.3. *Modified Etherington relation*

Finally, we are ready to study light propagation in this non-metric spacetime under the influence of gravity. One fundamental kinematical property of *any* Lorentzian metric spacetime is the Etherington distance duality relation, originally derived in Ref. 9, which connects the luminosity distance, angular diameter distance and redshift. Thus, we naturally expect this relation to be broken for a non-metric spacetime, such as our area metric spacetime, possibly even in a way such that the result is no longer merely kinematical but dynamical, that is, dependent on the gravity theory. This turns out to be the case here, as we shall see in the following.

First of all, recall from Sec. 3.1 that, at leading order in geometrical optics, area metric light propagation is essentially metric, so the standard ray optical derivation of the Etherington relation (e.g., Ref. 7, ch. 3) largely applies, yielding

$$D_L = \frac{(1+z)^2 D_A}{\sqrt{1+\Delta}}\,, \tag{8}$$

where $D_L,\ D_A,\ z$ are luminosity distance, angular diameter distance and redshift, respectively, as usual, and $\Delta$ is the photon excess fraction in a ray bundle domain $\mathcal{D}$ from the light source to the observer obeying

$$\Delta \cdot N = \int_{\mathcal{D}} \mathrm{d}^4x \sqrt{-\det P^\sharp}\ N^\mu{}_{;\mu}\,, \tag{9}$$

where the covariant derivative is with respect to metric $P^\sharp$ of (6), $N$ is the photon number and $N^\mu$ the photon current, as before. In standard metric theory, photon conservation implies the vanishing of $N^\mu{}_{;\mu}$ in Eq. (9) whence $\Delta = 0$ in Eq. (8), and

we recover the standard Etherington relation. However, in the area metric case, although conservation law (7) applies, the $N^{\mu}{}_{;\mu}$ does not vanish whence $\Delta \neq 0$, and we find a modification of the Etheringon relation. Concretely, for linearized area metric Schwarzschild of Sec. 3.2, Yukawa-type corrections emerge,

$$D_L = (1+z)^2 D_A \left(1 + \frac{3\kappa M}{8\pi}\left(\frac{e^{-\mu r_s}}{r_s} - \frac{e^{-\mu r_o}}{r_o}\right)\right),$$

with Euclidean distances $r_o, r_s$ between the mass $M$ and the observer and light source, respectively. Thus, although at the optical geometry (and hence image positions and time delays) is the same as for metric Schwarzschild at leading order in $M$, the area metric volume measure implies a difference in the image magnification.

## 4. Concluding remarks

The constructive gravity approach allows the derivation of predictive gravitational dynamics from bihyperbolic kinematics. This enables the study of spacetimes beyond metric geometry without additional assumptions about the modified gravity theory. Thus, qualitatively new gravitational lensing effects due to non-metricities can be derived to provide possible observational tests. The first concrete example of this was described here, resulting in a modified Etherington relation for the linearized area metric Schwarzschild spacetime. The interested reader is also referred to session AT5 *Constructive Gravity* of this Marcel Grossmann Meeting.

## References


1. K. Giesel, F. P. Schuller, C. Witte and M. N. R. Wohlfarth, *Phys. Rev. D* **85**, 104042 (2012).
2. M. Düll, F. P. Schuller, N. Stritzelberger and F. Wolz, *Phys. Rev. D* **97**, 084036 (2018).
3. F. P. Schuller and M. C. Werner, *Universe* **3**, 52 (2017).
4. D. Rätzel, S. Rivera and F. P. Schuller, *Phys. Rev. D* **83**, 044047 (2011).
5. F. W. Hehl and Y. N. Obukhov, *Foundations of Classical Electrodynamics* (Birkhäuser, Boston, 2003).
6. M. J. Gotay, J. Isenberg and J. E. Marsden, `arXiv:physics/9801019` (1998), `arXiv:math-ph/0411032` (2004).
7. P. Schneider, J. Ehlers and E. E. Falco, *Gravitational Lenses* (Springer, Berlin, 1992).
8. S. A. Hojman, K. Kuchař and C. Teitelboim, *Ann. Phys.* **96**, 88 (1976).
9. I. M. H. Etherington, *Phil. Mag.* **15**, 761 (1933).

# Review of tokamak physics and a way to get big bang GW conditions before the $10^{-26}$ reduction in frequency seen today, confirming *e*-fold values of 60 for inflation

Andrew Walcott Beckwith

*Physics Department, Chongqing University, Chongqing 401331, People's Republic of China*
*E-mail: Rwill9955b@gmail.com; abeckwith@uh.edu*
*english.cqu.edu.cn*

We reconstruct early-universe GW conditions, before the onset of inflation. The idea is as follows. e-Fold expansion of the universe due to inflation is of the order of $10^{26}$. Hence, the goal is to produce relic gravitational waves with a device having a signal with frequencies around $10^{10}$ Hz, which is in line with todays theoretical GW frequencies of $10^{16}$ Hz for relic big-bang GW detected in the present era. Measuring $10^{10}$ Hz in a laboratory would be a proof of the calculated e-fold value of, say, 60.

*Keywords*: Tokamak Physics, Plasma Confinement Time, GW Amplitude, Drift Current.

## 1. Introduction

The main justification of our work is that we could, with due diligence, reconstruct GW radiation from the start of inflationary physics. First, there is the question of what sort of polarization would be produced in initial processes. Second, if we were able to produce $10^{10}$ Hz gravitational waves via our laboratory arrangements and if we were later able to confirm, say, the existence of $10^{-16}$ Hz GW frequencies via LISA in the present era, this would be a stunning proof of the big-bang hypothesis. So, we summarize what this inquiry may answer.

(1) Determination of the fidelity of the e-fold value of 60 in the big bang.
(2) Determination if a drop of say $10^{-2}$ in strain value, $h$, for GW, from the initial configuration of the big bang, to our present era, is supportable.
(3) Unknown initial GW polarization from the start of the big bang. Knowing (1) and (2), we could at least say which GW polarization states may be permitted.

## 2. Comparison with Grishchuk and Sazhin Results

Russian physicists obtained the amplitude of a Gravitational wave (GW) in a plasma as $A = h \sim \frac{G}{c^4} \cdot E^2 \cdot \lambda_{\rm GW}^2$.[1] Note that a simple model of how to provide a current, $I$, in the Toroid is provided by a transformer core using the simple Ohm's Law derivation referred to in the first part of the text. Here, $E$ is the electric field whereas $\lambda_{\rm GW}$ is the gravitational wavelength for GW generated by the tokamak in our model. In the original Grishchuk model, we would have very small strain values, which requires the following relationship between GW wavelength and resultant frequency. Note, $\omega_{\rm GW} \sim 10^6$ Hz $\Rightarrow \lambda_{\rm GW} \sim 300$ m, so we will be assuming $\omega_{\rm GW} \sim 10^9$ Hz $\Rightarrow \lambda_{\rm GW} \sim 0.3$ m as a baseline measurement for GW detection above the tokamak.

Furthermore,

$$A \sim h \sim \frac{G \cdot W_E \cdot V}{c^4 \cdot \widetilde{a}} \tag{1}$$

where $W_E$ is the energy density, $V$ is the volume, and $\widetilde{a}$ is the toroid's inner radius.

Equation (1) is due to the two-part strain's first term, with the strain's second term significantly larger than the first term and due to ignition of the tokamak plasma. The strain's first term is largely due to Grishchuk and Sazhin's calculations.[1] The plasma-fusion contribution is due to nonequilibrium contributions to plasma ignition, which we will elaborate. Note that the first term in the strain derivation is due to the electric field within the toroid, not Plasma fusion, and we will first of all discuss how to obtain the requisite strain for the electric-field contribution to the current inside a tokamak via Ohm's Law.

## 3. Derivation of Strain Generated by an Electric Field

We examine the would-be electric field, contributing to small strain values similar in part to Ohm's Law. A generalized Ohm's Law ties in well.

$$J = \sigma \cdot E \tag{2}$$

To obtain an electric field detectable by 3DSR,[4,5] we use generalized Ohm's Law (see Ref. 3, p. 146), where $E$ and $B$ are electric and magnetic fields, and $v$ is velocity.

$$E = \sigma^{-1} J - v \times B \tag{3}$$

The term $v \times B$ in Eq. (3) deserves special commentary. If $v$ is perpendicular to $B$, as occurs in a simple equilibrium case, then of course, Eq. (3) would be, simply put, Ohm's Law, and spatial equilibrium averaging would then lead to

$$E = \sigma^{-1} J - v \times B \xrightarrow[v \text{ perpendicular to } B]{} E = \sigma^{-1} J. \tag{4}$$

What saves the contribution of Plasma burning as a contributing factor to the tokamak generation of GW, with far larger strain values, is that the ion velocity in the plasma is not perpendicular to $B$ in the beginning of tokamak generation. It is, fortunately for us, a nonequilibrium initial process, with thermal irregularities leading to both terms in Eq. (3) contributing to the electric field values. We will be looking for an application for radial free-electric fields being applied (see Ref. 3, p. 120).

$$n_j e_j \cdot (E_{\mathrm{r}} + v_{\perp j} B) = -\frac{\mathrm{d}P_j}{\mathrm{d}r} \tag{5}$$

Here, $n_j$ is ion density of the $j^{\mathrm{th}}$ species, $e_j$ is ion charge of the $j^{\mathrm{th}}$ species, $E_{\mathrm{r}}$ is the radial electric field, $v_{\perp j}$ is the perpendicular velocity of the $j^{\mathrm{th}}$ species, $B$ is the magnetic field, and $P_j$ is the pressure of the $j^{\mathrm{th}}$ species. The results of Eq. (3) and Eq. (5) are

$$\frac{G}{c^4} \cdot E^2 \cdot \lambda_{\mathrm{GW}}^2 \sim \frac{G}{c^4} \cdot \left(\frac{C}{R}\right)^2 \cdot \lambda_{\mathrm{GW}}^2 + \frac{G}{c^4} \cdot \left(\frac{J_b}{n \cdot e} + v_{\mathrm{r}}\right)^2 \lambda_{\mathrm{GW}}^2 = (\text{first}) + (\text{second}) \tag{6}$$

Here, the first term is due to $\nabla \times E = 0$, and the second term is due to $E_{\rm n} = \frac{{\rm d}P_j}{{\rm d}x_{\rm n}} \cdot \frac{1}{n_j \cdot e_j} - (v \times B)_{\rm n}$ with the first term generating $h \sim 10^{-38}$–$10^{-30}$ in terms of GW amplitude strain 5 m above the tokamak. In contrast, the second term has a strain of $h \sim 10^{-26}$ in terms of GW amplitude above the tokamak. The article has contributions from amplitude from the first and second terms separately. The second part will be tabulated separately from the first contribution assuming a minimum temperature of $T \sim 10$ keV.[3]

## 4. GW *h* Strain Values When the First Term of Eq. (6) Is Used

We now look at what we can expect with the simple Ohm's Law calculation for strain values. As it is, the effort led to unusable GW amplitude values of up to $h \sim 10^{-38}$–$10^{-30}$ for GW wave amplitudes 5 m above a tokamak, and $h \sim 10^{-38}$–$10^{-28}$ in the center of a tokamak. This would be the tokamak-generated GW amplitude using the first term of Eq. (6) to obtain the following value[1]

$$h_{\rm 1st} \sim \frac{G}{c^4} \cdot E^2 \cdot \lambda_{\rm GW}^2 \sim \frac{G}{c^4} \cdot \left(\frac{J}{\sigma}\right)^2 \cdot \lambda_{\rm GW}^2. \tag{7}$$

We summarize the results for when $\omega_{\rm GW} \sim 10^9$ Hz $\Rightarrow \lambda_{\rm GW} \sim 0.3$ m and with conductivity $\sigma \sim 10$ m$^2$/s and with the following provisions as to initial values. We observe is a range of tokamak values that are, even in the case of ITER (not yet built) beyond the reach of any technological detection devices in the coming decade. These results, assuming fixed conductivity values $\sigma \sim 10$ m$^2$/s as well as $\lambda_{\rm GW} \sim 0.3$ m, are why the author expects the second term of Eq. (3) to dominate the results for the Chinese tokamak in Hefei.[6]

$$h_{\rm 2nd} \sim \frac{G}{c^4} \cdot E^2 \cdot \lambda_{\rm GW}^2 \sim \frac{G}{c^4} \cdot \left(\frac{J}{n \cdot e} + v_{\rm r}\right)^2 \cdot \lambda_{\rm GW}^2, \tag{8}$$

or values 10000 larger than the results in ITER due to Eq. (6).

We are setting $\lambda_{\rm GW} \sim 0.3$ m and $\sigma \sim 10$ m$^2$/s, using Eq. (6) above for GW amplitude. What makes the second term of Eq. (4) mandatory is that even in the case of ITER, 5 m above the tokamak ring, the GW amplitude is $\frac{1}{10000}$ the size of any reasonable GW detection device, including the new 3DSR technology.[4,5] Hence, we need to come up with a better estimate, which the second term of Eq. (6) provides.

## 5. Enhancing GW Strain Amplitude via a Burning Plasma Drift Current: Eq. (4)

We next examine the normal-to-surface induced electric field contribution (see Ref. 3, p. 120):

$$E_{\rm n} = \frac{{\rm d}P_j}{{\rm d}x_{\rm n}} \cdot \frac{1}{n_j \cdot e_j} - (v \times B)_{\rm n}. \tag{9}$$

If one has for $v_{\rm r}$ as the radial velocity of ions in the tokamak from tokamak center to its radial distance, $R$, from center, and $B_\theta$ as the direction of a magnetic

field in the "face" of a Toroid containing the plasma, in the angular $\theta$ direction from a minimal toroid radius of $R = a$, with $\theta = 0$, to $R = a + r$ with $\theta = \pi$, one has $v_{\rm r}$ for radial drift velocity of ions in the tokamak, and $B_\theta$ having a net approximate value of with $B_\theta$ not perpendicular to the ion velocity, so then [3]

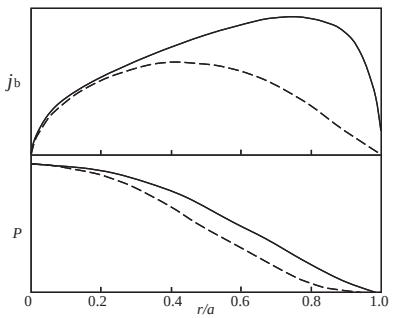



Fig. 1. Typical bootstrap currents with a shift due to $\frac{r}{a}$ where $r$ is the radial direction of the tokamak, and $a$ is the inner radius of the Toroid This figure is reproduced from Wesson. [3]

$$(v \times B)_{\rm r} \sim v_{\rm r} \cdot B_\theta. \tag{10}$$

Also, spatial change in pressure denoted (see Ref. 3, p. 167)

$$\frac{{\rm d}P_j}{{\rm d}x_{\rm n}} = -B_\theta \cdot j_b. \tag{11}$$

Here the drift current, using $\xi = \frac{a}{R}$, and drift current $j_b$ for plasma charges, is

$$j_b \sim -\frac{\xi^{\frac{1}{2}}}{B_\theta} \cdot T \cdot \frac{{\rm d}n_{\rm drift}}{{\rm d}r}. \tag{12}$$

Figure 1 introduces the role of the drift current, in terms of a tokamak. [3]

Then one has

$$B_\theta^2 \cdot \left(\frac{j_b}{n_j \cdot e_j}\right)^2 \sim \frac{B_\theta^2}{e_j^2} \cdot \frac{\xi^{\frac{1}{4}}}{B_\theta^2} \cdot \left(\frac{1}{n_{\rm drift}} \cdot \frac{{\rm d}n_{\rm drift}}{{\rm d}r}\right)^2 \sim \frac{\xi^{\frac{1}{4}}}{e_j^2} \cdot \left(\frac{1}{n_{\rm drift}} \cdot \frac{{\rm d}n_{\rm drift}}{{\rm d}r}\right)^2. \tag{13}$$

Now, the behavior of the numerical density of ions grows in the radial direction, so [3]

$$n_{\rm drift} = n_{\rm drift}|_{\rm initial} \cdot {\rm e}^{\widetilde{\alpha} \cdot r}. \tag{14}$$

This exponential behavior then will lead to the second term in Eq. (6) having in the center of the tokamak, for an ignition temperature of $T \geq 10$ keV a value of

$$h_{\rm 2nd} \sim \frac{G}{c^4} \cdot B_\theta^2 \cdot \left(\frac{j_b}{n_j \cdot e_j}\right)^2 \cdot \lambda_{\rm GW}^2 \sim \frac{G}{c^4} \cdot \frac{\xi^{\frac{1}{4}} \widetilde{\alpha} T^2}{e_j^2} \cdot \lambda_{\rm GW}^2 \sim 10^{-25}. \tag{15}$$

There is a critical ignition temperature [3] at its lowest point of the curve having $T \geq 30$ keV as an optimum value of the tokamak ignition temperature for $n_{\rm ion} \sim 10^{20}$ m$^{-3}$, with a still permissible temperature value of $T|_{\rm safe\ upper\ bound} \approx 100$ keV with a value of $n_{\rm ion} \sim 10^{20}$ m$^{-3}$ (see Ref. 3, p. 11). The relationship of Eq. (16), where $\tau_E$ is a tokamak confinement-of-plasma time of about one to three seconds. [3] Then $n_{\rm ion} \cdot \tau_E > 0.5 \times 10^{20}$ m$^{-3}$s. Also, if $T|_{\rm safe\ upper\ bound} \approx 100$ keV, then one could have at the tokamak center, even the Hefei-based tokamak, [3,6]

$$h_{\rm 2nd}|_{T \geq 100\ {\rm keV}} \sim \frac{G}{c^4} \cdot \frac{\xi^{\frac{1}{4}} \widetilde{\alpha} T^2}{e_j^2} \cdot \lambda_{\rm GW}^2 \sim 10^{-25}\text{–}10^{-26}. \tag{16}$$

This would lead to, for a GW reading 5 meter above the tokamak, [3,6]

$$\left[h_{\rm 2nd}|_{T \geq 100\ {\rm keV}}\right]_{\rm 5\ m\ above\ tokamak} \sim \frac{G}{c^4} \cdot \frac{\xi^{\frac{1}{4}} \widetilde{\alpha} T^2}{e_j^2} \cdot \lambda_{\rm GW}^2 \sim 10^{-25}\text{–}10^{-26}. \tag{17}$$

Note that the support for up to 100 keV for temperature can yield more stability in terms of thermal plasma confinement

## 6. Restating the Energy Density and Power

$$W_E \cdot V \sim \widetilde{\alpha} \cdot \lambda_{\mathrm{GW}}^2 \cdot \frac{\xi^{\frac{1}{4}} \widetilde{\alpha} T_{\mathrm{plasma}}^2}{e_j^2} \tag{18}$$

The temperature for plasma fusion burning is between 30 and 100 keV.[3] The corresponding power is then, for the tokamak,[3]

$$P_\Omega = E \cdot J \leq \frac{E}{\mu_0} \cdot \frac{B_\phi}{R} \tag{19}$$

The tie in with Eq. (18) by Eq. (20) can be seen by first setting the $E$ field as related to the $B$ field, via $E$ (electrostatic) $\approx 10^{12} V$ m$^{-1}$ as equivalent to a magnetic field $B \approx 10^4$ T.[2] In a one-second interval, if we use the input power as an experimentally supplied quantity, then the effective $E$ field is

$$E_{\mathrm{applied}} \sim \frac{\xi^{\frac{1}{8}} \cdot \widetilde{\alpha}}{e_j} \times T_{\mathrm{plasma}} \tag{20}$$

What is found is, that if Eq. (19) and Eq. (20) hold, then if $Z_{\mathrm{effective}} \sim 1.5$, $q_a Q_0 \sim 1.5$, and $\frac{R}{\widetilde{a}} \approx 3$, then the temperature of a tokamak, to a good approximation, would be between 30 and 100 keV (see Ref. 3, pp. 242–243). Then one has[3]

$$B_\phi^{\frac{4}{5}} \sim 0.87 \cdot (\widetilde{T} = T_{\mathrm{plasma}}). \tag{21}$$

Then the power for the tokamak is

$$P_\Omega = \frac{\xi^{\frac{1}{8}} \cdot \widetilde{\alpha}}{\mu_0 \cdot e_j \cdot R} \times \frac{T_{\mathrm{plasma}}^{\frac{9}{4}}}{0.87^{\frac{5}{4}}} \tag{22}$$

Then, $n$ $10^{-34}$ eV gravitons are produced per second, as, if $\widetilde{a} = \frac{R}{3}$:

$$n \propto \frac{3 \cdot \hbar \cdot e_j}{\mu_0 \cdot R^2 \cdot \xi^{\frac{1}{8}} \cdot \widetilde{\alpha}} \times \frac{T_{\mathrm{plasma}}^{\frac{1}{4}}}{0.87^{\frac{5}{4}} \cdot \lambda_{\mathrm{g}}^2 \cdot m_{\mathrm{g}}} \sim \frac{1}{\lambda_{\mathrm{g}}^2} \text{ scaling.} \tag{23}$$

## 7. Conclusion

Further elaboration of this matter in the experimental detection of experimental data sets for massive gravity lies in the viability of the expression derived, namely Eq. (19). $h \sim 10^{-27}$ for a GW detected 5 m above a tokamak represents a decrease in strain by a factor of about 100,[7] whereas in the center of the tokamak, we would have, say, $h_{2\mathrm{nd}} \sim 10^{-26}$–$10^{-27}$, a difference of two orders of magnitude. We state that our rough estimate is that we would see about the same strain values. In the initial starting point of the universe, we would have, say $h \approx 10^{-25}$ decreasing

to $h_{2\mathrm{nd}} \sim 10^{-26}$–$10^{-27}$ today, a comparatively small change in strain amplitude. Contrast this with e-folding issues,[8] wherein we would have a difference of $10^{26}$ in frequency magnitude, with $10^{10}$ Hz initially, for GW at the start of the big bang, decreasing to $10^{-16}$ Hz, due to inflation.[8] If we confirm that last statement observationally, we have confirmed the e-folding prediction.[8]

**Acknowledgements**

The author thanks Dr. Yang Xi for guidance at the U of Houston. This work is supported in part by National Natural Science Foundation of China grant No. 11375279.

**References**

1. L. P. Grishchuk and M. V. Sazhin, Excitation and detection of standing gravitational waves, *Zh. Eksp. Theor. Fiz.* **68**, 1569 (1975).
2. F. Li, M.-X. Tang, J. Luo and Y.-C. Li, Electrodynamical response of a high energy photon flux to a gravitational wave, *Phys. Rev. D* **62**, Article 044018 (2000), https://doi.org/10.1103/PhysRevD.62.044018.
3. J. Wesson, *Tokamaks*, International Series of Monographs on Physics, Vol. 149, 4th edn. (Oxford Science, Oxford, England, 2011).
4. F. Li, N. Yang, Z. Fang, R. M. L. Baker, Jr., G. V. Stephenson and H. Wen, Signal photon flux and background noise in a coupling electromagnetic detecting system for high frequency gravitational waves, *Phys. Rev. D* **80**, Article 064013 (2009), https://doi.org/10.1103/PhysRevD.80.064013.
5. R. C. Woods, R. M. L. Baker, F. Li, G. V. Stephenson, E. W. Davis and A. W. Beckwith, A new theoretical technique for the measurement of high-frequency relic gravitational waves, *J. Mod. Phys.* **2**, 498 (2011), https://doi.org/10.4236/jmp.2011.26060.
6. J. Li, H. Y. Guo, B. N. Wan, X. Z. Gong, Y. F. Liang, G. S. Xu, K. F. Gan, J. S. Hu, H. Q. Wang, L. Wang, L. Zeng, Y. P. Zhao, P. Denner, G. L. Jackson, A. Loarte, R. Maingi, J. E. Menard, M. Rack and X. L. Zou, A long-pulse high-confinement plasma regime in the Experimental Advanced Superconducting Tokamak, *Nat. Phys.* **9**, 817 (2013), https://doi.org/10.1038/nphys2795.
7. A. Beckwith, Part 2: Review of tokamak physics as a way to construct a device optimal for graviton detection and generation within a confined small spatial volume, as opposed to Dyson's "infinite astrophysical volume" calculations, *J. High Energy Phys. Gravitation Cosmol.* **3**, 138 (2017), https://doi.org/10.4236/jhepgc.2017.31015.
8. S. Chongchitnan, Inflation model building with an accurate measure of $e$-folding, *Phys. Rev. D* **94**, Article 043526 (2016), https://doi.org/10.1103/PhysRevD.94.043526.

# Spin in gravitational and electromagnetic fields

Yuri N. Obukhov

*Russian Academy of Sciences, Nuclear Safety Institute (IBRAE),*
*B. Tulskaya 52, 115191 Moscow, Russia*
*E-mail: obukhov@ibrae.ac.ru*

A unified approach to the study of classical and quantum spin in external fields is developed. Understanding the dynamics of particles with spin and dipole moments in arbitrary gravitational, inertial and electromagnetic fields is important in astrophysics and high-energy and heavy-ion experimental physics.

*Keywords*: Spin, dipole moments, nonminimal interactions.

## 1. Projection method for spin in arbitrary external fields

One can develop a natural extension of the classic Frenkel-Thomas spin model which is based on the definition of the "magnetic" and "electric" components of the relativistic forces and moments of forces acting on a particle (hence it is natural to call this approach a projection method). Let us recall that in Maxwell's theory, electrodynamic phenomena are described by the field strength tensor $F_{\alpha\beta}$, and the magnetic and electric fields can be introduced in a covariant way as the longitudinal and transversal projections on the velocity $U^\alpha$:

$$E_\alpha := F_{\alpha\beta}U^\beta, \qquad B^\alpha := \frac{1}{2c}\eta^{\alpha\mu\nu\beta}F_{\mu\nu}U_\beta. \tag{1}$$

Here $\eta_{\alpha\beta\mu\nu}$ is the totally antisymmetric Levi-Civita tensor. Thereby one gets an unique representation of the electromagnetic field strength tensor in terms of its projections (electric and magnetic fields):

$$F_{\alpha\beta} = \frac{1}{c^2}\left(E_\alpha U_\beta - E_\beta U_\alpha + c\eta_{\alpha\beta\mu\nu}U^\mu B^\nu\right). \tag{2}$$

In the model of a classical particle with internal degrees of freedom (the generalized Frenkel-Thomas model), the motion of a test particle is characterized by the 4-velocity $U^\alpha$ and the 4-vector of spin $S^\alpha$, which satisfy the normalization $U_\alpha U^\alpha = c^2$ and the orthogonality condition $S_\alpha U^\alpha = 0$. In general, the dynamic equations for these variables can be written as

$$\frac{dU^\alpha}{d\tau} = \mathcal{F}^\alpha, \qquad \frac{dS^\alpha}{d\tau} = \Phi^\alpha{}_\beta S^\beta. \tag{3}$$

External fields of different physical nature (electromagnetic, gravitational, scalar, etc.) determine the forces $\mathcal{F}^\alpha$ acting on the particle, as well as the spin transport matrix $\Phi^\alpha{}_\beta$ which affects the spin. Normalization and orthogonality of the velocity and spin vectors impose conditions on the right-hand sides of (3):

$$U_\alpha\mathcal{F}^\alpha = 0, \qquad U_\alpha\Phi^\alpha{}_\beta S^\beta = -S_\alpha\mathcal{F}^\alpha. \tag{4}$$

The spin transport matrix must be skew-symmetric, $\Phi_{\alpha\beta} = -\Phi_{\beta\alpha}$, which automatically guarantees $S_\alpha S^\alpha$ =const. The relativistic 4-velocity vector $U^\alpha$ is conveniently parametrized via the spatial velocity 3-vector $\widehat{v}^a$ and the Lorentz factor $\gamma$:

$$U^\alpha = \begin{pmatrix} \gamma \\ \gamma\widehat{v}^a \end{pmatrix}, \qquad \gamma = \frac{1}{\sqrt{1-\widehat{v}^2/c^2}}, \qquad \widehat{v}^2 = \delta_{ab}\widehat{v}^a\widehat{v}^b. \tag{5}$$

The vector of the generalized force $\mathcal{F}^\alpha$ acting on a particle has only a transversal projection, according to (4). As for the spin transport matrix $\Phi_{\alpha\beta}$, like any bivector, we can decompose it into a pair of 4-vectors. Namely, by analogy with (1), we define "electric" and "magnetic" projections

$$N_\alpha := \Phi_{\alpha\beta}U^\beta, \qquad Q^\alpha := \frac{1}{2c}\eta^{\alpha\mu\nu\beta}\Phi_{\mu\nu}U_\beta. \tag{6}$$

By construction, these 4-vectors are orthogonal to the velocity of the particle,

$$U_\alpha N^\alpha = 0, \qquad U_\alpha Q^\alpha = 0, \tag{7}$$

and the relations (4) are then recast into

$$U_\alpha \mathcal{F}^\alpha = 0, \qquad S_\alpha N^\alpha = S_\alpha \mathcal{F}^\alpha. \tag{8}$$

As a result, we have an unique decomposition analogous to (2):

$$\Phi_{\alpha\beta} = \frac{1}{c^2}\left(N_\alpha U_\beta - N_\beta U_\alpha + c\eta_{\alpha\beta\mu\nu}U^\mu Q^\nu\right). \tag{9}$$

The orthogonality conditions (7) can be resolved so that

$$N^0 = \frac{1}{c^2}(\boldsymbol{N}\cdot\widehat{\boldsymbol{v}}), \tag{10}$$

where as usual $\widehat{\boldsymbol{v}} = \{\widehat{v}^a\}$ and $\boldsymbol{N} = \{N^a\}$. Accordingly, we find in components:

$$\Phi^a{}_0 = \gamma\left(N^a - \widehat{v}^a N^0 - \epsilon^a{}_{bc}Q^b\widehat{v}^c\right), \tag{11}$$

$$\Phi^0{}_a = \frac{\gamma}{c^2}\left(N_a - \widehat{v}_a N^0 - \epsilon_{abc}Q^b\widehat{v}^c\right), \tag{12}$$

$$\Phi^a{}_b = \frac{\gamma}{c^2}\left(\widehat{v}^a N_b - \widehat{v}_b N^a\right) - \gamma\epsilon^a{}_{bc}\left(Q^c - \widehat{v}^c Q^0\right). \tag{13}$$

The physical spin, as an "internal angular momentum" of a particle, is defined with respect to particle's rest frame system, in which $u^\alpha = \delta^\alpha_0$. The transition to this system is carried out by the Lorentz transformation $U^\alpha = \Lambda^\alpha{}_\beta u^\beta$, where

$$\Lambda^\alpha{}_\beta = \left(\begin{array}{c|c} \gamma & \gamma\widehat{v}_b/c^2 \\ \hline \gamma\widehat{v}^a & \delta^a_b + \frac{\gamma-1}{\widehat{v}^2}\widehat{v}^a\widehat{v}_b \end{array}\right). \tag{14}$$

Therefore, the dynamic equation for the *physical spin* $s^\alpha = (\Lambda^{-1})^\alpha{}_\beta S^\beta$ reads:

$$\frac{ds^\alpha}{d\tau} = \Omega^\alpha{}_\beta s^\beta, \tag{15}$$

where the tensor of angular precession of spin is constructed as

$$\Omega^\alpha{}_\beta = (\Lambda^{-1})^\alpha{}_\gamma \Phi^\gamma{}_\delta \Lambda^\delta{}_\beta - (\Lambda^{-1})^\alpha{}_\gamma \frac{d}{d\tau}\Lambda^\gamma{}_\beta. \tag{16}$$

The 0th component of (15) vanishes, which is equivalent to the second condition (4). As a result, the spin evolution equation (15) reduces to the 3-vector form

$$\frac{ds^a}{d\tau} = \Omega^a{}_b s^b, \qquad \text{or} \qquad \frac{d\boldsymbol{s}}{d\tau} = \boldsymbol{\Omega} \times \boldsymbol{s}. \tag{17}$$

Here the components of 3-vectors are introduced through $\boldsymbol{s} = \{s^a\}$ and $\boldsymbol{\Omega} = \left\{-\frac{1}{2}\epsilon^{abc}\Omega_{bc}\right\}$. A direct calculation gives the precession angular velocity in terms of the magnetic and electric projections of the spin transport matrix:

$$\boldsymbol{\Omega} = \gamma\left\{\boldsymbol{Q} - \widehat{\boldsymbol{v}}\,Q^0 + \frac{\gamma}{\gamma+1}\frac{1}{c^2}\left[\widehat{\boldsymbol{v}}\,(\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q}) - \widehat{v}^2\boldsymbol{Q}\right]\right\} + \frac{\gamma}{\gamma+1}\frac{\boldsymbol{\mathcal{F}}\times\widehat{\boldsymbol{v}}}{c^2}. \tag{18}$$

It is worthwhile to note that $\boldsymbol{\Omega}$ actually depends only on the transversal part of the "magnetic" vector, namely $Q^\alpha_\perp = Q^\alpha - \frac{1}{c^2}U^\alpha U_\beta Q^\beta$. Explicitly in components

$$Q^0_\perp = \frac{\gamma^2}{c^2}\left[\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q} - \widehat{v}^2 Q^0\right], \qquad Q^a_\perp = Q^a + \frac{\gamma^2}{c^2}\left[\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q} - c^2 Q^0\right]\widehat{v}^a, \tag{19}$$

and by a direct computation one can check that

$$\boldsymbol{Q}_\perp - \widehat{\boldsymbol{v}}\,Q^0_\perp = \boldsymbol{Q} - \widehat{\boldsymbol{v}}\,Q^0, \qquad \widehat{\boldsymbol{v}}\,(\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q}_\perp) - \widehat{v}^2\boldsymbol{Q}_\perp = \widehat{\boldsymbol{v}}\,(\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q}) - \widehat{v}^2\boldsymbol{Q}. \tag{20}$$

From (19) one then finds

$$Q^0_\perp = \frac{1}{c^2}\,\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q}_\perp, \tag{21}$$

and therefore we recast (18) into a final form

$$\boldsymbol{\Omega} = \boldsymbol{Q}_\perp - \frac{\gamma}{\gamma+1}\frac{\widehat{\boldsymbol{v}}\,(\widehat{\boldsymbol{v}}\cdot\boldsymbol{Q}_\perp)}{c^2} - \frac{\gamma}{\gamma+1}\frac{\widehat{\boldsymbol{v}}\times\boldsymbol{\mathcal{F}}}{c^2}. \tag{22}$$

The new general equations (22), (17) are valid for a particle with spin that interacts with arbitrary external fields. The actual dynamics of the physical spin depends on the forces acting on the particle and on the spin transport law.

## 2. Application: Spin in the gravitational and electromagnetic fields

In order to consider the most general case, we assume that the gravitational field is described by an arbitrary coframe and an (independent) local Lorentz connection $(e_i^\alpha, \Gamma_i{}^{\alpha\beta} = -\Gamma_i{}^{\beta\alpha})$, whereas the electromagnetic field as usual is represented by the vector potential $A_i = (-\Phi, \boldsymbol{A})$. The corresponding field strengths are the torsion $T_{ij}{}^\alpha = \partial_i e_j^\alpha - \partial_j e_i^\alpha + \Gamma_{i\beta}{}^\alpha e_j^\beta - \Gamma_{j\beta}{}^\alpha e_i^\beta$, the curvature $R_{ij}{}^{\alpha\beta} = \partial_i\Gamma_j{}^{\alpha\beta} - \partial_j\Gamma_i{}^{\alpha\beta} + \Gamma_{i\gamma}{}^\beta\Gamma_j{}^{\alpha\gamma} - \Gamma_{j\gamma}{}^\beta\Gamma_i{}^{\alpha\gamma}$, and the Maxwell tensor $F_{ij} = \partial_i A_j - \partial_j A_i$. Accordingly, the spacetime geometry in general carries the Riemann-Cartan structure with the metric $g_{ij} = e_i^\alpha e_j^\beta g_{\alpha\beta}$ and the nontrivial torsion. The "deviation" of the spacetime geometry from the Riemannian structure can be conveniently measured by the contortion tensor which is defined in terms of the difference of the local Lorentz connection $\Gamma_i{}^{\alpha\beta}$ and the Riemannian (Levi-Civita or Christoffel) connection $\tilde{\Gamma}_i{}^{\alpha\beta}$:

$$K_i{}^{\alpha\beta} = \tilde{\Gamma}_i{}^{\alpha\beta} - \Gamma_i{}^{\alpha\beta}. \tag{23}$$

### 2.1. *Quantum spinning particle*

A fermion particle with spin $\frac{1}{2}$, charge $q$ and mass $m$ is described by the relativistic Dirac theory. The spinor field $\psi$ dynamics is determined by the Lagrangian

$$L = \frac{i\hbar}{2}\left(\overline{\psi}\gamma^{\alpha}D_{\alpha}\psi - D_{\alpha}\overline{\psi}\gamma^{\alpha}\psi\right) - mc\overline{\psi}\psi + \frac{1}{2c}M_{\alpha\beta}\,\overline{\psi}\sigma^{\alpha\beta}\psi + \frac{i\hbar\nu'}{12}\check{T}^{\alpha}\,\overline{\psi}\gamma_{\alpha}\gamma_{5}\psi, \quad (24)$$

where the spinor covariant derivative is defined as (with $\sigma_{\alpha\beta} = i\gamma_{[\alpha}\gamma_{\beta]}$)

$$D_{\alpha}\psi = e^{i}_{\alpha}\left(\partial_{i}\psi - \frac{iq}{\hbar}\,A_{i}\psi + \frac{i}{4}\Gamma_{i}{}^{\beta\gamma}\sigma_{\beta\gamma}\psi\right). \quad (25)$$

The first two terms in (24) describe minimal coupling of the spinor field to electromagnetism and gravity encoded in (25). In addition, we assume possible nonminimal interactions, which are described by two Pauli-type terms in (24), where

$$M_{\alpha\beta} = \mu' F_{\alpha\beta} + c\delta'\,\frac{1}{2}\,\eta_{\alpha\beta\mu\nu}F^{\mu\nu}, \quad (26)$$

is the generalized polarization tensor, and the axial torsion is defined as

$$\check{T}^{\alpha} = -\,\frac{1}{2}\,\eta^{\alpha\mu\nu\beta}T_{\mu\nu\beta} = \left\{\check{T}^{0}, \check{\boldsymbol{T}}\right\}. \quad (27)$$

The parameters $\delta', \mu', \nu'$ characterize the strength of nonminimal couplings.

### 2.2. *Classical spinning particle*

The classical theory of spin was developed soon after the concept of spin was proposed in particle physics (see Ref. 1 for introduction and history). This theory underlies the analysis of the dynamics of polarized particles in accelerators and storage rings.

Neglecting second-order spin effects[2], the dynamical equations for a spinning particle in external electromagnetic and gravitational fields are written as

$$\frac{DU^{\alpha}}{d\tau} = -\,\frac{q}{m}\,F^{\alpha}{}_{\beta}\,U^{\beta}, \quad (28)$$

$$\frac{DS^{\alpha}}{d\tau} = -\,(\nu'-2)\,U^{i}K_{i\beta}{}^{\alpha}S^{\beta} - \frac{q}{m}\,F^{\alpha}{}_{\beta}S^{\beta} - \frac{2}{\hbar}\left[M^{\alpha}{}_{\beta} + \frac{1}{c^{2}}U^{\gamma}\left(U^{\alpha}M_{\beta\gamma} - U_{\beta}M^{\alpha}{}_{\gamma}\right)\right]S^{\beta}. \quad (29)$$

Here we follow the notations and conventions introduced in the previous sections.

### 2.3. *Application of the projection method*

Comparing the system (28)–(29) with (3), we find explicitly the generalized force and the spin transport matrix:

$$\mathcal{F}_{\alpha} = -\,\frac{q}{m}\,\mathbb{F}_{\alpha\beta}\,U^{\beta}, \qquad \Phi_{\alpha\beta} = -\,\frac{q}{m}\,\mathbb{F}_{\alpha\beta}, \quad (30)$$

where we introduced the combined external field

$$\mathbb{F}_{\alpha\beta} = F_{\alpha\beta} + \frac{2m}{q\hbar}\,M^{\perp}_{\alpha\beta} - \frac{m}{q}\,U^{i}\Gamma_{i\alpha\beta}. \quad (31)$$

As we see, $\Phi_{\alpha\beta}U^\beta = \mathcal{F}_\alpha$. Here we denoted

$$M^\perp_{\alpha\beta} = \Big(\delta^\mu_\alpha - \frac{1}{c^2}U_\alpha U^\mu\Big)\Big(\delta^\nu_\beta - \frac{1}{c^2}U_\beta U^\nu\Big)M_{\mu\nu}. \tag{32}$$

The components of this tensor read explicitly

$$M^\perp_{0a} = \frac{\gamma^2}{c}\left\{-v^2\mathcal{P}_a + \widehat{v}_a(\widehat{\boldsymbol{v}}\cdot\boldsymbol{\mathcal{P}}) + c[\widehat{\boldsymbol{v}}\times\boldsymbol{\mathcal{M}}]_a\right\}, \tag{33}$$

$$M^\perp_{ab} = \epsilon_{abc}\frac{\gamma^2}{c^2}\left\{c^2\mathcal{M}^c - \widehat{v}^c(\widehat{\boldsymbol{v}}\cdot\boldsymbol{\mathcal{M}}) + c[\widehat{\boldsymbol{v}}\times\boldsymbol{\mathcal{P}}]^c\right\}. \tag{34}$$

Here we identify $c\boldsymbol{\mathcal{P}}_a = \{M_{\widehat{01}}, M_{\widehat{02}}, M_{\widehat{03}}\}, \boldsymbol{\mathcal{M}}^a = \{M_{\widehat{23}}, M_{\widehat{31}}, M_{\widehat{12}}\}$ as the polarization and magnetization 3-vectors, respectively.

Now we use the projection method developed above and extract the "electric" and "magnetic" vectors from the artificial electromagnetic field (31)

$$\mathbb{E}_\alpha := \mathbb{F}_{\alpha\beta}U^\beta, \qquad \mathbb{B}^\alpha := \frac{1}{2c}\eta^{\alpha\mu\nu\beta}\mathbb{F}_{\mu\nu}U_\beta, \tag{35}$$

which yields an unique representation

$$\mathbb{F}_{\alpha\beta} = \frac{1}{c^2}\left(\mathbb{E}_\alpha U_\beta - \mathbb{E}_\beta U_\alpha + c\eta_{\alpha\beta\mu\nu}U^\mu\mathbb{B}^\nu\right). \tag{36}$$

Comparing (6) with (9), we thus identify

$$N_\alpha = -\frac{q}{m}\,\mathbb{E}_\alpha, \qquad Q^\alpha = -\frac{q}{m}\,\mathbb{B}^\alpha. \tag{37}$$

Let us define effective "electric" and "magnetic" fields

$$\mathfrak{E}^{\rm eff}_a = \{\mathbb{F}_{\widehat{10}}, \mathbb{F}_{\widehat{20}}, \mathbb{F}_{\widehat{30}}\}, \qquad \mathfrak{B}^a_{\rm eff} = \{\mathbb{F}_{\widehat{23}}, \mathbb{F}_{\widehat{31}}, \mathbb{F}_{\widehat{12}}\}. \tag{38}$$

Or in compact form: $\mathfrak{E}^{\rm eff}_a = \mathbb{F}_{a0}$, and $\mathfrak{B}^a_{\rm eff} = \frac{1}{2}\epsilon^{abc}\mathbb{F}_{bc}$.

Consequently from (35) and (37) we derive the components

$$N^0 = \gamma\,\frac{q\widehat{\boldsymbol{v}}\cdot\mathfrak{E}_{\rm eff}}{mc^2}, \qquad \boldsymbol{N} = \gamma\,\frac{q}{m}\left(\mathfrak{E}_{\rm eff} + \widehat{\boldsymbol{v}}\times\mathfrak{B}_{\rm eff}\right), \tag{39}$$

$$Q^0 = -\gamma\,\frac{q\widehat{\boldsymbol{v}}\cdot\mathfrak{B}_{\rm eff}}{mc^2}, \qquad \boldsymbol{Q} = -\gamma\,\frac{q}{m}\Big(\mathfrak{B}_{\rm eff} - \frac{1}{c^2}\widehat{\boldsymbol{v}}\times\mathfrak{E}_{\rm eff}\Big). \tag{40}$$

By construction, $\boldsymbol{Q}_\perp = \boldsymbol{Q}$, and moreover we have $\boldsymbol{\mathcal{F}} = \boldsymbol{N}$. Inserting (39) and (40) into (22), we obtain the precession angular velocity (22)

$$\boldsymbol{\Omega} = \frac{q}{m}\left(-\mathfrak{B}_{\rm eff} + \frac{\gamma}{\gamma+1}\,\frac{\widehat{\boldsymbol{v}}\times\mathfrak{E}_{\rm eff}}{c^2}\right) \tag{41}$$

as a function of external fields which enter via the effective variables:

$$\mathfrak{E}_{\rm eff} = \mathfrak{E} - \frac{2m}{q\hbar}\,\gamma^2\,\widehat{\boldsymbol{v}}\times\boldsymbol{\Delta} + \frac{m}{q}\,\boldsymbol{\mathcal{E}}, \tag{42}$$

$$\mathfrak{B}_{\rm eff} = \mathfrak{B} + \frac{2m}{q\hbar}\,\gamma^2\Big[\boldsymbol{\Delta} - \frac{1}{c^2}\,\widehat{\boldsymbol{v}}\,(\widehat{\boldsymbol{v}}\cdot\boldsymbol{\Delta})\Big] + \frac{m}{q}\,\boldsymbol{\mathcal{B}}. \tag{43}$$

Recall that the components of the true electric and magnetic fields are introduced as $\mathfrak{E}_a = \{F_{\widehat{1}\widehat{0}}, F_{\widehat{2}\widehat{0}}, F_{\widehat{3}\widehat{0}}\}, \mathfrak{B}^a = \{F_{\widehat{2}\widehat{3}}, F_{\widehat{3}\widehat{1}}, F_{\widehat{1}\widehat{2}}\}$. The generalized polarization current,

$$\boldsymbol{\Delta} = \boldsymbol{\mathcal{M}} + \frac{1}{c}\,\widehat{\boldsymbol{v}} \times \boldsymbol{\mathcal{P}}, \tag{44}$$

accounts in the spin precession (41)–(43) for the electromagnetic nonminimal coupling effects, whereas the gravitoelectric and gravitomagnetic fields

$$\boldsymbol{\mathcal{E}} = \widetilde{\boldsymbol{\mathcal{E}}} - \frac{\gamma c(3-\nu')}{6}\,\widehat{\boldsymbol{v}} \times \check{\boldsymbol{T}}, \qquad \boldsymbol{\mathcal{B}} = \widetilde{\boldsymbol{\mathcal{B}}} + \frac{\gamma c(3-\nu')}{6}\,\big(\check{\boldsymbol{T}} - \widehat{\boldsymbol{v}}\check{T}^0\big) \tag{45}$$

encompass general-relativistic Riemannian gravitoelectric $\widetilde{\boldsymbol{\mathcal{E}}}$ and gravitomagnetic $\widetilde{\boldsymbol{\mathcal{B}}}$ contributions, as well as the post-Riemannian terms due to the spacetime torsion.

## 3. Discussion and conclusion

The classical and quantum spin dynamics are fully consistent. The physical contents of the relativistic quantum theory (24) is revealed when we recast the Dirac equation into a Schrödinger form and go to the Foldy-Wouthuysen (FW) representation. In the Schwinger gauge, the coframe is parametrized by the functions $V, \boldsymbol{K}, W^{\widehat{a}}{}_b$:

$$e_i^{\widehat{0}} = V\,\delta_i^0, \qquad e_i^{\widehat{a}} = W^{\widehat{a}}{}_b\,\big(\delta_i^b - cK^b\,\delta_i^0\big), \qquad a = 1,2,3, \tag{46}$$

and the resulting FW Hamiltonian in the semiclassical approximation then reads

$$\mathcal{H}_{FW} = \beta mc^2 V\gamma + q\Phi + \frac{c}{2}\,(\boldsymbol{K}\cdot\boldsymbol{\pi} + \boldsymbol{\pi}\cdot\boldsymbol{K}) + \frac{\hbar}{2}\boldsymbol{\Sigma}\cdot\boldsymbol{\Omega}. \tag{47}$$

Here the Lorentz factor (5) operator $\gamma$ and the precession velocity (41) operator $\boldsymbol{\Omega}$ are both expressed in terms of the velocity operator $\widehat{\boldsymbol{v}}$ which is related to the momentum operator $\boldsymbol{\pi} = -i\hbar\boldsymbol{\nabla} - q\boldsymbol{A}$ via $\beta\,W^b{}_{\widehat{a}}\pi_b = m\gamma\widehat{v}_a$.

The analysis of the spin dynamics in electromagnetic, inertial and gravitational fields is fundamentally important for the study of the geometrical structure of spacetime[3], as well as in the high-energy physics experiments[4], in the search of gravitational waves[5], in the neutrino physics in matter[6], and in the heavy-ion collisions[7].

## Acknowledgement

This work was partially supported by the Russian Foundation for Basic Research (Grant No. 18-02-40056-mega).

## References

1. H.C. Corben, *Classical and quantum theories of spinning particles* (Holden-Day, Inc: San Francisco, 1968).
2. C. Chicone, B. Mashhoon, B. Punsly, *Phys. Lett. A* **343**, 1 (2005).
3. Yu.N. Obukhov, A.J. Silenko, O.V. Teryaev, *Phys. Rev. D* **90**, 124068 (2014).
4. Yu.N. Obukhov, A.J. Silenko, O.V. Teryaev, *Phys. Rev. D* **94**, 044019 (2016).
5. Yu.N. Obukhov, A.J. Silenko, O.V. Teryaev, *Phys. Rev. D* **96**, 105005 (2017).
6. M. Dvornikov, *Phys. Rev. D* **99**, 035027 (2019).
7. G.Y. Prokhorov, O.V. Teryaev, V.I. Zakharov, *Phys. Rev. D* **98**, 071901(R) (2018).

# Frequency spectrum of an optical resonator in a curved spacetime

Dennis Rätzel

*Institut für Physik, Humboldt-Universität zu Berlin, Newtonstraße 15, 12489 Berlin, Germany*

Fabienne Schneiter and Daniel Braun

*Institut für Theoretische Physik, Eberhard-Karls-Universität Tübingen, 72076 Tübingen, Germany*

Tupac Bravo

*Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria*

Richard Howl

*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK*

Maximilian P.E. Lock

*Department of Physics, Imperial College, London SW7 2AZ, United Kingdom*
*Institute for Quantum Optics and Quantum Information (IQOQI), Boltzmanngasse 3, 1090 Vienna, Austria*

Ivette Fuentes

*School of Mathematical Sciences, University of Nottingham, University Park, Nottingham NG7 2RD, UK*

There is an ever growing number of proposals for high precision experiments to measure gravitational effects, from simple Newtonian gravity to gravitational waves and even precision tests of general relativity (GR). In particular, more and more researchers from the fields of quantum optics and quantum opto-mechanics are becoming interested in GR and propose metrological experiments. Usually, such proposals rely heavily on a notion of length. However, in GR, as coordinates have no physical meaning, there is no unique concept for the length of a matter system. In this proceedings article, we summarize the article [24], where the conceptual problem of length is addressed for a subset of experimental proposals. In particular, the effect of gravitational fields and acceleration on the frequency spectrum of an optical resonator is discussed in the framework of GR. The optical resonator is modeled as a deformable rod of matter connecting two mirrors. Explicit expressions for the frequency spectrum are given for the case of a small perturbation. Example situations are discussed and a connection is obtained to a relativistic concept of rigidity.

*Keywords*: Optical cavity, Fabry-Pérot, gravitational field, rigid body, elastic body.

## 1. Introduction

In general relativity (GR), as coordinates have no physical meaning, there is no unique concept for the length of a matter system. Some notion of length can be covariantly defined using geometrical quantities or properties of matter. The ambiguity in the notion of length poses a problem for high accuracy metrological

experiments, where gravitational fields or acceleration play a significant role. For example, the frequency spectrum of a resonator depends on its dimensions and hence knowledge of the precise values of these dimensions is of utmost importance. Cases in which the effects of gravitational fields and acceleration must be considered include those in which the gravitational field is to be measured, such as in proposals for the measurement of gravitational waves with electromagnetic cavity resonators[6,9,10,20–22,27] or other extended matter systems[1,7,8,12,18,25,26], tests of GR[3,11] or the expansion of the universe[13,14]. Other situations are those in which the metrological system is significantly accelerated[16,17,23]. A fundamental limit for the precision of a light cavity resonator as a metrological system can even be imposed by the gravitational field of the light inside the cavity[4].

The two most important concepts of length are the proper distance and the radar distance. The proper distance is a geometrical quantity usually associated with the length of a rod that is rigid in the sense of that given by Born[2]. The radar distance is the optical length that can be measured by sending light back and forth between two mirrors and taking the time between the two events as a measure of distance. It is this radar length that gives the resonance frequency spectrum of an optical resonator for large enough wave numbers. However, the resonators that are part of the metrological systems described in Refs. 1, 3, 4, 6–14, 16–18, 20–23, 25–27 are confined by solid matter systems, and therefore, the notion of proper length plays also a role.

We assume that all effects on the optical resonator can be described as small perturbations. We use a specific coordinate system $x^{\mathcal{M}}$ valid in a region around the world line of the resonator's support in which the spacetime metric takes the form $g_{\mathcal{MN}} = \eta_{\mathcal{MN}} + h_{\mathcal{MN}}$, where $\eta_{\mathcal{MN}} = \mathrm{diag}(-1, 1, 1, 1)$ is the Minkowski metric and $h_{\mathcal{MN}}$ is a perturbation. $h_{\mathcal{MN}}$ is considered as small in the sense that $|h_{\mathcal{MN}}| \ll 1$ for all $\mathcal{M}, \mathcal{N}$.

## 2. Frequency spectrum

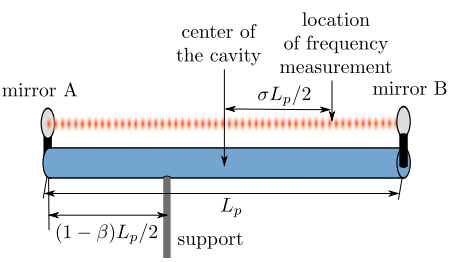



Fig. 1. Illustration of our model of an optical resonator consisting of two mirrors that are attached to the ends of a rod. We assume that the resonator is moved along a trajectory $\gamma(\varrho)$ by a support which is attached at a distance $(1-\beta)L_p/2$ from mirror A, where $L_p$ is the length of the rod without gravity or acceleration. Since proper time depends on the position in the gravitational field so does the measured frequency of a resonator mode. We assume the frequency to be measured at a distance $\sigma L_p/2$ from the center of the resonator towards mirror B.

In Ref. 24, we consider two fundamental models of an optical resonator. In both cases, there is a rod which is supported at a certain point along its length. In one model, the rod itself is the optical resonator. For example, it could consist of a homogeneous isotropic dielectric. The other model assumes two mirrors attached to the ends of the rod. The latter model is the focus of this article, and it is illustrated in Fig. 1. For both cases, we derived an expression for the resonance frequencies of the optical resonator in Ref. 24. As we are dealing with an extended object in GR, the obtained resonance frequencies are ambiguous as we cannot give a proper time to the whole resonator; the frequencies of the modes must depend on the point in the resonator where they are observed. Using the short wavelength expansion, we find that the frequencies of the modes of the resonator measured by an observer inside the cavity is given as

$$\omega_{\gamma,n} = \frac{cn\pi}{R_\gamma}\,, \tag{1}$$

where $R_\gamma$ is the radar distance which is defined as $c/2$ times the proper time for a round trip of a light pulse in the resonator as it would be measured by an observer at a fixed proper distance from the support of the resonator. It is clear that this is an approximate value; the notion of frequency means the rate of repetition of a signal. For this notion to make sense, it has to be constant at least for a few repetition cycles. Hence, the gravitational field has to change slowly in comparison to the round trip time $2R_\gamma/c$.

## 3. Rigid resonators

In GR, there exist different notions of rigidity as it turns out to be less than straightforward to formulate this basic concept of Newtonian mechanics in a relativistic way. Here, we use, as our starting point, a definition of a rigid rod that is Born rigid, and we undertake a perturbative analysis for small length scales, small accelerations, small velocities and small gravitational fields. We show that two types of effects are found; those due to spacetime properties alone and those due to small deformations of the rod which correspond to small deviations from Born rigidity. Since all effects can be considered as small, we remain in the linear regime, where the different effects are assumed to be independent.

Let us assume that we have a rod of very small diameter in comparison to its length, which means it is effectively 1-dimensional. We assume that the world lines of the segments of the rod form a family of curves $\gamma_\varsigma(\varrho)$ parametrized by $\varsigma$ which we assume to be in the interval $\varsigma \in [a, b]$. The end points of the rod are $\gamma_a(\varrho)$ and $\gamma_b(\varrho)$. The spacetime surface $F(\varrho, \varsigma) = \gamma_\varsigma(\varrho)$ can be called the world sheet of the rod. We assume that the curve parameter $\varrho$ is chosen so that the curves $s_\varrho(\varsigma) := F(\varrho, \varsigma)$ are space-like geodesics. We say that the rod is rigid if the proper distance between every two points on the curve $s_\varrho(\varsigma)$ is independent of the parameter $\varrho$.

In Ref. 24, we derive the resonance frequencies of a Born rigid resonator in terms of its constant proper length. For this purpose, we choose to work in a particular

coordinate system, the proper detector frame which is defined along the world line of the support of the optical cavity. In these coordinates, the spacetime metric seen by a non-rotating observer can be given simply in terms of the Riemann curvature tensor along the world line of the resonator's support $\gamma(\tau)$ denoted as $R_{\mathcal{MNKL}}(\tau)$ (caligraphic captical indeces run from 0 to 3) and the non-gravitational or proper acceleration with respect to a local freely falling frame, represented by the spatial vector $a^J$ $(J \in 1, 2, 3)$.

The proper detector frame of a non-rotating observer is accurate for proper distances much smaller than $l_{\text{var}}$, the length scale on which the properties of spacetime vary. Therefore, we assume that the proper length of the resonator in the absence of gravity and acceleration $L_p$ is small in comparison to the scale $l_{\text{var}}$. We consider $\gamma(\tau)$ to be the world line of the point at which the rod of the resonator is supported. We assume that this point is somewhere inside the resonator. We also assume that the resonator is not rotating in the frame of the observer. We orient the spatial geodesic representing the rigid rod along the $z$-direction at $\gamma(\tau)$ in the proper detector frame, i.e. $s'_\tau(\varsigma) = (0, 0, 0, 1)$. By construction of the proper detector frame, the geodesics $s_\tau(\varsigma)$ run along the $z$-coordinate. We consider only first order contributions of the proper acceleration, which means that we can consider the metric in the proper detector frame as a linearly perturbed flat spacetime metric.

To obtain the frequency of the rigid resonator measured by an observer at coordinates $(0, 0, z_0)$ in the proper detector frame based on Eq. (1), we have to calculate the corresponding radar distance between the mirrors. Then, the relative change of the resonance frequencies measured at $z_0 = (\sigma + \beta)L_p/2$ is given as

$$\delta_{\omega,\sigma} := \frac{\omega_n}{\bar{\omega}_n} - 1 \approx -\frac{a^z(\tau)}{2c^2}\sigma L_p - \frac{R_{0z0z}(\tau)}{24}\left(3\sigma^2 + 6\sigma\beta - 1\right)L_p^2\,, \tag{2}$$

where $\bar{\omega}_n = cn\pi/L_p$ is the $n$-th resonance frequency of the resonator for vanishing acceleration and curvature. We find that the only linear contribution of the acceleration $a^z$ to the resonance frequency spectrum in Eq. (2) is via a position-dependent red shift. It vanishes for $\sigma = 0$, which corresponds to a frequency measurement in the center of the resonator. The term $3\sigma^2$ corresponds to a pure red shift with respect to the center of the cavity. The term $6\beta\sigma$ is due to the displacement of the resonator's support from its center.

## 4. Deformable resonators

A realistic rod can never be rigid. Therefore, we consider the first order deviations from the rigid rod by taking the deformation of the rod due to small inertial and gravitational forces into account. In the proper detector frame, every segment of the rod has a world line with constant spatial components. The relative acceleration of a segment of the rod at $\mathbf{x} = (c\tau, 0, 0, z)$, in comparison to a freely falling test particle initially at rest at the same position as that segment, can be derived from the geodesic equation.

We consider the effect of relative acceleration on the resonator's end mirrors and the resulting deformation of the rod to be negligible in comparison to the direct effect of relative acceleration on the rod. Then, we obtain the inertial and tidal forces on the rod by multiplication with the mass density $\rho$. These forces give rise to stresses within the rod, represented by the stress tensor $\sigma_{KL}$. For static forces and forces that change very slowly, the stresses are related to the strain via Hooke's law. After an analysis of the resulting forces, we give conditions that have to be fulfilled for the longitudinal deformation of the rod to be the dominant effect on the frequency spectrum of the resonator. We find for the relative change of the resonance frequencies of the deformable resonator

$$\delta_{\omega,\sigma} \approx \frac{\mathbf{a}^z(\tau)}{2c^2}\left(\frac{c^2}{c_s^2}\beta - \sigma\right) L_p + \frac{R_{0z0z}(\tau)}{24}\left(2\frac{c^2}{c_s^2}(3\beta^2+1) - 3\sigma^2 - 6\sigma\beta + 1\right) L_p^2 \,. \tag{3}$$

We obtain the result in Eq. (2) for the Born rigid rod from Eq. (3) if the speed of sound $c_s$ in the material is infinite. This coincides with the observation that a Born rigid rod violates causality, as its segments would need to interact with an infinite speed. A more realistic definition of a rigid rod, which we call a causal rigid rod, was given in Ref. 19 as a rod in which the speed of sound is equivalent to the speed of light. In the appendix of Ref. 24, we show that the approach of Ref. 19 leads to Eq. (3) with the speed of sound replaced by the speed of light.

The speed of sound $c_s$ in every realistic material is always much smaller than the speed of light: for example the speed of sound in aluminum is of the order $5 \times 10^3$m/s. To date, the material with the highest ratio of Young's modulus and density $Y/\rho = c_s^2$ is carbyne, with a value of the order of $10^9 \mathrm{m}^2/\mathrm{s}^2$ [15], which would correspond to a speed of sound of the order of $3 \times 10^4$m/s. Therefore, we find that the effect of the deformation of matter is by far the most dominant and the rod is far from rigid (may it be Born rigid or causal rigid) in all realistic situations. However, the relativistic effect of gravitational redshift gives a fundamental limit on the definition of the frequency spectrum of an optical resonator as a property of the resonator alone; when resonance frequencies of an optical resonator are to be specified with a precision of the order of this relativistic effect, the position of the frequency measurement has to specified.

## 5. Example applications

To illustrate the applicability of our results, we consider some examples. A particularly straightforward example is the situation of a non-rotating resonator that is uniformly accelerated along the optical axis. We find

$$\delta_{\omega,\sigma} \approx \left(\frac{\beta}{c_s^2} - \frac{\sigma}{c^2}\right)\frac{a^z L_p}{2} \,. \tag{4}$$

For $\beta = \pm 1$, a length of the resonator of $L_p \sim 2\,\mathrm{cm}$, an acceleration of the order of $10\,\mathrm{ms}^{-2}$, which is similar to the gravitational acceleration of the Earth, and a

speed of sound in the rod of the order of $10^3\,\mathrm{ms}^{-1}$ (similar to the speed of sound in aluminum), we obtain a relative frequency shift of the order of $10^{-7}$. This frequency shift is given only by the first term in Eq. (4) as the second term is smaller by about 11 orders of magnitude. Since the first term is due to the deformation of the resonator it is a Newtonian effect. The second term is the gravitational red shift, which was already measured on the length scale of about 2 cm in the gravitational field of the Earth[5] and therefore for the acceleration in the example. The effect of gravitational red shift gives a limit on the validity of the concept of the frequency spectrum as a property of the optical resonator itself.

Our results are not limited to spacetimes that only bear weak gravitational effects. It is the spacetime metric seen by the optical resonator in its proper detector frame that has to be a linearized metric. This is ensured by the condition $l_{\mathrm{var}} \gg L_p$. To illustrate the applicability of our results to spacetimes with strong gravitational fields, we consider the situation of a non-rotating resonator that falls into a non-rotating black hole. For a vertically oriented causal rigid resonator supported at its center, we find the relative frequency shift at its center

$$\delta_{\omega,0}(\tau) \approx -\frac{r_S L_p^2}{8r(\varrho)^3}\,. \tag{5}$$

Note that there is no effect due to the crossing of the event horizon at $r_S$. This result can be applied as well to an optical resonator falling towards the Earth. For a distance from the center of the Earth of the same order as its radius, we find that the relative frequency shift in Eq. (5) is of the order of $10^{-27}$ for an optical resonator of 2 cm length. This relativistic effect is mostly gravitational red shift due to curvature. It is far from being observable with state of the art technology. However, it gives a fundamental limit of the validity of the concept of frequency spectrum as a property of the optical resonator without any reference as discussed above.

## Acknowledgments

D.R. thanks the Humboldt Foundation for funding his research with their Feodor-Lynen Felloship. R.H. and I.F. would like to acknowledge that this project was made possible through the support of the grant "Leaps in cosmology: gravitational wave detection with quantum systems" (No. 58745) from the John Templeton Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation. I.F. would like to acknowledge that this project was made possible through the support of the grant "Quantum Observers in a Relativistic World" from FQXi's Physics of the Observer program.

## References


1. Asimina Arvanitaki and Andrew A. Geraci. Detecting high-frequency gravitational waves with optically levitated sensors. *Phys. Rev. Lett.*, 110:071105, Feb 2013.
2. Max Born. Die Theorie des starren Elektrons in der Kinematik des Relativitätsprinzips. *Annalen der Physik*, 335(11):1–56, 1909.
3. V. B. Braginskii, C. M. Caves, and K. S. Thorne. Laboratory experiments to test relativistic gravity. *Phys. Rev. D*, 15:2047–2068, April 1977.
4. Daniel Braun, Fabienne Schneiter, and Uwe R Fischer. Intrinsic measurement errors for the speed of light in vacuum. *Classical and Quantum Gravity*, 34(17):175009, 2017.
5. Heiner Denker, Ludger Timmen, Christian Voigt, Stefan Weyers, Ekkehard Peik, Helen S. Margolis, Pacôme Delva, Peter Wolf, and Gérard Petit. Geodetic methods to determine the relativistic redshift at the level of $10^{-18}$ in the context of international timescales: a review and practical results. *Journal of Geodesy*, 92(5):487–516, May 2018.
6. G. Gemme, A. Chincarini, R. Parodi, P. Bernard, and E. Picasso. Parametric gravity wave detector. In *Electromagnetic probes of fundamental physics. Proceedings, Workshop, Erice, Italy, October 16-21, 2001*, pages 75–83, 2001.
7. Maxim Goryachev and Michael E. Tobar. Gravitational Wave Detection with High Frequency Phonon Trapping Acoustic Cavities. *Phys. Rev.*, D90(10):102005, 2014.
8. Peter W. Graham, Jason M. Hogan, Mark A. Kasevich, and Surjeet Rajendran. A New Method for Gravitational Wave Detection with Atomic Sensors. *Phys. Rev. Lett.*, 110:171102, 2013.
9. L.P. Grishchuk and M.V. Sazhin. Excitation and detection of standing gravitational waves. *JETP*, 41(5):787, 1975.
10. LP Grishchuk and MV Sazhin. Quantum electromagnetic oscillator in the field of a gravitational wave and the problem of nondemolition measurements. *JETP*, 53(4):639, 1981.
11. Richard Howl, Lucia Hackermuller, David Edward Bruschi, and Ivette Fuentes. Gravity in the Quantum Lab. *Advances in Physics: X*, 3(1):1383184, 2018.
12. L. Ju, D. G. Blair, and C. Zhao. Detection of gravitational waves. *Reports on Progress in Physics*, 63:1317–1427, September 2000.
13. S. M. Kopeikin. Einstein's equivalence principle in cosmology. *COSPAR Meeting*, 40, 2014.
14. S. M. Kopeikin. Optical cavity resonator in an expanding universe. *Gen. Rel. Grav.*, 47(2):5, 2015.
15. Mingjie Liu, Vasilii I Artyukhov, Hoonkyung Lee, Fangbo Xu, and Boris I Yakobson. Carbyne from first principles: chain of c atoms, a nanorod or a nanorope. *ACS nano*, 7(11):10075–10082, 2013.

16. Maximilian P. E. Lock and Ivette Fuentes. Relativistic quantum clocks. 2016.
17. Maximilian P. E. Lock and Ivette Fuentes. Dynamical Casimir effect in curved spacetime. *New J. Phys.*, 19(7):073005, 2017.
18. Michele Maggiore. *Gravitational Waves: Volume 1: Theory and Experiments*, volume 1. Oxford university press, 2008.
19. Jose Natario. Relativistic elasticity of rigid rods and strings. *Gen. Rel. Grav.*, 46(11):1816, 2014.
20. F Pegoraro, E Picasso, and L A Radicati. On the operation of a tunable electromagnetic detector for gravitational waves. *Journal of Physics A: Mathematical and General*, 11(10):1949, 1978.
21. F. Pegoraro, L.A. Radicati, Ph. Bernard, and E. Picasso. Electromagnetic detector for gravitational waves. *Physics Letters A*, 68(2):165 – 168, 1978.
22. C. E. Reece, P. J. Reiner, and A. C. Melissinos. A detector for high frequency gravitational effects based on parametric conversion at 10-GHz. *eConf*, C8206282:394–402, 1982.
23. Bartosz Regula, Antony R. Lee, Andrzej Dragan, and Ivette Fuentes. Generating entanglement between two-dimensional cavities in uniform acceleration. *Phys. Rev.*, D93(2):025034, 2016.
24. Dennis Rtzel, Fabienne Schneiter, Daniel Braun, Tupac Bravo, Richard Howl, Maximilian P E Lock, and Ivette Fuentes. Frequency spectrum of an optical resonator in a curved spacetime. *New Journal of Physics*, 20(5):053046, 2018.
25. Carlos Sabin, David Edward Bruschi, Mehdi Ahmadi, and Ivette Fuentes. Phonon creation by gravitational waves. *New J. Phys.*, 16:085003, 2014.
26. S. Singh, L. A. De Lorenzo, I. Pikovski, and K. C. Schwab. Detecting continuous gravitational waves with superfluid $^4$He. *New J. Phys.*, 19(7):073023, 2017.
27. Sergey P. Tarabrin. Interaction of plane gravitational waves with a Fabry-Perot cavity in the local Lorentz frame. *Phys. Rev.*, D75:102002, 2007.

# Effects of magnetic-like part of gravitational waves on spinning particles

Matteo Luca Ruggiero*

*IIS Russell-Moro-Guarini, Corso Molise 58, Torino*
*Politecnico di Torino, Corso Duca degli Abruzzi 24, Torino, Italy and INFN, Sezione di Pisa, Italy* **E-mail: matteo.ruggiero@polito.it*

Antonello Ortolan

*INFN - National Laboratories of Legnaro, Viale dell'Università, I-35020 Legnaro (PD), Italy*

Gianni Carugno

*Department of Physics and Astronomy, University of Padova and INFN, Via F. Marzolo 8, I-35131 Padova, Italy*

We discuss the interaction of gravitational waves with spinning particle. To obtain physically measurable quantities, we make use of Fermi coordinates and we show that, using these coordinates, the magnetic-like (or gravitomagnetic) part of the gravitational field of the wave is emphasized. Eventually, we evaluate the magnitude of the effects induced by the waves on spinning particles, and discuss some measurement possibilities.

*Keywords*: Gravitational waves; spin; gravitomagnetism.

## 1. Introduction

The observation of gravitational wave (GW) signals[1,2] is just one of the latest successes of Einstein's theory of gravitation, General Relativity (GR): as a matter of fact, its prediction were verified with great accuracy during last century, even though challenges to the Einsteinian paradigm come from cosmological observations[3]. Einstein's theory is based on the principle of general covariance, which requires physics laws to be expressed by tensorial equations in space-time; accordingly, physical measurements are meaningful only when the observer and the object of the observations are unambiguously identified[4]. Roughly speaking, the measurement process can be summarised as follows: (i) observers possess their own space-time, in the vicinity of their world-lines; (ii) covariant physics laws are then projected onto local space and time; (iii) predictions for the outcome of measurements in the local space-time of the observers are obtained. In practice, it is convenient for an observer to use a quasi-Cartesian coordinates system in his neighborhood to describe the effects of gravitation; a continuous set of quasi-Cartesian coordinates associated with the observer's world-line defines the so called Fermi coordinates. These coordinates have a concrete meaning, since they are the coordinates an observer would naturally use to make space and time measurements in the vicinity of his world-line. The definition of these coordinates is relevant to define measurements performed in the space-time of a gravitational wave, even though in current literature the interaction between gravitational waves and detectors is described in terms of a transverse and traceless tensor, which allows to introduce the so-called TT coordinates (see e.g. Ref. 5 and references therein for a thorough discussion on the various coordinates

used to describe the interaction with gravitational waves). Motion of spinning particles in the weak gravitational plane wave space-time has been investigated in the past [6]; indeed, gravitational waves carry angular momentum and, hence, spinning particles interact with gravitational waves differently from non spinning ones. In particular, writing the gravitational waves space-time using Fermi coordinate emphasizes these features, since gravitomagnetic terms are present in the metric (see e.g. Ref. 7): the term *gravitomagnetic* refers to the magnetic-like part of the gravitational field, which originates from the angular momentum of the sources. It is a well known fact (see e.g. Ref. 8) that Einstein equations, in weak-field approximation (small masses, low velocities), can be written in analogy with Maxwell equations for the electromagnetic field, where the mass density and current play the role of the charge density and current, respectively. Once that a set of coordinates adapted to the observer world-line is chosen, gravitomagnetic effects are determined by the off-diagonal terms in the space-time metric. Actually, these effects are very small, but there were many proposals in the past (see Ref. 8) and also more recently to test them (see e.g. Refs. 9, 10, 11, 12, 13, 14). Starting from the results described in Ref. 7, it is possible to test the effects of the gravitomagnetic field of a gravitational wave on a bunch of spinning particles; in particular, we are going to show that the response of a such a system to the interaction with the wave could lead to macroscopic measurable quantities.

## 2. Spinning Particles in the Gravitational Wave Space-Time

Using the transverse - traceless gauge (TT) the space-time metric of a gravitational wave propagating along the $x$ axis is given by the following line element:

$$ds^2 = -dt^2 + dx^2 + (1 - h_+)dy^2 + (1 + h_+)dz^2 - 2h_\times dydz\,, \tag{1}$$

where[a] the functions $h_{+,\times}$ depend on $t - x$. In particular, if we consider monochromatic waves with frequency $\omega$, we have

$$h_+(t-x) = A_+ \sin\omega\,(t-x)\ , \quad h_\times(t-x) = A_\times \cos\omega\,(t-x) \tag{2}$$

In what follows, analysis is carried out up to linear order in $h_+$, $h_\times$. It is easy to show [5] that in the TT metric (1) the geodesics are such that an inertial mass initially at rest remains at rest indefinitely; however the $x, y, z, t$ coordinates have no operational meaning, since they are not directly related to physical quantities: in fact, two masses at rest in the TT with (constant) coordinate distance $\Delta y$, have a physical distance varying with time $\ell = \Delta y \left(1 + \frac{h_+}{2}\right)$. In order to give a physical meaning to coordinates, we need to use a set of Fermi coordinates in the space-time of the gravitational wave. The transformation from TT coordinates $x, y, z, t$ to Fermi $X, Y, Z, T$ coordinates is given by [7]:

[a] We use units such that $c = 1$, however physical units are used to evaluate the order of magnitude of the effects; the space-time signature is (-1,1,1,1).

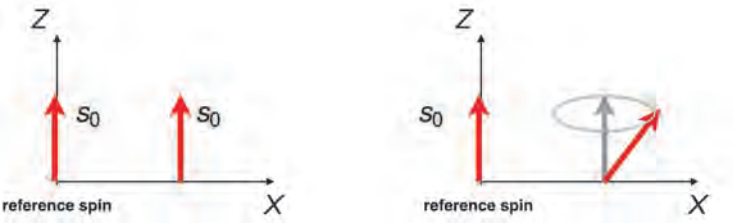



Fig. 1. Left: the reference spin and the other one before the passage of the wave; right: the two spin after the passage of the wave.

$$
\begin{aligned}
t &= T - \frac{1}{2}A_+\left\{\frac{Y^2-Z^2}{X}\left[\frac{\cos(\omega(T-X))-\cos(\omega T)}{\omega X}-\sin(\omega T)\right]\right.\\
&\quad \left.+\,\omega(S_2^0 Z + S_3^0 Y)[\cos(\omega T)-1]\right\}\\
&\quad +A_\times\left\{\frac{YZ}{X}\left[\frac{\sin(\omega(T-X))-\sin(\omega T)}{\omega X}+\cos(\omega T)\right]-\frac{1}{2}\omega(S_2^0 Y - S_3^0 Z)\sin(\omega T)\right\}\\
x &= X - \frac{1}{2}A_+\frac{Y^2-Z^2}{X}\left[\frac{\cos(\omega(T-X))-\cos(\omega T)}{\omega X}-\sin(\omega T)\right]\\
&\quad +A_\times\frac{YZ}{X}\left[\frac{\sin(\omega(T-X))-\sin(\omega T)}{\omega X}+\cos(\omega T)\right],\\
y &= Y + A_+\left\{Y\left[\frac{\cos(\omega(T-X))-\cos(\omega T)}{\omega X}-\frac{1}{2}\sin(\omega T)\right]-\frac{1}{2}S_3^0[\sin(\omega T)-\omega T]\right\}\\
&\quad -A_\times\left\{Z\left[\frac{\sin(\omega(T-X))-\sin(\omega T)}{\omega X}+\frac{1}{2}\cos(\omega T)\right]-\frac{1}{2}S_2^0[\cos(\omega T)-1]\right\},\\
z &= Z - A_+\left\{Z\left[\frac{\cos(\omega(T-X))-\cos(\omega T)}{\omega X}-\frac{1}{2}\sin(\omega T)\right]+\frac{1}{2}S_2^0[\sin(\omega T)-\omega T]\right\}\\
&\quad -A_\times\left\{Y\left[\frac{\sin(\omega(T-X))-\sin(\omega T)}{\omega X}+\frac{1}{2}\cos(\omega T)\right]+\frac{1}{2}S_3^0[\cos(\omega T)-1]\right\}.
\end{aligned}
\tag{3}
$$

The above expressions are obtained in the vicinity of the world-line of the spinning particle. It is based on the study of the motion of the particle and the evolution of its spin vector components in the gravitational field of the wave. It has been supposed that, before the passage of the wave, the particle is at rest with associated spin vector aligned along a given direction with constant magnitude; $S_i^0$ $(i = 1, 2, 3)$ are the initial values of the coordinate components of the particle's spin. In order to compare the spin of two different spinning particles, the deviation equation has been solved together with a suitable transport equation for the spin vector between two neighbouring world-lines. The space-time metric in Fermi coordinates, up to

the second order, is then given by

$$
\begin{aligned}
g_{TT} &= -1 - \left[S_2^0 Z + S_3^0 Y - \frac{1}{2}(Y^2 - Z^2)\right]\omega^2 A_+ \sin(\omega T) \\
&\quad + \left[S_2^0 Y - S_3^0 Z + YZ\right]\omega^2 A_\times \cos(\omega T) + O(3)\,, \\
g_{TX} &= -\frac{1}{3}\omega^2\left[(Y^2 - Z^2)A_+ \sin(\omega T) + 2YZA_\times \cos(\omega T)\right] + O(3)\,, \\
g_{TY} &= \frac{1}{3}\omega^2 X\left[YA_+ \sin(\omega T) + ZA_\times \cos(\omega T)\right] + O(3)\,, \\
g_{TZ} &= \frac{1}{3}\omega^2 X\left[-ZA_+ \sin(\omega T) + YA_\times \cos(\omega T)\right] + O(3)\,, \\
g_{XX} &= 1 + \frac{1}{6}\omega^2\left[(Y^2 - Z^2)A_+ \sin(\omega T) + 2YZA_\times \cos(\omega T)\right] + O(3)\,, \\
g_{XY} &= -\frac{1}{2}g_{TY} + O(3)\,, \qquad g_{XZ} = -\frac{1}{2}g_{TZ} + O(3)\,, \\
g_{YY} &= 1 + \frac{1}{6}\omega^2 X^2 A_+ \sin(\omega T) + O(3)\,, \qquad g_{YZ} = \frac{1}{6}\omega^2 X^2 A_\times \cos(\omega T) + O(3)\,, \\
g_{ZZ} &= 1 - \frac{1}{6}\omega^2 X^2 A_+ \sin(\omega T) + O(3)\,. \qquad (4)
\end{aligned}
$$

In particular, we see that the gravitomagnetic terms $g_{TX}, g_{TY}, g_{TZ}$ are not null, since a gravitational wave transports angular momentum. The world-line of a spinning particle and the evolution of its spin vector can be found in Ref. 7. Notice that, by inspection of the $g_{TT}$ component, we see that clocks at rest in the laboratory measure different time due to their spin. Let us suppose that a test body is spinning around the Z-axis before the passage of the wave with constant spin $S_1^0 = 0, S_2^0 = 0, S_3^0 = s_0$. Due to the passage of the wave, the particle spin vector has non-vanishing spatial components varying with time and displacements along the three spatial directions; in particular, close to the reference spinning particle, the spin vector components turn out to be:

$$
\begin{aligned}
\tilde{S}^X &= -\frac{1}{2}\omega s_0[YA_\times \sin(\omega T) + ZA_+ \cos(\omega T)] + O(2) \\
\tilde{S}^Y &= \frac{1}{2}\omega s_0 X A_\times \sin(\omega T) + O(2) \\
\tilde{S}^Z &= \frac{1}{2}\omega s_0 X A_+ \cos(\omega T) + O(2)
\end{aligned}
$$

If we consider a second spinning particle at coordinates (X,0,0) (see Figure 1) with the same spin vector of the reference one, the passage of the GW induces a relative precession:

$$
\tilde{S}^X = 0, \quad \tilde{S}^Y = \frac{1}{2}\omega s_0 X A_\times \sin(\omega T), \quad \tilde{S}^Z = \frac{1}{2}\omega s_0 X A_+ \cos(\omega T).
$$

To estimate the effect, we write the magnitude of the component of the spin vector on the wave front $\tilde{S}^\perp = \sqrt{(\tilde{S}^Y)^2 + (\tilde{S}^Z)^2} = \frac{1}{2}s_0 h\omega X$, where we set $A_+ = A_\times = h$. If the two spin vectors (for instance in two laboratories) are separated by a

distance $L$, we have: $\tilde{S}^{\perp}(L) = \frac{1}{2} s_0 h \omega L$, provided that $\omega L/c \lesssim 1$, that is to say the wavelength should be far greater than $L$. Then, measuring $\tilde{S}^{\perp}(L)$ allows to get information on the product of the amplitude $h$ and frequency $\omega$ of the GW. In summary, the passage of the GW induces the relative spin variation

$$\frac{\Delta S}{S} = \frac{\tilde{S}^{\perp}(L)}{s_0} \simeq h \frac{\omega L}{c} \simeq h \frac{L}{\lambda_{GW}} \tag{5}$$

Since $h \simeq 10^{-21}$, the effect can be increased by using a large separation $L$ (which needs to be in any case smaller than $\lambda_{GW}$). For instance, for terrestrial laboratories $L = 10^7$ m, so that $\lambda_{GW} = 10^8$ m or greater could be suitable (corresponding to $\omega \simeq 1$ Hz, which is out of the terrestrial interferometers range). Accordingly, we get the following estimate for the relative spin variation:

$$\frac{\Delta S}{S} = 10^{-22} \tag{6}$$

Possibile candidates for measuring such a small effect could be the optical magnetometers: in particular GNOME[15] (Global Network of Optical Magnetometers for Exotic Physics) is based on synchronous measurements of optical-magnetometer signals from several devices operating in magnetically shielded environments in distant locations: by synchronously detecting and correlating magnetometer signals, transient events of global character may be identified. These devices are accurate magnetic field sensors, and their measurements can be related to spin dynamics. Another possibility is the use of a ferromagnetic sample: in fact since the passage of the GW induces a spin variation, the magnetization due to electron spins in two ferromagnetic samples will change after the passage of the GW. Eventually, if we consider mechanical gyroscopes, in terms of angles, the gravitational wave induces a precession angle of $10^{-22}$ rad; by comparison, remember that in the Gravity Probe B mission[10] the (1-year) Lense-Thirring effect was $10^{-9}$ rad, hence it seems very hard to use these devices to measure the effect of the gravitational waves.

## 3. Conclusions

We have discussed the interaction between spinning particles and gravitational waves. To this end, we have used Fermi coordinates, because of their operational meaning in terms of space and time measurements. As a consequence, we have reported the expression of the spacetime metric in Fermi coordinates adapted to the world-line of a spinning particle. The form of the metric emphasizes the transport of angular momentum by gravitational waves, as it testified by the presence of gravitomagnetic terms. Hence, this approach allow to explicitly calculate the gravitomagnetic effects in the gravitational field of the wave; it is important to point out that current interferometers are not sensitive to these effects. In particular, the interaction with gravitational waves causes spinning particles to change their spin. Consequently, spinning particles could be used as a probe to measure the gravitomagnetic effects of gravitational waves. Furthermore, we showed also that clocks at rest in the laboratory measure different time due to their spin. Indeed, these

effects are very small, as it usually happens when dealing with gravitational waves, however in our very preliminary analysis we have suggested that magnetometers or ferromagnetic samples could be uses to explore these effects.

**Acknowledgments**

The authors acknowledge Donato Bini and Andrea Geralico for useful discussions, suggestions, and knowledge sharing.

**References**

1. B. P. Abbott, R. Abbott, T. D. Abbott *et al.*, Observation of gravitational waves from a binary black hole merger, *Phys. Rev. Lett.* **116**, p. 061102 (Feb 2016).
2. B. P. Abbott, R. Abbott, T. D. Abbott *et al.*, Gw170817: Observation of gravitational waves from a binary neutron star inspiral, *Phys. Rev. Lett.* **119**, p. 161101 (Oct 2017).
3. I. Debono and G. F. Smoot, General relativity and cosmology: Unsolved questions and future directions, *Universe* **2** (2016).
4. F. De Felice and D. Bini, *Classical measurements in curved space-times* (Cambridge University Press, 2010).
5. M. Rakhmanov, Fermi-normal, optical, and wave-synchronous coordinates for spacetime with a plane gravitational wave, *Classical and Quantum Gravity* **31**, p. 085006 (Apr 2014).
6. M. Mohseni, Spinning particles in gravitational wave space-time, *Phys. Lett.* **A301**, 382 (2002).
7. D. Bini, A. Geralico and A. Ortolan, Deviation and precession effects in the field of a weak gravitational wave, *Phys. Rev. D* **95**, p. 104044 (May 2017).
8. M. L. Ruggiero and A. Tartaglia, Gravitomagnetic effects, *Nuovo Cim.* **B117**, 743 (2002).
9. I. Ciufolini and E. C. Pavlis, A confirmation of the general relativistic prediction of the lense–thirring effect, *Nature* **431**, p. 958 (2004).
10. C. F. Everitt, D. DeBra, B. Parkinson *et al.*, Gravity probe b: final results of a space experiment to test general relativity, *Physical Review Letters* **106**, p. 221101 (2011).
11. F. Bosi, G. Cella, A. Di Virgilio *et al.*, Measuring gravitomagnetic effects by a multi-ring-laser gyroscope, *Phys. Rev. D* **84**, p. 122002 (Dec 2011).
12. A. Ortolan, J. Belfi, F. Bosi *et al.*, The ginger project and status of the gingerino prototype at lngs, in *Journal of Physics: Conference Series*, (7)2016.
13. A. Tartaglia, A. Di Virgilio, J. Belfi, N. Beverini and M. L. Ruggiero, Testing general relativity by means of ring lasers, *Eur. Phys. J. Plus* **132**, p. 73 (2017).
14. M. L. Ruggiero and A. Tartaglia, Test of gravitomagnetism with satellites around the Earth, *Eur. Phys. J. Plus* **134**, p. 205 (2019).
15. S. Pustelny, D. F. J. Kimball, C. Pankow *et al.*, Global Network of Optical Magnetometers for Exotic (GNOME): Physics Novel scheme for exotic physics searches, *arXiv e-prints*, arXiv:1303.5524 (Mar 2013).

# Atom Interferometry with the Sr optical clock transition for gravity measurements

Leonardo Salvi, Liang Hu*, Jonathan Tinsley, Enlong Wang, Nicola Poli†, and Guglielmo M. Tino‡

*Dipartimento di Fisica e Astronomia and LENS - Università di Firenze, INFN - Sezione di Firenze, Via Sansone 1, I-50019 Sesto Fiorentino, Italy*

Inertial sensors based on atom interferometers have demonstrated extremely high levels of sensitivity thus allowing for accurate measurements of gravitational interactions. Most such sensors employ multi-photon transitions induced by laser fields to manipulate the atomic wavepackets that enter an interferometer. However, this approach cannot totally suppress the contribution of the laser phase noise to the gravitational signal when two sensors are separated by a large distance. In this article, a proof-of-principle experiment for an interferometer operating on the single-photon clock transition of strontium atoms is described. This configuration can suppress the effect of laser phase noise and therefore provides a promising condidate for the detection of low-frequency gravitational waves.

*Keywords*: Atom interferometry; gravity measurements; gravitational waves; optical clock transition.

## 1. Introduction

Recent advances in the manipulation of atomic ensembles have made it possible to exploit the wave nature of matter to build interferometers that are highly sensitive to many fundamental interactions[1]. In particular, due to massive nature of atoms, these devices are sensitive to gravitational interactions and can be configured to measure, for example, gravitational acceleration[2,3], gravity gradients[4,5], gravity curvature[6], and can provide tests of general relativity at the quantum level[7–9]. As compared to macroscopic masses, atomic probes offer a number of key advantages. For example, it is possible to accurately manipulate their motion by using the interaction with laser light. This allows for the construction of analogues to beam splitters and mirrors for matter, thereby allowing for the implementation of atom interferometers. Moreover, atoms are made to travel in vacuum and are therefore immune to undesired forces such as friction or vibrations. Other parasitic effects such as electromagnetic interactions can also be largely suppressed by a careful choice of the atomic species and of the atomic levels employed. Finally, atoms are quantum objects and therefore allow for the testing of gravity in the broad context of quantum mechanics where gravitational effects are usually negligible.

Atom interferometers are also good candidates for the detection of gravitational waves. Indeed, while optical interferometers such as LIGO and Virgo are sensitive to gravitational waves with frequencies larger than about 10 Hz due to seismic noise affecting the motion of the interferometer mirrors, atomic sensors are expected

*Also at Shanghai Jiao Tong University, Minhang, Shanghai, China.
†Also at CNR-INO, Firenze, Italy.
‡Also at CNR-IFAC, Sesto Fiorentino, Italy.

to be sensitive in the 1 mHz-10 Hz band, where new interesting sources, such as the coalescence of intermediate and massive black holes and of neutron stars, are expected to appear [10].

Generally, the beam splitters and mirrors of atom interferometers are implemented through multi-photon transitions induced by a couple of counterpropagating laser fields. In a gradiometric configuration, where the gravity difference between two locations is measured by two simultaneous interferometers, the phase noise of the two lasers is largely common mode and is therefore very highly suppressed by the differential configuration. However, when the separation between the two sensors is large, the atomic clouds will sense two counterpropagating fields emitted at different times. During this interval, noise is accumulated in the gradiometer which cannot be canceled and which puts stringent requirements on the laser frequency stability and on the laser platform vibrations. Indeed, it can be shown [11] that the amplitude spectral density of the laser fractional frequency fluctuations $\sqrt{S_f}$ should be at most of the order of the gravitational-wave-induced strain $h$, or $\sqrt{S_f} \leq h/2$. Based on the predicted gravitational wave amplitudes, the relative frequency fluctuations should not exceed $10^{-21}$. Comparing this requirement with the state of the art of lasers locked to ultrastable cavities [12], where $\sqrt{S_f} = 10^{-16}/\sqrt{\text{Hz}}$ at the Fourier frequency of 1 Hz, one can conclude that an improvement of five orders of magnitude in laser frequency stability would be required.

The effects of laser phase noise can be largely suppressed from the output of the interferometer if the transitions are driven by a single laser. Intuitively, this immunity arises because the photon phase does not acquire additional noise in the path between the two sensors. In this situation, a single-photon transition connecting two optically separated stable states must be used. In principle, if an interferometer was operated on such a transition, the low frequency limit of the strain sensitivity would only be set by the coherence time of the transition [13,14]. Such an interferometer would therefore measure the gravitational-wave-induced shift of the time required to travel the distance between the two interferometers and record this interval as a phase shift, via the internal atomic oscillation. Even though there are in principle many atomic species that can fulfill the requirement of having two long-lived states connected by an optical transition, only a few cases have the required immunity to external perturbations. It follows that the atoms usually employed to realize optical atomic clocks, such as Sr or Yb, should also be suitable for this application. However, despite the phase noise immunity of gradiometers operated on a clock transition, several challenges emerge in terms of the required laser system. For example, in order to fully address the atoms in the interferometer, the interrogation laser must have a high frequency stability. Additionally, because the atoms are in free flight, the Doppler effect from different atomic velocities will cause a broadening of the transition so that a high efficiency of the interferometer pulses can only be attained with large optical power.

In this article, a proof-of-principle experiment of a gravimeter and a gravity gradiometer operating on the single-photon $^1S_0$-$^3P_0$ transition of $^{88}Sr$ is described[15]. In Section 2, the main experimental configuration, including the laser cooled and trapped ensemble of strontium atoms in a vacuum chamber and the interferometry laser, as well as the interferometer sequence, will be described. In Section 3, the main experimental results regarding the gravimeter and the gradiometer will be provided. The amount of observable interferometry laser frequency noise cancellation in the gradiometer will be established. Finally, in Section 4 conclusions and future perspectives for this new kind of sensor will be provided.

## 2. Experimental setup and sequence

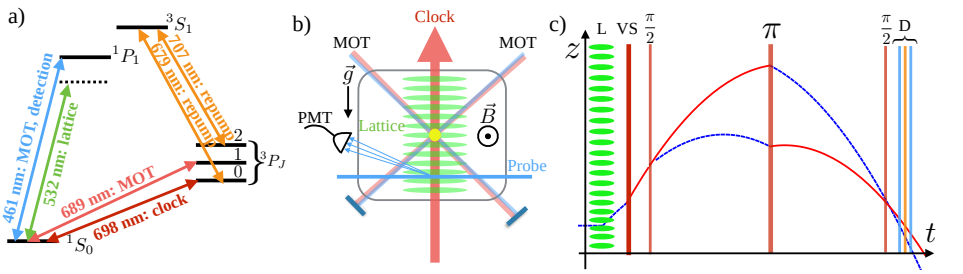



Fig. 1. a) Strontium level diagram showing the main optical transitions for cooling, trapping, interferometry, repumping and detection. b) Simplified drawing of the chamber, PMT: photomultiplier tube, MOT: magneto-optical trap, $\vec{B}$: applied magnetic field. c) Interferometer trajectories with the launch (L), velocity selection (VS), Mach-Zehnder interferometer ($\pi/2 - \pi - \pi/2$) and detection (D) consisting of probing on the blue transition and repumping (orange). The blue dashed (red) lines indicate atoms in the ground $^1S_0$ state (excited $^3P_0$ state).

In our experiment, the starting point is the ensemble of $^{88}Sr$ atoms. The atoms are laser cooled and trapped in a two-stage magneto-optical trap (MOT) first performed on the dipole-allowed $^1S_0$-$^1P_1$ transition and then on the $^1S_0$-$^3P_1$ intercombination transition (see Fig. 1(a)). This sequence produces almost $10^7$ atoms at a temperature of 1.2 $\mu$K with typical size of 100 $\mu$m at full width at half maximum.

While in the fermionic isotope of strontium, $^{87}Sr$, the clock transition is naturally allowed, in the $^{88}Sr$ boson it is strictly forbidden but can be induced with a static magnetic field. As a result, the laser cooled atoms are loaded in a 1D vertical optical lattice at 532 nm, where they are held for 65 ms. During this time, the current of one of the MOT coils, originally in the anti-Helmholtz configuration, is inverted to produce a homogeneous magnetic field of 330 G used to induce the clock transition. The lattice light is generated by a Coherent Verdi V6 laser which can deliver about 1 W of optical power to the atoms. When collimated to a beam waist of 350 $\mu$m, this laser yields a trap depth of $9E_r$, where $E_r = (k_B/2) \times 0.8$ $\mu$K is the recoil energy. This depth allows the atoms to be held and launched by accelerating the lattice

with reduced losses due to Landau-Zener tunneling. When a single interferometer is implemented, we load the atoms in the lattice and accelerate at $3g$ ($g$ is the gravitational acceleration) the cloud upwards by chirping the frequency of one of the lattice beams (Fig. 1(c)). In our gradiometer experiments, after launching a first cloud, the residual losses are used for a second launch at a slightly different final velocity.

The launching stage produces about $2.5 \times 10^5$ atoms in each cloud with a velocity difference $\Delta v = 8.5$ mm/s.

After the cloud preparation, the light from the interferometry (clock) laser is delivered to the atoms. The clock laser light, resonant with the $^1S_0$-$^3P_0$ transition at 698 nm, is produced by a grating-stabilized laser diode which is prestabilized by locking to an intermediate finesse ($F = 10000$) optical cavity and stabilized by locking to a high-finesse ($F = 500000$) ultrastable cavity with an ultra-low expansion spacer. When locked, this laser has a fast linewidth of 1 Hz and an estimated drift on the order of 1 Hz/s[16]. The light from the clock laser is preamplified by optically injecting another diode laser and finally amplified by a tapered amplifier. The light is then coupled into a 10-m single-mode polarization-maintaining fiber which can deliver up to 80 mW of light to the atoms. As opposed to multi-photon Raman or Bragg accelerometers, a single interferometer operated on the clock transition is sensitive to the single laser beam phase imprinted onto the atoms, rather than the phase difference of two counterpropagating beams. As a result, both the intrinsic laser noise and the optical components' vibrations will affect the gravimeter noise. The 10 m-long fiber contributes a large fraction of the noise which we can suppress with a fiber noise cancellation setup[17].

When we operate the gradiometer, we drive an acousto-optic modulator on the clock laser path with two frequency components in order to match the velocity difference of the two clouds after the double lattice launch. This method has the advantage of allowing the two clouds to be addressed separately and imprints a differential phase shift which can be used to characterize the gradiometer sensitivity.

The clock laser is first used to select a narrow velocity class of $10^4$ atoms, corresponding to an effective temperature of 1 nK[18]. Following velocity selection, the atoms undergo a Mach-Zehnder $\pi/2 - \pi - \pi/2$ sequence performed through clock laser pulses. A simplified view of the setup of the laser beams for the experiment is shown in Fig. 1(b) while the sequence for a single interferometer is shown in Fig. 1(c).

The interferometer output is detected by collecting the fluorescence induced by a thin sheet of light resonant with the $^1S_0$-$^1P_1$ transition onto a photomultiplier tube. In order to count the atoms in the excited $^3P_0$ state, two repumping lasers (Fig. 1(a)) are shone onto the atoms that transfer the population to the ground $^1S_0$ state where they are detected again by fluorescence collection.

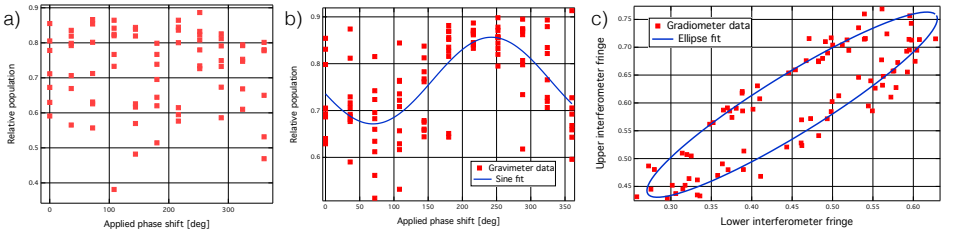



Fig. 2. Gravimeter and gradiometer results for interferometer durations of $2T = 10$ ms. a) Single interferometer fringe without the fiber noise cancellation b) Single interferometer fringe with fiber noise cancellation active. c) Plot of the upper interferometer fringe versus the lower one in the gradiometer configuration. The data show a clear correlation between the two fringes despite the presence of laser phase noise.

## 3. Gravimeter and gradiometer results

After releasing the atoms from the green optical lattice, we performed some preliminary measurements of the efficiency of the laser pulses resonant with the clock transition for untrapped atoms. The Rabi oscillations between the two interferometer states were measured by shining the clock laser onto velocity-selected atoms for a variable duration and then detecting the population of the two states. With our maximum optical power, we reached a Rabi frequency of $\Omega = 2\pi \times 750$ Hz and observed a damping time of 1.2 ms, with an efficiency limited to about 50% mainly due to the finite velocity distribution width. The consequence of this limited efficiency is a fringe contrast of 30%, which was measured to be constant as a function of the duration of the interferometer up to a time $2T = 10$ ms.

When gravimeter experiments were performed, using the sequence depicted in Fig. 1(c), we compared the fringe visibility with and without the fiber noise cancellation. As shown in Figs. 2(a) and (b), the phase noise induced by the fiber contributes a large amount of noise which can completely wash out the fringe for $2T = 10$ ms. The fringe acquires a clear visibility when the fiber noise cancellation is active. The remaining noise is due to the intrinsic laser frequency fluctuations and to the vibrations of the optical components in the laser's path from the source to the atoms. Although these noise sources appear relatively hard to suppress, it is in principle possible to improve these results by using a laser with better frequency stability and by reducing the vibrations of the optical components.

The potential of interferometry on the optical clock transition should appear in the gradiometer configuration, where a large suppression of the laser phase noise is expected. However, the relatively small separation of the two clouds and the small $T$ of our experiment would yield a negligible signal from the Earth's gradient. As already anticipated, the small velocity difference of the two clouds allows them to be addressed separately and to set an artificial differential phase shift. Using this feature, we performed gradiometric measurements with durations of $2T = 10$ ms. Plotting the upper interferometer fringe against the lower one displays the expected

correlation, as shown in Fig. 2(c). With a cycle time of 2.4 s, averaging this signal for $\tau = 400$ s results in an Allan deviation decaying as $\tau^{-1/2}$, an indication of white phase noise. The level of the recorded phase uncertainty lies only a factor of five above the estimated atom shot noise fluctuations. We estimate that this excess noise arises from detection noise.

## 4. Conclusions and perspectives

In this article, interferometry on the optical clock transition is presented. Results for the gravimeter show the expected sensitivity of the interference fringes to laser phase noise, while in the gradiometer configuration, the expected immunity to laser phase noise is established and operation close to the atom shot noise limit is observed. By further cooling the atoms, by using higher clock laser power and by using the fermion $^{87}$Sr for which the clock transition is naturally allowed, this scheme provides a potentially interesting avenue for the detection of low-frequency gravitational waves. Further improvements might include the implementation of large momentum transfer and of atomic squeezing to increase the sensitivity. Other applications are possible with this interferometer such as, for example, the test of the weak equivalence principle for atoms in coherent superpositions of optically separated states.

## Acknowledgments

We acknowledge financial support from INFN and the Italian Ministry of Education, University and Research (MIUR) under the Progetto Premiale “Interferometro Atomico” and PRIN 2015.

## References

1. Proceedings of the International School of Physics Enrico Fermi Course CLXXXVIII, Atom Interferometry, edited by G. M. Tino and M. A. Kasevich (SIF and IOS Press,Bologna, Amsterdam, 2014).
2. M. Kasevich and S. Chu, Phys. Rev. Lett. **67**, 181 (1991).
3. A. Peters, K. Y. Chung and S. Chu, Nature **400**, 849-852 (1999).
4. F. Sorrentino, Q. Bodart, L. Cacciapuoti, Y.-H. Lien, M. Prevedelli, G. Rosi, L. Salvi, and G. M. Tino, Phys. Rev. A **89**, 023607 (2014).
5. G. Rosi, F. Sorrentino, L. Cacciapuoti, M. Prevedelli and G. M. Tino, Nature **510**, 518-521 (2014).
6. G. Rosi, L. Cacciapuoti, F. Sorrentino, M. Menchetti, M. Prevedelli, and G. M. Tino, Phys. Rev. Lett. **114**, 013001 (2015).
7. M. G. Tarallo, T. Mazzoni, N. Poli, D. V. Sutyrin, X. Zhang, and G. M. Tino, Phys. Rev. Lett. **113**, 023005 (2014).
8. X.-C. Duan, X.-B. Deng, M.-K. Zhou, K. Zhang, W.-J. Xu, F. Xiong, Y.-Y. Xu, C.-G. Shao, J. Luo, and Z.-K. Hu, Phys. Rev. Lett. **117**, 023001 (2016).

9. G. Rosi, G. D'Amico, L. Cacciapuoti, F. Sorrentino, M. Prevedelli, M. Zych, C. Brukner and G. M. Tino, Nat. Commun. **8**, 15529 (2017).
10. S. Dimopoulos, P. W. Graham, J. M. Hogan and M. A. Kasevich, Phys. Lett. B **678**, 37 (2009).
11. J. G. Baker and J. I. Thorpe, Phys. Rev. Lett. **108**, 211101 (2012).
12. W. Zhang, J. M. Robinson, L. Sonderhouse, E. Oelker, C. Benko, J. L. Hall, T. Legero, D. G. Matei, F. Riehle, U. Sterr and J. Ye Phys. Rev. Lett. **119**, 243601 (2017).
13. N. Yu and M. Tinto, Gen. Relativ. Gravit. **43**, 1943 (2011).
14. P. W. Graham, J. M. Hogan, M. A. Kasevich and S. Rajendran, Phys. Rev. Lett. **110**, 171102 (2013).
15. L. Hu, N. Poli, L. Salvi and G. M. Tino, Phys. Rev. Lett. **119**, 263601 (2017).
16. M. G. Tarallo, N. Poli, M. Schioppo, D. V. Sutyrin and G. M. Tino, Appl. Phys. B **103**, 17 (2011).
17. L.-S. Ma, P. Jungner, J. Ye and J. L. Hall, Opt. Lett. **19**, 1777 (1994).
18. K. Moler, D. S. Weiss, M. Kasevich and S. Chu, Phys. Rev. A **45**, 342 (1992).

# Advances in high resolution inertial rotation sensing

Karl Ulrich Schreiber, André Gebauer, Jan Kodet

*Technical University of Munich, Research Unit Satellite Geodesy, Arcisstr. 21, 80333 München, Germany*

Jon-Paul Renee Wells

*Dodd-Walls Centre for Photonic and Quantum Technologies, School of Physical and Chemical Sciences, University of Canterbury, PB4800, Christchurch 8140, New Zealand*

Large ring lasers have improved significantly in recent years, such that we are now in the position to separate and mitigate error sources that are not directly related to the rotation sensing process from the Sagnac interferogram. As a result, we are now able to reduce the measurement error of the 16 $m^2$ G ring laser of the Geodetic Observatory Wettzell by a factor of two. Improvements in the measurement of the relevant parameters for the backscatter correction remove most of the sensor drift effects, so that the backscatter induced coupling is no longer a real concern. Now that we can separate the mechanisms of the error contributors much better, we can mitigate them in a more effective way. In this paper we report on the latest progress.

## 1. Introduction

The rotation rate of the Earth and the orientation of the rotational axis of the Earth in space are the quantities which link the terrestrial (ITRF) and the celestial (ICRF) reference frames. Until now, the only way of obtaining the rotational velocity of the Earth as well as the variation of the orientation of the respective rotational axis with sufficient accuracy has been to interferometrically observe a set of radio sources - quasars at the perimeter of the observable universe form an external set of stable markers. This can be used to link the respective reference frames together. Today residuals of about 10 $\mu$s for the measurement of the Length of Day (LoD) and 0.5 nrad (0.1 milli-arc-seconds) for measurement of the pole position are achieved by a network of VLBI radio telescopes and GNSS observations[1,2]. These efforts are carried out by the services of the International Association of Geodesy (IAG). The operation of such a large network requires a significant number of radio telescopes and a very substantial maintenance and processing effort. Huge amounts of data of up to 1 Tb of volume per radio telescope are recorded in each of the 24h measurement sessions, which until recently required physical transport over large intercontinental distances to perform the correlation process in the analysis centers. Data latency, combined with the fact that there is no continuous measurement coverage, suggests that complementary methods for the accurate estimation of Earth rotation and polar motion should be explored. Optical Sagnac interferometry offers an independent technical approach. This is desirable in order to identify any technique related biases if they exist. Large ring lasers are potential candidates for such an alternative measurement technique. This technology is widely used in inertial aircraft navigation and can measure angular velocities absolutely, i.e. independent

from an external reference frame. The beat note, $\delta$f, of two counter propagating mono-mode laser beams in the ring cavity is proportional to the experienced rate of rotation $\Omega$ of the entire apparatus and is described by the ring laser equation

$$\delta f = \frac{4A}{\lambda P}\mathbf{n} \cdot \mathbf{\Omega}, \tag{1}$$

where $A$ is the area circumscribed by the laser beams, $P$ the corresponding perimeter, $\lambda$ the optical wavelength and $\mathbf{n}$ the normal vector on the plane of the laser beams. However, the demands on such instruments for the application in geodesy are extremely high, and cannot be met by existing commercial devices. These requirements can be summarized as below:

- sensitivity to angular motion of less than 0.1 $prad/s$ over an integration time of about 1 h
- sensor stability of 1 part in $10^9$ over several months (requirement for the measurement of the Chandler and Annual Wobble with high temporal resolution)
- resolution for the sensor orientation of approximately 1 $nrad$, corresponding to a polar motion effect of around 1 cm at the pole

All these requirements demand a substantial improvement of ring laser technology over existing navigational instrumentation in all aspects of the sensor design. A significant upscaling of the physical parameters of the Sagnac interferometer is the most promising approach in order to make ring lasers a viable technique for applications in space geodesy. The design of the large ring laser gyro G (Grossring), located at the Geodetic Observatory Wettzell in eastern Bavaria is one way of approaching these demands[3]. This single axis gyro is a 16 meter perimeter square ring laser with a length of 4 m on each side, which utilizes a mixture of helium and neon as a gain medium, similar in many ways to the well established systems used in aircraft navigation. However, the vastly increased size of the G ring laser with a Q factor well in excess of $5 \times 10^{12}$ provides a measurement sensitivity to angular velocities, much higher than any other rotation sensing instrument on the ground. A nearly monolithic sensor block design from the low thermal expansion material ZERODUR located in an underground laboratory provides the necessary passive mechanical stability of the G ring laser body for the desired level of performance. The ring laser G has now matured to the point where small geophysical signals with very long periods such as the Chandler and Annual wobble of the rotating Earth are measurable[4]. This has not been achieved through higher sensitivity but rather through considerable improvements in the overall cavity stability[5]. Such long-term laser stabilization is one of the key requirements for the proposed terrestrial measurements of general relativistic phenomena such as the Lense–Thirring (or frame-dragging) effect[6–8].

## 2. Backscatter Correction

The most important contributor to the sensor drift arises from the backscatter coupling between the two counter-propagating laser beams, causing frequency pulling and pushing. This quantity is not constant over time, but strongly depends on the phase of the laser beam as it hits each mirror. Variations in ambient temperature and atmospheric pressure cause a tiny change in mirror separation and hence a slow variation in the backscatter coupling. While atmospheric pressure variations can be controlled with the installation of a pressure stabilizing vessel, isolating the ring laser from the rest of the laboratory, changes in temperature are not well enough controlled for a structure as large as G. While the temperature inside the ring laser vessel only varies at a level of 3 – 5 mK per day following the annual temperature cycle of the seasons, this is enough to cause a significant drift of the observed interferometer beat note. By observing the amount of backscattered light in each of the laser beams, a correction value $\Delta f_S$ can be computed according to Ref. 9 as

$$\Delta f_S = \frac{1}{2} f_S m_1 m_2 \cos\varphi, \tag{2}$$

where $f_S$ is the observed beat note, $m_1$ and $m_2$ are the fractional beam modulations and $\varphi$ the phase angle between them. For a given mirror quality, $m_1$ and $m_2$ scale approximately as $L^{-2.5}$ for a cavity of linear size $L$. The ratio of the correction to the observed interferometer beat note to the beat note itself scales approximately as $L^{-5}$. This shows how important it is to make the ring laser cavity as large as possible, while sufficient mechanical stability of the entire structure has still to be maintained. In the absence of active control mechanisms it appears that a symmetrical construction with a length of 6 – 10 m on a side seems to be feasible for a large ring laser structure.

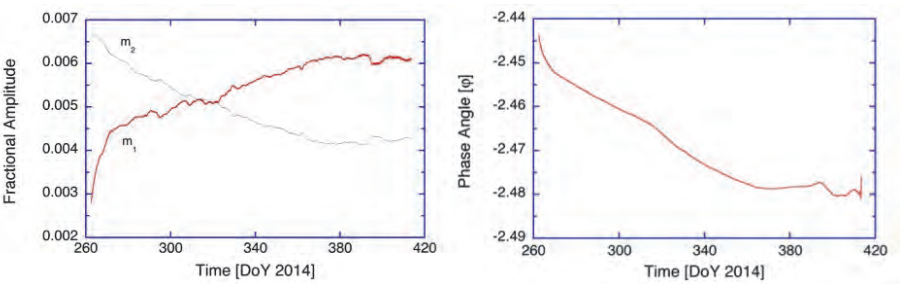



Fig. 1. The parameters $m_1$, $m_2$ and $\varphi$ of the backscatter correction (Eq. (2)) of a time series of 130 days of length. Due to the mechanically monolithic construction of the G ring laser and the active pressure stabilized environment of the gyroscope, there are only small changes in the backscatter pulling observable.

Figure 1 shows the observed variation of the fractional beam modulations and the phase angle between them over a period of more than 130 day. Since this dataset was taken with the atmospheric pressure stabilizing vessel in operation, the

observed variations of the correction quantities $m_1$, $m_2$ and $\varphi$ are entirely caused by temperature variation. When these corrections are applied to the gyro observations, we are able to remove the backscatter induced drift in the observation of the Earth rotation signal. However, some short-term fluctuations and some irregular steps still remain in the measurement signal. This signature compares well with the tiltmeter readings obtained from a high resolution tilt sensor, placed on top of the ring laser structure. Figure 2 depicts the observed Earth rotation rate after the backscatter correction has been applied to the full 130 days of observation.

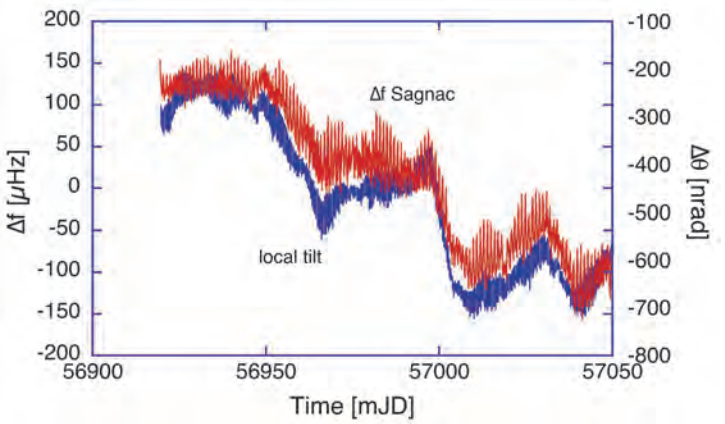



Fig. 2. 130 days of ring laser observation of Earth rotation velocity. After the backscatter correction procedure has been performed, the effects of a number of geophysical signals, like local tilt, diurnal polar motion, solid Earth tides, the Chandler and the Annual wobble remain.

The lower (blue) curve, corresponding to the variation of local $g$ over the same period of time, shows the changes experienced in the North/South tilt of the ring laser monument. The step like drops in $\Delta\Theta$ after day 56950 and around 57000 correspond to the response of the ring laser monument to significant rain falls at the observatory. The response of the tiltmeter corresponds well with the trend in the observed gyro signal represented in the same diagram in the upper (red) curve. After the tilt correction we have computed the known geophysical signals, which cause a change of the projection of $\Omega$ on the area vector of our Sagnac interferometer. Figure 3 represents the combined effect of all known signals over the length of the measurement. The section between the two dashed lines is shown enlarged in the inset in order to emphazise the fine structure of the signals. $\Delta\Omega$ is expressed in the units of the observed beat note in accordance with Eq. (1).

Once all the model corrections (diurnal polar motion, solid Earth tides, ocean loading and the variation of latitude) are applied, the Chandler and the Annual wobble, as measured by VLBI are also reduced from the data. The result is a continuous dataset of the rotation rate of the Earth with a spacing of 1 hour between

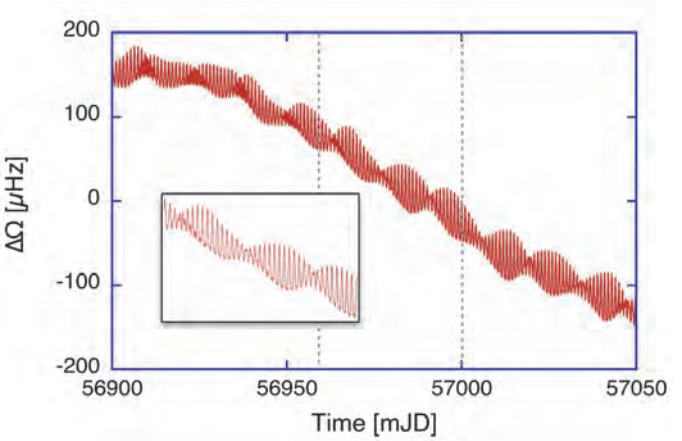



Fig. 3. The combined effect of the known geophysical signals, which have an effect on the North - South projection of the instantaneous Earth rotation vector onto the area vector of the laser gyro. The abscissa shows the variation of the measured angular velocity, expressed as a variation in the beat note of the Sagnac interferogram, derived from Eq. (1). The area between the two dashed lines is shown enlarged in the inset.

the data points as shown in Fig. 4. Although some small short-term variations at the level of less than 1 part in $10^7$ remain, a stable rotation rate is obtained over the full span of 130 days. The successful observation of a continuous long dataset of

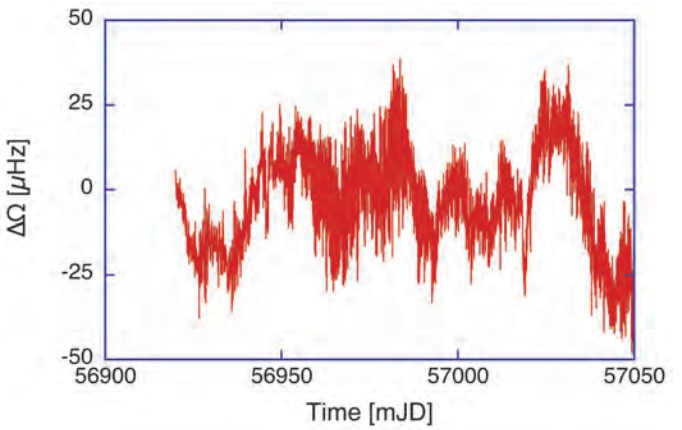



Fig. 4. The backscatter corrected Earth rotation rate over 130 days of observation. Each data point was averaged over 1 hour and the known geophysical signals have been removed from the measurements. The abscissa shows the variation of the measured angular velocity expressed as a variation in the beat note of the Sagnac interferogram, derived from Eq. (1).

high resolution ring laser observations of the Earth rotation velocity over 130 days marks a great progress in the application of Sagnac interferometers on geodetic

observations. We note in particular the stability of the measurements. Despite getting close to the target of a gyroscopic observation of the variations in the length of day (LoD), we are not quite there. However the short-term variation is within a factor of two of the order of the typical amplitude of LoD variations. It would appear that small variable non-reciprocal effects are playing a role.

## 3. Conclusion

Although the application of the backscatter correction has improved the long-term stability of our G ring laser to $\Delta\Omega/\Omega \leq 10^{-7}$ over the entire measurement series, we still consider this work in progress. We have identified the tiltmeters as one weak link in our correction process for a single component ring laser. In addition, the backscatter correction procedure itself will have limitations. The long term stability for the measurement of the fractional amplitudes as well as the corresponding phase angle, remains to be proven. In order to use laser gyros for fundamental physics, it is important to have a three component ring laser of sufficient size in operation. Furthermore it appears to be desirable to locate this gyro in a deep underground facility, such that seasonal temperature variations, excessive noise from the interaction between wind and the top soil and the variability of the hydrology are considerably reduced. Since the Lense-Thirring frame dragging effect appears as a DC quantity, it is also necessary to determine the absolute scale factor of the instrument in order to contribute to fundamental physics. However this ambitious goal appears to be much more practical now than several years ago.

## Acknowledgements

The authors acknowledge support by Deutsche Forschungsgemeinschaft (DFG) grant SCHR 645/6-1.

## References

1. R. Haas; *Investigating High-Frequency Earth Orientation Variations with Continuous Geodetic VLBI Campaigns*; IVS 2006 General Meeting Proc. (Concepcion, Chile), pp. 316–319, (2006).
2. M. Rothacher, G. Beutler, R. Weber, J. Hefty; *High-frequency variations in Earth rotation from Global Positioning System data*; J.Geophys.Res., **106** (B7), 13711, (2001).
3. K.U. Schreiber, T. Klügel, A. Velikoseltsev, W. Schlüter, G.E. Stedman, J.-P. R. Wells; *The Large Ring Laser G for Continuous Earth Rotation Monitoring*; J. Pure Appl. Geophys., **166**, 1485 (2009).
4. Schreiber, K. U., Klügel, T., Wells, J.-P. R., Hurst, R. B., Gebauer, A. *How to detect the chandler and the annual wobble of the earth with a large ring laser gyroscope*, Phys. Rev. Lett., **107**(17), 173904-1, (2011). http://doi.org/10.1103/PhysRevLett.107.173904.

5. K. U. Schreiber and J.-P. R. Wells; Rev. Sci. Instrum. **84**, 041101 (2013).
6. Everitt, C. W. F., DeBra, D. B., Parkinson, B. W., Turneaure, J. P., Conklin, J. W., et *Gravity Probe B: Final Results of a Space Experiment to Test General Relativity*, Phys. Rev. Lett., **106**(22), 221101, (2011); http://doi.org/10.1103/PhysRevLett.106.221101.
7. I. Ciufolini, E. C. Pavlis, F. Chieppa, E. Fernandes-Vieira, and J. Perez-Mercader; Science 279, 2100 (1998).
8. Bosi, F., Cella, G., Di Virgilio, A., Ortolan, A., Porzio, A., Solimeno, S., et al. *Measuring gravitomagnetic effects by a multi-ring-laser gyroscope*, Phys. Rev. D, **84**(12), 122002, (2011); http://doi.org/10.1103/PhysRevD.84.122002.
9. Hurst, R. B., Rabeendran, N., Schreiber, K. U. and Wells, J.-P. R.; *Correction of backscatter-induced systematic errors in ring laser gyroscopes*, Appl. Opt, **53**(31), 7610–7618, (2014), http://doi.org/10.1364/AO.53.007610.

# LAGRANGE: How to measure the angular momentum of the galactic dark halo

Angelo Tartaglia

*OATO, INAF, Strada Osservatorio 20*
*Pino Torinese, 10125, Italy*
**E-mail: angelo.tartaglia@inaf.it*
*www.oato.inaf.it*

The Lagrange points of the Sun/Earth pair form an interesting reference frame corotating with the Earth around the barycenter of the pair. Here we propose to use them as a base for a "rigid" physical loop allowing for the propagation of electromagnetic signals along a closed contour. If a gravito-magnetic field is concatenated with the loop, it produces an asymmetry of the times of flight in opposite directions, just as for the classical Sagnac effect. Due to the large scale of the loop, an experiment based on these premises could allow for a measurement of the angular momentum of the Sun and also the detection of the angular momentum of the dark halo of the Milky Way, if it exists.

*Keywords*: Gravito-magnetism; Sagnac effect; Lagrange points; Dark halo.

## 1. Introduction

Dark matter is today accepted as one of the ingredients that make up the universe as far as we can observe it. The need for more mass than that people could directly observe was recognized long time ago. Letting aside some early remarks going back even to the $19^{th}$ century, in 1932 Oort [1] observed that the Milky Way needed more mass than that of the visible stars in order to justify its internal kinematics. One year later Fritz Zwicky [2], studying the Coma galaxy cluster concluded that it had to be hundreds of times more massive than suggested by its appearance. The ensuing open problem was that of the "missing mass"; later the mysterious ingredient was named "dark matter". Today the evidence in favor of the existence of dark matter is multiple and varied: at the cosmological scale, observations tell us that dark matter in the universe should be approximately 5.5 times more massive than ordinary matter [3]; the temperature of the plasma trapped in the globular clusters is too high to be consistent with stability and suggests a deeper gravitational potential well; rotation curves of the stars in spiral galaxies correspond to peripheral speeds higher than the values deduced from the visible mass distribution [4]; often observed gravitational lensing and microlensing indicate mass distributions which do not manifest themselves in any other way.

In parallel with the search for evidence of dark matter a continuous theoretical investigation has been and is under way for conjecturing what kind of particles could be the constituents of dark matter. The main conviction of the present paper is, however, that an important task, before speculating on the nature of dark matter, should be to fully characterize its gravitational effects and, since the standard theory used to describe gravity is general relativity (GR), to explore all effects of that theory at the scale of the dark matter phenomenology. The simplest explanation for the

observed rotation curves of galaxies is to assume that the whole galaxy (including ours) is immersed into a huge dark matter halo[5]. Such halo extends beyond the visible disk of the galaxy, since we see that the anomalously high rotation speeds continue outside the disk, as shown by the radiation emitted by neutral hydrogen. The halo is more massive than the galaxy and we may also infer that it should also rotate together with the stars. If so, GR predicts that the dark halo, besides an additional gravitational pull on visible matter, should also generate a *gravito-magnetic* (GM) field, depending on its angular momentum; that GM field would in turn influence the motion of freely falling bodies within it. Why then not try and look for effects of that sort?

This is the simple idea which is the starting point of this work. An experiment at the scale of the internal solar system will be illustrated, aimed at evidencing possible GM effects of the angular momentum of the dark halo of the Milky Way.

## 2. Gravito-magnetism

A generic form for the metric of a curved space-time in arbitrary coordinates may be written:

$$ds^2 = g_{00}c^2dt^2 + 2g_{0i}cdtdx^i + g_{ij}dx^idx^j \tag{1}$$

The usual conventions have been adopted: index 0 is for time, Latin indices denote space variables; Einstein convention on index summation is applied.

The mixed components of the metric tensor, $g_{0i}$, tell us that the source of gravity is moving with respect to the observer. If the movement is a rotation no coordinate change can globally make the $g_0i$'s disappear and those terms correspond to physical effects we may hope to detect by suitable experiments.

In weak field conditions, i.e. when the difference between the actual metric tensor and Minkowski metric is very small (almost everywhere in the universe), the mixed space-time components of the metric tensor may be interpreted as the components of an ordinary vector potential similar to the one of Maxwell electromagnetism (EM)[6]:

$$cg_{0i} = h_i \tag{2}$$

Correspondingly we may describe the gravitational interaction as due to a *gravito-electric* field, causing the usual gravitational attraction, and a *gravito-magnetic* field, $\overline{B}_g$ stemming out of the $\overline{h}$ vector potential as in the case of electromagnetism:

$$\overline{B}_g = \overline{\nabla} \wedge \overline{h} \tag{3}$$

The influence of the GM field on freely falling particles is the analogue of the effect of a magnetic field on moving charges.

### 2.1. *Gravito-magnetism and Sagnac*

When describing EM signals, their line element is null and (1) may be read as a second order equation whose unknown is $cdt$:

$$0 = g_{00}c^2dt^2 + 2g_{0i}cdtdx^i + g_{ij}dx^idx^j \tag{4}$$

The coordinated time interval for the propagation of EM signals over a space displacement expressed by the $dx^i$'s is then:

$$cdt = -\frac{g_{0i}}{g_{00}}dx^i \pm \frac{1}{g_{00}}\sqrt{(g_{0i}dx^i)^2 - g_{00}g_{iy}dx^idx^j} \tag{5}$$

If now, by some appropriate device, we force the EM signal to move along a closed route (in the given reference frame) and come back at the starting point, we may calculated the proper time of flight $\tau$ integrating (4) along the whole path. What happens is that, like in the typical Sagnac measurement, right-handed and left-handed travels along the same circuit require different times. The proper time asymmetry $\delta\tau_G$ is[7]:

$$\delta\tau_G = 2\sqrt{g_{00}}\oint\frac{g_{0i}dx^i}{g_{00}} \tag{6}$$

## 3. LAGRANGE

The Lagrange points (L-points) of the Sun-Earth pair occupy positions which stay fixed with respect to the Earth, thus forming a natural co-rotating reference frame at the scale of the inner solar system. Their configuration is shown in Fig. 1.

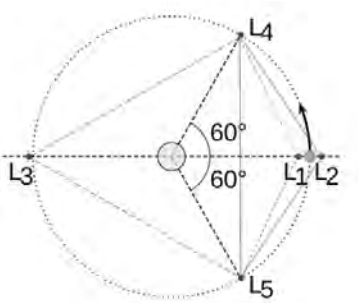



Fig. 1. Location of the Lagrange points of the Sun-Earth pair, along the terrestrial orbit. The straight lines outline possible closed paths for the EM signals of the LAGRANGE experiment.

Locating transponders in the $L$-points would allow to implement the closed paths required for a wide scale Sagnac-like experiment: let us call it LAGRANGE. Not considering $L_3$, which is beyond the Sun, a possible choice would be the triangle $L_2 - L_5 - L_4$. The typical scale of such configuration is the astronomic unit (AU), i.e. $1.5 \times 10^{11}$m: the sides $L_4 - L_2$ and $L_2 - L_5$ are just 1 AU long.

### 3.1. *Measuring the angular momentum of the Milky Way*

As written in the introduction, if the Milky Way (as the other spiral galaxies) is immersed in a huge dark halo and the latter, as it is reasonable to expect, rotates at the same rate as the visible stars, the solar system is in turn immersed into the GM field originated by the rotating halo.

The distance of the Sun from the center of the Milky Way, $R_G$, is 8 kpc i.e. $\sim 2.5 \times 10^{20}$m. The diameter of the orbit of the Earth, $D_E$, is $\sim 3 \times 10^{11}$m: as we see, the ratio $D_E/R_G$ is $\sim 10^{-9}$. We may then infer that the relative change of the galactic GM field met by the Lagrangian triangle during the year will reasonably be proportional to $10^{-9}$: in practice we are allowed to treat $\overline{B}_G$ as constant and uniform across the inner solar system.

The assumption above, together with Stokes theorem, simplifies the calculation of the formula for the time of flight (tof) asymmetry along the Lagrangean triangle:

$$\delta\tau_G = \frac{2}{c}\oint \overline{h}\cdot d\bar{l} = \frac{2}{c}\int_S \overline{B}_G \cdot \hat{u}_n dS \cong \frac{2}{c} B_G S\cos\alpha \tag{7}$$

Here of course $d\bar{l}$ is the elementary displacement vector along the trajectory of the EM signal; $\hat{u}_n$ is the unit vector perpendicular to the plane of the trajectory; $S$ is the contoured area. The EM analogy and the symmetry tell us that on the galactic plane $\overline{B}_G$ is perpendicular to the plane; angle $\alpha$ is the inclination of the ecliptic plane with respect to the galactic plane.

In order to fix orders of magnitude, consider that for the $L_2 - L_5 - L_4$ triangle it is $S_{425} \approx 9.9 \times 10^{21}\ m^2$ and $\alpha \approx 62.6°$. Assuming a sensitivity of the measurement of the tof asymmetry in the order of $10^{-15}$s, from (7) we deduce a lower limit for the weakest detectable GM field, as big as $3 \times 10^{-29}\ m^{-1}$. For comparison, I recall that the terrestrial GM field, at the surface of the planet, is $\sim 10^{-30}\mathrm{m}^{-1}$.

### 3.2. *Measuring the angular momentum of the Sun*

We have discussed above the (dark) GM field of the Milky Way, but we should not forget that LAGRANGE would also be sensible at least to one more contribution: the GM field of the Sun.

In the case of our star, the weak field approximation for the line element in a non-rotating reference frame and polar coordinates may be written:

$$ds^2 = (1-2\frac{GM}{c^2 r})c^2dt^2 - (1+2\frac{GM}{c^2 r})dr^2 - r^2 d\theta^2 - r^2\sin^2\theta d\phi^2 + 4\frac{GJ\sin\theta}{c^3 r^2}(cdt)(r\sin\theta d\phi) \tag{8}$$

Here $M$ is the mass of the Sun and $J$ its angular momentum; $\overline{J}$ is assumed to be perpendicular to the ecliptic plane. If motion (then the corresponding space part of the line element) is contained in the ecliptic plane, it is $\theta = \pi/2 =$ constant. It is also convenient to refer to $m = \frac{GM}{c^2}$ $(= 1475\mathrm{m})$ and $j = \frac{GJ}{c^3}$ $(= 4.71 \times 10^6\mathrm{m}^2)$. This simplified line element cannot be used in the case of the galactic dark halo, because there we are *inside* the mass distribution, whereas here we assume to be in vacuo[8].

The view point for the LAGRANGE measurement would be the one of a terrestrial observer, which means that we have to pass to a co-rotating reference frame at the orbital angular velocity of the Earth $\Omega$. At the approximation level we are using, the orbital velocity of the Earth (and the $L$-points) can be assumed to be Keplerian:

$$\Omega = c\sqrt{\frac{m}{r^3}} \tag{9}$$

Putting all the above assumptions together, we see that the line element for the terrestrial observer (co-moving with the LAGRANGE triangle) is:

$$ds^2 \approx (1-3\frac{m}{r})c^2dt^2 - (1+2\frac{m}{r})dr^2 - r^2d\phi^2 + 2(2\frac{j}{r^2} - \sqrt{\frac{m}{r}})(cdt)(rd\phi) \tag{10}$$

.

As can be seen, the mixed time-space component of the metric tensor now contains a contribution due to the angular momentum of the Sun (proper GM term) and a kinematical contribution due to the rotation of the frame. Both terms concur to the GM field perceived by the rotating observer, then to the right/left tof difference along the Lagrangean triangle. The calculation is now a bit more complicated than in the case of the galactic halo effect, since now the field cannot be assumed to be constant along the trajectory. In principle we should also consider that the EM rays which physically form the sides of the triangle are not really straight because of the gravitational lensing effect that bends them. Luckily that effect, besides being very small, can safely be neglected because it is symmetric (it does not depend on the rotation sense), then cancels when subtracting the left and right tof from one another.

Eq. (6), preserving a consistent approximation level, is now:

$$\delta\tau \approx \frac{2}{c}\oint (2\frac{j}{r^2} - 3(\frac{m}{r})^{3/2} + \frac{3m}{2R}\sqrt{\frac{m}{r}} - \sqrt{\frac{m}{r}})rd\phi \tag{11}$$

$R$ is the average radius of the terrestrial orbit ($R \approx 1.495 \times 10^{11}$m). For $L_4$ and $L_5$ it is $r_{L_4} = r_{L_5} = R$; $L_2$ is a bit further: $r_{L_2} = R + a = 1.51 \times 10^{11}$m. To do the calculation, the sides are assumed to be straight and the typical equation is:

$$r = \frac{b}{\cos\phi - \phi_b} \tag{12}$$

The minimum approach to the origin (the barycenter of the Sun-Earth pair) is $b$; the corresponding azimuth is $\phi_b$. It is:

$$b_{45} = \frac{R}{2} \tag{13}$$

$$b_{42} = b_{25} = \frac{R}{2}\sqrt{\frac{3(R+a)^2}{R^2+Ra+a^2}} \tag{14}$$

Performing the calculation and inserting numbers, the purely gravito-magnetic contribution of the Sun, $\delta\tau_j$ is:

$$\delta\tau_j = \sqrt{3}\frac{8j}{c}\frac{R+2a}{R(R+a)} \approx 1.47 \times 10^{-12}\text{s} \tag{15}$$

The other kinematical terms appearing in Eq. (11) provide bigger contributions, which are, in the order from smallest to biggest:

$$\delta\tau_k \approx (5.35 \times 10^{-9} - 1.428 \times 10^{-8} - 0.36172)\text{s} \tag{16}$$

## 4. Discussion and conclusion

We have presented a principle experiment based on the typical Sagnac approach applied to a rotating closed path for EM signals having three Lagrange points of the Sun-Earth pair as vertices. The experiment could evidence the presence of a galactic GM field, dominated by the dark component in the halo incorporating the Milky Way, and could measure the angular momentum of the Sun inferred from its GM field. The signals of different origin would be superposed to kinematical contributions coming from the pure rotation of the Earth along its orbit. Numbers are indeed small and the proper GM terms are the smallest, however comparable and even easier to attain than the ambitious goals of other space missions, such as LISA. A possible strategy to mark interesting weak contributions could be to produce a time modulation, which could evidence them when exploring the cumulated data: this approach will be explored in further work now in preparation.

I have not discussed here the practical problems that LAGRANGE should surmount in order to be implemented, but I am confident that they are all within the possibilities of existing technologies. Both on the side of improving technologies and methods, and of reducing unitary costs, a global strategy would be to associate LAGRANGE to other experiments and missions under way or in preparation.

## References


1. J. H. Oort, *Bulletin of the Astronomical Institutes of the Netherlands* **6**, 249 (1932).
2. F. Zwicky, *Helv. Phys. Acta* **6**, 110 (1933).
3. Planck collaboration, `arXiv:astro-ph.CO/1807.06205v1` *Planck 2018 results. I. Overview, and the cosmological legacy of Planck*, to appear in *Astronomy & Astrophysics* , pp. 161, 2018.
4. V. Rubin, N. Thonnard, W. K. Ford Jr., *The Astrophysical J. Lett.* **225**, L107-L111 (1978).
5. V. Rubin, N. Thonnard, W. K. Ford Jr., *The Astrophysical J.* **238**, 471 (1980).
6. M. L. Ruggiero & A. Tartaglia, *Nuovo Cimento B* **117**, 743-767 (2002).
7. A. Tartaglia & M. L. Ruggiero, *Am. J. Phys.* **83**, 427-432 (2015).
8. A. Tartaglia, *Int. J. Mod. Phys. D* **27**, 1847012 (2018).

# Proposal for laboratory generated gravitomagnetic field measurement

G. V. Stephenson,[†] W. Rieken,[*] and A. Bhargava[*‡]

[†]*Seculine Consulting,*
*Houston, TX 77005, USA*
*E-mail: seculine@gmail.com*

[*]*Graduate School of Materials Science,*
*Nara Institute of Science and Technology,*
*Ikoma, Nara 630-0192, Japan*

[‡]*Scotch College, Melbourne, Australia 3122*

We describe how to create a measurable unbalanced gravitational acceleration using a gravitomagnetic field surrounding a superconducting toroid. Such a gravitomagnetic toroid has been experimentally quite impractical. However recent advances in nanorod superconducting wire technology has enabled a new class of SMES devices operating at current densities and magnetic field strengths sufficient to develop measurable gravitomagnetic fields, while still maintaining mechanical integrity. In the present paper an experimental SMES toroid configuration is proposed that uses an absolute quantum gravimeter to measure acceleration fields along the axis of symmetry of a toroidal coil, thus providing experimental confirmation of the additive nature of the gravitomagnetic fields, as well as the production of a linear component of the overall acceleration field.

*Keywords*: Gravitational, Gravitomagnetic, Lense-Thirring, Superconducting Magnetic Energy Storage, SMES, nanorods, nanowires.

## 1. Introduction

When Forward[1] first proposed a gravitomagnetic toroid for unbalanced gravitational force production in 1962, any experimental realization was quite impractical. However recent advances in high-temperature superconducting (HTSC) nanorod wire (nanowire) technology[2] has enabled a new class of superconducting magnetic energy storage (SMES) devices operating at current densities and magnetic-field strengths sufficient to develop measurable gravitomagnetic fields, while still maintaining mechanical integrity. In the present study, an experimental SMES toroid configuration is proposed that uses a quantum gravitometer to measure acceleration fields along the axis of symmetry of a toroidal coil, thus providing experimental confirmation of the additive nature of the gravitomagnetic fields, as well as the production of a linear component of the overall acceleration field. See Fig. 1 for details.

In Forward's gravitational generation coil described in this paper, superconducting electron flow provides the change in mass current in the toroid.

## 2. Background

We summarize enabling developments in high-current-density nanorod conductors and the overall design and use of SMES devices, which are emerging as an alternate approach to energy storage that does not require chemical energy technologies.

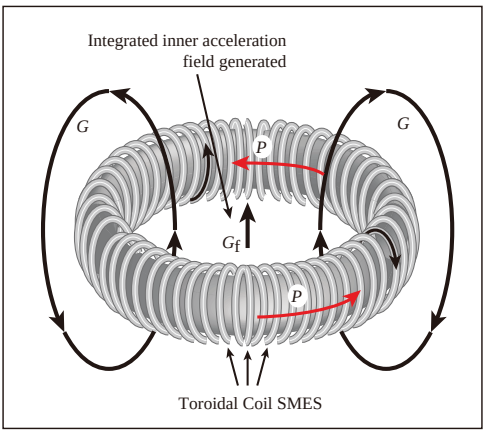



Figure 1. Gravitational force generation coil from Forward 1962 [1] with an inspiraling mass current, with a vector potential $P$, creating gravitomagnetic field $G$, which is additive in the center.

### 2.1. *Superconducting Nanorods*

New developments in superconductor nanomaterial processing [3] led to nanotubular superconductors [2] with a $T_c$ at 92 K. A uniqueness of the nanotubular and other geometric structures of HTSC makes for a practical wire form without using the melt texturing techniques which make for brittle thin films that are also difficult to shape into wire. Another aspect of these new HTSC materials is negation of post oxygenation at high temperatures. The elimination of this requirement makes room temperature forming and application of HTSC materials practical. The process has been demonstrated to be a low-cost and mass production method of superconductors which is scalable and without vacuum or cleanroom requirements. [2] These developments have led to the commercialization by True 2 Materials PTE, LTD (T2M) in Singapore, of a new HTSC wire using standard wire-making practices.

Although the critical temperature of the wire is 92 K, operation at 77 K in liquid nitrogen is more reasonable due to safety issues with gases and nitrogen's inertness, nonexplosive and nonflammable, as a cryogenic liquid. Currently T2M prototype wire is in the millimeter range and approaching the micron range. However, development of an HTSC wire, or filament at nanometer scale is on the roadmap [3] of T2M. The estimated diameter of the wire used in this study, currently theoretical, is 200 nm O.D. including insulation and a 30 nm O.D. HTSC core, with the total weight of the wire at approximately 0.001 $\frac{g}{m}$.

Individual wires make up a 19-nanofilament cable of 1 µm diameter as shown in Fig. 2, illustrating the compactness of the nanocable design. This allows the scaling up of the critical current limit, quenching aside, without adding significant weight to the toroidal coil. The individual nanofilament as described in Fig. 3 consists of a core (A), core sleeve (B), a highly insulated sleeve (C) with good heat-transport properties, and a high-strength gigapascal (GPa) outside sleeve (D), which also possesses

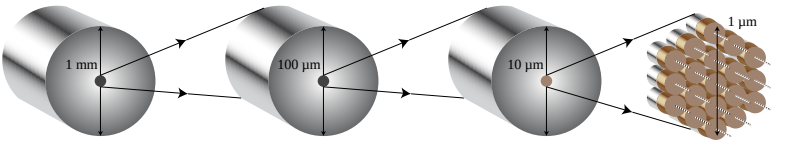



Figure 2. Representation of multiple filaments in a cable where each 1 $\mu$m yields 19 nanofilaments.

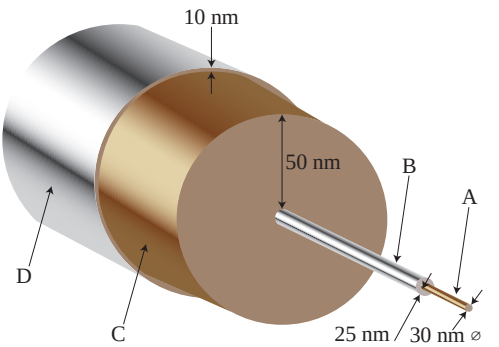



Figure 3. Detailed illustration of a nanowire filament.

good heat-transport properties at low temperatures. The main consideration for a candidate of the materials used in this study would be of carbon composition.

### 2.2. *Superconducting Magnetic Energy Storage*

SMES devices are an emerging battery replacement technology.[4] A typical application is shown in Fig. 4. The device is fed by a DC current, developing a magnetic field, typically in a toroidal geometry coil. When the need for emergency power is detected, an output switch is activated that provides DC current out, which may be converted to AC power by a power inverter.

Given the recent advances in nanowire technology, these devices are poised for remarkable improvements in capability in the very near term. With these coming improvements in this technology, and the similarity in geometry with the Forward design of Fig. 1, the present paper will study this technology at its limits for possible reapplication as a generator of an unbalanced DC gravitational force.

## 3. SMES Mass Current

The linear force, $G_{\mathrm{f}}$, developed by gravitomagnetic force in the mass flow toroid of Fig. 1 is given by Eq. (1):[1]

$$G_{\mathrm{f}} = \frac{\eta}{4\pi} \cdot \frac{N\dot{T}r^2}{R^2}, \tag{1}$$

where $G_{\mathrm{f}}$ = gravitomagnetic force, $\eta$ is gravitomagnetic permeability, $\eta = \eta_{\mathrm{o}}\eta_{\mathrm{r}}$, $\eta_{\mathrm{o}}$ = absolute gravitomagnetic permeability, $\eta_{\mathrm{r}}$ = relative gravitomagnetic permeability,

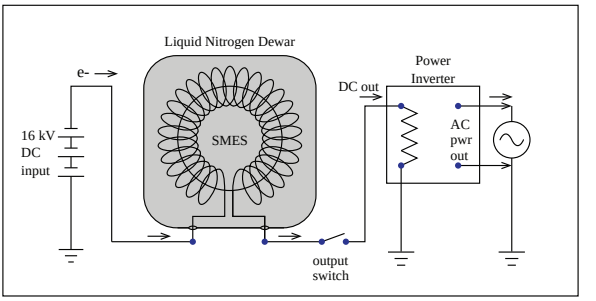



Figure 4. SMES in an energy-storage application.

$N$ = number of turns in the coil of the torus, $\dot{T}$ = change in mass flow, $r$ = the cross-sectional radius of the torus, and $R$ = the centerline radius of the torus.

Change in mass flow for a single electron flow is given by

$$\dot{T}_{\mathrm{e}} = \dot{p}_{\mathrm{e}} = a \times m_{\mathrm{e}} = (\Omega \times v) m_{\mathrm{e}}. \tag{2}$$

This is equivalent to a centripetal force of

$$\dot{T}_{\mathrm{e}} = F_{\mathrm{e}} = m_{\mathrm{e}} \frac{v^2}{r} = m_{\mathrm{e}} a_{\mathrm{e}} = m_{\mathrm{e}} (\omega^2 r), \tag{3}$$

We now attempt to estimate the possible mass currents enabled by emerging nanowire SMES technology as it relates to the core geometry constraints described by a torus geometry. We start with the assumptions needed to calculate the number of turns, $N$.

For the purposes of describing an idealized case with a realistic geometry, we develop a description of a device bounded by a 10 m toroid centerline diameter and with a cross-sectional core diameter of 1 m. We further assume a conductor winding depth of 0.5 m wound about the core. Assuming each nanowire conductor has a diameter $d_{\mathrm{c}} = 100\,\mu\mathrm{m}$, then the cross-sectional area of each conductor will be given by $A_{\mathrm{c}} = \pi r^2 = 7.854 \times 10^{-9}\,\mathrm{m}^2$. For packing the conductors described above to the winding depth described, it can be shown [5] that the total number of windings for the entire toroid can be as high as $N = 1.256 \times 10^9$.

What is $\dot{T}$ with the forgoing assumptions? In this idealized case electrons circulate about a coil of circumference $c_r$, or slightly larger, as described by $c_r = 2\pi r = 3.14\,\mathrm{m}$. Assume further a supply voltage of 16 kV, resulting in 16 keV of kinetic energy for each electron, which corresponds with the upper limit of a nonrelativistic case where, for idealized electron mobility, $v = 0.25c$, so that $\gamma = 1.06 \approx 1.0$. Then, from Eq. (3) for nonrelativistic circular motion, the vector change in DC current flow is

$$\dot{T}_{\mathrm{e}} = \frac{m_{\mathrm{e}} v^2}{r}, \tag{4}$$

which, for a single electron, has the following values: $m_{\mathrm{e}}$ = mass of the electron $= 9.11 \times 10^{-31}\,\mathrm{kg}$, $v$ = velocity of the electron $= 0.25c = 0.75 \times 10^8\,\frac{\mathrm{m}}{\mathrm{s}}$, $r = 0.5\,\mathrm{m}$

for the assumed geometry. The angular acceleration of the electron is

$$a_{\mathrm{e}} = \frac{v^2}{r} = 1.125 \times 10^{16}\,\frac{\mathrm{m}}{\mathrm{s}^2} \tag{5}$$

Thus, change in mass flow represents centripetal acceleration in the case of circular motion:

$$\dot{T}_{\mathrm{e}} = m_{\mathrm{e}} \times a_{\mathrm{e}} = 10.25 \times 10^{-15}\,\mathrm{N}. \tag{6}$$

Equation (6) corresponds to the change in mass flow for one electron in one loop of coil. Total mass-flow change is, therefore, the mass-flow change per electron times the number of electrons:

$$\dot{T} = \dot{T}_{\mathrm{e}} \times N_{\mathrm{e}}. \tag{7}$$

$N_{\mathrm{e}}$ in one loop can be described by the current $I$ times the period of a single-loop circulation $\Delta t$:

$$N_{\mathrm{e}} = I \times \Delta t, \tag{8}$$

where, for idealized mobility, the period of an orbit in a loop can be described by

$$\Delta t = \frac{c_r}{v} = 2\pi\frac{r}{v} = 41.89\,\mathrm{ns}. \tag{9}$$

What is the possible current inside the idealized device for the case where the entire winding is in series? We make the assumption about max current to stay below critical current density of 250 $\frac{\mathrm{MA}}{\mathrm{m}^2}$. Current is limited by the maximum permissible current density and the cross section of the conductor $t$:

$$I = J \times A_{\mathrm{c}} \tag{10}$$

where $J$ is material dependent. For the nanowire assumed in Ref. 2, $J = 250\,\frac{\mathrm{MA}}{\mathrm{m}^2}$. The assumed cross-sectional area $Ac = 7.85 \times 10^{-9}\,\mathrm{m}^2$ yields a maximum current of $I = 1.96\,\mathrm{A}$.

Expanding on Eq. (8), the number of electrons $N_{\mathrm{e}}$ in circulation in one loop may be calculated by noting that there are $6.24 \times 10^{18}$ electrons per Coulomb.

$$N_{\mathrm{e}} = \left(\frac{\mathrm{electrons}}{\mathrm{C}}\right) I \left(\frac{\mathrm{C}}{\mathrm{s}}\right) \times t = 5.12 \times 10^{11}\,\mathrm{electrons} \tag{11}$$

## 4. SMES Forces

Expressing Eq. (7) as force per electron times the number of electrons in motion in one loop:

$$\dot{T} = \dot{T}_{\mathrm{e}} \left(\frac{\mathrm{N}}{\mathrm{electrons}}\right) \times N_{\mathrm{e}}(\mathrm{electrons}) = 5.248\,\mathrm{mN}. \tag{12}$$

Thus, each loop experiences about 5 mN of integrated centripetal force ($\dot{T}$) due to the electrons in circulation within.

We now describe the scale factor to couple the centripetal force to the gravitomagnetic effect. Revisiting Eq. (1), which describes the overall linear force developed at the center of the toroidal coil, total gravitomagnetically developed force will be

$$G_{\mathrm{f}} = \eta_{\mathrm{o}}\eta_{\mathrm{r}} \cdot \frac{N\dot{I}r^2}{4\pi R^2} = \eta_{\mathrm{o}}\eta_{\mathrm{r}} \cdot 5{,}248\,\mathrm{N}. \tag{13}$$

grouping known variables on the right and unknown variables on the left. This raises the question, What are the correct values for $\eta_{\mathrm{o}}$ and $\eta_{\mathrm{r}}$? If $\eta_{\mathrm{o}}$ goes as $\frac{G}{c}$, as does gravitomagnetic potential (Ref. 6, Eq. (1.5)), then

$$\eta_{\mathrm{o}} = -\frac{G}{2c} = 1.11 \times 10^{-19}. \tag{14}$$

In this case $G_{\mathrm{f}} = 5.8 \times 10^{-16}\eta_{\mathrm{r}}$. Values of $\eta_{\mathrm{r}}$ are experimentally unknown at this time. However, if values of $\eta_{\mathrm{r}}$ track values of $\mu_{\mathrm{r}}$, then values as high as $\eta_{\mathrm{r}} = 10^6$ may be possible, yielding $G_{\mathrm{f}} = 5.8 \times 10^{-10}\,\mathrm{N} = 0.58\,\mathrm{nN}$.

Even with very sensitive quantum gravitometer, this would be a very difficult measurement. However, with additional current or winding count, a device scaled up from the idealized case considered in this paper may someday achieve a measurable DC gravitational field, even in the nonrelativistic case considered here.

## 5. Conclusion

An argument is made for using SMES to gravitomagnetically create an unbalanced force, possibly of measurable amplitude. Further research would be required to determine to what extent SMES devices could be operated into relativistic regimes to enhance relative mass flow change in the rest frame thus improving effect detectability.[5]

## Acknowledgments

The authors wish to acknowledge the Nara Institute of Science and Technology for their support of HTSC nanowire technology. The financial support of Seculine Consulting is also gratefully acknowledged.

## Bibliography

1. R. Forward, Guidelines to antigravity, in *Proceedings of the Gravity Research Foundation*, (New Boston, NH, 1962).
2. W. Rieken, R. H. Bhargava, A., J. Akimitsu and H. Daimon, $YBa_2Cu_3O_x$ superconducting nanorods, *Jpn. J. Appl. Phys.* **57**, p. 023101 (2018).
3. True 2 Materials, We are True 2 Materials (2018).
4. Wikipedia, Superconducting magnetic energy storage (2018).
5. G. V. Stephenson, W. Rieken and B. A., Extended cases of laboratory generated gravitomagnetic field measurement devices, *JHEPGC* **5**, 321 (2019).
6. B. Mashhoon, Gravitoelectromagnetism: A brief review (2008).

# Searching for dark matter through magnetized media: The QUAX proposal of a ferromagnetic axion haloscope

Antonello Ortolan*, Augusto Lombardi, Ruggero Pengo and Giuseppe Ruoso

*INFN National Labs of Legnaro, Viale dell'Università, 2 35020, Legnaro, PD, Italy*
**E-mail: ortolan@lnl.infn.it*

Caterina Braggio, Giovanni Carugno, Nicoló Crescini and Sebastiano Gallo

*Dept. of Physics and Astronomy, University of Padova and INFN Sez. di Padova, Via Marzolo, 8 35131, Padova, Italy*

David Alesini, Daniele Di Gioacchino, Claudio Gatti, Carlo Ligi, Alessio Rettaroli and Simone Tocci

*INFN National Laboratories of Frascati, Viale Enrico Fermi 40, I-00044, Frascati, Rome, Italy*

Paolo Falferi

*CNR Istituto di Fotonica e Nanotecnologie, via alla Cascata, 56/C 38100 Povo, TN and INFN TIPFA, Via Sommarive, 14, 38123 Povo, TN*

Renato Mezzena

*Dept. of Physics, University of Trento and INFN TIPFA, Via Sommarive, 14 38123 Povo, TN*

Umberto Gambardella, Gerardo Iannone and Sergio Pagano

*Dept. of Physics, University of Salerno, Via Giovanni Paolo II, 132, 84084 Fisciano SA*

Luca Taffarello

*INFN Sez. di Padova, Via Marzolo, 8 35131, Padova, Italy*

Gianluca Lamanna

*Dept. of Physics, University of Pisa and INFN Sez. di Pisa, L. Pontecorvo, 2 44100 Pisa Italy*

Clive C. Speake

*Dept. of Physics, University of Birningham, Birningha, United Kingdom*

Light stable axions, originally proposed to solve the strong CP problem of quantum chromodynamics (QCD), emerge now as leading candidates of WISP dark matter. The axion-electron coupling, explicitly predicted by some models, can be exploited to envisage novel detectors complementary to the "Sikivie haloscope". In fact, due to the Earth motion with respect the dark matter halo, the interaction of relic axions with electron spins results in an effective magnetic field that inject power in magnetized media. In this proceeding we present the QUAX proposal of a ferromagnetic haloscope and the related ongoing experimental activity at the National Laboratories of Legnaro (Italy). The experimental parameters required to achieve cosmologically relevant sensitivity with our detector will be discussed. Some preliminary results on the operation of the QUAX prototype are eventually presented.

*Keywords*: Axion, Dark Matter.

## 1. Introduction

An impressive result of modern cosmology is that a major fraction of the mass content of the universe is composed of dark matter (DM), i.e. particles not interacting significantly with electromagnetic radiation, ordinary matter, not even self-interacting (cold dark matter). The DM existence is inferred from its gravitational influence on galaxies, clusters and Cosmic Microwave Background, which suggests it dominates by a ratio of 5:1 over ordinary matter described by the Standard Model (SM) of particle physics. In the era of precision cosmology an important threshold has been crossed: the issue nowadays is no longer the existence of particle dark matter, but its nature as new particle(s) Beyond SM (BSM models). Moreover, according to Big Bang cosmology, dark matter particles must be produced thermally similarly to SM particles. Hypothetical Weakly Interacting Massive Particles (WIMPs), e.g. 100 Gev particles from supersymmetric extension of SM, have been a prime DM candidate because of their self-annihilation cross section $\langle \sigma v \rangle \simeq 3 \times 10^{-26} \mathrm{cm}^3 \mathrm{s}^{-1}$, that guarantees the correct abundance of dark matter today ("WIMP miracle")[1]. Recently, the WIMP paradigm has started to look less as the obvious solution to the dark matter problem, and the spotlight passed over Weakly Interacting Sub-eV Particles (WISP), that are very light bosons characterized by sub-eV masses and large occupation numbers, such as axions, axion-like particles (ALP) and hidden photons. In particular, the axion, originally introduced in the SM by Peccei and Quinn to solve the strong CP problem of QCD, is a good candidate for DM. The axion is the pseudo-Goldstone boson associated to an additional symmetry of SM Lagrangian which is spontaneously broken at an extremely high energy scale $F_a$. Current literature favours invisible axion mechanism with $F_a \gg 250$ GeV (weak scale), which has two main implementations: the KSVZ (Kim-Shifman-Vainshtein-Zakharov) models and the DFSZ (Dine-Fischler-Srednicki-Zhitnitsky) models. For scales $F_a \simeq 10^{12}$ GeV, corresponding to typical mass values $m_a \simeq 1$ meV, axions may account for the totality of DM[2]. As a consequence, many different detectors have been proposed over the last decades to search for relic axions. Although theory do not fix the value of $F_a$, cosmological considerations and astrophysical observations provide boundaries on $F_a$ and suggest a favoured axion mass range $1\mu eV < m_a < 10meV$, i.e. over $10^{15}$ times smaller than WIMPs mass. In addition, lattice results on QCD topological susceptibility, based on reliable computations of the axion relic density, indicate a preferred window for the axion mass in the range of tens of $\mu$eV[3]. With such a small mass, we would expect $n_a \simeq 3 \times 10^{13} (10\ \mu eV/m_a)$ axions per cubic centimeter in our Solar System to account for the observed DM density $\rho_{DM} \sim 0.3$ GeV/$cm^3$. Most of axion detectors rely on the conversion of axions into photons in a resonant cavity in the presence of a static magnetic field, following the detection scheme proposed by P. Sikivie in 1983 – called axion haloscope – which is based on the inverse Primakoff effect[4]. In particular, the ADMX experiment reached the cosmologically relevant sensitivity to exclude the axion mass range $2.66 < m_a < 2.81$ $\mu$eV for DFSZ models and $1.9 < m_a < 3.7$ $\mu$eV for KSVZ models[5]. On the other hand, the axion-fermion coupling, explicitly predicted in DFSZ models[6], allows for

designing new detectors that exploit the interaction between axions and fermionic spins[7]. In fact, the motion of the Solar System through the DM cloud surrounding the Galaxy results in a gradient of the axion field $\boldsymbol{\nabla} a$ pointing in the motion direction, and one can demonstrate that the effect of $\boldsymbol{\nabla} a$ on spins in a magnetized material plays the rôle of an effective oscillating magnetic field $\boldsymbol{B}_a$ with amplitude, direction and frequency determined by $\rho_{DM}$, $\boldsymbol{\nabla} a$ and $m_a$, respectively. The QUAX detector is precisely based on the resonant interaction of $\boldsymbol{B}_a$ with the homogeneous magnetization mode in ferrimagnetic samples. In this proceeding we summarize the main ideas of the QUAX proposal[7] and report on some recent results regarding the cryogenic operation of a QUAX demonstrator[8].

## 2. Axion detection by resonant interaction with electron spin

Assuming standard cosmology, we calculate the effective magnetic field $\boldsymbol{B}_a$ with the aim of estimating the power released in a magnetized sample as a function of QUAX parameters. As the DFSZ axion models do not suppress the axion and electron coupling at the tree level, the Lagrangian reads

$$L = \bar{\psi}(x)(i\hbar\gamma^{\mu}\partial_{\mu} - m_e)\psi(x) - ig_{ae}a(x)\bar{\psi}(x)\gamma_5\psi(x), \tag{1}$$

where $\psi(x)$ is the spinor field of an electron with mass $m_e$. Here $\gamma^{\mu}$ are the 4 Dirac matrices, $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, and $a(x)$ is coupled to matter by the dimensionless pseudo-scalar coupling constant $g_{ae}$. By taking the non-relativistic limit of $L$, the resulting interaction term can be written as

$$-\frac{g_{ae}\hbar}{2me}\boldsymbol{\sigma}\cdot\boldsymbol{\nabla} a = -2\frac{e\hbar}{2m_e}\boldsymbol{\sigma}\cdot\left(\frac{g_{ae}}{2e}\right)\boldsymbol{\nabla} a \equiv -2\mu_B\boldsymbol{\sigma}\cdot\boldsymbol{B}_a, \tag{2}$$

which has clearly the form of the interaction between the spin magnetic moment and an effective magnetic field, where $e$ is the electron charge, $\mu_B$ is the Bohr magneton, $\boldsymbol{\sigma}$ is the vector Pauli matrices, and $\boldsymbol{B}_a \equiv g_{ae}/(2e)\boldsymbol{\nabla} a$ is the effective magnetic field. For DFSZ axions, the coupling to electrons is $g_{ae} = m_e/(3F_a)\cos^2\beta \simeq 2.8 \times 10^{-15}(m_a/10^{-4}eV)$ assuming O(1) value for the free parameter $\cos^2\beta$. To calculate amplitude and frequency of the effective magnetic field due to the presence of the DM axion we use the standard model of galactic halo: a spherical distribution with a pseudo-isothermal density profile, local density $\rho_{DM} \sim 0.3GeV/cm^3$ corresponding to $n_a \sim 3 \times 10^{12}$ ($10^{-4}$ eV/$m_a$) axions per cubic centimeter, and local velocity distribution described by the Maxwell Boltzmann distribution with a dispersion $\sigma_v \approx 270$ km/sec. The Earth velocity with respect the galactic halo is $|\boldsymbol{v}_E| \simeq 220$ Km/sec. Therefore the equivalent oscillating rf field has an expected mean amplitude and central frequency

$$B_a = 2.0 \times 10^{-22}\left(\frac{m_a}{200\,\mu\text{eV}}\right)\ \text{T}, \quad \frac{\omega_a}{2\pi} = 48\left(\frac{m_a}{200\,\mu\text{eV}}\right)\ \text{GHz}, \tag{3}$$

with relative linewidth $\Delta\omega_a/\omega_a \simeq 5.2 \times 10^{-7}$ and direction $\boldsymbol{v}_E$. The explicit dependence on $\omega$ of the power spectrum of axion field has been calculated in Ref. 6.

Once the calculation is tailored for the gradient of the axion field, the coherence time and correlation length for the QUAX detector read

$$\tau_{\nabla a} \simeq 0.68\,\tau_a = 17\left(\frac{200\,\mu\text{eV}}{m_a}\right)\left(\frac{Q_a}{1.9\times 10^6}\right)\ \ \mu\text{s};$$
$$\lambda_{\nabla a} \simeq 0.74\,\lambda_a = 5.1\left(\frac{200\,\mu\text{eV}}{m_a}\right)\ \text{m}, \tag{4}$$

where the standard DM halo parameters were used[7].

### 2.1. *The QUAX proposal*

In the QUAX proposal we focused our analysis to frequencies in the $40 \div 50$ GHz range. To measure the power released by extremely small rf field $B_a$ we make use of the Electron Spin Resonance (ESR) in magnetic samples[7]. To enhance the interaction, the ferromagnetic resonance of the sample – i.e. the Larmor frequency $\omega_L = \gamma B_0$ of the electron spin precession in an external magnetic field $B_0$ – is tuned to the frequency $\omega_a$ associated with mass value of the searched for axion. Here $\gamma/2\pi = e/m_e = 28GHz/T$ is the electron gyromagnetic ratio. In the limit of weak rf field, the steady state solutions of Bloch Equations of the magnetization $M_a(t)$ relative to the ESR Kittel mode reads

$$M_a(t) = \gamma\mu_B B_a n_S \tau_{\min}\cos(\omega_a t), \tag{5}$$

where $n_S$ is the material spin density and $\tau_{\min} \equiv \min(\tau_{\nabla a}, \tau_2, \tau_r)$ is the shortest coherence time among the axion wind coherence $\tau_{\nabla a}$, magnetic material relaxation $\tau_2$ and radiation damping $\tau_r$, and it represent the coherence time of $M_a(t)$ oscillatiions. However, the radiation damping mechanism may result in an issue for the QUAX detector. In fact, in free space, the energy damping due to magnetic dipole emission $\tau_r = (c^3/\omega_L^3)/(\gamma\mu_0 M_0 V_s)$, and so high Larmor frequency, large magnetization $M_0$, and large volume $V_s$ of magnetized material imply $\tau_r << \tau_a$. However, the radiation damping mechanism can be inhibited by inserting the magnetized material inside a microwave resonant cavity in the strong coupling regime. In this case, the hybridization between Kittel magnetic mode of the sample and a suitable cavity mode occurs, and the limited phase space of the resonant cavity inhibits the radiation damping mechanism, thus providing a more favourable damping time equal to the cavity decay time. In the hybridization regime we have $\tau_{\min} = \min(\tau_{\nabla a}, \tau_2, \tau_c)$, where $\tau_c$ is the cavity decay time[7].

In the presence of the axion wind, the average amount of power absorbed by the magnetized sample in each cycle is

$$P_{\text{in}} = \mu_0 \mathbf{H}\cdot\frac{d\mathbf{M}}{dt} = B_a\frac{dM_a}{dt}V_s$$
$$= \gamma\mu_B n_S \omega_a B_a^2 \tau_{\min} V_s. \tag{6}$$

In a steady state condition, the power balance ensures that $P_{\text{in}}$ will be emitted as rf radiation, and so $P_{\text{in}}/2$ can be collected by using an antenna critically coupled to

the hybridized mode. The output power can be expressed in terms of the relevant experimental design parameters

$$P_{\rm out} = \frac{P_{\rm in}}{2} = 3.8 \times 10^{-26} \left(\frac{m_a}{200\,\mu{\rm eV}}\right)^3 \left(\frac{V_s}{100\ {\rm cm}^3}\right) \left(\frac{n_S}{2 \cdot 10^{28}/{\rm m}^3}\right) \left(\frac{\tau_{\rm min}}{2\,\mu{\rm s}}\right) {\rm W}, \tag{7}$$

where the chosen axion mass is determined by a magnetizing field $B_0 = 1.7$ T, and the value of the spin density is typical of paramagnets at low temperature or material as YIG (Yttrium Iron Garnet) even at room temperature. To measure such a low power is a very difficult experimental challenge. To reach the QUAX goal of detecting cosmological QCD axions, we need to improve the detector sensitivity along the following lines:

1) Study of materials: decrease linewidth of ferromagnetic resonance in the $10 \div 100$ GHz frequency range at low temperatures and increase spin density of magnetized materials (YIG, GaYIG, LiF or BDPA and other paramagnets); make use of ultrapure material (e.g. by trying to minimize rare earths contamination); highly accurate polishing and smoothing of surfaces.
2) Cavity design: design of a high-Q ($\sim 10^6$) cavity to be operated in few Tesla magnetic fields; cavity design should also maximize SNR and allow for housing the required amount of magnetized material.
3) Static magnetic field source: realization of a highly uniform magnetic field (up to 10 ppm for a few Tesla field to avoid inhomogeneous broadening).
4) Decrease of noise level: use a dilution refrigerator to lower thermal noise that also allow for the operation of a Josephson Parametric Amplifier (JPA); a crucial issue is the concurring development of a single photon counter in the microwave frequency range in order to overcome the Standard Quantum Limit of linear amplifiers.

It is worth noticing that only the component of the equivalent magnetic field $\boldsymbol{B_a}$ orthogonal to the magnetizing field $\boldsymbol{B_0}$ will drive the magnetization of the sample, therefore QUAX is a directional detector and it shows a daily modulation of the axion signal that can be exploited to get rid of spurious noise sources[7].

### 2.2. *The QUAX demonstrator*

We have addressed experimentally some issues that affect the sensitivity of a ferromagnetic haloscope, such as ferrimagnetic dissipation, cavity quality factor, magnetic field homogeneity[8]. Thus we have set up a demonstrator of the QUAX experiment which is made of 5 GaYIG (Gallium Yttrium Iron Garnet), 1 mm diameter spheres, placed in a cylindrical copper cavity (diameter $\sim$ 26 mm and length 50 mm), and immersed in a $\sim$ 0.5 T magnetic field. We use the TM110 mode with resonance frequency $f_c \simeq 13.98$ GHz and linewidth $k_c/2\pi \simeq 400$ kHz at liquid helium temperature. The degeneracy of this mode has been removed by digging

two symmetric grooves in the lateral surface of the cylinder. The TM110 mode is characterized by a uniform maximum magnetic rf field along the cavity axis, and therefore we can house more YIG spheres along it. Moreover, cavity volume can be increased by lengthening the cylinder without changing the TM110 mode resonance frequency. With cylindrical geometry we can also exploit the uniformity of magnetic field produced by a solenoid along the symmetry axis so as to avoid the inhomogeneous broadening of the resonances. In fact, the Larmor frequency $f_L$ of the GaYIG spheres is established by an highly uniform solenoidal magnetic field (1 part in $10^4$), and the hybridization condition $f_L \simeq f_c$ is met when $B_0 \simeq 0.5$ T. In the strong coupling regime, the hybrid mode frequencies are $f_+ = 14.061$ GHz and $f_- = 13.903$ GHz. The detection electronics (a cascade of a room temperature and a cryogenic amplifier) was calibrated with a Johnson noise source at different temperature. The output of the cavity is down-converted in its in-phase and quadrature components with respect to a local oscillator and sampled at 2 MHz. The power of hybridized modes was estimated with $\sim 2 \times 10^6$ FFTs of 8192 bins each (frequency resolution of 244 Hz), which were square averaged and rebinned into 7.8 kHz bandwidths (256 bins), close to the expected axion bandwidth. The standard deviation of the estimated power is $\sigma_P = (2.2 \pm 0.1) \times 10^{-22}$ W, compatible with the Dicke radiometer equation. No significant excess power consistent with DM axions was found. Therefore we can set the upper limit $\sigma'_P \sim 10^{-21}$ W within the 3 MHz band around 13.903 GHz at the 95% C.L.. This value translates to the upper limit $B_a \leq 1.6 \times 10^{-17}$ T of the equivalent axion field, corresponding to a axion electron coupling $g_{ae} \leq 5 \times 10^{-10}$ for axion masses $58.527 \leq m_a \leq 58.541$ $\mu$eV at 95% CL[8].

## 3. Conclusions

We have reported on the QUAX proposal, i.e. a ferromagnetic haloscope sensitive to DM axions through their interaction with electron spin. Our findings by means of the QUAX demonstrator indicate the possibility of performing ESR measurements of a sizable quantity of material inside a cavity cooled down to cryogenic temperatures. The sensitivity of the demonstrator reached the limit of the Dicke radiometer equation, the overall behavior of the apparatus is as expected, and we are confident that the QUAX detector can reach a cosmological relevant sensitivity.

## References

1. See e.g. G. Jungman et al, Physics Reports **267**, 5 (1996).
2. J. E. Kim and G. Carosi, Rev. Mod. Phys. **82**, 557 (2010).
3. S. Borsanyi et al, Nature **539**, 69 (2016).
4. P. Sikivie, Phys. Rev. Lett. **51**, 1415 (1983); *ibid.* Phys. Rev. **D 32**, 2988 (1985).
5. N. Du et al. Phys. Rev. Lett. **120**, 151301 (2018).
6. L.M. Krauss et al, Phys. Rev. Lett **55**, 1797 (1985).
7. R. Barbieri et al, Physics of the Dark Universe **15**, 135 (2017).
8. N. Crescini at al, Eur. Phys. J. **C 78**, 703 (2018).

# What could the value of the cosmological constant tell us about the future variation of the fine structure constant?

Antonio Enea Romano

[1] *YITP, Kyoto University, Kyoto 606-8502, Japan;*
[2] *Instituto de Fisica, Universidad de Antioquia, A.A.1226, Medellin, Colombia*

Motivated by reported claims of the measurements of a variation of the fine structure constant $\alpha$ we consider a theory where the electric charge, and consequently $\alpha$, is not a constant but depends on the Ricci scalar $R$. We then study the cosmological implications of this theory, considering in particular the effects of dark energy and of a cosmological constant on the evolution of $\alpha$. Some low-red shift expressions for the variation of $\alpha(z)$ are derived, showing the effects of the equation of state of dark energy on $\alpha$ and observing how future measurements of the variation of the fine structure constant could be used to determine indirectly the equation of state of dark energy and test this theory. In the case of a $\Lambda CDM$ Universe, according to the current estimations of the cosmological parameters, the present value of the Ricci scalar is $\approx 10\%$ smaller than its future asymptotic value determined by the value of the cosmological constant, setting also a bound on the future asymptotic value of $\alpha$.

## 1. Introduction

There have been different claims about the detection of the variation[1–5] of the fine structure constant. Different models have been proposed to account for such a phenomenon, in particular theories where the electric charge is treated as free field[6]. We consider the cosmological implications of a model in which the electric charge depends on the local value of the space-time curvature. This model can be considered a special case of the Bekenstein theory[6], where the electric charge depends on the Ricci scalar, and consequently it is not necessary to introduce an additional kinetic term in the Lagrangian to determine the dynamics of the field associated to the electric charge, as in the standard Bekenstein case. In this sense this model has the advantage of not introducing a new fully independent degree of freedom and can consequently give more definite predictions. At the same time the hypothesis that the electric charge depends only on the Ricci scalar is in agreement with the principle of general covariance and as such could be extended to the study of the variation of other natural constants without the need of introducing any new degree of freedom.

On the basis of general covariance it is natural to consider as an extension of the standard electromagnetism an action of the type

$$S = \int \sqrt{-g}(f(R)F_{\mu\nu}F^{\mu\nu} + e_0 u_\mu A^\mu)\,, \tag{1}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the Faraday tensor and $A\mu$ is the vector potential, as defined in the standard covariant formulation of electromagnetism.

The main difference respect to the classical formulation of electromagnetism corresponds to the fact that such an action implies a space time variation of the

fine structure constant $\alpha$. This can be seen easily by varying the action respect to the vector potential $A^\mu$, and after neglecting derivative terms of quantities varying only on very large scales we get

$$\partial_\mu F^{\mu\nu} = f(R)^{-1} e_0 u^\nu \,, \tag{2}$$

which shows how the effective electric charge appearing on the r.h.s. is not constant. The variation of the action respect to the metric is not expected to modify the standard Einstein's equations because the term $F_{\mu\nu}F^{\mu\nu} \propto (E^2 - B^2)$ should give an average vanishing contribution in the case of a homogenous electromagnetic field. The above action can be considered a particular case of the Bekenstein theory:

$$S_B = \int \sqrt{-g}(\epsilon^{-2} F_{\mu\nu}F^{\mu\nu} + e(x_\mu)\, u_\mu \tilde{A}^\mu) = \int \sqrt{-g}(\epsilon^{-2} F_{\mu\nu}F^{\mu\nu} + e_0 u_\mu A^\mu)\,, \tag{3}$$

$$A_\mu = \epsilon \tilde{A}_\mu \,, \tag{4}$$

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \,, \tag{5}$$

where the electric charge appearing in the Lagrangian is treated as a free field

$$e(x_\mu) = e_0 \epsilon(x^\mu) \,, \tag{6}$$

and the action (1) corresponds to the case in which

$$\epsilon(x^\mu) = f(R)^{-1/2} \,, \tag{7}$$

$$\alpha(z) = \alpha_0 f(R)^{-1} \,. \tag{8}$$

In the Bekenstein theory an additional kinetic term is introduced in the Lagrangian in order to determine the dynamics of $\epsilon$, while for the action S (1) this is not necessary since $f(R)$ entirely determines the variation of the fine structure constant. From a physical point of view this corresponds to assuming that the fine structure constant value depends only on the Ricci scalar, preserving the invariance under general coordinate transformations.

The Ricci scalar can be related to the trace of the energy momentum tensor by taking the trace of the Einstein tensor:

$$R = -T^\mu_\mu \,, \tag{9}$$

For a set of perfect fluids with equation of state

$$P_i = w_i \rho_i \,, \tag{10}$$

we get

$$R = \sum_i (3w_i - 1)\rho_i \,, \tag{11}$$

where the index i stands for the i-th component of the total energy density of the Universe. It is interesting to note that, according to the above equation, radiation does not contribute to the Ricci scalar at any time, since it vanishes for $w_i = 1/3$.

We can now derive a low red-shift expansion of the variation of the fine structure constant. Even if this would not be valid at higher re-shift, it can still give some useful insight about the features of the theory we are studying. In order to get the leading corrections to the value of $\alpha$ we expand the equation of state of dark energy $w(z)$ and the function $f(R)$ according to:

$$\Delta\alpha = \frac{\alpha(z) - \alpha_0}{\alpha_0} = \Delta\alpha_1 z + \Delta\alpha_2 z^2 + \cdots , \tag{12}$$

$$w(z) = w_0 + w_1 z + w_2 z^2 + \cdots , \tag{13}$$

$$f(R) = 1 + f_1 \left( \frac{R - R_0}{R_0} \right) + f_2 \left( \frac{R - R_0}{R_0} \right)^2 + \cdots , \tag{14}$$

where

$$R_0 = R(z = 0) . \tag{15}$$

is the present time value of the Ricci scalar, The above expansion for $f(R)$ is clearly dimensionally consistent and satisfy the normalization condition

$$\alpha(z = 0) = \alpha_0 . \tag{16}$$

It should also be noted that such a local expansion is only valid at relatively recent times otherwise, at early times, when $(R - R_0)/R_0 \to \infty$, $\alpha$ would diverge. Nevertheless it is a good approximation to estimate the future variation of $\alpha$ because $(R_\infty - R_0)/R_0 \approx -0.1$, as it will be shown later.

Assuming a flat Universe, after substituting in Eq. (11) the contributions from matter and dark energy

$$\rho_X(z) = 3H_0^2 \Omega_X \exp\left[ \int_0^z \frac{3[1 + w(x)]}{1 + x}\, dx \right] , \tag{17}$$

$$\rho_M(z) = 3H_0^2 (1 + z)^3 , \tag{18}$$

we get the following expression for the Ricci scalar as a function of the red-shift

$$R(z) = -3(1 + z)^3 H_0^2 \Omega_M + 3H_0^2 \Omega_X \exp\left[ \int_0^z \frac{3[1 + w(x)]}{1 + x}\, dx \right] [3w(z) - 1] . \tag{19}$$

The above expression is also valid at early times when the Universe was radiation dominated since, as noted previously, radiation does not contribute at any time to the Ricci scalar.

We can then compute the leading corrections to the value of the fine structure constant as a function of the red-shift:

$$\Delta\alpha_1 = \frac{3f_1(-1+2w_0\Omega_X+3w_0^2\Omega_X+w_1\Omega_X)}{3w_0\Omega_X-1}\,, \tag{20}$$

$$\Delta\alpha_2 = \frac{1}{2(1-3w_0\Omega_X)^2}\Bigg[3\Big((6f_1^2-3f_2)(-1+2w_0\Omega_X+3w_0^2\Omega_X+w_1\Omega_X)^2+\,, \tag{21}$$

$$-f_1(1-3w_0\Omega_X)(2-(-2+24w_0^2+18w_0^3+4w_1+2w_0(1+6w_1)+w_2)\Omega_X)\Big)\Bigg],$$

where we have used the relation

$$\Omega_\Lambda+\Omega_X=1\,, \tag{22}$$

following from the Friedman equation evaluated at present time. As it can be seen the linear contribution to the variation of the fine structure constant value is affected by the central value and first order derivative of the equation of state of dark energy and by the first derivative of function $f(R)$. Conversely the above expression shows how measurements of the variation of the fine structure constant could be used to determine indirectly the equation of state of dark energy.

Another important case to consider is that of a $\Lambda CDM$ Universe, for which the Ricci scalar ass a function of the red-shift is given by

$$R(z) = -\rho_M^0(1+z)^3-4\rho_\Lambda = -3H_0^2(\Omega_M^0(1+z)^3+4\Omega_\Lambda)\,, \tag{23}$$

which implies that at late times $R(z)$ tends to a constant value determined by the cosmological constant, while at earlier times it decreases.

We can then obtain a local expansion for $\alpha(z)$ as in the dark energy case:

$$\Delta\alpha_1 = \frac{3f_1(-1+\Omega_\Lambda)}{1+3\Omega_\Lambda}\,, \tag{24}$$

$$\Delta\alpha_2 = \frac{3(-1+\Omega_\Lambda)(6f_1^2(-1+\Omega_\Lambda)-3f_2(-1+\Omega_\Lambda)+f_1(2+6\Omega_\Lambda))}{2(1+3\Omega_\Lambda)^2}\,. \tag{25}$$

It can be easily checked that these last formulae are in agreement with the dark energy case when $\{w_n=0, w_0=-1\}$. In this case the variation of $\alpha$ is completely determined by the coefficients $\{f_1, f_2\}$ and the value of the cosmological constant.

Given the current estimation of cosmological parameters[7–9] the Ricci scalar is tending to an asymptotic value determined by the cosmological constant

$$\frac{R_\infty}{R_0} = \frac{4\Omega_\Lambda}{1+3\Omega_\Lambda} \approx 0.9\,, \tag{26}$$

and the present value is quite closed to the asymptotic one as it can be seen in Fig. 1. This shows how the Ricci scalar is not expected to vary largely in the future history of the Universe, and if the theory we are studying is correct, also the fine

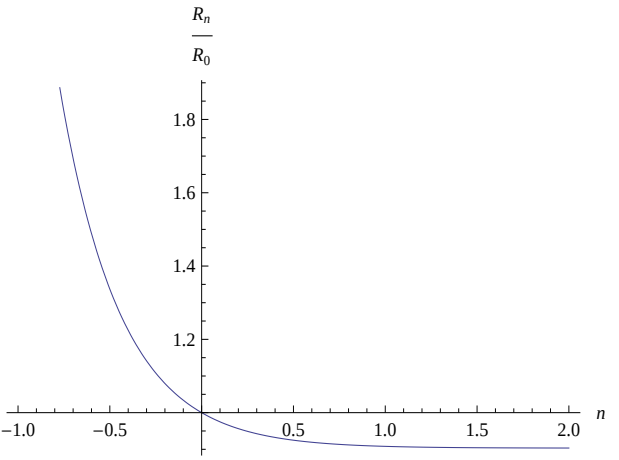



Fig. 1. The ratio $R_n/R_0$ is plotted as a function of the number $n$ of e-folds from today, where $R_n = -3H_0^2(\Omega_M^0(\frac{a_0}{a})^3 + 4\Omega_\Lambda)$, $a = a_0 e^n$ and $R_0$ is the value of the Ricci scalar today, when $n = 0$. As it can be seen the present value $R_0$ is quite closed to the asymptotic future value of R.

structure constant will reach a corresponding asymptotic value given by

$$\alpha_\infty \approx \alpha_0 f(R_\infty)^{-1} \approx \alpha_0[1 + f_1(-0.1) + f_2(-0.1)^2]^{-1} \approx \alpha_0[1 + f_1(0.1) - f_2(0.1)^2]\,. \tag{27}$$

According to this prediction the so called coincidence problem, i.e. the fact that we happen to live just around the time of transition between a matter dominated and a cosmological constant dominated stage of the evolution of the Universe, would also have implications on the future variation of the fine structure constant, which would be expected to tend to an asymptotic value not too different from the present one. After the coefficients $f_1, f_2$ have been determined by fitting available observational data of $\Delta\alpha(z)$ we could get a good estimation of $\alpha_\infty$, but we would have to wait a time interval of the order of the Hubble time $H_0^{-1}$ before the matter contribution to the Ricci scalar will be dominated by that of the cosmological constant and we could test such a prediction.

**Acknowledgements**

I thank M. Sasaki for insightful discussions and financial support for the participation to the YITP Gravity and Cosmology 2012 workshop.

**References**

1. M. T. Murphy, J. K. Webb, V. V. Flambaum and ,
2. J. A. King, J. K. Webb, M. T. Murphy, V. V. Flambaum, R. F. Carswell, M. B. Bainbridge, M. R. Wilczynska, F. E. Koch and ,

3. M. T. Murphy, J. K. Webb, V. V. Flambaum and , Mon. Not. Roy. Astron. Soc. **384**, 1053 (2008) [astro-ph/0612407].
4. P. Tzanavaris, M. T. Murphy, J. K. Webb, V. V. Flambaum, S. J. Curran and , Mon. Not. Roy. Astron. Soc. **374**, 634 (2007) [astro-ph/0610326].
5. J. -P. Uzan, Living Rev. Rel. **14**, 2 (2011) [arXiv:1009.5514 [astro-ph.CO]].
6. J. D. Bekenstein,
7. P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1303.5076 [astro-ph.CO].
8. P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1303.5062 [astro-ph.CO].
9. P. A. R. Ade *et al.* [Planck Collaboration], arXiv:1303.5081 [astro-ph.CO].

# Quintessence evolution of fundamental constants and cosmological parameters using the beta function formalism

Rodger I. Thompson

*Steward Observatory and Department of Astronomy, University of Arizona, Tucson, Arizona 85721, USA*
**E-mail: rit@email.arizona.edu*

A key lesson from the multi-messenger observations of colliding neutron stars is the importance of having accurate templates to compare with observations. This manuscript demonstrates the production of such templates for the evolution of cosmological parameters and fundamental constants using the beta function formalism. The beta function in this formalism is the derivative of the scalar field with respect to the natural log of the scale factor. This demonstration utilizes the quintessence cosmology with an inverse power law dark energy potential as an example but the formalism is expandable to other dark energy potentials and cosmologies. The advantage of the beta function formalism is that it produces analytic solutions of the evolutions as a function of the scale factor which is a fundamental observable quantity.

*Keywords*: Cosmology, Fundamental Constants.

## 1. Introduction

This article presents beta function formalism calculations of the evolution of fundamental constants and cosmological parameters in a quintessence cosmology for comparison with observational constraints. Quintessence is a well studied example of a dynamical cosmology where the dark energy equation of state is changing with time. The predicted evolution provides tests to determine whether dark energy is static or dynamically evolving. Interestingly the static ΛCDM cosmology is a special case of the dynamical quintessence cosmology, partially demonstrating the "Universality" of the beta function methodology. The epochs of interest are redshifts between 0 and 9 which correspond to scale factors $a$ between 1.0 and 0.1. These epochs are well covered by the upcoming JWST and Euclid space missions as well as the DESI and LSST ground based surveys.

## 2. The Potential and Related Parameters

For specificity a dark energy potential of the common form

$$V(\phi) = M^{4+p}\phi^{-p} \tag{1}$$

is considered here where $M$ has the units of the reduced Planck mass $M_p$. The potential $V(\phi)$ has units of $M_p^4$ and by the continuity equation

$$V(\phi) = 3\Omega_\phi \left(\frac{H(\phi)}{\kappa}\right)^2 \tag{2}$$

where $H(\phi)$ is the Hubble parameter, $\Omega\phi$ is the ratio of the dark energy density to the critical density and $\kappa = \frac{1}{M_p}$. This relation indicates that the value of $V_0$ is set by $H_0$ and $\Omega_{\phi_0}$ where the 0 subscript indicates the current value.

## 3. The Beta Function Formalism

The beta function formalism was recently developed to explore the universality of dynamical cosmologies thought previously to be independent of each other. The beta function formalism has its roots in a perceived correspondence between cosmological inflation and the Quantum Field Theory renormalization group flow equation (1), (2) and( 3). In that context it is valid as the solution for the slow evolution of a system approaching or leaving a critical (fixed) point (1). Both (1) and (2) have considered the formalism for the late time dark energy inflation where the critical point is in the infinite future. The descriptions here follow these references with particular dependence on (2) who have incorporated matter as well as dark energy in order to describe a real universe.

The beta function is defined as the derivative of the scalar field $\phi$ with respect to the natural logarithm of the scale factor of the universe $\ln(a)$.

$$\beta(\phi) \equiv \frac{\kappa d\phi}{d\ln(a)} = \kappa\phi' \tag{3}$$

where the prime in the right hand side denotes the derivative with respect to $\ln(a)$, Since $\phi$ has units of mass the product $\kappa\phi$ is dimensionless and $\phi$ is specified in units of reduced Planck masses. The reduced Planck mass is generally used in the cosmological literature since the factor of $8\pi$ simplifies several equations. Proper units are an important aspect of this article therefore $\kappa$ is retained throughout the article rather than being set to one as is common practice. Natural units are used where $G$, $c$ and $\hbar$ are set to one.

Specifying $\beta(\phi)$ provides a method to express the evolution of cosmological parameters and fundamental constants in terms of the observable scale factor $a$ rather than the unobservable scalar field $\phi$. References (4) and (5) give concrete examples of this process for quintessence. Where as the main purpose of (1), (2) and (3) was to demonstrate the universality of cosmologies associated with a given beta function the purpose of (4), (5) and this article is to calculate cosmological parameters for quintessence for a given dark energy potential $V(\phi)$. Here the beta function is determined by $V(\phi)$ rather than being prespecified.

Rather than being an arbitrary function of the scalar $\phi$, the beta function is determined by the potential, (1) and (2), such that

$$V_m(\phi) = V_0 \exp\left\{-\kappa \int \beta(\phi) d\phi\right\} \tag{4}$$

where $V_m(\phi)$ is the model potential given by Eq. (1). Equation (4) shows that $\beta(\phi)$ is the negative of the logarithmic derivative of the model potential.

The potential in Eq. (1) is an inverse power law, IPL, potential. The beta formalism for IPL potentials was considered in (4) which showed that the beta function for this type of potential is given by the simple form.

$$\beta(\phi) = \frac{p}{\kappa\phi} = \frac{\kappa d\phi}{d\ln(a)} \tag{5}$$

Integrating Eq. (5) yields

$$\kappa\phi = \sqrt{2p\ln(a) + (\kappa_0)^2} \tag{6}$$

where $\phi_0$ is the current value of the scalar which depends on the current values of cosmological parameters and the cosmology as shown in section 4.

## 4. Quintessence

The action for quintessence is given by

$$S = \int d^4x\sqrt{-g}\left[\frac{R}{2} - \frac{1}{2}g^{\mu\nu}\partial_\mu\partial_\nu\phi - V(\kappa\phi)\right] + S_m \tag{7}$$

where $R$ is the Ricci scalar, $g$ is the determinant of the metric $g^{\mu\nu}$, $V(\phi)$ is the dark energy potential, and, $S_m$ is the action of the matter fluid. The dark energy potential determines the form of quintessence being studied. The standard quintessence dark energy density, $\rho_\phi$, and pressure, $P_\phi$ equations are

$$\rho_\phi \equiv \frac{\dot{\phi}^2}{2} + V(\phi), \qquad P_\phi \equiv \frac{\dot{\phi}^2}{2} - V(\phi) \tag{8}$$

and the dark energy equation of state $w(\phi)$ is

$$w(\phi) = \frac{P_\phi}{\rho_\phi} = \frac{\frac{\dot{\phi}^2}{2} - V(\phi)}{\frac{\dot{\phi}^2}{2} + V(\phi)}. \tag{9}$$

Combining Eq. (8) gives

$$P_\phi + \rho_\phi = \dot{\phi}^2 \tag{10}$$

It follows that

$$\frac{P_\phi + \rho_\phi}{\rho_\phi} = w + 1 = \frac{\dot{\phi}^2}{\rho_\phi}. \tag{11}$$

From the definition of $H = \frac{\dot{a}}{a}$

$$\dot{\phi} = \frac{d\phi}{da}a\frac{da}{dt}\frac{1}{a} = \frac{d\phi}{d\ln(a)}H = \phi' H. \tag{12}$$

Using the following relations

$$\rho_\phi = \frac{3\Omega_\phi H^2}{\kappa^2}, \qquad \frac{\dot{\phi}^2}{\rho_\phi} = \frac{(\phi' H)^2}{\rho_\phi}, \tag{13}$$

a new relationship for $(w+1)$ is established.

$$\frac{\dot{\phi}^2}{\rho_\phi} = \frac{(\phi'\kappa)^2 H^2}{3\Omega_\phi H^2} = \frac{(\kappa\phi')^2}{3\Omega_\phi} = (w+1) \tag{14}$$

From Eq. (14)

$$\kappa\phi' = \sqrt{3\Omega_\phi(w+1)}. \tag{15}$$

From Eq. (3) $\beta(\phi) = \kappa\phi'$ coupled with Eq. (5) yields

$$\frac{p}{\kappa\phi} = \sqrt{3\Omega_\phi(w+1)}, \qquad \kappa\phi = \frac{p}{\sqrt{3\Omega_\phi(w+1)}} \tag{16}$$

Equation (16) establishes an important boundary condition on the scalar in terms of the current value of $\omega_\phi$ and $(w+1)$ for Quintessence IPL dark energy potentials.

$$\kappa\phi_0 = \frac{p}{\sqrt{3\Omega_{\phi_0}(w_0+1)}} \tag{17}$$

Equation (17) and Eq. (6) give the evolution of $\phi$ as a function of the scale factor $a$.

## 5. Parameter Relationships

A proper exploration of the parameter space of a cosmololgy such as quintessence requires an understanding of the relationships between the parameters. This section establishes the relations between the scalar $\phi$, the Hubble parameter $H$, the dark energy equation of state $w$ and the value of $M$ in Eq. (1). Combining Eq. (1) and Eq. (2) gives

$$M^{4+p}\phi^{-p} = 3\Omega_\phi \left(\frac{H}{\kappa}\right)^2. \tag{18}$$

Setting the current values of $H$, $\Omega$ and $w$ provides the boundary conditions for the remaining parameters. For a given value of the power $p$ in Eq. (1) Eq. (17) determines the value of $M$ as

$$M = \left[3\Omega_{\phi_0}\left(\frac{H_0}{\kappa}\right)^2 \phi_0^p\right]^{\frac{1}{4+p}}. \tag{19}$$

Equation (19) indicates that once $p$, $H_0$, $\Omega_{\phi_0}$ and $w_0$ are chosen the value of $M$ is set.

## 6. The Dark Energy Potentials

The beta function formalism utilizes a superpotential defined (2) simply as

$$W(\phi) = -2H(\phi) = -2\frac{\dot{a}}{a}. \tag{20}$$

It is obvious that calculation of the superpotential $W(\phi)$ is equivalent to the calculation of the Hubble parameter $H(\phi)$. In the absence of matter both the dark energy superpotential and the potential are defined by the beta function (2).

$$W(\kappa\phi) = W_0 \exp\left\{-\frac{1}{2}\int_{\kappa\phi_0}^{\kappa\phi} \beta(x)dx\right\} \tag{21}$$

and

$$V(\kappa\phi) = \frac{3}{4}W_0^2 \exp\left\{-\int_{\kappa\phi_0}^{\kappa\phi} \beta(x)dx\right\}\left(1 - \frac{\beta^2(\kappa\phi)}{6}\right) \tag{22}$$

where $W_0$ is the current value of $W$ equal to $-2H_0$. Equation (22) is the beta formalism equivalent of Eq. (4) and is called the beta potential $V_\beta(\phi)$ of the model potential $V_m(\phi)$. The beta potential is the model potential multiplied by $(1-\frac{\beta^2(\kappa\phi)}{6})$. If $\frac{\beta^2(\kappa\phi)}{6} << 1$ then the beta potential is an accurate but not exact representation of the model potential. In general the calculated evolution of the cosmological parameters and fundamental constants has the same accuracy as the beta potential representation of the model potential. Figure 1 shows the fractional deviation of the beta potential from the model potential for a sample case with a quartic potential, $p = 4$, $H_0 = 70$ for a flat Universe with $\Omega_\phi = 0.7$. It is clear that the error is less

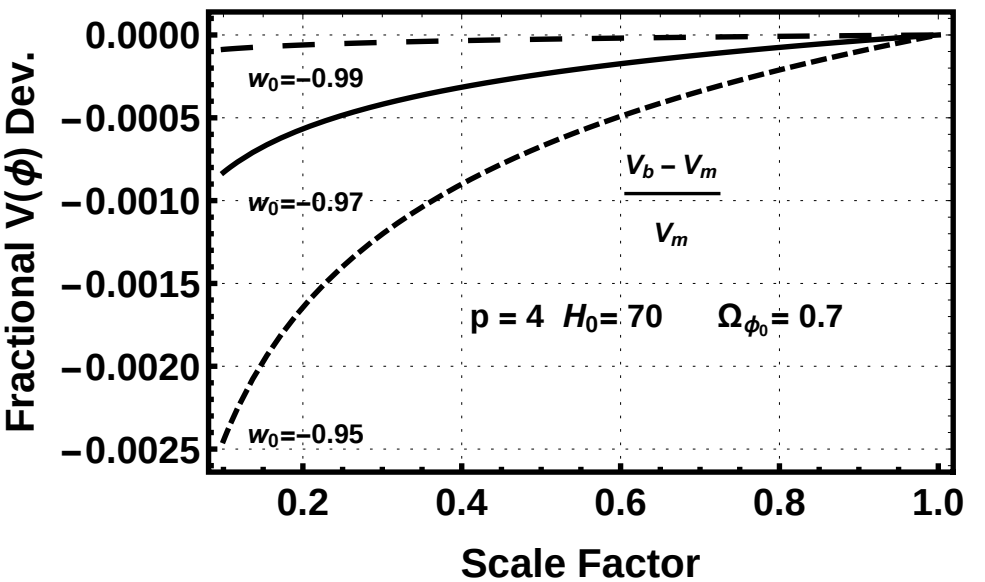



Fig. 1. The fractional error of the beta potential representation of the model potential for a quartic inverse power law model potential with $w_0 = -0.99$, $-0.97$ and $-0.95$.

than 0.25% for all cases and is less than 0.1% for the two $w_0$ values closest to minus one. This is a significantly higher accuracy than current observational cosmological parameter constraints.

## 7. A Universe with Matter

Since the goal of this article is the calculation of parameters for comparison with observations the dark and baryonic matter of the universe must be accounted for. In the beta function formalism the matter density $\rho_m(\kappa\phi)$ is given by

$$\rho_m(\kappa\phi) = \rho_{m0} \exp\left(-3 \int_{\kappa\phi_0}^{\kappa\phi} \frac{d\kappa\phi}{\beta(\kappa\phi)}\right) \tag{23}$$

where $\rho_{m0}$ is the current matter density. Using the definition of $\beta$ Eq. (23) can be written as

$$\rho_m(a) = \rho_{m0} \exp\left(-3\int_1^a dln(a)\right) = \rho_{m0} a^{-3} \tag{24}$$

as expected, independent of $\beta$.

### 7.1. *Introducing the superpotential to mass*

The presence of mass has a profound effect on the superpotential. Instead of the simple forms in section 6 there is a differential equation (2) for the superpotential $W$ that has the form

$$\frac{WW_{,\phi}}{\kappa^3} + \frac{\beta W^2}{2\kappa^2} = -2\frac{\rho_m}{\beta} \tag{25}$$

where $W_{,\phi}$ is the derivative of $W$ with respect to $\phi$. The solutions for $W$ are found through the use of integrating factors. Section 8 describes how the integrating factors can be found in general. The power and inverse power law integrating factors are $(\kappa\phi)^{-p}$ and $(\kappa\phi)^p$ respectively. Using the inverse power law integrating factor (4) shows that the the solution for $W(a)$ is

$$\begin{aligned} W_i(a) = -\Bigg\{ & -\frac{4\rho_{m_0}}{3}\left(\frac{2p}{3}\right)^{-\frac{p}{2}} \exp\left(\frac{3\phi_0^2}{2p}\right)(\phi(a))^{-p} \\ & \times\left[\Gamma\left(1+\frac{p}{2}, 3\ln(a)+\frac{3\phi_0^2}{2p}\right) - \Gamma\left(1+\frac{p}{2}, \frac{3\phi_0^2}{2p}\right)\right] \\ & + W_0^2\left(\frac{\phi_0}{\phi(a)}\right)^p\Bigg\}^{1/2} \end{aligned} \tag{26}$$

where $\Gamma$ is the incomplete Gamma function. The Hubble parameter is of course just $-\frac{1}{2}W$. Figure 2 shows the evolution of the Hubble parameter with $p = 4$.

### 7.2. *Onset of the acceleration of the expansion of the Universe*

An important check on the validity of the Hubble parameter is whether it initiates the late time acceleration of the universe at the correct epoch. Figure 3 shows the evolution of the speed of the expansion of the universe, $\dot{a}$, as a function of the scale factor $a$. The acceleration initiates at a scale factor of 0.6 consistent with the observations.

## 8. The Dark Energy Density

In a flat universe the dark energy density is just the total density minus the matter density therefore the ratio of the dark energy density to the critical density is

$$\Omega_\phi = \frac{3(H/\kappa)^2 - \rho_{m_0} a^{-3}}{3(H/\kappa)^2} \tag{27}$$

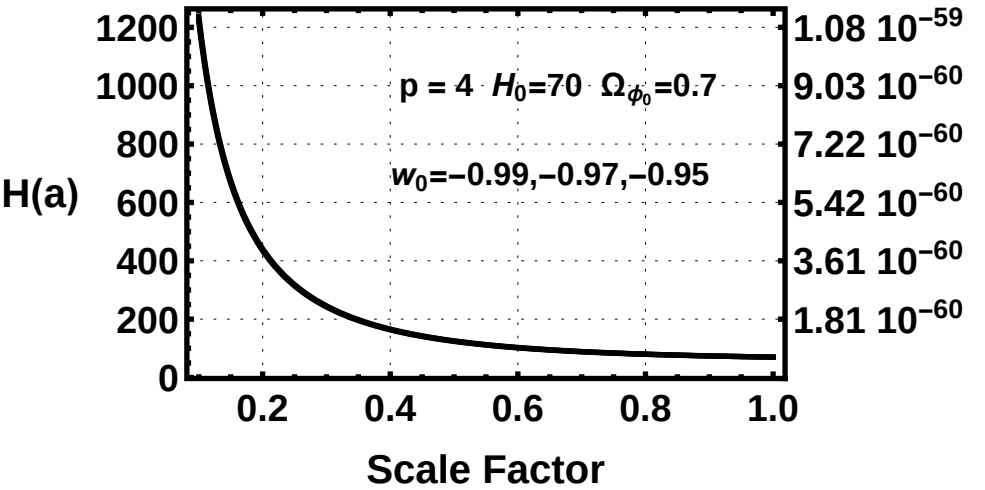



Fig. 2. The evolution of the Hubble parameter $H(a)$ with $p = 4$ and $w_0 = -0.97$. The left ordinate is labeled in $(km/sec)/Mpc$ and the right ordinate in reduced Planck mass $M_p$.

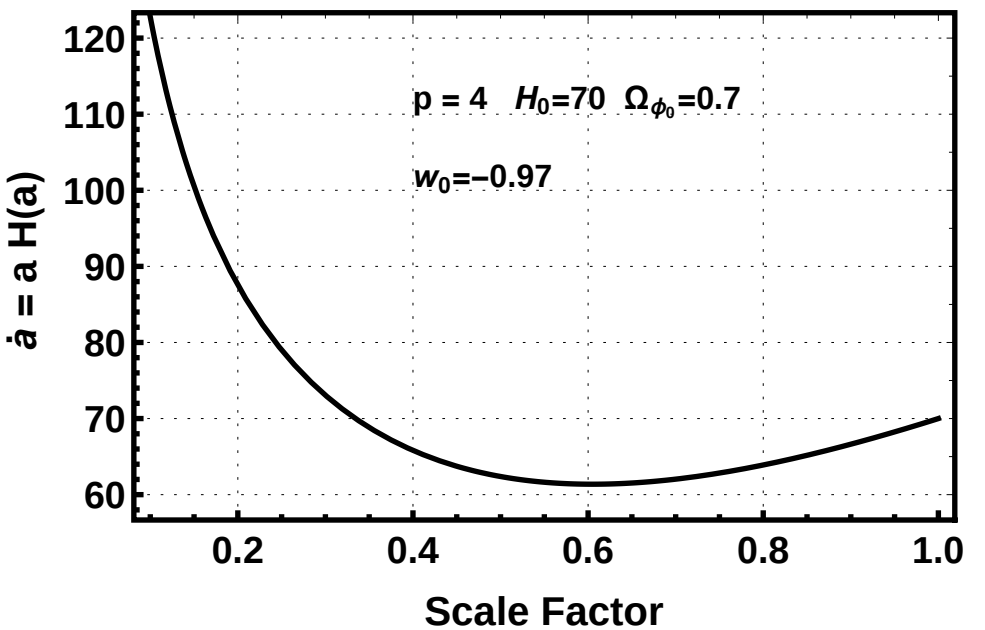



Fig. 3. The evolution of the speed of expansion of the universe $\dot{a}$ with $p = 4$ and $w_0 = -0.97$. The acceleration initiates at a scale factor of 0.6, consistent with observations.

Figure 4 shows the evolution of $\Omega_\phi(a)$ for $p = 4$ and the three values of $w_0$. The figure show that $\Omega_\phi$ is not a strong function of $w_)$ but some separation of the three plots is apparent near $a = 0.1$. $\Omega_\phi$ approaches zero at that point which is relevant in considering the evolution of $w$.

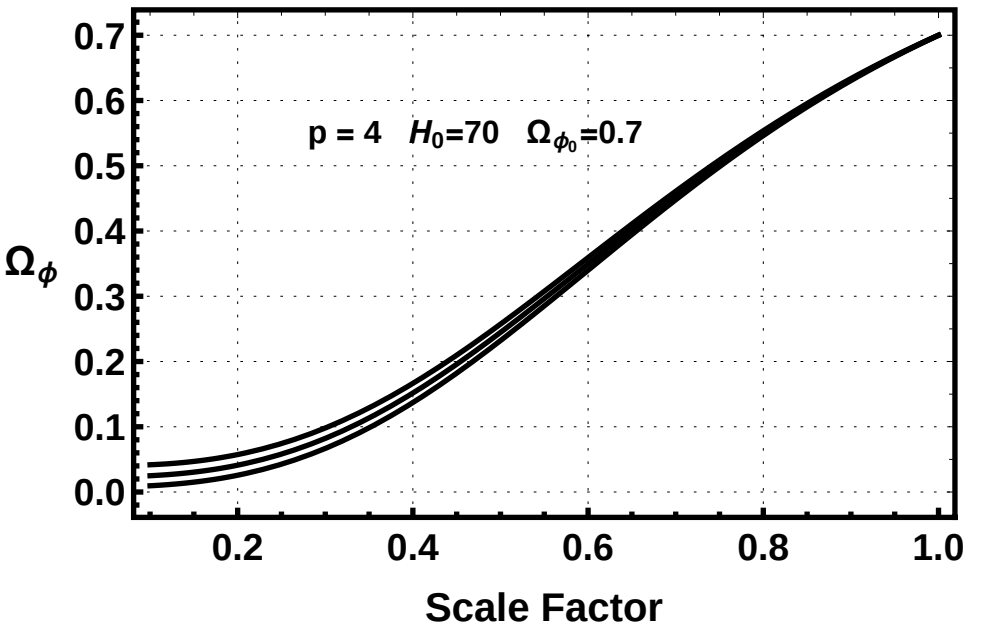



Fig. 4. The evolution of $\Omega_\phi$, the ratio of the dark energy density to the critical density, with $p = 4$ and $w_0 = -0.97$.

### 9. The Dark Energy Equation of State

The dark energy equation of state, DEoS, is an important cosmological parameter whose value discriminates between static and dynamic cosmologies. From (4) the equations for the equation of state with matter included are

$$1 + w(\phi) = \frac{\beta^2}{3} \frac{1}{(1 - \frac{4\rho_{m_0} a^{-3}}{3(W/\kappa)^2})} = \frac{\beta^2}{3} \frac{1}{(1 - \Omega_m)} = \frac{\beta^2(\phi)}{3\Omega_\phi} \tag{28}$$

This differs slightly from (4) because $\kappa$ is included in the equations rather than being set to one. Since, as shown in section 8, $\Omega_\phi$ approaches very close to zero at scale factors near $a = 0.1$ the more accurate first term to the right in Eq. (28) is used rather than the later terms. Figure 5 shows the evolution of $w(a)$ for $p = 4$ and the three $w_0$ values. Note that although the evolution of $w(a)$ is relatively well fit by the linear CPL model for scale factors larger than 0.5, the early evolution of $w(a)$ is very non-linear.

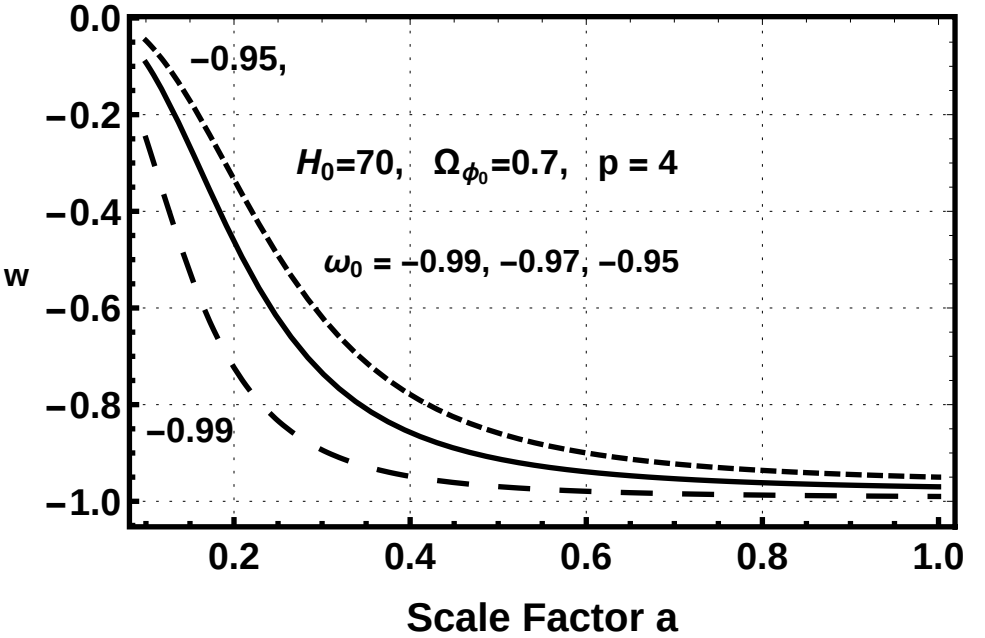



Fig. 5. The evolution of $w(a)$ with $p = 4$ and $w_0 = -0.99, -0.97$ and $-0.95$.

## 10. The Evolution of the Fundamental Constants $\alpha$ and $\mu$

Up until now we have concentrated on the evolution of cosmological parameters as a result of the rolling scalar field. Now attention is centered on two of the dimensionless fundamental constants, the fine structure constant $\alpha$ and the proton to electron mass ratio $\mu$. The same scalar field that interacts with the gravitational sector is also assumed to interact with the other sectors. It is difficult without fine tuning to limit the interaction of a scalar field to only the gravitational sector (6). The relationship between the variation of $\mu$ or $\alpha$ and $\phi$ is given simply by

$$\frac{\Delta x}{x} = \zeta_x \kappa(\phi - \phi_0) = \zeta_\mu \int_1^a \beta(a') d\ln(a') \tag{29}$$

where $x$ is either $\mu$ or $\alpha$ and $\zeta_x$ is the dimensionless coupling constant for the interaction. This can be taken as the first term in a Taylor series expansion of a more complicated coupling. $\beta(a)$ is equal to $\beta(\phi(a))$. From now on we consider the proton to electron mass ratio $\mu$ as it has the strongest limit on its variation $\Delta\mu/\mu \leq 10^{-7}$ at $z = 0.89$( 7) and (8) a look back time greater than half the age of the universe. Figure 6 shows the evolution of $\Delta\mu/\mu$ with a coupling constant $\zeta_\mu$ set to $10^{-6}$, a relatively large value set to highlight the nature or the evolution. All of the tracks in Fig. 6 except for $w_0 = -0.99$ violate the limit on $\Delta\mu/\mu$ at $z = 0.89$.

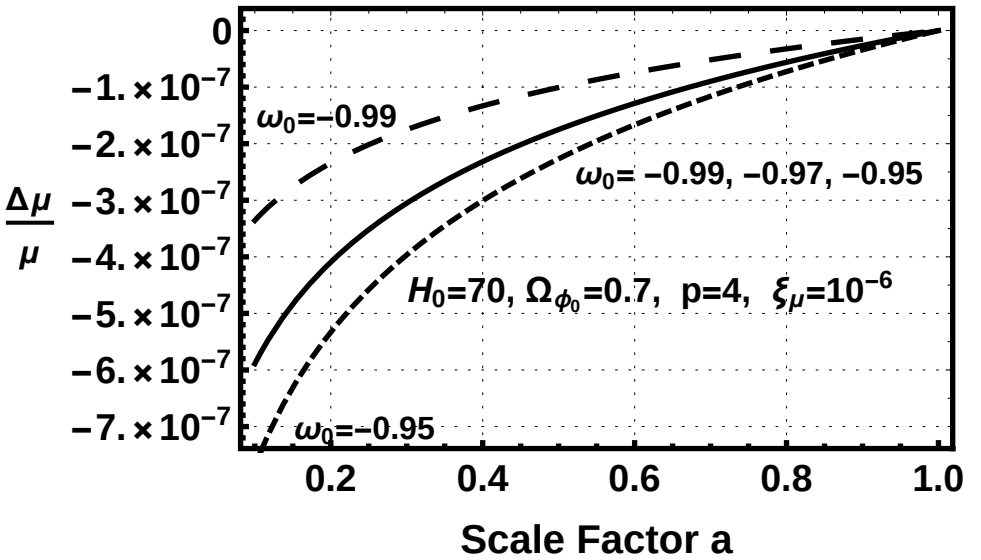



Fig. 6. The evolution of $\Delta\mu/\mu$ with $p = 4$ and $w_0 = -0.99, -0.97$ and $-0.95$.

## 11. Limiting the Parameter Space with Fundamental Constants

Equations (29) and (28) established a two dimensional parameter space in terms of $\zeta_\mu$ and $(w_0 + 1)$. The limits on $\Delta\mu/\mu$ or $\Delta\alpha/\alpha$ can be satisfied either by lowering the coupling constant or by moving $w_0$ closer to minus one. The explicit relation between $w_0$, $\Delta\mu/\mu$ and $\zeta_\mu$ is (5)

$$\zeta_\mu = \frac{\Delta\mu/\mu}{\kappa(\phi(a_{ob}) - \phi_0)}. \tag{30}$$

where $a_{ob}$ is the scale factor of the observation of the limiting constraint. Figure 7 shows the allowed and forbidden regions in the parameter space established by the limit on $\Delta\mu/\mu$ at $z = 0.89$. Future improvements on the variational limits of the fundamental constants can further reduce the allowable space. Only a verified observation of a change in $\mu$ or $\alpha$ can eliminate the ΛCDM and Standard Model point at $(0, 0)$.

## 12. Summary

This work is a demonstration of the beta function formalism using the quintessence cosmology with an inverse power law dark energy potential as an example. The beta function is defined as the derivative of the scalar field $\phi$ with respect to the natural log of the scale factor $a$, $\beta = d\phi/d\ln(a)$. The beta function is the negative of the logarithmic derivative of the dark energy potential and has potential the accurately but not exactly reproduces the true dark energy potential. When the criterion of $\beta^2/6 \ll 1$ is satisfied the beta potential to an accuracy of $(1 - \beta^2/6)$.

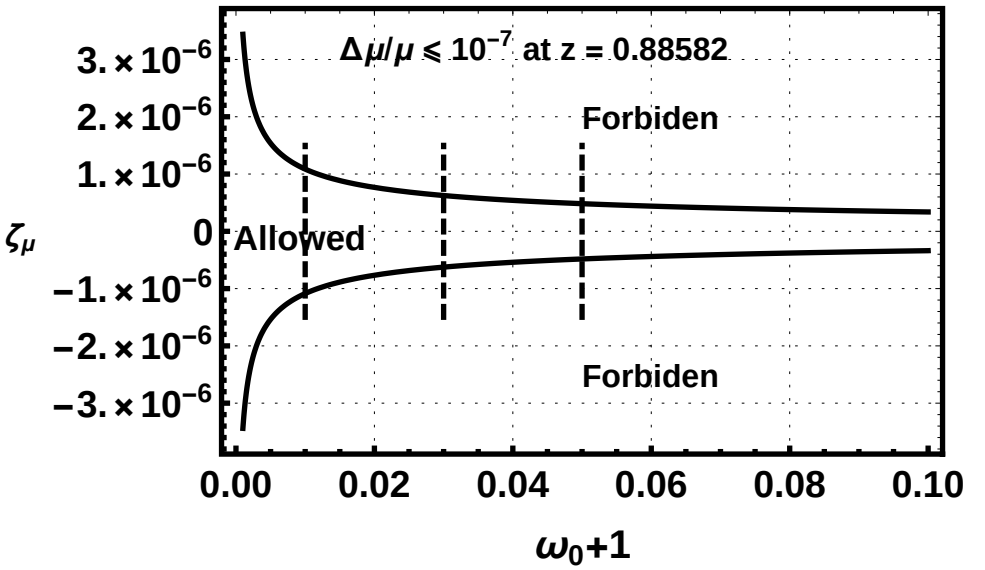



Fig. 7. The allowed and fforbidden areas in the $\zeta_\mu - (w_0 + 1)$ parameter space established by the observational limits on $\Delta\mu/\mu$. The vertical dashed lines indicate the three values of $w_0$ considered here.

An important feature of the beta function is that it provides a mechanism of determining $\phi(a)$ to produce analytic functions of the cosmological parameters and the evolution of fundamental constants as a function of the scale factor. These functions provided templates for comparison with observations that can discriminate between dynamical and static cosmologies along with the parameter space allowed for the cosmologies. The analysis reveals that ΛCDM is included in the quintessence cosmology power law potential as the zero power case.

**Acknowledgments**

The author wishes to acknowledge informative and helpful discussion with Peter Behroozi during the preparation of this manuscript.

**References**

1. Binetruy, P., Kiritisis, E. Mabillard, J., Pieroni, M. and Rosset, C. 2015 JCAP, 1504, no. 04, 033.
2. Cicciarella, F. and Pieroni, M. 2017 JCAP, 1708, no. 8, 010 arXiv:1611.10074v1 [gr-qc].
3. Kohri, K. and Matsui, H. 2017, JCAP, 06, 006.
4. Thompson, R. 2018, MNRAS, 477, 4104.
5. Thompson, R. 2018, MNRAS, 482, 5448.
6. Carroll, S.M. 1998, Phys. Rev. Let., 81, 3067.

7. Bagdonaite, J.,Dapra, Jansen, P., Bethlem, H.L., Ubachs, W., Henkel, C., and Menten, K.M. 2013, Phys. Rev. Letters, 111, 231101.
8. Kanekar, N., Ubachs, W., Menten, K.M., Bagdonaite, J., Brunthaler, A., Henkel, Muller, C.S., Bethlem, H.L. and Dapra, M. 2015, MNRAS 448, L104.

# Quantum satellites and tests of relativity

Matteo Schiavon, Giuseppe Vallone, Francesco Vedovato and Paolo Villoresi

*Dipartimento di Ingegneria dell'Informazione, Università degli Studi di Padova, Padova 35131, Italy*

*Istituto Nazionale di Fisica Nucleare (INFN) — Sezione di Padova, Italy*

Piergiovanni Magnani

*Department of Physics, Politecnico di Milano, Milano 20133, Italy*

Alexander R. H. Smith

*Department of Physics and Astronomy, Dartmouth College, Hanover, New Hampshire 03755, USA*

Sai Vinjanampathy

*Department of Physics, Indian Institute of Technology-Bombay, Powai, Mumbai 400076, India*

*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore, Singapore*

Daniel R. Terno*

*Department of Physics and Astronomy, Macquarie University, Sydney, NSW 2109, Australia*
*E-mail: daniel.terno@mq.edu.au*

Deployment of quantum technology in space provides opportunities for new types of precision tests of gravity. On the other hand, the operational demands of such technology can make previously unimportant effects practically relevant. We describe a novel optical interferometric red-shift measurement and a measurement scheme designed to witness the possible spin-gravity coupling effects.

*Keywords*: Equivalence principle; satellite-based experiments; optics; spin.

## 1. Introduction

Amazing progress in quantum sensing and quantum, together with satellite deployment of quantum technologies have ushered in a new era of experimental physics in outer space. The success of the first space based quantum key distribution experiments performed with the Micius satellite[1], is expected to be soon followed by European and North American missions. Current missions, such as LAGEOS-2, BEACON-C and LCT on Alphasat I-XL, are adapted for quantum optics experiments[2,3]. While the primary goal of the space-based platforms is to provide links for global quantum key distribution, the missions also envisage substantial scientific programs. These experiments have the exciting potential to enable novel searches for signatures of quantum gravity and/or physics beyond the standard model[4].

Here we describe the interplay of these technologies with the Einstein Equivalence Principle (EEP). The principle comprises three statements[5,6]. The first — *Weak Equivalence Principle* (WEP) — states that the trajectory of a freely falling test body is independent of its internal composition. Closely related to the WEP

is the Einstein elevator: if all bodies fall with the same acceleration in an external gravitational field, then to an observer in a small freely falling lab in the same gravitational field, they appear unaccelerated[6]. The remaining two statements deal with outcomes of non-gravitational experiments performed in freely falling laboratories where self-gravitational effects are negligible. The second statement — *Local Lorentz Invariance* — asserts that such experiments are independent of the velocity of the laboratory where the experiment takes place. The third statement — *Local Position Invariance* (LPI) — asserts that "the outcome of any local non-gravitational experiment is independent of where and when in the universe it is performed"[5].

We consider a novel all-optical test of LPI and the inertial and aspects of the spin-gravity coupling, commenting on implications for the validity of the WEP.

## 2. Optical test of position invariance

Tests of the "when" part of the LPI bound the variability of the non-gravitational constants over cosmological time scales[7]. The "where" part was expressed in Einstein's analysis of what in modern terms is a comparison of two identical frequency standards in two different locations in a static gravitational field. The so-called *red-shift* implied by the LPI affects the locally measured frequencies of a spectral line that is emitted at location 1 with $\omega_{11}$ and then detected at location 2 with $\omega_{12}$. The red-shift can be parametrized at the leading post-Newtonian order as

$$\Delta\omega/\omega_{11} = (1+\alpha)(U_2 - U_1) + \mathcal{O}(c^{-3}), \tag{1}$$

where $\Delta\omega := \omega_{12} - \omega_{11}$, $U_i := -\phi_i/c^2$ has the opposite sign of the Newtonian gravitational potential $\phi_i$ at the emission (1) and detection (2), while $\alpha \neq 0$ accounts for possible violations of LPI. In principle, $\alpha$ may depend on the nature of the clock that is used to measure the red-shift[5]. The standard model extension includes variously constrained parameters that predict LPI violation[8,9]. Alternative theories of gravity not ruled out by current data also predict $\alpha \neq 0$[5,10].

A typical red-shift experiment involves a pair of clocks, naturally occurring[11] or specially-designed[12–14], whose readings are communicated by electromagnetic (EM) radiation. The resulting estimates of $\alpha$ are based on comparison of fermion-based standards. Hence, different types of experiments, which employ a single EM-source and compare optical phase differences between beams of light traversing different paths in a gravitational field, provide a complementary test of LPI.

Such an all-optical experiment was proposed as a possible component of the QEYSSAT mission[4]. A photon time-bin superposition is sent from a ground station on Earth to a spacecraft, both equipped with an interferometer of imbalance $l$, in order to temporally recombine the two time-bins and obtain an interference pattern depending on the gravitational phase-shift:

$$\varphi_{\rm gr} = \frac{\Delta\omega}{\omega}\frac{2\pi}{\lambda}l \approx (1+\alpha)\frac{2\pi}{\lambda}\frac{ghl}{c^2}\ , \tag{2}$$

where $g$ is the Earth's gravity, $h$ the satellite altitude and $\lambda = 2\pi c/\omega$ the sent wavelength. For $\alpha = 0$, the order of magnitude of the gravitational red-shift is about 1 rad supposing $\lambda = 1550$ nm, l=1.2 km and an altitude $h = 1500$ km (which corresponds to $\Delta U \approx -1.3 \times 10^{-10}$).

However, a careful analysis of this optical COW-like experiment[15] revealed that first-order Doppler effects are roughly $10^5$ times stronger than the desired signal $\varphi_{\rm gr}$ from which $\alpha$ would be estimated. This first-order Doppler effect is an unavoidable part of the interferometric phase[16] was recently measured by exploiting large-distance precision interferometry along space channels[17], which constitute a resource for performing fundamental tests of quantum mechanics in space and space-based quantum cryptography.

We propose[18] a new gravitational red-shift experiment, which uses a single EM-source and a double large-distance interferometric measurement performed at two different gravitational potentials. By comparing the phase-shifts obtained at a satellite and on Earth, it is possible to cancle the first-order Doppler effect. Thus, this experimental proposal allows for a bound on $\alpha$ quantifying the violation of LPI in the EM-sector with a precision on the order of $10^{-5}$.

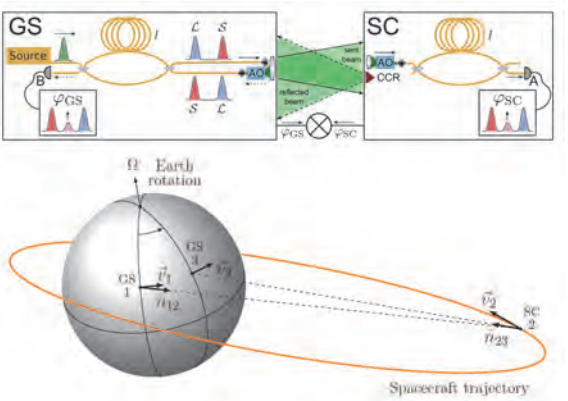



Fig. 1. Top: A schematic diagram of the proposed experiment. Both the ground station (GS) and spacecraft (SC) are equipped with a MZI of equal delay line $l$ and an adaptive optics system for fibre injection. Bottom: The geometry of the GS and SC used in the experiment, where $\vec{v}_1$ is the velocity of the GS at the emission location and potential $U_1$; $\vec{v}_2$ is the velocity of the SC at the detection location on the satellite and potential $U_2$; $\vec{v}_3$ is the velocity of the ground station at the detection of the beam retro-reflected by the SC, which occurs at a potential $U_3 = U_1$.

This proposal[17] is comprised of an interferometric measurement obtained by sending a light pulse through a cascade of two fiber-based Mach-Zehnder interferometers (MZI) of equal temporal imbalance $\tau_l$. After the first MZI, the pulse is split into two temporal modes, called *short* ($\mathcal{S}$) and *long* ($\mathcal{L}$) depending on the path taken in the first MZI. The equal imbalance of the two MZIs guarantees that the two pulses are recombined at the output of the second MZI, where they are detected. Such a satellite interferometry experiment setup is sketched in Fig. 1.

The combination of the possible paths the pulses may take leads to a characteristic detection pattern comprised of three possible arrival times for each pulse, as depicted in the insets of the upper picture in Fig. 1. The first (third) peak corresponds to the pulses that took the $\mathcal{S}$ ($\mathcal{L}$) path in both the MZIs, while the mid peak is due to the pulse that took the $\mathcal{S}$ path in the first interferometer and the $\mathcal{L}$ path in the second interferometer, or vice versa. Hence, interference is expected only in the central peak due to the indistinguishability of these latter two possibilities. A successful realization of the experiment depends on a number of important technical aspects that are described in detail in Ref. 18.

A bound on $\alpha$ is retrieved from the difference of the two phase-shifts, $\varphi_{\rm SC}$ and $\varphi_{\rm GS}$, that are obtained from interferometric measurements on the spacecraft and ground station, respectively. As just described, the interfering beams take different paths in the passage through the two MZIs. At the satellite, the beam that took the $\mathcal{L}$ path on Earth and the $\mathcal{S}$ path on the spacecraft interferes with the beam that passed took the $\mathcal{S}$ path on Earth and then took the $\mathcal{L}$ one on the spacecraft. This interference is a result of the phase difference $\varphi_{\rm SC}$. Analogously, at the ground station (GS) the beams that were delayed on the Earth before and after their round trip to the spacecraft (SC) will also interfere because of the phase difference $\varphi_{\rm GS}$.

The first-order Doppler terms are eliminated by manipulating the corresponding data sets from the GS and SC in a manner similar to the time-delay interferometry techniques[19] and those used in the Gravity Probe A experiment[13]. The basis for the derivation is the observation[16,18] that the phase difference at each detector is proportional to the proper time difference at the emission, that is related to the different travel times over $\mathcal{S}$ and $\mathcal{L}$. The key feature allowing for this elimination is that the ratio of the terms stemming from the first-order Doppler effect the two signals, $\varphi_{\rm SC}$ and $\varphi_{\rm GS}$ is exactly two[18]. Hence the target signal is $S = \varphi_{\rm SC} - \frac{1}{2}\varphi_{\rm GS}$, leading to

$$\frac{S}{\omega_0 \tau_l} = (1+\alpha)(U_2 - U_1) + \tfrac{1}{2}(\beta_2^2 - \beta_1^2) - \vec{\beta}_1 \cdot (\vec{\beta}_1 - \vec{\beta}_2) - (\mathfrak{d}_2^2 - \mathfrak{d}_1^2) - T(\hat{n}_{12} \cdot \vec{a}_1) - \left( (\vec{\beta}_2 - \vec{\beta}_1)^2 - (\mathfrak{d}_2 - \mathfrak{d}_1)^2 \right) \frac{\tau_l}{4T}, \qquad (3)$$

where $\alpha$ parameterizes the violation of LPI, $\vec{\beta}_i := \vec{v}_i/c$, $\mathfrak{d}_i := \hat{n}_{12} \cdot \vec{\beta}_i$, $T$ is the zeroth order time-of-flight between the GS and the SC, $\vec{a}_1$ is the centripetal acceleration of the GS at 1.

## 3. Weak equivalence principle and orbiting clocks

Matter of the Standard Model is characterized by two parameters of the irreducible representations of the Poincaré group: mass and spin (or helicity). General relativity is a universal interaction theory about masses[20], as the Newtonian gravity is, with polarization effects implicitly omitted from the WEP.

Regardless of their origins, spin-gravity coupling terms provide effective corrections to the Hamiltonian in the limit of weak gravity and non-relativistic motion.

The leading terms of the Hamiltonian of a free spin-$\frac{1}{2}$ particle that take into account the effects of rotation of the reference frame with angular velocity $\vec{\omega}$ and acceleration $\vec{a}$ (or a uniform gravitational field) can be represented as

$$H = H_{\rm cl} + H_{\rm rel} + H_\sigma + H_{\rm ext}. \tag{4}$$

The first three terms on the right hand side are obtained by performing the standard Foldy-Wouthuysen transformation and taking the non-relativistic limit [21]. The term $H_{\rm cl}$ represents the standard Hamiltonian of a free non-relativisitc particle in a non-inertial frame, $H_{\rm rel}$ describes the higher-order relativistic corrections that do not involve spin, and

$$H_\sigma = -\tfrac{1}{2}\hbar\vec{\omega}\cdot\vec{\sigma} + \frac{\hbar}{4mc^2}\vec{\sigma}\cdot(\vec{a}\times\vec{p}). \tag{5}$$

Finally, the term

$$H_{\rm ext} = \frac{\hbar k}{2c}\vec{a}\cdot\vec{\sigma} \tag{6}$$

represents the spin-accelerating (or spin-gravity) coupling. It is a limiting form of the simplest phenomenological addition to the Dirac equation that breaks the WEP [22]. For the value $k = 1$ it results from a particular version of the Foldy-Wouthuysen transformation [23]. While commonly considered a mathematical artefact of this transformation, the term naturally arises in gravitationally inspired Standard Model extensions. Only model-dependent bounds on $k$ in $H_{\rm ext}$ were obtained by a variety of techniques [24], including the optical magnetometery [25].

The spin dependent terms are small under normal conditions. On the Earth's surface $\hbar g/c = 2.15\times10^{-23}$ eV, which is equivalent to an effective magnetic field of $3.7\times10^{-19}$ Tl, still several orders of magnitude below the peak sensitivity of optical magnetometery. The spin-rotation term is significantly larger, since already on the ground $\omega c/g = 2.22\times10^{3}$. It will be about an order of magnitude stronger for low-orbit satellites that are planned to carry entangled optical clocks aiming to establish the precision of $10^{-18}-10^{-20}$, making it a factor to consider in the clock design.

A potentially promising way of detecting these effects is via so-called weak amplification [26]. Weak value amplification involves two systems (typically referred to as "system" and "meter") that can interact via an interaction Hamiltonian of the form $q\delta(t-t_0)\hat{A}\otimes\hat{p}$. The bipartite system-meter is prepared in an initial state $|s_i\rangle\otimes|m_i\rangle$, following which the two are allowed to interact for a small time that includes $t_0$. Following this, the system is measured and measurements corresponding to a post-selected system state $|s_f\rangle$ are considered. This pre- and post-selection induces a "kick" in the meter state, given by the evolution $e^{-iq\mathcal{A}_w\hat{p}}|m_i\rangle$, where $\mathcal{A}_w \equiv \langle S_f|A_s|S_i\rangle/\langle S_f|S_i\rangle$.

The key insight here is that since $\langle S_f|S_i\rangle$ can be a small number, the measurement of $q$ is influenced by a large multiplicative factor $\mathcal{A}_w$. A subsequent measurement of the meter reveals the desired parameter $q$. Trapped atoms are potentially promising system to implement this scheme [26].

The simplest model of such a set-up consists of a species of spin interacting with a harmonic mode, subject to an interaction of the form

$$H \approx \hbar\omega_t(a^\dagger a + 1/2) + \frac{\hbar\omega_g}{2}\sigma_z + \hbar\lambda[(1 - i\delta)\sigma_+ a + (1 + i\delta)\sigma_- a^\dagger], \tag{7}$$

$\omega_g$ being the gravitational term and constants $\delta, \lambda$ are related to the interaction with the harmonic mode[26]. Analysis of the unitary evolution that is followed by post-selection indicates that for realistic parameter values the inertial and gravitational effects are within the sensitivity range of the optical magnetometery.

## References


1. J. Yin *et al.*, *Science* **356**, 1140 (2017).
2. G. Vallone *et al.*, *Phys. Rev. Lett.* **116**, 253601 (2016).
3. K. Günthner *et al.*, `arXiv:1608.03511` (2016).
4. A. Rideout *et al.*, *Class. Quant. Grav.* **29**, 224011 (2012).
5. C. M. Will, *Living Rev. Relativity* **17**, 4 (2014).
6. E. Poisson and C. W. Will, *Gravity: Newtonian, Post-Newtonian, Relativistic*, (Cambridge University Press, 2014).
7. J.-P. Uzan, *Living Rev. Relativity* **14**, 2 (2011).
8. V. A. Kostelecký and N. Russell, *Rev. Mod. Phys.* **83**, 11 (2011).
9. S. Liberatti, *Class. Quantum Grav.* **30**, 133001 (2013).
10. J. Sakstein, *Phys. Rev. D* **97**, 064028 (2018).
11. J. C. LoPresto, C. Schrader, and A. K. Pierce, *Astrophys. J* **376**, 757 (1991).
12. R. V. Pound and G. A. Rebka, Jr., *Phys. Rev. Lett.* **4**, 337 (1960).
13. R. F. C. Vessot *et al.*, *Phys. Rev. Lett.* **45**, 2081 (1980).
14. N. Ashby, T. E. Paker, and B. R. Patla, *Nature Phys.* **14**, 822-826 (2018).
15. A. Brodutch *et al.*, *Phys. Rev. D* **91**, 064041 (2015).
16. D. R. Terno, G. Vallone, F. Vedovato and P. Villoresi, *Phys. Rev. D* **101**, 104052 (2020).
17. G. Vallone *et al.*, *Phys. Rev. Lett.* **116**, 253601 (2016).
18. D. R. Terno, F. Vedovato *et al.*, `arXiv:1811.04835` (2018).
19. M. Tinto and S. V. Dhurandhar, *Living Rev. Relativity* **17**, 6 (2014).
20. W.-T. Ni, *Rep. Prog. Phys.* **73**, 056901 (2010).
21. F. W. Hehl and W.-T. Ni, Phys. Rev. D **42**, 20145 (1990).
22. A. Peres, *Phys. Rev. D* **18**, 2739 (1978).
23. Y. N. Obukhov, *Phys. Rev. Lett.* **86**, 192 (2001).
24. M. S. Safronova *et al.*, *Rev. Mod. Phys.* **90** 025008 (2018).
25. D. F. J. Kimball *et al.*, *Ann. Phys. (Berlin)* **525**, 514 (2013).
26. S. Ghosh, L.-C. Kwek, D. R. Terno, and S. Vinjanampathy, `arXiv:1912.10693` (2019).

# GINGERINO and the GINGER project

Filippo Bosi, Angela Di Virgilio, Umberto Giacomelli, Andrea Simonelli and Giuseppe Terreni
*INFN Sez. di Pisa, Largo B. Pontecorvo 3, Pisa, Italy*

Andrea Basti, Nicolò Beverini, Giorgio Carelli, Donatella Ciampini, Francesco Fuso, Enrico Maccioni and Fabio Stefani
*Dept. of Physics of the Univ. of Pisa, Largo B. Pontecorvo 3, Pisa, italy*

Antonello Ortolan
*INFN-National Laboratories of Legnaro, viale dell'Università 2, I-35020, Legnaro (PD), Italy*

Alberto Porzio
*CNR - SPIN and INFN, Napoli, Complesso Univ. Monte Sant'Angelo, via Cintia, 80126 Napoli, Italy*

Carlo Altucci, Salvatore Capozziello and Raffele Velotta
*Dept. of Physics, Univ. "Federico II" and INFN, Napoli, Complesso Univ. Monte Sant'Angelo, via Cintia, 80126 Napoli, Italy*

Matteo Luca Ruggiero and Angelo Tartaglia
*Politecnico di Torino, Corso Duca degli Abruzzi, 24 - 10129 Torino, Italy*

GINGER (Gyroscopes IN GEneral Relativity) is a project based on an array of large dimension ring laser gyroscopes, which is aiming at measuring in a ground laboratory the gravito-electric and gravito-magnetic effects (also known as De Sitter and Lense-Thirrings effect), foreseen by General Relativity, and proposed at the underground Gran Sasso laboratory (LNGS). The geometry control to keep constant the scale factor and the optimal orientation of the array have been studied. GINGERINO, a square ring-laser prototype built inside LNGS, has shown the advantages of an underground location for GINGER. At present, it is the only high sensitivity laser gyro running unattended in a seismically active area. It recorded the large signals of the sequence of the central Italy 2016 earthquakes and microseismic signals of the Mediterranean area five orders of magnitude smaller. The analysis of 90 days of continuous operation shows that its duty cycle is higher than 95%, with noise limit of the order of $10^{-10}$ (rad/s)/$\sqrt{Hz}$.

## 1. Introduction

The Sagnac effect has been discovered by Georges Sagnac more than 100 years ago, and states that the difference of time of flight of two light beams counter-propagating inside a closed path, is proportional to the angular rotation rate of the frame.[1–3] Usually the closed path is an optical fiber coil or a ring Fabry-Perot cavity composed by $3-4$ mirrors. The device can be passive or active: passive is when the resonant cavity is interrogated injecting light from the outside; active when it contains an active medium, and the device is itself a laser emitting along two counter propagating modes, in this case it is called Ring Laser Gyro (RLG).[4] RLGs based on large frame ring cavities, with perimeters of several meters, are at

present the most sensitive angular rotation sensors based on Earth. They have been built for geodesy, geophysics and for General Relativity (GR) tests[1] (for geophysics see *https : //www.geophysik.uni − muenchen.de/ROMY/*). Since 2011 we are studying the feasibility of the Lense Thirring test at the level of 1%, with an array of large frame RLGs,[5–7] studying in details the orientation of the RLGs in order to define the specifications of such apparatus. At the same time several tests have been pursued in order to define and test in details the geometry control and the data analysis. It is necessary to push the relative sensitivity of the Earth rotation rate measurement in the range from 3 parts in $10^9$ (present record[8]) up to 1 part in $10^{12}$, this is equivalent to reach the sensitivity level of $10^{-16}$rad/s. It is necessary to point out that this measurement is even more demanding since it requires very high accuracy. At present, the accuracy limit is certainly at least 10 times worst than the sensitivity limit. RLG consists of a laser with a cavity comprising of three or four mirrors, depending if the cavity is triangular or square, rigidly attached to a frame; the two counter-propagating cavity modes have slightly different frequency, and the beat note $f_s$ of the two beams is proportional to the angular rotation rate $\Omega$ felt by the ring cavity.

$$f_s = S\Omega\cos\theta \qquad (1)$$
$$S = 4\frac{A}{\lambda L}$$

where $A$ is the area of the ring cavity, $L$ is its perimeter, $\lambda$ the wavelength of the light, and $\theta$ the angle between the area versor of the ring and the rotation axis. Considering RLG attached to the Earth crust, horizontally aligned (*i.e.* area versor vertical) $\theta$ is the colatitude angle, while for RLGs aligned at the maximum Sagnac signal (*i.e.* area versor along the Earth rotation axes) $\theta = 0$. Eq. (1) connects $\Omega$ with the scale factor $S$, which depends on the geometry, and the wavelength $\lambda$. RLGs have large interest in fundamental physics to study the property of the gravito-magnetic field, and they have been proposed in the past for axion search and Lorenz invariance violation.[9] GINGER would provide *the first measurement* of a *GR* dynamic effect of the gravitational field on the Earth surface (not considering the gravitational redshift). Though not in free fall condition, it would be a direct local measurement, independent from the global distribution of the gravitational field and not an average value, as in the case of space experiments. It is important to point out that this kind of measurement depends on the angular momentum of the Earth, so it provides a way to check whether dark matter is nearby rotating. There are space based project to further develop the gravito-magnetic measurements in space to test the presence of dark matter.[10] The measurement is based on properties of the velocity of light and on the measurement of the frequency, ensuring fast response, large bandwidth, and a huge dynamical range. For example GINGERINO, our RLG prototype at LNGS, has recorded microseismic events in the range of fraction of nrad/s and high magnitude nearby earthquakes[11] 5 orders of magnitude higher. Fig. 1 shows the typical lay-out of a square cavity RLG. The

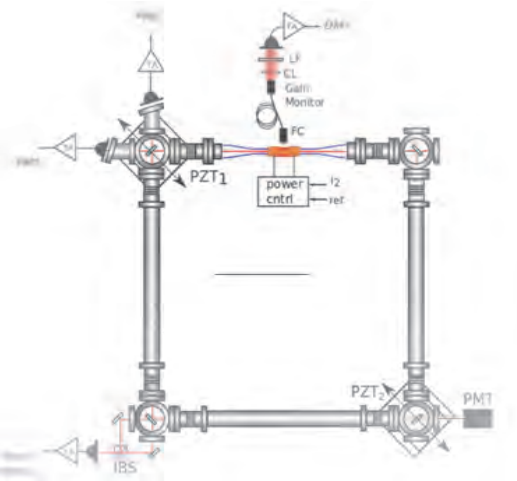



Fig. 1. Typical scheme of RLG with a square ring cavity.

four mirrors are placed at the corners of the square ring, each contained inside small vacuum chambers connected by steel pipes. The whole setup is vacuum tight and filled with Helium and an isotopic 50/50 mixture of $^{20}$Ne and $^{22}$Ne. The r.f. discharge, located in one of the side, generates the laser plasma. In most of the cases, piezoelectric actuators are utilised to translate the mirrors, allowing a control of the RLG perimeter length. The Sagnac beat note signal is observed at one corner (bottom-left) by superimposing the two output beams on a photodiode. At the top left corner two amplified photodiodes monitor the clock-wise (cw) and the counter clock-wise (ccw) output beams optical power (mono-beam photodiodes). Another photodiode monitors the fluorescence from the discharge, filtered around 633nm by an interference filter, providing a rough indication of the density of excited atoms in the laser upper state. In normal operation, plasma discharge is electronically controlled in order to keep as much as possible constant the optical power of one of the two mono-beams. All these signals are acquired by an ADC card at a frequency rate of a 5 kHz, suitable to allow their reliable reconstruction from DC up to the Sagnac frequency.

## 2. GINGER

The effects of the gravito-magnetic (Lense-Thirring, LT) and gravito-electric (de Sitter, dS) fields are observable by the gyro as angular velocity $\mathbf{\Omega_{LT}}$ and $\mathbf{\Omega_{dS}}$, that combine with $\mathbf{\Omega}_\oplus$, the Earth angular velocity, seen in the Cosmic inertial frame. *GR* is extremely predictive, allowing to compute the orientation and the amplitude of $\Omega_{LT}$ and $\Omega_{dS}$, in function of the latitude of the observer. Note also that several alternative theories of gravitation predict different dependence on latitude; so that a direct observation would, perhaps, discriminate between different theories.[12] Fig. 2

shows the mutual orientation of the various vectors for $\simeq 45^o$ of latitude (as e.g. at the GranSasso laboratories, LNGS).

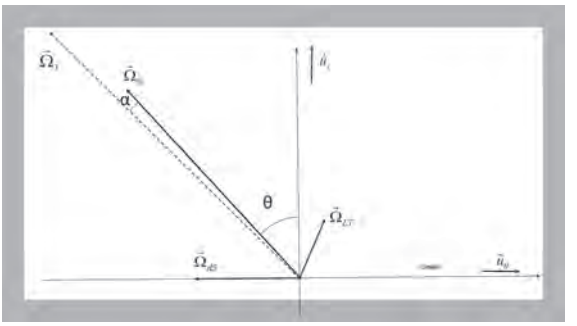



Fig. 2. The three axial vectors $\vec{\Omega}_\oplus$, $\vec{\Omega}_{LT}$, and $\vec{\Omega}_{dS}$ are shown, with the relative orientation at the latitude of the underground laboratory of GranSasso (LNGS), following General Relativity. The angle $\alpha$ and $\Omega_T$ (dashed line) are shown in the picture. The graph is not to scale, it gives just a pictorial view of the relative orientations of the different components. In reality, the modulus of $\vec{\Omega}_\oplus$ is 9 orders on magnitude bigger than the $GR$ terms, and the angle $\alpha$ is of the order of $\sim 3.5{\cdot}10^{-10}$ rad at the latitude of $45^o$.

The idea is to measure the total angular velocity seen by the gyros array and to evaluate the GR terms comparing the results with the measurement done by the international system IERS (International Earth Rotation and reference System), which independently measures the kinematic term $\Omega_\oplus$. The problem of the Lense-Thirring test is a very general one, similar to any effect induced by geophysical phenomena; it is particularly challenging since it is a DC effect, and not only sensitivity is required, but also accuracy. The main points of the GINGER apparatus are the following:

- reconstruct with very high precision a vector in the space using the information of the projectors, Eq. 2 shows that the geometry of the apparatus must be controlled, in particular the ratio $\frac{A}{P}$, the wavelength $\lambda$ and the variation of the absolute inclination $\delta\theta$;
- in general at least 3 independent gyros, 4 or more would be better for redundancy;
- it is necessary to study the noise related to any kind of variations of the apparatus, in particular all geophysical signals.

The main difficulty is the required accuracy, because it is necessary to distinguish a DC signal $9-10$ orders of magnitude smaller than the dominant signal $\Omega_\oplus$. It has been shown that in general the signal can be reconstructed combining the different RLG of the array,[6] but the relative angles between the different RLGs must be known with adequate accuracy. This means accuracy in the relative alignment of the RLGs of the order of nrad or better. One technique suitable to monitor with nanometer accuracy the distance between the mirrors has been tested.[13] Utilising

the natural symmetry of the problem it is possible to mitigate the requirement of the relative alignment. It has been shown that aligning one of the RLG at the maximum signal and the other horizontally or vertically it is possible to evaluate the angle between the two and solve the problem.[7] The RLG aligned at the maximum signal, so with area versor along the Earth rotational axis, has the interesting property that delivers the modulus of the total angular velocity and is at first order insensitive to local tilts. In principle it delivers the variations of the modulus of $\Omega_{\oplus}$, or equivalently the Length of Day (LoD). The details of the specification have been published.[6,7] Fig. 3 shows a picture of GINGER with 3 independent RLG, the location may be node B of LNGS, an area far apart from the main experiments. The exact dimension of the ring laser resonators has not been decided yet, but they will be squared structures with side length of $5 - 6$ meters.

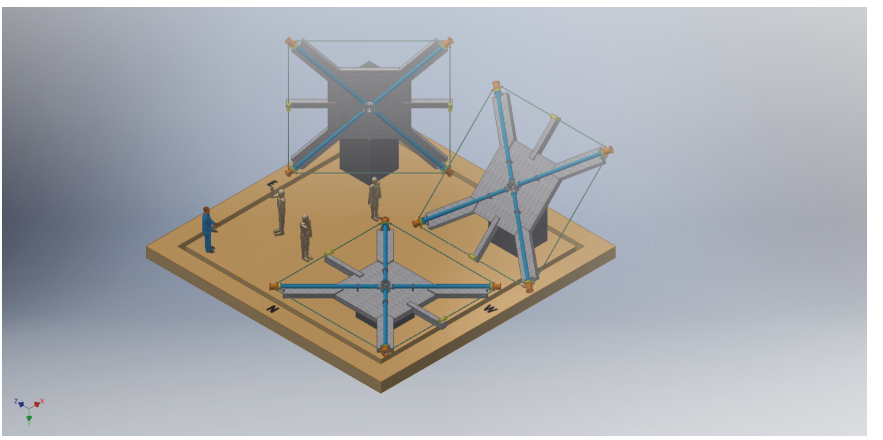



Fig. 3. Pictorial view of the GINGER project.

## 3. GINGERINO and work toward GINGER

Large effort has been pursued in Italy toward GINGER. Presently the most performing RLG is "G", situated at the Geodetic Observatory of Wettzell in Germany; it is composed of a monolithic structure, a rigid block of Zerodur with mirrors optically contacted in order to have a geometry fixed by construction. It is evident that such a structure cannot be extended to form a three axial array. Hetero-lithic mechanical structures have to be used instead, and in this case an electronic control is necessary in order to keep constant the geometry. A suitable control scheme to keep constant the scale factor has been studied and tested.[14–16] It is a well known fact that RLGs are affected by non linearity of the laser, and, in particular, optical back-scattered noise. In 2012-2014 Kalman filters have been successfully applied,[17,18] more recently we have developed a novel analytical technique to reconstruct the Sagnac frequency avoiding laser systematic.[19] Fig. 4 shows the time

behaviour of the data of GP2, analysed with the standard method and with the new one. GP2 is a square ring laser with side of 1.6m, aligned at the maximum Sagnac signal and presently working in Pisa for test purposes.[16] Gingerino is a test RLG

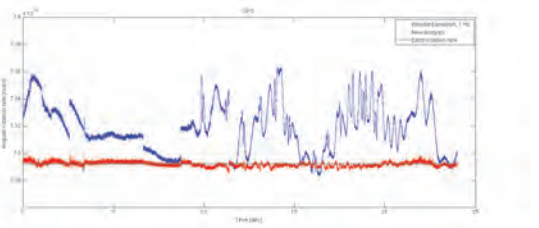


Fig. 4. One day of data of the prototype GP2 analysed with the standard method and the new one, the black line is the expected Earth angular rotation rate. GP2 is the RLG prototype with 6.4m perimeter, aligned at the maximum Sagnac signal.

3.6 m in side, installed at LNGS since 2014. In seismology, translation and strain are routinely observed by seismometer and strain meters. However, a full description of the ground movements requires also the acquisition of a third type of information, namely rotations. In particular, co-located translation and rotation sensors allow to estimate the local underground velocity structure, which is an information of high interest in geophysical survey. Rotational signals induced by seismic waves have a quite small amplitude. A strong seismic wave with a linear acceleration of 1 mm/$s^2$ produces a rotation velocity amplitude of some $10^{-7}$ rad/s, while micro-seismic rotational background noise (around 0.1 Hz) is smaller than $10^{-10}$ rad/s.[11,20] GINGERINO has shown the advantage of the underground location, since it is the first large frame hetero-lithic RLG, which operates continuously unattended with high sensitivity and with 95% duty cycle.[21] It has also shown that the local tilts on long time are below $\mu$rad,[7] this is an important parameter for the installation of a RLG at the maximum signal. GINGERINO is taking data for seismology, so at present inside LNGS there is one of the few seismic station allowing to record at the same time both the three translation components and the rotation angle.[11,20] Besides, GINGERINO is a very useful tool to develop new analysis technique, to investigate the very low frequency signals and to test all the equipment to remotely control and operate on the apparatus.

## 4. Conclusion

A large experimental work has been pursued toward GINGER, the array of RLGs designed to detect on ground the Lense-Thirring effect at 1% precision. Geometry control of each independent RLG, study of the optimal orientation of each RLG of the array, and reconstruction of the angular rotation rate taking into account the systematic of the laser. GINGERINO has been realized to validate LNGS for this experiment, and it has shown the advantage of an underground location

far from atmospheric perturbations and large thermal variation. In fact it is the first high sensitivity RLG based on a hetero-lithic mechanical structure operating with 95% duty cycle and with a sensitivity of the order of $10^{-10}$rad/s. It has detected microseismic signals of the order of $10^{-10}$ rad/s and large teleseismic events 5 orders of magnitude larger. It is the first and only RLG operative in one of the Mediterranean most important seismically active area.

## References


1. Dossier sommaire. The Sagnac effect: 100 years later. *Comptes Rendus Physique*, 15(10):iii–iv, 2014.
2. K.U. Schreiber, A. Gebauer *et al.*, *Comptes Rendus Physique*, 15(10):859–865, 2014.
3. A. Di Virgilio, M. Allegrini *et al.*, *Comptes Rendus Physique*, 15(10):866–874, 2014.
4. K.Ul. Schreiber and J-P Wells, *Review of Scientific Instruments*, 84(4):041101, 2013.
5. F. Bosi, G. Cella, A. Di Virgilio *et al.*, *Phys. Rev. D*, 84:122002, Dec 2011.
6. Angelo Tartaglia, Angela Di Virgilio *et al.*, *The European Physical Journal Plus*, 132(2):73, Feb 2017.
7. A.D. V. Di Virgilio, J. Belfi *et al.*, *The European Physical Journal Plus*, 132(4):157, Apr 2017.
8. K. U. Schreiber, A. Gebauer, and J.-P. R. Wells, *Opt. Lett.*, 38(18):3574–3577, Sep 2013.
9. G. E. Stedman. *Reports on Progress in Physics*, 60(6):615, 1997.
10. Ruggiero M.L. and Tartaglia A. *arXiv:1810.11785[gr-qc]*, 2019.
11. A Simonelli, H Igel *et al.*, *Geophysical Journal International*, 214(1):705–715, 2018.
12. S. Capozziello, G. Lambiase *et al.*, *Phys. Rev. D*, 91:044012, Feb 2015.
13. A. Donazzan, G. Naletto *et al.*, *IEEE Trans. on Instr. and Meas.*, 68, 4 1165–1177, 2019.
14. R Santagata, A Beghi *et al.*, *Classical and Quantum Gravity*, 32(5):055013, Feb 2015.
15. J Belfi, N Beverini, D Cuccato *et al.*, *Classical and Quantum Gravity*, 31(22):225003, Oct 2014.
16. J Belfi, N Beverini *et al.*, *submitted to Optics Letters*, 2019.
17. Alessandro Beghi, Jacopo Belfi *et al.*, *Appl. Opt.*, 51(31):7518–7528, Nov 2012.
18. D. Cuccato, A. Beghi *et al.*, *Metrologia*, 51:97–107, 2014.
19. Di Virgilio A.D.V., Beverini N. *et al.*, *submitted to Optica*, 2019.
20. Simonelli A. *et al.*, *in Preparation*, 2019.
21. J. Belfi, N. Beverini *et al.*, *Appl. Opt.*, 57(20):5844–5851, Jul 2018.

# A spacetime structure exploration plan in the earth-moon system

Dong Jiang*, Wang Peng, Hu Xiao Gong, Du Yuan Jie, Cao Zhou Jian, Zhang Yan, Xu Ming, Lee Ming, Zhang Zhong Ying, Shuai Tao and Liu Liang

*Yunnan Astronomical Observatories,*
*396 Yangfangwang, Guandu District, Kunming, 650216, P.R. China;*
*Key Laboratory for the Structure and Evolution of Celestial Objects, CAS;*
*Center for Astronomical Mega-Science, Chinese Academy of Sciences*
** E-mail: dj@ynao.ac.cn*
*https://github.com/JiangDong*

With the technology development in the cold atomic clock etc. and the gyposcope etc., many relativistic geodetic project and the plan of Atomic Gravitational wave Interferometric Sensor (AGIS) was proposed. Here we introduce one plan of spacetime structure exploration in the earth-moon system by the above mentioned thecniques, which focus on surveying the gravitational potential and gravitational first order redshift in spacetime geometry of the earth-moon system. We also discussed that apply those satellites and instrumenmts to find Geometrodynamic field moment which include gravitomagnetic clock, the possible CPT violation (Lorentz Invariance Violation) from Gravitional second order Redshift.

*Keywords*: Gereral relativity experiments; the clock and gyposcope; gravity; spacetime; difference geometry.

## 1. The Introduction

With the technology development in the clock and the gyroscope, we propose our plan of relativistic geodetic project which based on spacetime geometry of the earth-moon system. In Sec. 2, we introduce the instrument onboard, then we explain the major scientific objectives in Sec. 3, and the other scientific objectives in Sec. 4.

## 2. A Spacetime structure exploration plan in the earth-moon system

### 2.1. *The technology development in the clock and the gyposcope*

Recently, Space cold atomic clock (SCAC) had been test in the orbit[1], and ACES (Atomic Clock Ensemble in Space) also been installed on the ISS (International Space Station). An atomic clock with $10^{-18}$ instability on the ground had been about ten years[2,3]. Optical clock also arrive to high-accuracy[4].

In another side, the gyroscope from continuous cold-atom inertial sensor had reach 1 nrad/sec rotation stability[5]. Ring laser gyroscope have can detect earth tide and tilt detection[6]. Those technology provide the chance to check the relativity theory.

### 2.2. *Our Spacetime structure exploration plan in the earth-moon system*

Our plan is that the spacetime structure exploration in the earth-moon system by the Full Tensor Gradiometry (FTG) from the above mentioned thecniques inlcude the clock and gyroscope (more than 5 units), which focus on surveying the gravitational potential and gravitational first order redshift in the spacetime geometry of the earth-moon system. The satellite (or the detector) will sent to the Earth-Moon Lagrangian points which difference with Italy project [7–9] after finished relativistic geodetic mission.

And we will study the Psedo-Newtonian potenial in Kerr Spacetime like Paczyski-Wiita potential [10,11] et al. Although all Psedo-Newtonian potenial can back to PPN (Parameterized Post-Newtonian formalism) in some order, we will check it in the experiment. we also will check the effectness of the diffecence local quasi-inertial frame defined of PPN in the experiment.

Gyroscope in Local and nonlocal measurements of the Riemann tensor had study [12]. We will try to study the relation between the precision of the clock and gyroscope array and the large scale structure of spacetime, and the later with local and nonlocal measurements. We will try to find time dilation symmetry violations in the Gravitational secondary Red-shift in the data of the clock and the gyroscope.

## 3. The major scientific objectives

### 3.1. *Gravitational potential and Gravitational first order redshift for Relativitic geodesic*

Now atomic clocks is enough to measure relativity [2], many relativistic experiments with atomic clocks has been proposed [13]. A precision measurement of the gravitational red-shift by the interference of matter waves has been done [14], even a clock can directly linking time to a particle's mass [15].

Based on the sample relation between the gravitational potential and gravitational first order redshift in general relativity, many relativitic geodesic plan has been proposed, for example, A spaceborne gravity gradiometer concept based on cold atom interferometers for measuring Earth's gravity field [16]. Atomic clocks and gravity field determination for geodesy by many researcher [17,18]. In USA, many theory work [19–21] was done with the plan of NASA [22]. In German, the plan of gravitational clock compass [23] based on the early direction study of the curvature [24] was done. Compared with them, our theoretical basis is gravitoelectromagnetism analogy based on tidal tensor and the Terrae proper reference frame, and the instrument is FTG.

### 3.2. *The symmetry violations of Gravitational second order redshift*

After measure the gravitational first order red-shift[25], Jaffe and Vessot want to study the Second-Order gravitational redshift in the earth[26,27]. Then the symmetry violations of gravitational redshift is exist that was pointed out in the frame of parametrized post-Newtonian[28]. The similar result also had from the frame of gravitoelectromagnetism[29]. If the precision of the instrument is enough, we will try to find it from the data with the orbital raw data.

## 4. The other scientific objectives

### 4.1. *Gravitomagnetic effect in the clock and gyroscope*

Now many plan had been done for explore relativistic phenomena like LenseThirring (LT) effect and gravitational redshift et al. In clock, many gravitational redshift detecting plan for geodesy was proposed[2,13]. In gyroscope, there are cold atomic gyroscope[5] and superfluid gyroscope[30] for LT precession like the project of Gravty Probe B[31]. A multi-ring-laser gyroscope also designed for measuring gravitomagnetic effects[32]. In especial, OPTIS (`http://www.exphy.uni-duesseldorf.de/OPTIS/optis.html`) design the plan include the core technologies for OPTIS are optical cavities, highly stabilized lasers, capacitive gravitational reference sensors, dragfree control, ion clocks, frequency combs, and laser tracking systems for the experiments include Special Relativity, and General Relativity, and constancy of speed of light, and gravitational redshift, and LenseThirring effect et al.[33].

Recently, generalized gravitomagnetic clock effect was calculated for the GPS near the earth[34] after Mashoon give the gravitomagnetic clock effect for the circular orbits[35]. And some result shown it is in the order of $10^{-6}$ [36]. consider our instrument is FTG which difference with GNSS, and the orbit is in the large ellipse, we will try to find it from the data with the orbital raw data.

### 4.2. *Gravitational moment and wave detecting by the clock and atomic interferometer*

With the technology development in the clock which include cold atomic interferometric and optical frequency comb et al., and in the gyroscope which include laser gyroscopes and fiber optic gyroscopes based on the Sagnac effect and the matter wave gyroscopes et al., the plan of Atomic Gravitational wave Interferometric Sensor (AGIS) was proposed. A comparison between matter wave and light wave interferometers for the detection of gravitational waves had been done[37]. If our FTG instrument is well in the stability at long term, we will try to find the singal of gravitational wave from the data with the orbital raw data.

## 5. The Summary

Our plan is that using high-precision FTG based the atomic clock or (and) gyroscope to finishing relativistic geodetic mission and exploring spacetime geometry. About full tensor gravitational gradiometer (array) in measurement of gravitational theory and cosmology, we are applying the patent in China and in PCT (Patent Cooperation Treaty)[38–40].

## Acknowledgments

DJ thanks Project U1838104 supported by NSFC, and Project 201902 supported by Shanghai Space Navigation and Positioning Laboratory.

## References

1. L. Liu, D.-S. Lü, W.-B. Chen *et al.*, In-orbit operation of an atomic clock based on laser-cooled $^{87}$Rb atoms, *Nature Communications* **9**, p. 2760 (July 2018).
2. C. W. Chou, D. B. Hume, T. Rosenband and D. J. Wineland, Optical Clocks and Relativity, *Science* **329**, p. 1630 (September 2010).
3. N. Hinkley, J. A. Sherman, N. B. Phillips *et al.*, An Atomic Clock with $10^{-18}$ Instability, *Science* **341**, 1215 (September 2013).
4. C. W. Chou, D. B. Hume, J. C. J. Koelemeij, D. J. Wineland and T. Rosenband, Frequency Comparison of Two High-Accuracy $Al^{+}$ Optical Clocks, *Phys. Rev. Lett.* **104**, p. 070802 (February 2010).
5. I. Dutta, D. Savoie, B. Fang *et al.*, Continuous cold-atom inertial sensor with 1 nrad/sec rotation stability, *Phys. Rev. Lett.* **116**, p. 183003 (2016).
6. K. U. Schreiber, T. Klügel and G. E. Stedman, Earth tide and tilt detection by a ring laser gyroscope, *Journal of Geophysical Research (Solid Earth)* **108**, p. 2132 (February 2003).
7. A. Tartaglia, E. C. Lorenzini, D. Lucchesi, G. Pucacco, M. L. Ruggiero and P. Valko, How to use the Sun-Earth Lagrange points for fundamental physics and navigation, *General Relativity and Gravitation* **50**, p. 9 (January 2018).
8. F. de Marchi and G. Congedo, Space tests of the strong equivalence principle: BepiColombo and the Sun-Earth Lagrangian points opportunity, *International Journal of Modern Physics D* **26**, p. 1741021 (2017).
9. G. Congedo and F. De Marchi, Testing the strong equivalence principle with spacecraft ranging towards the nearby Lagrangian points, *Phys. Rev. D* **93**, p. 102003 (May 2016).
10. B. Paczyński and P. J. Wiita, Thick accretion disks and supercritical luminosities, *A&A* **88**, 23 (August 1980).
11. M. Abramowicz, The paczyski-wiita potential. a step-by-step "derivation". commentary on: Paczynsky and wiita, 1980, aa, 88, 23, *A&A* **500**, 213 (2009).
12. J. Audretsch and C. Lmmerzahl, Local and nonlocal measurements of the riemann tensor, *General Relativity and Gravitation* **15**, 495 (1983).

13. S. Reynaud, C. Salomon and P. Wolf, Testing General Relativity with Atomic Clocks, *Space Sci. Rev.* **148**, 233 (December 2009).
14. H. Müller, A. Peters and S. Chu, A precision measurement of the gravitational redshift by the interference of matter waves, *Nature* **463**, 926 (February 2010).
15. S.-Y. Lan, P.-C. Kuan, B. Estey *et al.*, A Clock Directly Linking Time to a Particle's Mass, *Science* **339**, p. 554 (February 2013).
16. O. Carraz, C. Siemes, L. Massotti, R. Haagmans and P. Silvestrin, A Space-borne Gravity Gradiometer Concept Based on Cold Atom Interferometers for Measuring Earth's Gravity Field, *Microgravity Science and Technology* **26**, 139 (October 2014).
17. J. Müller, D. Dirkx, S. M. Kopeikin *et al.*, High Performance Clocks and Gravity Field Determination, *Space Sci. Rev.* **214**, p. 5 (February 2018).
18. T. Mehlstäubler, G. Grosche, C. Lisdat, P. Schmidt and H. Denker, Atomic Clocks for Geodesy, *ArXiv e-prints* (March 2018).
19. N. Ashby, Relativistic theory of the falling retroreflector gravimeter, *Metrologia* **55**, p. 1 (February 2018).
20. B. Mashhoon, General Relativistic Gravity Gradiometry, *ArXiv e-prints* (June 2018).
21. D. Bini and B. Mashhoon, Relativistic gravity gradiometry, *Phys. Rev. D* **94**, p. 124009 (December 2016).
22. T. A. Ely, E. A. Burt, J. D. Prestage, J. M. Seubert and R. L. Tjoelker, Using the deep space atomic clock for navigation and science, *IEEE Trans Ultrason Ferroelectr Freq Control* **PP**, 950 (2018).
23. D. Puetzfeld, Y. N. Obukhov and C. Lämmerzahl, Gravitational clock compass in general relativity, *Phys. Rev. D* **98**, p. 024032 (July 2018).
24. P. Szekeres, The Gravitational Compass, *Journal of Mathematical Physics* **6**, 1387 (September 1965).
25. R. V. Pound and G. A. Rebka, Apparent Weight of Photons, *Phys. Rev. Lett.* **4**, 337 (April 1960).
26. J. Jaffe and R. F. C. Vessot, The Second-Order Gravitational Redshift, in *Bulletin of the American Astronomical Society*, , BAASVol. 6March 1974.
27. J. Jaffe and R. F. C. Vessot, The second-order gravitational redshift, *General Relativity and Gravitation* **6**, 55 (February 1975).
28. T. Krisher, Parametrized post-Newtonian gravitational redshift, *Phys. Rev. D* **48**, 4639 (November 1993).
29. L. I. Yuan-Hong and L. U. Jun-Li, The doppler effect in curved space-time, *Journal of Natural Science of Hunan Normal University* (2011).
30. B. Barrett, R. Geiger, I. Dutta *et al.*, The Sagnac effect: 20 years of development in matter-wave interferometry, *Comptes Rendus Physique* **15**, 875 (December 2014).
31. C. W. F. Everitt, D. B. Debra, B. W. Parkinson *et al.*, Gravity Probe B: Final Results of a Space Experiment to Test General Relativity, *Phys. Rev. Lett.* **106**, p. 221101 (June 2011).

32. F. Bosi, G. Cella, A. di Virgilio *et al.*, Measuring gravitomagnetic effects by a multi-ring-laser gyroscope, *Phys. Rev. D* **84**, p. 122002 (December 2011).
33. C. Lmmerzahl, I. Ciufolini, H. Dittus, L. Iorio, H. Mller, A. Peters, E. Samain, S. Scheithauer and S. Schiller, Optisan einstein mission for improved tests of special and general relativity, *General Relativity and Gravitation* **36**, 2373 (2004).
34. E. Hackmann and C. Lämmerzahl, Generalized gravitomagnetic clock effect, *Phys. Rev. D* **90**, p. 044059 (August 2014).
35. J. M. Cohen and B. Mashhoon, Standard clocks, interferometry, and gravitomagnetism, *Physics Letters A* **181**, 353 (October 1993).
36. F. de Felice and D. Bini, *Classical Measurements in Curved Space-Times* July 2010.
37. P. Delva, M. C. Angonin and P. Tourrenc, A comparison between matter wave and light wave interferometers for the detection of gravitational waves, *Physics Letters A* **357**, 249 (2006).
38. D. Jiang, W. Peng *et al.*, 3-d autonomous navigation based on high precision atomic clock array (2019).
39. D. Jiang, Application of interferometric method based on full tensor gravitational gradiometer array in measurement of gravitational effect curved spacetime and cosmology, cn and pct applying (2019).
40. D. Jiang, Application of relativistic effect in gravity matching navigation and guidance, external ballistic correction and spacecraft orbit prediction, and gravity survey in space and sea-air (2019).

# Thermal dimension of quantum spacetime: Comparison with the spectral dimension and application in cosmology

Francesco Brighenti[1], Giovanni Amelino-Camelia[2], Giulia Gubitosi[3], João Magueijo[4] and Grasiele Santos[5]

[1] *Dipartimento di Scienze Pure E Applicate, Università di Urbino "Carlo Bo", and INFN, Sezione di Firenze, Via Santa Chiara 27, I-61029 Urbino, Italy*

[2] *Dipartimento di Fisica Ettore Pancini, Universitá di Napoli "Federico II", and INFN, Sezione di Napoli, Complesso Univ. Monte S. Angelo, I-80126 Napoli, Italy*

[3] *Departamento de Física, Universidad de Burgos, E-09001 Burgos, Spain*

[4] *Theoretical Physics, Blackett Laboratory, Imperial College, London, SW7 2BZ, United Kingdom*

[5] *Departamento de Matemática e Computação, Universidade Federal de Itajubá, Av. BPS 1303, Itajubá MG, Brazil*

The dimensionality of the quantum spacetime is often understood in terms of the spectral dimension; a different notion of dimensionality, the thermal dimension, has recently been proposed by some of us in Ref. 1. We showed through the study of specific models of quantum gravity that, in those cases where the spectral dimension has puzzling physical properties, the thermal dimension gives a different and more meaningful picture. In Ref. 2, we applied the statistical mechanics developed to define the thermal dimension to the study of the production of primordial cosmological scalar perturbations, assuming a running Newton constant and Rainbow Gravity. Here, we briefly review the main arguments and results of these studies.

## 1. Thermal dimension of quantum spacetime

In the last decade it has been found that many different quantum gravity models share the common feature of "dynamical dimensional reduction" of spacetime: the familiar four-dimensional picture of spacetime in the IR is replaced by a quantum picture with an effective number of spacetime dimensions smaller than four in the UV.

This phenomenon has been studied mostly in terms of the spectral dimension, that provides a valuable characterization of classical Riemannian geometries[3], but its proposed applicability to a quantum spacetime involves some adaptations that are arguably responsible for some of its inadequacies.

When the IR Hausdorff dimension of spacetime is $D+1$, and the Euclidean d'Alembertian of the theory is represented on momentum space as $\mathcal{C}_p^{Euc}(E,p)$, the spectral dimension is computed via $d_S(s) = -2\frac{d\ln P(s)}{d\ln s}$ where $P(s)$ is given by $P(s) \propto \int dE\, dp\, p^{D-1}\, e^{-s\,\mathcal{C}_p^{Euc}(E,p)}$. Reasons to be concerned can be found in the use done of Euclidean d'Alembertian (it is in fact well known that the Euclidean version of a quantum-gravity model can be profoundly different from the original model in Lorentzian spacetime[4]), in the role played by off-shell modes and in the

invariance of $P(s)$ under active diffeomorphisms on momentum space (an active diffeomorphism on momentum space amounts to an irrelevant change of integration variable for $P(s)$). Since an active diffeomorphism can map a given physical theory into a very different one, this degeneracy of the spectral dimension is worrisome.

While these concerns are very serious, it must be acknowledged that several analyses centered on the spectral dimension give rather meaningful results. Therefore, the guiding idea is that it is necessary to replace the spectral dimension with some other fully physical notion of dimensionality of a quantum spacetime, with the requirement that in most cases the new notion should agree with the spectral dimension. Only when the unphysical content of the spectral dimension plays a particularly significant role should the new notion differ significantly from the spectral dimension.

The guidance adopted in searching for such a new notion is the observation, reported in recent studies[5–8], that in some instances the Stefan-Boltzmann law and the equation of state (EOS) parameter give indications on the dimensionality of spacetime that are consistent with the spectral dimension. One can view the Stefan-Boltzmann law and EOS parameter as indicators of spacetime dimensionality since for a gas of radiation in a classical spacetime with $D+1$ dimensions they take the form $U \propto T^{D+1}$ and $w \equiv \frac{p}{\rho} = \frac{1}{D}$. These observations inspire the proposal of assigning a "thermal dimension" to a quantum spacetime.

Consider a class of generalized Hořava-Lifshitz (HL) scenarios[9,10]. These are cases where the momentum-space representation of the deformed d'Alembertian takes the form $\mathcal{C}_{\gamma_t\gamma_x}(E,p) = E^2 - p^2 + \ell_t^{2\gamma_t} E^{2(1+\gamma_t)} - \ell_x^{2\gamma_x} p^{2(1+\gamma_x)}$, where $E$ is the energy, $p$ is the modulus of the spatial momentum, $\gamma_t$ and $\gamma_x$ are dimensionless parameters, and $\ell_t$ and $\ell_x$ are parameters with dimension of length (usually assumed to be of the order of the Planck length). For this model it is known[10] that the UV value of the spectral dimension, obtained from the Euclidean version of the above d'Alembertian ($E^2 + p^2 + \ell_t^{2\gamma_t} E^{2(1+\gamma_t)} + \ell_x^{2\gamma_x} p^{2(1+\gamma_x)}$), is $d_S(0) = \frac{1}{1+\gamma_t} + \frac{D}{1+\gamma_x}$.

In deriving the thermal dimension for this case one can start from the logarithm of the thermodynamical partition function, written as

$$\log Q_{\gamma_t\gamma_x} = -\frac{2V}{(2\pi)^3} \int dE\, d^3p \left[ \delta(\mathcal{C}_{\gamma_t\gamma_x})\, \Theta(E) \cdot \right. \\ \left. \cdot\, 2E \log\left(1 - e^{-\beta E}\right) \right]. \tag{1}$$

where $\beta = \frac{1}{k_B T}$, and the delta function $\delta(\mathcal{C}_{\gamma_t\gamma_x})$ enforces the on-shell relation $\mathcal{C}_{\gamma_t\gamma_x} = 0$. From (1) one obtains the energy and pressure densities respectively as $\rho_{\gamma_t\gamma_x} \equiv -\frac{1}{V}\frac{\partial}{\partial\beta} \log Q_{\gamma_t\gamma_x}$ and $p_{\gamma_t\gamma_x} \equiv \frac{1}{\beta}\frac{\partial}{\partial V} \log Q_{\gamma_t\gamma_x}$. For the UV/high-temperature values of $\rho_{\gamma_t\gamma_x}$ and $w_{\gamma_t\gamma_x}$ one can easily establish the following behaviours at high temperature,

$$\rho_{\gamma_t\gamma_x} \propto T^{1+3\frac{1+\gamma_t}{1+\gamma_x}}\,, \quad w_{\gamma_t\gamma_x} = \frac{1+\gamma_x}{3(1+\gamma_t)}\,. \tag{2}$$

One sees that both of these results give a consistent prediction for the "thermal dimension" at high temperature, which is

$$d_T = 1 + 3\frac{1+\gamma_t}{1+\gamma_x}\,. \tag{3}$$

In the context of generalized HL scenarios it is simple to show (see Ref. 1 for more details) that while the spectral dimension is insensitive to the difference between active and passive diffeomorphisms on momentum space, the thermal dimension takes generally different values for different theories connected by an active diffeomorphism. Since physics is not invariant under such diffeomorphisms, this difference of sensitivity gives the thermal dimension a stronger discrimination power than the spectral dimension.

Another case which might deserve special interest from the quantum-gravity perspective, as stressed in Ref. 11, is that in which deformed d'Alembertian takes the form $\mathcal{C}_\gamma(E,p) = E^2 - p^2 - \ell^{2\gamma}\left(E^2-p^2\right)^{1+\gamma}$ . For this case one easily finds that the UV spectral dimension is $d_S(0) = \frac{4}{1+\gamma}$, but the fact that the UV dimensionality of spacetime depends on $\gamma$ is puzzling and points very clearly to the type of inadequacies of the spectral dimension previously mentioned. In fact, in the UV limit the parameter $\gamma$ has no implications for the on-shell/physical properties of the (massless) theory. In general, massless particles governed by $\mathcal{C}_\gamma$ will be on-shell only either when $E^2 = p^2$ or when $E^2 = p^2 + \frac{1}{\ell^2}$ , independently of the value of $\gamma$. Also, in the UV limit the two possibilities become indistinguishable, all particles are governed by $E \simeq p$ just like in any 4-dimensional spacetime. So without any need to resort to complicated analyses one knows that this theory in the UV limit must behave like a 4-dimensional theory, in contradiction with the mentioned result for the UV spectral dimension.

Indeed, exact computation confirms that the value of the thermal dimension is 4, as it is computed strictly on-shell. The disastrous failures of the spectral dimension in this case is to be attributed to a combination of its sensitivity to off-shell properties and its reliance on the Euclidean d'Alembertian. It is noteworthy that for the Euclidean d'Alembertian $\mathcal{C}_\gamma^{Euc} = E^2 + p^2 + \ell^{2\gamma}(E^2+p^2)^{1+\gamma}$ , in the UV limit one can neglect $E^2 + p^2$ with respect to $\ell^{2\gamma}(E^2+p^2)^{1+\gamma}$. Instead for on-shell modes of the original Lorentzian $\mathcal{C}_\gamma$ one can never neglect $E^2 - p^2$ with respect to $\ell^{2\gamma}(E^2-p^2)^{1+\gamma}$.

In summary, it should be evidently seen as advantageous for the thermal dimension the fact that it assigns different UV dimension to very different theories (connected by an active diffeomorphism) and not being dominated by off-shell modes.

## 2. Primordial perturbations in a rainbow universe with running Newton constant

It is interesting to investigate the implications that the proposed modified themrodynamics could have in certain models of primordial Universe. In fact, recent

results suggest that the properties of primordial cosmological fluctuations might be a consequence of quantum-gravitational effects, which are relevant in the early universe[12,13], with no need for inflation. In particular it has been shown[12,14–18] that a scale invariant power spectrum can be obtained in those theories for which the dimension of spacetime runs to 2. These results rely on rigid assumptions[a] that are can be relaxed in order to find a mechanism that would produce the observed small departure from exact scale invariance[2].

We assume that gravity is described by an energy-dependent metric as proposed in Ref. 19, an energy-dependent Newton constant[b] and we consider both perturbations of quantum origin for a vacuum state and perturbations that are originated in a thermal state[24–28]. In the latter case it is assumed that the universe is filled with radiation and that both the background and the fluctuations are thermalized, so that they share the same (modified) thermodynamical properties[29].

For a deformed dispersion relation $E^2 = p^2(1 + (\ell p)^\gamma)$ the associated rainbow metric is $ds^2 = dt^2 - \frac{a^2(t)}{1+(\ell p)^\gamma}\delta_{ij}dx^i dx^j$. After lengthy but straightforward computations[2] one finds the evolution equation for perturbations modes $v$

$$v'' - \left(\left(\ell\frac{k}{a}\right)^{2\gamma} k^2 + \frac{a''}{a}\right) v = 0\,. \tag{4}$$

where $\frac{d}{d\eta} \equiv a\frac{d}{dt}$ and $k = ap$ is the comoving momentum. If the Newton constant has a power-law dependence on energy in the UV regime, $G(E) = \ell^2(\ell E)^\alpha \sim \ell^2\left(\frac{\ell k}{a}\right)^{(1+\gamma)\alpha}$, then $\alpha$ must satisfy

$$-\frac{2+\gamma}{1+\gamma} < \alpha < \frac{\gamma-2}{1+\gamma} \tag{5}$$

in order to solve the horizon paradox and having cosmological expansion (with positive conformal time $\eta$). The spectral index of vacuum perturbations is

$$n_s^V - 1 = \frac{(\gamma+4)(2-\gamma)}{2-\gamma+\alpha(1+\gamma)}\,. \tag{6}$$

Clearly $\gamma = 2$ gives a scale invariant power spectrum for any value of $\alpha$ allowed by the constraint[c] (5). One can require $n_s^V$ to match the present observational constraint from Planck[30] $n_s^V = 0.968 \pm 0.006$, obtaining the allowed range of values shown in Fig. 1 of Ref. 2. The values of $\alpha$ that are selected by observational constraints are all negative, suggesting a vanishing Newton constant in the deep UV regime. On the other hand, from the modified Hubble equation[2] one can see that observational constraints allow for both an accelerated or decelerated expansion.

[a]Such as that the second order action for perturbations is the one of Einstein gravity and that the perturbations are produced in a quantum vacuum state.

[b]This is motivated by results in Hořava-Lifshitz gravity and in Asymptotic Safety[20–23], where the Newton constant tends to zero at super-Planckian energies.

[c]The fact that scale invariance is achieved independently of how the Newton constant scales with energy is due to the time perturbations being already scale-invariant and proportional to the scale factor $a$ inside the horizon.

Including the thermal contribution, the spectral index of perturbations becomes

$$n_s^T = n_s^V - 1 - \gamma. \tag{7}$$

Using the value of the vacuum spectral index found in the previous section, Eq. (6), the thermal spectral index can be written as

$$n_s^T = \frac{4(2-\gamma) - \alpha\gamma\,(1+\gamma)}{2-\gamma+\alpha(1+\gamma)}\,. \tag{8}$$

For $\alpha \neq 0$, asking that the perturbations are scale invariant leads to a constraint linking $\alpha$ and $\gamma$. Asking in addition that the horizon problem is solved, Eq. (5), introduces an inferior bound $\gamma > 2$ on the allowed values of $\gamma$. Then the values of $\alpha$ that are compatible with scale invariance and which allow to solve the horizon problem fall in the range $-1/4 < \alpha < 0$. It is also possible to match the spectral index to the Planck observed value giving the constraints shown in Fig. 2 of Ref. 2. According to he modified Hubble equation[2], these observational constraint on $\alpha$ and $\gamma$ only allow for a decelerating expansion of the universe.

## 3. Concluding remarks

The exciting realization that the UV dimension of spacetime might be different from its IR dimension adds significance to the old challenge of describing the dimension of a quantum spacetime and it is argued that it is crucial to link this issue to observable properties. The notion of thermal dimension reviewed here is free from the shortcomings of the spectral dimension, since it relies on the analysis of observable thermodynamical properties of radiation in the quantum spacetime. We showed how it is not only physical but also particularly useful, at least for studies of the early universe, which is anyway the context where the UV dimension of spacetime should find its most significant applications[12,17].

We saw how a rainbow universe with running Newton constant can accommodate primordial perturbations whose spectral index matches current constraints, without relying on inflation to solve the horizon problem. Crucially, it has been assumed a power-law dependence of the Newton constant on energy and that the background satisfies the thermodynamical relations peculiar to radiation subject to deformed dispersion relations. For both kinds of initial conditions for the perturbations (vacuum and thermal) the running of the Newton constant is essential in achieving a viable picture. In particular, the Newton constant is constrained to be decreasing with energy in the ultraviolet regime. This is consistent with intuition from quantum gravity theories, such as Hořava-Lifshitz gravity and Asymptotic safety. It also resonates with the conjecture put forward in Ref. 17.

## References


1. G. Amelino-Camelia, F. Brighenti, G. Gubitosi and G. Santos,
2. F. Brighenti, G. Gubitosi, J. Magueijo, *Phys. Rev.* **D 95**, 063534 (2017); *Phys. Lett.* **B 767**, 58 (2017).
3. D. Benedetti, *Phys. Rev. Lett.* **102** (2009) 111303.
4. S. Carlip, *Class. Quant. Grav.* **32** (2015) 23, 232001.
5. V. Husain, S. S. Seahra and E. J. Webster, *Phys. Rev.* **D 88** (2013) 2, 024014.
6. G. Santos, G. Gubitosi and G. Amelino-Camelia, *JCAP* **1508** (2015) 08, 005.
7. K. Nozari, V. Hosseinzadeh and M. A. Gorji, *Phys. Lett.* **B 750** (2015) 218.
8. J. J. Atick and E. Witten, *Nucl. Phys.* **B 310** (1988) 291.
9. P. Hořava, *Phys. Rev. Lett.* **102** (2009) 161301.
10. G. Amelino-Camelia, M. Arzano, G. Gubitosi and J. Magueijo, *Phys. Lett.* **B 736** (2014) 317.
11. R. Percacci, In Oriti, D. (ed.): *Approaches to quantum gravity* 111-128.
12. G. Amelino-Camelia, M. Arzano, G. Gubitosi and J. Magueijo, *Phys. Rev.* **D 87** (2013) 12, 123532.
13. S. Mukohyama, *JCAP* **0906** (2009) 001.
14. G. Amelino-Camelia, M. Arzano, G. Gubitosi and J. Magueijo, *Phys. Rev.* **D 88** (2013) no. 4, 041303.
15. G. Amelino-Camelia, M. Arzano, G. Gubitosi and J. Magueijo, *Phys. Rev.* **D 88** (2013) no. 10, 103524.
16. G. Amelino-Camelia, M. Arzano, G. Gubitosi and J. Magueijo, *Phys. Lett.* **B 736** (2014) 317.
17. G. Amelino-Camelia, M. Arzano, G. Gubitosi and J. Magueijo, *Int. J. Mod. Phys.* **D 24** (2015) 12, 1543002.
18. M. Arzano, G. Gubitosi, J. Magueijo and G. Amelino-Camelia, *Phys. Rev.* **D 91** (2015) 12, 125031.
19. J. Magueijo, L. Smolin, *Class. Quant. Grav.* **21** (2004) 1725-1736.
20. D. Benedetti and F. Guarnieri, *JHEP* **1403** (2014) 078.
21. G. D'Odorico, J. W. Goossens and F. Saueressig, *JHEP* **1510** (2015) 126.
22. O. Lauscher and M. Reuter, *Class. Quant. Grav.* **19** (2002) 483.
23. M. Niedermaier and M. Reuter, *Living Rev. Rel.* **9** (2006) 5.
24. T. Biswas, R. Brandenberger, T. Koivisto and A. Mazumdar, *Phys. Rev.* **D 88** (2013) no. 2, 023517.
25. P. Ferreira and J. Magueijo, *Phys. Rev.* **D 78** (2008) 061301.
26. J. Magueijo and L. Pogosian, *Phys. Rev.* **D 67** (2003) 043518.
27. J. Magueijo and P. Singh, *Phys. Rev.* **D 76** (2007) 023510.
28. S. Koh and R. H. Brandenberger, *JCAP* **0706** (2007) 021.
29. J. Magueijo and J. Noller, *Phys. Rev.* **D 81** (2010) 043509.
30. P. A. R. Ade *et al.* [Planck Collaboration], *Astron. Astrophys.* **594** (2016) A13.

# Counting initial conditions and degrees of freedom in nonlocal gravity

Gianluca Calcagni

*Instituto de Estructura de la Materia, CSIC, Serrano 121, 28006 Madrid, Spain*
*E-mail: g.calcagni@csic.es*

Nonlocal quantum gravity is a class of fundamental theories whose classical and quantum dynamics is specified by "form factors", operators with infinitely many derivatives. After briefly reviewing this paradigm and its role in the resolution of big-bang and black-hole singularities, we count the number of nonpertubative field degrees of freedom as well as the number of initial conditions to be specified to solve the Cauchy problem. In particular, in four dimensions and for the string-related form factor, there are 8 degrees of freedom (2 graviton polarization modes, which propagate, and 6 nonpropagating rank-2 tensor modes) and 4 initial conditions. The method to obtain this result is illustrated for the case of a nonlocal scalar field.

## 1. Introduction

"Nonlocal quantum gravity" is an umbrella name including at least two different settings. The first group, which includes general relativity, consists in classically local gravitational theories which receive quantum corrections such that the effective one-loop action is nonlocal (i.e., it is made of operators with infinitely many derivatives). In this brief review of recent results, we will confine ourselves to the second meaning of the term, where nonlocality is fundamentally present already at the classical level and, thanks to the suppression of the graviton propagator in the ultraviolet (UV), the theory is renormalizable or finite. Nowadays, nonlocal quantum gravity has achieved a high degree of independence both from these antecedents and from other proposals, to the point where it can be considered as one of the most promising and accessible candidates for a theory where the gravitational force consistently obeys the laws of quantum mechanics. In particular, there exist several renormalization results, both at finite order and to all orders in a perturbative Feynman-diagram expansion, which showed that the good UV properties guessed at the level of power-counting indeed hold rigorously (e.g., Ref. 1).

Despite the investment of much effort in taming fundamental nonlocality, several questions remain open to date: (i) Is the Cauchy problem well-defined? (ii) If so, how many initial conditions must one specify for a solution? (iii) How many degrees of freedom are there? (iv) How to construct nontrivial solutions? (iv) Is causality violated? (v) Are singularities resolved at the classical or quantum level?

Here we will give the following answers to some of these issues: (i) Yes, for the form factors appearing in fundamental theories (not for all conceivable form factors). (ii) Two or higher for a scalar field theory and four or higher for gravity (depending on the form factor), but finite. (iii) Eight (in $D = 4$ dimensions), but only the graviton modes propagate. (iv) Via the diffusion method. Let us now examine where these cryptic responses come from. The main results can be found in Refs. 2, 3.

## 2. Action and form factors

The classical fundamental (not effective) action of the theory is

$$S = \frac{1}{2\kappa^2}\int d^D x\sqrt{-g}\,[R + G_{\mu\nu}\gamma(\Box)R^{\mu\nu}]\ , \tag{1}$$

where $G_{\mu\nu}$ is the Einstein tensor and $\gamma(\Box)$ is a weakly or quasi-polynomial nonlocal form factor. Formally, $\gamma$ is an analytic function that can be expressed as an infinite series with infinite convergence radius, $\gamma(\Box) = \sum_n c_n\Box^n$, although this expansion does not actually span the whole space of solutions in general.[4] As long as we require good properties at the quantum level (in particular, locality of the counterterms), the coefficients $c_n$ can be selected in a subclass of entire functions having special asymptotic properties.[1,5–9] There are several form factors that preserve perturbative unitarity in (1) (see Table 1). In general, we can parametrized $\gamma$ as

$$\gamma(\Box) = \frac{e^{H(\Box)} - 1}{\Box}\,, \qquad H(\Box) := \alpha\int_0^{P(\Box)} d\omega\,\frac{1 - f(\omega)}{\omega}\,, \tag{2}$$

where $H(\Box)$ depends on the dimensionless combination $l^2\Box$ and $l$ is a fixed length scale. The profile $H(\Box)$ can be defined through an integral where $\alpha > 0$ is real and $P(\Box)$ is a generic function of $l^2\Box$. The parameter $\alpha$ will not play any important role in what follows, but we included it to reproduce some form factors in the literature.

Table 1. Form factors in nonlocal gravity.

| $H(\Box)$ | $P(\Box)$ | $f(\omega)$ | Form factor name |
|---|---|---|---|
| $H^{\rm pol}(\Box) := \alpha\{\ln P(\Box) + \Gamma[0, P(\Box)] + \gamma_{\rm E}\}$ | $-l^2\Box$<br>$O(\Box^n)$ | $e^{-\omega}$ | Kuz'min[6]<br>Tomboulis[5,8] |
| $H^{\rm exp}(\Box) := \alpha P(\Box)$ | $-l^2\Box$<br>$l^4\Box^2$ | $1 - \omega$ | string-related[10]<br>Krasnikov[11] |

All these form factors share the common property of blowing up in the UV in momentum space. In general, this implies asymptotic freedom, i.e., interactions are subdominant at short scales.

## 3. The wild beast of nonlocality

Consider the scalar field theory on Minkowski spacetime

$$S_\phi = \int d^D x\,\left[\frac{1}{2}\phi\Box\gamma(\Box)\phi - V(\phi)\right]\ , \tag{3}$$

where $V$ is an interaction potential. How could one explore its classical dynamics?

As a first attempt, one can try to truncate the nonlocal operator up to some finite order, $\gamma(\Box) \simeq \sum_{n=0}^{N} c_n\Box^n = c_0 + c_1\Box + \cdots + c_N\Box^N$. However, the resulting

finite-order dynamics is physically inequivalent to the original one and there is no smooth transition between them. The free ($V = 0$) case illustrates the point well:

$$\begin{aligned} \gamma(\Box) = e^{-l^2\Box} &\Rightarrow \text{dispersion relation:} \quad -k^2 e^{l^2 k^2}\phi_k = 0 \\ &\Rightarrow \text{propagator:} \quad -\frac{e^{-l^2k^2}}{k^2} \quad \Rightarrow \quad 1 \text{ DOF}\,, \qquad (4) \\ g(\Box) \simeq 1 - l^2\Box &\Rightarrow \text{dispersion relation:} \quad -k^2(1+l^2k^2)\phi_k = 0 \\ &\Rightarrow \text{propagator:} \quad -\frac{1}{k^2} + \frac{1}{l^{-2}+k^2} \quad \Rightarrow \quad 2 \text{ DOF, } 1 \text{ ghost}\,. \qquad (5) \end{aligned}$$

This example provides a good occasion to comment also on how to determine the number of field degrees of freedom (DOF). In the free case, there is only one double pole, corresponding to 1 DOF. This can also be seen by making a nonlocal field redefinition $\tilde{\phi} := \sqrt{\gamma(\Box)}\phi$, so that the free Lagrangian reads $\tilde{\mathcal{L}}_\phi = (1/2)\tilde{\phi}\Box\tilde{\phi}$. This operation is safe if $\gamma$ is an entire function; if $\gamma$ is not entire (for instance, $\gamma = \Box^{-n}$), then the field redefinition may result in the elimination of physical modes or the introduction of spurious ones. The Lagrangian $\tilde{\mathcal{L}}_\phi$ is second-order in spacetime derivatives and features one local field $\tilde{\phi}$, hence there is only 1 DOF and solutions are specified by two initial conditions. However, when $V$ is nonlinear of cubic or higher order the field redefinition does not absorb nonlocality completely and one is left with a possibly intractable problem, with extra nonperturbative degrees of freedom[12] and an infinite tower of Ostrogradsky modes.

In fact, the Cauchy problem can be naïvely stated as the assignment of an infinite number of values at some initial time $t_{\rm i}$,

$$\phi(t_{\rm i}),\ \dot{\phi}(t_{\rm i}),\ \ddot{\phi}(t_{\rm i}),\ \dddot{\phi}(t_{\rm i}),\ \ldots\,. \qquad (6)$$

Thus, paradoxically, we can solve the dynamics only if we already know the solution:[13]

$$\phi(t) = \sum_{n=0}^{+\infty} \frac{\phi^{(n)}(t_{\rm i})}{n!}(t-t_{\rm i})^n\,. \qquad (7)$$

If we do not specify all the initial conditions, the solution may be non-unique.

Finally, a word on how to rewrite nonlocalities as a convolution. It is well known that infinitely many derivatives can be traded for integrated kernel functions:

$$\gamma(\Box)\,\phi(x) = \int d^D y\, F(y-x)\,\phi(y)\,. \qquad (8)$$

However, by itself this operator bears no practical advantage. Hiding infinitely many derivatives into integrals does not help in solving the Cauchy problem, *unless* the kernel $F$ could be found by solving some auxiliary, finite-order differential equations. This is precisely the leverage point we will focus on.

## 4. Diffusion method

The diffusion method was proposed some years ago[4,12,14,15] to solve nonlocal scalar field theories with exponential form factor (4), a very specific nonlocal operator that arises in string theory. By trading nonlocal operators with shifts in a fictitious extra direction $r$, the method allows one to count the number of DOF and of initial conditions (which are finite) and to find nonperturbative solutions. All these features can be easily illustrated by the scalar field theory

$$S_\phi = \int d^D x \, \left[ \frac{1}{2} \phi(x) \Box e^{-l^2 \Box} \phi(x) - V(\phi) \right] , \tag{9}$$

where $l^2$ is a constant. The equation of motion is

$$\Box e^{-l^2 \Box} \phi - V'(\phi) = 0 \, . \tag{10}$$

Define now a *localized system*, *a priori* independent of (9), living in $D+1$ dimensions and featuring two scalars $\Phi(r,x)$ and $\chi(r,x)$:

$$\mathcal{S}[\Phi,\chi] = \int d^D x \, dr \, (\mathcal{L}_\Phi + \mathcal{L}_\chi) \, , \tag{11}$$

$$\mathcal{L}_\Phi = \frac{1}{2} \Phi(r,x) \Box \Phi(r - l^2, x) - V[\Phi(r,x)] \, , \tag{12}$$

$$\mathcal{L}_\chi = \frac{1}{2} \int_0^{l^2} dq \, \chi(r-q,x)(\partial_{r'} - \Box) \Phi(r',x), \qquad r' = r + q - l^2. \tag{13}$$

The equations of motion are

$$0 = (\partial_r - \Box)\Phi(r,x), \qquad 0 = (\partial_r - \Box)\chi(r,x), \tag{14}$$

$$0 = \frac{1}{2}[\Box\Phi(r-l^2,x) + \chi(r-l^2,x)] + \frac{1}{2}[\Box\Phi(r+l^2) - \chi(r+l^2)] - V'[\Phi(r,x)]. \tag{15}$$

The first line is telling us that the fields are diffusing along the extra direction. At this point, one assumes that there exists a constant $\beta$ such that the equation of motion (15) coincides with the one of the nonlocal system (9), Equation (10), on the slice $r = \beta l^2$ (the *physical slice*). This is achieved provided the following conditions hold:

$$\Phi(\beta l^2, x) = \phi(x), \qquad \chi(\beta l^2, x) = \Box \Phi(\beta l^2, x) \, . \tag{16}$$

The conclusion is that the localized system has 4 initial conditions $\Phi(r, t_\mathrm{i}, \mathbf{x})$, $\dot{\Phi}(r, t_\mathrm{i}, \mathbf{x})$, $\chi(r, t_\mathrm{i}, \mathbf{x})$, $\dot{\chi}(r, t_\mathrm{i}, \mathbf{x})$ and 2 field DOF $\Phi$ and $\chi$. On the physical slice, because of (16) the number of DOF reduces to 1 and the initial conditions are on the field $\phi$ and its first two derivatives (the initial conditions of $\chi$ are not independent):

$$\phi(t_\mathrm{i}, \mathbf{x}), \quad \dot{\phi}(t_\mathrm{i}, \mathbf{x}) \, . \tag{17}$$

With traditional methods, only *perturbative* solutions of the linearized EOM or what we call "static" (in the extra direction $r$) solutions are available to inspection. By this name, we mean solutions where nonlocality is, in one way or

another, trivialized, such as when $\Box\phi = \lambda\phi$. In contrast, the diffusion method gets access to nonperturbative solutions valid in the presence of nonlinear interactions and nontrivial nonlocality. These solutions are, in general, only approximate, and are found by searching for the value of $\beta$ minimizing the equations of motion[4,15,16]. Examples are: $\phi(t) = \text{Kummer}(t)$ on a Friedmann–Robertson–Walker (FRW) background; rolling tachyon $\phi(t) = \sum_n a_n e^{nt}$ in string field theory, $V = \phi^3$; kink $\phi(x) = \text{erf}(x)$, $V = \phi^3$; $\phi(t) = \gamma(\alpha, t)$ (incomplete gamma function), $V = \phi^n$ on FRW background; instanton $\phi(x) = \text{erf}(x)$, $V = \phi^4$ (brane tension recovered at 99.8% level); kink $\phi(x) = \text{erf}(x)$, $V = (e^{\Box}\phi^2)^2$; various profiles $\phi(t)$ in bouncing and singular cosmologies.

The main reason why diffusion works is that nonlocal operators are represented as a shift in an extra direction rather than as an infinite sum of derivatives. The latter representation does not span the whole space of solutions, as one can see by a toy example.[4] Consider a $D = 4$ FRW background with Hubble expansion $H = \dot{a}/a = H_0/t$, the Laplace–Beltrami operator $\Box = -\partial_t^2 - 3H\partial_t$, and the homogeneous power-law profile $\phi(t) = t^p$. If we try to calculate the object $e^{r\Box}\phi$ as a series, the result diverges: $e^{r\Box}\phi = \sum_{n=0}^{\infty}(r\Box)^n\phi/n! = \infty$. On the other hand, with the diffusion method one interprets $\phi(t) = \Phi(t, 0)$ as the initial condition in the diffusion scale $r$ and the profile $e^{r\Box}\Phi(t, 0) = \text{Kummer}(t, r)$ is a linear superposition of well-defined Kummer functions.

## 5. Initial conditions and degrees of freedom

The diffusion method has been extended to the case of gravity in Ref. 2 for the string-related and Krasnikov exponential form factors and in Ref. 3 for the asymptotically polynomial (Kuz'min and Tomboulis–Modesto) form factors. The reader can consult those papers for technical details; here we only quote the bottom line, which is that, for the string-related form factor, the localized system associated with (1) has 6 initial conditions $g_{\mu\nu}(t_\text{i}, \mathbf{x})$, $\dot{g}_{\mu\nu}(t_\text{i}, \mathbf{x})$, $\Phi_{\mu\nu}(r, t_\text{i}, \mathbf{x})$, $\dot{\Phi}_{\mu\nu}(r, t_\text{i}, \mathbf{x})$, $\chi_{\mu\nu}(r, t_\text{i}, \mathbf{x})$, $\dot{\chi}_{\mu\nu}(r, t_\text{i}, \mathbf{x})$, two for each rank-2 symmetric tensor field (the metric $g_{\mu\nu}$ and the tensors $\Phi_{\mu\nu}$ and $\chi_{\mu\nu}$). Since, on the physical slice, $\Phi_{\mu\nu}(\beta l^2, x) = G_{\mu\nu}$ and $\chi_{\mu\nu}(\beta l^2, x) = R_{\mu\nu}$, the initial conditions on these fields are not independent. However, while $\chi_{\mu\nu}$ is an auxiliary field of the localized system and depends on the dynamical degrees of freedom of the nonlocal system, $\Phi_{\mu\nu}$ is an auxiliary field already at the level of nonlocal dynamics and it encodes the two derivatives hidden in the Ricci tensor and scalar. Therefore, the solutions of the nonlocal system (1) are characterized by 4 initial conditions:

$$g_{\mu\nu}(t_\text{i}, \mathbf{x}), \quad \dot{g}_{\mu\nu}(t_\text{i}, \mathbf{x}), \quad \ddot{g}_{\mu\nu}(t_\text{i}, \mathbf{x}), \quad \dddot{g}_{\mu\nu}(t_\text{i}, \mathbf{x})\,. \tag{18}$$

Regarding the degrees of freedom, the counting for the exponential form factor is the following. (i) Graviton $g_{\mu\nu}$: symmetric $D \times D$ matrix with $D(D+1)/2$ independent entries, to which one subtracts $D$ Bianchi identities $\nabla^\mu G_{\mu\nu} = 0$ and $D$ diffeomorphisms (the theory is fully diffeomorphism invariant). Total: $D(D-3)/2$.

In $D = 4$, there are 2 degrees of freedom. (ii) Tensor $\phi_{\mu\nu}$: symmetric $D \times D$ matrix with $D(D+1)/2$ independent entries, to which one subtracts $D$ transverse conditions $\nabla^\mu \phi_{\mu\nu} = 0$. Total: $D(D-1)/2$. In $D = 4$, there are 6 degrees of freedom. Similar results hold for the asymptotically polynomial Kuz'min form factor, although in that case the diffusion method requires more elaboration.[3]

The grand total is $D(D-2)$. In $D = 4$, there are 8 DOF. Two of them (the graviton) are visible already at the perturbative level, while the other 6 are of nonperturbative origin. Their role in phenomenology[3] has been determined only recently.[17] It was shown that the extra $D(D-1)/2$ tensor degrees of freedom do not propagate on Ricci-flat backgrounds, at any perturbative order.

## 6. Conclusions

The number of degrees of freedom and of initial conditions of fundamentally nonlocal gravitational theories with "well-behaved" form factors is finite. In the diffusion method, infinitely many initial conditions are traded for boundary conditions in an extra direction. Solving the diffusion equation and algebraic relations is way simpler than solving nonlocal equations. By making sense of the Cauchy problem in this class of theories, the doors of classical top-down phenomenology may open up.

## Acknowledgment

The author is supported by the I+D grant FIS2017-86497-C2-2-P.

## References

1. Modesto, L.; Rachwał, L. *Nucl. Phys. B* **889**, 228 (2014); *Nucl. Phys. B* **900**, 147 (2015).
2. Calcagni, G.; Modesto, L.; Nardelli, G. *JHEP* **1805**, 087 (2018).
3. Calcagni, G.; Modesto, L.; Nardelli, G. `arXiv:1803.07848`.
4. Calcagni, G.; Montobbio, M.; Nardelli, G. *Phys. Rev. D* **76**, 126001 (2007).
5. Tomboulis, E.T. `arXiv:hep-th/9702146`.
6. Kuz'min, Y.V. Yad. Fiz. 50 (1989) 1630 [Sov. J. Nucl. Phys. 50 (1989) 1011].
7. Asorey, M.; López, J.L.; Shapiro, I.L. *Int. J. Mod. Phys. A* **12**, 5711 (1997).
8. Modesto, L. *Phys. Rev. D* **86**, 044005 (2012).
9. Tomboulis, E.T. *Mod. Phys. Lett. A* **30**, 1540005 (2015).
10. Biswas, T.; Mazumdar, A.; Siegel, W. *JCAP* **0603**, 009 (2006); Calcagni G.; Modesto, L. *Phys. Rev. D* **91**, 124059 (2015).
11. Krasnikov, N.V. *Theor. Math. Phys.* **73**, 1184 (1987) [*Teor. Mat. Fiz.* **73**, 235 (1987)].
12. Calcagni, G.; Montobbio, M.; Nardelli, G. *Phys. Lett. B* **662**, 285 (2008).
13. N. Moeller and B. Zwiebach, *JHEP* **0210**, 034 (2002).
14. Mulryne, D.J.; Nunes, N.J. *Phys. Rev. D* **78**, 063519 (2008).
15. Calcagni, G.; Nardelli, G. *JHEP* **1002**, 093 (2010).

16. Calcagni, G.; Nardelli, G. *Phys. Rev. D* **78**, 126010 (2008); *Phys. Lett. B* **669**, 102 (2008); *Nucl. Phys. B* **823**, 234 (2009); *Int. J. Mod. Phys. D* **19**, 329 (2010); *Phys. Rev. D* **82**, 123518 (2010).
17. Briscese, F.; Calcagni, G.; Modesto, L. *Phys. Rev. D* to appear [arXiv: 1901.03267].

# Emerging Hawking-like radiation in gravitational scattering beyond the Planck scale

Dimitri Colferai

*Department of Physics and Astronomy, University of Florence, Sesto Fiorentino, 50019, Italy*
**E-mail: colferai@fi.infn.it*

We generalize the semiclassical treatment of graviton radiation to gravitational scattering at very large energies $E \gg M_P$ and finite scattering angles $\theta_s$, so as to approach the collapse regime. Our basic tool is the extension of the recently proposed, unified form of radiation to the Amati-Ciafaloni-Veneziano reduced-action model. By resumming eikonal scattering diagrams, we are able to derive the corresponding (unitary) coherent-state operator. The resulting graviton spectrum, tuned on the gravitational radius $R$, fully agrees with previous calculations for small angles $\theta_s \ll 1$ but, for sizeable angles $\theta_s \sim 1$ acquires an exponential cutoff of the large frequency region ($\omega > 1/R$), due to energy conservation, so as to emit a finite fraction of the total energy. In the approach-to-collapse regime we find a radiation enhancement due to large tidal forces, so that the whole energy is radiated off, with a large multiplicity $\langle N \rangle \sim Gs \gg 1$ and a well-defined frequency cutoff of order $1/R$. The latter corresponds to the Hawking temperature for a black hole of mass somewhat smaller than $E$.

*Keywords*: Gravitational scattering; Gravitational radiation; Quantum gravity.

## 1. Introduction

One of the main unsolved problems in physics is to reconcile quantum mechanics (QM) with general relativity (GR). Its resolution is a formidable task, but important informations can be obtained by investigating some processes lying at the interface between QM and GR with the theories and tools at our disposal.

This work aims at studying gravitational scattering of particles at very high energies $E \gg M_p$ beyond the Planck scale, in order to investigate graviton bremmsstrahlung and possibly macroscopic black-hole (BH) formation and evaporation at (semiclassical) quantum level, and therefore to shed light on the so-called "information paradox". What we would like to understand is whether a semiclassical picture of collapse emerges from the quantum theory, or if something different happens.

In this talk I present a model that describes gravitational processes in trans-planckian particle collisions. I shall show you that a quantum description of particle scattering at extreme energies gives rise to graviton radiation

- whose properties are similar to Hawking radiation, thus suggesting a sort of gravitational collapse.
- but can still be described by a unitary S-matrix, thus pointing towards a resolution of the information paradox.

## 2. The Eikonal Model

Following the pioneering works of Amati-Ciafaloni-Veneziano (ACV)[1] and others, we consider two massless particles (which could be strings) colliding at center of mass energy much larger than the Planck scale $2E = \sqrt{s} \gg M_P$ with some impact parameter $b$. Such a simple initial state allows us to use the formalism of QM, while the condition on the energy should imply the formation of a macroscopic BH of radius $R = 4GE$, according to GR.

In this high-energy regime the adimensional parameter $\alpha_G \equiv Gs/\hbar \gg 1$ is very large and plays the role of effective gravitational coupling. It also measures the typical action of such processes, in units of $\hbar$.

Actually, we are interested in the semiclassical regime where both $R, b \gg \lambda_P, \lambda_s$ (the latter being the Planck and string lengths, respectively).

### 2.1. *Elastic scattering*

When $b \gg R$, scattering is essentially elastic. Let us consider the expansion of the elastic amplitude in terms of string diagrams. In the high-energy limit $s \gg |t|$ and in the semiclassical regime $R, b \gg l_s$, the $N$-loop amplitude is factorized in the convolution of $N$ single-graviton-exchange amplitudes between the colliding particles, represented by ladder-like effective diagrams (Fig. 1(a)).

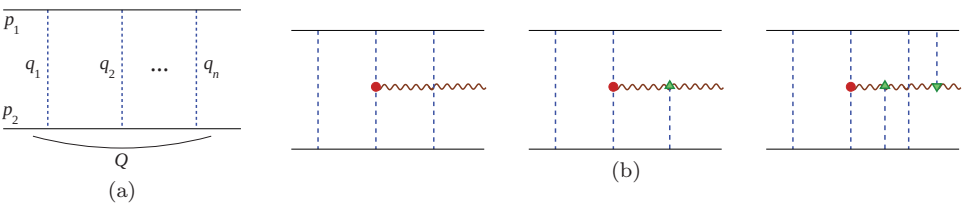



Fig. 1. Leading eikonal diagrams: (a) elastic scattering; (b) graviton emission and rescattering. The red disk represents the Lipatov vertex.

Such convolution in transverse momentum can be diagonalized by a Fourier transform in impact parameter space, and the ladder diagrams can be resummed in exponential series, yielding the elastic $S$-matrix in eikonal form, $S = \mathrm{e}^{\mathrm{i}2\delta(b,s)}$. The quantity $\delta(b,s) \equiv Gs\Delta(b) = Gs\log(L/b)$ can be interpreted as the (very large) phase shift in potential scattering. Here $L$ is an IR cutoff needed to regularize the Coulomb-like infinite phase shift which occurs in long-range interactions like gravity.

Note that the phase shift is real in this eikonal approximation, therefore the $S$-matrix is unitary. From the phase shift one can compute various observables, like the deflection angle, proportional to the gradient of the phase shift. One finds perfect agreement with the Einstein deflection angle $\theta_s = 2R/b$.

### 2.2. *Graviton emission*

Graviton emission at high-energy (Regge kinematics) is described by the so-called effective Lipatov vertex[2] (Fig. 1(b)), which is accurate in the region where the graviton emission angle $\theta$ is much larger than the deflection angle $\theta_s$ of the incoming particles. However such vertex is inaccurate when $\theta \sim \theta_s$. In this case, the emission amplitude for soft gravitons whose energy $\omega \ll E$ is well described by the soft-graviton emission theorem in terms of the Weinberg current[3]. It turns out that, for spin-2 particles, a very simple and elegant unifying amplitude can be derived which interpolates both angular regimes for all the relevant graviton energies, even $\omega \gg M_P$, provided $\omega \ll E$.

In addition we have to take into account the so-called rescattering diagrams[1], which describe the gravitational interaction of the radiated gravitons with the original sources, as shown in the last two diagrams of Fig. 1(b).

By neglecting correlations such as emissions from the same rung, we were able to resum all diagrams of Fig. 1(b) and we obtained a remarkably simple result[4]:

$$\mathcal{M}_{2\to 2+N} = \mathrm{e}^{2\mathrm{i}\delta(\boldsymbol{b},s)} \prod_{j=1}^{N} \mathfrak{M}_{\lambda_j}(\boldsymbol{b},\omega_j,\boldsymbol{q}_j) \times \left[1 + \mathcal{O}\left(\omega_j^2/E^2\right)\right] \tag{1}$$

$$\mathfrak{M}_\lambda(\boldsymbol{b},\omega,\boldsymbol{q}) = \sqrt{Gs}\frac{R}{\pi}\mathrm{e}^{\mathrm{i}\lambda\phi_{\boldsymbol{q}}} \int \frac{\mathrm{d}^2\boldsymbol{x}}{2\pi|\boldsymbol{x}|^2\mathrm{e}^{\mathrm{i}\lambda\phi_{\boldsymbol{x}}}}\, \mathrm{e}^{\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}} \frac{\mathrm{e}^{2\mathrm{i}\omega R\Phi_B(\boldsymbol{x})} - \mathrm{e}^{2\mathrm{i}\omega R\Phi_A(\boldsymbol{x})}}{2\mathrm{i}\omega R} \tag{2}$$

$$\Phi_A(\boldsymbol{x}) \equiv \frac{E}{\omega}\left[\Delta(\boldsymbol{b} - \tfrac{\omega}{E}\boldsymbol{x}) - \Delta(\boldsymbol{b})\right], \qquad \Phi_B(\boldsymbol{x}) \equiv \left[\Delta(\boldsymbol{b}-\boldsymbol{x}) - \Delta(\boldsymbol{b})\right] \tag{3}$$

In words, the total emission amplitude is given by the elastic amplitude times an emission factor $\mathfrak{M}$ for each emitted gravition. $\mathfrak{M}$ represents the basic object for the description of graviton radiation.

Soft infrared divergencies coming from phase-space integrations of emitted soft gluons are cancelled by virtual correction, that we incorporate by means of the Weinberg method[3], which amounts to write the final state of graviton in Fock space as a coherent state operator $S$ acting on the vacuum with creation (real emission) and destruction (virtual corrections) operators:

$$S = \mathrm{e}^{2\mathrm{i}\delta} \exp\left\{\int \frac{\mathrm{d}^3q}{\sqrt{2\omega}}\, 2\mathrm{i} \sum_\lambda \left[\mathfrak{M}_\lambda a_\lambda^\dagger(q) + \mathfrak{M}_\lambda^* a_\lambda(q)\right]\right\} \tag{4}$$

We are therefore able to construct an explicitly unitary $S$ matrix for particle scattering and associated bremmsstrahlung. From Eq. (4) we can compute the energy spectrum of graviton radiation $\mathrm{d}E^{\mathrm{GW}}/\mathrm{d}\omega$ which is given by $2\omega|\mathfrak{M}|^2$, integrated over all angular directions. The result is depicted in fig. Fig. 2(a). We notice that the spectrum at low frequencies tends to a constant, in agreement with the zero-frequency-limit[5].

It then shows a $\log(1/\omega)$ behaviour in the intermediate range $1/b < \omega < 1/R$, while for $\omega > 1/R$ the spectrum decreases as $1/\omega$, almost independently of the impact parameter $b$, i.e., of the deflection angle $\theta_s$ of the hard particles. Note

that $R$ increases with energy $E$. This means that if we increase the energy of the process, the typical frequencies of the radiation decrease, in a way reminiscent of how Hawking radiation depends upon the BH mass.

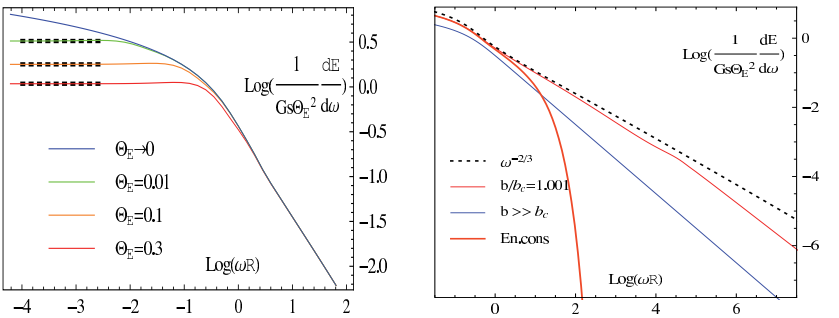



Fig. 2. Rescaled frequency spectrum of graviton radiation. (a) Spectrum in the leading eikonal approximation ($b \gg R$) for various values of $\theta_s$; the black dashes on the left represent the zero-frequency limit. (b) Resummed spectrum with subleading contribution at $b \sim b_c \sim R$, showing the intermediate enhanced region and the exponential fall-off when including energy conservation.

Remarkably, this result agrees with the classical limit ($\hbar\omega/E \to 0$) frequency spectrum of Ref. 6. Of course, in our framework we can compute also the quantum corrections.

## 3. Strong Gravity Regime

By decreasing the impact parameter $b$ towards the gravitational radius $R$, we enter the strong gravity regime, the one we are really interested in. In this regime we have to consider subleading corrections.

### 3.1. *Elastic scattering*

ACV were able to identify and resum also the class of subleading diagrams, the so-called H diagrams[7]. These are diagrams with a trilinear interaction (the Lipatov vertex) which is responsible for graviton emissions from exchanged gravitons.

As a consequence, a very interesting feature appears: the phase shift $\delta(b, s)$ acquires an imaginary part for impact parameters $b < b_c \simeq 1.6R$ so that the elastic $S$-matrix is suppressed. The question then arises: is this unitarity deficit in the elastic channel compensated by inelastic graviton production? Or is this critical value a signal of gravitational collapse? Let us note that the unitarity deficit has a fractional critical exponent $= 3/2$, resembling a sort of phase transition.

### 3.2. *Graviton emission*

By including graviton production from the multi-H diagrams we argue that the resummed emission amplitude maintains the form of Eq. (1) by replacing the leading phase shift $\delta(b,s)$ with the resummed one. For $b \to b_c^+$ the non-analytic behaviour of $\delta(b)$ severely affects the energy spectrum. In particular, for $\omega > 1/R$ there is an intermediate region where the frequency spectrum decreases less rapidly than $1/\omega$, actually $\omega^{-2/3}$ (Fig. 2(b)), implying a much higher radiation due to strong tidal forces[8].

### 3.3. *Energy Conservation*

Because of the logarithmic behaviour of $\mathrm{d}E^{GW} \sim Gs\theta_s^2 \mathrm{d}\omega/\omega$, and also because of the enhanced emission of radiation at large angles, it is mandatory to take into account energy conservation, i.e., to require that the energy carried by emitted gravitions be less than the total CM energy.

Following the proposal of Ref. 9, we impose energy conservation event by event, by requiring $\sum_j \omega_j < E$ in each hemisphere, and extending this bound to virtual corrections on the basis of AGK cutting rules[10]. Without entering here into technical details, such procedure amounts to multiply each amplitude by $\Theta(E - \sum_j \omega_j)$ and maintaining phase coherence them. By a saddle point evaluation of the usual integral representation of the $\Theta$-function one can compute the radiated energy distribution[8]. We observe that, while $b$ approaches the critical value $b_c = 1.6\,R$, the energy distribution shows an intermediate regime where it is enhanced by tidal forces and decreases only as $\sim \omega^{-2/3}$, before the eventual exponential decrease $\sim \mathrm{e}^{-\omega/T}$ at large $\omega$'s. Such exponential behaviour is typical of a thermal radiation. But here the radiation is coherent since the $S$ matrix is unitary. The "quasi-temperature" parameter $T$ is of order of the Hawking temperature $T_H = 1/(4\pi R)$ for $b \to b_c$: $T \simeq 0.8/R$. In particular it decreases with increasing energy of the collision.

## 4. Conclusions

To summarize, we claim we have a reliable model of gravity which allows us to compute graviton interactions and radiation in transplanckian collisions.

Within this model, we now see the role of the gravitational radius as the inverse of the characteristic frequency for all values of CM energies and impact parameters.

When the impact parameter $b$ is smaller than the gravitational radius $R$, tidal forces cause a dramatic enhancement of graviton radiation and all energy is radiated off. By taking into account energy-conservation constraints, we obtain a coherent radiation pattern with an exponentially decreasing spectrum, resembling a thermal radiation with a temperature of the order of Hawking's one $T \sim T_H$ of a BH somewhat lighter than $\sqrt{s}$.

All in all, these results suggest a possible mechanism of solving the information paradox.

## References


1. D. Amati, M. Ciafaloni and G. Veneziano, Higher Order Gravitational Deflection and Soft Bremsstrahlung in Planckian Energy Superstring Collisions, *Nucl.Phys.* **B347**, 550 (1990).
2. L. Lipatov, Multi - Regge Processes in Gravitation, *Sov.Phys.JETP* **55**, 582 (1982).
3. S. Weinberg, Infrared photons and gravitons, *Phys.Rev.* **140**, B516 (1965).
4. M. Ciafaloni, D. Colferai, F. Coradeschi and G. Veneziano, Unified limiting form of graviton radiation at extreme energies, *Phys. Rev.* **D93**, p. 044052 (2016).
5. L. Smarr, Gravitational Radiation from Distant Encounters and from Headon Collisions of Black Holes: The Zero Frequency Limit, *Phys.Rev.* **D15**, 2069 (1977).
6. A. Gruzinov and G. Veneziano, Gravitational Radiation from Massless Particle Collisions, *Class. Quant. Grav.* **33**, p. 125012 (2016).
7. D. Amati, M. Ciafaloni and G. Veneziano, Towards an S-matrix description of gravitational collapse, *JHEP* **0802**, p. 049 (2008).
8. M. Ciafaloni and D. Colferai, Unitarity restoring graviton radiation in the collapse regime of gravitational scattering, *Phys. Rev.* **D96**, p. 126008 (2017).
9. G. Veneziano, String-theoretic unitary S-matrix at the threshold of black-hole production, *JHEP* **0411**, p. 001 (2004).
10. V. A. Abramovsky, V. N. Gribov and O. V. Kancheli, Character of Inclusive Spectra and Fluctuations Produced in Inelastic Processes by Multi - Pomeron Exchange, *Yad. Fiz.* **18**, 595 (1973), [Sov. J. Nucl. Phys.18,308(1974)].

# Observables from modified dispersion relations on curved spacetimes: Circular orbits, redshift and lateshift

Christian Pfeifer

*Laboratory of Theoretical Physics, Institute of Physics, University of Tartu, W. Ostwaldi 1 50411 Tartu, Estonia*
*E-mail: christian.pfeifer@ut.ee*

The Hamiltonian formulation of modified dispersion relations (MDRs) allows for their implementation on generic curved spacetimes. In turn it is possible to derive phenomenological effects. I will present how to construct the kappa-Poincare dispersion relation on curved spacetimes, its spherically symmetric realizations, among them the kappa deformation of Schwarzschild spacetime, and its implementation on Friedmann-Lemaître-Robertson-Walker spacetimes with arbitrary scale factor. In addition we will construct the general first order modifications of the general relativistic dispersion relation. Afterwards we will use the perturbative MDRs to calculate specific observables such as the redshift, lateshift and photon circular orbits.

*Keywords*: Quantum gravity phenomenology, modified dispersion relations, observables, lateshift, redshift, photon orbits.

## 1. Traces of quantum gravity in point particle motion

A large source of information about the gravitational interaction come from the observation of trajectories of freely falling particles, such as for example neutrinos or photons, through spacetime. Traces of the expected quantum nature of gravity and its influence on the particles' paths therefore should manifest itself in the data collected by telescopes and observatories. One approach to derive effects, one can search for in the available data, is provided by quantum gravity phenomenology [1]. Due to a still missing fundamental theory of quantum gravity, we proceed along the following convincing pictorial idea to set up an effective model for the interaction between point particles and gravity on small distances or high energies: When test particles propagate through spacetime they probe spacetime on lengths scales inverse proportional to their energy. Thus particles with larger energy probe smaller length scales then lower energetic ones. The quantum nature of gravity is expected to become relevant at the Planck scale, i.e. at the Planck energy $E_{pl}$ respectively the Planck length $\ell_{pl}$. Hence particles with larger energies, closer to the Planck energy, probe length scales closer to the Planck length and should thus interact stronger with the quantum features of gravity.

This pictorial idea can be seen analogue to how a medium can be probed with photons of different energies, to obtain insights about its constituents. For very low energetic photons the medium may be invisible, while photons within a certain energy range may interact with the medium, get scattered and transport information about the medium to an observer. For the interaction between photons and the elementary constituents of the medium we know, that they are fundamentally explained by the standard model of particle physics. However effectively one can

describe several aspects of the system by an energy dependent propagation of the photons through the medium. The latter can be derived from an effective theory of electrodynamics, such premetric electrodynamics[2], and leads to a non local lorentz invariant (LLI) dispersion relation of the photons. Thus even though a fundamental interaction is LLI, observables may be described effectively by a non LLI theory.

Regarding the structure of spacetime and the quantum nature of gravity, we do not know the fundamental theory and of quantum gravity and its properties yet. In analogy to the effective description of the propagation of photons through media we will consider modifications of the general relativistic dispersion relation of point particles on spacetime. From such modifications we can calculate observables and look for their signatures in the available data. Any evidence or non-detection of such a signal then reveals properties, suitable semi-classical limits of quantum gravity must have.

Most famous effects searched for are an energy dependent redshift and time of arrival (lateshift) of photons[3]. Preliminary analyses of the ICECUBE and Fermi Gamma-Ray Space telescope observations for a lateshift effect have recently been performed[4–7] and shall be extended as soon as additional data is available.

We demonstrate how modified dispersion relations (MDRs) are realized on generically curved spacetimes we will derive observables such as the redshift, the lateshift and photon orbits.

## 2. Dispersion relations as Hamilton functions

The study of non LLI effects and MDRs as effective description of the interaction of point particles with the quantum nature of gravity has a long history in the literature, see for example[8–12] and references therein.

In Ref. 13 we demonstrated how to realize modified dispersion relations on curved spacetimes covariantly: *A dispersion relation of a point particle on curved spacetime is given by a level set of a Hamilton function $H(x,p)$ on the point particle phase space (cotangent bundle) of spacetime. The particle's motion is determined by the corresponding Hamilton equations of motion.* Moreover the Hamiltonian defines the geometry of phase space in a canonical way. In general spacetime and momentum space are curved and their curvatures depend on positions and momenta.

Using this covariant approach to dispersion relations we recall the Hamilton functions of general relativistic particles $H_{GR}$ and construct its first order perturbations $H_\ell$ as well as its $\kappa$-Poincaré deformations $H_\kappa$ [14,15] on general[16], static spherically symmetric[16] and homogeneous and isotropic spacetimes[16–18].

On a generically curved spacetime the Hamiltonians are build from a spacetime metric $g$, a perturbation function $h$ and unit timelike vector field of the metric $Z$

$$H_{GR}(x,p) = g^{ab}(x)p_a p_b\,, \tag{1}$$

$$H_\ell(x,p) = g^{ab}(x)p_a p_b + \ell h(x,p)\,, \tag{2}$$

$$H_\kappa(x,p) = -\tfrac{4}{\ell^2}\sinh\left(\tfrac{\ell}{2}Z^c(x)p_c\right)^2 + e^{\ell Z^c(x)p_c}\left(g^{ab}(x)p_a p_b + (Z^d(x)p_d)^2\right). \tag{3}$$

In static spherical symmetry the dependence of the Hamiltonians on the positions and momenta is restricted to $H(x,p) = H(r,p_t,p_r,v)$ with $v^2 = p_\theta^2 + \sin\theta^{-2} p_\phi^2$ [16],

$$H_{GR}(x,p) = -a(r)p_t^2 + b(r)p_r^2 + r^{-2}v^2\,, \tag{4}$$

$$H_\ell(x,p) = -a(r)p_t^2 + b(r)p_r^2 + r^{-2}v^2 + \ell h(r,p_t,p_r,v)\,, \tag{5}$$

$$\begin{aligned} H_\kappa(x,p) = &-\tfrac{4}{\ell^2}\sinh\left(\tfrac{\ell}{2}(c(r)p_t + d(r)p_r)\right)^2 + e^{\ell(c(r)p_t + d(r)p_r)} \\ &\times\left((-a(r)+c(r)^2)p_t^2 + 2c(r)d(r)p_tp_r + (b(r)+d(r)^2)p_r^2 + \tfrac{1}{r^2}v^2\right). \end{aligned} \tag{6}$$

Setting here $a(r)^{-1} = b(r) = 1 - \frac{r_s}{r}$ we obtain the $\kappa$-Poincaré deformations of Schwarzschild spacetime parametrized by two free function $c(r)$ and $d(r)$.
On homogeneous and isotropic spacetimes the form of the Hamiltonian is further restricted to $H(x,p) = H(t,p_t,w)$ with $w^2 = p_r^2(1-kr^2) + r^{-2}w^2$

$$H_{GR}(x,p) = -p_t^2 + A(t)^{-2}w^2\,, \tag{7}$$

$$H_\ell(x,p) = -p_t^2 + A(t)^{-2}w^2 + \ell h(t,p_t,w)\,, \tag{8}$$

$$H_\kappa(x,p) = -\tfrac{4}{\ell^2}\sinh\left(\tfrac{\ell}{2}p_t\right)^2 + e^{\ell p_t}A(t)^{-2}w^2\,. \tag{9}$$

## 3. Observables

A detection of traces of quantum gravity is most likely with high energetic photons and every MDR of interest can be expanded to first order around $H_{GR}$. Thus we focus on the derivation of observables for massless particles from $H_\ell(x,p)$.

### 3.1. *Redshift*

The redshift of photons involves a description of how observers measure the frequency of a photon. A classical observer, not subject to the MDR, on a worldline $\mathrm{x}(\tau)$ with tangent $\dot{\mathrm{x}}(\tau)$, momentum $p_\mathrm{x}(\tau)$ and mass $m_\mathrm{x}^2 = -H_{GR}(\mathrm{x},p_\mathrm{x})$, associates to a photon, subject to the MDR, on a worldline $\mathrm{y}(\tau)$ ,with momentum $p_\mathrm{y}(\tau)$ and satisfying $H_\ell(\mathrm{y},p_\mathrm{y}) = 0$, the frequency

$$\nu_\mathrm{x}(\mathrm{y}) = \dot{\mathrm{x}}^a p_{\mathrm{y}a} = \tfrac{1}{2m_\mathrm{x}}\partial_{p_a}H_{GR}(\mathrm{x},p_\mathrm{x})\; p_{\mathrm{y}a} = \tfrac{1}{m_\mathrm{x}}g^{ab}(\mathrm{x})\; p_{\mathrm{x}b}\; p_{\mathrm{y}a}\,. \tag{10}$$

The redshift of a photon between two observers $\mathrm{x}_i$ and $\mathrm{x}_f$ then is

$$z(t_i,r_i,t_f,r_f) \equiv z = \tfrac{\nu_{\mathrm{x}_i}(\mathrm{y})}{\nu_{\mathrm{x}_f}(\mathrm{y})} - 1. \tag{11}$$

Thus, two observers at rest at $\mathrm{x}_i = (t_i, r_i, \frac{\pi}{2}, 0)$ and $\mathrm{x}_f = (t_f, r_f, \frac{\pi}{2}, 0)$ find for a radially freely falling photon in spherical symmetry from (5) the redshift [16]

$$z = \sqrt{\tfrac{a(r_f)}{a(r_i)}} - 1 + \mathcal{O}(\ell^2)\,, \tag{12}$$

while for the homogeneous and isotropic case (8) they find that the redshift becomes dependent of the constant of motion $w$, which is related to the photons frequency [18]

$$\begin{aligned} z &= \left(\tfrac{A(t_f)}{A(t_i)} - 1\right) - \tfrac{\ell}{2w^2}\tfrac{A(t_f)}{A(t_i)}\left(A(t_f)^2 h(t_f, p_t^0(t_f,w), w) - A(t_i)^2 h(t_i, p_t^0(t_i,w), w)\right) \\ &\to \left(\tfrac{A(t_f)}{A(t_i)} - 1\right) - \ell w \tfrac{A(t_f) - A(t_i)}{2A(t_i)^2}\,. \end{aligned} \tag{13}$$

The first order momentum is $p_t^0(t,w) = -\frac{w}{A(t)}$ and the last line displays the result for a first order in $\ell$ expansion of the $\kappa$-Poincaré dispersion relation (9).

### 3.2. *Lateshift*

For the so called lateshift consider two massless particles with different frequency parameters $w_1$ and $w_2$ on radial trajectories $r_1(t,w_1)$ and $r_2(t,w_2)$ solving the Hamilton equations of motion of (8). They shall be emitted at the same coordinates $(t_i, R_i)$, i.e. $R_i = r_1(t_i,w_1) = r_2(t_i,w_2)$, and we ask when do they reach the same radial distance $R_f = r_1(t_1,w_1) = r_2(t_2,w_2)$. The difference in their time of arrival $\Delta t = t_2 - t_1$ is the so called lateshift. For the solutions of the radial Hamilton equation of motion of (8), derived in Ref. 18, we find,

$$\Delta t = \ell A(t_1) \int_{t_i}^{t_1} \mathrm{d}\tau \ \tfrac{f(\tau,p_t^0(\tau,w_2),w_2) - f(\tau,p_t^0(\tau,w_1),w_1)}{A(\tau)} \tag{14}$$

$$\to -\ell \tfrac{A(t_1)(w_1-w_2)}{2} \int_{t_i}^{t_1} d\tau \tfrac{1}{A(\tau)^2}\,. \tag{15}$$

where again $p_t^0(t,w) = -\frac{w}{A(t)}$, the last line is the first order in $\ell$ result of (9), and

$$f(t,p_t^0(t,w),w) = \tfrac{1}{2(p_t^0)^2}\big[h(t,p_t^0,w) - p_t^0 \partial_{p_t} h(t,p_t^0,w) - w\partial_w h(t,p_t^0,w)\big]\,. \tag{16}$$

### 3.3. *Innermost circular photon orbits*

To calculate the innermost circular photon orbits we solve the Hamilton equations of motion of (5) with the assumption that $\dot{r} = 0$. Due to the spherical symmetry we can without loss of generality consider orbits in the equatorial plane $\theta = \frac{\pi}{2}$. The Hamilton equations of motion $\partial_{p_r} H_\ell = \dot{r} = 0$ and $\partial_r H = -\dot{p}_r$ with $p_r = p_r^0 + \ell p_r^1$, together with the fact that $p_t$ and $v$ are constants of motion, then imply to first order in $\ell$

$$p_r^0 = 0, \quad p_r^1 = -\tfrac{\partial_{p_r} h(r,p_t,p_r^0,v)}{2b(r)} \quad \Rightarrow \quad \dot{p}_r^0 = 0 = \dot{p}_r^1\,. \tag{17}$$

Using $H_\ell(x,p) = 0$ to express $p_t$ as function of $r$ and $v$, and solving $\partial_r H = -\dot{p}_r = 0$ yields,

$$0 = \tfrac{a'(r_0)}{a(r_0)} + \tfrac{2}{r_0}, \quad r_1 = \frac{\partial_r h(r_0,p_t^0(r_0,v),0,v) - \frac{a'(r_0)}{a(r_0)} h(r_0,p_t^0(r_0,v),0,v)}{\left(\frac{a''(r_0)}{a(r_0)r_0^2} - \frac{a'(r_0)^2}{a(r_0)^2 r_0^2} - \frac{2a'(r_0)}{a(r_0)r_0^3} - \frac{6}{r_0^4}\right)v^2}\,, \tag{18}$$

which determine the allowed circular photon orbits $r = r_0 + \ell r_1$ to first order in $\ell$.

Considering MDRs on Schwarzschild spacetime by specifying the function $a(r)^{-1} = 1 - \frac{r_s}{r}$ the zeroth order becomes $r_0 = \frac{3}{2} r_s$, as it must be, and the equation which determines the first order correction becomes

$$r_1 = \frac{27 r_s^4 (\partial_r h(r_0,p_t^0(r_0,v),0,v) - \frac{a'(r_0)}{a(r_0)} h(r_0,p_t^0(r,v),0,v))}{32 v^2}\,. \tag{19}$$

Specifying to the first order in $\ell$ of the spherically symmetric $\kappa$-Poincaré dispersion relation (6) with $d(r) = 0$ and $c(r) = \frac{1}{\sqrt{1-\frac{r_s}{r}}}$ (19) yields $r_1 = \frac{v}{6}$ [16].

### 3.4. *Conclusion*

The implementation of modified dispersion relation on curved spacetimes allows for the prediction of traces of the quantum nature of gravity, which can be confirmed or constraint by observations. Particularly promising are time of arrival measurements of high energetic gamma-rays, which can be compared against equation (14) and the search for rainbow effects in lensing images, for example in the shadows of black holes, which would correspond to a non-trivial $v$ dependence in (18).

To bridge the the gap between fundamental approaches to quantum gravity and the phenomenological one presented here, it is necessary to derive the perturbation function $h(x, p)$ in (2) in some semi-classical limit from the fundamental approaches. This is one of the major tasks for future research.

## Acknowledgments

This talk is based on a series of articles[13,16,17] written in collaboration with Leonardo Barcaroli, Lukas Brunkhorst, Giulia Gubitosi and Niccoló Loret.

## References

1. G. Amelino-Camelia, Living Rev. Rel. **16** (2013) 5 doi:10.12942/lrr-2013-5 [arXiv:0806.0339 [gr-qc]].
2. F. W. Hehl and Y. N. Obukhov, "Foundations of Classical Electrodynamics," Birkhäuser, 2003.
3. G. Amelino-Camelia, L. Barcaroli, G. Gubitosi and N. Loret, Class. Quant. Grav. **30** (2013) 235002 doi:10.1088/0264-9381/30/23/235002 [arXiv:1305.5062 [gr-qc]].
4. G. Amelino-Camelia, G. D'Amico, G. Rosati and N. Loret, Nat. Astron. **1** (2017) 0139 doi:10.1038/s41550-017-0139 [arXiv:1612.02765 [astro-ph.HE]].
5. H. Xu and B. Q. Ma, JCAP **1801**, no. 01, 050 (2018) doi:10.1088/1475-7516/2018/01/050 [arXiv:1801.08084 [gr-qc]].
6. H. Xu and B. Q. Ma, Phys. Lett. B **760** (2016) 602 doi:10.1016/j.physletb.2016.07.044 [arXiv:1607.08043 [hep-ph]].
7. H. Xu and B. Q. Ma, Astropart. Phys. **82** (2016) 72 doi:10.1016/j.astropartphys.2016.05.008 [arXiv:1607.03203 [hep-ph]].
8. H. J. Matschull and M. Welling, Class. Quant. Grav. **15** (1998) 2981 doi:10.1088/0264-9381/15/10/008 [gr-qc/9708054].
9. L. Freidel, J. Kowalski-Glikman and L. Smolin, Phys. Rev. D **69** (2004) 044001 doi:10.1103/PhysRevD.69.044001 [hep-th/0307085].
10. Y. Ling, B. Hu and X. Li, Phys. Rev. D **73** (2006) 087702 doi:10.1103/PhysRevD.73.087702 [gr-qc/0512083].
11. F. Girelli, S. Liberati, R. Percacci and C. Rahmede, Class. Quant. Grav. **24** (2007) 3995 doi:10.1088/0264-9381/24/16/003 [gr-qc/0607030].

12. S. Liberati, Class. Quant. Grav. **30** (2013) 133001 doi:10.1088/0264-9381/30/13/133001 [arXiv:1304.5795 [gr-qc]].
13. L. Barcaroli, L. K. Brunkhorst, G. Gubitosi, N. Loret and C. Pfeifer, Phys. Rev. D **92** (2015) no.8, 084053 doi:10.1103/PhysRevD.92.084053 [arXiv:1507.00922 [gr-qc]].
14. G. Gubitosi and F. Mercati, Class. Quant. Grav. **30**, 145002 (2013) doi:10.1088/0264-9381/30/14/145002 [arXiv:1106.5710 [gr-qc]].
15. J. Kowalski-Glikman, Mod. Phys. Lett. A **17** (2002) 1 doi:10.1142/S0217732302006175 [hep-th/0107054].
16. L. Barcaroli, L. K. Brunkhorst, G. Gubitosi, N. Loret and C. Pfeifer, Phys. Rev. D **96** (2017) no.8, 084010 doi:10.1103/PhysRevD.96.084010 [arXiv:1703.02058 [gr-qc]].
17. L. Barcaroli, L. K. Brunkhorst, G. Gubitosi, N. Loret and C. Pfeifer, Phys. Rev. D **95** (2017) no.2, 024036 doi:10.1103/PhysRevD.95.024036 [arXiv:1612.01390 [gr-qc]].
18. C. Pfeifer, Phys. Lett. B **780** (2018) 246 doi:10.1016/j.physletb.2018.03.017 [arXiv:1802.00058 [gr-qc]].

# A linear equation of state for trapped gravitons, logarithmic corrections to the black hole entropy and the dark energy

Stefano Viaggiu

*Dipartimento di Matematica, Universietà di Roma "Tor Vergata",*
*Via della Ricerca Scientifica, 1, I-00133 Roma, Italy*
*E-mail: viaggiu@axp.mat.uniroma2.it,*
*Universitá Guglielmo Marconi, Via Plinio 44, I-00193 Rome,*
*INFN, Sezione di Napoli, Complesso Universitario di Monte S. Angelo,*
*Via Cintia Edificio 6, 80126 Napoli, Italy*

I present recent results concerning the statistical description of the Black Hole entropy in terms of trapped gravitons. In particular, I present a theorem depicting a possible physical mechanism, from finite size effects induced by Planckian fluctuations, giving trapped gravitons with a linear equation of state. As a consequence, logarithmic corrections to the Black Hole entropy naturally arise. Finally, it is also shown that such a mechanism allows a statistical description of macroscopic configurations made of cosmological constant in terms of massless excitations within a spherical box.

*Keywords*: Black Hole Thermodynamics; Black Hole Entropy; Trapped Gravitons.

## 1. Corrections to the black hole entropy

As well known[1], black holes (BH) emit a thermal radiation with no information within and depending only from the mass (charge and angular momentum, no-hair theorem) and at the temperature $(r \to \infty)$ $T_h = \frac{c\hbar}{4\pi K_B R}$. The entropy is $S_H = \frac{K_B A}{4L_P^2}$. Many open questions are still unsolved:

- What is the physical origin of the degrees of freedom implying $S_h$?
- How can we obtain the log. corrections to the BH entropy?

The physical origin of the BH entropy degrees of freedom is still obscure. In Refs. 2, 3 we have proposed that the degree of freedom leading to the BH entropy are represented by trapped gravitons inside the event horizon. After imposing Dirichlet boundary conditions on the Zerilli or Regge-Wheeler functions (Gravitational waves in the vacuum) at $r = R$: $Z_{\ell m}^{(p,a)}(R, \omega) = 0$, we obtain the following expression for the allowed frequencies:

$$\omega_{\ell n} \simeq \frac{c}{2R}\left(2 + \ell + 2n\right)\pi, \;\; \ell \geq 2, \;\; n \in \mathbb{N}. \tag{1}$$

We assume, as usual, that the internal energy $U_i$ is provided by the ADM mass $M$, i.e. $U_i = Mc^2$. A statistical entropy can thus be derived from (1) in the usual way by means of the Shannon (von Newmann) entropy:

$$\begin{aligned} S = &-NK_B\left[\ln\left(1 - e^{-\frac{X}{2}}\right) + \ln\left(1 - e^{-X}\right)\right] \\ &+ \frac{c\pi\hbar N e^{-\frac{X}{2}}\left[1 + 3e^{-\frac{X}{2}}\right]}{2TR\left(1 - e^{-X}\right)}, \;\; X = \frac{c\pi\beta\hbar}{R}. \end{aligned} \tag{2}$$

The equation of state is the one of a radaition field $P_i V = \frac{U_i}{3}$. For the temperature we have $T_i = \alpha T_h = \frac{\alpha c \hbar}{4\pi K_B R}$, $\alpha \in (0, \infty)$. We find $S = K_B Y(\alpha) \frac{A_h}{4L_P^2}$, with $Y(\alpha) = 1$ with $\alpha \simeq 2.2 \sim 2$, where $\alpha = 2$ is the theretical value: our model it gives viable results within the approximation made for the (1). How can be obtained logarithmic corrections to the BH entropy? BH entropy is semi-classical, so we expect some quantum (Planckian) effect leading to the well known $\ln(A)$ corrections. First of all we expect modifications for the semi-classical formula $U(R)$. In a quantum spacetime[6–8], spacetime uncertainty relations are satured with $\Delta E \sim \frac{1}{R}$, and the following modification of $U_i$ emerges with $U_i(R) = \frac{c^4}{2G} R + \frac{C_1 c^4 L_P^2}{2GR}$, $\{C_1\} \in \mathbf{R}^+$. Unfortunately, the added term does not lead to logarithmic entropy corrections. To overcome this issue, in Ref. 4 and after in Ref. 5 for any massless excitation, we have proposed a physical mechanism, motivated by solid state physics, capable to depict a system with a gamma linear equation of state, i.e. quantum fluctuations change the equation of state $PV = \gamma U$, $\gamma \in \mathbf{R}$. The following theorem holds:

*Theorem:* Let $\omega_0 = \omega_{\ell n}$ denote the angular frequency of $N$ trapped gravitons with internal energy $U^{(0)}$ (radiation). The so trapped gravitons with angular frequency $\omega = \omega_0 + \frac{\phi(R)}{N}$ have a linear equation of state with $U = U^{(0)} + \hbar\, \phi(R)$ iff there exists a differentiable function $\phi(R)$, satisfying the following equation

$$\hbar \left[ R\, \phi_{,R}(R) + \phi(R) \right] = U(R)(1 - 3\gamma),$$

together with the condition

$$U - \hbar\, \phi(R) > 0.$$

To start with, consider the semi-classical BH case $U = \frac{c^4 R}{2G}$; for $\Phi(R)$ we have $\phi(R) = \frac{c}{4L_P^2} R(1 - 3\gamma)$. with limitation $\gamma > -\frac{1}{3}$ (independently on $R$), i.e. in the semi-classical case the active gravitational mass is positive and gravity is attractive, a reasonable result. Consider now $U = \frac{c^4}{2G} R + \frac{C_1 c^4 L_P^2}{2GR}$, we obtain

$$\phi(R) = \frac{cR}{4L_P^2}(1 - 3\gamma) + \frac{c\, C_1\, (1 - 3\gamma)}{2R} \ln\left( \frac{R}{s L_P} \right). \tag{3}$$

As a first important consequence of (2), since entropy, independently on the equation of state, is proportional to $N$, is that $N \sim R(U - \hbar\phi)$ and as a result logarithmic corrections emerge. We must discuss te existence condition $U - \hbar\, \phi(R) > 0$. After posing $W = \frac{R}{L_P}$, we get:

$$W + \frac{C_1}{W} - (1 - 3\gamma) \left[ \frac{W}{2} + \frac{C_1}{W} \ln\left( \frac{W}{s} \right) \right] > 0.$$

The results are:

- $\gamma = -1$, Cosmological constant: A maximum radius $R_M$ is present of the order (greater) of Planck length $L_P$. This implies that BH filled with gravitons with a cosmological constant equation of state are only possible with microscopic Planckian sizes.

- By supposing $C_1 < 0$ we obtain a minimum radius $R_m$ and a maximum one of the order of $\sim L_P$.
- For $\gamma \in [0, \frac{1}{3}]$ we have no restriction for $R$, while for $\gamma > \frac{1}{3}$ a minimum allowed value for $R$ is present. In the stiff case ($\gamma = 1$), we have $R_m \sim \frac{L_P}{2}$.
- For $\gamma < 0$ a maximum allowed radius does appear $R_M \sim L_P$. As an example, for $\gamma = -\frac{1}{3}$ ($s \sim 1$) we have $R_M = \sqrt{e} L_P$.

## 2. The dark energy

The treatment presented above can be extended to any massless excitation in a spherical box[5]. In this frame, the theorem proven is still valid. In particular, by considering the positive cosmological constant case $\rho = B, B \in \mathbf{R}^+, \gamma = -1$ with $U = BV + \frac{C}{R}$ (Planck effects again), macroscopic dark energy configurations can be obtained. It is worth to be noted that with $C = 0$ we obtain the constraint $\gamma > -1$. The case $\gamma = -1$ is only possible with the quantum correction term $\sim 1/R$ in $U(R)$. This fact, in my opinion, strongly suggests the quantum-gravity nature of the cosmological comstant as due to strong Planckian fluctuations. In fact, in a cosmological de Sitter spacetime, an effective cosmological constant $\overline{\Lambda}$ emerges after applying the technology depicted above:

$$\overline{\Lambda} = \Lambda + \frac{3\xi L_P^2}{L^4}, \tag{4}$$

where $L$ denotes the physical scale and $\xi$ is a positive constant expected to be less than unity and $\Lambda$ denotes the bare cosmological constant without Planckian fluctuations. It is important to stress that the (4) does not represent a quintessence dark energy, but rather the effective value of the cosmological constant depends on the scale at which the physics is considered that in turn depends on the volume $V = 4\pi L^3/3$. In my view the cosmological constant is averaged over bigger and bigger scales up to a scale $L_D$, decoherence scale, such that the quantum Planckian fluctuations are negligible and a de Sitter universe emerges thus representing the crossover to classicality. Exactly at $L = L_D$ we have an absolute minimum for $U(L)$ with zero specific heat: the cosmological constant is thus fixed (frozen) at this critical scale. This scale depends crucially on the phenomenological parameter $\xi$. As an example, for $\xi \sim 1$ at scales $L \sim 10^{-5} m \ \rightarrow \overline{\Lambda} \sim 10^{-52}/m^2$. This seems a rather huge value, but no arguments to reject this value are at our disposal, although a scale of decoherence near the Planck length could be more appropriate. In any case, the important fact is that a new view concerning the origin of the cosmological constant emerges in the context depicted above.

## References

1. S. Hawking, *Commun. Math. Phys.* **43**, 199 (1975).
2. S. Viaggiu, *Physica A* **473**, 412 (2017.
3. S. Viaggiu, *Physica A* **488**, 72 (2017).

4. S. Viaggiu, *Int. J. Mod. Phys. D* **27**, 1850061 (2018).
5. S. Viaggiu, *Int. J. Mod. Phys. A* **33**, 1850074 (2018).
6. S. Doplicher, K. Fredenhagen and J. E. Roberts, *Comm. Math. Phys.* **172**, 187 (1995).
7. L. Tomassini and S. Viaggiu, *Class. Quantum Grav.* **28**, 075001 (2011).
8. L. Tomassini and S. Viaggiu, *Class. Quantum Grav.* **31**, 185;001 (2014).

# Fermionic backreaction in hybrid Loop Quantum Cosmology

Beatriz Elizaga de Navascués

*Institute for Quantum Gravity, Friedrich-Alexander University Erlangen-Nürnberg, Staudstraße 7, 91058 Erlangen, Germany*

Guillermo A. Mena Marugán and Santiago Prado

*Instituto de Estructura de la Materia, CSIC, Serrrano 121, Madrid, Spain*

We study how one can avoid some ultraviolet problems in the description of a perturbative inhomogeneous Dirac field in inflationary hybrid Loop Quantum Cosmology (LQC). Within a canonical approach for the entire cosmology, we consider different separations between the canonical variables that describe the homogeneous geometry and the fermionic perturbations. Each of such choices leads to a different fermionic contribution to the total Hamiltonian. In the hybrid quantization scheme, a quantum version of that contribution contains a backreaction term on the homogeneous part of the wavefunction. We provide the conditions on the choice of fermionic variables that lead to a finite backreaction.

*Keywords*: Loop Quantum Cosmology, Fermionic Perturbations, Quantum Backreaction.

## 1. Introduction

In conventional Quantum Field Theory (QFT) in curved spacetimes, one is forced to employ, in one way or another, some type of regularization procedure to cure the infinities that generically arise in possible observables, such as the Hamiltonian, constructed out of products of the field operators[1]. It is commonly believed that the reasons behind these issues can be traced to the treatment of the spacetime as a classical, continuum background.

We investigate here whether these problems can be avoided in cosmology by treating the entire system, including the spacetime background, as a dynamical entity to be quantized. For that purpose, we adopt the so-called hybrid approach in LQC to the quantization of a homogeneous and isotropic cosmology minimally coupled to a Dirac field, that we treat as a perturbation[2]. The approach is based on a splitting of the phase space into a homogeneous sector, that is represented with LQC techniques[3], and an inhomogeneous one, which is treated with conventional QFT methods. In this context, we exploit the freedom in performing canonical transformations that assign different dynamical roles to the homogeneous sector and the inhomogeneities. We show that this freedom permits a fermionic contribution to the Hamiltonian constraint that, in particular, gives rise to a non-divergent backreaction on the homogeneous background in the quantum theory.

## 2. The Cosmological System

The starting point for the construction of the cosmology under study is a Friedmann-Lemaître-Robertson-Walker (FLRW) spacetime with flat and compact spatial

hypersurfaces (isomorphic to a three-torus, $T^3$). We employ spatial coordinates adapted to the homogeneity. The matter content is given by a homogenous scalar (inflaton) field subject to a potential, and a Dirac field, both of them minimally coupled. The Dirac field is treated entirely as a perturbation. We truncate our perturbed system so that its Einstein-Dirac action is at most quadratic in all the perturbations[4]. Our canonical formulation is obtained from the symplectic structure and from the Hamiltonian associated with this truncated action.

In view of the symmetries of the homogeneous FLRW background, it is convenient to expand each of the two chiral components of the Dirac field in a complete set of eigenspinors of the Dirac operator on $T^3$, after imposing the time gauge on the homogeneous tetrads[4]. The spectrum of that operator is discrete and characterized by eigenvalues $\pm\omega_k$ that grow asymptotically as $2\pi|\vec{k}|/l_0$, where $\vec{k} \in \mathbb{Z}^3$ and $l_0$ is the compactification length of the tori. In the expansion of the left-handed component of the Dirac field, we call $e^{-3\alpha/2}m_{\vec{k}}$ and $e^{-3\alpha/2}\bar{r}_{\vec{k}}$ the time-dependent coefficients of the eigenspinors of the Dirac operator on $T^3$ with respective eigenvalues $\omega_k$ and $-\omega_k$. Here, $\alpha$ is, up to an additive constant, the logarithm of the scale factor of the FLRW cosmology. Similarly, $e^{-3\alpha/2}\bar{s}_{\vec{k}}$ and $e^{-3\alpha/2}t_{\vec{k}}$ respectively denote the coefficients of the complex conjugates of the eigenspinors with eigenvalues $\omega_k$ and $-\omega_k$ in the expansion of the right-handed component of the Dirac field. All of these eigenspinor coefficients are taken as Grassmann variables. Besides, each of them forms a canonical pair with its complex conjugate, with a Dirac bracket equal to $-i$, and vanishing anticommutation relations with the rest of coefficients. Introducing these mode decompositions in the action, one obtains the fermionic contribution to the total Hamiltonian. This contribution is quadratic in the fermionic variables, and is given by a sum over all modes, which decouple from each other. It appears multiplied by the homogeneous lapse function $N_0$, so we call it $N_0 H_D$, hence it only affects the zero mode of the Hamiltonian constraint.

## 3. Fermionic Annihilation and Creationlike Variables

To describe the fermionic degrees of freedom, in this section we will consider a rather generic family of definitions of annihilation and creationlike variables for the Dirac field. For this purpose, we will exploit the freedom to perform linear canonical transformations of the fermionic variables that depend on the homogeneous background geometry described by $\alpha$ and its momentum $\pi_\alpha$. Specifically, we restrict to definitions that respect the symmetries of the fermionic Hamiltonian[5]:

$$\begin{aligned} a^{(x,y)}_{\vec{k}} &= f^k_1(\alpha,\pi_\alpha)\, x_{\vec{k}} + f^k_2(\alpha,\pi_\alpha)\, \bar{y}_{-\vec{k}-2\vec{\tau}}, \\ \bar{b}^{(x,y)}_{\vec{k}} &= g^k_1(\alpha,\pi_\alpha)\, x_{\vec{k}} + g^k_2(\alpha,\pi_\alpha)\, \bar{y}_{-\vec{k}-2\vec{\tau}}, \end{aligned} \tag{1}$$

where $(x_{\vec{k}}, y_{\vec{k}})$ is any of the ordered pairs $(m_{\vec{k}}, s_{\vec{k}})$ or $(t_{\vec{k}}, r_{\vec{k}})$, the constant tuple $2\vec{\tau}$ in $\mathbb{Z}^3$ fixes the spin structure[6] on $T^3$, and $a^{(x,y)}_{\vec{k}}$ and $\bar{b}^{(x,y)}_{\vec{k}}$ correspond to annihilation-like variables for particles and creationlike variables for antiparticles, respectively.

These variables only satisfy the standard canonical anticommutation relations if[5]:

$$g_1^k = e^{iJ_k}\bar{f}_2^k, \qquad g_2^k = -e^{iJ_k}\bar{f}_1^k, \tag{2}$$

$$f_2^k = e^{iF_2^k}\sqrt{1-\left|f_1^k\right|^2}, \tag{3}$$

with $J_k$ and $F_2^k$ being some (possibly background-dependent) phases.

One can restrict the selection of annihilation and creationlike variables to a privileged family by imposing some physically desirable properties. A satisfactory criterion is the imposition that, in the context of QFT in curved spacetimes, the dynamics of the annihilation and creationlike variables can be implemented as unitary nontrivial transformations in Fock space. This condition, together with the invariance of the vacuum under the symmetries of the fermionic Hamiltonian, and a standard convention for the notions of particles and antiparticles, leads to a family of unitarily equivalent Fock representations[7]. The family of fermionic variables (1)–(3) that satisfies the explained selection criterion is totally specified by the following asymtpotic behavior in the limit of large $\omega_k$:

$$f_1^k = \sqrt{\frac{\xi_k-\omega_k}{2\xi_k}} + \frac{\tilde{M}e^{\alpha}}{2\omega_k}\left[e^{iF_2^k}-1\right] + \theta_k \quad \text{with} \quad \sum_{\vec{k}\in\mathbb{Z}^3}|\theta_k|^2 < \infty, \tag{4}$$

where we have defined

$$\xi_k = \sqrt{\omega_k^2 + \tilde{M}^2 e^{2\alpha}}, \tag{5}$$

and $\tilde{M} = 2M\sqrt{\pi/(3l_0^3)}$ where $M$ is the mass of the Dirac field.

The family of annihilation and creationlike variables defined by (1)–(3), together with condition (4), is obtained by means of an $(\alpha,\pi_\alpha)$-dependent transformation that is canonical within the fermionic sector of the phase space. In order to be canonical in the entire truncated system, the geometric variables $(\alpha,\pi_\alpha)$ of the homogeneous sector must be replaced with a new, corrected, canonical pair $(\tilde{\alpha},\tilde{\pi}_\alpha)$. These new homogeneous variables differ from the old ones by some known corrections that are quadratic in the fermionic perturbations[5]. At the quadratic order of our perturbative truncation, the expression of the total Hamiltonian of the cosmological system in terms of these new variables can be obtained by directly substituting the new pair $(\tilde{\alpha},\tilde{\pi}_\alpha)$ in its functional dependence on $(\alpha,\pi_\alpha)$, and replacing the Dirac Hamiltonian $N_0H_D$ by $N_0\tilde{H}_D$, where $\tilde{H}_D$ is the sum over all $\vec{k}\in\mathbb{Z}^3$ of some functions $\tilde{H}_{\vec{k}}$ that, when $\omega_k\to\infty$, behave as[5,8]:

$$\begin{aligned}\tilde{H}_{\vec{k}} = \sum_{(x,y)}\Bigg[&\left(e^{-\tilde{\alpha}}\xi_k + 2h_D^k\right)\left(\bar{a}_{\vec{k}}^{(x,y)}a_{\vec{k}}^{(x,y)} + \bar{b}_{\vec{k}}^{(x,y)}b_{\vec{k}}^{(x,y)}\right) + 2h_J^k\left(\bar{b}_{\vec{k}}^{(x,y)}b_{\vec{k}}^{(x,y)}\right)\\ &+ e^{i(J_k-F_2^k)}e^{-\tilde{\alpha}}\left(2\omega_k\bar{\theta}_k + \bar{h}_I^k\right)a_{\vec{k}}^{(x,y)}b_{\vec{k}}^{(x,y)} + \text{H.c.}\Bigg],\end{aligned} \tag{6}$$

where $\xi_k$ is evaluated at $\tilde{\alpha}$, H.c. stands for Hermitian (complex) conjugate, and

$$h_D^k = \mathcal{O}(1), \quad h_J^k = \mathcal{O}(1), \quad h_I^k = i\tilde{\pi}_\alpha\frac{\tilde{M}e^{-\tilde{\alpha}}}{2\omega_k}e^{iF_2^k} + \mathcal{O}\left(\text{Max}[\theta_k,\omega_k^{-2}]\right). \tag{7}$$

Here, a contribution is $\mathcal{O}(.)$ when it is of the asymptotic order of the corresponding argument (or smaller), and the function Max[.,.] picks out the argument of dominant asymptotic order. To arrive at these asymptotic expressions, we have further restricted to functions $f_1^k, f_2^k, g_1^k$, and $g_2^k$ with asymptotic expansions (when $\omega_k \to \infty$) such that their partial derivatives with respect to $\alpha$ and $\pi_\alpha$ maintain the asymptotic order of each term of those expansions.

## 4. Backreaction and Unitarity in Hybrid LQC

The two sectors of the phase space to be quantized are the homogeneous FLRW background [with geometry described by $(\tilde{\alpha}, \tilde{\pi}_\alpha)$] and the fermionic degrees of freedom described above. For the former, we select a LQC-inspired representation[3,5] with kinematical Hilbert space that we denote by $\mathcal{H}_{\rm kin}^{\rm grav}$. On the other hand, for the inflaton field $\phi$ and its momentum we choose a standard Schrödinger representation, with Hilbert space given by $L^2(\mathbb{R}, d\phi)$. Finally, for the fermionic perturbations we consider the Fock representation associated with any choice of annihilation and creationlike variables within the family defined by (1)-(4). We call $\mathcal{F}_D$ the corresponding Fock space. Besides $\hat{a}_{\vec{k}}^{(x,y)}$ and $\hat{b}_{\vec{k}}^{(x,y)\dagger}$ respectively denote the annihilation operators of particle excitations and the creation operators of antiparticle excitations, with their adjoints acting reversely. Both of these sectors are jointly subject to the zero mode of the Hamiltonian constraint, formed by the sum of the constraint of the unperturbed inflationary model and the new fermionic contribution formed by $\tilde{H}_D$. In the hybrid approach, one attempts to represent this constraint on the resulting tensor product space, and then imposes that it annihilates the physical states of the system[5].

In order to find physical states of interest, we consider wave functions of the form $\Gamma(V,\phi)\psi_D(\mathcal{N}_D,\phi)$, where $V$ is the physical volume of the homogeneous sector of our cosmology, and $\mathcal{N}_D$ refers to the fermionic occupation numbers. Moreover, we restrict our considerations to normalized states $\Gamma$ in $\mathcal{H}_{\rm kin}^{\rm grav}$ with a unitary evolution in $\phi$, which furthermore is generated by certain positive operator[5]. The only restriction that we impose on this operator is to be such that the partial state $\Gamma$ differs from being an exact solution of the unperturbed FLRW cosmology at most in a quadratic contribution of the perturbations. With this ansatz for the states, we impose the Hamiltonian constraint. If in the partial state $\Gamma$ we can ignore any transition in the homogeneous geometry mediated by the action of our quantum Hamiltonian constraint, and the contribution of the perturbations to the momentum of the inflaton is negligible with respect to that of $\Gamma$, a sort of Born-Oppenheimer approximation can be applied. Then, one arrives at the following Schrödinger-like equation for the fermionic part of the wave function[5]:

$$i\partial_\phi \psi_D(\mathcal{N}_D,\phi) = \frac{l_0\langle \widehat{V^{2/3}e^{\tilde{\alpha}}\tilde{H}_D}\rangle_\Gamma - C_D^{(\Gamma)}(\phi)}{\langle \hat{\tilde{\mathcal{H}}}_0\rangle_\Gamma}\,\psi_D(\mathcal{N}_D,\phi). \tag{8}$$

Here, the hat over classical observables indicates their corresponding representation as operators. Besides, the brackets $\langle . \rangle_\Gamma$ stand for the expectation value in $\Gamma$, taken with respect to the inner product in $\mathcal{H}^{\rm grav}_{\rm kin}$. Since the momentum of $\phi$ does not appear in $\tilde{H}_D$, the right-hand side of (8) represents a $\phi$-dependent Hamiltonian operator acting on the fermionic sector. On the other hand, the function $C_D^{(\Gamma)}(\phi)$ is a backreaction term that provides, in mean value, how much $\Gamma$ departs from an exact solution of the unperturbed system[8].

The solutions of (8) can be studied by considering its associated Heisenberg equations for the annihilation and creation operators. Taking initial data $\hat{a}_{\vec{k}}^{(x,y)}$ and $\hat{b}_{\vec{k}}^{(x,y)\dagger}$ at $\phi = \phi_0$, these are generally solved by the linear combinations:

$$\begin{aligned}
\hat{a}_{\vec{k}}^{(x,y)}(\phi,\phi_0) &= \alpha_k(\phi,\phi_0)\hat{a}_{\vec{k}}^{(x,y)} + \beta_k(\phi,\phi_0)\hat{b}_{\vec{k}}^{(x,y)\dagger},\\
\hat{b}_{\vec{k}}^{(x,y)\dagger}(\phi,\phi_0) &= -e^{i\chi_k(\phi,\phi_0)}\bar{\beta}_k(\phi,\phi_0)\hat{a}_k^{(x,y)} + e^{i\chi_k(\phi,\phi_0)}\bar{\alpha}_k(\phi,\phi_0)\hat{b}_{\vec{k}}^{(x,y)\dagger},
\end{aligned} \tag{9}$$

where $|\alpha_k|^2 + |\beta_k|^2 = 1$ and, in the asymptotic limit $\omega_k \to \infty$[8],

$$\chi_k = \mathcal{O}(1), \quad \alpha_k = \mathcal{O}(1), \quad \beta_k = \mathcal{O}(G_k^{(\Gamma)}\omega_k^{-1}) = \mathcal{O}\left(\mathrm{Max}[\theta_k, \omega_k^{-2}]\right) \tag{10}$$

with:

$$G_k^{(\Gamma)} = \frac{2i\omega_k\langle \widehat{e^{i(F_2^k - J_k)}V}^{2/3}\theta_k\rangle_\Gamma + i\langle \widehat{e^{i(F_2^k - J_k)}V}^{2/3}h_I^k\rangle_\Gamma}{\langle \hat{V}^{2/3}\rangle_\Gamma}. \tag{11}$$

The Heisenberg evolution is thus described by a $\phi$-dependent family of Bogoliubov transformations, which are all unitarily implementable in $\mathcal{F}_D$ since the sequence of beta coefficients is clearly square summable[9] over all $\vec{k} \in \mathbb{Z}^3$. It is actually possible to explicitly construct the corresponding unitary operators, which are completely specified by the alpha and beta coefficients, up to an arbitrary phase. The state obtained from evolving the vacuum in $\mathcal{F}_D$ with these operators then indeed solves (8), with a backreaction term given by[8]:

$$C_D^{(\Gamma)} = l_0\langle \hat{V}^{2/3}\rangle_\Gamma \sum_{\vec{k},(x,y)} \Im(G_k^{(\Gamma)}\bar{\Delta}_k), \quad \Delta_k = \mathcal{O}(\beta_k), \tag{12}$$

up to a term coming from the $\phi$-derivative of the mentioned arbitrary phase in the evolution operator. Our aim is to obtain a backreaction contribution which needs not any regularization techniques by means of, e.g., a choice of the arbitrary phase that absorbs any possible divergences in (12), or an specific conditional summation rule that ensures the convergence of such equation. We therefore impose that the sum that specifies $C_D^{(\Gamma)}$ is absolutely convergent. Besides, we naturally require that this contribution to the backreaction converge independently of the choice of homogeneous state $\Gamma$, and for all $\phi$. This univocally leads to impose[8]:

$$\theta_k = -i\frac{\tilde{M}e^{-\alpha}}{4\omega_k^2}\pi_\alpha e^{iF_2^k} + \vartheta_k, \quad \sum_{\vec{k}\in\mathbb{Z}^3} \omega_k|\vartheta_k|^2 < \infty. \tag{13}$$

Finally, let us comment that one can restrict even further the choice of fermionic variables in order to guarantee that the Hamiltonian operator that appears in the Schrödinger equation (8) has, with normal ordering, a well-defined action on the Fock vacuum. This amounts to require that the coefficients of the $a_{\vec{k}}^{(x,y)} b_{\vec{k}}^{(x,y)}$ terms in $\tilde{H}_D$ form a square summable sequence over all $\vec{k} \in \mathbb{Z}^3$. This happens if and only if (13) is again satisfied, but now requiring the stronger condition: $\sum_{\vec{k}} \omega_k^2 |\vartheta_k|^2 < \infty$[8].

## 5. Conclusions

We have studied how to split the phase space of an inflationary cosmology minimally coupled to a perturbative Dirac field in such a way that, within the hybrid approach to LQC, one obtains a well-defined fermionic contribution to the quantum backreaction without the need of any regularization scheme. This criterion leads to a very specific choice of annihilation and creationlike variables for the fermionic degrees of freedom, something that amounts to a restriction on the dynamical characterization of their quantum excitations, as well as on the choice of their vacua.

## Acknowledgments

This work was supported by Grants No. FIS2014-54800-C2-2-P and No. FIS2017-86497-C2-2-P from Spain.

## References

1. R.M. Wald, *Quantum field theory in curved spacetimes and black hole thermodynamics* (University of Chicago Press, Chicago, 1994).
2. M. Martín-Benito, L.J. Garay, and G.A. Mena Marugán, *Phys. Rev. D* **78**, 083516 (2008); B. Elizaga Navascués, M. Martín-Benito, and G.A. Mena Marugán, *Int. J. Mod. Phys. D* **25**, 1642007 (2016).
3. M. Bojowald, *Living Rev. Relativity* **11**, 4 (2008); A. Ashtekar, T. Pawłowski, and P. Singh, *Phys. Rev. D* **74**, 084003 (2006); M. Martín-Benito, G.A. Mena Marugán, and J. Olmedo, *Phys. Rev. D* **80**, 104015 (2009).
4. P.D. D'Eath and J.J. Halliwell, *Phys. Rev. D* **35**, 1100 (1987).
5. B. Elizaga Navascués, M. Martín-Benito, and G.A. Mena Marugán, *Phys. Rev. D* **96**, 044023 (2017).
6. H.B. Lawson and M.L. Michelson, *Spin geometry* (Princeton University Press, Princeton, 1989); Th. Friedrich, *Colloq. Mathematicum* **48**, 57 (1984).
7. J. Cortez, B. Elizaga Navascués, M. Martín-Benito, G.A. Mena Marugán, and J.M. Velhinho, *Ann. Phys. (Amsterdam)* **376**, 76 (2017).
8. B. Elizaga Navascués, G.A. Mena Marugán, and S. Prado Loy, *Phys. Rev. D* **98**, 063535 (2018).
9. D. Shale, *Trans. Am. Math. Soc.* **103**, 149 (1962); J. Dereziński, *Lect. Notes Phys.* **695**, 63 (2006).

# Detailed background dynamics and trans-planckian effects in loop quantum cosmology

Killian Martineau

*Laboratoire de Physique Subatomique et de Cosmologie, Université Grenoble-Alpes, Grenoble, 38000, France*
*E-mail: killian.martineau@lpsc.in2p3.fr*
*http://lpsc.in2p3.fr/index.php/en/*

Cosmology appears as the most promising way to test and constrain quantum gravity theories. Loop quantum gravity is among the most advanced attempts to perform a non-perturbative quantization of general relativity. Its cosmological counterpart, loop quantum cosmology, has clear predictions both for the cosmological background and for the perturbations. In particular, the initial Big Bang singularity is replaced by a bounce due to quantum geometry effects. In this proceeding I will focus on new results obtained in loop quantum cosmology: i) the prediction of the duration of inflation as a function of all the unknown parameters of the model and ii) new primordial power spectra obtained with modified dispersion relations accounting for trans-planckian effects.

*Keywords*: Loop quantum cosmology, inflation, trans-planckian problem, primordial power spectra.

## 1. Introduction

Loop quantum gravity [1, 2] (LQG) might be the most advanced non-perturbative and background-invariant quantization of general relativity. Loop quantum cosmology [3] (LQC) is a quantum theory inspired by the LQG quantization scheme that takes into account the cosmological symmetries. Its background dynamics, governed by the effective modified Friedmann equation

$$H^2 = \frac{\kappa}{3}\rho\left(1 - \frac{\rho}{\rho_c}\right), \tag{1}$$

leads to a bounce and is now well established. It has been confirmed by numerical simulations [4], but also by calculations in group field theory [5], underlining the robustness of the model. In this equation $H$ is the Hubble parameter, $\kappa = 8\pi$, $\rho$ stands for the energy density, and $\rho_c \sim \rho_{Pl}$ for its value at the bounce.

The situation is however less clear for perturbations. Not only because different settings are being considered (mainly the *deformed algebra* [6] and the *dressed metric* [7] approaches) but most importantly because trans-planckian effects are often neglected or not treated as a dominant process. It seems that too much emphasis has been put on *density* effects and not enough on *length* effects. Even if the former do trigger the bounce the latter cannot be ignored in a theory that modifies the space structure at the Planck length. In this proceeding I will therefore focus on two new results in LQC : i) the prediction of the duration of inflation as a function of all the unknown parameters of the model and ii) new primordial power spectra obtained with modified dispersion relations accounting for trans-planckian effects.

## 2. Exhaustive investigation of the inflation duration in the loop quantum cosmological framework

Remarkably, inflation is a strong attractor and appears naturally in the LQC model as long as the matter content of the universe is assumed to be a massive scalar field. One of the most interesting consequence of this cosmological framework is therefore that the inflation duration can, to some extent, be predicted.

However, even at the background level, three main uncertainties remain to be addressed systematically. The first one is related with the way to set initial conditions. There are mainly two schools of thought : the first one sets them in the remote past of the contracting branch whereas the other one sets them at the bounce. Since those initial conditions are in a one-to-one correspondence with each other the key point is to look for a variable to which a known probability distribution function (PDF) can be assigned. The second uncertainty is associated with the amount of anisotropic shear at the bounce, which is expected to play an important role in any bouncing model. Finally the third main uncertainty is the inflaton potential shape, as the matter content of the universe is not predicted by LQG (but can be partially experimentally determined).

Within this framework we have performed an exhaustive investigation of the duration of inflation by varying those three unknowns [8].

When dealing with the initial conditions issue, it is in our opinion more appealing to set them in the remote past of the classical contracting branch of the Universe. In this case, a flat PDF can easily be associated to the initial phase of the scalar field. If we do so, and if the inflaton potential is confining, the duration of inflation appears to be severely constrained and, most interestingly, the predicted number of e-folds is not much higher than the minimum value required by observations. This result is an important feature of the loop quantum cosmological model and underlines its strong predictive power for a massive scalar field. This predictive power is even increased when anisotropies are taken into account as the number of e-folds associated with the PDF mean value decreases. However the LQC predictive power is basically lost if initial conditions are set at the bounce as there is apparently no variable to which a known PDF can be assigned.

In summary, once the inflaton potential is determined and if initial conditions are set in the contracting branch, in agreement with the intuition of causality, there is an obviously interesting predictive power of LGC for the inflation duration. This predictive power is stronger when the potential is confining and weaker if it features a plateau-like shape. The initial amount of shear remains however unknown but this is not necessarily a problem if the inflaton potential is confining as the predicted $e$-folds number is then restricted to a small interval, bounded from below by observations and from above by the isotropic case.

## 3. A first step towards the inflationary trans-planckian problem treatment in loop quantum cosmology

The trans-planckian problem is a well-known cosmological issue [9] that cannot be ignored in the LQG framework. As soon as the number of e-folds of inflation is higher than 70, all modes of physical interest were highly trans-planckien at the bounce time.

As a first elementary step to account for trans-planckian effects, we have suggested the use of modified dispersion relations (MDRs) in the LQC framework [10]. This is obviously not the final word on this question and trans-planckian effects should ideally be considered in the full theory setting but we believe that MDRs are a meaningful first step.

Modified relations need to be applied to physical quantities and are therefore introduced in the Mukhanov-Sasaki equation by the replacement: $k_\varphi \to \mathcal{F}(k_\varphi)$. The $\mathcal{F}(k_\varphi)$ function depends on the considered dispersion relation, the standard case corresponding to $\mathcal{F}(k_\varphi) = k_\varphi$. The Mukhanov-Sasaki equation in the Deformed Algebra approach then becomes:

$$v_k''(\eta) + \left(\Omega(\eta)a^2(\eta)\mathcal{F}(k_\varphi)^2 - \frac{z_{T/S}''(\eta)}{z_{T/S}(\eta)}\right) v_k(\eta) = 0, \tag{2}$$

where $\Omega(\eta) = 1 - 2\rho(\eta)/\rho_c$. It is usually thought that the Deformed Algebra approach is excluded by data because of an exponential increase of the power spectra in the ultraviolet regime due to the $\Omega$ factor. One should however keep in mind that what is excluded is just a very specific scenario in which anisotropies, backreaction, and more importantly the trans-planckian problem, have been neglected. Namely we have shown that considering the MDR

$$\mathcal{F}(k_\varphi) = k_0 \tanh\left[\left(\frac{k_\varphi}{k_0}\right)^p\right]^{\frac{1}{p}} = k_0 \tanh\left[\left(\frac{k_c}{a(t)k_0}\right)^p\right]^{\frac{1}{p}}, \tag{3}$$

where $k_0$ is the transition scale in physical coordinates and $p$ determines the sharpness of the transition, cures the pathological behaviour of the deformed algebra, and leads to almost scale-invariant spectra in the observed regime both for tensor and scalar perturbations, as it can be seen in Figs. 1 and 2.

This is a significant result in the sense that the Deformed Algebra approach is now shown, when trans-planckian effects are taken into account, to be possibly in agreement with data. This also shows that the spectra are actually highly sensitive to quantum gravity effects and a better understanding of the trans-planckian behaviour from the full LQG theory is necessary to go ahead in this direction.

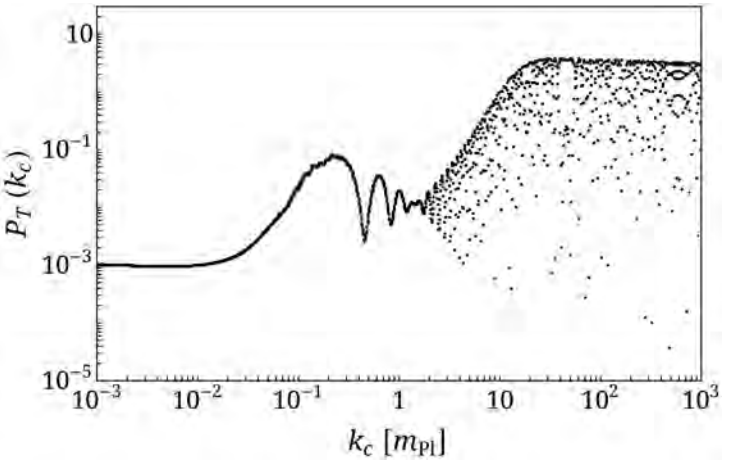



Fig. 1. Primordial tensor power spectrum in the deformed algebra approach with the Unruh-like MDR $\mathcal{F}(k_\varphi) = k_0 \tanh\left[(k_c/(a(t)k_0))^p\right]^{1/p}$, $k_0 = 10$ and $p = 1$.

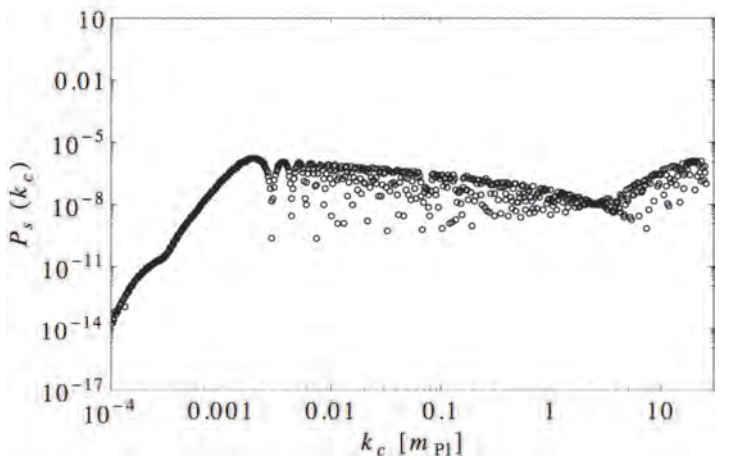



Fig. 2. Primordial scalar power spectrum in the deformed algebra approach with the Unruh-like MDR $\mathcal{F}(k_\varphi) = k_0 \tanh\left[(k_c/(a(t)k_0))^p\right]^{1/p}$, $k_0 = 10$ and $p = 1$.

## Acknowledgments

K.M is supported by a grant from the C.F.M foundation.

## References

1. C. Rovelli and F. Vidotto, *Covariant Loop Quantum Gravity* (Cambridge University Press, 2014).
2. R. Gambini and J. Pullin, *A First Course in Loop Quantum Gravity* (Oxford University Press, 2011).

3. A. Ashtekar and S. Parampreet, *Loop Quantum Cosmology: A Status Report* (Class. Quant. Grav, vol. 28, 2011).
4. P. Diener and B. Gupt and S. Parampreet, *Chimera: A hybrid approach to numerical loop quantum cosmology* (Class. Quant. Grav, vol. 31, 2014).
5. D. Oriti and L. Sindoni and E. Wilson-Ewing, *Emergent Friedmann dynamics with a quantum bounce from quantum gravity condensates* (Class. Quant. Grav, vol. 33, 2016).
6. M. Bojowald and G.M. Paily, *Deformed General Relativity and Effective Actions from Loop Quantum Gravity* (Phys. Rev., vol. D86, 2012).
7. I. Agullo and A. Ashtekar and W. Nelson, *Extension of the quantum theory of cosmological perturbations to the Planck era* (Phys. Rev., vol. D87, 2013).
8. K. Martineau and A. Barrau and S. Schander, *Detailed investigation of the duration of inflation in loop quantum cosmology for a Bianchi-I universe with different inflaton potentials and initial conditions* (Phys. Rev, vol. D95, 2017).
9. J. Martin and R.H. Brandenberger, *The TransPlanckian problem of inflationary cosmology* (Phys. Rev., vol. D63, 2001).
10. K. Martineau and A. Barrau and J. Grain, *A first step towards the inflationary trans-planckian problem treatment in Loop Quantum Cosmology* (Int. J. Mod. Phys., vol. D27, 2018).

# The mass of cosmological perturbations in Loop Quantum Cosmology

Guillermo A. Mena Marugán*, Daniel Martín de Blas
*Instituto de Estructura de la Materia, CSIC, Serrrano 121, Madrid, Spain*

and Beatriz Elizaga de Navascués
*Institute for Quantum Gravity, Friedrich-Alexander University Erlangen-Nürnberg, Staudstraße 7, 91058 Erlangen, Germany*

Loop Quantum Cosmology has recently been applied to extend the analysis of primordial perturbations to the Planck era. Two approaches to Loop Quantum Cosmology leading to predictions that can be compatible with observations are the so-called hybrid and dressed metric approaches. In spite of their similarities, we show that the effective equations that they provide for the evolution of the perturbations are different. When backreaction is neglected, the discrepancy appears only in the time-dependent mass term of the corresponding field equations. We explain the origin of this difference.

*Keywords*: Loop Quantum Cosmology, Cosmological Perturbations.

## 1. Introduction

The present era of precision cosmology is opening potential windows to falsify alternative theories of gravity, including those that contain quantum effects. One of these theories is Loop Quantum Cosmology (LQC).[1] By applying the methods of Loop Quantum Gravity, LQC enables us to extend the study of cosmological perturbations beyond the onset of inflation. In particular, in LQC there exist physical states that are peaked on effective trajectories which avoid the Big Bang singularity by means of a quantum bounce, known as the Big Bounce.[1]

Two approaches to LQC that have a conventional ultraviolet behaviour and have led so far to predictions compatible with the observations are the so-called hybrid[2–5] and dressed metric[6] approaches. In both of them, quantum geometry effects around the bounce change not only the cosmological evolution, but also the propagation of the perturbations and their vacuum. Moreover, in both cases the primordial perturbations can be described by a collection of "harmonic" oscillators with a time-dependent mass. The aim of this work is to show that those masses are different, and therefore the evolution of the perturbations and the imprints left on them by quantum effects can differ. We use units with $c = \hbar = 1$.

## 2. Cosmological Perturbations in Loop Quantum Cosmology

We start with a homogeneous and isotropic cosmology with flat compact sections, with metric $ds^2 = -N_0^2(t)dt^2 + a^2(t)\ {}^0h_{ij}d\theta_i d\theta_j$. The angular coordinates $\theta_i$ have a period $l_0$, $a$ is the scale factor, ${}^0h_{ij}$ the Euclidean metric, and $N_0$ the lapse. We include a homogeneous scalar field, the inflaton, subject to a potential $V(\phi)$.

The phase space of this cosmology can be described with the canonical pairs $(a, \pi_a)$ and $(\phi, \pi_\phi)$. In LQC it is usual to replace $(a, \pi_a)$ with the pair[1] $(v, b/2)$,

where $v$ is the volume of the spatial sections, up to a constant, and $b$ is proportional to the Hubble parameter; $a^3 = 2\pi G\gamma\sqrt{\Delta_g}|v|/l_0^3$ and $a\pi_a = -3vb/2$. The sign of $v$ determines the orientation of the triad, $\gamma$ is the Immirzi parameter, and $\Delta_g$ is the area gap (the minimum positive area eigenvalue allowed by Loop Quantum Gravity).[1]

To consider small anisotropies and inhomogeneities, we next introduce perturbations in the geometry and the inflaton.[2] We expand them in scalar, vector, and tensor harmonics (of the spatial Laplacian) taking advantage of the background symmetries.[3] This perturbed cosmology is the system that we are going to analyze.

In the hybrid approach, one starts the analysis by truncating the action at quadratic order in perturbations. Zero modes are treated exactly at this order. After the truncation, the system can be shown to be symplectic and subject to constraints, inherited from General Relativity (GR).[3] Moreover, at our perturbation order, one can introduce canonical variables that include a gauge-invariant set for the perturbations [namely, the tensor modes and the Mukhanov-Sasaki (MS) scalar] and corrected zero modes, the rest of variables being perturbative constraints and gauge degrees of freedom.[3] The corrected zero modes differ from the original ones in quadratic perturbative contributions that can be understood as backreaction terms. In the following, we will adopt the same notation for the original and the corrected zero modes, to simplify the discussion and because we will assume that backreaction does not play an important role in our states.

Physical states depend only on zero modes and perturbative gauge invariants. In the hybrid approach, these states are still subject to a global constraint. It is the zero mode of the Hamiltonian constraint, obtained from that of the homogenous cosmology, $H_0$, by adding to it the Hamiltonians of the tensor and MS modes. To find solutions, it is common to adopt an ansatz with separation of variables in the different gauge invariants and the homogeneous geometry, with the inflaton viewed as an internal time (so that each partial state can vary with it). Moreover, as part of the ansatz, we consider states of the homogeneous geometry that are: 1) supported on the superselection sectors of homogenous LQC;[3] 2) solutions to homogeneous LQC (thus, we ignore the possible backreaction of the perturbations); and 3) peaked on trajectories of the so-called effective LQC dynamics.[1]

We recall that in LQC the connection variable $b$ does not admit a straightforward operator representation, and needs to be represented in terms of holonomies, which have elements of the form $e^{\pm ib/2}$. In practice, this has the effect of replacing the powers of $b$ in the Hamiltonian constraint by powers of sine functions. For the homogeneous cosmology, instead of $H_0$, the resulting effective Hamiltonian is $H_0^{eff} = [\pi_\phi^2 - 3\pi G v^2 \sin^2 b + 8\pi^2 G^2\gamma^2\Delta_g v^2 V(\phi)]/(4\pi G\gamma\sqrt{\Delta_g}v)$. The corresponding dynamics corrects the trajectories of GR.[1] In the rest of our discussion, we are going to consider precisely such effective trajectories.

If we define the density and pressure of the inflaton as in the homogenous case, i.e. $\rho = V(\phi) + \pi_\phi^2/(8\pi^2 G^2\gamma^2\Delta_g v^2)$ and $P = \rho - 2V(\phi)$, from the fact that the

volume has a minimum at the bounce one can check that the density reaches there a maximum,[1] known as the critical density: $\rho_{max} = 3/(8\pi G\gamma^2\Delta_g)$. Note also that, since the kinetic energy of the inflaton is positive, the potential contribution can never be larger than the density, and hence $V(\phi) \leq \rho_{max}$.

On effective trajectories of LQC, one gets in the hybrid approach a constraint for the perturbations that, in conformal time (with the corresponding derivative indicated with a prime), leads to the following mode equations:[4,5]

$$d''_{\vec{k},\epsilon} + \left[k^2 - \frac{4\pi G}{3}a^2(\rho - 3P)\right] d_{\vec{k},\epsilon} = 0, \quad v''_{\vec{k}} + \left[k^2 - \frac{4\pi G}{3}a^2(\rho - 3P) + \tilde{U}\right] v_{\vec{k}} = 0,$$

for the tensor modes $d_{\vec{k},\epsilon}$ and the modes $v_{\vec{k}}$ of the MS scalar, respectively. Here, $\vec{k}$ is the (constant) wave vector, $k$ is its norm, and $\epsilon$ indicates the two polarizations of the tensor perturbations. Besides, for the MS scalar, we have defined

$$\tilde{U} = a^2\left[V_{,\phi\phi} + \frac{8\pi G\gamma\sqrt{\Delta_g}}{l_0^3 a^3}\frac{\sin(2b)}{\sin^2 b}\pi_\phi V_{,\phi} + 48\pi G V(\phi) - \frac{128\pi^2 G^2\gamma^2\Delta_g}{\sin^2 b}V^2(\phi)\right].$$

We will refer to this function as the MS potential. The ratio of sine functions in $\tilde{U}$ appears because, in the hybrid quantization, one must adopt a prescription that respects the homogeneous superselection sectors, as we will further explain below.[5] The comma followed by $\phi$ denotes the derivative with respect to the inflaton.

We will call time-dependent mass the k-indepedent term that multiplies the perturbation mode in these equations of harmonic-oscillator type. In particular, for $d_{\vec{k},\epsilon}$ this mass can be rewritten as $4\pi G a^2(3P - \rho)/3 = -(\{a\{a, H_0\}, H_0\})_{eff}$, where the result of the double Poisson bracket with the generator of the homogenous dynamics, $H_0$, must be evaluated on effective trajectories after its computation.[8]

The dynamical equations for the gauge invariant perturbations in GR are actually very similar to those displayed above. In the case of GR, the MS potential is provided by the following slightly different function

$$U = a^2\left[V_{,\phi\phi} - 12\frac{\pi_\phi}{a\pi_a}V_{,\phi} + 48\pi G V(\phi) - \frac{72 l_0^6}{\pi_a^2}a^4 V^2(\phi)\right],$$

and the time-dependent mass of the tensor modes, though given by the same double Poisson bracket, is not evaluated on effective trajectories, but on Einsteinian ones.

Let us now discuss the dressed metric approach. In this case, one first imposes as a constraint the zero mode of the Hamiltonian in homogeneous LQC, $H_0$. Therefore, no backreaction is contemplated. Peaked solutions of this constraint determine a dressed metric, which incorporates quantum corrections. This homogeneous peak trajectory is then lifted to the truncated phase space, obtaining dynamical equations for the tensor and MS modes.[6] In the description of these gauge-invariant perturbations in the dressed metric approach, one often uses mode coefficients $T^{(\epsilon)}_{\vec{k}}$ and $Q_{\vec{k}}$ that are related by a scaling to the variables that satisfy equations of the harmonic-oscillator type: $T^{(\epsilon)}_{\vec{k}} = \sqrt{32\pi G l_0^3}\, d_{\vec{k},\epsilon}/a$ and $Q_{\vec{k}} = l_0^{3/2} v_{\vec{k}}/a$ (this scaling can be completed into a canonical transformation; for details see Ref. 5). After the

scaling one gets, in the corresponding conformal time, propagation equations that are similar to those of the hybrid approach, but with the time-dependent mass of the tensor modes and the MS potential replaced, respectively, with[8]

$$-\{a\{a,H_0^{eff}\},H_0^{eff}\} = 4\pi G a^2 P\left(1-2\frac{\rho}{\rho_{max}}\right) - \frac{4\pi G}{3}a^2\rho\left(1+2\frac{\rho}{\rho_{max}}\right),$$

$$W = a^2\left[V_{,\phi\phi} - 2\sqrt{f}V_{,\phi} + fV(\phi)\right], \quad f = \frac{48\pi G\pi_\phi^2}{\pi_\phi^2 + l_0^6 a^6 V(\phi)}.$$

We end with a caveat. Compatibility of the expression of $W$ with the value of the MS potential in GR requires that the sign of $\sqrt{f}$ coincides with the sign of $\pi_\phi/\pi_a$, which changes from collapsing to expanding branches in effective trajectories.[7,8]

## 3. Differences in The Time-Dependent Mass

We can now specify and explain the difference between the time-dependent masses of the hybrid and the dressed metric approaches. Let us start with the mass of the tensor perturbations. This mass is in fact minus the second derivative of the scale factor with respect to the conformal time, divided by the scale factor.[8] The difference arises because in the hybrid approach one obtains this second derivative in terms of canonical expressions, which are evaluated on the effective trajectory only at the very end of the calculation, whereas in the dressed metric approach the derivatives are computed with the generator of the effective dynamics. Thus, the difference comes from the fact that $(\{a\{a,H_0\},H_0\})_{eff} \neq \{a\{a,H_0^{eff}\},H_0^{eff}\}$.

Consider now the time-dependent mass of the MS scalar which, compared to the mass of the tensor modes, includes the MS potential as an additional contribution. Differences in this potential appear only in the term proportional to $V_{,\phi}$. On the one hand, as we have commented, its coefficient must have the sign of $-\pi_\phi/\pi_a$, something that is granted in the hybrid approach, but that in the dressed metric approach requires a cautious definition of the square root of $f$. On the other hand, this coefficient is not exactly the same in the two approaches. This is due to the different prescriptions adopted to represent the function $1/b$.[8] In fact, in GR the coefficient is $-12a\pi_\phi/\pi_a = 8a^2\pi_\phi/(vb)$. This becomes $8\pi G\gamma\sqrt{\Delta_g}\pi_\phi \sin(2b)/(l_0^3 a\sin^2 b)$ in the hybrid approach to respect the homogeneous superselection sectors (and assure a non-negative denominator).[8] In the dressed metric approach, however, it has the effective counterpart $16\pi G\gamma\sqrt{\Delta_g}\pi_\phi/(l_0^3 a\sin b)$, obtained with the direct promotion of $b$ to $\sin b$ without having to comply with any superselection issue.

It is worth noting that the differences that affect the MS potential are not relevant for kinetically dominated solutions, on which the inflaton potential has a small influence. This type of solutions are of the greatest interest in LQC, because they may lead to power spectra for the fluctuations that are compatible with the observations while retaining quantum features at large scales. These features were considered e.g. in Ref. 5 for the case of a quadratic potential, with a mass of $1.2\times10^{-6}$ Planck units. In the following, we will call $s^{(t)}$, $s^{(s)}$, $\breve{s}^{(t)}$, and $\breve{s}^{(s)}$ the time-dependent masses of the tensor (t) and scalar (s) modes in the hybrid and

the dressed metric approaches, respectively. With kinetic dominance, therefore, the tensor and scalar masses are almost identical. However, the masses of the hybrid and the dressed metric approaches are in general significantly different around the bounce, owing to the effects of the two distinct generators of the homogenous dynamics.[8]

On the other hand, in the effective trajectories of LQC the Big Bounce provides a distinguished spatial section that can be used to set initial conditions. The positivity of the time-dependent mass at the bounce becomes then a relevant issue, because the initial oscillatory or exponential behavior of the modes far from the ultraviolet sector depends just on it. Moreover, a negative mass can be an obstacle to impose initial conditions corresponding to adiabatic vacuum states, determined by a frequency $\Omega_k$ that is defined iteratively as a solution of a differential equation, starting with the value $\sqrt{k^2+s}$, where $s$ is the time-dependent mass.[4,5]

Since the density of the inflaton equals its critical value $\rho_{max}$ at the bounce, the pressure is determined there by the potential. Therefore, the tensor mass depends only on the potential at the bounce, $V_B$. We get[8] $s_B^{(t)}/(8\pi G a_B^2) = -V_B + \rho_{max}/3$ and $\breve{s}_B^{(t)}/(8\pi G a_B^2) = V_B - \rho_{max}$, where $a_B$ is the scale factor at the bounce. From this, we see that the mass is positive in the hybrid approach for $V_B \in [0, \rho_{max}/3]$, which includes the interesting sector of kinetic dominance. However, the mass of the dressed metric approach can never be positive at the bounce, since $V_B \leq \rho_{max}$.

We analyze now the scalar mass in the hybrid approach. The corresponding MS potential at the bounce has a quadratic expression in terms of $V_B$, namely $U_B = a_B^2[V^B_{,\phi\phi} + 48\pi G V_B - 128\pi^2 G^2\gamma^2\Delta_g V_B^2]$. Let us consider non-negative inflaton potentials, and treat the second derivative at the bounce $V^B_{,\phi\phi}$ as a non-negative parameter. Actually, in the case of a quadratic potential, this derivative is just the inflaton mass. The expression of $s^{(s)}$ is also a quadratic polynomial, with only one positive root, $x_+ = (5 + \sqrt{33 + 8\gamma^2\Delta_g V^B_{,\phi\phi}})\rho_{max}/12$. It is not difficult to check[8] that, with our assumptions, the mass is positive for $V_B \leq x_+$. This includes the region of kinetic dominance. Besides, since $x_+ \geq x_+(V^B_{,\phi\phi} = 0) \approx 0.895$, we can assure that the mass is always positive for $V_B \in [0, 0.895\,\rho_{max}]$.

For the case of the dressed metric approach, the MS potential at the bounce is $W_B = a_B^2[V^B_{,\phi\phi} \pm 16\pi G\gamma\sqrt{\Delta_g}(|\phi'_B|/a_B)V^B_{,\phi} + 48\pi G V_B - 128\pi^2 G^2\gamma^2\Delta_g V_B^2]$. Let us assume that $\left|V^B_{,\phi}\right| \leq C\sqrt{2V^B_{,\phi\phi}V_B}$, for a certain quantity $C \equiv C(V^B_{,\phi\phi})$ of the order of the unit. Actually, for a mass term, we have $C = 1$. Then we can bound the term of $V^B_{,\phi}$ and, defining the two polynomials of $V_B$ (with $V^B_{,\phi\phi}$ regarded as a parameter) $P_\pm = \breve{s}_B^{(t)} + a_B^2[V^B_{,\phi\phi} + 48\pi G V_B \pm 6C\sqrt{V^B_{,\phi\phi}}/(\gamma\sqrt{\Delta_g}) - 128\pi^2 G^2\gamma^2\Delta_g V_B^2]$, we conclude[8] that $P_- \leq \breve{s}_B^{(s)} \leq P_+$. Moreover, for small $\gamma^2\Delta_g V^B_{,\phi\phi}$, the roots of these polynomials satisfy $y_-(P_+) \leq y_-(P_-) \leq y_+(P_-) \leq y_+(P_+)$. Indeed, for a mass term as the inflaton potential, and values of the inflaton mass close to that considered in Ref. 5, one gets $\gamma^2\Delta_g V^B_{,\phi\phi} \approx 10^{-12}$, which is certainly very small. For such negligible values, the lowest root of the analyzed polynomials is $y_-(P_+) \approx \rho_{max}/6$.

As a consequence, the time-dependent mass of the scalar modes in the dressed metric approach is negative for $V_B \in [0, y_-(P_+)] \approx [0, \rho_{max}/6]$. In addition, for negligible $\gamma^2 \Delta_g V^B_{,\phi\phi}$, one can show that $\breve{s}^{(s)}_B$ is negative in the interval $[y_+(P_+), \rho_{max}]$ around the region of potential dominance.[8]

## 4. Conclusions

We have seen that the hybrid and the dressed metric approaches lead to different time-dependent masses for the cosmological perturbations, and therefore they predict different departures from GR in the early evolution of the primordial fluctuations. The most relevant difference is a consequence of the tension between the canonical methods and the effective dynamics when computing time derivatives in the expression of the tensor mass. There exist other differences arising from the distinct quantization prescriptions adopted for the term proportional to the derivative of the inflaton potential in the scalar mass, but this are not important in scenarios with kinetic dominance. Particularizing our study to the Big Bounce, we have seen that the time-dependent mass is positive in the hybrid approach in an ample sector of solutions that contains all kinetically dominated ones. Around those solutions, however, the mass is negative for the dressed metric approach. This negativity may be an obstruction for the definition of initial conditions for certain vacua.

## Acknowledgments

This work was supported by Project. No. MINECO FIS2014-54800-C2-2-P and Project. No. MINECO FIS2017-86497-C2-2-P from Spain.

## References

1. A. Ashtekar and P. Singh, *Classical Quantum Gravity* **28**, 213001 (2011).
2. B. Elizaga Navascués, M. Martín-Benito, and G.A. Mena Marugán, *Int. J. Mod. Phys. D* **25**, 1642007 (2016); M. Fernández-Méndez, G.A. Mena Marugán, and J. Olmedo, *Phys. Rev. D* **86**, 024003 (2012); *Phys. Rev. D* **88**, 044013 (2013).
3. L. Castelló Gomar, M. Fernández-Méndez, G.A. Mena Marugán, and J. Olmedo, *Phys. Rev. D* **90**, 064015 (2014); L. Castelló Gomar, M. Martín-Benito, and G.A. Mena Marugán, *JCAP* **1506** (2015) 045.
4. D. Martín de Blas and J. Olmedo, *JCAP* **1606** (2016) 029.
5. L. Castelló Gomar, G.A. Mena Marugán, D. Martín de Blas, and J. Olmedo, *Phys. Rev. D* **96**, 103528 (2017).
6. I. Agullo, A. Ashtekar, and W. Nelson, *Phys. Rev. Lett.* **109**, 251301 (2012); *Phys. Rev. D* **87**, 043507 (2013); *Classical Quantum Gravity* **30**, 085014 (2013).
7. T. Zhu, A. Wang, K. Kirsten, G. Cleaver, and Q. Sheng, *Phys. Rev. D* **96**, 083520 (2017).
8. B. Elizaga Navascués, D. Martín de Blas, and G.A. Mena Marugán, *Phys. Rev. D* **97**, 043523 (2018).

# Some aspects of black hole physics in loop quantum gravity

Flora Moulin

*Laboratoire de Physique Subatomique et Cosmology, Université Grenoble-Alpes, CNRS/IN2P3*
*53, avenue des Martyrs 38026 Grenoble, France*
** E-mail: moulin@lpsc.in2p3.fr*
*www.univ-grenoble-alpes.fr*

Some recent results on black holes in effective loop quantum gravity are presented. Quantum gravity effects might allow the transition of a black hole into a white hole, when the Planck density is reached. I briefly review previous studies and focus on the random nature of the bouncing lifetime which was not previously taken into account. I show that, when we consider a stochastic lifetime, the signal emitted by bouncing black holes might explain fast radio bursts. Then, I present new results on the absorption cross sections calculated for a quasi Schwarzschild black hole including loop quantum gravity corrections. The black hole geometry deformation due to quantum effects has consequences for the cross sections and the Hawking spectrum.

*Keywords*: Black hole, quantum gravity, quantum field theory.

## 1. Bouncing black holes

This part is based on [1]. A bouncing black hole (BH) is described by a classical collapsing solution which is linked to the classical exploding one by a quantum tunneling. It was argued [2, 3] that a BH of mass M would have a lifetime in the order of $M^2$. Because of the quantum nature of the process, the lifetime of a bouncing BH should be considered as a random variable. The mean lifetime of a BH would be $\tau = kM^2$, with $k$ chosen to be around 0.05 [2]. The probability that a BH has not bounced yet after a time $t$ is given by $P(t) = \frac{1}{\tau}e^{-\frac{t}{\tau}}$. This is similar to the usual nuclear decay behavior. We focus on local effects and consider primordial black holes (PBHs) as we are interested in BH bouncing in the contemporary universe. The number of BHs bouncing after the Hubble time $t_H$ in a time interval $dt$ is

$$dN = \frac{N_0}{kM^2}e^{-\frac{t_H}{kM^2}}dt, \tag{1}$$

where $N_0$ is the initial abundance. The initial differential mass spectrum of the considered PBHs is given by $dN/dM$. Photons emitted have a characteristic wavelength in the order of the size of the BH. The shape of the signal emitted by a single BH is assumed to be given by a Gaussian function:

$$\frac{dN_\gamma^{BH}}{dE} = Ae^{-\frac{(E-E_0)^2}{2\sigma_E^2}}, \tag{2}$$

where $E_0 = 1/(2R_S) = 1/(4M)$, with $R_S$ the Schwarzschild radius. The width is fixed to be $\sigma_E = 0.1E_0$ but the results do not critically depend on this value.

The full signal due to a local distribution of bouncing BHs is given by

$$\frac{dN_\gamma}{dE} = \int_{M_{Pl}}^{\infty} Ae^{-\frac{(E-E_0)^2}{2\sigma_E^2}} \cdot \frac{dN}{dM}(M) \cdot \frac{1}{kM^2} e^{-\frac{t_H}{kM^2}}. \tag{3}$$

We considered two types of mass spectra for the PBHs. The first one is peaked around a value $M_0$ (4). It is justified if the PBHs are created during a phase transition [4], for exemple. The second one is quite wide (5). It could be explain if PBHs are produced by a scale-invariant density pertubations in a perfect fluid [5]. In that case the spectrum is proportional to $M^\alpha$. Here we considered $\alpha$ as an unknown parameter.

$$\frac{dN}{dM} \propto e^{-\frac{(M-M_0)^2}{2\sigma_M^2}}, \tag{4}$$

$$\frac{dN}{dM} \propto M^\alpha. \tag{5}$$

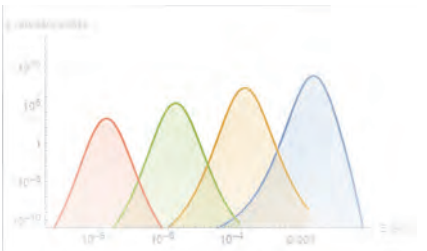


Fig. 1. Differential electromagnetic flux emitted by bouncing PBHs for a central mass $M_0$ equal (from right to left) to $M_{t_H}$, $10M_{t_H}$, $100M_{t_H}$, and $1000M_{t_H}$.

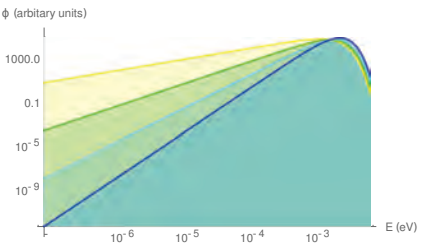



Fig. 2. Signal expected from a wide mass spectrum, with $\alpha = \{-3, -2, -1, 0\}$ from the lower curve to the upper curve at $10^{-6}$ eV.

In Fig. 1, the expected emitted flux is shown for different values of the central mass $M_0$. Actually the mean energy of the signal might not be the one naively expected: $E \sim 1/4M_{t_H}$ with $M_{t_H}$ the mass satisfying $t_H = kM_{t_H}^2$ (corresponding to BHs with a mean lifetime equal to the age of the universe). Because of the stochastic nature of the process, the mean energy of the emitted signal can be different from the one obtained in previous studies. The central mass $M_0$ can in principle be different than $M_{t_H}$. If the mass spectrum is peaked around masses higher than $M_{t_H}$, it is possible to have a signal in the energy range of the fast radio bursts (FRBs): the curve on the left in Fig. 1 is peaked around $10^{-6}$ eV. It was not the case when the lifetime was taken to be deterministic. Then, in Fig. 1, we observe that the mean signal highly depends on the value of the peak mass and at this stage, there is no obvious motivation for choosing a specific value. In Fig. 2, we present the expected signal for a wide spectrum. The shape of the mass spectrum does influence the shape of expected signal as the probabilistic nature of the lifetime is now taken into account: BHs with masses smaller or larger than $M_{t_H}$ do also contribute to the emitted radiation and changing their relative weights influences the result. Because of the probabilistic lifetime, we could observed a

signal in FRBs energy range. As a conclusion, one should expect a higher flux as the energy increases (up to the infrared band). The slope of the signal spectrum reflects that of the mass spectrum.

The key point of this study was to show that the randomness of the lifetime of BHs in quantum gravity can drastically change the spectral characteristic of the expected signal and can lead to predictions.

## 2. Emission cross section for loop black holes

The Hawking evaporation process [6] predicts that BHs emit a blackbody spectrum at temperature $T_H = 1/(8\pi M)$, with M the mass of the BH. However, the real spectrum is slightly more complicated because the emitted particles have to cross a potential barrier before escaping to infinity. This induces a modification, captured by the cross section $\sigma$, it encode information on the spacetime structure considered. The spectrum reads as:

$$\frac{dN}{dt} = \frac{1}{e^{\frac{\omega}{T_H}} \pm 1}\sigma(M, s, \omega)\frac{d^3k}{(2\pi)^3}, \tag{6}$$

with $s$ the particle spin and $\omega$ its energy. We have calculated this cross section for an effective loop BH [7]. The resulting metric (7) is obtained by considering a minisuperspace, based on polymerization procedure. It take into account that, in LQG, there is a minimal area.

$$ds^2 = -G(r)dt^2 + \frac{dr^2}{F(r)} + H(r)d\Omega^2, \tag{7}$$

$$G(r) = \frac{(r - r_+)(r - r_-)(r + r_*)^2}{r^4 + a_0^2}, \tag{8}$$

$$F(r) = \frac{(r - r_+)(r - r_-)r^4}{(r + r_*)^2(r^4 + a_0^2)}, \tag{9}$$

$$H(r) = r^2 + \frac{a_0^2}{r^2}, \tag{10}$$

with $a_0 = \frac{\sqrt{3}\gamma l_{Pl}^2}{2}$. There are two horizons $r_+ = 2M$ and $r_- = 2MP^2$, $r_* = \sqrt{r_+ r_-} = 2MP$ and $d\Omega^2 = d\theta^2 + sin^2\theta d\phi^2$. $\gamma$ is the Barbero-Immirzi parameter, $P = (\sqrt{1+\epsilon^2} - 1)$ is the polymeric function, $\epsilon = \gamma\delta$, with $\delta$ the polymeric parameter. With $a_0 = 0$, the metric (7) tends to the Schwarzschild one when $\delta$ tends to zero. According to the optical theorem [8], the cross section reads

$$\sigma(\omega) = \sum_{l=0}^{\infty} \frac{(2j+1)\pi}{\omega^2}|A_l|^2, \tag{11}$$

where $A_l$ is the transmission coefficient of the angular momentum mode $l$, and $j = l + s$ is the total angular momentum.

### 2.1. *Massless scalar field*

Considering the underlying symmetries, the scalar field can be written as $\Phi(r, \theta, \phi, t) = R(r)A(\theta)e^{i(\omega t + m\phi)}$, where $\omega$ is the frequency and $m$ is an integer.

The dynamics of a massless scalar field minimally coupled to a gravitational field is described by the generalized Klein-Gordon equation. With the metric (7) we obtain the radial equation (12):

$$\frac{\sqrt{GF}}{H}\partial_r\left(H\sqrt{GF}\partial_r R\right)+\left(\omega^2-\frac{G}{H}l(l+1)\right)R=0. \tag{12}$$

We define the tortoise coordinate $r^*$: $dr^*=dr/\sqrt{GF}$. At the horizon $r_+$, the radial part of the wave function $R^h$ will be a plane wave (13) with respect to $r^*$. At infinity, the radial wave function $R^\infty$ is a spherically wave (14):

$$R^h(r^*)=A^h_{in}e^{-i\omega r^*}+A^h_{out}e^{i\omega r^*}, \tag{13}$$

$$R^\infty(r)=\frac{A^\infty_{in}}{r}e^{-i\omega r}+\frac{A^\infty_{out}}{r}e^{i\omega r}, \tag{14}$$

with $A^h_{in}$ (respectively $A^h_{out}$) the probability amplitude for the incoming modes (resp. outgoing modes) at the horizon. Same definitions for $A^\infty$ at the infinity. For convenience, we choose the absorption point of view. With this convention, there are incoming and outgoing modes infinitively far from the BH and only incoming ones at the horizon. We therefore impose $A^h_{out}=0$. With the asymptotic solutions and by solving the differential equation (12), we obtained numerically the transmission amplitude from which the cross section can be obtained:

$$|A_l|^2=1-\left|\frac{A^\infty_{out}}{A^\infty_{in}}\right|^2. \tag{15}$$

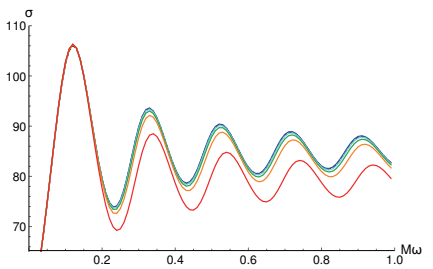



Fig. 3. Emission cross section for a scalar field with energy $\omega$ in the background spacetime of a loop BH of mass $M$ for different values of $\epsilon$ ($\epsilon=\gamma\delta$ measures the "quantumness" of spacetime). From bottom to top: $\epsilon=10^{\{-0.3,-0.6,-0.8,-1,-3\}}$. The blue line, corresponding to $\epsilon=10^{-3}$, is superposed with the cross section for a Schwarzschild BH.

In Fig. 3 we observe that the cross section does decrease when $\epsilon$ increases. When $\epsilon<10^{-0.8}$ it is hard to distinguish between the solutions. It therefore seems that taking into account the quantum correction does not influence the cross section of a scalar field for reasonable value of $\epsilon$, that is $\epsilon<<1$. However, the main trend is clear.

### 2.2. *Massless spin $\frac{1}{2}$ field*

For spin-$\frac{1}{2}$ fields, we have used the Newman-Penrose formalism [9]. We choose the following basis:

$$l^i = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{G}}, -\sqrt{F}, 0, 0\right), \quad n^i = \frac{1}{\sqrt{2}}\left(\frac{1}{\sqrt{G}}, \sqrt{F}, 0, 0\right),$$
$$m^i = \frac{1}{\sqrt{2}}\left(0, 0, \frac{1}{\sqrt{H}}, \frac{i}{\sqrt{H}sin\theta}\right), \quad \overline{m}^i = \frac{1}{\sqrt{2}}\left(0, 0, \frac{1}{\sqrt{H}}, \frac{-i}{\sqrt{H}sin\theta}\right)$$

Then we calculate the spin coefficients. The wavefunction is represented by a pair of spinors, $P^A$ and $\overline{Q}^{A'}$ with $A = 0, 1$ and $A' = 0, 1$. We use the ansatz

$$P^0 = \frac{e^{i(\omega t + m\phi)}}{\sqrt{H(r)}(G(r)F(r))^{\frac{1}{8}}} R_+(r)S_+(\theta), \quad P^1 = \frac{e^{i(\omega t + m\phi)}}{\sqrt{H(r)}(G(r)F(r))^{\frac{1}{8}}} R_-(r)S_-(\theta),$$
$$\overline{Q}^{0'} = -\frac{e^{i(\omega t + m\phi)}}{\sqrt{H(r)}(G(r)F(r))^{\frac{1}{8}}} R_+(r)S_-(\theta), \quad \overline{Q}^{1'} = \frac{e^{i(\omega t + m\phi)}}{\sqrt{H(r)}(G(r)F(r))^{\frac{1}{8}}} R_-(r)S_+(\theta).$$

Following the Chandrasekhar procedure [10], we obtain the radial equation

$$\sqrt{HF}\mathcal{D}\left(\sqrt{HF}\mathcal{D}^\dagger R_+\right) - \lambda^2 R_+ = 0, \quad \text{with} \quad \mathcal{D} = \partial_r + \left(\frac{G'}{8G} - \frac{F'}{8F}\right) + \frac{iw}{\sqrt{GF}}, \tag{16}$$

and $\lambda = l + 1$ a separation constant. At the horizon, the solution at the horizon reads as:

$$R_+(r^*) = A_1 e^{x_1 r^*}, \tag{17}$$

with $A_1$ a complex number and $x_1$ a complex number with a positive imaginary part. As before, we normalize such that $R_+(r_{ini}) = 1$, which leads to $\frac{dR_+(r_{ini})}{dr} = \frac{x_1}{\sqrt{G(r_{ini})F(r_{ini})}}$. At spacial infinity, the solution is a plane wave.

It has been shown in [11] that the transmission coefficient for spin 1/2 fields is given by:

$$|A_l|^2 = \left|\frac{A^h_{in}}{A^\infty_{in}}\right|^2. \tag{18}$$

By the same procedure described in the scalar field section, we obtain the cross section

In Fig. 4 we observe that the general trend is to decrease the cross section when the quantum correction increases. This should leave a footprint through a distortion of the instantaneous Hawking spectrum which will exhibit slight suppression of its UV tail because the relative effect is getting bigger with an increasing energy of the emitted particle. For scalar and fermionic fields, the trend is clear but the effects are

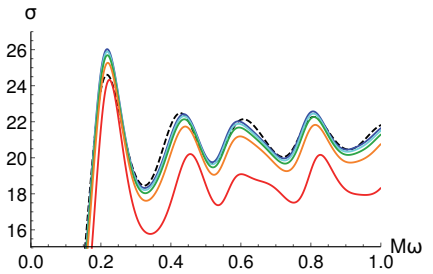



Fig. 4. Emission cross section for a fermionic field, with energy $\omega$, in the background spacetime of a LBH of mass $M$. From bottom to top: $\epsilon = 10^{\{-0.3,-0.6,-0.8,-1,-3\}}$. The dashed dark curve corresponds to the Schwarzschild cross section.

generically small. From the phenomenological point of view, if the polymerisation parameter is large we could probe a decreased cross section.

## References


1. A. Barrau, F. Moulin and K. Martineau, *Phys. Rev. D* **97**, 066019 (2018).
2. H. Haggard and C. Rovelli, *Phys. Rev. D* **92**, 104020 (American Physical Society, 2015).
3. C. Barcel, R. Carballo-Rubio and L. J. Garay, *J. High Energy Phys.* (2016).
4. J. L. G. Sobrinho, P. Augusto and A. L. Gonalves, *Monthly Notices of the Royal Astronomical Society*, **463**, 3, (2016).
5. B. G. Carr, *Astrophys. J.*, **201**, 3, (2016).
6. S. Hawking, *Commun. math. Phys*, **43**, 199, (1975).
7. L. Modesto, *Int. J. Theor. Phys.*, **49**, 1649, (2010).
8. J. J. Sakurai, *Advanced quantum mechanics* (1967).
9. E. Newman and R. Penrose, *J. Math. Phys.*, **3**, 566 (1962).
10. S. Chandrasekhar, *The Mathematical Theory of Black Holes*, (1998).
11. M. Cvetic and F.Larsen, *Phys. Rev.*, **D57**, 6297, (1998).

# Computation of non-Gaussianity in loop quantum cosmology

Vijayakumar Sreenath

*Inter-University Centre for Astronomy and Astrophysics, Ganeshkhind, Pune, Maharashtra 411007, India*
*E-mail: vsreenath@iucaa.in*

Ivan Agullo

*Department of Physics and Astronomy, Louisiana State University, Baton Rouge, LA 70803, U.S.A.*
*E-mail: agullo@lsu.edu*

Boris Bolliet

*Jodrell Bank Centre for Astrophysics, The University of Manchester, Alan Turing Building, Oxford Road, Manchester, M13 9PL.*
*E-mail: boris.bolliet@manchester.ac.uk*

We summarize our investigations of the second-order perturbations in loop quantum cosmology (LQC). We shall discuss, primarily, two aspects. Firstly, whether the second-order contributions arising from the cosmic bounce, occurring at Planck scale, could be large enough to break the validity of perturbation theory. Secondly, the implications of the upper bounds on primordial non-Gaussianity, arrived at by the Planck collaboration, on the LQC phenomenology.

*Keywords*: Loop quantum cosmology; Primordial non-Gaussianity.

## 1. Introduction

Loop quantum cosmology (LQC) provides an extension of the inflationary paradigm to the Planck era (see, for instance, Ref. 1). Over the past decade or so, there has been a research program aimed at investigating the viability of LQC as a theory of the pre-inflationary universe. Until now, investigations of primordial perturbations generated in LQC have focused mainly at the level of the power spectrum. In this work, we extend the analysis to the level of three-point functions, namely, the bispectrum of curvature perturbations. We will analyze primarily two aspects. Firstly, we check whether next-to-leading order corrections to the power spectrum are sub-leading. Secondly, we verify that the amount of non-Gaussianity as quantified by the dimensionless quantity $f_{\rm NL}$ is compatible with the observations of cosmic microwave background (CMB) and investigate new predictions.

## 2. Computation of the Bispectrum in the Dressed Metric Approach

The system of interest is scalar perturbations $\delta\phi$ living on a Friedmann-Lemaître-Robertson-Walker (FLRW) metric sourced by a scalar field $\phi$. In LQC, such a system is described by a wavefunction $\Psi(v, \phi, \delta\phi)$, where $v \equiv a^3\,\mathcal{V}_0\,4/\kappa$ with $a$ being the scale factor and $\mathcal{V}_0$, the volume of the universe, introduced to regulate

infrared divergence. The dynamics is governed by the constraint equation, $\hat{\mathcal{H}}\,\Psi = 0$, where the Hamiltonian operator can be split in to the background and perturbed part as $\hat{\mathcal{H}} = \hat{\mathcal{H}}_{FLRW} + \hat{\mathcal{H}}_{pert}$. We are interested in solutions wherein $\Psi(v,\,\phi,\,\delta\phi) = \Psi_0(v,\,\phi,)\otimes\delta\Psi(v,\,\phi,\,\delta\phi)$, where $\Psi_0$ describes a quantum FLRW geometry and $\delta\Psi$ describes the scalar perturbations.

The states $\Psi_0$ satisfies the equation $\hat{\mathcal{H}}_{FLRW}\,\Psi_0 = 0$. It has been shown that, for states that are sharply peaked in the volume $v$ during the entire evolution, the background geometry can be described by an effective classical Hamiltonian (see e.g. Ref. 1 and references therein). In the dressed metric approach, we are interested in quantum states $\delta\Psi(v,\,\phi,\,\delta\phi)$ that are a small perturbation around such a quantum FLRW state $\Psi_0(v,\,\phi)$. A detailed analysis shows that $\delta\Psi(v,\,\phi,\,\delta\phi)$ are solutions to the Schrödinger equation, $i\hbar\,\partial_\phi\delta\Psi = \langle\Psi_0|\hat{\mathcal{H}}_{\rm pert}[N_\phi]|\Psi_0\rangle\,\delta\Psi$, where $\hat{\mathcal{H}}_{pert} = \hat{\mathcal{H}}^{(2)} + \hat{\mathcal{H}}^{(3)}$, namely the Hamiltonian at second and third order in perturbations respectively and $N_\phi$ is the lapse associated with relational time $\phi$.[2]

We are interested in computing the correlation functions of these scalar perturbations. The first step is to expand the perturbations in Fourier space and introduce creation and annihilation operators

$$\hat{\delta\phi}(\vec{x},\eta) = \int \frac{{\rm d}^3k}{(2\pi)^3}\left(\hat{A}_{\vec{k}}\;\varphi_k(\eta) + \hat{A}^\dagger_{-\vec{k}}\;\varphi^*_k(\eta)\right)e^{i\vec{k}\cdot\vec{x}}, \tag{1}$$

where $[\hat{A}_{\vec{k}},\hat{A}^\dagger_{\vec{k}'}] = \hbar\,(2\pi)^3\,\delta^{(3)}(\vec{k}+\vec{k}')$ and $[\hat{A}_{\vec{k}},\hat{A}_{\vec{k}'}] = 0$. The dynamics of perturbations are governed by the second-order Hamiltonian with the background quantities determined using the effective background Hamiltonian. The scalar power spectrum of $\hat{\delta\phi}$ is defined as

$$\langle 0|\hat{\delta\phi}_{\vec{k}}(\eta)\hat{\delta\phi}_{\vec{k}'}(\eta)|0\rangle \equiv (2\pi)^3\delta^{(3)}(\vec{k}+\vec{k}')\frac{2\pi^2}{k^3}\mathcal{P}_{\delta\phi}(k,\eta)\,, \tag{2}$$

where $|0\rangle$ is the vacuum annihilated by the operators $\hat{A}_{\vec{k}}$ for all $\vec{k}$. For the purpose of relating perturbations to the late time physics, it is convenient to express the power spectrum in terms of comoving curvature perturbations. The power spectrum of curvature perturbation, $\mathcal{R}$, in terms of inflaton perturbation $\delta\phi$, evaluated at the end of inflation is $\mathcal{P}_{\mathcal{R}}(k) \equiv \left(\frac{a(\eta_{\rm end})}{z(\eta_{\rm end})}\right)^2 \frac{\hbar\,k^3}{2\pi^2}\,|\varphi_k(\eta_{\rm end})|^2$, where $z = -\frac{6}{\kappa}\frac{p_\phi}{\pi_a}$ with $\kappa = 8\,\pi\,G$ and $p_\phi$ and $\pi_a$ are momenta conjugate to $\phi$ and $a$ respectively.

The self-interaction of scalar perturbation, at lowest order, is described by the third-order interaction Hamiltonian, $\hat{\mathcal{H}}_{\rm int} \equiv \langle\Psi_0|\hat{\mathcal{H}}^{(3)}[N_\phi]|\Psi_0\rangle$.[3] The perturbations at this order are quantified using the scalar bispectrum, $B_{\mathcal{R}}(k_1,k_2,k_3)$, that is defined in terms of curvature perturbations by

$$\langle 0|\hat{\mathcal{R}}_{\vec{k}_1}\hat{\mathcal{R}}_{\vec{k}_2}\hat{\mathcal{R}}_{\vec{k}_3}|0\rangle \equiv (2\pi)^3\delta^{(3)}(\vec{k}_1+\vec{k}_2+\vec{k}_3)\,B_{\mathcal{R}}(k_1,k_2,k_3)\,. \tag{3}$$

It is often convenient to quantify the bispectrum using a dimensionless function, $f_{\rm NL}$, which can be defined as,

$$B_{\mathcal{R}}(k_1,k_2,k_3) \equiv -\frac{6}{5}\,f_{\rm NL}(k_1,k_2,k_3)\;\times(\Delta_{k_1}\Delta_{k_2} + \Delta_{k_1}\Delta_{k_3} + \Delta_{k_2}\Delta_{k_3})\,, \tag{4}$$

where $\Delta_k \equiv \frac{2\pi^2}{k^3}\,\mathcal{P}_{\mathcal{R}}(k)$ is the dimensionful power spectrum.

In order to compute bispectrum, we need to express it in terms of $\delta\phi$ as follows,

$$\langle 0|\hat{\mathcal{R}}_{\vec{k}_1}\hat{\mathcal{R}}_{\vec{k}_2}\hat{\mathcal{R}}_{\vec{k}_3}|0\rangle = \left(-\frac{a}{z}\right)^3 \langle 0|\hat{\delta\phi}_{\vec{k}_1}\hat{\delta\phi}_{\vec{k}_2}\hat{\delta\phi}_{\vec{k}_3}|0\rangle + \left(-\frac{3}{2} + 3\frac{V_\phi\, a^5}{\kappa\, p_\phi\, \pi_a} + \frac{\kappa}{4}\frac{z^2}{a^2}\right)\left(-\frac{a}{z}\right)^4$$
$$\times\Big[\int \frac{d^3p}{(2\pi)^3}\,\langle 0|\hat{\delta\phi}_{\vec{k}_1}\hat{\delta\phi}_{\vec{k}_2}\hat{\delta\phi}_{\vec{p}}\hat{\delta\phi}_{\vec{k}_3-\vec{p}}|0\rangle + (\vec{k}_1 \leftrightarrow \vec{k}_3) + (\vec{k}_2 \leftrightarrow \vec{k}_3) + \cdots\Big], \tag{5}$$

where the symbols $(\vec{k}_i \leftrightarrow \vec{k}_j)$ indicate terms obtained by replacing $k_i$ with $k_j$ in the first term of the second line and the dots indicate higher order terms. At leading order in perturbations, the first term on RHS can be evaluated using time dependent perturbation theory

$$\langle 0|\hat{\delta\phi}_{\vec{k}_1}(\eta)\hat{\delta\phi}_{\vec{k}_2}(\eta)\hat{\delta\phi}_{\vec{k}_3}(\eta)|0\rangle = \langle 0|\hat{\delta\phi}^{\rm I}_{\vec{k}_1}(\eta)\hat{\delta\phi}^{\rm I}_{\vec{k}_2}(\eta)\hat{\delta\phi}^{\rm I}_{\vec{k}_3}(\eta)|0\rangle$$
$$- i/\hbar \int d\eta' \langle 0|\left[\hat{\delta\phi}^{\rm I}_{\vec{k}_1}(\eta)\hat{\delta\phi}^{\rm I}_{\vec{k}_2}(\eta)\hat{\delta\phi}^{\rm I}_{\vec{k}_3}(\eta), \hat{\mathcal{H}}^{\rm I}_{\rm int}(\eta')\right]|0\rangle\,, \tag{6}$$

where the superscript $I$ indicates fields in the interaction picture. Since, $\delta\phi^I$ is a Gaussian field, the first term vanishes and only the second term contributes. The second term in the RHS of Eq. (5) can be evaluated using Wick's theorem and Eq. (2). Using Eqs. (5) and (6), one can compute the bispectrum and hence the function $f_{\rm NL}$ using Eq. (4).

## 3. Numerical Method and Results

In this section, we will briefly describe our implementation of the formalism for computing $f_{\rm NL}$ and the results we obtain. In order to compute $f_{\rm NL}$ at the end of inflation, one needs to evolve the perturbations from an early time before the bounce until the perturbations leave the horizon during inflation at which point their amplitude freezes in time. We need to make three choices to do this computation. Firstly, we need to specify the potential governing the field $\phi$. We choose the quadratic potential, $V(\phi) = m^2\,\phi^2/2$, where $m = 6.4\times10^{-6} M_{P\ell}$. Secondly, we need to choose a background geometry by specifying the value of $\phi$ and energy density, $\rho$, at the bounce. We work with $\phi_{\rm B} = 7.62\, M_{P\ell}$ and $\rho_B = 1\, M^4_{P\ell}$, where subscript $B$ denotes the bounce, so that the effects due to LQC appear at observable scales while respecting the Planck constraints on power spectrum. Finally, we need to choose an initial state for perturbations, which we choose to be a Minkowski initial state. More specifically, we choose $\varphi_k(\eta_0) = \frac{1}{a(\eta_0)\sqrt{2\,k}}$ and $\varphi'_k(\eta_0) = [-i\,k + \frac{a'(\eta_0)}{a(\eta_0)}]\,\varphi_k(\eta_0)$ as initial data for the modes, at conformal time $\eta_0 = -2.8\times10^3\, T_{P\ell}$ (the bounce takes place at $\eta = 0$). The initial time was chosen so that all the modes of interest, namely those between $k_{\rm min} = k_*/10$, and $k_{\rm max} = 1000 k_*$, where $k_*/a(t_{today}) = 0.002\,{\rm Mpc}^{-1}$ is the pivot scale, were in the adiabatic regime. We have investigated the effects of varying these choices in detail in Ref. 2.

To perform this computation, we use the platform provided by `class`.[4] The computation was done in two stages. In the first stage we evolve the background from very early times to the end of inflation. In the second step, we convert the time integral in Eq. (6) to a differential equation and evolve it together with the differential equation for the fourier modes. In the remaining part of this section, we will discuss the various results.

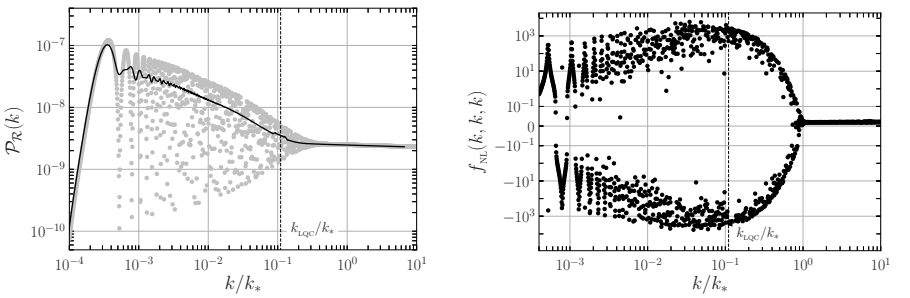



Fig. 1. The power spectrum and $f_{\rm NL}(k_1, k_2, k_3)$ evaluated in the equilateral limit.

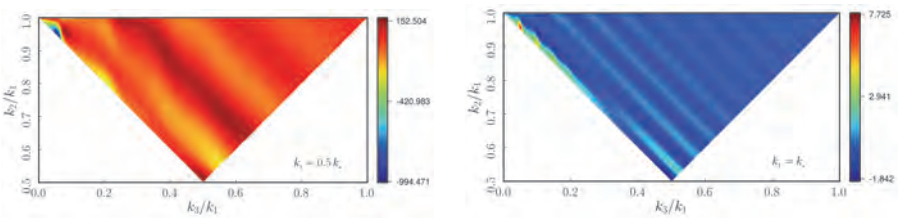



Fig. 2. The shape of the non-Gaussianity, $f_{\rm NL}(k_1, k_2, k_3)$, evaluated with $k_1 = k_*/2$ and $k_*$. The figure shows the portions allowed by the condition $\vec{k}_1 + \vec{k}_2 + \vec{k}_3 = 0$.

Fig. 1 depicts the scalar power spectrum and $f_{\rm NL}$ in the equilateral limit. One can see that for $k \leq k_{\rm LQC}$, where $k_{\rm LQC}$ is the scale set by the spacetime curvature at the bounce, the spectra are strongly scale dependent while for $k >> k_{\rm LQC}$, the spectra approach their slow roll values. At low wave numbers, the figure shows that $f_{\rm NL}$ is oscillatory. We have depicted the $f_{\rm NL}$ for all configurations in Fig. 2. In this figure, we have fixed the value of $k_1$ and varied $k_2$ and $k_3$ in such a way that they obey the triangle condition. This figure illustrates the shape of the non-Gaussianity. It can be seen that, in both the figures, the $f_{\rm NL}$ peaks in the squeezed($k_3 \ll k_1 \simeq k_2$) - flattened ($k_1 \simeq k_2 + k_3$) limit.

The primordial non-Gaussianity generated due to LQC has a characteristic enhancement of amplitude at scales comparable to $k_{\rm LQC}$. By analyzing the integrals involved in the computation of $f_{\rm NL}$, we can estimate the contribution to $f_{\rm NL}$ from the epoch around the bounce. For modes, $k \geq k_{\rm LQC}$, we can approximate the mode function as $\varphi_k \sim e^{-ik\eta}$. Then the contribution to the integral from time around

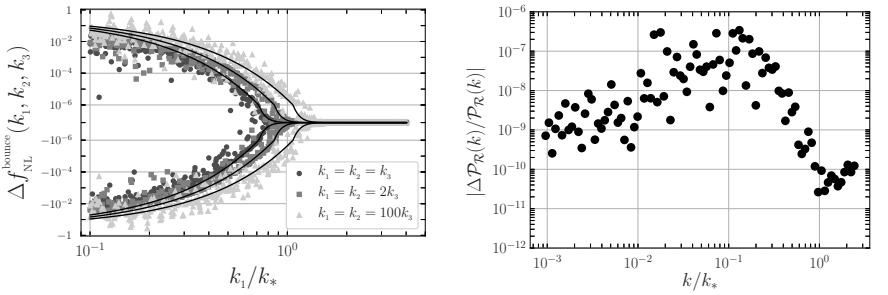



Fig. 3. On left, we have plotted a comparison of the analytical expression, $e^{-\alpha k_t/k_{\rm LQC}}$, for contribution to $f_{\rm NL}$ from the bounce to the numerical result in different configurations. On right, we compute the relative amplitude of the leading order correction to the power spectrum.

the bounce can be schematically written as,

$$I(k_1,k_2,k_3) \sim \int_{-\Delta\eta}^{\Delta\eta} {\rm d}\eta\, g(\eta)\, e^{i(k_1+k_2+k_3)\,\eta} \approx \int_{-\infty}^{\infty} {\rm d}\eta\, g(\eta)\, e^{ik_t\,\eta}\, W(\eta,\Delta), \qquad (7)$$

where, $g(\eta)$ is a combination of the functions depending on the background and the wavenumbers, $k_t \equiv k_1 + k_2 + k_3$ and $W(\eta, \Delta\eta)$ is a window function which selects only the contribution from the time range $-\Delta\eta < \eta < \Delta\eta$. This integral can be computed using Cauchy's residue theorem and the spectral dependence of the integral can be written as $e^{-\alpha k_t/k_{\rm LQC}}$. In Fig. 3, we have compared the analytical expression with numerical result and we find a good match between the two.

Finally let us compute the contribution to the power spectrum from the bispectrum. For the perturbation theory to be valid, this contribution has to be subdominant. The first perturbative correction to the two-point function of curvature perturbation is given by

$$\langle 0|\hat{\mathcal{R}}_{\vec{k}_1}\hat{\mathcal{R}}_{\vec{k}_2}|0\rangle = (2\pi)^3\delta^{(3)}(\vec{k}_1+\vec{k}_2)\,\frac{2\pi^2}{k_1^3}\,\hbar\,\left[\mathcal{P}_\mathcal{R}(k_1) + \Delta\mathcal{P}_\mathcal{R}(k_1)\right], \qquad (8)$$

where

$$\Delta\mathcal{P}_\mathcal{R}(k_1) = \hbar\,\frac{k_1^3}{\pi^2}\left[\left(-\frac{a}{z}\right)^3\left[-\frac{3}{2}+3\frac{V_\phi\,a^5}{\kappa\,p_\phi\,\pi_a}+\frac{\kappa}{4}\frac{z^2}{a^2}\right]\int\frac{{\rm d}^3p}{(2\pi)^3}\,B_{\delta\phi}(\vec{k}_1,\vec{p},-\vec{k}_1-\vec{p})\right.$$
$$\left.+\left(-\frac{a}{z}\right)^4\left[-\frac{3}{2}+3\frac{V_\phi\,a^5}{\kappa\,p_\phi\,\pi_a}+\frac{\kappa}{4}\frac{z^2}{a^2}\right]^2\int\frac{{\rm d}^3p}{(2\pi)^3}\,|\varphi_p|^2\,|\varphi_{|\vec{k}_1-\vec{p}|}|^2\right], \qquad (9)$$

where $B_{\delta\phi}(\vec{k}_1,\vec{p},-\vec{k}_1-\vec{p})$ is the bispectrum of inflaton perturbations and all the quantities on the right are evaluated at the end of inflation. We have numerically ploted the relative amplitude of the first order correction, $|\Delta\mathcal{P}_\mathcal{R}/\mathcal{P}_\mathcal{R}|$, in Fig. 3. We find that, as expected, the magnitude of first-order correction to the power spectrum is negligible. This result can be qualitatively understood as follows. The leading order contribution to $\Delta\mathcal{P}_\mathcal{R}(k_1)$ is given by the first term in Eq. (9) and it is

given by $\epsilon\, f_{\rm NL}\, \mathcal{P}_{\mathcal{R}}^2$, where $\epsilon$ is the slow roll parameter of $\mathcal{O}(10^{-2})$. Since, $f_{\rm NL} \leq 10^4$ and $\mathcal{P}_{\mathcal{R}} \leq 10^{-7}$, we obtain $\Delta\mathcal{P}_{\mathcal{R}}/\mathcal{P}_{\mathcal{R}} \leq 10^{-4}$ as in Fig. 3.

## 4. Discussion

Let us conclude by making some remarks on the robustness of the results and its implication in the light of Planck data. We have verified the robustness of the results to a variation of the basic assumptions discussed in Sec. 3.[2] For instance, we find that the effect of changing $\phi_b$ is only a shift in the scale which is sensitive to the effect of the bounce with respect to the scales observable today. An increase in $\rho_B$ also leads only to a similar shift in the scales sensitive to the curvature of the bounce, in addition, to an increase in amplitude of $f_{\rm NL}$. The Planck mission has put strong constraints on certain models of scale invariant non-Gaussianity, but, it provides little information on the scale dependent non-Gaussianity as produced in LQC.[5] Moreover, since the error bar on $f_{\rm NL}$ goes as $1/\sqrt{\ell}$, at low multipoles, where the non-Gaussianity due to LQC is expected to be large, the error bar would be large. Considering the Planck error bars at low multipoles and demanding that the enhancement in $f_{\rm NL}$ due to the LQC bounce appears at $\ell \lesssim 50$, one could try to arrive at constraints on the minimum value of scalar field at the bounce, for a given value of $\rho_{\rm B}$. Furthermore, by demanding that the imprint of the bounce should be at observable scales, we can arrive at an upper bound on $\phi_B$. For instance, for $\rho_{\rm B} = 1\, M_{\rm Pl}^4$, we obtain $7.46 M_{\rm Pl} \leq \phi_{\rm B} \leq 7.82 M_{\rm Pl}$. It should be kept in mind that the constraint described above is a very conservative estimate. Most probably, the oscillations in $f_{\rm NL}$ will relax the constraint on $\phi_{\rm B}$ discussed above. A more detailed account of this work has been published in Ref. 2.

## Acknowledgments

We thank the organizers for giving VS the opportunity to present this work. VS would also like to thank Inter-University Centre for Astronomy and Astrophysics, Pune for financial support to attend MG XV.

## References

1. I. Agullo and P. Singh, Loop Quantum Cosmology, in *Loop Quantum Gravity: The First 30 Years*, eds. A. Ashtekar and J. Pullin (WSP, 2017) pp. 183–240.
2. I. Agullo, B. Bolliet and V. Sreenath, Non-Gaussianity in Loop Quantum Cosmology, *Phys. Rev.* **D97**, p. 066021 (2018).
3. J. M. Maldacena, Non-Gaussian features of primordial fluctuations in single field inflationary models, *JHEP* **05**, p. 013 (2003).
4. D. Blas, J. Lesgourgues and T. Tram, The Cosmic Linear Anisotropy Solving System (CLASS). Part II: Approximation schemes, *JCAP* **7**, p. 034 (July 2011).
5. P. A. R. Ade *et al.*, Planck 2015 results. XVII. Constraints on primordial non-Gaussianity, *Astron. Astrophys.* **594**, p. A17 (2016).

# Effect of Schwinger pair production on the evolution of the Hubble constant in de Sitter spacetime

Ehsan Bavarsad*

*Department of Physics, University of Kashan, Kashan, 8731753153, Iran*
**E-mail: bavarsad@kashanu.ac.ir*

Sang Pyo Kim

*Department of Physics, Kunsan National University, Kunsan, 54150, Korea*
*Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing, 100190, China*
*E-mail: sangkim@kunsan.ac.kr*

Clément Stahl

*Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla, Valparaíso, 4950, Chile*
*E-mail: clement.stahl@pucv.cl*

She-Sheng Xue

*ICRANet, Piazzale della Repubblica, Pescara, 10, 65122 ,Italy*
*Dipartimento di Fisica, Universita di Roma "La Sapienza", Piazzale Aldo Moro, Rome, 5, 00185, Italy*
*E-mail: xue@icra.it*

In this proceeding we consider a massive charged scalar field in a uniform electric field background in a de Sitter spacetime (dS). We compute the in-vacuum expectation value of the trace of the energy-momentum tensor for the created Schwinger pairs, and using adiabatic subtraction scheme the trace is regularized. The effect of the Schwinger pair creation on the evolution of the Hubble constant is investigated. We find that the production of the semiclassical pairs leads to a decay of the Hubble constant. Whereas, the production of a light scalar field in the weak electric field regime leads to a superacceleration phenomenon.

*Keywords*: de Sitter Spacetime; Schwinger Effect; Trace of the Energy-Momentum Tensor; Evolution of the Hubble Constant.

## 1. Introduction

In a flat spacetime the phenomenon of the pair creation in a strong electric field background is referred to as the Schwinger effect.[1] The presence of the strong electromagnetic and gravitational fields in the early Universe motivates the study of the Schwinger effect in dS.[2] The Schwinger effect and the induced current of the created pairs were investigated in a uniform electric field for various dimensions of the de Sitter spacetime in Refs. 3, 4, 5, 6, 7. Recently, the effect of a conserved flux magnetic field on the Schwinger effect in dS has been investigated in Ref. 8, see also Ref. 9. Without considering an electromagnetic field, the energy-momentum tensor of the created scalar particles in the gravitational background field of dS has been investigated in Refs. 10, 11, 12, 13, 14, 15, 16, 17, 18. The results of Refs. 10, 15, 16 show that the Hubble constant of dS decays due to the

particle creation. Furthermore, the gravitational backreaction effects of the quantum fluctuations may lead to a superacceleration phase, where the Hubble constant increase.[15] Using a semiclassical approach the energy-momentum tensor of the created Schwinger scalars in a dS has been computed in Refs. 6, 19, which showed that the Hubble constant decays. The trace of the induced energy-momentum tensor of the Schwinger scalars in a three dimensional dS has been investigated in Ref. 20, and the authors found that the creation of the semiclassical pairs may lead to an increase of the Hubble constant. With the aim of developing the renormalization theory in the curved spacetime, recently the Schwinger effect and the conformal anomaly for both of the scalar and spinor de Sitter QED cases have been investigated in Ref. 21. With the aim of completing the work started in Ref. 6 and with cosmological applications in mind, in this proceeding we compute the trace of the energy-momentum tensor for the Schwinger scalars created in a uniform electric field background in a four-dimensional dS. The proceeding is organised as follows: in Sec. 2 the preliminary explanation of our model is introduced. We compute the regularized trace of the induced energy-momentum tensor in Sec. 3. In Sec. 4 we give some conclusions.

## 2. Preliminaries

We begin our study by considering the scalar QED action in a four-dimensional dS as

$$S = \int d^4x\sqrt{|g|}\Big\{g^{\mu\nu}\big(\partial_\mu + ieA_\mu\big)\varphi\big(\partial_\nu - ieA_\nu\big)\varphi^* - \big(m^2 + \xi R\big)\varphi\varphi^* - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}\Big\}, \quad (1)$$

where $\varphi(x)$ is a complex scalar field with mass $m$ and electric charge $e$ which is coupled to an electromagnetic vector potential background $A_\mu(x)$. The parameter $\xi$ is a dimensionless nonminimal coupling of the scalar field to the Ricci scalar $R$, and in this proceeding from now on we set $\xi = 1/6$. The $|g|$ denotes the absolute value of the metric determinant. In Secs. 2 and 3, we assume that the gravitational and electromagnetic fields are not affected by the pair creation. This assumption on the gravitational field does not hold when we discuses backreaction effects in Sec. 4. We consider the Poincaré patch of dS, whose metric reads

$$ds^2 = \Omega^2(\tau)\big(d\tau^2 - d\mathbf{x}^2\big), \qquad \Omega(\tau) = -\frac{1}{\tau H}, \qquad \tau \in (-\infty, 0), \qquad \mathbf{x} \in \mathbb{R}^3, \quad (2)$$

where $H$ is the Hubble constant and $\tau$ is the conformal time which relates to the cosmological proper time $t$ as

$$\tau = -\frac{1}{H}e^{-Ht}. \quad (3)$$

Having a uniform electric field with the constant energy density in the metric background given by Eq. (2), we choose the electromagnetic vector potential in the gauge

$$A_\mu(\tau) = -\frac{E}{H^2\tau}\delta^1_\mu, \quad (4)$$

where $E$ is a constant. The the Klein-Gordon equation for the scalar field reads from the action (1). We impose that the mode functions of the Klein-Gordon equation have the asymptotic behavior similar to those mode functions in the Minkowski spacetime at the early times, $\tau \to -\infty$. Then, these mode functions that describe the Hadamard[4] in vacuum state are given by[5,6]

$$U_{\text{in}\mathbf{k}}(x) = (2k)^{-\frac{1}{2}} e^{\frac{i\pi\kappa}{2}} \Omega^{-1}(\tau) e^{+i\mathbf{k}\cdot\mathbf{x}} W_{\kappa,\gamma}\big(2ik\tau\big), \tag{5}$$

$$V_{\text{in}\mathbf{k}}(x) = (2k)^{-\frac{1}{2}} e^{-\frac{i\pi\kappa}{2}} \Omega^{-1}(\tau) e^{-i\mathbf{k}\cdot\mathbf{x}} W_{\kappa,-\gamma}\big(-2ik\tau\big), \tag{6}$$

where $U_{\text{in}\mathbf{k}}$ and $V_{\text{in}\mathbf{k}}$ are positive and negative frequency mode functions, respectively. The function $W$ is the Whittaker function, and the variables are defined as

$$k := \big|\mathbf{k}\big|, \quad \mu := \frac{m}{H}, \quad \lambda := -\frac{eE}{H^2}, \quad r := \frac{k_x}{k}, \quad \kappa := -i\lambda r, \quad \gamma^2 := \frac{1}{4} - \mu^2 - \lambda^2. \tag{7}$$

Then we can expand the scalar field operator $\phi(x)$ as

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3} \Big[ U_{\text{in}\mathbf{k}}(x) a_{\text{in}\mathbf{k}} + V_{\text{in}\mathbf{k}}(x) b^\dagger_{\text{in}\mathbf{k}} \Big], \tag{8}$$

where $a_{\text{in}\mathbf{k}}$ and $b^\dagger_{\text{in}\mathbf{k}}$ are the annihilation and creation operators for particle and antiparticles with comoving momentum $\mathbf{k}$, respectively, which satisfy the commutation relations. Then, the in vacuum state is defined by

$$a_{\text{in}\mathbf{k}}|\text{in}\rangle = b_{\text{in}\mathbf{k}}|\text{in}\rangle = 0, \qquad \forall\, \mathbf{k}. \tag{9}$$

## 3. Trace of the Induced Energy-Momentum Tensor

The induced energy-momentum tensor of the created Schwinger pairs is necessary to study evolution of the de Sitter spacetime. We will compute all of the induced energy-momentum tensor's components in a future work. However, in this proceeding we assume that the created Schwinger pairs take the form of a perfect fluid with the *vacuum* equation of state. Consequently, the energy-momentum tensor $T_{\mu\nu}$ is related to its trace $T$ as

$$T_{\mu\nu} = \frac{1}{4} T g_{\mu\nu}. \tag{10}$$

Hence, it is sufficient to compute the trace of the induced energy-momentum tensor which is defined by the variation of the action (1) with respect to the variation of the metric $\delta g_{\mu\nu}$ as

$$T := -\frac{2}{\sqrt{|g|}} g_{\mu\nu} \frac{\delta S}{\delta g_{\mu\nu}}. \tag{11}$$

In the action (1) the Maxwell term, $F_{\mu\nu}F^{\mu\nu}$, is the pure electromagnetic part of the action which is constant, and according to our assumption is not affected by the Schwinger pair creation, hence to obtain the trace of the induced energy-momentum tensor of the Schwinger pairs we do not consider this term. We then obtain

$$T = 2m^2 \varphi\varphi^*. \tag{12}$$

Substituting the scalar field operator (8) into the expression (12) and using Eq. (9) leads to

$$\langle \mathrm{in}|T|\mathrm{in}\rangle = 2m^2 \int \frac{d^3k}{(2\pi)^3}\Big|U_{\mathrm{in}\mathbf{k}}\big(x\big)\Big|^2. \tag{13}$$

Equation (13), in terms of a dimensionless integral variable $p = -k\tau$ can be written as

$$\langle \mathrm{in}|T|\mathrm{in}\rangle = \frac{H^4\mu^2}{4\pi^2}\int_{-1}^{1} dr e^{\lambda\pi r}\int_0^{\Lambda} dp\, p\Big|W_{-i\lambda r,\gamma}\big(-2ip\big)\Big|^2, \tag{14}$$

where the dimensionless momentum cutoff $\Lambda$ is defined to regularize the ultraviolet divergency. Following the integration procedure introduced in Refs. 4, 5 we obtain the unregularized expression for the trace of the induced energy-momentum tensor

$$\begin{aligned}\langle \mathrm{in}|T|\mathrm{in}\rangle = \frac{H^4\mu^2}{16\pi^2}\Big\{&4\Lambda^2 - 4\mu^2\log(2\Lambda) + 2\mu^2 - \frac{8\lambda^2}{3} + 2 + 2i\pi\mu^2 - \frac{6\gamma}{\pi}\csc(2\pi\gamma)\\ &\times\Big(\cosh(2\pi\lambda) - \frac{1}{2\pi\lambda}\sinh(2\pi\lambda)\Big) + i\csc(2\pi\gamma)\int_{-1}^{1} dr\big(3r^2\lambda^2 - \lambda^2 - \mu^2\big)\\ &\times\Big[\big(e^{2\pi\lambda r} + e^{2\pi i\gamma}\big)\psi\Big(\frac{1}{2} + i\lambda r + \gamma\Big) - \big(e^{2\pi\lambda r} + e^{-2\pi i\gamma}\big)\psi\Big(\frac{1}{2} + i\lambda r - \gamma\Big)\Big]\Big\},\end{aligned} \tag{15}$$

where $\psi$ is the digamma function. In order to remove the ultraviolet divergences from the expression (15) we apply the adiabatic subtraction scheme as introduced in Ref. 22. However, a new condition for renormalization the vacuum expectation value of the quantities in the context of de Sitter QED has been introduced in Ref. 23. The positive frequency mode function of adiabatic zeroth order is given by

$$U_{\mathrm{A}}(x) = \Omega^{-1}(\tau)\Big(2\omega(\tau)\Big)^{-\frac{1}{2}}\exp\Big(i\mathbf{k}.\mathbf{x} - i\int\omega(\tau)d\tau\Big), \tag{16}$$

where the time dependent frequency is

$$\omega(\tau) = H\Omega(\tau)\sqrt{k^2\tau^2 + 2\lambda r k\tau + \lambda^2 + \mu^2}. \tag{17}$$

Using the mode function (16) the zeroth order of the adiabatic expansion of the trace of the energy-momentum tensor is obtained

$$T_{\mathrm{A}} = \frac{H^4\mu^2}{16\pi^2}\Big(4\Lambda^2 - 4\mu^2\log(2\Lambda) + 2\mu^2 - \frac{8\lambda^2}{3}\Big). \tag{18}$$

The adiabatic regularized trace of the induced energy-momentum tensor is given by subtracting the counterterm (18) from the unregularized expression (15). We then obtain the regularized trace as

$$\begin{aligned}T = \langle \mathrm{in}|T|\mathrm{in}\rangle - T_{\mathrm{A}} = \frac{H^4\mu^2}{16\pi^2}\Big\{&2 + 2i\pi\mu^2 - \frac{6\gamma}{\pi}\csc(2\pi\gamma)\Big(\cosh(2\pi\lambda)\\ &- \frac{1}{2\pi\lambda}\sinh(2\pi\lambda)\Big) + i\csc(2\pi\gamma)\int_{-1}^{1} dr\big(3r^2\lambda^2 - \lambda^2 - \mu^2\big)\Big[\big(e^{2\pi\lambda r} + e^{2\pi i\gamma}\big)\\ &\times\psi\Big(\frac{1}{2} + i\lambda r + \gamma\Big) - \big(e^{2\pi\lambda r} + e^{-2\pi i\gamma}\big)\psi\Big(\frac{1}{2} + i\lambda r - \gamma\Big)\Big]\Big\}.\end{aligned} \tag{19}$$

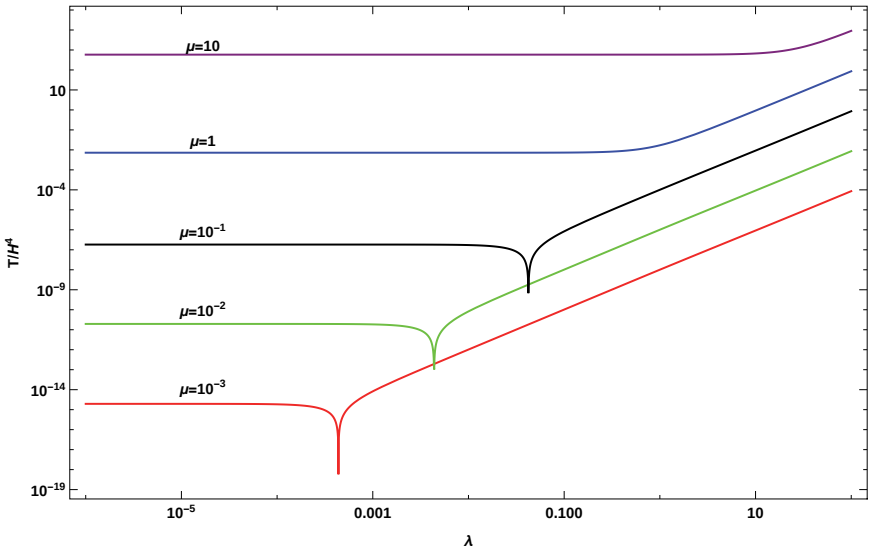



Fig. 1. The normalized trace $T/H^4$ is plotted as functions of the normalized electric field $\lambda = |-eE|/H^2$, for different values of the normalized scalar field mass $\mu = m/H$.

In Fig. 1 the regularized trace (19) is plotted as function of the electric field for different values of the scalar field mass. A numerical investigation shows that for a massive scalar field $\mu \gtrsim 1$, the sign of the trace is positive. However, for a light scalar field $\mu \lesssim 1$, the trace vanishes at $\lambda = L$. In the domain $\lambda < L$ the sign of the trace is negative, whereas in the domain $\lambda > L$ the sign of the trace is positive. Further numerical investigations illustrate that $L \simeq 0.4\mu$. Therefore in the semiclassical regime $\lambda^2 + \mu^2 \gg 1$ the sign of the trace is positive, whereas in the infrared regime $\lambda^2 + \mu^2 \ll 1$ the sign of the trace is negative.

## 4. Conclusions

These results for the trace of the energy-momentum tensor would be important for discussing the gravitational backreaction effect of the Schwinger pair creation. As a consequence of considering the *vacuum* equation of state for the created Schwinger pairs, see Eq. (10), the Einstein equation leads to the time evolution equation for the Hubble constant as[10]

$$\frac{dH}{dt} = -\frac{\pi T}{3M_{\rm P}^2}, \tag{20}$$

where $t$ is the proper cosmological time, see Eq. (3), and $M_{\rm P}$ is the Planck mass. Considering the results of Sec. 3 for the trace (19) and Eq. (20), we conclude that in the semiclassical regime $\lambda^2 + \mu^2 \gg 1$, that the sign of the trace is positive the Hubble constant decays $dH/dt < 0$. This result is in agreement with the results obtained in Ref. 6 for the decay of the Hubble constant due to the semiclassical

Schwinger pair creation. We find that in the infrared regime $\lambda^2 + \mu^2 \ll 1$, that the sign of the trace is negative, Eq. (20) implies that $dH/dt > 0$. In Ref. 15 the authors found that the gravitational backreaction effects of the quantum fluctuations may lead to a similar behavior of the Hubble constant, i.e., a period of superacceleration with $dH/dt > 0$.

## Acknowledgments

E. B. is supported by the University of Kashan.

## References

1. J. S. Schwinger, *Phys. Rev.* **82**, 664 (1951).
2. J. Martin, *Lect. Notes Phys.* **738**, 193 (2008).
3. J. Garriga, *Phys. Rev. D* **49**, 6343 (1994).
4. M. B. Frb, J. Garriga, S. Kanno, M. Sasaki, J. Soda, T. Tanaka and A. Vilenkin, *J. Cosmol. Astropart. Phys.* **04**, 009 (2014).
5. T. Kobayashi and N. Afshordi, *J. High Energy Phys.* **10**, 166 (2014).
6. E. Bavarsad, C. Stahl and S. S. Xue, *Phys. Rev. D* **94**, no. 10, 104011 (2016).
7. S. P. Kim and D. N. Page, *Phys. Rev. D* **78**, 103517 (2008).
8. E. Bavarsad, S. P. Kim, C. Stahl and S. S. Xue, *Phys. Rev. D* **97**, 025017 (2018).
9. E. Bavarsad, S. P. Kim, C. Stahl and S. S. Xue, *Eur. Phys. J. Web Conf.* **168**, 03002 (2018).
10. E. Mottola, *Phys. Rev. D* **31**, 754 (1985).
11. L. Parker and S. A. Fulling, *Phys. Rev. D* **9**, 341 (1974).
12. S. A. Fulling and L. Parker, *Annals Phys.* **87**, 176 (1974).
13. J. S. Dowker and R. Critchley, *Phys. Rev. D* **13**, 3224 (1976).
14. S. Habib, C. Molina-Paris and E. Mottola, *Phys. Rev. D* **61**, 024010 (2000).
15. T. Markkanen and A. Rajantie, *J. High Energy Phys.* **01**, 133 (2017).
16. T. Markkanen, *Eur. Phys. J. C* **78**, no. 2, 97 (2018).
17. G. W. Gibbons and S. W. Hawking, *Phys. Rev. D* **15**, 2738 (1977).
18. Y. Zhang, X. Ye and B. Wang, arXiv:1903.10115 [gr-qc].
19. E. Bavarsad and N. Margosian, *Journal of Research on Many-body Systems*, **8**, number 17, 1 (2018).
20. E. Bavarsad and M. Mortezazadeh, *Iranian Journal of Physics Research*, **18**, No. 1, 91 (2018).
21. A. Ferreiro and J. Navarro-Salas, *Phys. Rev. D* **97**, no. 12, 125012 (2018).
22. L. Parker and D. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity* (Cambridge University Press, Cambridge, 2009).
23. T. Hayashinaka and S. S. Xue, *Phys. Rev. D* **97**, no. 10, 105010 (2018).

# Photon polarization oscillations

Fabio Briscese*

*Department of Physics, Southern University of Science and Technology, Shenzhen, 518055, China*
*and Istituto Nazionale di Alta Matematica Francesco Severi, Gruppo Nazionale di Fisica Matematica, Città Universitaria, P.le A. Moro 5, 00185 Rome, Italy*
**E-mail:briscese.phys@gmail.com, briscesef@sustc.edu.cn*

Nicolò Burzillà and Andrea Dosi

*Department of Physics, Southern University of Science and Technology, Shenzhen, 518055, China*

Quantum corrections to the Maxwell equations induced by light-by-light (LbL) scattering can significantly modify the propagation of light in vacuum. Studying the Heisenberg-Euler Lagrangian, it can be shown that, in some configurations, the polarization of plane monochromatic waves oscillates periodically between different helicity states, due to LbL scattering. We discuss the physical implications of this finding, and the possibility of measuring this effect in optical experiments.

*Keywords*: Nonlinear optics in vacuum, light-by-light scattering, multiscale perturbative approach.

Despite the fact that the equations of the classical electromagnetic field are linear, quantum corrections due to photon-photon scattering introduce nonlinear effects in vacuum. The quantum corrections due to photon-photon scattering were calculated a long time ago by Heisenberg and Euler [1], and extensively studied by other authors [2–5]. The effective Lagrangian of the electromagnetic field, obtained retaining only one electron loop corrections, is [5]

$$L = \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \epsilon^2\left[(F_{\mu\nu}F^{\mu\nu})^2 - \frac{7}{16}\left(F_{\mu\nu}\tilde{F}^{\mu\nu}\right)^2\right], \tag{1}$$

where $F^{\mu\nu} = A^{\mu,\nu} - A^{\nu,\mu}$ is the electromagnetic field,[a] $A^\mu$ is the electromagnetic four-potential, $\tilde{F}^{\mu\nu} \equiv \epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}$, and $\epsilon^2 = \alpha^2(\hbar/m_e c)^3/90m_e c^2$, $\alpha = e^2/4\pi\epsilon_0\hbar c \simeq 1/137$ is the fine structure constant, $\epsilon_0$ the dielectric permeability of vacuum, and $m_e$ the electron mass. The Lagrangian (1) is accurate when it is possible to neglect other quantum effects. For instance, for low energetic photons of energies $E_\gamma \ll m_e c^2$, particles creation is inhibited, and the photon-photon scattering is the only process inducing quantum correction to the Maxwell equations.

The terms $\propto \epsilon^2$ in the Lagrangian (1) take into account photon-photon scattering, and induce cubic corrections in the equations for the four-potential $A^\mu$. Since

[a] In this paper we use the covariant formalism, so that the zeroth coordinate is defined as $x^0 = c\,t$.

$\epsilon^2 \simeq 4 \times 10^{-31} m^3/J$, so that $\epsilon^2 F_{\mu\nu}F^{\mu\nu}$ is extremely small in realistic laboratory conditions, such corrections are usually negligible with high accuracy.

However, in many physical situations, tiny perturbations produce huge effects on a system, due to the action of hidden resonances. Indeed, resonances can be used to amplify the effect of extremely small perturbations. Exploiting this idea, in Refs. 6, 7 it has been shown that the dynamics described by the Lagrangian (1) is unstable for some configurations of the electromagnetic field, due to resonances.

For instance, considering two plane counterpropagating electromagnetic waves in vacuum, the nonlinear terms in (1) generate resonant (or secular) corrections in the equations of the electromagnetic field. Introducing a slow time variable, the secular terms can be treated in a multiscale scheme. In fact, the amplitudes of the two counterpropagating waves satisfy a system of nonlinear coupled ordinary differential equations in the slow time. The analysis of this system shows that, for some initial conditions, the effect of photon-photon scattering is unexpectedly relevant, and consists of a continuous oscillation in the polarization of the two beams between different helicity states.

Without loss of generality, we use the Lorentz gauge $\partial_\alpha A^\alpha = 0$; and we express the polarization vectors in terms of left and right polarizations $\hat{e}_L = (1, i, 0)/\sqrt{2}$ and $\hat{e}_R = (1, -i, 0)/\sqrt{2}$. Let us express the four-potential $A$ of the classical electromagnetic field in the form

$$A = a^\alpha + b^\alpha + c.c., \quad \text{with} \quad a = \left(a_L \hat{e}^L + a_R \hat{e}^R\right) e^{ikx}\,,\; b = \left(b_L \hat{e}^L + b_R \hat{e}^R\right) e^{ihx}, \tag{2}$$

where c.c. stands for complex conjugate. The wave vectors in (2) are given by

$$k = (k_0, 0, 0, k_3), \quad h = (h_0, 0, 0, h_3), \tag{3}$$

with $k_0/k_3 = -h_0/h_3 = 1$, so that the two waves $a$ and $b$ are counterpropagating.

When nonlinearities are neglected, (1) reduces to the Maxwell Lagrangian, and (2) is a solution when polarization vectors are constant. When nonlinearities are considered, the solutions for the Lagrangian (1) are still in the form (2), but with the polarization vectors depending on a slow time

$$y^0 \equiv \epsilon^2 x^0. \tag{4}$$

The dependence of the polarization vectors from the slow time is given by the following system

$$\begin{aligned}
&i\partial_{y^0} a_L + 16 k_0 h_0^2 \left(-3\, a_L \left(|b_L|^2 + |b_R|^2\right) + 22\, a_R b_L \bar{b}_R\right) = 0 \\
&i\partial_{y^0} a_R + 16 k_0 h_0^2 \left(-3\, a_R \left(|b_L|^2 + |b_R|^2\right) + 22\, a_L b_R \bar{b}_L\right) = 0 \\
&i\partial_{y^0} b_L + 16 k_0^2 h_0 \left(-3\, b_L \left(|a_L|^2 + |a_R|^2\right) + 22\, b_R a_L \bar{a}_R\right) = 0 \\
&i\partial_{y^0} b_R + 16 k_0^2 h_0 \left(-3\, b_R \left(|a_L|^2 + |a_R|^2\right) + 22\, b_L a_R \bar{a}_L\right) = 0\,.
\end{aligned} \tag{5}$$

that has been obtained by a multi-scale perturbative expansion of the equations of motion[6,7]. We recall that the multiscale approach is useful when the dynamics evolves on widely different scales. In this case, the time dependence of the electromagnetic field is split into fast and slow time variables $x^0$ and $y^0$.

Let us study (5) in detail. It is quite immediate to recognize that the energy densities $< \rho_a >= k_0^2 \left(|a_L|^2 + |a_R|^2\right)$ and $< \rho_b >= h_0^2 \left(|b_L|^2 + |b_R|^2\right)$ are constant. Therefore, the intensities of the two plane waves $a^\mu$ and $b^\mu$ are conserved separately. Furthermore, the spin conservation implies that the quantity $S = k_0 \left(|a_L|^2 - |a_R|^2\right) + h_0 \left(|b_L|^2 - |b_R|^2\right)$ is also constant. Exploiting these relations, the system can be simplified and then integrated, see Refs. 6, 7. However, to understand the dynamics under study, it is sufficient to solve (5) numerically.

We choose the initial conditions in such a way that, at least one of the products $a_L\, a_R$ or $b_L\, b_R$ is nonzero at the initial time. Numerical solutions show that, in such case, the polarizations of the two counterpropagating waves change periodically. In Fig. 1, we plot $|a_L|^2/|a_L^0|^2 + |a_R^0|^2$ and $|a_R|^2/|a_L^0|^2 + |a_R^0|^2$ for the solution of (5) with realistic initial values of the electromagnetic potential, that is $|a_L^0|^2 = 10^3 J/m$, $a_R^0 = 0$, $|b_L^0|^2 = |b_R^0|^2 = 10^3 J/m$, $k_0 = h_0 = 10^7\, m^{-1}$, corresponding to a laser of intensity $I \simeq 10^{23}\, W/cm^2$ and wavelenght $\lambda \sim 1\, \mu m$. We see that $|a_R|$ is initially zero, but it grows to $|a_R| = |a_L^0|$, while $|a_L|$ goes from $|a_L^0|$ to zero. Thus, the $a$ beam, initially in the left-handed polarization, switches to the right-handed polarization. It remains in this state most of the time, until it returns to its initial left-handed configuration. The variation of the $b$ beam is depicted in Fig. 2, where we plot $|b_L|^2/|b_L^0|^2 + |b_R^0|^2$ and $|b_R|^2/|b_L^0|^2 + |b_R^0|^2$ for the same initial values. The beam $b$ is initially circularly polarized, since $|b_L^0|^2 = |b_R^0|^2$, but it rapidly goes to a left-handed configuration with $|b_R| = 0$ and $|b_L|^2 = 2|b_L^0|^2$. It remains in this state most of the time, until it returns to its initial state. This dynamics is repeated periodically.

Thus, the beam $a$ oscillates between left and right polarizations, while the beam $b$ switches periodically between linear and right-handed helicities. The period of such polarization oscillations is $\Delta y^0 \simeq 10^{-26}\, m^4/J$ in the slow time $y^0$. This value is in good agreement with theoretical estimations $\Delta y^0 \sim inf\left\{1/k_0 h_0^2 |b_0|^2, 1/k_0^2 h_0 |a_0|^2\right\}$ obtained in Refs. 6, 7. The corresponding period in the physical time $x^0 = t$ is $T = \Delta y^0/\epsilon^2\, c \simeq 10^{-4}\, s$.

Finally, we consider the possibility of observing the polarization oscillations in optical experiments. The search for signatures of the photon-photon scattering in optics is in progress[8–32]. We can estimate the time of recurrence of the polarization oscillations for light beams produced in petawatt class lasers, which will be available in the near future. The intensities attainable in these lasers reach $I \sim 10^{23}\, W/cm^2$ [33,34], giving a recurrence time $T_i \sim 4 \times 10^2\, (\lambda/m)\, s$, where $\lambda/m$ is the laser wavelength in meters (we used $k \sim h \sim 2\pi/\lambda$ and $k^2 a^2 \sim k^2 b^2 \sim < \rho > \sim I/c$). Therefore, for realistic lasers with $\lambda \sim 1\, \mu m$, observation times can be of the order of $10^{-3}\, s$; to be compared with those estimated

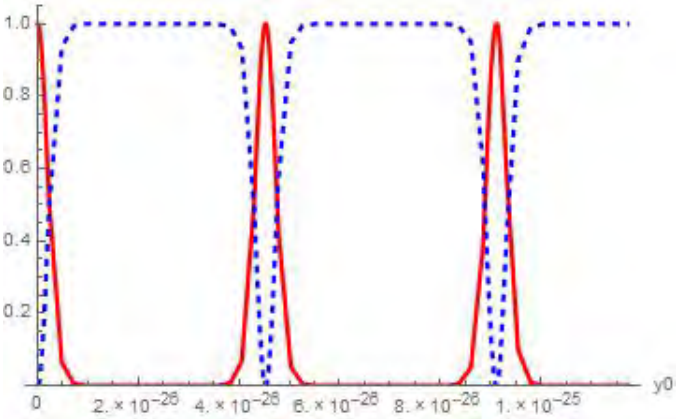



Fig. 1. We plot the evolution of $|a_L|^2/|a_L^0|^2+|a_R^0|^2$ (solid red line) and $|a_R|^2/|a_L^0|^2+|a_R^0|^2$ (dashed blue line) against $y^0$ (in units of $m^4/J$) for $|a_L^0|^2 = 10^3 J/m$, $a_R^0 = 0$, $|b_L^0|^2 = |b_R^0|^2 = 10^3 J/m$, $k_0 = h_0 = 10^7\, m^{-1}$.

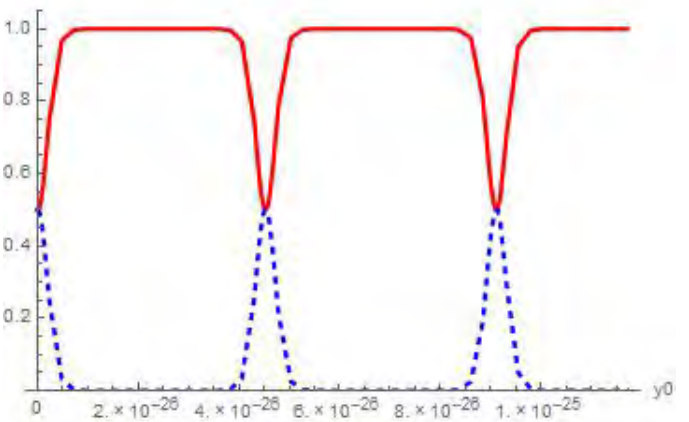



Fig. 2. We plot the evolution of $|b_L|^2/|b_L^0|^2+|b_R^0|^2$ (solid red line) and $|b_R|^2/|b_L^0|^2+|b_R^0|^2$ (dashed blue line) against $y^0$ (in units of $m^4/J$) for $|a_L^0|^2 = 10^3 J/m$, $a_R^0 = 0$, $|b_L^0|^2 = |b_R^0|^2 = 10^3 J/m$, $k_0 = h_0 = 10^7\, m^{-1}$.

in Ref. 14. This lets us hope to be able to observe polarization oscillations in two counterpropagating petawatt laser beams.

In conclusion, the extremely weak photon-photon interaction might be responsible for surprisingly strong deviations from the free dynamics of electromagnetic waves. In the case of two counterpropagating laser beams, the evolution of the electromagnetic waves can change dramatically with respect to the linear Maxwell equations, entailing slow oscillations of the polarizations of the beams.

## References


1. W. Heisenberg and H. Euler, *Folgerungen aus der Diracschen Theorie des Positrons*, Z. Phys. **98**, 714 (1936).
2. D. A. Dicus, C. Kao, W. W. Repko, *Effective Lagrangians and low energy photon-photon scattering*, Phys.Rev. D **57** (1998) 2443-2447, arXiv:hep-ph/9709415.
3. R. Karplust, M. Neuman, *The Scattering of Light by Light*, Phys. Rev. **83**, 4, 776-784.
4. R.A. Leo, G. Soliani, *Tensor Amplitudes for Elastic Photon-Photon Scattering*, Il Nuovo Cimento **30** A, N. 2, 1975.
5. J. Shwinger, *On Gauge Invariance and Vacuuwn Polarization*, Phys. Rev. E **82**, 5 (1951).
6. F. Briscese, *Collective behavior of light in vacuum*, Phys. Rev. A **97** (March 2, 2018) 033803.
7. F. Briscese, *Light polarization oscillation induced by photon-photon scattering*, Phys. Rev. A **96** (November 1, 2017) 053801.
8. G. O. Schellstede, V. Perlick, C. Lammerzahl, *Testing non-linear vacuum electrodynamics with Michelson interferometry*, Phys. Rev. D **92** (2015), arXiv:1504.03159 [gr-qc].
9. P. Gaete, J. A. Helayel-Neto, *A note on nonlinear electrodynamics*, arXiv: 1709.03869 [physics.gen-ph].
10. O.J. Pike, F. Mackenroth, E.G. Hill, S.J. Rose, *A photon–photon collider in a vacuum hohlraum*, Nature Photon. 8 (2014) 434-436.
11. V. Dinu, T. Heinzl, A. Ilderton, M. Marklund and G. Torgrimsson, *Photon polarization in light-by-light scattering: Finite size effects*, Phys. Rev. D **90** (2014) no.4, 045025 [arXiv:1405.7291 [hep-ph]].
12. V. Dinu, T. Heinzl, A. Ilderton, M. Marklund and G. Torgrimsson, *Vacuum refractive indices and helicity flip in strong-field QED*, Phys. Rev. D **89** (2014) no.12, 125003 [arXiv:1312.6419 [hep-ph]].
13. B. King, N. Elkina, *Vacuum birefringence in high-energy laser-electron collisions*, Phys. Rev. A **94**, 062102 (2016), arXiv:1603.06946 [hep-ph].
14. B. King, *et al., A matterless double slit*, Nature Photonics **4**, 92–94 (2010).
15. S. Bragin, *et al., High-energy vacuum birefringence and dichroism in an ultrastrong laser field*, arXiv:1704.05234 [hep-ph].
16. S. Shakeri, S. Z. Kalantari, and S. Xue, *Polarization of a probe laser beam due to nonlinear QED effects*, Phys. Rev. A **95**, 012108 (2017).
17. H. Schlenvoigt, *et al., Detecting vacuum birefringence with x-ray free electron lasers and high-power optical lasers: a feasibility study*, Phys. Scripta **91**, 023010 (2016).
18. F. Karbstein and C. Sundqvist, *vacuum birefringence using x-ray free electron and optical highintensity lasers*, Phys. Rev. D **94**, 013004 (2016).

19. G. Zavattini, *et al., A polarisation modulation scheme for measuring vacuum magnetic birefringence with static fields*, Eur. Phys. J. C **76**, 294 (2016).
20. D. M. Tennant, *wave mixing as a probe of the vacuum*, Phys. Rev. D **93**, 125032 (2016).
21. H. Gies, *et al., Quantum reflection of photons off spatio-temporal electromagnetic field inhomogeneities*, New J. Phys. **17**, 043060 (2015).
22. F. Fillion-Gourdeau, *et al., for the detection of mixing processes in vacuum*, Phys. Rev. A **91**, 031801 (2015).
23. F. Karbstein and R. Shaisultanov, *Photon propagation in slowly varying inhomogeneous electromagnetic fields*, Phys. Rev. D **91**, 085027 (2015).
24. H. Hu and J. Huang, *Modified light-cone condition via vacuum polarization in a time-dependent field*, Phys. Rev. A **90**, 062111 (2014).
25. Y. Monden and R. Kodama, *Interaction of two counterpropagating laser beams with vacuum*, Phys. Rev. A 86, 033810 (2012).
26. B. King and C. H. Keitel, *Photon-photon scattering in collisions of intense laser pulses*, New J. Phys. **14**, 103002 (2012).
27. G. Yu. Kryuchkyan and K. Z. Hatsagortsyan, *Bragg Scattering of Light in Vacuum Structured by Strong Periodic Fields*, Phys. Rev. Lett. **107**, 053604 (2011).
28. K. Homma, D. Habs, and T. Tajima, *Probing vacuum birefringence by phase-contrast Fourier imaging under fields of high-intensity lasers*, Appl. Phys. B **104**, 769 (2011).
29. D. Tommasini, *et al., Detecting photon-photon scattering in vacuum at exawatt lasers*, Phys. Rev. A **77**, 042101 (2008).
30. E. Lundstrom, *et al., Using High-Power Lasers for Detection of Elastic Photon-Photon Scattering*, Phys. Rev. Lett. **96**, 083602 (2006).
31. A. Di Piazza, K. Z. Hatsagortsyan, C. H. Keitel, *Light Diffraction by a Strong Standing Electromagnetic Wave*, Phys. Rev. Lett. **97**, 083603 (2006).
32. T. Heinzl, *et al., On the observation of vacuum birefringence*, Opt. Commun. **267**, 318-321 (2006).
33. C. Danson, D. Hillier, N. Hopps, and D. Neely, *Petawatt class lasers worldwide*, High Power Laser Science and Engineering **3**, e3 (2015).
34. T. M. Jeong and J. Lee, *Femtosecond petawatt laser*, Ann. Phys. **526**, 157–172 (2014).

# Charge emission from near-extremal charged black holes

Chiang-Mei Chen*

*Department of Physics, National Central University, Chungli 320, Taiwan*
** E-mail: cmchen@phy.ncu.edu.tw*

Sang Pyo Kim§

*Department of Physics, Kunsan National University, Kunsan 54150, Korea*
*and Institute of Theoretical Physics, Chinese Academy of Sciences, Beijing 100190, China*
*§ E-mail: sangkim@kunsan.ac.kr*

Jia-Rui Sun†

*School of Physics and Astronomy, Sun Yat-Sen University, Guangzhou 510275, China*
*† E-mail: sunjiarui@sysu.edu.cn*

Using the symmetry of the near-horizon geometry and applying quantum field theory of a complex scalar field, we study the spontaneous pair production of charged scalars from near-extremal rotating, electrically and/or magnetically charged black holes. Analytical expressions for pair production, vacuum persistence and absorption cross section are found, and the spectral distribution is given a thermal interpretation. The pair production in near-extremal black holes has a factorization into the Schwinger effect in AdS and Schwinger effect in Rindler space, measuring the deviational from extremality. The associated holographical correspondence is confirmed at the 2-point function level by comparing the absorption cross section ratio as well as the pair production rate both from the gravity and the conformal field theories. The production of monopoles is discussed.

*Keywords*: Charged black holes; Schwinger pair production; AdS/CFT correspondence.

## 1. Introduction

Black holes, as a consequence of vacuum fluctuations, create all species of particle pairs near their horizons, half of which radiate to the spatial infinity, leading to the so-called Hawking radiation.[1,2] Schwinger mechanism has also been known for long that virtual pairs created from vacuum fluctuations could be physically separated by a strong external electric field and substantiated as real pairs leading to spontaneous pair production.[3–5] Charged black hole thus provide both strong gravitational and electric fields resulting in either the Hawking radiation and/or Schwinger pair production, which intermingles part of quantum gravity effects and strong field quantum electrodynamics (QED) effects. The emission of charges from charged black holes has been studied independently of the Hawking radiation[6–10] and intensively studied since then both for nonextremal and extremal cases[11–14] (for a review and references, see Refs. 15 and 16).

In this paper, we review the emission of scalar charges from near-extremal charged black holes from the view point of quantum field theory in near-horizon geometry. A near-extremal black hole has an extremely small Hawking temperature, and therefore its Hawking radiation is exponentially suppressed. On the other

hand, the electric field on the horizon becomes strong enough to trigger the emission of charges dominantly through the Schwinger mechanism. This study differs from those early works in the following aspect. A conventional wisdom was to calculate the tunneling probability of virtual particles in negative energy states of the Dirac sea into positive energy states, whose mass gap is modified by the charged black hole to allow finite tunneling probability.[8–14] In the phase-integral method the leading contribution to the pair production comes from the poles near the horizon.[17] In our approach, we make use of the local geometry near the horizon: the geometry of a near-extremal charged black hole has the structure of $\mathrm{AdS}_2 \times \mathrm{S}^2$ for a nonrotating one and it has the warped $\mathrm{AdS}_3$ for a rotating one, while the geometry of a non-extremal charged black hole is $\mathrm{Rindler}_2 \times \mathrm{S}^2$ for a nonrotating one. Hence, the symmetry of near-horizon geometries allows the separation of quantum fields, which leads to analytical expressions for the in-vacuum and the out-vacuum and thereby the spectral distribution of spontaneously created pairs. The Schwinger mechanism and QED vacuum polarization in $\mathrm{dS}_2$ and $\mathrm{AdS}_2$ and their relation to black holes are discussed in Ref. 17.

The Schwinger mechanism and the Hawking radiation as quantum tunneling[18] are, however, intertwined for the spontaneous pair production occurring in charged black holes, whose dominant contribution comes from the near-horizon region. Here we review and study the scalar particle emission in the spacetime of the near-horizon region of the near-extremal charged black holes. The simplest model is the scalar production in Reissner-Nordström (RN) black holes[19] (see Ref. 20 for spinor production). In this case the spacetime has an $\mathrm{AdS}_2 \times \mathrm{S}^2$ structure and the electric field is constant near the horizon. Using the symmetry of geometry, one can analytically solve the Klein-Gordon (KG) equation and give an exact expression for the production rate. The analysis can be generalized to Kerr-Newman (KN) black holes[21] and also include a magnetic charge.[22] Then, the angular momentum deforms the near-horizon spacetime geometry to be a warped $\mathrm{AdS}_3$, but the KG equation still can be separated and solved analytically. The pair production has a remarkable thermal interpretation based on the discussions in Refs. 23, 24, 25, 26 and 17. Moreover, the scalar production has a nice conformal field theory (CFT) dual picture supporting the KN/CFTs correspondence.[27,28]

## 2. Emission of Charges from near-extremal KN black holes

The near-horizon geometry of a near-extremal dyonic KN black hole has the structure of a warped $\mathrm{AdS}_3$ as in Ref. 22

$$ds^2 = \Gamma(\theta)\left[-(\rho^2 - B^2)d\tau^2 + \frac{d\rho^2}{\rho^2 - B^2} + d\theta^2\right] + \gamma(\theta)(d\varphi + b\rho d\tau)^2, \tag{1}$$

$$A_{[1]} = -\frac{Q(r_0^2 - a^2\cos^2\theta) - 2Pr_0 a\cos\theta}{\Gamma(\theta)}\rho d\tau - \frac{Qr_0 a\sin^2\theta - P(r_0^2 + a^2)\cos\theta \pm P\Gamma(\theta)}{\Gamma(\theta)}d\varphi, \tag{2}$$

where

$$\Gamma(\theta) = r_0^2 + a^2\cos^2\theta, \qquad \gamma(\theta) = \frac{(r_0^2+a^2)^2\sin^2\theta}{r_0^2 + a^2\cos^2\theta},$$
$$b = \frac{2ar_0}{r_0^2+a^2}, \qquad r_0 = \sqrt{Q^2+P^2+a^2}. \tag{3}$$

Here, $a$ is the angular momentum parameter and $Q, P$ are the electric and magnetic charges of the original dyonic KN black holes, and $B$ measures a deviation from the extremal limit and acts as the new horizon $\rho_H = B$ in this geometry. The zero magnetic charge ($P = 0$) corresponds to the KN black hole while the zero angular momentum ($a = 0$) corresponds to the RN black hole with both electric and magnetic charges and the gauge potential for the electric charge and Dirac monopole in the near-horizon geometry

$$A_{[1]} = -Q\rho d\tau + P\big(\cos\theta \mp 1\big)d\varphi. \tag{4}$$

Associated to the black hole thermodynamics, the Hawking temperature, entropy, angular velocity, and chemical potentials ($\bar{\Phi}_{\rm H}$ is given from the Hodge dual of Maxwell field $dA_{[1]}$ in the original KN metric) are

$$T_{\rm H} = \frac{B}{2\pi}, \qquad S_{\rm BH} = \pi(r_0^2 + a^2 + 2Br_0), \qquad \Omega_{\rm H} = -\frac{2ar_0B}{r_0^2+a^2},$$
$$\Phi_{\rm H} = \frac{Q(Q^2+P^2)B}{r_0^2+a^2}, \qquad \bar{\Phi}_{\rm H} = \frac{P(Q^2+P^2)B}{r_0^2+a^2}. \tag{5}$$

The thermodynamical variables (5), as the metric (1) and the gauge field (2) do, contain the KN black hole, RN black hole as well as dyonic RN black hole with both electric and magnetic charges in the limit of $P = 0$, $a = P = 0$, and $a = 0$, respectively. The KG equation for scalar dyons can be exactly solved. An effective potential due to the electromagnetic and gravitational interactions induces tunneling processes for the pair production. By imposing a suitable boundary condition, one can obtain the production rate from the incoming and outgoing fluxes on the horizon and the asymptotic boundary of the near-horizon geometry.[19] For instance, without an incoming flux at the asymptotic outer boundary, the relative ratio of the outgoing (transmitted) flux in the asymptotic region to the incoming (reflected) flux at the horizon counts the spontaneously produced particles while the ratio of the outgoing (incident) flux to the incoming flux at the horizon gives the vacuum persistence amplitude due to vacuum fluctuations. On the other hand, the group velocity for created fermions leads to the relative ratio of the outgoing flux in the asymptotic region to the outgoing flux at the horizon for the fermion production probability.

We now focus on production of bosonic particles. Following Ref. 29, the flux conservation

$$|D_{\rm incident}| = |D_{\rm reflected}| + |D_{\rm transmitted}|, \tag{6}$$

is related to the Bogoliubov relation

$$|\mathcal{A}|^2 - |\mathcal{B}|^2 = 1, \tag{7}$$

where the vacuum persistence amplitude $|\mathcal{A}|^2$ and the mean number of produced pairs $|\mathcal{B}|^2$ are given by the ratios of the flux components

$$|\mathcal{A}|^2 \equiv \frac{|D_{\text{incident}}|}{|D_{\text{reflected}}|}, \qquad |\mathcal{B}|^2 \equiv \frac{|D_{\text{transmitted}}|}{|D_{\text{reflected}}|}. \tag{8}$$

Moreover, from the viewpoint of scattering of an incident flux from the asymptotic boundary, we can define the absorption cross section ratio as

$$\sigma_{\text{abs}} \equiv \frac{|D_{\text{transmitted}}|}{|D_{\text{incident}}|} = \frac{|\mathcal{B}|^2}{|\mathcal{A}|^2}. \tag{9}$$

In comparison, the group velocity for fermions leads to the vacuum persistence amplitude and the mean number of produced fermion pairs

$$|\mathcal{A}|^2 \equiv \frac{|D_{\text{reflected}}|}{|D_{\text{incident}}|}, \qquad |\mathcal{B}|^2 \equiv \frac{|D_{\text{transmitted}}|}{|D_{\text{incident}}|}, \tag{10}$$

where the fluxes $D_{\text{incident}}, D_{\text{reflected}}$ and $D_{\text{transmitted}}$ are computed from the spin-diagonal equations.

Using the following ansatz (hereafter parameters $m, q$, and $p$ are the mass, electric and magnetic charges of a scalar field)

$$\Phi(\tau, \rho, \theta, \varphi) = \mathrm{e}^{-i\omega\tau + i[n \mp (qP - pQ)]\varphi} R(\rho) S(\theta), \tag{11}$$

a straightforward calculation leads to the Bogoliubov coefficients and the absorption cross section ratio as (the derivation in detail can be found in Ref. 22)

$$|\mathcal{A}|^2 = \frac{\cosh(\pi\kappa - \pi\mu)\cosh(\pi\tilde{\kappa} + \pi\mu)}{\cosh(\pi\kappa + \pi\mu)\cosh(\pi\tilde{\kappa} - \pi\mu)}, \tag{12}$$

$$|\mathcal{B}|^2 = \frac{\sinh(2\pi\mu)\sinh(\pi\tilde{\kappa} - \pi\kappa)}{\cosh(\pi\kappa + \pi\mu)\cosh(\pi\tilde{\kappa} - \pi\mu)}, \tag{13}$$

$$\sigma_{\text{abs}} = \frac{|\mathcal{B}|^2}{|\mathcal{A}|^2} = \frac{\sinh(2\pi\mu)\sinh(\pi\tilde{\kappa} - \pi\kappa)}{\cosh(\pi\kappa - \pi\mu)\cosh(\pi\tilde{\kappa} + \pi\mu)}, \tag{14}$$

where three essential parameters are ($\lambda_l$ being a separation constant)

$$\tilde{\kappa} = \frac{\omega}{B}, \quad \kappa = \frac{(qQ+pP)(Q^2+P^2) - 2nar_0}{r_0^2 + a^2}, \quad \mu = \sqrt{\kappa^2 - m^2(r_0^2 + a^2) - \lambda_l - \frac{1}{4}}, \tag{15}$$

in which $\mu^2$ is positive due to the BF bound violation in the $\mathrm{AdS}_2$ spacetime.[22]

Following our previous studies,[19,21] the mean number of produced pairs (13) can be rewritten as

$$\mathcal{N} = |\mathcal{B}|^2 = \left(\frac{\mathrm{e}^{-2\pi\kappa + 2\pi\mu} - \mathrm{e}^{-2\pi\kappa - 2\pi\mu}}{1 + \mathrm{e}^{-2\pi\kappa - 2\pi\mu}}\right)\left(\frac{1 - \mathrm{e}^{-2\pi\tilde{\kappa} + 2\pi\kappa}}{1 + \mathrm{e}^{-2\pi\tilde{\kappa} + 2\pi\mu}}\right). \tag{16}$$

Note that the mean number (16) has a similar form as those of charged scalars with $\lambda_l = l(l+1)$ for spherical harmonics in a near-extremal RN black hole[19] and with

the same quantum number in KN black hole[21] since the near-horizon geometry has an $\mathrm{AdS}_2 \times S^2$ and a warped $\mathrm{AdS}_3$ for the near-extremal RN and KN black holes, respectively. According to Refs. 24 and 25, one may introduce an effective temperature and its associated counterpart

$$T_{\mathrm{KN}} = \frac{\bar{m}}{2\pi\kappa - 2\pi\mu} = T_U + \sqrt{T_U^2 + \frac{\mathcal{R}}{8\pi^2}}, \quad \bar{T}_{\mathrm{KN}} = \frac{\bar{m}}{2\pi\kappa + 2\pi\mu} = T_U - \sqrt{T_U^2 + \frac{\mathcal{R}}{8\pi^2}}, \tag{17}$$

where the effective mass $\bar{m}$ is

$$\bar{m} = \sqrt{m^2 - \frac{\lambda + 1/4}{2}\mathcal{R}}, \tag{18}$$

and the corresponding Unruh temperature $T_U$ and AdS curvature $\mathcal{R}$ are

$$T_U = \frac{\kappa}{2\pi\bar{m}(r_0^2 + a^2)} = \frac{(qQ + pP)(Q^2 + P^2) - 2nar_0}{2\pi\bar{m}(r_0^2 + a^2)^2}, \quad \mathcal{R} = -\frac{2}{r_0^2 + a^2}. \tag{19}$$

Note that an AdS space binds a pair and thus increases the effective mass in contrast to a dS space which separates a pair and reduces the effective mass.

Finally, we may introduce a thermal interpretation by factorizing the mean number (16) as

$$\mathcal{N} = \mathrm{e}^{\frac{\bar{m}}{T_{\mathrm{KN}}}} \times \underbrace{\left( \frac{\mathrm{e}^{-\frac{\bar{m}}{T_{\mathrm{KN}}}} - \mathrm{e}^{-\frac{\bar{m}}{T_{\mathrm{KN}}}}}{1 + \mathrm{e}^{-\frac{\bar{m}}{T_{\mathrm{KN}}}}} \right)}_{\text{Schwinger effect in } \mathrm{AdS}_2} \times \underbrace{\left\{ \frac{\mathrm{e}^{-\frac{\bar{m}}{T_{\mathrm{KN}}}} \left( 1 - \mathrm{e}^{-\frac{\omega - q\Phi_{\mathrm{H}} - p\bar{\Phi}_{\mathrm{H}} - n\Omega_{\mathrm{H}}}{T_{\mathrm{H}}}} \right)}{1 + \mathrm{e}^{-\frac{\omega - q\Phi_{\mathrm{H}} - p\bar{\Phi}_{\mathrm{H}} - n\Omega_{\mathrm{H}}}{T_{\mathrm{H}}}} \mathrm{e}^{-\frac{\bar{m}}{T_{\mathrm{KN}}}}} \right\}}_{\text{Schwinger effect in Rindler}_2}. \tag{20}$$

The first parenthesis is the Schwinger effect with the effective temperature $T_{\mathrm{KN}}$ in $\mathrm{AdS}_2$[23] and the second parenthesis is the Schwinger effect in the Rindler space,[30] in which the Unruh temperature is given by the Hawking temperature and charges have the chemical potentials $\Phi_{\mathrm{H}}, \bar{\Phi}_{\mathrm{H}}$, and $\Omega_{\mathrm{H}}$ while the effective temperature for the Schwinger effect due to the electric field on the horizon is still given by $T_{\mathrm{KN}}$. The extremal KN black hole can be obtained by taking the limit $T_{\mathrm{H}} = 0$ and $\tilde{\kappa} = 0$, in which the second parenthesis multiplied by the prefactor becomes unity. An interesting physics is the monopole production: upon the Dirac quantization $(ep) = 2\pi$ of charges and a small electric coupling $\alpha = e^2 = 1/137$, the symmetry of the pair production of electric and magnetic charges from dyonic black holes is broken due to $e/p = \alpha/2\pi$. Monopole production and related physics go beyond the scope of this paper and will be addressed elsewhere.

## 3. KN/CFT duality

According to the KN/CFTs duality,[28,31] the absorption cross section ratio of scalar field in Eq. (14) corresponds to that of its dual operator in the dual two-dimensional

CFT with left- and right-hand sectors

$$\sigma_{\rm abs}\sim\frac{T_{\rm L}^{2h_{\rm L}-1}T_{\rm R}^{2h_{\rm R}-1}}{|\Gamma(2h_L)||\Gamma(2h_R)|}\sinh\left(\frac{\tilde{\omega}_{\rm L}}{2T_{\rm L}}+\frac{\tilde{\omega}_{\rm R}}{2T_{\rm R}}\right)\left|\Gamma\left(h_{\rm L}+i\frac{\tilde{\omega}_{\rm L}}{2\pi T_{\rm L}}\right)\right|^2\left|\Gamma\left(h_{\rm R}+i\frac{\tilde{\omega}_{\rm R}}{2\pi T_{\rm R}}\right)\right|^2, \tag{21}$$

where $T_{\rm L},T_{\rm R}$ are the temperatures, $h_{\rm L},h_{\rm R}$ are the conformal dimensions of the dual operator, $\tilde{\omega}_{\rm L}=\omega_{\rm L}-q_{\rm L}\Phi_{\rm L}$ and $\tilde{\omega}_{\rm R}=\omega_{\rm R}-q_{\rm R}\Phi_{\rm R}$ are the total excited energy in which $(q_{\rm L},q_{\rm R})$ and $(\Phi_{\rm L},\Phi_{\rm R})$ are respectively the charges and chemical potentials (both including the electric and the magnetic contributions for the dyonic KN black hole case) of the dual left and right-hand operators. The complex conformal dimensions $(h_{\rm L},h_{\rm R})$ of the dual operator be read out from the asymptotic expansion of the bulk dyonic charged scalar field at the AdS boundary[22]

$$h_{\rm L}=h_{\rm R}=\frac{1}{2}\pm i\mu. \tag{22}$$

For dyonic KN black holes, there are in general three different pictures, namely $J$-, $Q$- and the $P$-pictures, in the dual CFTs descriptions. Here, we only show the result of $J$-picture (for the other two pictures, see Ref. 22).

In the $J$-picture, the left- and right-hand central charges of the dual CFT are determined by the angular momentum[28,31]

$$c_{\rm L}^J=c_{\rm R}^J=12J, \tag{23}$$

and the associated left- and right-hand temperatures for the near-extremal dyonic KN black hole are

$$T_{\rm L}^J=\frac{r_0^2+a^2}{4\pi a r_0},\qquad T_{\rm R}^J=\frac{B}{2\pi a}. \tag{24}$$

The CFT microscopic entropy from the Cardy formula

$$S_{\rm CFT}=\frac{\pi^2}{3}(c_{\rm L}^JT_{\rm L}^J+c_{\rm R}^JT_{\rm R}^J)=\pi(r_0^2+a^2+2r_0B), \tag{25}$$

agrees with the macroscopic entropy (5) of the near-extremal KN black hole.

Besides, by matching the first law of black hole thermodynamics with that of the dual CFT, i.e., $\delta S_{\rm BH}=\delta S_{\rm CFT}$, the following relation holds

$$\frac{\delta M-\Omega_{\rm H}\delta J-\Phi_{\rm H}\delta Q-\bar{\Phi}_{\rm H}\delta P}{T_{\rm H}}=\frac{\tilde{\omega}_{\rm L}}{T_{\rm L}}+\frac{\tilde{\omega}_{\rm R}}{T_{\rm R}}, \tag{26}$$

where the angular velocity and chemical potentials at $\rho=B$ are given in Eq. (5). To probe the rotation we need to turn off the charges of the probe scalar field and set $T_{\rm L}=T_{\rm L}^J$ and $T_{\rm R}=T_{\rm R}^J$, then for the dyonic KN black hole $\delta M=\omega,\ \delta J=n,\ \delta Q=0,\ \delta P=0$. Thus, we have

$$\tilde{\omega}_{\rm L}^J=n\quad\text{and}\quad\tilde{\omega}_{\rm R}^J=\frac{\omega}{a}\quad\Rightarrow\quad\frac{\tilde{\omega}_{\rm L}^J}{2T_{\rm L}^J}=-\pi\kappa\quad\text{and}\quad\frac{\tilde{\omega}_{\rm R}^J}{2T_{\rm R}^J}=\pi\tilde{\kappa}, \tag{27}$$

where $q,p$ are set to zero. Consequently, the agreement between the absorption cross section ratio (14) of the scalar field (with $q=p=0$) in the near-extremal dyonic KN black hole and that of its dual scalar operator in Eq. (21) is confirmed in the $J$-picture.

## 4. Conclusion

We have reviewed and studied the emission of charges from near-extremal dyonic KN black holes within the framework of the quantum field in the near-horizon geometry.[22] The symmetry of the near-horizon geometry of near-extremal black holes leads to analytic expressions for the solutions and therefrom the in-vacuum and the out-vacuum and pair production rate. This makes a drastic difference from the conventional wisdom which computes the tunneling probability of virtual particles from the Dirac sea in charged black holes.[6–16] In fact, the phase-integral formulation or the Hamilton-Jacobi equation for spinless charges has simple poles near the horizon, whose residues determine the leading term of pair production. On the other hand, the quantum field theory in the near-horizon geometry exhibits a rich structure of the spectral distribution of produced particles. A caveat, however, is that the exact solutions in the near-horizon are not the solutions in the original global spacetime of black holes but carry the essential information of pair production that occurs the near horizon and that the back-reaction due to produced pairs and the induced current is not considered.

The near-horizon region contains a causal horizon and the electric field effect near the horizon dominantly gives rise to both the Hawking radiation and the Schwinger mechanism. By imposing the proper boundary condition on the exact solutions, the pair production, vacuum persistence amplitude and absorption cross section can be obtained from the relative ratios of the fluxes on the asymptotic boundary and horizon. This approach provides one with a systematic method for the pair production and vacuum polarization in a proper-time integral representation.[17,23] The thermal interpretation of the pair production rate for the near-extremal dyonic KN black holes consists of the Schwinger effect in the $AdS_2$ space mainly due to the electromagnetic field of black holes and the Schwinger effect in the Rindler space due to the Hawking temperature. The dual CFTs descriptions of the Schwinger pair production of the near-extremal KN black hole[21] can be generalized into the threefold dual CFTs pictures for the dyonic KN black hole which includes an additional magnetic charge. The third $P$-picture is associated with the dual gauge potential, a new "magnetic hair" of the dyonic KN black hole and a $U(1)$ fiber on the base manifold. Based on the threefold dyonic KN/CFTs duality, the dual CFTs descriptions of the absorption cross section ratios and the pair production rate of the dyonic charged scalar field can be confirmed in the $J$-, $Q$-, and $P$-pictures, respectively.

The near-extremal dyonic KN black hole is the most general model in the sense that the zero magnetic charge limit corresponds to the KN black hole and the zero angular momentum (nonrotating) limit recovers the RN black hole with both electric and magnetic charge. Further restricting to zero angular momentum and magnetic charge reduces to the RN black hole. The formulae in this paper have such limits by replacing the separation parameter $\lambda_l$ by $l(l+1)$ for spherical harmonics for nonrotating black holes. The magnetic charge has the Dirac monopole for the

black hole, and the quantization condition $ep = 2\pi$ of electric charges and the fine structure constant $e^2 = \alpha$ leads to $e/p = \alpha/2\pi$, which implies the magnetic charge is larger by order of three than the electric charge. Then, the monopole production from a magnetic black hole is suppressed since the Unruh temperature is proportional to $p/P$ compared to $e/Q$ of an electric black hole. The physics related to production of electric charges and magnetic monopoles from dyonic black holes will be addressed in a future publication.

## Acknowledgments

The authors would like to thank Gary Gibbons and Remo Riffini for pointing out that Schwinger mechanism was studied independently before Hawking radiation was discovered and also to express thanks to Alexei Gaina for informing us many relevant works related to this review article. The work of C.M.C. was were supported by the Ministry of Science and Technology of the R.O.C. under the grant MOST 107-2119-M-008-013. The work of S.P.K. was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2015R1D1A1A01060626). The work of J.R.S. was supported by the NSFC under Grant No. 11675272 and the Fundamental Research Funds for the Central Universities.

## References

1. S. W. Hawking, *Nature* **248**, 30 (1974).
2. S. W. Hawking, *Commun. Math. Phys.* **43**, 199 (1975) Erratum: [*Commun. Math. Phys.* **46**, 206 (1976)].
3. F. Sauter, *Z. Phys.* **69**, 742 (1931).
4. F. Sauter, *Z. Phys.* **73**, 547 (1932).
5. J. S. Schwinger, *Phys. Rev.* **82**, 664 (1951).
6. W. T. Zaumen, *Nature* **247**, 531 (1974).
7. B. Carter, *Phys. Rev. Lett.* **33**, 558 (1974).
8. T. Damour and R. Ruffini, *Phys. Rev. Lett.* **35**, 463 (1975).
9. G. W. Gibbons, *Commun. Math. Phys.* **44**, 245 (1975).
10. T. Damour and R. Ruffini, *Phys. Rev. D* **14**, 332 (1976).
11. I. M. Ternov, A. B. Gaina and G. A. Chizhov, *Sov. J. Nucl. Phys.* **44**, 343 (1986) [*Yad. Fiz.* **44**, 533 (1986)].
12. I. B. Khriplovich, *J. Exp. Theor. Phys.* **88**, 845 (1999) [*Zh. Eksp. Teor. Fiz.* **115**, 1539 (1999)] [gr-qc/9812060].
13. I. B. Khriplovich, *Phys. Rept.* **320**, 37 (1999).
14. S. P. Kim and D. N. Page, *Nuovo Cim. B* **120**, 1193 (2005) [gr-qc/0401057].
15. I. B. Khriplovich, *Phys. Atom. Nucl.* **65**, 1259 (2002) [*Yad. Fiz.* **65**, 1292 (2002)].
16. R. Ruffini, G. Vereshchagin and S. S. Xue, *Phys. Rept.* **487**, 1 (2010) [arXiv:0910.0974 [astro-ph.HE]].

17. S. P. Kim, *Int. J. Mod. Phys. D* **25**, 1645005 (2016) [arXiv:1602.05336 [hep-th]].
18. M. K. Parikh and F. Wilczek, *Phys. Rev. Lett.* **85**, 5042 (2000) [hep-th/9907001].
19. C.-M. Chen, S. P. Kim, I.-C. Lin, J.-R. Sun and M.-F. Wu, *Phys. Rev. D* **85**, 124041 (2012) [arXiv:1202.3224 [hep-th]].
20. C.-M. Chen, J.-R. Sun, F.-Y. Tang and P.-Y. Tsai, *Class. Quant. Grav.* **32**, 195003 (2015) [arXiv:1412.6876 [hep-th]].
21. C.-M. Chen, S. P. Kim, J.-R. Sun and F.-Y. Tang, *Phys. Rev. D* **95**, 044043 (2017) [arXiv:1607.02610 [hep-th]].
22. C.-M. Chen, S. P. Kim, J.-R. Sun and F.-Y. Tang, *Phys. Lett. B* **781**, 129 (2018) [arXiv:1705.10629 [hep-th]].
23. R. G. Cai and S. P. Kim, *JHEP* **1409**, 072 (2014) [arXiv:1407.4569 [hep-th]].
24. S. P. Kim, H. K. Lee and Y. Yoon, arXiv:1503.00218 [hep-th].
25. S. P. Kim, *Int. J. Mod. Phys. A* **30**, 1545017 (2015) [arXiv:1506.03990 [hep-th]].
26. S. P. Kim, doi:10.1142/9789814759816_0011 [arXiv:1509.05532 [hep-th]].
27. C.-M. Chen and J.-R. Sun, *J. Phys. Conf. Ser.* **330**, 012009 (2011) [arXiv:1106.4407 [hep-th]].
28. C.-M. Chen and J.-R. Sun, *Int. J. Mod. Phys. Conf. Ser.* **07**, 227 (2012) [arXiv:1201.4040 [hep-th]].
29. S. P. Kim and D. N. Page, *Phys. Rev. D* **73**, 065020 (2006) [hep-th/0301132].
30. Cl. Gabriel and Ph. Spindel, *Annals Phys.* **284**, 263 (2000) [gr-qc/9912016].
31. C.-M. Chen, Y.-M. Huang, J.-R. Sun, M.-F. Wu and S.-J. Zou, *Phys. Rev. D* **82**, 066004 (2010) [arXiv:1006.4097 [hep-th]].

# Pair creation in electric fields, renormalization, and backreaction

Antonio Ferreiro*, Jose Navarro-Salas**, and Silvia Pla†

*Departamento de Fisica Teorica and IFIC, Universidad de Valencia-CSIC, Facultad de Fisica, Burjassot, 46100, Spain*

* *E-mail: antonio.ferreiro@ific.uv.es*

** *E-mail: jnavarro@ific.uv.es*

† *E-mail: silvia.pla@uv.es*

We consider pair production phenomena in spatially homogeneous strong electric fields. We focus on spinor QED in two-dimensions and discuss the potential ambiguity in the adiabatic order assignment for the electromagnetic potential required to fix the renormalization subtractions. This ambiguity can be univocally fixed by imposing, at the semiclassical level, stress-energy conservation when both electric and gravitational backgrounds are present.

*Keywords*: Schwinger effect; pair creation; adiabatic renormalization; semiclassical Maxwell equations.

## 1. Introduction

A time-dependent gravitational field yields the creation operators of quantum fields to evolve into a superposition of creation and annihilation operators. This produces the spontaneous creation of particle-antiparticle pairs out of the vacuum. This effect was first discovered, including the precise probability distribution of the produced particles, in the physical context of an expanding universe [1]. A similar superposition of creation and annihilation operators takes place if the quantized field is coupled to a time-varying gauge field background [2]. In this contribution we want to focus on the particle creation phenomena induced by electric fields in presence of gravity, which are of major relevance in astrophysics and cosmology [3].

A fundamental problem in the physical understanding of gravitational particle creation processes is the calculation of the expectation values $\langle T_{\mu\nu} \rangle$. The computations require methods of regularization and renormalization to deal with new ultraviolet divergences (UV) not present in Minkowski space. Equivalently, in a time-dependent electric field the fundamental problem is to evaluate the renormalized electric current $\langle j_\mu \rangle$, which acts as the proper source of the semiclassical Maxwell equations.

In this work we discuss an improvement of the adiabatic regularization method, originally introduced in cosmological scenarios and for quantized scalar fields [4], to include homogeneous electric fields and Dirac quantized fields. We will reexamine the method to consistently deal with both electric and gravitational fields. By doing this we will fix an inherent ambiguity of the method. The adiabatic order assignment of the vector potential has been traditionally assumed in the literature of order zero. We will argue that the correct adiabatic order assignment is one, instead of zero, at least if a gravitational field is present. This problem has been discussed

for scalar fields in Refs. 5, 6. Here we extend the discussion to Dirac fields. To focus on the main ideas we will restrict the analysis to two-dimensional spacetime.

## 2. Spinor QED$_2$ and the adiabatic renormalization scheme

We consider two-dimensional spinor QED in an expanding spacetime described by the metric $ds^2 = dt^2 - a^2(t)dx^2$. The classical action is given by

$$\mathcal{S} = \int dx^2 \sqrt{-g} \left( -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + i\bar{\psi} \not{D} \psi - m\bar{\psi}\psi \right) , \tag{1}$$

and the corresponding Dirac equation reads

$$(i\underline{\gamma}^{\mu} D_{\mu} - m)\psi = 0 , \tag{2}$$

where $D_\mu \equiv \partial_\mu - \Gamma_\mu - iqA_\mu$ and $\Gamma_\mu$ is the spin connection. $\underline{\gamma}^{\mu}(x)$ are the spacetime-dependent Dirac matrices satisfying the anticommutation relations $\{\underline{\gamma}^{\mu}, \underline{\gamma}^{\nu}\} = 2g^{\mu\nu}$. These gamma matrices are related with the Minkowskian ones by $\underline{\gamma}^{0}(t) = \gamma^0$ and $\underline{\gamma}^{1}(t) = \gamma^1/a(t)$, and the components of the spin connections are $\Gamma_0 = 0$ and $\Gamma_1 = (\dot{a}/2)\gamma_0\gamma_1$. Therefore, $\underline{\gamma}^{\mu}\Gamma_\mu = -\frac{\dot{a}}{2a}\gamma_0$. It is very convenient to fix the gauge for the vector potential as $A_\mu = (0, -A(t))$. The Dirac equation (2) becomes[5,7]

$$\left( i\gamma^0\partial_0 + \frac{i}{2}\frac{\dot{a}}{a}\gamma^0 + \left( \frac{i}{a}\partial_1 + \frac{qA_1}{a} \right) \gamma^1 - m \right) \psi = 0. \tag{3}$$

From now on we will use the Weyl representation (with $\gamma^5 \equiv \gamma^0\gamma^1$)

$$\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \gamma^1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad \gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} .$$

We expand the quantized field in momentum modes

$$\psi(t,x) = \int_{-\infty}^{\infty} dk \left[ B_k u_k(t,x) + D_k^{\dagger} v_k(t,x) \right] , \tag{4}$$

where the two independent spinor solutions are

$$u_k(t,x) = \frac{e^{ikx}}{\sqrt{2\pi a}} \begin{pmatrix} h_k^I(t) \\ -h_k^{II}(t) \end{pmatrix}, \qquad v_k(t,x) = \frac{e^{-ikx}}{\sqrt{2\pi a}} \begin{pmatrix} h_{-k}^{II*}(t) \\ h_{-k}^{I*}(t) \end{pmatrix} , \tag{5}$$

and $B_k$ and $D_k$ the creation and annihilation operators which fulfill the usual anti-commutation relations. The field equations (3) are now converted into

$$\begin{aligned} \dot{h}_k^I - \frac{i}{a}\left(k + qA\right) h_k^I - imh_k^{II} = 0 \\ \dot{h}_k^{II} + \frac{i}{a}\left(k + qA\right) h_k^{II} - imh_k^{I} = 0 . \end{aligned} \tag{6}$$

We also assume the normalization condition $|h_k^I|^2 + |h_k^{II}|^2 = 1$. The classical electric current is given by $j^\nu = -q\bar{\psi}\underline{\gamma}^{\nu}\psi$, and the formal expression for the vacuum expectation value of $\langle j^x \rangle$ is

$$\langle j^x \rangle \ = \frac{q}{2\pi a^2} \int_{-\infty}^{\infty} dk \left( \left| h_k^{II} \right|^2 - \left| h_k^{I} \right|^2 \right) . \tag{7}$$

The above expression possesses, as expected, UV divergences. To obtain the finite, physical values, we have to perform appropriate subtractions

$$\langle j^x \rangle_{ren} = \frac{q}{2\pi a^2} \int dk \left( |h_k^{II}|^2 - |h_k^I|^2 - SUBSTRACTIONS \right) . \tag{8}$$

The semiclassical Maxwell equations for $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ are

$$\nabla_\mu F^{\mu\nu} = \langle j^\nu \rangle_{ren}. \tag{9}$$

Equations (6) and (9) determine the continuous interchange of energy between the electric field and matter, via charged pair production and backreaction. Now, the main problem is to obtain the required subtractions consistently. In this context, the most natural way to determine the renormalization subtractions is the adiabatic regularization method. The basic principles for scalar fields can be borrowed from Ref. 8, while for spin-1/2 fields one can see Refs. 7, 9. The main idea here is to consider an adiabatic expansion of the mode functions. For scalar fields, this expansion is based on the WKB-type ansatz , namely

$$h_k(t) = \frac{1}{\sqrt{\Omega_k(t)}} e^{-i\int^t \Omega_k(t')dt'}, \qquad \Omega_k(t) = \omega_k^{(0)} + \omega_k^{(1)} + \omega_k^{(2)} + \cdots \tag{10}$$

where the order of the expansion is determined by the number of derivatives of the background fields. A different type of expansion is required for fermionic fields[7,9]. A very crucial point to properly define the adiabatic expansion, for both scalar and Dirac fields, is to fix the leading order. The zeroth order adiabatic term is determined by $\omega_k^{(0)}$.

## 3. Two attempts toward the backreaction equations in Minkowski space

Let us first assume for simplicity that we are in Minkowski space and $a(t) = 1$. Therefore, the gauge field is now the only background field. In this scenario it is very natural to define $\omega_k^{(0)} = \sqrt{(k + qA)^2 + m^2}$. This means that the adiabatic order assignment for the gauge field $A(t)$ has been implicitly chosen as 0, as first assumed in Ref. 10 and in all subsequent papers on this topic. Therefore, $\dot{A}(t)$ should be of order 1, etc. The proposed renormalized current is then given by

$$\langle j^x \rangle_{ren}^I = q \int_{-\infty}^{\infty} \frac{dk}{2\pi} \left( |h_k^{II}|^2 - |h_k^I|^2 - \frac{(k + qA)}{\omega_k^{(0)}} \right) \tag{11}$$

Plugging this expression into the semiclassical Maxwell equations (9), we get the following backreaction equation,

$$-\dot{E} = \ddot{A} = q \int_{-\infty}^{\infty} \frac{dk}{2\pi} \left[ |h_k^{II}|^2 - |h_k^I|^2 - \frac{(k + qA)}{\sqrt{(k + qA)^2 + m^2}} \right] , \tag{12}$$

together with the equation for the field modes (6). It is important to stress that the above semiclassical Maxwell equations are compatible with the conservation of the

energy if the renormalized stress-energy tensor is constructed by subtracting up to the zero adiabatic order

$$\langle\rho\rangle^I_{ren} = \langle T_{00}\rangle^I_{ren} = \frac{1}{2\pi}\int_{-\infty}^{\infty} dki\left[h_k^{II}\dot{h}_k^{II*} + h_k^I\dot{h}_k^{I*}\right] + \omega_k^{(0)} \ . \tag{13}$$

It is easy to check that $\partial_\mu\langle T^{\mu\nu}\rangle_{ren} + \partial_\mu T^{\mu\nu}_{elec} = 0$, where $T^{elec}_{\mu\nu} = \frac{1}{2}E^2\eta_{\mu\nu}$.

Alternatively, we can define

$$\omega_k^{(0)} = \sqrt{k^2+m^2} \equiv \omega, \tag{14}$$

which assumes that $A$ is of adiabatic order 1, and proceed according to the rules of the adiabatic expansion. With this choice we obtain

$$\langle j^x\rangle^{II}_{ren} = q\int_{-\infty}^{\infty}\frac{dk}{2\pi}\left(|h_k^{II}|^2 - |h_k^I|^2 - \frac{k}{\omega} - \frac{qm^2}{\omega^3}A\right) \ . \tag{15}$$

Hence, the semiclassical Maxwell equation reads $-\dot{E} = \ddot{A} = \langle j^x\rangle^{II}_{ren}$, and the mode functions $h_k^I$ and $h_k^{II}$ satisfy again Eq. (6). The choice (14) is also compatible with the conservation of energy, defined now by subtracting up to second adiabatic order

$$\langle T_{00}\rangle^{II}_{ren} = \frac{1}{2\pi}\int dki\left[h_k^{II}\dot{h}_k^{II*} + h_k^I\dot{h}_k^{I*}\right] + \omega + \frac{kqA}{\omega} + \frac{m^2q^2A^2}{2\omega^3} \ . \tag{16}$$

It is immediate to see that the fermionic currents (11) and (15) are equivalent,

$$\triangle\langle j^x\rangle_{ren} = q\int_{-\infty}^{-\infty}\frac{dk}{2\pi}\left[\frac{k}{\omega} - \frac{(k+qA)}{\sqrt{(k+qA)^2+m^2}} + \frac{m^2qA}{\omega^3}\right] = 0 \ . \tag{17}$$

The first two terms of the equation above correspond to linearly divergent integrals, differing by a constant shift, and hence, their difference is finite $-q^2A(t)/\pi$. On the other hand, the last term is a finite integral, which cancels out with the previous quantity. In conclusion, in Minkowski space the choice of the adiabatic order of the background field $A(t)$ does not affect the physical observables and, therefore, both options are equally valid. However, this conclusion is misleading as soon as one introduces the gravitation field.

## 4. The role of gravity

Now, let us assume that our space-time is a two-dimensional expanding universe with metric $ds^2 = dt^2 - a^2(t)dx^2$. In this case, the prescription for $a(t)$ is to fix it at adiabatic order 0, and the rules for the adiabatic subtraction terms are univocally fixed according to the scaling dimension of the relevant operators. The stress-energy tensor must be renormalized at second adiabatic order[11], while the electric current should be renormalized at adiabatic order 1. With this restrictions, the only possibility is to choose $A(t)$ of adiabatic order 1, in the same footing as $\dot{a}(t)$. That is, we should have a hierarchy between the two background fields. The leading order corresponds to gravity, and consequently one should replace the definition (14) by

$$\omega_k^{(0)} = \sqrt{k^2/a^2+m^2} \equiv \omega \ , \tag{18}$$

instead of the naive generalization $\omega_k^{(0)} = \sqrt{(k+qA)^2/a^2+m^2}$ . This point has been overlooked in the literature, as recently stressed in Ref. 5. The gauge field should enter at the next to leading order in the adiabatic expansion.

Therefore, the renormalized expression for the electric current should be

$$\langle j^x\rangle_{ren} = \frac{q}{2\pi a^2}\int_{-\infty}^{\infty} dk\left(|h_k^{II}|^2 - |h_k^I|^2 - \frac{k}{a\omega} - \frac{qm^2}{a\omega^3}A\right) , \tag{19}$$

and the backreaction equations are

$$-\dot{E} = \frac{\ddot{A}}{a} - \frac{\dot{A}}{a}\frac{\dot{a}}{a} = \frac{q}{2\pi a}\int_{-\infty}^{\infty} dk\left(|h_k^{II}|^2 - |h_k^I|^2 - \frac{k}{a\omega} - \frac{qm^2}{a\omega^3}A\right) \quad , \tag{20}$$

together with the equation for the field modes (6).

As in Minkowski space, one should expect stress-energy conservation, namely $\nabla_\mu\langle T^{\mu\nu}\rangle_{ren} + \nabla_\mu T^{\mu\nu}_{elec} = 0$, with $T^{elec}_{\mu\nu} = \frac{1}{2}E^2 g_{\mu\nu}$. After the required adiabatic subtractions, one obtain the following expressions for the diagonal components of $\langle T_{\mu\nu}\rangle$:

$$\langle T_{00}\rangle_{ren} = \frac{1}{2\pi a}\int_{-\infty}^{\infty} dki\left[h_k^{II}\dot{h}_k^{II*} + h_k^I\dot{h}_k^{I*}\right] + \omega + \frac{kqA}{\omega} + \frac{m^2q^2A^2}{2a^2\omega^3} - \frac{k^2m^2\dot{a}^2}{8a^4\omega^5}, \tag{21}$$

$$\begin{aligned}\langle T_{11}\rangle_{ren} = \ & \frac{1}{2\pi}\int_{-\infty}^{\infty} dk(k+qA(t))\left(|h_k^I|^2 - |h_k^{II}|^2\right) + \frac{k^2}{a\omega} + \frac{km^2qA}{a\omega^3} + \frac{kqA}{a\omega} - \frac{m^4\ddot{a}}{4\omega^5} \\ & + \frac{m^2\ddot{a}}{4\omega^3} + \frac{5m^6\dot{a}^2}{8a\omega^7} - \frac{3m^4\dot{a}^2}{4a\omega^5} + \frac{m^2\dot{a}^2}{8a\omega^3} + \frac{3m^4q^2A^2}{2a\omega^5} - \frac{m^2q^2A^2}{2a\omega^3} \ .\end{aligned} \tag{22}$$

The conservation equation for the zeroth component can be decomposed as

$$\nabla_\mu\langle T^{\mu 0}\rangle_{ren} + \nabla_\mu T^{\mu 0}_{elec} = \partial_0\langle T_{00}\rangle_{ren} + \frac{\dot{a}}{a}\langle T_{00}\rangle_{ren} + \frac{\dot{a}}{a^3}\langle T_{11}\rangle_{ren} + \partial_0 T^{elec}_{00} = 0. \tag{23}$$

Plugging (21) and (22) into (23) and using the equations (6) we get

$$\nabla_\mu\langle T^{\mu 0}\rangle_{ren} + \nabla_\mu T^{\mu 0}_{elec} = \frac{\dot{A}}{a}\left(\frac{\ddot{A}}{a} - \frac{\dot{A}\dot{a}}{a^2} - \langle j^x\rangle_{ren}\right) = 0 \ . \tag{24}$$

Note that the factor in parentheses is precisely the semiclassical Maxwell equation for the electric field (20). A similar result is trivially obtained for the remaining component: $\nabla_\mu\langle T^{\mu 1}\rangle_{ren} + \nabla_\mu T^{\mu 1}_{elec} = 0$. Therefore, the semiclassical Maxwell equations must be satisfied to ensure stress-energy conservation.

If one assumes that $A(t)$ is of adiabatic order 0, the stress-energy conservation does not hold anymore. With this alternative adiabatic assignment, and renormalizing $\langle T^{\mu\nu}\rangle$ up to and including the second adiabatic order, one obtains

$$\nabla_\mu\langle T^{\mu 0}\rangle_{ren} + \nabla_\mu T^{\mu 0}_{elec} \neq 0 \ . \tag{25}$$

In this case, the left-hand side in the above equation is proportional to $E\langle j^x\rangle^{(2)}$, where $\langle j^x\rangle^{(2)}$ is the second adiabatic order of the electric current, which cannot be properly absorbed into the renormalization substractions of the electric current. We note that another inconsistency is the disagreement with the trace anomaly [5] in the massless limit.

## 5. Conclusions

The role of gravity is essential to fix the ambiguity in the adiabatic order assignment of the electromagnetic potential $A(t)$, and therefore to obtain the correct renormalization subtractions of physical observables. Although in absence of a gravitational background the two different prescriptions for the adiabatic subtraction terms are equivalent, when gravity is incorporated into the game only the adiabatic order assignment 1 for the gauge field $A(t)$ is compatible with stress-energy conservation.

## References


1. L. Parker, *Phys. Rev. Lett.* **21**, 562 (1968); *Phys. Rev. D* **183**, 1057 (1969).
2. E. Brezin and C. Itzykson, *Phys. Rev. D* **2**, 1191 (1970).
3. R. Ruffini, G. Vereshchagin and S. S. Xue, *Phys. Rept.* **487**, 1 (2010). C. Stahl, E. Strobel and S.-S. Xue, *Phys. Rev. D* **93**, 025004 (2016).
4. L. Parker and S. A. Fulling, *Phys. Rev. D* **9**, 341 (1974). S. A. Fulling and L. Parker, *Ann. Phys.* (N.Y.) **87**, 176 (1974). S. A. Fulling, L. Parker and B. L. Hu, *Phys. Rev. D* **10**, 3905 (1974). T. S. Bunch, *J. Phys. A* **13**, 1297 (1980). P. R. Anderson and L. Parker, *Phys. Rev. D* **36**, 2963 (1987). I. Agullo, J. Navarro-Salas, G. J. Olmo and L. Parker, *Phys. Rev. Lett.* **103**, 061301 (2009). *Phys. Rev. D* **81**, 043514, (2010).
5. A. Ferreiro and J. Navarro-Salas, *Phys. Rev. D* **97**, 125012 (2018).
6. A. Ferreiro, J. Navarro-Salas and S. Pla, *Phys. Rev. D***98**, 045015 (2018).
7. F. J. Barbero G., A. Ferreiro, J. Navarro-Salas and E. J. S. Villaseñor, *Phys. Rev. D***98**, 98, 025016 (2018).
8. L. Parker and D. J. Toms, *Quantum Field Theory in Curved Spacetime: Quantized Fields and Gravity*, CUP, Cambridge, UK (2009).
9. A. Landete, J. Navarro-Salas and F. Torrenti, *Phys. Rev. D* **88**, 061501 (2013); *Phys. Rev. D* **89** 044030 (2014). A. del Rio, J. Navarro-Salas and F. Torrenti, *Phys. Rev. D* **90**, 084017 (2014). S. Ghosh, *Phys. Rev. D* **91**, 124075 (2015); *Phys. Rev. D* **93**, 044032 (2016). A. del Rio and J. Navarro-Salas, *Phys. Rev. D* **91**, 064031 (2015). A. del Rio, A. Ferreiro, J. Navarro-Salas and F. Torrenti, *Phys. Rev. D* **95**, 105003 (2017).
10. F. Cooper and E. Mottola, *Phys. Rev. D* **40**, 456 (1989). Y. Kluger et al. *Phys. Rev. Lett.* **67**, 2427 (1991).
11. N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space*, CUP, Cambridge, UK (1982).

# The gravitational field of a laser beam

Fabienne Schneiter

*Eberhard-Karls-Universität Tübingen, Institut für Theoretische Physik, 72076 Tübingen, Germany*

Dennis Rätzel

*Institut für Physik, Humboldt-Universität zu Berlin, Newtonstrae 15, 12489 Berlin, Germany, dennis.raetzel@physik.hu-berlin.de*

Daniel Braun

*Eberhard-Karls-Universität Tübingen, Institut für Theoretische Physik, 72076 Tübingen, Germany*

In this proceedings article, we review the results presented in [Fabienne Schneiter et al. 2018 Class. Quantum Grav. 35 195007] on the gravitational field of light in a laser beam, modeled as a solution to Maxwell's equations perturbatively expanded in the beam divergence. Using this approach, wave properties of light, such as diffraction, are taken into account that have been neglected in earlier studies. Interesting features of the gravitational field of laser beams become apparent like frame-dragging due to the intrinsic angular momentum of light and the deflection of parallel co-propagating test beams for short distances to the source beam.

*Keywords*: Gravitational properties of light, laser beams, linearized gravity.

## 1. Introduction

The gravitational field of a light beam has first been studied in Ref. 15 by Tolman, Ehrenfest and Podolski in 1931, who described the light beam as a one-dimensional "pencil of light". Later in Ref. 3 by Bonnor, a description for the gravitational field of a cylindrical beam of light of a finite radius was presented. Light was modeled as a continuous fluid moving at the speed of light. A central feature of the models of Ref. 15 and Ref. 3 is the lack of diffraction; the beams do not diverge. This corresponds to the short wavelength limit where all wavelike properties of light are neglected. Further studies to the gravitational field of light that share this feature include the investigation of two co-directed parallel cylindrical light beams of finite radius[2,10], spinning non-divergent light beams[9], non-divergent light beams in the framework of gravito-electrodynamics[5] and the gravitational field of a point like particle moving with the speed of light[1,17].

In contrast, the wavelike properties of light were taken into account in Ref. 16, where the gravitational field of a plane electromagnetic wave was investigated. An approach to take finite wavelengths into account for the case of a laser pulse was given in Refs. 11, 12, where, however, diffraction was neglected. Here, we describe the laser beam as a solution to Maxwell's equations. This is done perturbatively by an expansion in the beam divergence, which is considered to be small. The zeroth order of the expansion corresponds to the paraxial approximation and coincides

with the result of Ref. 3. In the first order in the beam divergence, frame-dragging due to the internal angular momentum of circularly polarized beams occurs. In the forth order, a parallel co-propagating test beam of light is found to be deflected by the gravitational field of the laser beam.

## 2. The model

In the following, we will employ dimensionless coordinates by dividing the Cartesian coordinates corresponding to the lab reference by the beam waist $w_0$ as $\tau = ct/w_0$, $\xi = x/w_0$, $\chi = y/w_0$ and $\zeta = z/w_0$, where $c$ is the speed of light. A laser beam in its simplest mode is accurately described by a Gaussian beam. The Gaussian beam is an almost monochromatic electromagnetic plane wave which has the property that its intensity distribution decays with a Gaussian factor with the distance to the beamline. It is obtained as a perturbative solution of Maxwell's equations, an expansion in the beam divergence $\theta$, the opening angle of the beam, which is assumed to be small. When the beam divergence is small, the beam may be thought of as a bunch of almost, but not exactly, parallel propagating rays of light. The electromagnetic four-vector potential describing the Gaussian beam is obtained by a plane wave multiplied by an envelope function, which is assumed to be varying slowly in the direction of propagation, in agreement with the property that the divergence of the beam is small. Corresponding to these features, we make the ansatz for the four-vector potential $A_\alpha = \mathcal{A} v_\alpha(\xi, \chi, \theta\zeta) \exp(2i(\zeta - \tau)/\theta)$, were $\mathcal{A}$ is the amplitude, $v_\alpha$ the envelope function, and the exponential factor describes a plane wave propagating in $\zeta$-direction with angular wave number $k = 2/\theta$, where $w_0$ is the beam waist at its focal point, a measure of the radius of the beam (see Fig. 1).

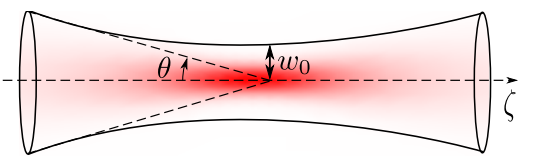



Fig. 1. Schematic illustration of the Gaussian beam, the beam waist $w_0$ and the beam divergence $\theta$. More specifically, the figure illustrates the scalar envelope function $v_0$ of the vector potential of the Gaussian beam in a plane that contains the optical axis (represented by the dashed horizontal line). Due to the rotational symmetry of the envelope function around the optical axis, the vertical axis can be any direction transversal to the optical axis. The thick curved lines mark the distance $w(\zeta) = w_0\sqrt{1 + (\theta\zeta)^2}$ from the optical axis at which the absolute value of the envelope function reaches $1/e$ times its maximum.

In the following, the beam waist is kept constant. In this case, since the beam divergence is small, the angular wave number is large. As for any beam of radiation in flat spacetime, the four-vector potential satisfies the wave equations

$\left(\partial_\xi^2+\partial_\chi^2+\partial_\zeta^2-\partial_\tau^2\right)A_\mu = 0$, which follow from Maxwell's equations when the Lorenz gauge is chosen. The wave equations for the four-vector potential reduce to a Helmholtz equation for the envelope function, $\left(\partial_\xi^2+\partial_\chi^2+\theta^2\partial_{\theta\zeta}^2+4i\partial_{\theta\zeta}\right)v_\alpha = 0$. We assume that the envelope function is slowly varying in the direction of propagation, i.e. $v_\alpha$ is a function of $\theta\zeta$. Then, the Helmholtz equation can be solved iteratively by writing the envelope function as a power series in the divergence angle $\theta$ (see Ref. 4). For each term in the expansion of the envelope function, one obtains a Helmholtz equation with a source term. These source terms are proportional to terms of lower order in the expansion of the envelope function, and even and odd orders do not mix. We start from the zeroth order solution, which we assume to have a Gaussian profile in the focus plane, and derive the vector potential up to fourth order in the divergence angle.

To obtain the gravitational field of the laser beam, the energy momentum tensor $T_{\alpha\beta}$ has to be calculated from the field strength $F_{\alpha\beta} = \partial_\alpha A_\beta - \partial_\beta A_\alpha$. We only consider vector potentials with field strength tensors that are eigenfunctions with eigenvalue $\lambda = \pm 1$ of the generator of the duality transformation of the electromagnetic field, which in our case is given by $F_{\alpha\beta} \mapsto -i\epsilon_{\alpha\beta\gamma\delta}F^{\gamma\delta}/2$, where $\epsilon_{\alpha\beta\gamma\delta}$ is the completely anti-symmetric tensor with $\epsilon_{0123} = -1$. In that case the rapidly oscillating contributions of the plane wave factor in the vector potential drop out and the energy momentum tensor assumes the simple form $T_{\alpha\beta} = c^2\epsilon_0 \,\mathrm{Re}\left(F_\alpha{}^\sigma F^*_{\beta\sigma} - \frac{1}{4}\eta_{\alpha\beta}F^{\delta\rho}F^*_{\delta\rho}\right)/2$. We interpret the eigenvalues $\lambda$ of the generator duality transformation as the helicity of the beam since the standard notion of circular polarization is recovered in the zeroth order in $\theta$.

## 3. The gravitational field

Since the energy of a laser beam is small, we may expect its gravitational field to be weak. The spacetime metric describing the gravitational field is thus assumed to consist of the metric for flat spacetime in the rescaled coordinates $\eta = w_0^2\mathrm{diag}(-1,1,1,1)$ plus a small perturbation $h_{\alpha\beta}$, where small means $|h_{\alpha\beta}| \ll w_0^2$ for all $\alpha$ and $\beta$. Terms quadratic in the metric perturbation are neglected; this is the linearized theory of general relativity. When the Lorenz-gauge $\partial^\alpha h_{\alpha\beta} = \frac{1}{2}\partial_\beta h^\alpha{}_\alpha$ is chosen, Einstein's field equations reduce to wave equations for the metric perturbation[8]: $\left(\partial_\xi^2+\partial_\chi^2+\partial_\zeta^2-\partial_\tau^2\right)h_{\alpha\beta} = -\kappa w_0^2 T_{\alpha\beta}$, where $\kappa = 16\pi G/c^4$ and $G$ is Newton's constant. Solving the linearized Einstein equations for the energy momentum tensor of the laser beam with emitter and absorber at general positions can be quite cumbersome. Therefore, we consider two different limiting situations instead; the distance between emitter and absorber being very large in one case and very small in the case.

For a large distance between emitter and absorber, we can neglect the rapid change of the field strength at the emitter and the absorber. Then we can take into account that $T_{\alpha\beta}$ is changing slowly in $\zeta$. In particular, we have $T^\lambda_{\alpha\beta} = T^\lambda_{\alpha\beta}(\xi,\chi,\theta\zeta)$.

Therefore, the metric perturbation can be expanded in orders of the beam divergence angle as $h_{\alpha\beta} = \sum_{n=0}^{\infty} \theta^n h_{\alpha\beta}^{(n)}$. The linearized Einstein equations lead to the differential equations

$$\left(\partial_\xi^2 + \partial_\chi^2\right) h_{\alpha\beta}^{\lambda(0)} = -w_0^2 \kappa\, t_{\alpha\beta}^{\lambda(0)} \,, \tag{1}$$

$$\left(\partial_\xi^2 + \partial_\chi^2\right) h_{\alpha\beta}^{\lambda(1)} = -w_0^2 \kappa\, t_{\alpha\beta}^{\lambda(1)} \,, \tag{2}$$

$$\left(\partial_\xi^2 + \partial_\chi^2\right) h_{\alpha\beta}^{\lambda(n)} = -w_0^2 \kappa\, t_{\alpha\beta}^{\lambda(n)} - \partial_{\theta\zeta}^2 h_{\alpha\beta}^{\lambda(n-2)} \,, \text{ for n} > 1 \,, \tag{3}$$

where the $t_{\alpha\beta}^{\lambda(n)}(\xi, \chi, \theta\zeta)$ are given by $T_{\alpha\beta}^{\lambda}(\xi, \chi, \theta\zeta) = \sum_{n=0}^{\infty} \theta^n t_{\alpha\beta}^{\lambda(n)}(\xi, \chi, \theta\zeta)$. The solutions $h_{\alpha\beta}^{\lambda(n)}$ of Eqs. (1), (2) and (3) can be found by direct calculation as we did in Ref. 13. These solutions have to be constructed such that the components of the Riemann curvature tensor vanish at infinite distance from the beamline. The Riemann curvature tensor governs the spread and the contraction of the trajectories of test particles. This means, if the Riemann tensor vanishes, parallel geodesics stay parallel and there is no physical effect as the only reference for a test particle in linearized gravity can be another test particle. As the energy distribution of the laser beam decays like a Gaussian function with the distance from the beamline, no gravitational effect should remain at infinite spatial distances from the beamline. Therefore, we have to ensure that the Riemann curvature tensor $R^{\mu}{}_{\rho\sigma\alpha}$ vanishes for $\rho \to \infty$. General solutions of Eqs. (1), (2) and (3) can be given in terms of the free space Green's function for the Poisson equation in two dimensions as

$$h_{\alpha\beta}^{\lambda(n)}(\xi, \chi, \theta\zeta) = \frac{1}{4\pi} \int_{-\infty}^{\infty} d\xi' d\chi' \log\left((\xi - \xi')^2 + (\chi - \chi')^2\right) Q_{\alpha\beta}^{\lambda(n)}(\xi', \chi', \theta\zeta') \,, \tag{4}$$

where $Q^{\lambda(n)}$ are the right hand sides of Eqs. (1), (2) and (3), respectively.

In the second situation, where we assume a short distance between emitter and absorber, the rapid change of the field strength at emitter and absorber cannot be neglected. Then, we solve the linearized Einstein equations by making use of the corresponding Green's function.

Iteratively solving the Poisson equations for the terms in the expansion of the metric perturbation up to a given order in the divergence angle, we obtain an expression for the spacetime metric which contains all information about the gravitational field of the laser beam to the given order. Knowing the spacetime metric allows us to study the motion of test particles in the gravitational field of the laser beam; their world lines $\gamma^\mu$ satisfy the geodesic equations which depend on the Christoffel symbols, and the distance between two initially parallel geodesics is governed by the geodesic deviation equation which is given in terms of the Riemann curvature tensor. Both Christoffel symbols and curvature tensor are expanded in orders of the divergence angle $\theta$.

## 4. Conclusions

In zeroth order in $\theta$, all non-vanishing components of the metric perturbation are equal up to a sign. Explicitly, we have $h_{\tau\tau}^{(0)} = h_{\zeta\zeta}^{(0)} = -h_{\zeta\tau}^{(0)} = -h_{\tau\zeta}^{(0)} =: I^{(0)}$. For

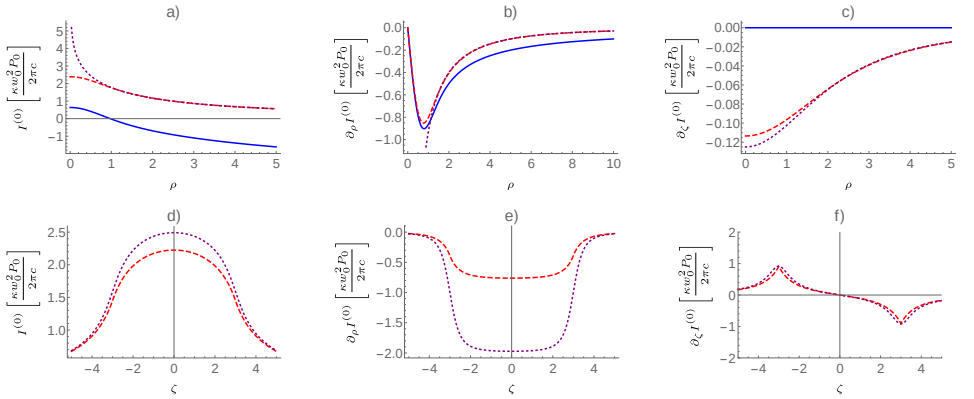



Fig. 2. These plots show the value of the leading order of the metric perturbation $I^{(0)}$ and its first derivatives for the Gaussian beam with infinite distance between emitter and absorber (plain, blue), the Gaussian beam with short distance between emitter and absorber (dashed, red), and the infinitely thin beam (dotted, purple). $\rho$ is defined as $\sqrt{\xi^2+\chi^2}$. In the second and the third case, the distance between emitter and absorber is chosen to be $6w_0$. In the first row, the functions are plotted for $\zeta = 1$ and in the second row for $\rho = 1/2$. The second row does not contain plots for large distances between emitter and absorber as there is no dependence of $I^{(0)}$ on $\zeta$ in that case. We find that the values for $I^{(0)}$ and its first derivatives are usually larger for the infinitely thin beam than for the other two cases. This is due to the divergence at the beamline for the case of the infinitely thin beam. In the other two cases, the gravitational field is spread out as the sources are. In b), we see that the absolute value of the first $\rho$-derivative of $I^{(0)}$ reaches a maximum at a finite distance from the beamline. Note that $\partial_\rho I^{(0)}$ is proportional to the acceleration that a test particle experiences if it is initially at rest at a given distance $\rho$ to the beamline. We see that the acceleration is always directed towards the beamline. It is larger in the case of an infinite distance between emitter and absorber than in the case of a finite distance, which we can attribute to the larger extension of the source (and thus the larger amount of energy) in the former than in the latter. In d), which shows plots for finite distance between emitter and absorber, we see that $\partial_\rho I^{(0)}$ still is the largest at the center between emitter and absorber and decays quickly once their positions at $\zeta = \pm 3$ are passed.

small values of the beam waist and for $\theta = 0$, which corresponds to the paraxial approximation in our case, our solution for the laser beam corresponds to the solution for the infinitely thin beam[15]. If we consider the laser beam to be infinitely long and assume $\theta = 0$, we recover the solution for an infinitely long cylinder[3]. In Fig. 2, the function $I^{(0)}$ and its derivatives are illustrated for the three cases of the infinitely long Gaussian beam, the Gaussian beam with short distance between emitter and absorber, and the infinitely thin beam.

In first order in the divergence angle, we find frame dragging due to spin angular momentum of the circular polarized laser beam. This is similar to the result of Ref. 14 for beams with intrinsic orbital angular momentum. In contrast to frame dragging induced by orbital angular moment, the effect that we find decays like a Gaussian with the distance from the beamline.

The statement of Ref. 15 by Tolman et al. that a non-divergent light beam does not gravitationally deflect a co-directed parallel light beam has been recovered

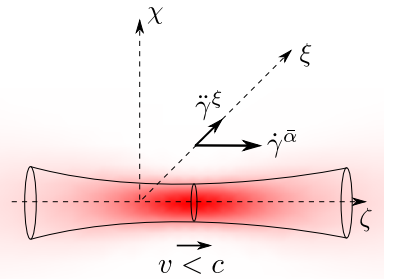



Fig. 3. Schematic illustration of the laser beam and the parallel co-propagating test ray of light: We look at the deflection of the test ray of light due to the gravitational field of the laser beam.

in different contexts: two co-directed parallel cylindrical light beams of finite radius[2,3,10], spinning non-divergent light beams[9], non-divergent light beams in the framework of gravito-electrodynamics[5] and parallel co-propagating light-like test particles in the gravitational field of a one-dimensional light pulse[11]. In fourth order in the divergence angle, we find a deflection of parallel co-propagating test beams. This shows that the result of Ref. 15 and Ref. 3 only holds up to the third order in the divergence angle. This could have been expected from the fact that the group velocity of light in a Gaussian beam along the beamline is not the speed of light[6,7]. However, the deflection of parallel co-propagating light beams by light in a focused laser beam decays like a Gaussian with the distance from the beamline. This means that the effect does not persist outside of the distribution of energy given by the laser beam.

In Ref. 13, we compare the result to the deflection that one obtains from a rod of matter boosted to a speed close to the speed of light. We conclude that focused light does not simply behave like massive matter moving with the reduced velocity identified in Refs. 11, 14. We argue that this difference is due to the divergence of the laser beam along the beamline which leads to additional non-zero components of the metric perturbation which do not appear in the case of the boosted rod. These additional contributions cancel the effect of the reduced propagation speed of light in the focused beam for large distances from the beamline.

## References

1. P. C. Aichelburg and R. U. Sexl. On the gravitational field of a massless particle. *General Relativity and Gravitation*, 2(4):303–312, 1971.
2. A. Banerjee. Cylindrically symmetric stationary beam of electromagnetic radiation. *Journal of Mathematical Physics*, 16:1188, 1975.
3. W. B. Bonnor. The gravitational field of light. *Communications in Mathematical Physics*, 13(3):163–174, 1969.
4. L. W. Davis. Theory of electromagnetic beams. *Phys. Rev. A*, 19:1177–1179, Mar 1979.

5. Valerio Faraoni and RM Dumse. The gravitational interaction of light: from weak to strong fields. *General Relativity and Gravitation*, 31(1):91–105, 1999.
6. M. V. Federov and S. V. Vintskevich. Diverging light pulses in vacuum: Lorentz-invariant mass and mean propagation speed. *Laser Phys. 27, 036202*, 2017.
7. Daniel Giovannini, Jacquiline Romero, Václav Potoček, Gergely Ferenczi, Fiona Speirits, Stephen M. Barnett, Daniele Faccio, and Miles J. Padgett. Spatially structured photons that travel in free space slower than the speed of light. *Science*, 347(6224):857–860, 2015.
8. C. W. Misner, K. S. Thorne, and J. A. Wheeler. *Gravitation*. W.H. Freeman and Company, San Francisco, 1973.
9. Nikolai Vsevolodovich Mitskievic and Krishnadeva K. Kumaradtya. The gravitational field of a spinning pencil of light. *Journal of Mathematical Physics*, 30(5):1095–1099, 1989.
10. Raymond W. Nackoney. The gravitational influence of a beam of light. *Journal of Mathematical Physics*, 14(9), 1973.
11. Dennis Rätzel, Martin Wilkens, and Ralf Menzel. Gravitational properties of light – the gravitational field of a laser pulse. *New Journal of Physics*, 18(2):023009, 2016.
12. Dennis Rätzel, Martin Wilkens, and Ralf Menzel. Gravitational properties of light: The emission of counter-propagating laser pulses from an atom. *Physical Review D*, 95(8):084008, 2017.
13. Fabienne Schneiter, Dennis Rtzel, and Daniel Braun. The gravitational field of a laser beam beyond the short wavelength approximation. *Classical and Quantum Gravity*, 35(19):195007, 2018.
14. J. Strohaber. Frame dragging with optical vortices. *General Relativity and Gravitation*, 45(12):2457–2465, 2013.
15. Richard C. Tolman, Paul Ehrenfest, and Boris Podolsky. On the gravitational field produced by light. *Phys. Rev.*, 37:602–615, Mar 1931.
16. J. W. van Holten. The gravitational field of a light wave. *Fortschr. Phys. 59, No. 3-4, 284-295*, 2011.
17. N.A. Voronov and I. Yu. Kobzarev. On the gravitational field of a massless particle. *JETP, Vol. 39, No. 4, p. 575*, 39(4):575, 1973.

# Schwinger effect impacting primordial magnetogenesis

Clément Stahl

*Instituto de Física, Pontificia Universidad Católica de Valparaíso, Casilla 4950, Valparaíso, Chile*

*clement.stahl@pucv.cl*

We explore the enhancement of an electromagnetic field in an inflationary background with an anti-conductive plasma of scalar particles. The scalar particles are created by Schwinger effect in curved spacetime and backreact to the electromagnetic field. The possibility of a negative conductivity was recently put forward in the context of the renormalization of the Schwinger induced current in de Sitter spacetime. While a negative conductivity enhances the produced magnetic field, we find that it is too weak to seed the observed intergalactic magnetic field today. This result on pair creation in inflationary scenario is however important for primordial scenarii of magnetogenesis as the presence of a conductivity alters the spectral index of the magnetic field. This also shows on a specific example that backreaction can increase the electromagnetic field and not only suppress it.

*Keywords*: Cosmology, Quantum field theory in curved spacetime, Inflation, Magnetogenesis, Backreaction.

## 1. Introduction

The generation of cosmological scale magnetic fields is an open question nowadays[1]. From blazar observation, it is possible to put a lower bound on the intergalactic magnetic field today: $B_0 > 10^{-18}$ G[2]. A primordial (inflationary) generation of magnetic fields is appealing because it would explain the large coherence length of the present magnetic field. However, most scenarii are plagued with the so-called backreaction problem and strong coupling problem[3]. On the other hand, if strong electromagnetic fields are present during inflation, they may trigger particle production also known as Schwinger effect. In inflationary spacetime, the production of particles has very peculiar effects: negative conductivity, infrared-hyperconductivity that could lead to exotic phenomenology.

In this proceeding, we report on Ref. 4, which investigated the backreaction of particles produced via Schwinger mechanism to the electromagnetic field. It is found that: provided that the particules are produced in a weak electric field regime (assumption required in order to have a negative conductivity), the electromagnetic field is enhanced. But not enough to seed the observed present magnetic field. Our model is specific in the sense that it does *not* break the conformal invariance by hand but with backreaction effects. Besides, it shows on a specific example how the electromagnetic fields is impacted by particle production. A more complete numerical investigation was carried out in Ref. 5.

In section 2, we review the latest results on Schwinger effect in de Sitter space. In section 3, we present the basic equations for our setup and calculate the electromagnetic power spectra at the end of inflation. Eventually in section 4, we give some concluding remarks and discuss some possible extensions of this work.

## 2. Schwinger effect

We consider the production of charged scalar particles $\varphi$ of mass $m$ in $1+3$ D under the influence of a constant electric field in dS:

$$S = \int \sqrt{-g} d^4x \left[ -\frac{1}{4\pi} F_{\mu\nu} F^{\mu\nu} + g^{\mu\nu} (\partial_\mu - ieA_\mu) \varphi^* (\partial_\nu + ieA_\nu) \varphi - m^2 \varphi^2 \right]. \quad (1)$$

The scale factor $a$ of the de Sitter metric reads:

$$a(\eta) = -\frac{1}{1 - H\eta}, \quad (2)$$

where $\eta$ is the conformal time running from $-\infty$ to $H^{-1}$ and $H$ is the Hubble constant. The gauge field $A_\mu$ ($\mu = 0, 1, 2, 3$) is considered pointing in the z-direction with a constant amplitude $E^2$. We will discuss in more detail in section 3 the relevance of modeling the electric field as a constant one. In dS, (2), it reads:

$$A_\mu = -\frac{E}{H} \left( a(\eta) - 1 \right) \delta^z_\mu. \quad (3)$$

With such as choice, the Lorentz and Coulomb gauge conditions are satisfied: $A_0 = \partial_i A_i = 0$. In order to quantize, the rescaled complex scalar field $\chi = a\varphi$ is expanded on modes function:

$$\chi(\eta, \mathbf{x}) = \int \frac{d^3\mathbf{k}}{(2\pi)^3} e^{i\mathbf{x}.\mathbf{k}} \left( \chi_{\mathbf{k}}(\eta) b_{\mathbf{k}} + \chi^*_{\mathbf{k}}(\eta) d^\dagger_{-\mathbf{k}} \right) \quad (4)$$

The commutation relation for the annihilation and creation operators are:

$$\left[ b_{\mathbf{k}}, b^\dagger_{\mathbf{k}'} \right] = \left[ d_{\mathbf{k}}, d^\dagger_{\mathbf{k}'} \right] = (2\pi)^3 \delta^{(3)}(\mathbf{k} - \mathbf{k}'), \quad (5)$$

the other commutators being 0. Varying the action (1) with respect to the scalar field gives the Klein-Gordon equation for each modes:

$$\left[ \partial^2_\eta + \omega^2_{\mathbf{k}}(\eta) \right] \chi_{\mathbf{k}} = 0, \quad (6)$$

where the effective pulsation is: $\omega^2 = \mathbf{p}^2 + \frac{\beta}{1-H\eta} + \frac{\gamma}{(1-H\eta)^2}$, the shifted momentum is: $\mathbf{p} = \mathbf{k} + (0, 0, \frac{eE}{H})$, $\alpha = m^2 + \left(\frac{eE}{H}\right)^2$ and $\beta = -2\frac{eE}{H} p_z$. Changing the time variable to $z = -2ip\left(\frac{1}{H} - \eta\right)$, two independent solutions to (6) are the Whittaker functions $\mathcal{W}_{\kappa,\mu}$ and $\mathcal{M}_{\kappa,\mu}$ which are hypergeometric functions. The Maxwell current is defined as:

$$J_\mu = ie \left[ (D_\mu \varphi)^\dagger \varphi - \varphi^\dagger D_\mu \varphi \right], \quad (7)$$

where $D_\mu = \partial_\mu + ieA_\mu$. This current can be shown to be conserved and the only non-vanishing component of it is the z-component. It is possible to calculate it in term of the solutions to the Klein-Gordon equation (6):

$$< J_i >= \frac{2e}{a^2} \int \frac{d^3\mathbf{k}}{(2\pi)^3} (k_i + eA_i) |\chi_{\mathbf{k}}|^2. \quad (8)$$

With the ingredients given above, it is in principle a mechanical task to calculate the current (8). But this current is UV divergent and needs to be renormalized. The issue of renormalization is not as well understood in curved spacetime than in flat spacetime, as Prof. Navarro-Salas also discussed during his talk of the session[6,7], see also Ref. 8 for related issues. While, the exact details of the renormalization scheme are still debated, it seems that most of the groups[9–11] agree that the current (8) in the weak electric field regime ($\frac{eE}{H^2} \ll 1$) is *negative*:

$$\mathcal{J} = -\frac{8}{9\pi} M^3 \exp(-2\pi M) L, \tag{9}$$

where we further defined the dimensionless electric field and mass: $M = \frac{m}{H}$, $L = \frac{eE}{H^2}$. This results will be our starting point to investigate how magnetogenesis is impacted by the presence of such a negative current. A negative current is somehow counter-intuitive because it would means that charged particles move opposite to the electric field, which gives intuitively an anti-screening of the electromagnetic field. This effect exist only in the infrared regime of de Sitter and therefore deserve special attention because of its peculiarity.

## 3. Inflationary magnetogeneis

We now also use the action (1) but focus on the electromagnetic side of the theory. We will be using the results of section 2 for Schwinger effect in dS: at each Hubble time, pairs are created. Since those pairs were derived for a constant electric field, our model is not valid for any situation: the typical time of pair creation $t_{\rm pair} = \frac{1}{H}$ has to be much smaller that the typical time of evolution of the electromagnetic field $t_E = \frac{1}{k}$. In other words, our model is valid, as soon as one focuses large scales. Intuitively, particle creation is a local quantum process, therefore the time variation of the electric field have to be on scale larger than the scale of production of the pairs. Besides, as we investigate the possibility of a negative conductivity as in equation (9) which appears in the weak electric field limit: $L \ll 1$, this avoids, in passing, the backreaction problem present in many models of magnetogenesis. Inspired by (9), we assume that a Maxwell current of the form

$$J_i = H\sigma E_i. \tag{10}$$

$\sigma$ is a dimensionless negative constant. It depends on the details of the pair creation process. The conformal invariance prevents perturbations of a gauge-field to develop, however here, the Schwinger mechanism breaks it with quantum fluctuations inducing a current and leading to the possibility of parametric amplification of the gauge field. The case of a strong electric field ($L \gg 1$) leads to a positive conductivity which screens the amplification of the EM field[12,13].

Varying (1), including the Maxwell current and changing variable to the canonical one, one gets[4]:

$$\mathcal{A}''(k,\eta) + \left[k^2 - H^2 a^2 \left(\frac{\sigma}{2}\right)\left(1 - \frac{\sigma}{2}\right)\right] \mathcal{A}(k,\eta) = 0 \tag{11}$$

which is the equation of an harmonic oscillator with a time dependent frequency. The detailed calculation for the quantization procedure can found in Ref. 14. It is possible to solve explicitly (11) to eventually find the electric and magnetic power spectra.

$$P_E \propto H^4 \left(\frac{k}{aH}\right)^{\sigma+4} \tag{12}$$

$$P_B \propto H^4 \left(\frac{k}{aH}\right)^{5-|\sigma+1|} \tag{13}$$

**Discussion of the parameters** We define the electric and magnetic spectral index $n_E$ and $n_B$ as $n_{E/B} \equiv \frac{\partial \log P_{E/B}}{\partial \log k}$. To avoid the backreaction problem and validate the scheme of pair production considered by taking an induced current of the form (10), one needs to consider a scenario where the electric field does not dominate the dynamics: $n_E > 0$ implying: $\sigma \in ]-4, 0]$. The strong coupling problem is also solved for these values. In that case $n_B \in ]2, 5]$: the magnetic spectrum is blue, dominated by UV contributions and is more damped than its electric counter part. For those values, it has been shown however that the maximal magnetic field today at Mpc scales would be $10^{-32}$ G[3], which is lower than the observed lower bound[2]. In more standard scenarii, where conformal symmetry is broken by hand, for instance by a non-minimal kinetic term $I$ parameterized as $I(\eta) \propto a^n$, the same equation of motion for the gauge field (11) would hold provided that $2n = \sigma$. This is an important result of this work as particle creation is expected as soon as the gauge field is present and our result is a clear backreaction results on how the gauge field dynamics would change in the presence of this particle: the spectral index of the magnetic field for instance would be shifted by $\frac{\sigma}{2}$. Similar results were found in Ref. 5.

## 4. Conclusion

In this proceeding, we reported on the backreaction of the particles produced by Schwinger effect in de Sitter to the electromagnetic field[4]. We showed that Schwinger effect impacts on the amplification of the EM quantum fluctuations. In Ref. 12, it was already remarked that strong electric fields constrain a branch of the IFF scenarios in a similar manner than the backreaction problem. In this work, we completed this statement by considering negative conductivity (for weak electric fields) as a way to break the conformal invariance and to amplify EM quantum fluctuations. The presence of a negative conductivity is either motivated by a new implementation of the adiabatic subtractions in the renormalization of the electric current presented in Ref. 10 or by the presence of non-linear terms generalizing the Einstein-Maxwell equation and a logarithmic running of the coupling constant $e$[11]. We found models where both the strong coupling and the backreaction problem are avoided for $\sigma \in ]-4, 0]$ however these models do not fulfill the lower bound of observed magnetic fields[2].

Those results are also relevant for primordial magnetogenesis model builder as the pair creation process depicted here matters as soon as EM fields are present during inflation. While in the case, of strong (backreacting) electric fields, one expects a dwindling of the EM field[12], for weaker fields this effects could add up to any other way to break the conformal invariance. The global effect of Schwinger pair creation on the power spectrum of quantum fluctuations would be to have a redder spectral index, more concretely: $n_B \to n_B + \frac{\sigma}{2}$.

Several directions are fruitful for the future. Prof. Navarro-Salas showed during his talk that the electric field (as any connection) has to be taken to be adiabatic order 1 in order to recover the trace anomaly[6]. Prof. Bavarsad discussed during his talk that backreaction can also happen to the gravitational field and Schwinger pairs could challenge the stability of de Sitter space[15], see also Ref. 16 for recent results. Producing Schwinger particles from a SU(2) gauge field and not U(1) leads to a different phenomenology[17] and in particular, it is known that SU(2) particles typically produce huge amount of gravitational waves, which may be constrained by the standard cosmological observables (CBM...). Finally, directly coupling the particle produced to the inflaton, with the Schwinger-Keldysh formalism, allow to produce a unique type of non-gaussianities[18] which in principle allow to measure Schwinger effect in the CMB or in LSS.

## Acknowledgments

It is a pleasure to thank all the participants and speakers of this session for stimulating scientific discussions ans interaction during the meeting and during the nights. I am also very grateful to all the people who organized the Marcel Grossmann meeting: professors, staffs, secretaries, students.

## References

1. R. Durrer and A. Neronov, Astron. Astrophys. Rev. **21** (2013) 62 [arXiv:1303.7121 [astro-ph.CO]].
2. A. Neronov and I. Vovk, Science **328** (2010) 73 [arXiv:1006.3504 [astro-ph.HE]].
3. V. Demozzi, V. Mukhanov and H. Rubinstein, JCAP **0908** (2009) 025 [arXiv:0907.1030 [astro-ph.CO]].
4. C. Stahl, Nucl. Phys. B **939** (2019) 95 [arXiv:1806.06692 [hep-th]].
5. O. O. Sobol, E. V. Gorbar, M. Kamarpour and S. I. Vilchinskii, Phys. Rev. D **98** (2018) no.6, 063534 [arXiv:1807.09851 [hep-ph]].
6. A. Ferreiro and J. Navarro-Salas, Phys. Rev. D **97** (2018) no.12, 125012 [arXiv:1803.03247 [hep-th]].
7. A. Ferreiro, J. Navarro-Salas and S. Pla, Phys. Rev. D **98** (2018) no.4, 045015 [arXiv:1807.10361 [gr-qc]].
8. J. F. Barbero G., A. Ferreiro, J. Navarro-Salas and E. J. S. Villaseor, Phys. Rev. D **98** (2018) no.2, 025016 [arXiv:1805.05107 [gr-qc]] ; A. Ferreiro and J. Navarro-Salas, arXiv:1812.05564 [gr-qc].

9. T. Hayashinaka, T. Fujita and J. Yokoyama, JCAP **1607** (2016) no.07, 010 [arXiv:1603.04165 [hep-th]].
10. T. Hayashinaka and S. S. Xue, Phys. Rev. D **97** (2018) no.10, 105010 [arXiv:1802.03686 [gr-qc]].
11. M. Banyeres, G. Domnech and J. Garriga, JCAP **1810** (2018) no.10, 023 [arXiv:1809.08977 [hep-th]].
12. T. Kobayashi and N. Afshordi, JHEP **1410** (2014) 166 [arXiv:1408.4141 [hep-th]].
13. C. Stahl and S. S. Xue, Phys. Lett. B **760** (2016) 288 [arXiv:1603.07166 [hep-th]].
14. J. Martin and J. Yokoyama, JCAP **0801** (2008) 025 [arXiv:0711.4307 [astro-ph]].
15. E. Bavarsad, C. Stahl and S. S. Xue, Phys. Rev. D **94** (2016) no.10, 104011 [arXiv:1602.06556 [hep-th]] ; E. Bavarsad and N. Margosian, Journal of Research on Many - body Systems, Volume 9, Number 17, Summer 2018, 223
16. G. Moreau and J. Serreau, Phys. Rev. Lett. **122** (2019) no.1, 011302 [arXiv:1808.00338 [hep-th]].
17. K. D. Lozanov, A. Maleknejad and E. Komatsu, JHEP **1902** (2019) 041 [arXiv:1805.09318 [hep-th]] ; A. Maleknejad and E. Komatsu, arXiv:1808.09076 [hep-ph].
18. W. Z. Chua, Q. Ding, Y. Wang and S. Zhou, arXiv:1810.09815 [hep-th].

# Surface electrical charge distribution in white dwarfs

J. D. V. Arbañil*

*Departamento de Ciencias, Universidad Privada del Norte, Avenida el Sol 461 San Juan de Lurigancho, Lima, Peru.*
** E-mail: jose.arbanil@upn.pe*
*http://www.upn.edu.pe/es*

G. A. Carvalho, R. M. Marinho Jr and M. Malheiro

*Departamento de Física, Instituto Tecnológico de Aeronáutica, São José dos Campos, SP, 12228-900, Brazil*
*E-mail: araujogc@ita.br, marinho@ita.br, malheiro@ita.br*

We investigate the influence of the surface electrical charge in the static equilibrium configuration of white dwarfs, this is possible by solving numerically the hydrostatic equilibrium equation for the charged case. We consider that the fluid in the star is described by a fully degenerate electron gas and that the electric charge is distributed close to the surface of the white dwarf. We found that super-Chandrasekhar mass white dwarfs are found for a large surface electrical charge.

*Keywords*: White dwarfs; stellar structure.

## 1. Introduction

### 1.1. *Super-Chandrasekhar white dwarfs*

In the last few years have been disclosed observations concerning the existence of a particular event in the universe, the super-luminous type Ia supernovae (SNIa)[1–4]. These events are very interesting because their explosion are the more foreseeables and frequently the brightest incidents in the sky. Some authors argue that the possible progenitor of the event SNIa is the super-Chandrasekhar white dwarf, star which ultrapass the standard Chandrasekhar mass limit, $1.44\,M_\odot$[5,6]. It is estimated that the super-Chandrasekhar white dwarfs' masses are in the range $2.1-2.8\,M_\odot$[7,8].

Several authors hint models to explain super-Chandrasekhar white dwarfs. In literature, for example, we found white dwarf models where are considered a strong magnetic field[9,10], in rotation and with different topologies for magnetic field[11–14] and with a electric charge distribution[15].

It is found that white dwarfs with an uniform and very strong magnetic field can attain masses of $\sim 2.9\,M_\odot$. Although these objects exceeds significantly the Chandrasekhar limit, they suffer from severe stability[16–20]. The effect of the rotation and different topology for the magnetic field help to reach white dwarfs' masses around $5M_\odot$[11–14]. A similar effect to the ones aforementioned are produced by the electric charge, which produces a force which helps to the one generated by the fluid pressure to counteract more mass and thus avoid gravitational collapse[15].

### 1.2. *This work*

Such as is realized in Ref. 15, in this article we also studied the influence of the electric charge in the structure of white dwarfs. The main difference between these two works is the profile of the electric charge. Instead of consider the electric charge density proportional to the energy density, $\rho_e = \alpha\rho$, with $\alpha$ being a dimensionless constant, such as is considered in Ref. 15, here we consider that the electric charge is distributed at the star's surface[21] of the form:

$$\rho_e = k\, exp\left[-\frac{(r-R)^2}{b^2}\right], \tag{1}$$

where $r$ and $R$ represent the radial coordinate and the total radius of the uncharged star. Moreover, $b$ represents the width of the electric charge distribution, here we consider that $b = 10\,[\mathrm{km}]$. The charge profile (1) is taking into account since electrons and ions at white dwarfs could important producing surface's strong electric fields.

With the aim of found $k$, we use the equality:

$$\sigma = \int_0^\infty 4\pi r^2 \rho_e dr, \tag{2}$$

where $\sigma$ bears the magnitude directly proportional to the electric charge distribution. Considering Eq. (1) in Eq. (2), it is found:

$$k = \frac{\sigma}{4\pi}\left(\frac{\sqrt{\pi}\, b\, R^2}{2} + \frac{\sqrt{\pi}\, b^3}{4}\right)^{-1}. \tag{3}$$

As can be seen, Eq. (3) connects $k$ and $\sigma$.

For the fluid, we consider that the fluid pressure and fluid energy density within the object are given by the relations:

$$p(k_F) = \frac{1}{3\pi^2\hbar^3}\int_0^{k_F} \frac{k^4}{\sqrt{k^2+m_e^2}} dk, \tag{4}$$

$$\rho(k_F) = \frac{1}{\pi^2\hbar^3}\int_0^{k_F} \sqrt{k^2+m_e^2}\, k^2 dk + \frac{m_N \mu_e}{3\pi^2\hbar^3} k_F^3, \tag{5}$$

with $m_e$, $m_N$, $\hbar$, $\mu_e$ and $k_F$ being respectively the electron mass, the nucleon mass, the reduced Planck constant, the ratio between the nucleon number and atomic number for ions and the Fermi momentum of the electron[22]. It is important to say that we consider $\mu_e = 2$.

## 2. Stellar structure equations

We shall study the stellar equilibrium configuration of charged white dwarfs using the profile of the electric charge and the equation of state showed in the previous section. We consider that within the star the charged fluid is described by the energy-momentum tensor (EMT):

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu + p g_{\mu\nu} + \frac{1}{4\pi}\left(F^{\mu\gamma}F_{\varphi\gamma} - \frac{1}{4} g_{\mu\nu} F_{\gamma\beta}F^{\gamma\beta}\right), \tag{6}$$

where $\rho$ and $p$ represent respectively the energy density and fluid pressure. In addition, $u_\mu$ beings the fluid's four velocity, $g_{\mu\nu}$ stands the metric tensor and $F^{\mu\gamma}$ represents the Faraday-Maxwell tensor.

The interior space-time of the charged star is described, in Schwarzschild coordinates, by the following line element:

$$ds^2 = -e^{\nu(r)}dt^2 + e^{\lambda(r)}dr^2 + r^2\left(d\theta^2 + \sin^2\theta d\phi^2\right), \tag{7}$$

as indicated, the functions $\nu(r)$ and $\lambda(r)$ dependent on the radial coordinate $r$ only.

For the line element (7), the nonzero components of the Maxwell-Einstein equation are given by the equalities:

$$q'(r) = 4\pi\rho_e(r)r^2e^{\lambda(r)/2}, \tag{8}$$

$$m'(r) = 4\pi r^2\rho(r) + \frac{q(r)q'(r)}{r}, \tag{9}$$

$$p'(r) = -(p(r)+\rho(r))\left(4\pi rp(r) + \frac{m(r)}{r^2} - \frac{q^2(r)}{r^3}\right)e^{\lambda(r)} + \frac{q(r)q'(r)}{4\pi r^4}, \tag{10}$$

with the potential metric of the form:

$$e^{-\lambda(r)} = 1 - \frac{2m(r)}{r} + \frac{q^2(r)}{r^2}, \tag{11}$$

primes indicate the derivations with respect to $r$. The functions $q$ and $m$ represent the electric charge and the mass inside the sphere of radial coordinate $r$. Moreover, Eq. (10) represents the hydrostatic equilibrium equation, known also as the Tolman-Oppenheimer-Volkoff equation[23,24]. This equation is modified from its original version to include the electrical part[25].

## 3. Results and conclusions

The static equilibrium configurations of charged white dwarfs are presented in Fig. 1. On the left and on the right hand side are shown the behavior of total mass $M/M_\odot$

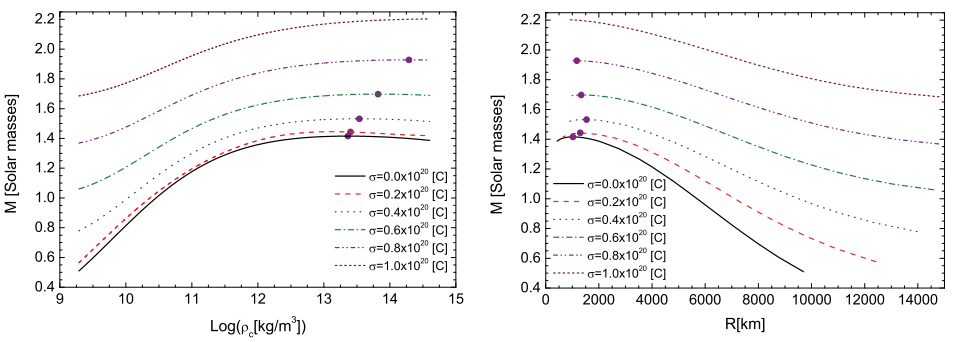



Fig. 1. The mass as a function of the central energy density and against the radius are plotted on the left and right hand side, respectively. The full circles represent the maximum mass points. In both figures are used different values of $\sigma$.

versus with central energy density $\rho_c$ and with the radius $R$, respectively, for different values of $\sigma$. For $\sigma$ lower than $0.8 \times 10^{20}$ [C], it can be observed that $M/M_\odot$ grows with $\rho_c$ until to attain $M_{\max}/M_\odot$ (marked with a full circles in purple), after that point, $M/M_\odot$ begins to decreases with the increment of $\rho_c$. At the same time, it is important to say that for $\sigma = 1.0 \times 10^{20}$ [C], $M/M_\odot$ increases monotonically with $\rho_c$, in this way, not maximum mass point is found. In addition, we highlight that the maximum mass found in this model is $2.199 M_\odot$, which is obtained using $\sigma = 1.0 \times 10^{20}$[C]. This total mass in the interval masses estimated for the super-Chandrasekhar white dwarfs, 2.1-2.8$M_\odot$, check Refs. 8, 7.

In Fig. 1 we can also note that in the uncharged case we obtain a maximum mass is found in a central energy density $\rho_c^*$, in turn, in the charged cases this points are attained for $\rho_c > \rho_c^*$. Additionally, such as happen for the total mass, the radius change with the total charge.

From the results aforementioned, we can conclude that the electric charge affects notably the static equilibrium configuration of white dwarfs. For a larger $\sigma$, white dwarfs with larger masses are found. This can be understand since $\sigma$ is directly related with the total charge in the star. An increment of the electric charge produce an increment of a force which helps to the fluid pressure to support more mass, thus avoiding the gravitational collapse. More details are presented in Ref. 26.

## Acknowledgments

Authors thanks to Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-CAPES and Fundação de Amparo à Pesquisa do Estado de São Paulo-FAPESP, thematic project 2013/26258-4, for grants.

## References

1. R. A. Scalzo *et al.*, *Astrophys. J.* **713**, 1073 (2010).
2. D. A. Howell *et al.*, *Nature (London)* **443**, 308 (2006).
3. M. Hicken *et al.*,*Astrophys. J.* **669**, L17 (2007).
4. M. Yamanaka *et al.*, *Astrophys. J.* **707**, L118 (2009).
5. S. Chandrasekhar, *Mon. Not. R. Astron. Soc.* **91**, 456 (1931).
6. S. Chandrasekhar, *Mon. Not. R. Astron. Soc.* **95**, 207 (1935).
7. J. M. Silverman *et al.*, *Mon. Not. R. Astron. Soc.* **410**, 585 (2011).
8. S. Taubenberger *et al.*, *Mon. Not. R. Astron. Soc.* **412**, 2735 (2011).
9. U. Das and B. Mukhopadhyay, *Phys. Rev.* **D86**, 042001 (2012).
10. U. Das and B. Mukhopadhyay, *Phys. Rev. Lett.* **110**, 071102 (2013).
11. K. Boshkayev, J. A. Rueda, R. Ruffini and I. Siutsou, *Astrophys. J.* **762**, 117 (2013).
12. B. Franzon and S. Schramm, *Phys. Rev.* **D92**, 083006 (2015).
13. S. Subramanian and B. Mukhopadhyay, *Mon. Not. R. Astron. Soc.* **454**, 752 (2015).

14. P. Bera and D. Bhattacharya, *Mon. Not. R. Astron. Soc.* **456**, 3375 (2016).
15. H. Liu, X. Zhang, and D. Wen, *Phys. Rev.* **D89**, 104043 (2014).
16. J. M. Dong, W. Zuo, P. Yin, and J. Z. Gu, *Phys. Rev. Lett.* **112**, 039001 (2014).
17. N. Chamel, A. F. Fantina, and P. J. Davis, *Phys. Rev.* **D88**, 081301(R) (2013).
18. J. G. Coelho *et al.*, *Astrophys. J.* **794**, 86 (2014).
19. R. Nityananda and S. Konar, *Phys. Rev.* **D91**, 028301 (2015).
20. G. A. Carvalho, R. M. Marinho Jr and M. Malheiro, *Gen. Relativ. Gravit.* **50**, 38 (2018).
21. R. P. Negreiros, F. Weber, M. Malheiro and V. Usov, *Phys. Rev.* **D80**, 083006 (2009).
22. S. Shapiro and S. Teukolsky, *Black Holes, White Dwarfs and Neutron Stars: The Physics of Compact Objects*, (Wiley, 2008)
23. R. C. Tolman, *Phys. Rev.* **D55**, 364 (1939).
24. J. R. Oppenheimer and G. Volkoff, *Phys. Rev.* **D55**, 374 (1939).
25. J. Bekenstein, *Phys. Rev.* **D4**, 2185 (1971)
26. G. A. Carvalho, J. D. V. Arbañil, R. M. Marinho Jr. and M. Malheiro, *Eur. Phys. J.* **C78**, 411 (2018).

# Possible formation of lowly luminous highly magnetized white dwarfs by accretion leading to SGRs/AXPs

B. Mukhopadhyay[1*], M. Bhattacharya[2], A. R. Rao[3], S. Mukerjee[1] and U. Das[4]

*1. Indian Institute of Science, Bangalore 560012, India*
**E-mail: bm@iisc.ac.in*
*2. University of Texas, Austin, USA*
*3. Tata Institute of Fundamental Research, Mumbai, India*
*4. University of Colorado, Boulder, USA*

We sketch a possible evolutionary scenario by which a highly magnetized super-Chandrasekhar white dwarf could be formed by accretion on to a commonly observed magnetized white dwarf. This is an exploratory study, when the physics in cataclysmic variables (CVs) is very rich and complex. Based on this, we also explore the possibility that the white dwarf pulsar AR Sco acquired its high spin and magnetic field due to repeated episodes of accretion and spin-down. We show that strong magnetic field dramatically decreases luminosity of highly magnetized white dwarf (B-WD), letting them below the current detection limit. The repetition of this cycle can eventually lead to a B-WD, recently postulated to be the reason for over-luminous type Ia supernovae. A spinning B-WD could also be an ideal source for continuous gravitational radiation and soft gamma-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs). SGRs/AXPs are generally believed to be highly magnetized, but observationally not confirmed yet, neutron stars. Invoking B-WDs does not require the magnetic field to be as high as for neutron star based model, however reproducing other observed properties intact.

*Keywords*: White dwarfs; strong magnetic fields; CVs; pulsars; SGRs/AXPs.

## 1. Introduction

Several independent observations repeatedly argued in recent past for the existence of highly magnetized white dwarfs (B-WDs). Examples are overluminous type Ia supernovae[1,2], white dwarf pulsars[3,4] etc. Also soft gamma-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs) could be explained as B-WDs[5–7], while they are generally believed to be highly magnetized neutron stars[8] without however any direct detection of underlying required high surface field $B_s \sim 10^{15}$ G. Interestingly, explaining SGR/AXP by a magnetized white dwarf requires a lower $B_s \lesssim 10^{12}$ G, which may however correspond to central field $\gtrsim 10^{14}$ G. Nevertheless, the origin of such fields in a white dwarf remains a question, when the observed confirmed surface field is $\lesssim 10^9$ G.

Here we explore a possible evolution of a conventionally observed magnetized white dwarf to a B-WD by accretion, which may pass through a phase exhibiting currently observed AR Sco. This is an exploratory study, and the present venture is based more on an idealized situation, when the physics in accreting white dwarfs, i.e. cataclysmic variables (CVs), is very rich and complex. We also show, based on some assumption, that the thermal luminosity of such a B-WD could be very small, below their current detection limit. However, due to high field and rotation, their spin-down luminosity could be quite high. Hence, they could exhibit SGRs/AXPs.

## 2. Accretion induced evolution

The detailed investigation of the accretion induced evolution faces several difficulties including nova eruptions (hence nonsteady increase of mass) and the eruption and ejection of accumulated shells. Nevertheless, the discovery of AR Sco, which is a fast rotating magnetized white dwarf, argues for the possibility of episodic increase of mass in a CV. Hence, we sketch a tentative evolutionary scenario with repeated episodes of accretion phase leading to the high magnetic field via flux freezing and spin-power phase decreasing field. Eventually this mechanism can plausibly lead to a B-WD. Note that there are already observational evidences for transitions between spin-power and accretion-power phases in a binary millisecond pulsar[9]. The conservation laws controlling the accretion-power phase around the stellar surface of radius $R$ and mass $M$, which could be inner edge of accretion disk, are given by

$$l\Omega(t)^2 R(t) = \frac{GM(t)}{R(t)^2},\ I(t)\Omega(t) = \text{constant},\ B_s(t)R(t)^2 = \text{constant}, \tag{1}$$

where $l$ takes care of inequality due to dominance of gravitational force over the centrifugal force in general, $I$ is the moment of inertia of star and $\Omega$ the angular velocity of the star which includes the additional contribution acquired due to accretion as well. Solving the conservation laws given by equation (1) simultaneously, we obtain the time evolution of radius (or mass), magnetic field and angular velocity during accretion. Accretion stops when

$$-\frac{GM}{R^2} = \frac{1}{\rho}\frac{d}{dr}\left(\frac{B^2}{8\pi}\right)|_{r=R} \sim -\frac{B_s^2}{8\pi R\rho}, \tag{2}$$

where $\rho$ is the density of inner edge of disk.

For a dipolar fixed field, $\dot{\Omega} \propto \Omega^3$ [7], where over-dot implies time derivative. Generalizing for the present purpose it becomes $\dot{\Omega} = k\Omega^n$ with $k$ being constant. Therefore, during the phase of spin-power pulsar (when accretion inhibits), the time evolution of angular velocity and surface magnetic field may be given by

$$\Omega = \left[\Omega_0^{1-n} - k(1-n)(t-t_0)\right]^{\frac{1}{1-n}},\ B_s = \sqrt{\frac{5c^3 Ik\Omega^{n-m}}{R^6 \sin^2\alpha}}, \tag{3}$$

where $\Omega_0$ is the angular velocity when accretion just stops at the beginning of spin-power phase at time $t = t_0$, $k$ is fixed to constrain $B_s$ at the beginning of first spin-powered phase, which is determined from the field evolution in the preceding accretion-power phase, $\alpha$ is the angle between magnetic and spin axes. Note that $n = m = 3$ corresponds to dipole field.

Figure 1 shows a couple of representative possible evolutions of angular velocity and magnetic field with mass (time). It is seen that initial larger $\Omega$ with accretion drops significantly during spin-power phase (when accretion stops and hence no change of mass), followed by its increasing phase. Similar trend is seen in $B_s$ profiles with a sharp increasing trend (with value $\sim 10^{11}$ G) at the last cycle leading to the increase of $B_c$ as well, forming a B-WD. At the end of evolution, it could be left

out as a super-Chandrasekhar B-WD and/or a SGR/AXP candidate with a higher spin frequency. Of course, in reality they may depend on many other factors and the current picture does not match exactly with what is expected in AR Sco itself.

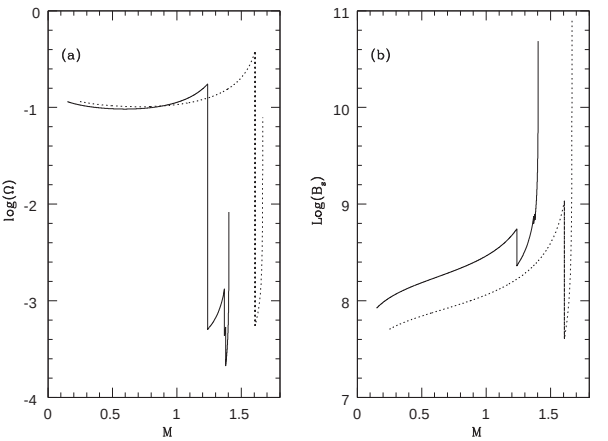



Fig. 1. Time evolution of (a) angular velocity in $s^{-1}$, (b) magnetic field in G, as functions of mass in units of solar mass. The solid curves correspond to $n = 3$, $m = 2.7$, $\rho = 0.05$ gm cm$^{-3}$, $l = 1.5$ and dotted ones to $n = 3$, $m = 2$, $\rho = 0.1$ gm cm$^{-3}$, $l = 2.5$; $k = 10^{-14}$ CGS, $\dot{M} = 10^{-8} M_{\odot} \mathrm{Yr}^{-1}$, $\alpha = 10$ degree and $R = 10^4$ km at $t = 0$. This is reproduced from a previous work[7].

## 3. Luminosity

With the increase of mass, the radius of white dwarfs, hence B-WDs, becomes very small[4,10]. Indeed, the increase of magnetic field is due to decreasing radius via flux-freezing. Now due to smaller radius, UV-luminosity of B-WDs turns out to be very small if the surface temperature is same as their nonmagnetic counterpart[7]. However more interestingly, from the conservation of energy, it is expected that the presence of strong magnetic field enforces decreasing thermal energy and hence lowering luminosity in stable equilibrium.

Combining the magnetostatic and photon diffusion equations in the presence of magnetic field but ignoring tension, we obtain

$$\frac{d}{dT}\left(P + P_B\right) = \frac{4ac}{3}\frac{4\pi GM}{L}\frac{T^3}{\kappa}, \tag{4}$$

which we solve to obtain the envelop properties. Here $P$ is the matter pressure, $P_B$ the magnetic pressure, $\kappa$ the opacity, $T$ the temperature, $a$ the radiation constant, $M$ the mass of white dwarf within the core radius $r$, which is practically the whole mass of white dwarf because the envelop is very thin, and $L$ is the luminosity. For the strong field considered here, the radiative opacity variation with $B$ can be modelled similarly to neutron stars as $\kappa = \kappa_B \approx 5.5 \times 10^{31} \rho T^{-1.5} B^{-2}$ cm$^2$g$^{-1}$ [11].

We use a field profile proposed earlier for neutron stars[12] to enumerate the field magnitude at a given density (radius), irrespective of other complicated effects, given by

$$B\left(\frac{\rho}{\rho_0}\right) = B_s + B_0\left[1 - \exp\left(-\eta\left(\frac{\rho}{\rho_0}\right)^{\gamma}\right)\right], \tag{5}$$

where $B_0$ (similar to central field) is a parameter with the dimension of $B$, other parameters are set as $\eta = 0.1$, $\gamma = 0.9$, $\rho_0 = 10^9$ g cm$^{-3}$ for all the calculations. Further equating the electron pressure for the non-relativistic electrons on both sides of the core-envelop interface gives

$$\rho_*(B_*) \approx 1.482 \times 10^{-12}\, T_*^{1/2} B_s \tag{6}$$

at interface. Now we solve equation (4) along with the photon diffusion equation

$$\frac{dT}{dr} = -\frac{3}{4ac}\frac{\kappa(\rho + \rho_B)}{T^3}\frac{L}{4\pi r^2}, \tag{7}$$

with boundary conditions $\rho(T_s) = 10^{-10}$ g cm$^{-3}$, $r(T_s) = R = 5000$ km and $M = M_\odot$, where $T_s$ is the surface temperature, and obtain $\rho - T$ and $r - T$ profiles. Further $T_*$ and $\rho_*$ can be obtained by solving for the $\rho - T$ profile along with equation (6), as shown in Fig. 2, and knowing $T_*$, we can obtain $r_*$ from the $r - T$ profile. We see that interface moves inwards ($r_*$ decreases) with increasing $B$ and $L$. But $\rho_*$ increases with increasing $L$ and/or $B$, as $\rho_* \propto T_*^{1/2} B$.

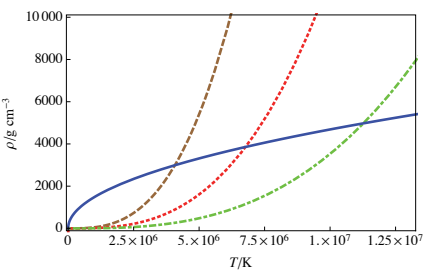



Fig. 2. Variation of density with temperature for $B \equiv (B_s, B_0) = (10^{12}\,\mathrm{G}, 10^{14}\,\mathrm{G})$ and $L = 10^{-5}L_\odot$ (dashed line), $10^{-4}L_\odot$ (dotted line) and $10^{-3}L_\odot$ (dot-dashed line). The solid line represents equation (6). This is reproduced from a previous work[13].

With above benchmarking, we now explore, based energy conservation, if the luminosity of a magnetized white dwarf or B-WD changes. For $B = 0$ and $L = 10^{-5}L_\odot$, we have $r_* = 0.9978R$, $\rho_* = 170.7$ g cm$^{-3}$ and $T_* = 2.332 \times 10^6$ K. Using the same boundary condition as described above, we now solve equations (4) and (7) with $B \neq 0$, but vary $L$ in order to fix $r_* = 0.9978R$. We find interestingly that $L$ decreases for $B \neq 0$, as shown in Table 1. Physically this corresponds to increasing $B$, and thence magnetic energy, is compensated by decreasing thermal energy (decreasing $T_*$) and thence $L$, when total energy is conserved. Similarly,

increasing $B$ may be compensated by decreasing gravitational energy (decreasing $r_*$). In either of the cases, $L$ decreases.

Table 1. Variation of luminosity with magnetic field for fixed $r_* = 0.9978R$

| $B/\mathrm{G} = (\mathrm{B_s/G}, \mathrm{B_0/G})$ | $L/L_\odot$ | $T_*/\mathrm{K}$ | $\rho_*/\mathrm{g\,cm^{-3}}$ | $T_\mathrm{s}/\mathrm{K}$ |
|---|---|---|---|---|
| $(0, 0)$ | $1.00 \times 10^{-5}$ | $2.332 \times 10^{6}$ | $1.707 \times 10^{2}$ | $3.85 \times 10^{3}$ |
| $(10^9, 6 \times 10^{13})$ | $2.53 \times 10^{-7}$ | $4.901 \times 10^{5}$ | $1.037 \times 10^{0}$ | $1.53 \times 10^{3}$ |
| $(5 \times 10^9, 2 \times 10^{13})$ | $3.96 \times 10^{-8}$ | $3.262 \times 10^{5}$ | $4.232 \times 10^{0}$ | $9.65 \times 10^{2}$ |
| $(10^{10}, 10^{13})$ | $1.02 \times 10^{-6}$ | $7.189 \times 10^{5}$ | $1.257 \times 10^{1}$ | $2.17 \times 10^{3}$ |
| $(2 \times 10^{10}, 8 \times 10^{12})$ | $4.40 \times 10^{-9}$ | $2.063 \times 10^{5}$ | $1.346 \times 10^{1}$ | $5.57 \times 10^{2}$ |
| $(5 \times 10^{10}, 4 \times 10^{12})$ | $2.59 \times 10^{-8}$ | $3.185 \times 10^{5}$ | $4.182 \times 10^{1}$ | $8.68 \times 10^{2}$ |
| $(5 \times 10^{11}, 10^{12})$ | $2.93 \times 10^{-9}$ | $2.206 \times 10^{5}$ | $3.480 \times 10^{2}$ | $5.03 \times 10^{2}$ |

## 4. SGRs/AXPs as B-WDs

Paczynski[5] and Usov[6] independently proposed that SGRs and AXPs are moderately magnetized white dwarfs but following Chandrasekhar's mass-radius relation[14]. Many features of SGRs/AXPs are explained by their model at relatively lower magnetic fields, while the more popular magnetar model[8] requires field $\gtrsim 10^{15}$ G, which is not observationally well established yet. Nevertheless, such a white dwarf based model suffers from a deep upper limit on the optical counterparts of some AXPs/SGRs, e.g. SGR 0418+5729, due to their larger moment of inertia.

Now B-WDs established here could be quite smaller in size and hence have smaller moment of inertia. Therefore, their optical counterparts, with very low UV-luminosities, are quite in accordance with observation. Hence, the idea of B-WD brings a new scope of explain SGRs/AXPs at smaller magnetic fields, which are observationally inferable, compared to highly magnetized magnetar model. For details see the work by Mukhopadhyay & Rao[7].

## 5. Continuous Gravitational Radiation

Due to smaller size compared to their regular counterpart, B-WDs rotate relatively faster. Now if the rotation and magnetic axes are misaligned, they serve as good candidates for continuous gravitational radiation due to their quadrupole moment, characterized by the amplitude[15]

$$h_+(t) = \frac{h_0}{2}(1 + \cos^2\alpha_0)\cos\Phi(t),\ h_\times(t) = h_0\cos\alpha_0\sin\Phi(t),\ h_0 = \frac{4\pi^2 G I_{zz}\epsilon}{c^4 P_s^2 D}, \quad (8)$$

where $\alpha_0$ is the inclination of the star's rotation axis with respect to the observer, $\Phi(t)$ is the signal phase function, $\epsilon$ amounts the ellipticity of the star, $I_{zz}$ is the moment of inertial about z-axis, and $D$ is the distance between the star and detector.

A B-WD of mass $\sim 2M_\odot$, polar radius $\sim$ 700 km, spin period $P_s \sim 1$ s[16], $\epsilon \sim 5 \times 10^{-4}$ and $D \sim 100$ pc would produce $h_0 \sim 10^{-22}$, which is within the

sensitivity of the Einstein@Home search for early Laser Interferometer Gravitational Wave Observatory (LIGO) S5 data[15]. However, DECIGO/BBO would give a firm confirmation of their gravitational wave because they are more sensitive in their frequency range. In fact, if the polar radius is $\sim 2000$ km with $P_s \sim 10$ s and other parameters intact, DECIGO/BBO can detect it with $h_0 \sim 10^{-23}$. Nevertheless, (highly) magnetized rotating white dwarfs approaching B-WDs are expected to be common and such white dwarfs of radius $\sim 7000$ km, $P_s \sim 20$ sec and $D \sim 10$ pc could produce $h_0 \gtrsim 10^{-22}$ which is detectable by LISA.

## 6. Summary

The idea of B-WD has been proposed early this decade, mainly to explain observed peculiar type Ia supernovae inferring super-Chandrasekhar progenitor mass. Lately it has been found with various other applications, e.g. SGRs/AXPs, white dwarf pulsars like AR Sco, continuous gravitational wave etc. Here we have attempted to sketch a plausible evolution scenario to explain the formation of such a highly magnetized, smaller size white dwarf. In our simplistic picture, ignoring many complicated CV features, we are able to show that a commonly observed magnetized white dwarf could be evolved to a B-WD via accretion. Hence, the existence of highly magnetized, rotating, smaller white dwarfs is quite plausible.

## References


1. D. A. Howell, et al., *Nature* **443**, 308 (2006).
2. U. Das and B. Mukhopadhyay, *Phys. Rev. Lett.* **110**, 071102 (2013).
3. T. R. Marsh , et al., *Nature* **537**, 374 (2016).
4. B. Mukhopadhyay, A. R. Rao, and T. S. Bhatia, *MNRAS* **472**, 3564 (2017).
5. B. Paczynski, *ApJ* **365**, L9 (1990).
6. V. V. Usov, *ApJ* **427**, 984 (1994).
7. B. Mukhopadhyay and A. R. Rao, *JCAP* **05**, 007 (2016).
8. R. C. Duncan and C. Thompson, *ApJ* **392**, L9 (1992).
9. A. Papitto, et al., *Nature* **501**, 517 (2013).
10. U. Das, B. Mukhopadhyay, and A. R. Rao, *ApJ* **767**, L14 (2013).
11. A. Y. Potekhin and D. G. Yakovlev, *A&A* **374**, 213 (2001).
12. D. Bandyopadhyay, S. Chakrabarty, and S. Pal, *Phys. Rev. Lett.* **79**, 2176 (1997).
13. M. Bhattacharya, B. Mukhopadhyay, and S. Mukerjee, *MNRAS* **477**, 2 (2018).
14. S. Chandrasekhar, *MNRAS* **95**, 207 (1935).
15. C. Palomba, in *LIGO Scientific Collaboration and the Virgo Collaboration – Proceedings of the Recontres de Moriond*, 2011; arXiv:1201.3176
16. S. Subramanian and B. Mukhopadhyay, *MNRAS* **454**, 752 (2015).

# The effect of positrons in hot white dwarfs

Sílvia P. Nunes* and Manuel M. Malheiro

*Technological Institute of Aeronautics,*
*São José dos Campos, SP, Brazil*
**E-mail: silviapn@ita.br*

White dwarf stars are widely studied as being composed of an ion lattice embedded in a degenerate fermion gas. However, at the beginning of their lives these stars are subject to temperatures that can reach up to $T = 10^9 K$. In this limit there is no longer total degeneracy and a temperature-dependent equation of state (EOS) is needed to take in account the Fermi-Dirac occupation factor of fermion levels. In this article, we will study this regime, in particular the effect of positrons in the EOS and in the hot white dwarf structure.

*Keywords*: White Dwarfs, Hot White Dwarfs.

## 1. Introduction

White dwarfs (WDs) are known as stars that have mass of about one solar mass, radii of about 5000 km and mean densities around $10^6$ g/cm$^3$. These stars are also associated, after thermonuclear detonation in their interior, with supernova Ia formation [2]. As any other star, the WD also have a life, they are born with AGBs stars or even in mergers [2] and pass their lives cooling until they reach a totally degenerate state. Until this cold matter state is reached, the star can reach temperatures of about $10^8 K$ in the core, as already studied in Ref. 3, 4. Furthermore, temperatures of about $2 \times 10^9$ K have already been discussed in the white dwarf interior. [5]

In this article we propose the WD star may reach temperatures of $10^9$K and study the consequences of this value for the stellar interior, as the degeneracy breaks down. Also, we ended up struggling with a possible positron presence in the edge of the star, and analyzed their consequences in the star structure.

## 2. The relativistic frame

The Schwarzschild metric describes a spherical body with no rotation and can be used for WDs [6]

$$ds^2 = c^2\left(1 - \frac{2GM}{c^2 r}\right)dt^2 - \left(1 - \frac{2GM}{c^2 r}\right)^{-1} dr^2 - r^2 d\Omega^2 \tag{1}$$

from which we can derive the Tolman-Oppheimer-Volkoff equation and the mass conservation in the relativistic frame

$$\frac{dP(r)}{dr} = -\frac{4\pi r^3 G}{c^2}\frac{P(r)/c^2 + M(r)}{r^2\left(1 - \frac{2GM(r)}{c^2 r}\right)}\left[\varepsilon(r) + P(r)\right], \tag{2}$$

$$\frac{dM(r)}{dr} = 4\pi\rho r^2. \tag{3}$$

## 3. Chandrasekhar EOS at T=0

In the final state of a WD, the composition of this star can be described as a Fermi gas having degenerate matter, also known as the ideal Fermi gas. This was first described by Chandrasekhar in 1964 and the equation of state, responsible for the matter description as a function of the Fermi momentum $p_F$ is

$$P = P_e = \frac{1}{3\pi^2\hbar^3}\int_0^{p_F}\frac{p^4}{\sqrt{(mc^2)+(pc)^2}}, \tag{4}$$

$$n_e = \frac{1}{\pi^2\hbar^3}\int_0^{p_F} p^2 dp, \tag{5}$$

$$\varepsilon \simeq \varepsilon_N = \frac{A}{Z} m_u n_e, \tag{6}$$

where $P_e$ is the electron pressure, $A$ the mass number, $Z$ the atomic number, and $m_u = 1.6604 \times 10^{-24}$g the unified atomic mass.

Although the degenerate approximation is good for temperatures below $10^6$ K, there is a limit where this can not be used. In Fig. 1 we show the limit where the matter can be degenerate in the hatched area, with the condition $\rho_\star = (2.4 \times 10^{-8} g\ cm^{-3})\mu_e T_\star^{3/2}$. As the temperature arises, more the star will be nondegenerate. For temperatures below about $10^6 K$ a small part of the star is nondegenerate, which does not interfere with the star structure.

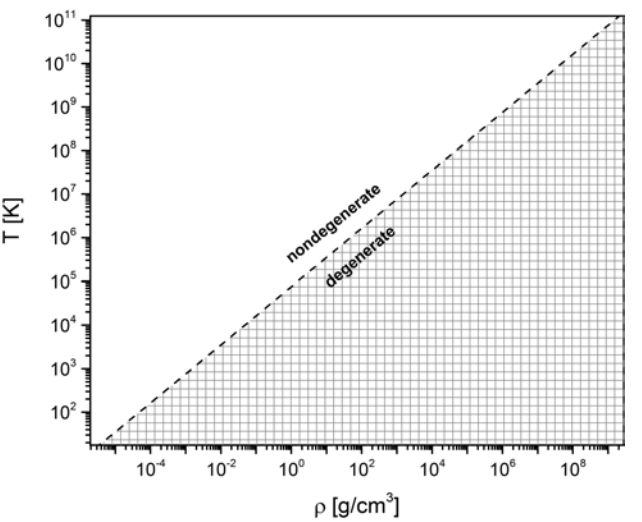



Fig. 1. The effects of nondegeneracy inside the WD star with a Fermi gas matter.

## 4. The equation of state for hot stars

For temperatures above $10^6$ the degenerate matter does not represent the whole star. So, the thermodynamics parameters have the Fermi-Dirac distribution instead of using the one particle per quantum level until the Fermi momentum. The pressure and the number density of the Fermi gas will be

$$P = P_e = \frac{1}{3\pi^2\hbar^3}\int_0^{p_F} \frac{p^4}{\sqrt{(mc^2)+(pc)^2}} \frac{1}{\exp\frac{(E-\mu)}{kT}+1}, \tag{7}$$

$$n_e = \frac{1}{\pi^2\hbar^3}\int_0^{p_F} p^2 \frac{1}{\exp\frac{(E-\mu)}{kT}+1} dp, \tag{8}$$

where $\frac{1}{\exp\frac{(E-\mu)}{kT}+1}$ is the Fermi-Dirac occupation factor. The same approximation can be used as Chandrasekhar for the energy density, as there are no new particles contributing. In Fig. 2 we can see the degeneracy behavior at different temperatures. For temperatures until $10^6 K$ the curves converge for densities above $10^2$ g/cm$^3$ which tells us that this may not influence on the star structure. The difference between the curves for small densities increases with the temperature. Using the

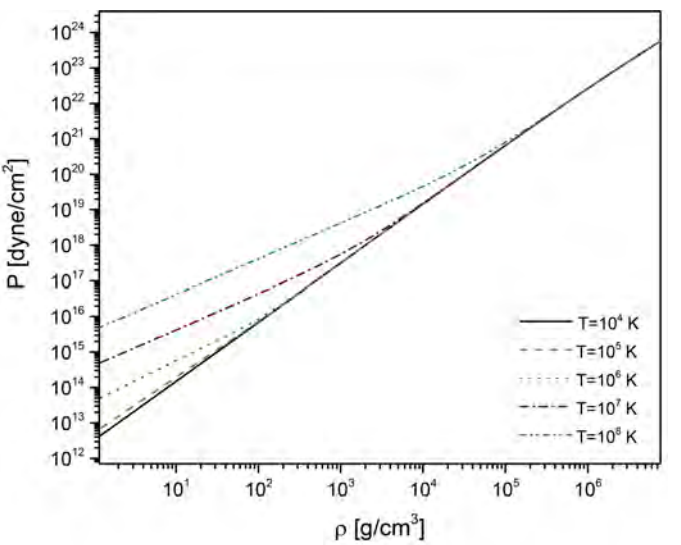



Fig. 2. The equation of state of a $C^{12}$ Fermi gas for various temperatures.

TOV equation, it is possible to find the mass as a function of the radius and the central density as shown in Fig. 3. In the left figure we can see that the radius of the star increases as the temperature increases. Furthermore, this effect is more noticeable for higher temperatures as $T = 10^8 K$.

The stability criteria can be determined considering a density $\rho^\star$, where $dM/d\rho = 0$, then, configurations with densities $\rho_c > \rho^\star$ are considered stable against

perturbation. We can notice in the right panel of Fig. 3 that the star stable region decreases with the temperature increasing.

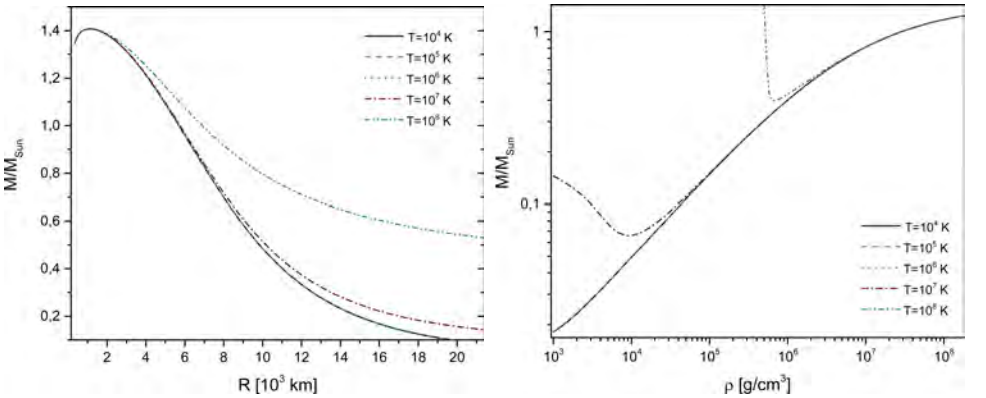



Fig. 3. Mass as a function of radius and mass as a function of central density for selected values of temperature in several stars.

### 4.1. *High temperatures inside white dwarfs*

The above treatment was done for temperatures above $10^8$, where the degenerate region still consists of a great part of the star. Considering now temperatures in the range $10^8$ to $10^9 K$ we show the equation of state in Fig. 4.

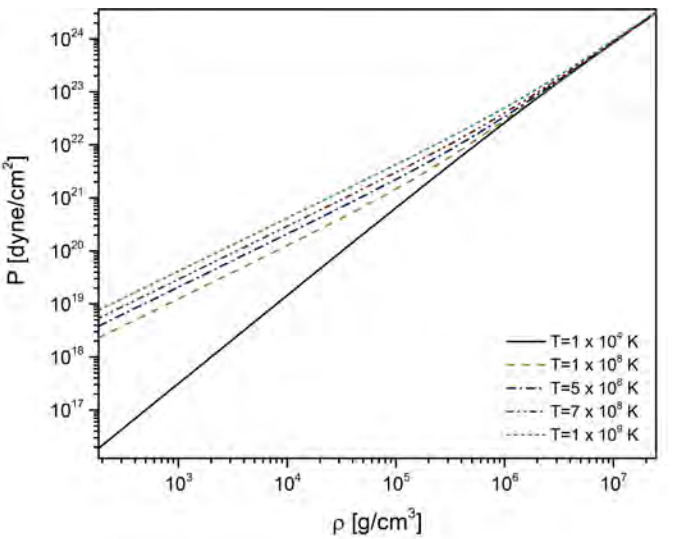



Fig. 4. The equation of state for a $C^12$ WD with high temperatures.

Increasing the white dwarf temperature we are struggling with a negative chemical potential in the end of the star. This can be understood if we consider a photon annihilation, where $\gamma \leftrightarrows e^- + e^+$. Since the photon has zero chemical potential, the positron chemical potential is related with the electron one as $\mu_{e^+} = -\mu_{e^-}$. So, the negative chemical potential indicates an antimatter (positron) presence.

To investigate the effect of positrons, we calculated the same equation of state with the negative chemical potential. In Fig. 5 we present the mass as function of the star radius and central density. The same conclusion can be made in relation to the stability, the increasing temperature generates a smaller stability region. The increase of the radius to a fixed mass is more noticeable when the thermal effects arise.

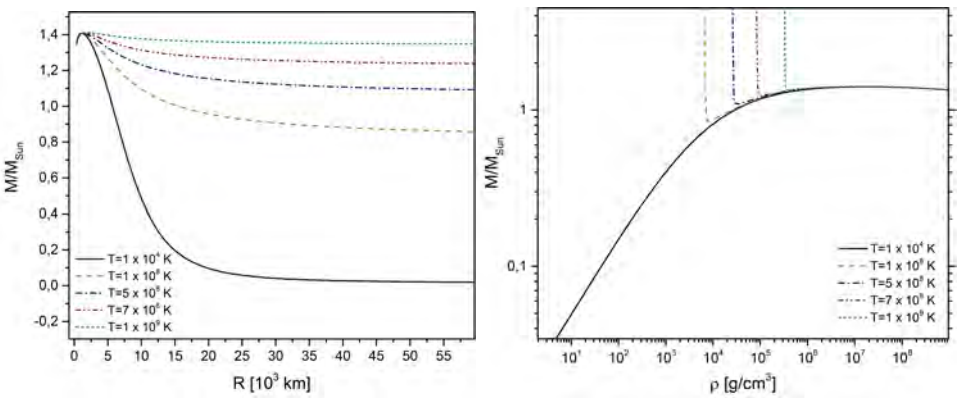



Fig. 5. Mass as a function of radius and mass as a function of central density for selected values of temperature in several stars.

### 4.2. *The effect of positrons in hot white dwarfs*

The negative chemical potential, as already discussed, is interpreted as a positron presence. This particle presence will give a residual positive charge in the last layers of the star. Instead of using equations 2 and 3, we will use a modified TOV, derived in the Reissner-Nordström spacetime solution, that includes new terms related to the electrostatic contribution

$$\frac{dq}{dr} = 4\pi\rho_e r^2 e^{\lambda/2}, \tag{9}$$

$$\frac{dm}{dr} = 4\pi r^2\rho + \frac{q}{r}\frac{dq}{dr}, \tag{10}$$

$$\frac{dP}{dr} = -(P+\rho)\left(4\pi r\rho + \frac{m}{r^2} - \frac{q^2}{r^3}\right)e^{\lambda} + \frac{q}{4\pi r^4}\frac{dq}{dr}, \tag{11}$$

where $q$ is the charge and $\rho_e$ the charge density.

We considered the charge density as

$$\rho_e = -en_{e^+}, \tag{12}$$

where $e$ is the electron charge and $n_{e^+}$ the positron number density. Here we find a hindrance, since the metric term $e^{\lambda} = (1 - \frac{2m}{r} + \frac{q^2}{r^2})^{-1}$ reaches negative values, and so the square root has an imaginary value.

## Conclusion

In this preliminary study we concluded that the effects of positrons in the high temperature regime need to be better studied, since they are complex. In the next steps of this work we are going to study carefully the degenerate effects and the cooling of these stars under the antimatter presence.

## Acknowledgments

We would like to thank the Marcel Grossman Meeting for the opportunity to present this work, and the CNPq and ITA for financial support. This work is supported in thematic project "Superdensity matter in the universe," FAPESP number 2013/26258–4.

## References

1. P. Hoyng, *Relativistic Astrophysics and Cosmology: A Primer*, (Springer Netherlands, 2007).
2. C. Raskin and E. Scannapieco and C. Fryer and G. Rockfeller and F.X. Timmes, *The Astrophysical Journal* **62**, (2012).
3. S.M. de Carvalho and M, Rotondo and J.A. Rueda and J. A. and R. Ruffini, *American Physical Society* **89**, (2014).
4. K. Boshkayev and J.A. Rueda and R. Ruffini and B. Zhami and Z. Kalymova and G. Balgimbekov, *Fourteenth Marcel Grossmann Meeting - MG14* **4287**, (2018).
5. R. F. Tooper, *The Astrophysical Journal* **156**, (1969).
6. G. A. Carvalho and R.M. Marinho and M. Malheiro,*General Relativity and Gravitation* **38**, (2018).

# Polarized emission from strongly magnetized sources

Roberto Taverna*, Sergio Fabiani, Denis González Caniulef, Roberto Mignani, Fabio Muleri, Paolo Soffitta, Roberto Turolla and Silvia Zane

** Department of Mathematics and Physics, University of Roma Tre, Rome, Italy 00146*
*E-mail: taverna@fis.uniroma3.it*
*www.matfis.uniroma3.it*

Anomalous X-ray pulsars (AXPs) and soft-gamma repeaters (SGRs) form together a single class of astrophysical sources characterized by the emission of strong X-ray bursts and persistent emission with luminosity $10^{31}$–$10^{36}$ erg/s in the 0.2–10 keV energy range. These objects are commonly associated to magnetars, i.e. neutron stars endowed with ultra-strong magnetic fields. New-generation X-ray polarimeters like IXPE (NASA SMEX program), to be launched in 2021, will play a key role in assessing the nature of these sources by directly probing the star magnetic field. In fact, in the highly magnetized environment radiation is expected to be strongly polarized and such a measure will be easily within reach of IXPE. Polarization measurements will eventually confirm the presence of ultra-strong magnetic fields, probing the magnetar scenario. In this work I will discuss theoretical expectations, within the magnetar scenario, for the polarization signature of AXPs and SGRs and present numerical simulations for the response of the new-generation polarimeters currently under construction. I will also show how these sources can be used to test vacuum birefringence, a QED effect predicted by Heisemberg and Euler in the '30s and not experimentally verified as yet.

*Keywords*: Polarization – instrumentation: polarimeters – X-rays: stars – stars: magnetars.

## 1. Introduction

Soft gamma repeaters (SGRs) and anomalous X-ray pulsars (AXPs) are commonly considered to be different manifestations of the same class of isolated neutron stars (NSs), called magnetars.[1] They show a number of common distinctive properties, such as long (2–12 s) spin periods and large ($10^{-13}$–$10^{-10}$ s/s) spin-down rates, which lead to magnetic dipole fields up to $10^{14}$–$10^{15}$ G, and persistent X-ray luminosities ($10^{31}$–$10^{36}$ erg/s in the 0.2–10 keV range) that exceed the rotational energy loss rate (see e.g. Ref. 2 for a review). Persistent spectra are usually fitted by the superposition of a thermal (blackbody) component (with temperature $\approx 0.5$ keV) at lower energies and a power-law like tail (with photon index $\approx 2$–4) at higher ones. However, purely thermal spectra, characterized by the superposition of two BBs with similar temperatures, have been observed in the case of transient sources,[3] which undergo quite long ($\sim 1$ yr) outburst phases in which the persistent flux can be enhanced up to a factor of 1000 with respect to quiescence.

The most peculiar observational manifestation of AXPs and SGRs is the emission of short (0.1–1 s), energetic ($10^{38}$–$10^{41}$ erg/s) X-ray bursts and longer (1–100 s), even more energetic ($10^{41}$–$10^{43}$ erg/s) flares. In the case of three SGRs, giant flares,

characterized by peak luminosities up to $\approx 10^{47}$ erg/s and long-lasting (100–1000 s) pulsating tails, modulated at the spin frequency of the star, have been detected. This phenomenology has been successfully explained in terms of the twisted magnetosphere model,[4] which invokes the occurrence of resonant Compton scattering (RCS) onto electrons flowing along the NS twisted magnetic field lines. Moreover, according to this model short bursts and intermediate/giant flares can originate either in plastic displacements of the crust (deformed under the action of the internal field toroidal component) or in sudden reconnections of the field lines high in the magnetosphere.[5,6]

In the presence of magnetar-like magnetic fields, radiation emitted from the NS surface and propagating in vacuo is expected to be linearly polarized at a high degree.[7,8] The polarization state of propagating photons (which can change due to interactions with magnetospheric particles) depends on the dielectric and magnetic properties of the vacuum around the star. These are affected by vacuum birefringence, a quantum electro-dynamics (QED) effect predicted by Heisenberg & Euler in the 1930's[9] and never tested yet, which becomes important for magnetic fields in excess of the quantum critical field $B_{\rm Q} \simeq 4.414 \times 10^{13}$ G. New polarimetric techniques developed in the last years promise to open a new window in astrophysical observations of compact objects like magnetars,[10] improving the contributions of previous X-ray polarimetry missions[11] which failed in achieving conclusive results. In this work we discuss in particular the observational prospects of the NASA SMEX mission IXPE,[12] scheduled for launch in 2021, which will allow to test vacuum birefringence and probe the existence of ultra-strong magnetic fields around magnetars.

In section 2 we illustrate the theoretical model. The issue of the polarization properties of magnetar emission is addressed in section 3. Finally, we discuss the results of some numerical simulations performed until now for IXPE in section 4 and present our conclusions in section 5.

## 2. Theoretical model

According to the twisted magnetosphere model,[4] the strong (up to $\approx 10^{16}$ G) internal magnetic field of magnetars is expected to be highly wound-up, with a toroidal component of the same order of magnitude of the poloidal one. This can exert a magnetic stress on the conductive surface, that can deform the crust displacing small portions of the star surface, at which the external magnetic field lines are anchored. As a result, the external field acquires in turn a toroidal component, becoming twisted. Although the twist is reasonably localized into bundles of field lines, for the sake of simplicity we will consider in the following the case of a global twist, which is characterized by the twist angle

$$\Delta\phi_{\rm N-S} = \lim_{\theta\to 0} \int_{\theta}^{\pi} \frac{B_{\phi}}{\sin\theta B_{\theta}} {\rm d}\theta \,, \tag{1}$$

where $\boldsymbol{B} = (B_{\rm r}, B_\theta, B_\phi)$ is the star magnetic field expressed in polar components ($\theta$ is the magnetic colatitude). Magnetar bursts and flares could also originate in the crustal displacements induced by the strong internal magnetic field. In fact, according to the model discussed in Ref. 5 these deformations would inject in the magnetosphere an Alfvén pulse, which dissipates into an electron-positron pair plasma, magnetically confined within the closed field line region (the so-called "trapped fireball"). The presence of this fireball can also explain the pulsating tails observed in the spectra of many burst-emitter sources (see e.g. Ref. 6 and references therein).

Since a twisted magnetic field is non-potential (i.e. $\boldsymbol{\nabla} \times \boldsymbol{B} \neq 0$), currents must flow along the closed magnetic field lines, contrary to the case of ordinary NSs, where potential (dipolar) fields are believed to be more likely. For simplicity, we assumed that the charge carriers are electrons and ions lifted from the star surface (uni-directional flow, see Refs. 4, 13, 14). These particles make the medium in which photons propagate optically thick for RCS: a scattering occurs once the resonance condition is met, i.e. when the photon energy equals the particle cyclotron energy in the particle rest frame. In particular, for a surface magnetic field strength $\approx 10^{14}$ G one finds that photons scatter onto electrons typically at a distance $r_{\rm res} \approx 10\, R_{\rm NS}$. Numerical simulations[15] show that RCS onto magnetospheric electrons can indeed account for the power-law tails observed in the soft X-ray spectra of many objects.[2,16] Resonant scatterings onto ions may also occur; however, ions are much heavier than electrons and they are lifted at much smaller heights above the star surface. At variance with electrons, one can expect that scattering onto ions can rather give origin to narrow absorption features in the spectra.[15]

## 3. Polarization of radiation in strong magnetic fields

Photons emitted from the surface of an ultra-magnetized NS and propagating in vacuo turn out to be linearly polarized in two normal modes, the ordinary (O) and the extraordinary (X) ones, in which the photon electric field oscillates either parallel or perpendicular to the $\boldsymbol{k}$-$\boldsymbol{B}$ plane, respectively,[7,8] where $\boldsymbol{k}$ is the photon propagation direction. The evolution of the polarization state of photons during propagation is determined by the wave equation, which can be written as

$$\boldsymbol{\nabla} \times (\bar{\boldsymbol{\mu}} \cdot \boldsymbol{\nabla} \times \boldsymbol{E}) = \frac{\omega^2}{c^2} \boldsymbol{\epsilon} \cdot \boldsymbol{E}\,, \tag{2}$$

with $\boldsymbol{E}$ the photon electric field, $\omega$ their frequency, $\boldsymbol{\epsilon}$ the dielectric tensor of the medium in which photons propagate, $\bar{\boldsymbol{\mu}}$ the inverse of the magnetic permeability tensor and $c$ the speed of light. In general, $\boldsymbol{\epsilon}$ is expected to deviate from unity due to terms which account for plasma effects (e.g. collisions, radiation damping, etc., see Ref. 8). Photons can also change their polarization state upon scatterings with particles which flow along the closed magnetic field lines.[13–15] However, in the presence of strong magnetic fields further additional terms appear in both $\boldsymbol{\epsilon}$ and $\bar{\boldsymbol{\mu}}$, due to the vacuum polarization effect.[9] In fact, virtual electron-positron pairs, that are present in the vacuum around the star, are polarized by the strong NS field,

modifying the dielectric and magnetic properties of the vacuum. It can be shown that, close to the surface of magnetars, vacuum contributions in the dielectric tensor are by far dominant with respect to those of plasma (see e.g. Ref. 13 and references therein). Hence, neglecting the plasma terms in $\boldsymbol{\epsilon}$ one can reduce equation (2) to a simpler differential equation system,

$$\begin{aligned}\frac{\mathrm{d}Q}{\mathrm{d}z} &= -\frac{k_0\delta}{2}(2PV)\\ \frac{\mathrm{d}U}{\mathrm{d}z} &= -\frac{k_0\delta}{2}(N-M)V\\ \frac{\mathrm{d}V}{\mathrm{d}z} &= \frac{k_0\delta}{2}\left[2PQ+(N-M)U\right],\end{aligned} \tag{3}$$

where $k_0 = \omega/c$, $\delta = \alpha_\mathrm{F}\,(B/B_\mathrm{Q})^2/(45\pi)$ (with $\alpha_\mathrm{F}$ the fine-structure constant), $Q$, $U$ and $V$ are the photon Stokes parameters and $M$, $N$ and $P$ are coefficients that depend on the star magnetic field and the $\boldsymbol{\epsilon}$ and $\bar{\boldsymbol{\mu}}$ components (see e.g. Ref. 14).

Equations (3) show that, due to vacuum polarization effects, the photon electric field evolves along a typical scale-length $\ell_\mathrm{A} = 2/(k_0\delta)$, which depends on $B^{-2}$ and, through it, on the distance $r$ from the star. So, it should be compared with the scale-length $\ell_B = B/|\boldsymbol{k}\cdot\boldsymbol{\nabla}B|$ along which the star magnetic field itself changes along the photon trajectory. This comparison allows to divide the photon propagation region into three zones:[13,17] close to the star surface, in the *adiabatic region*, $\ell_\mathrm{A} \ll \ell_\mathrm{B}$, so that the photon electric field can instantaneously adapt to the star magnetic field (which is in turn evolving along the photon trajectory), leaving unchanged the original polarization state; outward, in the *intermediate region*, $\ell_\mathrm{A} \simeq \ell_\mathrm{B}$ and the photon electric field cannot adapt so promptly as before to the star magnetic field, causing a deviation of the polarization states with respect to the emission ones; finally, in the *external region* (far from the star surface) $\ell_\mathrm{A} \gg \ell_\mathrm{B}$ and the photon electric field is practically frozen with respect to the star magnetic field, so that the polarization states can change (also dramatically) with respect to those at the emission. In fact, if the magnetic field topology in the external region is still rather chaotic the observed polarization degree turns out to be strongly reduced (even for 100% intrinsic polarization degree), due to the Stokes parameter rotation.[17]

This scenario can be further simplified by introducing the adiabatic radius,[17] or polarization limiting radius,[13] i.e. the distance at which $\ell_\mathrm{A} = \ell_\mathrm{B}$,

$$r_\mathrm{a} \simeq 4.8\, B_{\mathrm{p},11}^{2/5} E_1^{1/5} R_{10}^{1/5} R_\mathrm{NS}\,, \tag{4}$$

where $B_{\mathrm{p},11}$ is the polar magnetic field strength in units of $10^{11}$ G, $E_1$ is the photon energy in units of 1 keV and $R_{10}$ is the NS radius in units of 10 km. Looking at equation (4) it can be seen that the adiabatic radius is larger for NSs with a stronger magnetic field and for photons with higher energy. Since the farther from the NS surface the more uniform the magnetic field topology, one can expect to observe more polarized X-ray photons for stronger magnetized NSs, with $r_\mathrm{a}$ placed at a sufficiently large distance from the star surface to prevent the original polarization pattern to be washed out by geometrical effects (see e.g. Ref. 17).

## 4. Numerical simulations

To implement our numerical simulations, we used the Monte Carlo code developed by Nobili, Turolla & Zane,[15] with the addition of a specific module to account for the polarization transport described in section 3. We used a globally twisted magnetic field, with the polar strength $B_{\rm p}$, the twist angle $\Delta\phi_{\rm N-S}$ and the corresponding velocity $\beta$ of the magnetospheric particles as input parameters. Since for the ultra-strong magnetic fields expected in the case of magnetars the free-free opacity for X-mode photons turns out to be much suppressed with respect to that of the O-mode ones, we assumed that the emitted photons are 100% polarized in the X-mode at the surface.[13] Furthermore, owing to the fact that, in the case of interest, the adiabatic radius largely exceeds the typical radius $r_{\rm res}$ at which RCS onto magnetospheric electrons may occur, we treated separately the change in polarization state of photons due to scatterings and the evolution of the Stokes parameters determined by vacuum effects. The outputs of the Monte Carlo code are eventually reprocessed in an IDL script to account for the viewing geometry, i.e. the angles $\chi$ and $\xi$ that the NS rotation axis makes with the observer line-of-sight (LOS) and the star magnetic axis, respectively.

We computed the observed flux, polarization degree and polarization angle (as functions of the photon energy and the rotational phase) in the case of the AXP 1 RXS J170849.0-400910 ($B_{\rm p} = 4.7 \times 10^{14}$ G, $\Delta\phi_{\rm N-S} = 0.5$, $\beta = 0.34$,

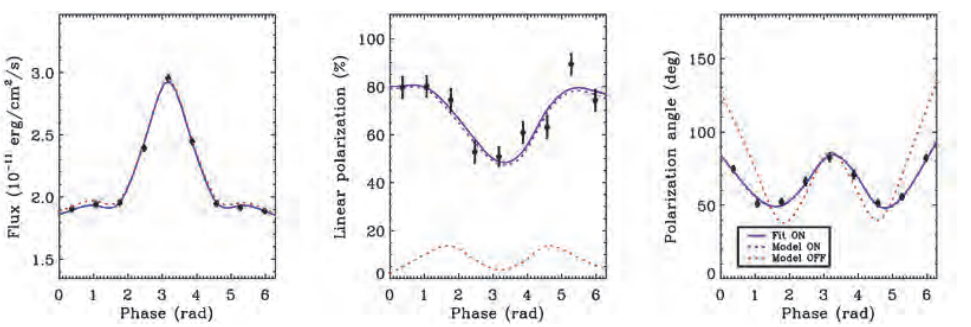



Fig. 1. Flux (left), linear polarization fraction (center) and polarization angle (right) simulated for a 250 ks observation with IXPE of the AXP 1 RXS J170849.0-400910, for $\chi = 90°$, $\xi = 60°$ and in the 2–6 keV range (filled circles with error bars). The model from which data are extracted (blue-dashed line) and that obtained for the same values of the input parameters but without accounting for QED effects (red-dashed line) are also shown. The best fit of data obtained using the entire model archive is marked by a blue, solid line.

Table 1. Input and best fit parameters of the simulation shown in Figure 1.

| | $\chi$ (deg) | $\xi$ (deg) | $\Delta\phi_{\rm N-S}$ (rad) | $\beta$ | $\chi^2_{\rm red}$ |
|---|---|---|---|---|---|
| Input | 90 | 60 | 0.5 | 0.34 | – |
| Best fit | $90.36 \pm 2.16$ | $59.12 \pm 1.21$ | $0.49 \pm 0.02$ | $0.37 \pm 0.07$ | 1.25 |

see Refs. 18, 19), and simulated the response of the IXPE polarimeter for this source (for $\chi = 90°$ and $\xi = 60°$) exploiting the instrumental effective area and modulation factor. The results are shown in Figure 1: the simulated data in the 2–6 keV energy range (filled circles with error bars) are plotted as functions of the rotational phase together with the model from which they are extracted (blue-dashed line) and the same model but with QED effects turned off (red-dashed line).[14,17] A best fit of the simulated data has been then performed, using an archive of models obtained leaving the input parameters $\chi$, $\xi$, $\Delta\phi_{N-S}$ and $\beta$ free to vary over a wide range of values. The fit (blue-solid line in Fig. 1) recovers the original input parameters within an acceptable degree of accuracy (within $1\sigma$, see Table 1), showing the strength of X-ray polarization measurements in extracting the physical and geometrical information of the source. Moreover, the QED-off case turns out to be ruled out at a high confidence level; this would allow us to say to have tested vacuum birefringence effects for the source at hand.

## 5. Discussion and conclusions

In this work we illustrated the commonly accepted theoretical model to explain the phenomenology of AXPs and SGRs, isolated NSs endowed with ultra-strong magnetic fields (magnetars). We focused in particular on the polarization of the X-ray photons emitted from these sources. In fact, radiation from magnetars is expected to be highly polarized and, due to their strong magnetic fields, these NSs are believed to be the best targets to test vacuum polarization effects, which has not been detected at the magnetic field strengths currently achievable in laboratory.

After having presented the twisted magnetosphere model and discussed the polarization evolution for photons propagating in the magnetized vacuum, we showed the output of our code in the case of the AXP 1 RXS J170849.0-400910, simulating the response of the NASA SMEX polarimeter IXPE, scheduled for the launch in 2021. Throughout our investigation, we assumed, for the sake of simplicity, that photons are emitted from the star surface 100% polarized in the X-mode. However, we point out that our code can also account for different intrinsic polarization patterns. A detailed analysis based on different, more realistic surface emission models is currently under development.

Our plots (see Fig. 1) show that X-ray polarization measurements are potentially able to distinguish between the expectations of our model with and without accounting for vacuum polarization effects. Besides to extract important information on the physics and geometry of the source (such as the values of the viewing angles $\chi$ and $\xi$ and the parameters $\Delta\phi_{\rm N-S}$ and $\beta$, see Table 1), X-ray polarimetry will allow to test QED effects in strong magnetic fields for the first time. Furthermore, since vacuum birefringence can prevent the geometrical depolarization of the collected radiation more effectively for stronger magnetized NSs, (see Sec. 3), the detection of a high degree of polarization can be considered as an indirect evidence for ultra-strong magnetic fields in magnetars.

## References


1. R. C. Duncan and C. Thompson, Formation of very strongly magnetized neutron stars - Implications for gamma-ray bursts, *ApJL* **392**, p. L9 (1992).
2. R. Turolla,S. Zane and A. L. Watts, Magnetars: the physics behind observations. A review, *Rep. Prog. Phys.* **78**, p. 116901 (2015).
3. N. Rea and P. Esposito, Magnetar outbursts: an observational review, *ASS Proc.* **21**, p. 247 (2011).
4. C. Thompson, M. Lyutikov and S. R. Kulkarni, Electrodynamics of Magnetars: implications for the persistent X-ray emission and spin-down of the soft gamma repeaters and anomalous X-ray pulsars, *ApJ* **574**, p. 332 (2002).
5. C. Thompson and R. C. Duncan, The soft gamma repeaters as very strongly magnetized neutron stars - I. Radiative mechanism for outbursts, *MNRAS* **275**, p. 255 (1995).
6. R. Taverna and R. Turolla, On the spectrum and polarization of magnetar flare emission, *MNRAS* **469**, p. 3610 (2017).
7. Yu. N. Gnedin and G. G. Pavlov, The transfer equations for normal waves and radiation polarization in an anisotropic medium , *JETP* **38**, p. 903 (1974).
8. A. K. Harding and D. Lai, Physics of strongly magnetized neutron stars, *Rep. Prog. Phys.* **69**, p. 2631 (2006).
9. W. Heisenberg and H. Euler, Consequences of Dirac's theory of the positron, *Zeitschr. Phys.* **98**, p. 714 (1936).
10. R. Bellazzini, E. Costa, G. Matt and G. Tagliaferri, X-ray polarimetry: A new Window in astrophysics, *Cambridge University Press, Cambridge* (2010).
11. M. C. Weisskopf, E. H. Silver, H. L. Kestenbaum, K. S. Long and R. Novick, A precision measurement of the X-ray polarization of the Crab Nebula without pulsar contamination, *ApJ* **220**, p. L117 (1978).
12. M. C. Weisskopf et al., in UV, X-Ray, and Gamma-Ray Space Instrumentation for Astronomy XVIII, *Proc. SPIE* **8859** (2013).
13. R. Fernández and S. W. Davis, The X-ray polarization signature of quiescent magnetars: effect of magnetospheric scattering and vacuum polarization, *ApJ* **730**, p. 131 (2011).
14. R. Taverna, F. Muleri, R. Turolla, P. Soffitta, S. Fabiani and L. Nobili, Probing magnetar magnetosphere through X-ray polarization measurements, *MNRAS* **438**, p. 1686 (2014).
15. L. Nobili, R. Turolla and S. Zane, X-ray spectra from magnetar candidates - I. Monte Carlo simulations in the non-relativistic regime, *MNRAS* **386**, p. 1527 (2008).
16. S. Mereghetti, The strongest cosmic magnets: soft gamma-ray repeaters and anomalous X-ray pulsars, *A&ARev* **15**, p. 225 (2008).
17. R. Taverna, R. Turolla, D. Gonzalez Caniulef, S. Zane, F. Muleri and P. Soffitta, Polarization of neutron star surface emission: a systematic analysis, *MNRAS* **454**, p. 3254 (2015).
18. S. A. Olausen and V. M. Kaspi, The McGill Magnetar Catalog, *ApJs* **212**, p. 6 (2014).
19. S. Zane, N. Rea, R. Turolla and L. Nobili, X-ray spectra from magnetar candidates - III. Fitting SGR/AXP soft X-ray emission with non-relativistic Monte Carlo models, *MNRAS* **398**, p. 1403 (2009).

# A white dwarf accretion model for the anomalous X-ray pulsar 4U 0142+61

Sarah V. Borges* and Claudia. V. Rodrigues

*Instituto Nacional de Pesquisas Espaciais, INPE, São José dos Campos, 12227-010, SP, Brazil*
** E-mail:villanovaborges@gmail.com*

Jaziel G. Coelho

*Departamento de Física, Universidade Tecnológica Federal do Paraná, UTFPR, Medianeira, 85884-000, PR, Brazil*

Manuel Malheiro

*Departamento de Física, Instituto Tecnológico de Aeronáutica, ITA, São José dos Campos, 12228-900, SP, Brazil*

Manuel Castro

*Instituto Nacional de Pesquisas Espaciais, INPE, São José dos Campos, 12227-010, SP, Brazil*

The persistent emission of the anomalous X-ray pulsar 4U 0142+61 extends over a broad range of energy, from mid-infrared up to hard X-rays. In particular, this object is unique among soft gamma-ray repeaters (SGRs) and anomalous X-ray pulsars (AXPs) in presenting simultaneously mid-infrared emission and also pulsed optical emission. In spite of having many propositions to explain this wide range of emission, it is still lacking one that reproduces simultaneously all the observations. Filling this gap, we present a model that is able, for the first time, to reproduce simultaneously the entire spectral energy distribution of 4U 0142+61 using plausible physical components and parameters. We propose that the persistent emission comes from an accreting white dwarf (WD) surrounded by a debris disk. This model is thoroughly discussed at Ref. 2 and assumes that: (i) the hard X-rays are due to the bremsstrahlung emission from the post-shock region of the accretion column; (ii) the soft X-rays are originated by hot spots on the WD surface; and (iii) the optical and infrared emissions are caused by an optically thick dusty disk, the WD photosphere, and the tail of the post-shock region emission. In this scenario, 4U 0142+61 harbors a fast-rotator near-Chandrasekhar WD, which is highly magnetized. Such a WD can be formed by a merger of two less massive WDs.

*Keywords*: Accretion, magnetic field, rotation, white dwarfs.

## 1. Introduction

4U 0142+61 is an Anomalous X-ray Pulsar (AXP) that presents quiescent emission in a broad range of energy, from mid-infrared up to hard X-rays. In particular, this object is unique among SGR/AXPs in presenting simultaneously mid-infrared emission and pulsed optical emission, which are rare features for the class. Its period is 8.68 s, the spin-down is around $2.0 \times 10^{-12}$ s.s$^{-1}$ and the soft X-rays luminosity is about $10^{35}$ erg.s$^{-1}$ [23].

The emission nature of AXP/SGRs is still reason for debate and several scenarios have been proposed to explain their observed spectra and properties, such as the magnetars[8], accreting NSs[28], quark stars[24], or WD pulsars[5,20]. All the

current models fail to explain the entire spectral range of 4U 0142+61. That problem is not exclusive of 4U 0142+61, since no scenario presents a complete model for the SGR/AXPs class. In this context, we propose that the persistent emission of 4U 0142+61 comes from an accreting isolated WD surrounded by a debris disk, having gas and dusty regions. This scenario is inspired by the periodic flux modulation and by the presence of mid-infrared emission, which is rare for NSs. In fact, apart from the SGR/AXP class in which only 1E 2259+586 and 4U 0142+61 have mid-infrared;[19,30], only three isolated NSs have detected mid-infrared: the radio pulsars Crab, Vela, and Geminga[6,27]. Thus, mid-infrared appears in about 0.3% of all isolated NSs. On the other hand, the presence of mid-infrared in WDs is quite common. Ref. 7 found that about 7% of all isolated WDs presents mid-infrared excess detected by *WISE*, which reinforces the WD origin for 4U 0142+61. This proceedings presents a study of 4U 0142+61 emission in the context of a WD nature. It is organized as follows. In section 2, we describe the WD accreting model we use to fit 4U 0142+61 data. In section 3, we show the spectral fit of 4U 0142+61. In section 4, we summarize our findings.

## 2. An accreting WD model for 4U 0142+61

We propose that the components of 4U 0142+61 persistent emission are the WD photosphere, a disk, and an accretion column. The disk is formed by a dusty external region and a gaseous internal region. The dusty disk is optically thick and emits such as a multi-temperature blackbody. The temperature of its inner radius is the grain sublimation temperature, which is about 1500 K for silicates. Conversely, the internal gaseous disk is optically thin and its emission can be neglected. The inner radius of the gaseous disk is equal to the magnetosphere radius. From that point on, the matter flows into the WD surface following the magnetic field lines and the debris disk ceases to exist.

Close to the WD photosphere, the in-falling flow of matter produces a shock, forming an extremely hot region, the so called post-shock region that emits bremsstrahlung. About half of that energy reaches the WD surface, where it is reprocessed, forming a hot spot. Once the high-energy emission for 4U 0142+61 is pulsed, with two peaks per phase, we assume that there are two accreting regions. Thus, we can express the total flux by:

$$F_{total} = F_{disk} + F_{wd} + F_{spot} + F_{brem}. \quad (1)$$

The WD photosphere ($F_{wd}$) and the hot spots ($F_{spot}$) emit as blackbodies, in which the intensity for a given wavelength $\lambda$ and temperature $T$ is the Planck function, $B(\lambda, T)$, whereas the post-shock region emits by thermal bremsstrahlung ($F_{brem}$). According to Ref. 21, the bremsstrahlung emitted power is:

$$P(\lambda, T_{brem}) = 2.051 \times 10^{-22} g_{ff} n_e^2 \lambda^{-1} T_{brem}^{-1/2} exp\left(\frac{-143.9}{\lambda T_{brem}}\right). \quad (2)$$

Assuming that the region is cylindrical, with a height $H_{brem}$, the optical depth of the bremsstrahlung emission, $\tau_{brem}$, is:

$$\tau_{brem} = \frac{H_{brem} P(\lambda, T_{brem})}{4\pi B(\lambda, T_{brem})}. \quad (3)$$

Assuming that the radius is $R_{brem}$, and the electron number density is $n_e$, the flux of the bremsstrahlung emission can be written as

$$F_{brem}(n_e, R_{brem}, H_{brem}, T_{brem}, d) = \\ = (1 - e^{-\tau_{brem}}) B(\lambda, T_{brem}) \pi \left(\frac{R_{brem}}{d}\right)^2. \quad (4)$$

At last, we have the multi-temperature disk component $(F_{disk})^4$:

$$F_{disk}(\nu, T_{in}, T_{out}, T_{wd}, R_{wd}, d) = 12\pi^{1/3} \cos(i) \left(\frac{R_{wd}}{d}\right)^2 \\ \times \left(\frac{2kT_{wd}}{3h\nu}\right)^{8/3} \left(\frac{h\nu^3}{c^2}\right) \int_{x_{in}}^{x_{out}} \frac{x^{5/3}}{e^x - 1} dx. \quad (5)$$

In this equation, $x = h\nu/kT$, where $T$ is the debris disk temperature, which ranges from $T_{out}$ to $T_{in}$ and $T_{wd}$ is the WD effective temperature. The model, as well as the parameters for each flux component, are described thoroughly in Ref. 2.

## 3. Fitting 4U 0142+61 SED

As the model parameters for each spectral region are not the same, we opted to fit spectral regions separately. To fit the SED of 4U 0142+61, we use the data presented in Figure 1. These data are dereddened and deabsorbed. We have used Markov Chain Monte Carlo MCMC -[14] to estimate the parameters and their uncertainties. The results are shown in Table 1 and Figure 1. We consider a distance of 3.78 kpc[2] and $N_H = 6.4 \times 10^{21}$ cm$^{-2}$.[9]

The fit quality of the hard X-rays increases for high bremsstrahlung temperatures, which can only be achieved for near-Chandrasekhar white dwarfs. We adopt a WD having mass of 1.41 $M_\odot$ and radius of 1021 km[3]. This corresponds to a bremsstrahlung temperature of 674.5 keV, which was considered as a fixed parameter.

After modelling the hard X-rays, we find the best fit for soft X-rays. The bremsstrahlung component is also included in the fit of the soft X-ray SED. To be consistent with the double peak in the soft X-rays light curve, we use two blackbodies components, which can have different temperatures and radii.

To fit the optical and infrared emission, we use the WD photosphere blackbody and the debris disk. The high-energy components are included in the fit and the tail of bremsstrahlung component from the post-shock region also contributes to the optical emission as shown in Figure 1. The flux of the disk is given by Eq. (5). We use the same values of $R_{wd}$ derived from the bremsstrahlung fit.

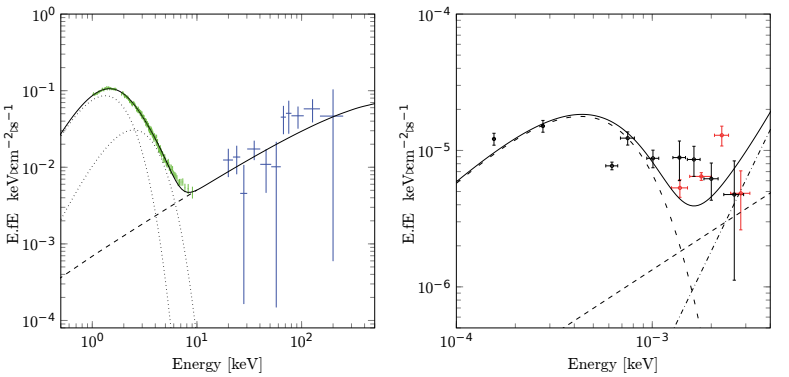



Fig. 1. Derreddened and deabsorbed SED of 4U 10142+61 along with the best fit. The continuous black curve is the complete fit, the long-dashed curve is the disk component, the dot-dashed curve is the WD photosphere, the two dotted curves are the hot spots components, and the short-dashed curve is the bremsstrahlung component. The black crosses are the data in low-energies used to fit the model and are from *Spitzer* (mid-infrared)[30], *Gemini* (near-infrared)[10] and *GTC* (optical)[22]; green crosses represent the soft X-ray data from *Suzaku*[11]; the blue crosses are the *INTEGRAL* data. *Left-bottom panel:* Zoom at the high-energy end. *Right-bottom panel:* Optical and infrared region. The red data[15,16] is displayed for comparison.

## 4. Conclusions

We obtained a good fit for the entire SED of 4U 0142+61. The optical/infrared emission of 4U 0142+61 comes from the WD itself and from the debris disk with a non-negligible contribution from the low-energy tail of the post-shock region. The hard X-rays is emitted by the accretion column and the soft X-rays by two hot spots in the WD photosphere.

The hard X-rays bremsstrahlung implies a near-Chandrasekhar WD, assumed to have a mass of 1.41 $M_{\odot}$ and a radius of 1021 km. Moreover, from the optical/infrared emission, we obtain an WD effective temperature of $9.4 \times 10^4$ K. Those radius and temperature point out to an young WD. The inner and outer disk temperatures are 1 991 K and 285 K.

In short, we were able to present a model that explains all the quiescent emission of 4U 0142+61, as well as the observed spin-down. Such a WD can be understood as the result of a recent merger of two less massive WDs.

Table 1. Parameters of the fitting of 4U 0142+61 in the accreting WD model.

| Parameter | Description | Value |
|---|---|---|
| | **X-rays** | |
| | Fixed parameters | |
| $d$ | distance of 4U 0142+61 | 3.78 kpc |
| $N_H$ | columnar density of hydrogen | 6.4 $10^{21}$ cm$^{-2}$ |
| $T_{brem}$ | temperature of the emission for the accretion column | 674.5 keV |
| $M_{WD}$ | WD's mass | 1.41 $M_\odot$ |
| $R_{WD}$ | WD's radius | 1,021 $10^5$ cm |
| | Fitted parameters | |
| $\dot{M}$ | accretion rate | 3.43 $10^{17}$g s$^{-1}$ |
| $R_{brem}$ | radius of the hard X-ray emission | 14.03 $10^5$ cm |
| $H_{brem}$ | height of the accretion column | 1.27 $10^5$ cm |
| $n_e$ | electrons number density | 2.05 $10^{19}$ cm$^{-3}$ |
| $\chi^2_{brem}/dof$ | reduced chi square for the hard X-rays | 0.85 |
| $T_{spot1}$ | temperature of the spot 1 | 0.632 $\pm$ 0.033 keV |
| $R_{spot1}$ | radius of the spot 1 | 2.35$\pm$0.45 $10^5$ cm |
| $T_{spot2}$ | temperature of the spot 2 | 0.337$\pm$0.012 keV |
| $R_{spot2}$ | radius of the spot 2 | 13.83$\pm$0.73 $10^5$ cm |
| $\chi^2/dof$ | reduced chi square for the soft X-rays | 1.06 |
| | **Optical/Infrared** | |
| | Fitted parameters | |
| $T_{WD}$ | WD's effective temperature | 9.4$\pm$7.3 $10^4$ K |
| $T_{in}$ | inner temperature of the debris disk | 1,991$\pm$16 K |
| $T_{out}$ | outer temperature of the debris disk | 285 $\pm$ 200 K |
| $R_{in}$ | inner radius of the debris disk | 2.35 $\pm$ 0.03 $R_\odot$ |
| $R_{out}$ | outer radius of the debris disk | $31^{+127}_{-16}$ $R_\odot$ |

**Note.** The fixed parameters were derived before the fit by independent methods. For the infrared/optical fit all the X-rays parameters are considered fixed, therefore, $R_{wd}$ is not a fitted parameter for this range of energy. The $1\sigma$ uncertainties for the last digit for the fitted parameters are in parenthesis.

## Acknowledgments

We acknowledge T. Enoto for kindly provide us with Suzaku soft X-ray data of 4U 0142+61.

The authors thank the grant #2013/26258-4, São Paulo Research Foundation (FAPESP). CVR acknowledge CNPq (Proc. 303444/2018-5). JGC is likewise grateful to the grant #2013/15088-0, São Paulo Research Foundation (FAPESP), and CNPq (Proc. 421265/2018-3 and 305369/2018-0). MC acknowledges financial support under grant #2015/25972-0 from São Paulo Research Foundation (FAPESP).

## References

1. Althaus, L. G., Córsico, A. H., Isern, J., & García-Berro, E. 2010, A&A Rev., 18, 471.
2. Borges S. V., Rodrigues C. V., Coelho J. G., Malheiro M., Castro M., 2020, Astrophysical Journal, 895, 26.

3. Carvalho, G. A., Marinho, R. M., & Malheiro, M. 2018, GeneralRelativity and Gravitation, 50, 38.
4. Chiang, E. I., & Goldreich, P. 1997, ApJ, 490, 368.
5. Coelho, J. G., & Malheiro, M. 2014, PASJ, 66, 14.
6. Danilenko, A. A., Zyuzin, D. A., Shibanov, Y. A., & Zharikov, S. V. 2011, MNRAS, 415, 867.
7. Debes, J. H., Hoard, D. W., Wachter, S., Leisawitz, D. T., & Cohen, M. 2011, ApJS, 197, 38.
8. Duncan, R. C., & Thompson, C. 1992, ApJ Lett., 392, L9.
9. Durant, M., & van Kerkwijk, M. H. 2006 ApJ, 650, 1082.
10. Durant, M., & van Kerkwijk, M. H. 2006 ApJ, 652, 576.
11. Enoto, T., Nakazawa, K., Makishima, K., et al. 2010, ApJ Lett.,722, L162.
12. Ferrario, L., Vennes, S., Wickramasinghe, D. T., Bailey, J. A., & Christian, D. J. 1997, MNRAS, 292, 205.
13. Girven, J., Brinkworth, C. S., Farihi, J., et al. 2012, ApJ, 749, 154.
14. Goodman, J., & Weare, J. 2010, Communications in Applied Mathematics and Computational Science, Vol. 5, No. 1, p. 65-80, 2010, 5, 65.
15. Hulleman, F., van Kerkwijk, M. H., & Kulkarni, S. R. 2000, Nature, 408, 689.
16. Hulleman, F., van Kerkwijk, M. H., & Kulkarni, S. R. 2004, A&A, 416, 1037.
17. Hurley, J. R., & Shara, M. M. 2003, ApJ, 589, 179.
18. Lodders, K. 2003, ApJ, 591, 1220.
19. Kaplan, D. L., Chakrabarty, D., Wang, Z., & Wachter, S. 2009, ApJ, 700, 149.
20. Malheiro, M., Rueda, J. A., & Ruffini, R. 2012, PASJ, 64, 56.
21. Mewe, R., Lemen, J. R., & van den Oord, G. H. J. 1986, A&AS, 65, 511.
22. Muñoz-Darias, T., de Ugarte Postigo, A., & Casares, J. 2016, MNRAS, 458, L114.
23. Olausen, S. A., & Kaspi, V. M. 2014, ApJS, 212, 6.
24. Ouyed, R., Leahy, D.,& Niebergal, B. 2011, MNRAS, 415, 1590.
25. Rafikov, R. R., & Garmilla, J. A. 2012, ApJ, 760, 123.
26. Rueda, J. A., Boshkayev, K., Izzo, L., et al. 2013, ApJ Lett., 772, L24.
27. Sandberg, A., & Sollerman, J. 2009, A&A, 504, 525.
28. van Paradijs, J., Taam, R. E., & van den Heuvel, E. P. J. 1995, A&A, 299, L41.
29. Wang, W., Tong, H., & Guo, Y.-J. 2014, Research in Astronomy and Astrophysics, 14, 673.
30. Wang, Z., Chakrabarty, D., & Kaplan, D. L. 2006, Nature, 440, 772.