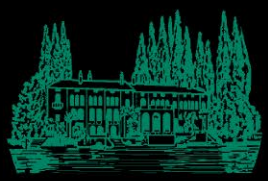


PROCEEDINGS
INTERNATIONAL SCHOOL OF PHYSICS «ENRICO FERMI»

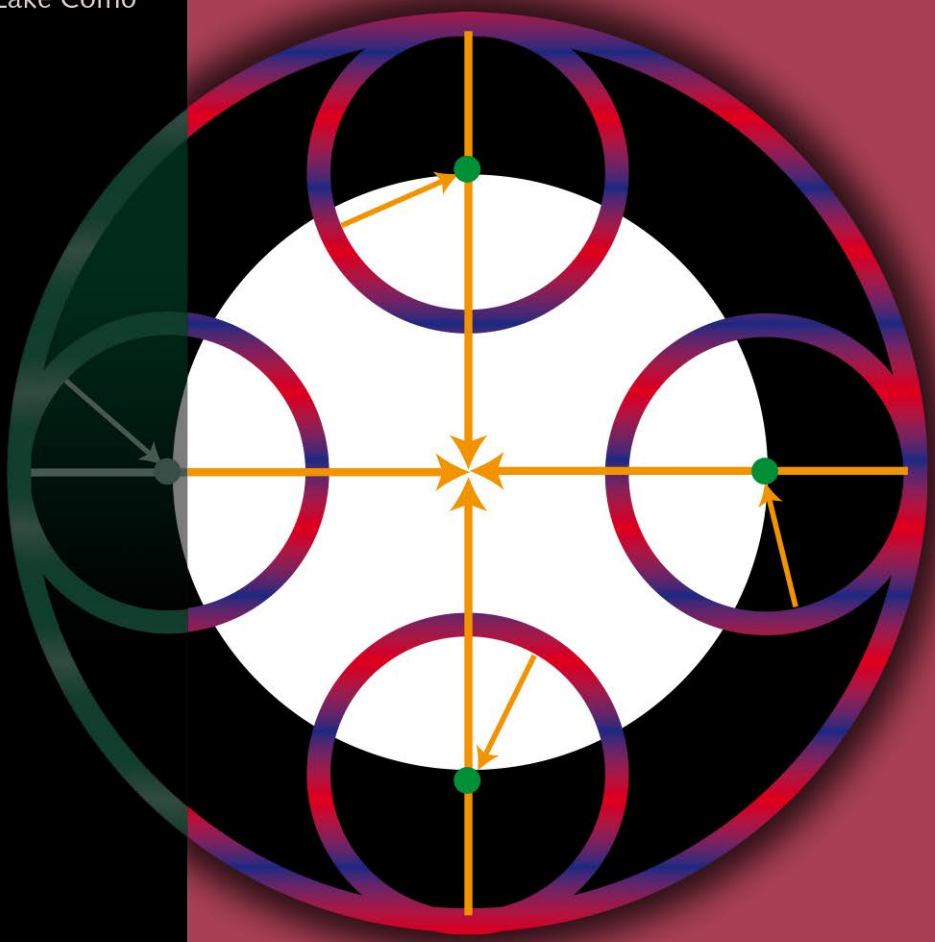
COURSE 200

Gravitational Waves and Cosmology

edited by E. Coccia, J. Silk and N. Vittorio

3 - 12 July 2017

Villa Monastero
Varenna, Lake Como



SOCIETÀ ITALIANA DI FISICA

SOCIETÀ ITALIANA DI FISICA

RENDICONTI

DELLA

SCUOLA INTERNAZIONALE DI FISICA

"ENRICO FERMI"

CC Corso

a cura di E. Coccia, J. Silk e N. Vittorio

Direttori del Corso

VARENNA SUL LAGO DI COMO

VILLA MONASTERO

3 – 12 Luglio 2017

Onde gravitazionali e cosmologia

2020

SOCIETÀ ITALIANA DI FISICA
BOLOGNA-ITALY

ITALIAN PHYSICAL SOCIETY

PROCEEDINGS

OF THE

INTERNATIONAL SCHOOL OF PHYSICS

"ENRICO FERMI"

COURSE 200

edited by E. COCCIA, J. SILK and N. VITTORIO

Directors of the Course

VARENNA ON LAKE COMO

VILLA MONASTERO

3 – 12 July 2017

Gravitational Waves and Cosmology

2020

IOS Press

AMSTERDAM - WASHINGTON DC

ISSN 0074-784X (print)
ISSN 1879-8195 (online)

ISBN 978-1-64368-094-1 (print) (IOS Press)
ISBN 978-1-64368-095-8 (online) (IOS Press)
ISBN 978-88-7438-121-0 (SIF)

LCCN 2020941150

DOI 10.3254/ENFI200

jointly published and distributed by:

IOS PRESS
Nieuwe Hemweg 6B
1013 BG Amsterdam
The Netherlands
fax: +31 20 687 0019
info@iospress.nl

SOCIETÀ ITALIANA DI FISICA
Via Saragozza 12
40123 Bologna
Italy
fax: +39 051 581340
order@sif.it

For book sales in the USA and Canada
IOS Press, Inc.
6751 Tepper Drive
Clifton, VA 20124
USA
Tel.: +1 703 830 6300
Fax: +1 703 830 2300
sales@iospress.com

Supported by

Camera di Commercio di Lecco

Istituto Nazionale di Fisica Nucleare (INFN)

Istituto Nazionale di Geofisica e Vulcanologia (INGV)

Univerlecco

Museo Storico della Fisica e Centro Studi e Ricerche “Enrico Fermi”

Istituto Nazionale di Ricerca Metrologica (INRiM)

Università di Roma Tor Vergata

Gran Sasso Science Institute (GSSI)

Produced by the SIF Editorial Staff

Production Editor: Marcella Missiroli

Cover: see Wayne Hu, *CMB polarization theory*, p. 217.

Graphic elaboration by Simona Oleandri

Printed in Italy by nuova MONOGRAF snc - Bologna

CONTENTS

The electronic version of volumes from 124 is available online at the IOS Press web site

http://ebooks.iospress.nl/bookseries/proceedings-of-the-international-school-of-physics-enrico-fermi

Figures with colour source files will appear in colour in the online version.

SIF Members are granted free access to these volumes at
https://members.sif.it

For a complete list of published courses see
https://en.sif.it/books/series/proceedings_fermi

This page intentionally left blank

Preface

The past twenty years have witnessed a number of breakthroughs in Astrophysics and Cosmology, that were awarded Nobel prizes: in 2019 "for theoretical discoveries in physical cosmology" and "for the discovery of an exoplanet orbiting a solar-type star"; in 2017 "for decisive contributions to the LIGO detector and the observation of gravitational waves"; in 2011 "for the discovery of the accelerating expansion of the Universe through observations of distant supernovae". How could we not mention the other two related breakthroughs in physics, also awarded with Nobel prizes: in 2015 "for the discovery of neutrino oscillations, which shows that neutrinos have mass"; in 2013 "for the theoretical discovery of a mechanism that contributes to our understanding of the origin of mass of subatomic particles, and which recently was confirmed through the discovery of the predicted fundamental particle, by the ATLAS and CMS experiments at CERN's Large Hadron Collider".

These physics triumphs serve to highlight our awareness that a Varenna School in "Gravitation and Cosmology" should provide the students with a solid and broad knowledge of the fundamentals of astrophysics and cosmology, without losing sight of the basics of the fundamental interactions in Physics. Thus, the aim of the School, held in the beautiful location of Villa Monastero from the 3rd to 12th of July 2017, was to expose students to state-of-the-art research in the field of Gravitational Waves and Cosmology, from both theoretical and experimental points of view. The choice of the subjects seemed particularly timely, given the discovery of gravitational waves by the LIGO-Virgo collaboration and the release of the Planck data in 2015, that so strongly contributed to the era of precision cosmology and to the adoption of a standard model of cosmology.

Twenty-four speakers participated in the School. We are very grateful to all of our colleagues, both to those who contributed written papers to this volume, as well as to those who participated in the School by just providing lectures. The level of the speakers was very high, with their presentations offering a broad overview of the subject matter.

The lectures were organized in a way designed to foster the interactions between two

different communities. This has been, in our opinion, one of the most notable added values of the School. The School admitted more than 50 PhD students, selected from all over the world, and also some Master students of the Erasmus Joint Master Program "Astromundus", funded by the EU and jointly provided by five European universities (Innsbruck, Gottingen, Padua, Tor Vergata and Belgrade). The environment of Villa Monastero naturally facilitates informal interactions between students and teachers. This added value of the "Enrico Fermi" Schools has been particularly important in enabling the case for integrating groups of students working in different fields and communities.

As will be obvious by going through the volume, there was a wide range of topics addressed by our lecturers. For the Gravitational Waves section, the lectures covered the experimental issues connected with gravitational wave detection and the new field of multi-messenger astronomy, as well as more astrophysical aspects. Lectures in the first category were provided by Fulvio Ricci, Viviana Fafone and Francesco Fidecaro, while the contributions of Marica Branchesi on multi-messenger astronomy and of Michela Mapelli about the open questions of black hole binaries covered the complementary aspects. In the realm of Cosmology, there are contributions on the early universe, on the Cosmic Microwave Background (CMB) and on redshift surveys. For the Early Universe, Jerome Martin provided a review of the inflationary scenarios, Sabino Matarrese discussed the non-Gaussian features of primordial density fluctuations, while Jens Chluba concentrated on the physical mechanisms responsible for the spectral distortions of the blackbody spectrum of the CMB. Regarding the CMB proper, the results of the Planck Legacy were presented by Carlo Burigana, and their theoretical implications were discussed by Douglas Scott, while the polarization modes of the CMB and the physical mechanisms underlying them were introduced by Wayne Hu. On the redshift survey side, Will Percival presented the wealth of information encoded in the large-scale structure of the universe, while David Mota reviewed the effects of modified gravity theories, with a particular focus on the effects on the non-linear regime on the structure formation process.

In conclusion, we are indebted to all our colleagues for their availability to participate in the School and for delivering outstanding lectures. We are also indebted to all the students that participated to the School with great enthusiasm and for expressing their positive opinion about the choice of the topics to which they were exposed. We are particularly thankful to the Italian Physical Society for hosting our School in the context of the program of the "Enrico Fermi" Schools, and for the professional support we received during our stay in Varenna.

EUGENIO COCCIA, JOE SILK and NICOLA VITTORIO

This page intentionally left blank

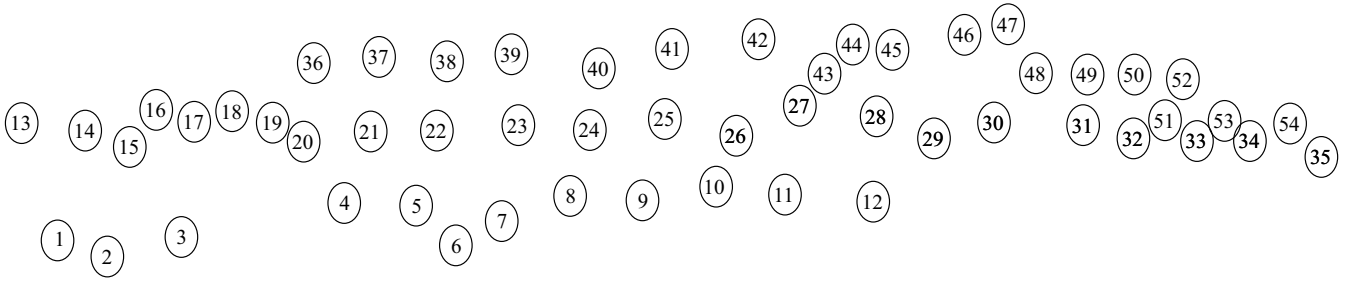
Italian Physical Society

INTERNATIONAL SCHOOL OF PHYSICS «E. FERMI»

COURSE 200

3 - 12 July 2017

VILLA MONASTERO − VARENNA, LAKE COMO



1 Irene Sartini
2 Barbara Alzani
3 Ramona Brigatti
4 Fulvio Ricci
5 Marica Branchesi
6 Alicia Sintes Jr.
7 Alicia Sintes
8 Eugenio Coccia
9 Gianluca Maria Guidi
10 Patrick Sutton
11 Douglas Scott
12 Wayne Hu
13 Andrew Miller
14 Sandeep Balakrishna Haridasu
15 Diana Lumaca
16 Vladimir Luković
17 Michael Kramer
18 Francesco Di Renzo
19 Lorenzo Galante
20 Francesco Fidecaro
21 Odysse Halim
22 Claudio Casentini
23 Francesco Cipriano
24 Lorenzo Aiello
25 Giampaolo Benevento
26 Maria Chiara Guzzetti
27 Matthias Dahlmanns
28 Andrea Ravenni
29 Oleg Kalinin
30 Marcus C. Werner
31 Alessandro Buzzelli
32 Arpine Kozmanyan
33 Huanchen Hu
34 Marco Celoria
35 Kirtika Juhi Hurgobin
36 Robert Hagala
37 Alessio Cirone
38 Rishikesh Pandit
39 Matteo Bonetti
40 Eugenio Noda
41 Jay Vijay Kalinani
42 Jose Maria Ezquiaga
43 Eleonora Castelli
44 Carlo Burigana
45 Chris Pattison
46 William Wright
47 Jacques Delabrouille
48 Mark Linton
49 Federico Ferrini
50 Joseph Silk
51 Jerome Martin
52 Nicola Vittorio
53 Viviana Fafone
54 Sabino Matarrese

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200002

Principles of gravitational wave detection

F. Fidecaro

Dipartimento di Fisica "Enrico Fermi", Università di Pisa e INFN, Sezione di Pisa
Largo Bruno Pontecorvo 3, 56127 Pisa, Italy

Summary. — The principles of gravitational wave detection are presented. The measurement of the propagating minute deformation of space-time remains an extraordinary experimental challenge. After a brief introduction to signals and noise, fundamental disturbances in the measurement are discussed, outlining how these have been sufficiently reduced to achieve detection.

1. – The detection of gravitational waves

The General Theory of Relativity [1], formulated by Albert Einstein more than one century ago, has proven to be very successful in describing gravitation as a property of space-time. As discussed in previous lectures one of the outstanding predictions of the theory is that the curvature of space-time due to the presence of mass does not establish itself instantaneously everywhere but propagates from the source outwards at the speed of light. This prediction was formulated by Einstein [2] discussing solutions of the General Relativity equations for small deviations from the flat (Minkowski) space-time, that is the one of Special Relativity. However only in 1957 at a conference in Chapel Hill, Feynman, elaborating an idea by Pirani, convinced the audience that gravitational

waves were carrying energy and thus were real and detectable [3]. The extremely small amplitude of the waves led Einstein to conclude

> [...] so sieht man daß A in allen nur denkbare Fällen ein praktisch verschwindenden Wert haben muß, [...] A. Einstein [2]
>
> [...] one sees that [the amplitude] A must have in all cases, even only thinkable, a practically vanishing value [...]

With the knowledge available that was the only conclusion to be drawn. Crucial discoveries were yet to come: with the neutron very compact astrophysical objects could be imagined, able to irradiate waves of significant amplitude. Then the availability of coherent light from lasers suggested gains of orders of magnitude in measurement precision. This stimulated a 50 year long experimental effort that culminated in the detection in 2015 by the LIGO interferometers of the inspiral and merge of two black holes of 36 and 29 solar masses [4].

The text begins with a discussion on how variations in the geometry of space-time can be detected, continuing with the basic properties of gravitational radiation in relation with the expected sources. Then follows a discussion on how to characterize noise when measuring signals. The final part describes the main problems encountered in interferometric gravitational wave detection and how they have been addressed, outlining the evolution of present and future instruments.

1˙1. *Gravitational waves.* – The Einstein field equations state how masses generate the curvature of space-time, connecting the Einstein tensor of curvature $G^{\alpha\beta}$ to the stress-energy tensor $T^{\alpha\beta}$ through the Newton constant $\mathrm{G} = 6.674 \times 10^{-11}\,\mathrm{kg}^{-1}\,\mathrm{m}^3\,\mathrm{s}^{-2}$:

$$G^{\alpha\beta} = \frac{8\pi\mathrm{G}}{\mathrm{c}^4} T^{\alpha\beta}.$$

These equations contain derivatives to be computed in space-time, whose curvature is determined by the equations themselves. Only a few exact solutions are known and in the general case one has to rely on approximations or on numerical solutions. In spite of the strong action of gravitation on macroscopic bodies, the common experience is that Euclidean geometry is verified at the highest level of precision.

This leads naturally to consider small perturbations to a flat space-time whose invariant distance element would be otherwise

$$\mathrm{d}s^2 = -\mathrm{c}^2\mathrm{d}t^2 + \mathrm{d}x^2 + \mathrm{d}y^2 + \mathrm{d}z^2 = \eta_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta.$$

In curved space-time

$$\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha \mathrm{d}x^\beta,$$

having the metric tensor $g_{\alpha\beta}$ deviating only slightly from the Minkowski space one $\eta_{\alpha\beta}$:

$$g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta}, \qquad |h_{\alpha\beta}| \ll 1.$$

It is worth noting that while $g_{\alpha\beta}$ is a tensor in curved space-time, $h_{\alpha\beta}$ is not. However $h_{\alpha\beta}$ behaves like a Lorentz tensor in a "background" Minkowski space.

Einstein's equations can be expanded to first order in h: following for example [5], one sets the trace of $h^{\alpha\beta}$

$$h = h^{\alpha}{}_{\alpha}$$

using the Minkowski metric tensor $\eta_{\alpha\beta}$ to lower and raise indices. It is useful to define

$$\bar{h}^{\alpha\beta} = h^{\alpha\beta} - \frac{1}{2}\eta^{\alpha\beta}h.$$

Within the freedom in coordinate definition given by General Relativity, it is possible to define coordinate transformations, called gauge transformations, that keep invariant the tensor describing the curvature. By setting

$$\bar{h} = 0, \qquad \bar{h}^{\alpha\beta}{}_{,\alpha} = \partial_{\alpha}\bar{h}^{\alpha\beta} = 0,$$

one chooses a "Transverse Traceless gauge" that allow to simplify considerably Einstein's equations. Using the D'Alembert operator $\Box = \partial^{\alpha}{}_{\alpha} = \partial_x^2 + \partial_y^2 + \partial_z^2 - 1/\mathrm{c}^2\partial_t^2$, the result is a wave equation

$$\Box\,\bar{h}^{\alpha\beta} = -\frac{16\pi\mathrm{G}}{\mathrm{c}^4}T^{\alpha\beta},$$

with source given by the stress-energy tensor $T^{\alpha\beta}$. The metric perturbation $\bar{h}^{\alpha\beta}$ originated by $T^{\alpha\beta}$ propagates at the speed of light c.

In this coordinate system, a plane wave travelling in the z direction with frequency Ω and corresponding wave number k is described by

$$\bar{h}^{TT}_{\alpha\beta} = A^{TT}_{\alpha\beta}e^{-i(\Omega t - kz)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & A_{xx} & A_{xy} & 0 \\ 0 & A_{yx} & -A_{yy} & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} e^{-i(\Omega t - kz)}.$$

There are two independent amplitudes $A_+ = A_{xx} = -A_{yy}$ and $A_\times = A_{xy} = A_{yx}$ corresponding to two polarization states + and $\times$. $A^{TT}_{\alpha\beta}$ is invariant under a rotation of π radians, which means in terms of rotation representations that a spin 2 is to be assigned to $A^{TT}_{\alpha\beta}$ or to the graviton, the quantum of the gravitational field. Furthermore, while a spin 2 massive particle has 5 different values for the spin projection, with only two degrees of freedom available, corresponding to opposed helicities, a zero mass has to be assigned to the graviton.

1˙2. *Effect on a single mass.* – The choice of the Transverse Traceless (TT) gauge determines the coordinate system used to describe the motion of a mass. It is interesting and useful to see how a mass behaves in presence of a gravitational wave. Considering the equation for a body that follows a geodetic line, or geodetic equation,

$$\frac{\partial^2}{\partial\tau^2}x^\alpha + \Gamma^\alpha_{\beta\gamma}\frac{\partial x^\beta}{\partial\tau}\frac{\partial x^\gamma}{\partial\tau} = 0,$$

one can compute the Cristoffel symbol $\Gamma^\alpha_{\beta\gamma}$ using

$$\Gamma^\alpha_{\beta\gamma} = \frac{1}{2}g^{\alpha\rho}\left[g_{\beta\rho,\gamma} + g_{\gamma\rho,\beta} - g_{\beta\gamma,\rho}\right]$$

With a mass initially at rest, $\vec{u} \to (-1, 0, 0, 0)$, only the value of Γ^α_{00} is useful. With the GW metric,

$$\Gamma^\alpha_{00} = 0.$$

A mass initially at rest, or better, with constant coordinates, will maintain the same coordinates.

1˙3. *Effect on a pair of masses.* – Consider two masses at some distance one from the other. Their separation can be measured using a light ray going from one mass toward the other and bouncing back to the first mass. A clock in the local frame of the first mass can then measure the round trip time, giving the distance. For events of emission and reception of a light ray

$$\mathrm{d}s^2 = g_{\alpha\beta}\mathrm{d}x^\alpha\mathrm{d}x^\beta = 0.$$

This can be applied to the case of a gravitational wave propagating along the z-direction in an otherwise flat space by writing:

$$\mathrm{d}s^2 = -\mathrm{c}^2\mathrm{d}t^2 + (1+h_+)\mathrm{d}x^2 + 2h_\times\mathrm{d}x\,\mathrm{d}y + (1-h_+)\mathrm{d}y^2 + \mathrm{d}z^2 = 0.$$

Specializing further to the wave + polarization and masses along the x-direction, always in the TT gauge, this becomes:

$$0 = \mathrm{c}^2\mathrm{d}t^2 - (1+h_+)\mathrm{d}x^2$$

For an infinitesimal separation, $h_+(t)$ can be taken constant and the clock measures a round trip time

$$\mathrm{c}\mathrm{d}t = 2\sqrt{1+h_+(t)}\mathrm{d}x = 2(1+h_+(t)/2)\mathrm{d}x.$$

Repeating the measurement allows to follow the value of $h_+(t)$ with time, a varying metric being the signal of a wave passage.

For a finite separation L between masses, the previous expression can be integrated along the unperturbed ray path, given by

$$x(t) = x_0 + \mathrm{c}t$$

neglecting second order effects. For $h_+(t)$ approximately constant during the round trip, that is for a gravitational wavelength $\lambda_{GW} \gg L$

$$t_1 - t_0 = \frac{2L}{\mathrm{c}} \left[1 + \frac{h_+(t)}{2}\right].$$

Without loss of generality one can consider a monochromatic plane wave $h_+(t, z) = h\cos(\Omega t + \phi)$. The integration along the previous path gives

$$t_1 - t_0 = \int_{x=0}^{x=L} \left(1 + \frac{h_+}{2}\right) \mathrm{d}x = \frac{2L}{\mathrm{c}} \left\{1 - \frac{h}{2} \frac{\sin(\Omega L/\mathrm{c})}{\Omega L/\mathrm{c}} \cos\left[\Omega\left(t_1 - L/\mathrm{c}\right) + \phi\right]\right\}.$$

The resulting time measurement reproduces the shape of the wave at the time light reaches the second mass but for a frequency-dependent attenuation factor

$$\mathrm{sinc}\left(\Omega L/\mathrm{c}\right) = \frac{\sin(\Omega L/\mathrm{c})}{\Omega L/\mathrm{c}} = \frac{\sin(2\pi L/\lambda_{GW})}{2\pi L/\lambda_{GW}}.$$

If $\Omega L/\mathrm{c} \ll 1$, or $2\pi L/\lambda_{GW} \ll 1$ the round trip time reproduces faithfully the signal shape. Otherwise, the high frequency components are attenuated.

Within the limits just discussed it is possible to increase the response to h increasing L, as the effect is in first instance equal to

$$L\frac{h}{2}.$$

In a multiple delay line one can have N round trips and the same results apply, with NL in place of L.

Note that if $\Omega 2NL/\mathrm{c} = k\pi$, with k integer, that is $2NL = k\lambda_{GW}$, the response is zero: the variation in round trip time accumulated when $h_+(t) > 0$ is reduced when $h_+(t) < 0$ and becomes zero when the above condition is verified, that is when $f_{GW} = k\mathrm{c}/2NL$; the response remains within an envelope proportional to $1/\Omega$. It will be shown when discussing the Fabry-Perot optical cavity, where *on average* light travels a distance $2NL$, the response to a monochromatic gravitational wave does not have zeroes any more, as they are smoothed out in the averaging process, but the envelope remains. The light stays on average a time $\tau_{FP} = 2NL/\mathrm{c}$ so the cavity acts as a low-pass filter with pole at $f_P = 1/4\pi\tau_{FP}$.

The method can be described in graphical form by drawing the world lines of the two masses in the TT gauge. Since the coordinates of each mass do not change, the world lines are parallel to the $x = 0$ line, but the light trajectory has a slope that depends on

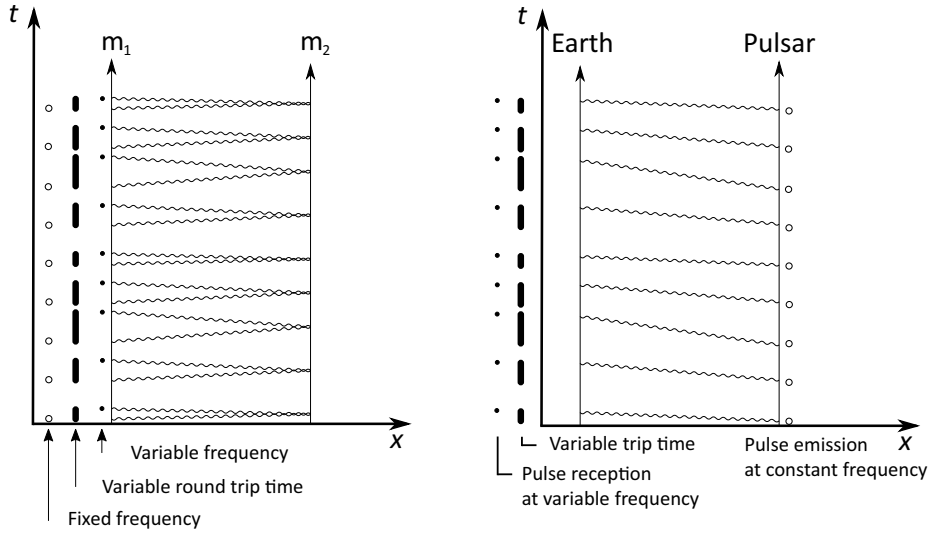


Fig. 1. – Left: space-time diagram for two free masses and a light ray bouncing back and forth between them. Coordinates are in the TT gauge. m_1 and m_2 have fixed spatial coordinates and light is emitted by m_1 at constant frequency. The apparent speed of light, the slope of the rays, changes with $h(t)$, resulting in a variable round trip time and a variable reception frequency. Right: one way light trip from a pulsar. The result is a variable reception frequency.

the value of h_x. As shown in fig. 1, left, timings are measured in a single place so that comparisons with a reference can be easily made. In a Michelson interferometer, the light is sent out in perpendicular directions and returns to the same place. There interference by spatial superposition of the beams allows to compare travel times looking at optical phase differences.

By having stable clocks attached to both masses it is possible to make measurements using one way light transmission. The timing relies on having pulses sent in a precisely known way, for example with a well defined period. Variations in mass separation are then measured comparing the time of arrival with a highly stable clock (fig. 1, right). Radiotelescopes on Earth are receiving signals from pulsars that show a stability comparable to the best atomic clocks built. Correlations between different sources of timing noise could point to low-frequency gravitational waves crossing our Galaxy. The use of Pulsar Timing Arrays (PTA) is discussed for example in [6] and references therein.

1˙4. *The laboratory frame*. – The detection method illustrated above is easy to understand for free falling masses that follow their world line. But in the ground-based laboratory masses are suspended and are anyway subject to forces from the environment through chemical bonds, van der Waals forces, electric or magnetic fields. Ultimately these are all electromagnetic interactions which are more easily described in the absence of gravitation, that is in a flat Minkovski space. For example the rest length of a spring

is determined by the electromagnetic interaction between its components that will result in a difference in coordinates for the two extremities in flat space. Hooke's law is then easily described in that space, and this is true for other forces too. For an observer in free fall, space-time is flat, so this is the natural choice for the laboratory reference frame (static gravity is not a concern, once masses are suspended).

This brings to choose a reference frame in free fall, where one of the masses is at rest. In that frame and for slow motion Newton's second law can be used, determining the motion of the other test masse in that frame, including the "apparent force" (in classical mechanics language) from gravitation. For that purpose one starts from the expression that describes how two nearby test masses change their separation while following their world line in a curved space-time:

$$\mathrm{D}_U \mathrm{D}_U \xi^\alpha = R^\alpha{}_{\mu\nu\beta} U^\mu U^\nu \xi^\beta,$$

where $U^\mu \to (-c, 0, 0, 0)$ is the 4-velocity of the local reference frame, D_U is the covariant derivative along U and $R^\alpha_{\mu\nu\beta}$ is the Riemann curvature tensor. In flat space the covariant derivatives become normal derivatives with respect to the proper local time τ. The equation becomes

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau} \xi^\alpha = R^\alpha{}_{\mu\nu\beta} U^\mu U^\nu \xi^\beta = R^\alpha{}_{00\beta} \xi^\beta.$$

For a separation from the origin of the second mass $\xi \to (0, \varepsilon, 0, 0)$ one has

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau} \varepsilon = R^\alpha{}_{00x} \varepsilon = -R^\alpha{}_{0x0} \varepsilon.$$

The components of the Riemann tensor, being gauge invariant, can be computed in the TT gauge, giving

$$\begin{aligned}
R^x{}_{0x0} &= R^{x0x0} = -\frac{1}{2} h^{TT}_{xx,00}, \\
R^y{}_{0x0} &= R^{y0x0} = -\frac{1}{2} h^{TT}_{xy,00}, \\
R^y{}_{0y0} &= R^{y0y0} = -\frac{1}{2} h^{TT}_{yy,00} = R^{x0x0}.
\end{aligned}$$

In Newton dynamics terms an apparent force changes the separation of the second mass from the first according to

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} \xi^x &= \frac{1}{2} h^{TT}_{xx,00} \varepsilon, \\
\frac{\mathrm{d}^2}{\mathrm{d}\tau^2} \xi^y &= \frac{1}{2} h^{TT}_{xy,00} \varepsilon.
\end{aligned}$$

If the mass m_1 feels a non-gravitational force F_1 while the mass m_2 is at $(0, \varepsilon, 0, 0)$, these can be written as an equation of motion for the mass separation

$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\xi^x = \frac{1}{2}h^{TT}_{xx,00}\varepsilon + \frac{F_1^x}{m_1} - \frac{F_2^x}{m_2}\,,$$
$$\frac{\mathrm{d}^2}{\mathrm{d}\tau^2}\xi^y = \frac{1}{2}h^{TT}_{yx,00}\varepsilon + \frac{F_1^y}{m_1} - \frac{F_2^y}{m_2}\,.$$

The rigid rulers of the laboratory frame that are used as coordinate axes can be modeled by masses kept at a fixed distance one from the other by electric interaction and quantum mechanics, which are Lorentz invariant. So the length of the spring must be the proper length in any reference frame. Changes in the separation between the extremities of the spring are changes in the proper length that call for a restoring force according to Hooke's law. Since the electric forces that bind atoms together are way much stronger than the apparent force of geodetic deviation, the ruler stays rigid and the second mass moves against it.

In the lab frame of the first mass (not in TT gauge) the passage of the gravitational wave corresponds to a displacement of the second mass

$$\delta x = L\frac{h}{2}\frac{\sin \Omega L/c}{\Omega L/c}\cos\left[\Omega\left(t_1 - L/c\right) + \phi\right].$$

For a wave with amplitude $h \approx 10^{-21}$ and a separation of 3 km

$$\delta x = L\frac{h}{2} \approx 1.5 \times 10^{-18}\,\mathrm{m}.$$

This length is three orders of magnitude smaller than the proton diameter, and has to be measured between the center of mass of macroscopic objects, that should be as far apart as possible to increase the effect of the gravitational wave.

2. – Essential properties

Among fundamental interactions, gravitation is by far the weakest, when compared to the interactions of components of the standard model of particle physics. The ratio of the electric to gravitational force between two electrons is

$$\frac{\mathrm{G}m_e^2}{\mathrm{k}e^2} = \frac{6.67 \times 10^{-11}}{8.99 \times 10^9}\left(\frac{9.1 \times 10^{-31}}{1.6 \times 10^{-19}C}\right)^2 \approx 10^{-42}.$$

Gravitation becomes dominant when the total mass involved becomes larger; there are no "negative" masses that would give "gravity-neutral" bodies, while these same bodies are stable and have zero charge.

The solution of the wave equation in the non-relativistic case is

$$h_{jk} = \frac{\mathrm{G}2}{\mathrm{c}^4 R} \frac{\mathrm{d}^2 Q_{jk}}{\mathrm{d}t^2} ,$$

where Q_{jk} is the traceless quadrupole moment of the system computed at the retarded time $t - R/\mathrm{c}$. In a way similar to electromagnetism, gravitational radiation is produced by accelerated masses, with an amplitude proportional to the coupling constant G, the radiating mass and its acceleration. A pair of rotating masses in the laboratory irradiates gravitationally, and as consequence of quadrupole emission, the frequency of gravitational radiation is twice the rotation speed of the system. When using real numbers and pushing stress in materials at the limit of disruption, the computed amplitude is many orders of magnitude lower than the noise in current detectors. In addition, the requirement to be in the radiation zone at a distance much larger than one wavelength lowers further the amplitude at the detector.

Possible sources have to be much more massive and their motion can only be driven by gravitation. High acceleration is obtained only if the object is compact, that is has a size much lower than the radiation wavelength. Systems of compact objects such as black holes and neutron stars were already known and their gravitational radiation has been recently detected, but these systems are rare and the rate of these observed events is low. To accumulate even a few events per year the volume of Universe hosting potential sources has to be increased, attempting to record events that occurred at large distance R: the volume grows as R^3, neglecting cosmological effects and variation in the population with time. But the gain in mass and acceleration from solar mass objects is much reduced by the distance R from the source as the wave amplitude goes as R^{-1}. A discussion of detectable sources of gravitational waves needs a brief description of the scale of distances in the Universe.

2˙1. *Distance ladder*. – While angles in the sky are measured with increasing precision and larger telescopes achieve impressive angular resolutions, the measurement of astronomical distances is a difficult problem. Indeed, by observing the variation of the position in the sky of a star while the Earth orbits around the Sun, the effect of the parallax allows to reconstruct geometrically its distance. The parsec is the unit of distance corresponding to the distance of an object that undergoes a parallax error of one arcsecond when the Earth changes its position by one Astronomical Unit. To extend the distance measurements astronomers rely on indirect methods like measuring the energy flow from of standard candles whose luminosity is evaluated by measuring some other observable parameter. A single method is not able to cover the full range of observed distances, the distance ladder is extended in steps, leading to systematic errors. As of today, depending on which "ladder" is used, there are appreciable differences in distance estimation. Further the possible contribution of gravitational wave detection to the cosmic distance ladder determination will be discussed. For the present introduction a rough view is presented in table I.

TABLE I. – *The cosmic distance ladder.*

Quantity	Value	Note
1 Astronomical Unit (AU)	1.50×10^{11} m exactly 149597870700 m	Approximately Sun-Earth distance
1 light-year	0.946×10^{16} m	
1 parsec (pc)	3.086×10^{16} m	3.262 light-year: observed parallax error of one arcsecond when the Earth moves transversally by 1 AU
Galactic center	10 kpc	High stellar density, massive black hole
Galaxy diameter	$\sim$ 30 kpc	There are of the order of 10^9 neutron stars in our Galaxy
Galaxies of the Local Group	50 kpc–1.4 Mpc	50 galaxies
Virgo galaxy cluster	$\sim$ 20 Mpc	2500 galaxies
GW170817	$\sim$ 40 Mpc	Host galaxy NGC 4993
GW150914	$\sim$ 410 Mpc	36–29 $M_\odot$ binary black holes
Horizon	$\sim$ 5 Gpc	15 billion light-year

The distance of the binary black hole event GW150914 was determined from the signal amplitude, once the masses were determined from the signal shape, but in absence of an electromagnetic counterpart, this events and others of the same type provide information on black hole population only. The distance of the binary neutron star event GW170817 has been determined by associating the event to an optical transient found in the three dimensional error box given by the gravitational wave network of LIGO and Virgo and associated to the host galaxy NGC 4993. In this case gravitational waves provide an independent measurement of distance that can be used to improve Hubble's law.

2˙2. *Expected amplitude*. – When considering the static case of the Schwarzschild metric, the deviation h from flat space-time generated by the mass M is

$$h = \frac{2\mathrm{G}}{\mathrm{c}^2}\frac{M_\odot}{R}\,,$$

or about 4×10^{-6} at the surface of the Sun ($R = 3.5 \times 10^8$ m). The same mass generates at 10 Mpc, or 3×10^{23} m, a static $h \approx 4 \times 10^{-21}$.

For gravitational wave emission by accelerated masses in a non-relativistic system a rough estimate is obtained through [5]:

$$\begin{aligned} h_{jk} &= \frac{2\mathrm{G}}{\mathrm{c}^4 R}\frac{\mathrm{d}^2 Q_{jk}}{\mathrm{d}t^2} \\ &\lesssim \frac{2\mathrm{G}}{\mathrm{c}^2 R}\frac{2Mv^2}{\mathrm{c}^2 r} \\ &\sim \frac{2\mathrm{G}M}{\mathrm{c}^2 R}\frac{2\mathrm{G}M}{\mathrm{c}^2 2r} \\ &= \frac{r_S}{R}\frac{r_S}{d} \\ &= \phi_{ext}\phi_{int}. \end{aligned}$$

The second line is obtained going from the second time derivative of the quadrupole moment to the expression of the force for circular motion. Then the kinetic energy term, divided by the mass, is substituted by the gravitational potential $\phi_{int} = \mathrm{G}M/r$ of one mass making use of the virial theorem. r_S is a typical Schwarzschild radius of the system and d is the distance between the two masses. ϕ_{ext} is the gravitational potential of the system at the observer, which scales inversely with the distance, resulting in a simple formula for order of magnitude estimates.

Compact binary white dwarf systems have been identified by astronomers, who can measure directly the orbital frequency. From these data the amplitude of the wave can be deduced, as discussed in [7].

$$h \simeq \frac{4(\mathrm{G}\mathcal{M})^{5/3}}{\mathrm{c}^4 R}\left(\frac{\omega}{2}\right)^{2/3},$$

where $\mathcal{M} = \mu^{3/5}M^{2/5}$ is the chirp mass, and M is the total mass of the system.

For binary black holes of $30\,M_\odot$ ($r_S = 3\,\mathrm{km}$) at 400 Mpc, forming a system of size 600 km

$$h \sim 10^{-21}.$$

One has to consider such a distance because finding a binary system of two black holes that coalesce right at the time of observation is very rare at the level of a single galaxy. The order of magnitude of the rate of Binary Black Hole mergers on the base of the observed events has been estimated to be 2–600 $\mathrm{Gpc}^{-3}\,\mathrm{yr}^{-1}$.

2˙3. *Compact objects.* – The previous discussion has indicated that gravitational attraction is the only force that can make a system emitting significant gravitational radiation. Going in more detail, in the simple case of a binary system, the size of the object is a limiting factor for the emission amplitude. In order to achieve high accelerations the bodies need to be at a short distance, as the force between them goes as R^{-2}. Their size

sets a lower limit to R: tidal forces disrupt the orbiting masses, even before contact of their outer layers.

White dwarfs are objects where matter resists to gravitational collapse through electron degeneracy pressure with a typical mass of a fraction of one solar mass and a radius comparable with Earth's. For a system in our Galaxy at 10 kpc consisting of two white dwarfs of 1 solar mass each orbiting at 10^5 km one from the other, the estimate for the wave amplitude is

$$h \simeq \frac{R_S}{R}\frac{R_D}{r_D} = \frac{3\,\mathrm{km}}{10^5\,\mathrm{km}}\frac{3\,\mathrm{km}}{10^{20}\,\mathrm{m}} \cong 3 \times 10^{-22}.$$

The wave generated by such a system could be detected by current interferometers, that are able to measure displacements of that order. Whether the test masses faithfully reproduce the varying geometry of space-time depends much on the frequency of the signal. Treating this system in a classical way the frequency of the emitted wave can be computed. The second Newton law states for a circular motion

$$\frac{\mathrm{G}M_\odot^2}{4R^2} = \mu\omega^2(2R),$$

or

$$\omega^2 = \frac{\mathrm{G}M_\odot}{4r^3},$$

where μ is the reduced mass. The resulting angular frequency is

$$\omega = 4 \times 10^{-3}\,\mathrm{rads}^{-1}$$

or a period of about 600 s, resulting in a signal with a period of 300 s. As will be discussed later, the local gravity fluctuations present on Earth generate at this frequency metric changes that are orders of magnitude higher than the wave generated by the binary white dwarf system. Indicatively current ground based interferometers aim at recording gravitational waves down to 10 Hz, while future project are attempting to push this limit to 1 Hz.

The discovery of the neutron in 1932 by Chadwick induced Baade and Zwicky to propose the existence of neutron stars, as the result of the core collapse of a star. Neutron stars are formed when a massive star undergoes core collapse when a supernova event occurs. Matter gets compressed under gravitational pressure and the inverse β decay reaction

$$e^- + p \rightarrow n + \nu_e$$

is energetically favoured leading to a compact body composed mostly of neutrons. The typical mass is $1.4\,M_\odot$ and the radius 10 km, resulting in a density of the order of

$10^{15}\,\mathrm{g/cm^3}$. Models point to the existence of a solid crust made of atomic nuclei with a sea of electrons, while the inner layers contain nuclei that are richer in neutrons but would show radioactive decay if not bound by gravity. Current understanding is that the progenitor star with a mass between ~ 8–$30\,M_\odot$ is likely to have in its remnants a neutron star.

To estimate the final orbital frequency one can consider a system of $1.4\,M_\odot$ bodies of 10 km radius. Then

$$\omega^2 = \frac{\mathrm{G}1.4\,M_\odot}{4r^3}$$

with the radius of the single object being now 10^4 m,

$$\omega = 1.6 \times 10^{-2}\,\mathrm{rads}^{-1}$$

or a period of the order of 1 ms. The gravitational wave has itself double frequency, or a period of 0.5 ms, or 2 kHz.

Encountering a system of two neutron stars, about to coalesce, is going to be very rare for a single Galaxy. A sphere of 100 Mpc radius might show in one year one such event.

For binary black holes of $30\,M_\odot$ ($r_S = 3$ km) at 400 Mpc, forming a system of size 600 km

$$h \sim 10^{-21}.$$

Neutron stars and black holes are the heaviest and most compact astrophysical objects known. For neutron stars the typical mass is $1.4\,M_\odot$ with a radius of 10 km. If the progenitor star has a mass greater than $\sim 25\,M_\odot$, matter does not resist to compression and a black hole forms. Black holes dimensions are given by their Schwarzschild radius $r_S = 2\mathrm{G}M/\mathrm{c}^2$: for a $30\,M_\odot$ black hole $r_S = 90$ km. From these sizes the highest frequency emitted can be estimated or, the other way around, a lower limit on the frequency of the signal to be detected sets an upper limit to the mass of the black hole system to be observed.

2˙4. *Single compact objects*. – The collapse of a stellar core sees the motion of large quantities of matter driven by gravity toward the star center, giving birth to a Supernova. These catastrophic events were among the first sources of gravitational waves being looked for with resonant bars, that could detect signals around 1 kHz. The amount of energy emitted was estimated to be quite high, of the order of 0.01 solar mass equivalent. Extremely complex hydrodynamic simulations showed that this energy was significantly overestimated, as only the non-axisymmetric component of the quadrupole derivative is emitting radiation. The current estimate is that only $10^{-8}\,M_\odot$ or less energy equivalent is emitted in a stellar core collapse, consequently the emitted amplitude for a source 10 kpc away is $\simeq 10^{-21}$. Unfortunately the rate of supernovae in our Galaxy is about 1 every

40 years, which requires being able to detect sources from the surrounding galaxy clusters. The observation of such an event, in particular if accompanied with the detection of the associated neutrinos and the observation of the light curves at various wavelengths would provide otherwise inaccessible information about these processes.

Angular-momentum conservation during the core collapse leads to a rapidly rotating object. With a sufficiently high mass progenitor, the remnant is likely to be a rapidly spinning neutron star. This in turn may turn out not to be axisymmetric and thus emit gravitational waves during very long times. The quadrupole magnitude depends on the formation and evolution history. On top of an axisymmetric surface it is expected that asperities, or "mountains", of a few mm stand gravity thanks to the solid crust, or hydrodynamically generated deformations result in a small but non-zero ellipticity:

$$\epsilon = \frac{I_{yy} - I_{xx}}{I_{xx}} < 10^{-6}$$

I_{xx} and I_{yy} are the principal moments of inertia perpendicular to the one nearly aligned with the angular momentum vector. The expected amplitude is very low

$$h \sim 4 \cdot 10^{-27} \left(\frac{f}{100\ \mathrm{Hz}}\right)^2 \left(\frac{10\ \mathrm{kpc}}{r}\right) \left(\frac{I_{zz}}{10^{38}\,\mathrm{kg\,m^2}}\right) \left(\frac{\epsilon}{10^{-6}}\right).$$

The different scale of distances with respect to binary systems places the source candidates inside our Galaxy. Currently only upper limits have been placed on the emitted amplitude of electromagnetically identified pulsar, as no periodic source has been detected yet.

The fascinating point about these sources is that they are continuous and can be observed in principle for years, allowing to average noise.

From stellar evolution it is estimated that there are of the order of 10^9 neutron stars in our Galaxy; on the other hand about two thousand pulsar have been identified by radiotelescopes and more recently by gamma ray satellites. This can be explained by the strong beaming of the electromagnetic radiation present in pulsars, that makes most of them going undetected. For gravitational radiation the antenna pattern is much broader, giving access to a much larger number of rotating neutron stars, provided the emitted amplitude is sufficient. The number of neutron stars with a significant emission amplitude is very uncertain, a year long observing period with the improved sensitivity of second generation interferometers might lead to identify the first continuous source of gravitational waves.

2˙5. *Supernovae.* – During stellar-core collapse a large amount of mass undergoes high acceleration. However most of this motion is axisymmetric and not coherent over the whole star volume, resulting in low gravitational wave emission. An intense theoretical and numerical effort is going on since many years to model the stellar-core collapse and supernova explosion. This is a complex and computer time consuming task that has

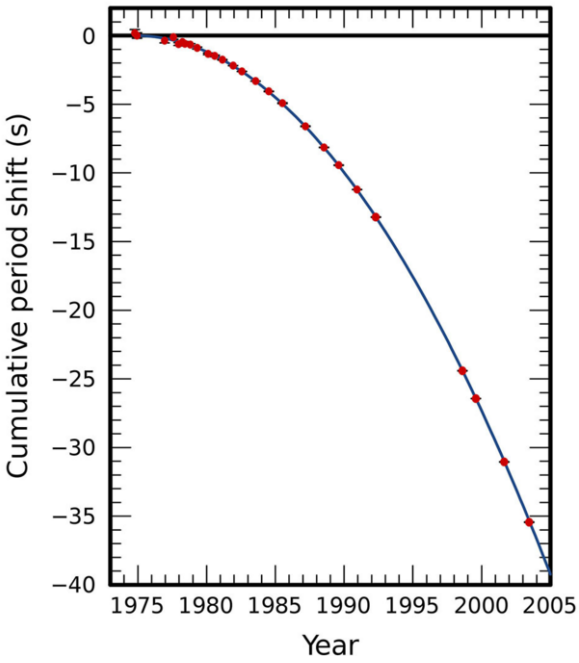


Fig. 2. – Cumulative shift in the periastron period in seconds for the binary star system PSR B1913+16.

to include tens and tens of fundamental processes and nuclear reactions, magnetohydrodynamics and relativistic effects. The radiated energy is of the order of 10^{-8} solar masses, still detectable if the source is in our Galaxy, with a very low rate (1 every 40 years). First searches for a supernova signal were tuned for a bipolar spike lasting a few milliseconds. Simulations are now showing that together with the spike there is emission at lower frequency lasting up to one second.

Observation of gravitational waves from a stellar core collapse, together with associated neutrinos and full spectrum electromagnetic radiation can enable the study of a wealth of phenomenta, from relativistic magnetohydrodynamics to nuclear physics to probing neutrino mass and structure. Present and future detectors have to keep as high as possible the probability of recording such a rare event.

2˙6. *The indirect evidence for gravitational radiation: PSR 1913+16.* – Nowadays gravitational waves have been observed on Earth, recording precisely the signal emitted during the coalescence of two compact objects. The first evidence for gravitational radiation came much earlier, in 1975, with the observation of periodic electromagnetic pulses, as expected from a pulsar, showing a frequency modulation on regular base of 7.7 hours. This was immediately interpreted as coming from a tight eccentric binary system containing a pulsar and another neutron star, with an orbital period decreasing

TABLE II. – *Physical parameters of PSR1913+16.*

Parameter	Symbol (units)	Value
i) "Physical" Parameters		
Right Ascension	α	$19^h15^m28.^s00018(15)$
Declination	δ	$16°06'27.''.4043(3)$
Pulsar Period	P_p (ms)	59.029997929613(7)
Derivative of Period	$\dot{P}_p$	$8.62713(8) \times 10^{-18}$
ii) "Keplerian" Parameters		
Projected semimajor axis	$a_p \sin \iota$ (s)	2.3417592(19)
Eccentricity	e	0.6171338(4)
Orbital Period	P_b (day)	0.322997462727(5)
iii) "Post-Keplerian" Parameters		
Mean rate of periastron advance	$\langle\dot{\omega}\rangle(^\circ \mathrm{yr}^{-1})$	4.226607(7)
Redshift/time dilation	γ' (ms)	4.294(1)
Orbital period derivative	$\dot{P}_b(10^{-12})$	−2.4211(14)

with time, in contrast with a classical Keplerian motion. Figure 2 shows the difference of periastron epoch with respect to what expected for a constant orbital period [8]. The parabolic behaviour comes from a linear variation of the period.

The results of a detailed analysis are summarized in table II. Other non-Newtonian effects are worth noting: the high periastron precession of $4^\circ\,\mathrm{yr}^{-1}$ to be compared with $0.44\,\mathrm{arcsec\,yr}^{-1}$ of the orbit of Mercury, and the variable periodic time delay due to the redshift due to the companion.

3. – Signals and noise

The first conclusion drawn about gravitational waves was that "in all thinkable cases the amplitude must have a practically vanishing value" [2]. This was in 1916, the atomic nucleus had been discovered 5 years earlier, the Schwarzschild solution of Einstein's field equations had just been published, and it would have taken more than one decade to learn of the existence of neutrons.

These discoveries would pave the way toward identifying the sources of gravitational waves discussed previously. Important experimental discoveries allow now to achieve extremely precise measurements of length and time using lasers, but gravitational wave detection is still limited by noise in the detectors. A full treatment of noise is beyond the scope of these lectures, but the main questions encountered when designing a grav-

itational wave detector and analyzing the data it produces can be discussed in terms of a few general ideas: the noise power spectrum and Wiener optimal filtering.

3˙1. *Noise power spectrum.* – A signal from the transducer of a noisy physics instrument can be represented as a random process, a function $s(t, X)$ of time and of a vector of random variables $X = (X_1, X_2, \ldots, X_n)$. A single observation, called a realization of the process, is obtained by drawing a value $\mathbf{x}$ for the random vector X. Signal properties like the average value and the variance can then be obtained using the relevant ensemble averages of $s(t, \mathbf{x})$ obtained drawing the random vector X.

Besides computing average and variance at a fixed time, it is important to consider correlations between different times. The *autocorrelation* is obtained from the ensemble average

$$R(t, \tau) = \langle n(t, K) n(t + \tau, K) \rangle.$$

It is assumed in the following that these properties do not vary with time, that is the random process is stationary (in a wide sense [9]).

In this case the autocorrelation is an even function of the delay τ only:

$$R_n(\tau) = \langle n(t) n(t + \tau) \rangle = \langle n(t - \tau) n(t) \rangle.$$

The zero delay correlation

$$R_n(0) = \langle n(t) n(t) \rangle = \mathrm{Var}[n(t)]$$

is the variance of $n(t)$.

If the signal shows no correlation with what happened immediately before,

$$R_n(\tau) = 0, \quad \tau \neq 0.$$

The system that produces this signal has no memory: such is the autocorrelation for independent radioactive decays. When the decay product interacts with matter producing an electromagnetic signal, the dynamics of the physical system has memory of the past states. There are time constants in the system making $R_n(\tau) \neq 0$ for $|\tau| > 0$.

The Fourier transform of the autocorrelation results in what is called the power spectrum $S_n(f)$ of the signal $n(t)$:

$$S_n(f) = \int_{-\infty}^{+\infty} R_n(\tau) \exp(-i 2\pi f \tau) \mathrm{d}\tau, \quad -\infty < f < +\infty.$$

As $R_n(\tau)$ is even, $S_n(f)$ is real. The autocorrelation function can be obtained from the power spectrum as

$$R(\tau) = \int_{-\infty}^{+\infty} S(\omega) \exp(i 2\pi f \tau) \mathrm{d}f.$$

TABLE III. – *Some properties of autocorrelation functions and power spectra* ($\omega = 2\pi f$).

$n(t)$	$R(\tau)$	$S(\omega)$
$an(t)$	$\lvert a\rvert^2 R(\tau)$	$\lvert a\rvert^2 S(\omega)$
$\frac{\mathrm{d}n(t)}{\mathrm{d}t}$	$-\frac{\mathrm{d}^2 R(\tau)}{\mathrm{d}\tau^2}$	$\omega^2 S(\omega)$
$\frac{\mathrm{d}^k n(t)}{\mathrm{d}t^k}$	$(-1)^k \frac{\mathrm{d}^k R(\tau)}{\mathrm{d}\tau^k}$	$\omega^{2k} S(\omega)$
$n(t)\exp[\pm i\omega_0 t]$	$R(\tau)\exp[\pm i\omega_0 \tau]$	$S(\omega \mp \omega_0)$
$n(t)\cos\omega_0 t$	$R(\tau)\cos\omega_0 \tau$	$(S(\omega+\omega_0)+S(\omega-\omega_0))/2$

For $\tau = 0$, using Parseval theorem:

$$\int_{-\infty}^{+\infty} S(f)\mathrm{d}f = R(0) = E[n^2(t)] = \mathrm{Var}[n(t)] \geq 0.$$

The integral of the power spectrum is the average power of $n(t)$, thinking of $n(t)$ being the current flowing through a $1\,\Omega$ resistor. Table III shows some useful properties of autocorrelation functions and power spectra.

Electric circuits that process a signal are causal, usually time invariant, linear systems. Their output $y(t)$ is the convolution of the input $n(t)$ with the impulse response of the system $L(t)$:

$$y(t) = \int_0^{+\infty} L(\tau)n(t-\tau)\mathrm{d}\tau = \int_{-\infty}^{0} L(t-\tau)n(\tau)\mathrm{d}\tau.$$

From the properties of the Fourier transform it can be shown [9] that

$$S_y(f) = |H(f)|^2 S_n(f),$$

where $H(f) = \int_0^{+\infty} L(t)\exp(-i2\pi ft)\mathrm{d}t$ is the complex transfer function of the linear system, which is causal.

A few autocorrelation functions and power spectra frequently encountered are listed in table IV. The case of a memoryless process results in a flat power spectrum, and one speaks of white noise, all frequency components are present, as in white light.

Using the transfer function of a band pass filter $H_{BP}(f) = 1$ for $f_0 < |f| < f_0 + 1\,\mathrm{Hz}$ and zero otherwise, it is possible to compute the power "localized" in a given frequency band and justify the name of power spectrum given to $S(f)$. Consider the voltage output of a channel $n(t)$ and distribute it to a battery of bandpass filters 1 Hz wide as in fig. 3.

TABLE IV. – *A few autocorrelation functions and power spectra* ($\omega = 2\pi f$).

$R(\tau)$	$S(\omega)$
$\exp[-\alpha t]$	$2\alpha/(\alpha^2+\omega^2)$
$\exp[-\alpha t]\cos\omega_0 t$	$\alpha/(\alpha^2+(\omega-\omega_0)^2)$
$\lvert t\rvert \le T : 1-\lvert t\rvert/T\rvert$, $\lvert t\rvert > T : 0$	$4\sin^2(\omega\tau/2)/T\omega^2$
$\delta(\tau)$	1
1	$2\pi\delta(\omega)$
$\cos\omega_0 t$	$\pi\delta(\omega+\omega_0)+\pi\delta(\omega-\omega_0)$

For each filter the power dissipated on a $1\,\Omega$ resistor is

$$\int_{-\infty}^{+\infty} S_n(f)|H(f)|^2 \mathrm{d}f = 2\int_{f_0}^{f_0+1\,\mathrm{Hz}} S_n(f)\mathrm{d}f \simeq 2S_n(f_0).$$

If $[n(t)]$ indicates the units in which the signal is measured, from the definition of power spectrum the units of $S_n(f)$ are

$$[S_n(f)] = [n(t)]^2\,\mathrm{s} = \frac{[n(t)]^2}{\mathrm{Hz}}.$$

For example, if the output of the detector is in Volt, the power spectrum is measured in V^2/Hz.

To compare the noise of an instrument to the amplitude of a signal, one uses the *Linear Power Spectrum* (LPS) which is defined as

$$\tilde{n}(f) = \sqrt{S_n(f)}.$$

Instead of indicating the variance of the noise in a 1 Hz bin, one quotes the standard deviation or rms of the signal in that same band, allowing a direct comparison between signal and noise. With $S_n(f)$ having units $[n(t)]^2/\mathrm{Hz}$, the units of $\tilde{n}(f)$ are

$$[\tilde{n}(f)] = \frac{[n(t)]}{\sqrt{\mathrm{Hz}}}.$$

For example, the noise in an operational amplifier is measured in $\mathrm{nV}/\sqrt{\mathrm{Hz}}$.

If there is more than one noise source that contributes to the output of a device, and if these sources are uncorrelated, to find the total noise the *power* from each of them has to be summed, like the dissipated power in a resistor is the sum of the powers of the various contributions. The power spectra can be summed, the LPS cannot, these should be summed in quadrature.

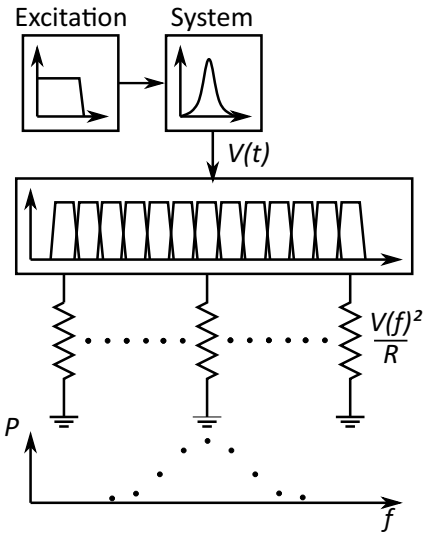


Fig. 3. – The noisy channel $n(t)$ is distributed to 1 Hz wide pass band filters. The power dissipated on a 1 Ω resistor is measured.

The previous discussion is about the level of noise in a 1 Hz band, however expected signals are not confined in a band of one frequency unit. A signal of characteristic duration τ_c has frequency components up to $f_c \simeq 1/\tau_c$. To obtain the instrumental noise, the power spectrum over the frequency band must be integrated up to f_c. In the approximation of a flat spectrum the total power is $f_c S_n(f)$. For a signal of amplitude A to be compared with noise, one obtains an estimate

$$\mathrm{SNR} \sim \frac{A}{\sqrt{S_n(f_c) f_c}}\,.$$

For a 1 ms duration signal of amplitude A and detector with linear noise n, SNR $\simeq A/30n$.

3˙2. *Power spectra in practice.* – The above definitions make use of Fourier transforms that are defined for positive and negative frequencies. On the other hand, physical measurements involve real quantities and only positive frequencies. For a noise measurement at a frequency f there are contributions from both equal power spectra $S_n(f)$ and $S_n(-f)$, giving as total noise power twice the power spectrum. The result is the so-called one-sided spectrum. When manipulating power spectra it is advised to verify whether the factor of 2 is already included: the LPS defined using only positive frequencies is $\sqrt{2}$ times the LPS defined for $-\infty < f < +\infty$.

In practical cases the estimate of the power spectrum is not based on the autocorrelation but uses directly the power in the various frequency bands, averaged over time. This is under the hypothesis that the system is ergodic, implying that an average over an ensemble of realizations of the stochastic process is equal to that same average over

time. The estimate of the power spectrum $\hat{S}_n(f)$ is then computed as

$$\hat{S}_n(f) = \frac{1}{N}\sum_1^N \frac{1}{T}\left|\int_{-T/2}^{+T/2} n(t)\exp(i2\pi ft)\mathrm{d}t\right|^2 .$$

Two scales of time are involved in this process: the duration of the Fourier transform T which determines the frequency resolution $\Delta f = 1/T$, and the overall duration of the measurement $\tau = NT$, which is determined by the number of desired averages. Increasing the number of averages reduces the fluctuations of the measured $S_n(f)$ but increases the total measurement time τ, which can be a problem in the presence of non-stationarity, as can happen in practice.

This procedure considers signals that start and end sharply at $-T/2$ and $T/2$ respectively: one says a rectangular time window $W(t)$ is used. This creates problems in the frequency domain and windows with specific weights (*e.g.*, Hann, Hamming, ...) are applied. Data at the beginning and at the end of the window are weighted less and to improve statistical accuracy time windows are overlapped. To estimate in practice a power spectrum it is necessary to specify the duration of the single Fourier transform T, the number of averages N, the window used and the amount of overlap between windowed data.

3˙3. *Power spectrum in digitized signals*. – Nowadays most signals from transducers are recorded digitally and then treated numerically. These signals are sampled at some frequency f_s and converted from analog to an integer number with some rounding off. This introduces an additional white noise that has rms error $\Delta A/\sqrt{12}$, where ΔA is the minimal step in the analog-to-digital conversion. In addition, if the signal hits the maximum amplitude that can be converted distortion can occur.

Sampling at some frequency f_s causes further information loss: f_s determines what are the fastest variations that can be recorded. Simply speaking, it is possible to keep track of an oscillating signal if a positive variation is followed by a negative one. At least three samples are needed and the maximum recordable frequency, called *Nyquist frequency*, is $1/2\tau_s = f_s/2$. If the analog signal oscillates with $f > f_N$ it will not be recorded faithfully. An oscillating signal with a frequency lower than f_N appears. This is called frequency aliasing.

Assume to sample a sinusoidal signal of frequency $0 < f < f_N$. Between one sample and the other the phase changes by

$$\Delta\phi = 2\pi f\tau_s = 2\pi f/f_s, \quad |\Delta\phi| < \pi$$

but the frequency of the sampled function is the same if the original phase changes instead by

$$\Delta\phi = \pm(2\pi f/f_s + 2k\pi) = \pm(2\pi(f + kf_s)/f_s),$$

k integer (positive and negative frequencies are not distinguishable).

Any signal with frequency $\pm(f+kf_s)$ gives a sampled signal with frequency f. In other words, a sampled signal with frequency f contains contributions from signal components at kf_s+f and kf_s-f, for any integer k. Signals with frequency $kf_s+f_N < f < (k+1)f_s$ appear at $f_N - f$ (note the minus sign in front of f), signals with frequency $kf_s < f < kf_s + f_N$ appear instead at frequency f.

To obtain the correct power spectrum of a signal from its samples there should be no contribution from the frequency band above the Nyquist frequency. If present it must be strongly attenuated by a low pass filter BEFORE sampling. Failure to do so can lead to a noise increase in the sampled signal spoiling dramatically the gravitational wave detector sensitivity.

3˙4. *Signal and noise.* – The present gravitational wave detectors look for rare transient signals among noise or long-lasting faint signals. For a signal present among the noise, the detector output is

$$s(t) = n(t) + h(t).$$

One procedure to detect such a signal is to compute the likelihood of observing $s(t)$ in both hypotheses of presence and absence of a signal. As in the case of discrete measurements, this can be expressed in terms of a distance between the observed signal and the hypothetical waveform. For digital signals this amounts to minimizing a quadratic form defined by the inverse of the covariance matrix $C(i\tau_s, j\tau_s)$ between samples at different times:

$$\chi^2 = \sum_{i,j} [s((i-j)\tau_s) - h((i-j)\tau_s)] \left(C^{-1}\right)_{i-j,i} [s((i)\tau_s) - h((i)\tau_s)].$$

In the continuous case the distance is obtained dividing by the inverse of the autocorrelation function:

$$\int_0^t \int_0^\tau \frac{(s(t-\tau) - \alpha h(t-\tau))(s(t) - \alpha h(t))}{R(t, t-\tau)} \mathrm{d}\tau \, \mathrm{d}t.$$

This is an example with a fixed waveform $h(t)$ with unknown amplitude α. The value of α that minimizes this distance designates the best waveform that fits the data.

Minimizing with respect to α amounts to solve

$$\int_0^t \int_0^\tau \frac{(s(t-\tau) - \alpha h(t-\tau))h(t)}{R(t, t-\tau)} \mathrm{d}\tau \, \mathrm{d}t = 0.$$

The solution for α is

$$\alpha = \frac{\displaystyle\int_0^t \int_0^\tau s(t-\tau)h(t)/R(t, t-\tau) \mathrm{d}\tau \, \mathrm{d}t}{\displaystyle\int_0^T \int_0^t h(t-\tau)h(t)/R(t, t-\tau) \mathrm{d}\tau \, \mathrm{d}t} = 0.$$

3·5. *Optimal filtering.* – The above procedure can be translated to the frequency domain. In the general case the known signal $h(t)$ is a function of several parameters (coalescence time, masses, ...). Then the optimal filter for a deterministic signal, fully determined by its parameters, requires to compute

$$\begin{aligned} s_W &= 2\int_0^{+\infty} \frac{\tilde{s}(f)\tilde{h}^*(f)}{S_n(f)}\,\mathrm{d}f \\ &= \int h(t) \int s(\tau) w(t-\tau)\,\mathrm{d}\tau\,\mathrm{d}t, \end{aligned}$$

$\tilde{h}^*(f)$ is the complex conjugate of the Fourier transform of $h(t)$. $w(t)$ weighs $s(t)$ more at frequencies where the detector is less noisy.

s_W is Gaussian distributed with mean zero and standard deviation 1 if the template $h(t)$ is properly normalized. A simple detection procedure would then consist in defining an appropriate threshold for $|s_W|$ for a given choice of probability of false alarm.

The expected signal-to-noise ration SNR is given by

$$\mathrm{SNR}^2 = 4\int_0^{+\infty} \frac{|\tilde{h}(f)|^2}{S_n(f)}\,\mathrm{d}f.$$

The SNR^2 builds up integrating over frequency, weighted with the inverse of the detector noise power spectrum. To achieve a reliable detection, the noise has to be low, possibly over the full frequency band spanned by the signal. Once a detection is claimed the parameters of the system can be estimated as those giving the maximum $|s_W|$.

The subject of signals and noise saw a rapid development during World War II in particular thanks to Norbert Wiener, after whom the optimal filter is named. More detailed treatment is to be found in textbooks on Stochastic Processes, Digital Filters, Statistical Signal Processing, Detection theory.

4. – Primary noise sources in gravitational wave interferometers

As discussed at the beginning, gravitational waves are tiny perturbations of the space-time metric that are observed looking at the separation between two free masses. The waveform is recorded by measuring, as a function of time, a variation of the order of 10^{-18} m of their distance. A ground-based detector uses pairs of suspended masses free to move in the horizontal plane and as much as possible insensitive to local perturbations. The apparatus has then to ensure that:

- local effects do not move the masses,
- mass separation is well measured.

The size of the expected effect is hard to imagine, ten billionths of the size of an atom, to be measured between macroscopic masses. As will be seen, this compares badly with

Earth ground motion, and hits physical limits coming from thermal motion and local gravity fluctuations.

In current detectors the distance measurement relies on the use of light with wavelength of the order of 10^{-6} m. Optical interference is used to measure the phase of the electromagnetic field returning from the remote mass at the level of 10^{-11} rad. However phase is obtained from the light intensity in the interference pattern, with precision limited by quantum mechanics.

Noise in test mass position is discussed first, starting form seismic noise, that is definitely prevalent at low frequency. Then other fundamental noise sources will be encountered. The description of the optical measurement concludes the section. Needless to say that an ideal world is described, current instruments receive many other contributions to the observed noise, called generically technical noise, coming from the associated control equipment or the residual environmental coupling. There the best textbook is the laboratory floor.

5. – Position noise

5˙1. *Seismic noise.* – Gravitational wave detection asks for masses that move less than 10^{-18} m in the laboratory reference frame in the detection frequency band. Seismic activity, on the other hand, results in the surface of the Earth moving by 0.1–1 μm over a 1 s time scale. This is a random superposition of many elastic waves that travel through the Earth crust and its interior, originated by earthquakes, tectonic plate motion, but also by pressure variation on soil due to sea motion, ocean swell, wind. On longer time scales, tides of the solid Earth change by tens of microns the distance between two points a few km apart. Human activity, from car traffic to manufacturing plants, when not planes, helicopters, trains and underground tunnels, is also an important source of vibrations, that may have varying intensity depending on time. Different locations may exhibit different vibration characteristics, depending on general seismic activity and mechanical soil response (speed of sound, attenuation).

The microseismic spectrum at the Virgo location near Pisa, in Central Italy, is usually parametrized as

$$\tilde{x}_s(f) = 10^{-7} \left(\frac{1\,\mathrm{Hz}}{f} \right)^2 \mathrm{m}/\sqrt{\mathrm{Hz}},$$

for $f > 1$ Hz. Below 1 Hz, wind and sea introduce a seismic component around 140 mHz, corresponding to a period of 7 s. The amplitude of the microseism is then 11 orders of magnitude higher than the expected signal. This is a rather unusual situation: for example the human ear can stand only 6 orders of magnitude in pressure above the earing threshold, or typical analog electronic devices can manipulate signals that span 6–7 orders of magnitude. The amplitude of the motion of the mass with respect to an inertial frame must be strongly reduced, but the ground moves enormously. Signals in

the instrument have to cover a large dynamic range, and this is one of the main practical difficulties of gravitational wave detectors.

5˙2. *Seismic attenuation.* – The reduction of seismic motion is achieved by applying a cascade of mechanical filters that give in steps the required attenuation [10]. These are essentially harmonic oscillators, that operate as low-pass filters above the resonance frequency, using the inertia of the attached mass to attenuate motion. Consider a pendulum with a moving suspension point $x_s(t)$ and neglecting damping. The position $x(t)$ of the pendulum mass follows:

$$\ddot{x} + \frac{g}{l}x = \frac{g}{l}x_s(t),$$

where dots indicate time derivatives. The pendulum length is l and g is the acceleration of gravity. In the frequency domain the solution for an excitation $x_s(t) = x_0 \exp(i\omega t)$ has an amplitude

$$x(\omega) = x_0 \frac{\omega_0^2}{|\omega^2 - \omega_0^2|}\,, \quad \omega_0^2 = \frac{g}{l}\,.$$

At high frequency with $\omega \gg \omega_0$, the amplitude becomes

$$x(\omega) = x_0 \frac{\omega_0^2}{\omega^2} = x_0 \frac{k}{m\omega^2}\,,$$

attenuation is achieved through inertia. Dissipation should be avoided, as it introduces a correlation between the pendulum and the suspension point or the environment, decreasing the filter performance. Furthermore, as discussed below, it is intrinsically noisy.

By cascading n pendula one has a system with n normal modes. Above the frequency of the highest normal mode the back action of pendulum i on pendulum $i-1$ becomes negligible. Each filter acts independently and attenuations can be multiplied. The total attenuation A can be estimated: for a 1 m pendulum

$$\omega_0^2 = \frac{g}{l} = 10\,\mathrm{rad^2\,s^{-2}}$$

and for n filters

$$A \sim \left(\frac{\omega_0^2}{\omega^2}\right)^n = \left(\frac{10}{4\pi^2 f^2}\right)^n.$$

The spectrum has a power dependence f^{-2n} making it a "seismic wall" toward low frequency.

As an example, seven filters provide at 4 Hz an attenuation factor

$$A = \left(\frac{\omega_0^2}{(2\pi f)^2}\right)^7 = \left(\frac{10}{640}\right)^7 = 2.3 \times 10^{-13}.$$

The last filter is the suspended test mass, so a six filter suspension allows to attenuate seismic noise down to 4 Hz. Below that frequency attenuation is greatly reduced and, worse, at the resonant frequency of the normal modes, seismic motion is amplified.

The system must be able to precisely position each mirror while damping all the resonant modes. These two functions deal with completely different ranges of displacement: the position must be controlled at the level of 10^{-12} m for adequate interferometer tuning, and at the level of 10^{-6} m when controlling seismically excited modes, without introducing noise at the level of 10^{-18} m. The dynamic range required cannot be provided by a single actuator: well designed analog electronics covers typically seven orders of magnitude when at least twelve are needed. Moreover the Earth elastic deformation due to the Moon and Sun tides introduces a distance change of several tens of microns over a distance of 3 km and a timescale of hours. The development of mirror suspensions able to meet these and other requirements has taken a long time. Different approaches have been pursued, the major conceptual difference being the selected solution to control the motion of the suspension point. Here the solution developed for Virgo is presented, based on a mostly passive approach, while LIGO adopted for the first attenuation stage an active inertial platform.

5˙3. *The Virgo Superattenuator.* – The Virgo Superattenuator (see [11]) was designed to provide sufficient seismic attenuation at a frequency as low as 10 Hz, avoiding as much as possible elements that could reintroduce noise or represent a bypass for vibrations. It was realized very early that attenuation should occur in all six degrees of freedom of the suspended mirror, even if the measurement involves only displacement along the beam axis. Indeed finite mechanical precision results in residual coupling between degrees of freedom, so high motion in one degree of freedom may then leak into the longitudinal direction, nullifying the filter action. Attenuation in the vertical direction is particularly important: the ultimate vertical-to-horizontal coupling is given by the Earth curvature: at a distance of 3 km the directions of the local vertical differ by 4.5×10^{-4} rad; motion in that direction couples directly to the horizontal axis. As a matter of fact, due to construction imperfections, mechanical coupling is usually much higher than this.

Vertical attenuation is a challenge: attenuation is higher with a low resonance frequency, obtained using a soft spring. But the suspension must be stiff enough to hold several hundred kilograms, so a specific solution, called antispring, was developed. The suspension spring is softened by introducing an element that has, over a small region, a "negative" elastic constant, producing a force

$$F = +k(y - y_0), \quad |y - y_0| < d.$$

In the Virgo Superattenuator this is achieved using the repulsive force of magnets [12]. If magnets are perfectly aligned, the transverse force is zero, otherwise a component proportional to displacement appears. A similar behaviour is obtained with compressed springs perpendicular to the vertical axis.

The Virgo Superattenuator includes an inverted pendulum [13] that supports the suspension point of the filter chain. Elastic joints keep vertical the legs that support

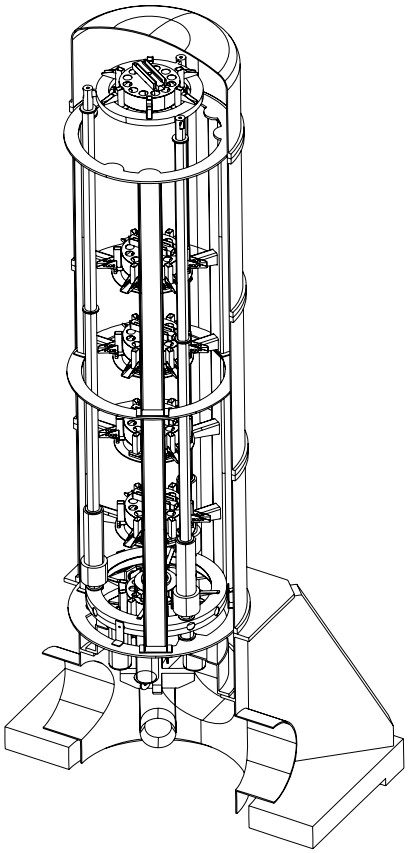


Fig. 4. – Perspective view of the Virgo Superattenuator. The inverted pendulum legs are 6 m long.

the top table of the suspension. Gravity acts as destabilising element, reducing at will the resonance frequency which is tuned around 30 mHz ensuring significant horizontal attenuation. The resulting elastic constant is very low, and only small forces are needed to move the top platform by tens of microns, as required to follow crust tidal deformations. These forces are applied over hours, moving the average position of the mirror. Being applied upstream in the filter chain, any noise component at higher frequency is filtered and does not reach the mirror. This split of the control signal into a high-amplitude and low-frequency component, and a low-amplitude high-frequency one, gives the dynamic range needed. The resulting structure of a long superattenuator is shown in fig. 4.

The transfer function

$$A(f) = \left| \frac{x_m(f)}{x_s(f)} \right|$$

for the Virgo Superattenuator is shown in fig. 5. $A(f)$ is 1 in the band below the resonance of the inverted pendulum, at 30 mHz. At increasing frequency further resonance peaks are

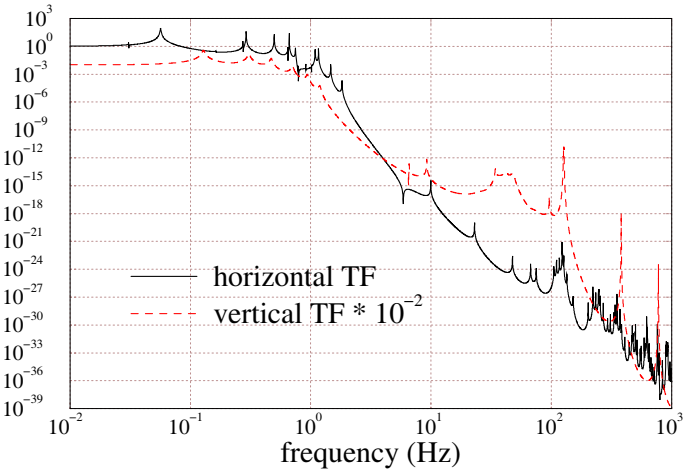


Fig. 5. – Superattenuator transfer function.

visible, with the attenuation slope increasing. At high frequency the internal mechanical resonances of the filters reduce the performance of the suspension. In ultimate analysis, the low-frequency performance relies on the distance between filters, which is proportional to the square root of the oscillation period. Pushing the seismic wall down to 2 or 1 Hz may turn out to be very difficult for practical reasons.

5˙4. *Thermal noise*. – One of the limitations that are present in measurement devices is thermal noise, coming from being in thermal equilibrium with the environment. The test mass suffers no exception and its position and velocity degrees of freedom store on average an energy of

$$k_B T = 3.9 \times 10^{-21}\,\mathrm{J}$$

at room temperature. For a 40 kg mirror suspended by a 1 m pendulum, the elastic constant of the equivalent harmonic oscillator is

$$mg/l = 400\,\mathrm{N/m}$$

and the corresponding oscillation amplitude x_T is

$$x_T = \sqrt{\frac{k_B T}{k}} = \sqrt{\frac{k_B T l}{mg}} = 3.1 \times 10^{-12}\,\mathrm{m}.$$

This is a root mean square value which can be thought to be averaged over all frequency components of the motion. Intuitively one understands that different frequency bands, for example on and off resonance, contribute differently to the total kinetic energy. This is discussed in more detail in the following section.

5˙5. *Fluctuation-Dissipation theorem.* – There is a connection between the dissipation of the energy stored in a system and its fluctuations when in equilibrium with a thermal bath: the dissipation mechanism lets energy flow at some rate toward the environment, but it also connects the heat bath to the system, creating a single body subject to thermodynamic fluctuations. The *Fluctuation-Dissipation* theorem [14] links the power spectrum of noise from thermal origin to the dynamics of the system and the amplitude of noise is determined by the "width" of the dissipation channel. The theorem states [15] that if $Y(\omega)$ is the mechanical admittance, or the inverse of the impedance $Z(\omega)$,

$$Y(\omega) = \frac{1}{Z(\omega)} = \frac{\dot{x}}{F}$$

the power spectrum $S_{\dot{x}}(\omega)$ of the quantity $\dot{x}$, proportional to the force and to the admittance, is

$$S_{\dot{x}_T}(\omega) = 2k_B T \Re Y(\omega),$$

where T is the equilibrium temperature, and $\Re Y(\omega)$ the real part of $Y(\omega)$.

Consider a pendulum in vacuum, with dissipation proportional to displacement rather than to velocity, as is internal friction in solids. The equation of motion is

$$\ddot{x} + \omega_0^2[1 + i\phi(\omega)]x = F/m$$

with the so-called loss angle $\phi \sim 10^{-3} - 10^{-6}$ for materials typically used in gravitational wave interferometers. The loss angle, times 2π, measures the fraction of mechanical energy dissipated during one oscillation.

The motion is given by

$$x(\omega) = \frac{F}{m} \frac{\omega_0^2 - \omega^2 - i\phi(\omega)\omega_0^2}{(\omega_0^2 - \omega^2)^2 + \phi^2(\omega)\omega_0^4},$$

so

$$\Re Y(\omega) = \frac{\omega}{m} \frac{\phi(\omega)\omega_0^2}{(\omega_0^2 - \omega^2)^2 + \phi^2(\omega)\omega_0^4}.$$

Applying the Fluctuation-Dissipation theorem, the linear power density for the position thermal noise is given by

$$\tilde{x}_T(\omega) = \frac{\sqrt{S_{\dot{x}_T}(\omega)}}{\omega} = \sqrt{\frac{4k_B T\phi(\omega)\omega_0^2}{m\omega[(\omega_0^2 - \omega^2)^2 + \phi^2(\omega)\omega_0^4]}},$$

for $\omega \geq 0$. The position frequency spectrum has a very high and narrow peak at the normal mode resonance frequency and low tails away from it. If ϕ has little frequency dependence, below resonance the noise goes as $f^{-1/2}$ while at high frequency the dependence is as $f^{-5/2}$.

5˙6. *Thermal noise mitigation.* – To mitigate the effects of thermal noise one can act on the temperature T, which enters as $T^{1/2}$ in the linear spectrum. Cooling a mass suspended using thin wires is complex and is a field of active research. Moreover additional complications are introduced, mainly due to the vibrations generated by the refrigeration equipment. Currently the large interferometers LIGO and Virgo operate at room temperature, but the Japanese detector KAGRA features test masses cooled to around 20 K.

Otherwise thermal noise can be reduced working on the dissipation itself, for example selecting suitable materials. For a test mass suspension, energy dissipation is through the silica suspension fibers that bend during oscillations. The bending energy is a few percent of the total energy of the pendulum, so the overall loss angle can be extremely low.

Current test masses are also made out of fused silica which ensures low thermal noise. However these masses are used as mirrors with a reflective surface obtained with thin coating layers of other materials, like oxides, that have much higher dissipation. This thermal noise from the mirror coatings is currently the main limitation in sensitivity in most of the detection band and intense research of new materials is in progress.

5˙7. *Newtonian noise.* – Variations with time of the local gravitational field cannot be shielded and can mimic perfectly the effect of gravitational waves; the resulting noise is called Newtonian noise. These fluctuations are generated by the motion of large masses, or of small masses in the proximity of the mirrors, or by density variations in the volume surrounding the mirrors. For a mirror suspended near the ground surface, seismic waves are particularly relevant: ground is raised on the crest and lowered in the through. There is more gravitational mass associated with the crest, and it attracts the mirror more than the through. Density variations can also be found in the atmosphere, due to wind, humidity and rain. In general the noise increases with decreasing frequency, as higher masses are involved, setting a limit to the lowest detectable frequency for Earth detectors.

If the mirror is surrounded by material in a symmetric way the effect can be cancelled, and this motivates locating detectors underground, although compression seismic waves would still vary the density asymmetrically. Also, massive objects at short distance can introduce additional Newtonian noise. Research is going on to understand how, by monitoring the ground or air motion with arrays of sensors, the effect of local gravity fluctuations can be predicted and appropriately subtracted.

6. – Measurement noise

6˙1. *Michelson-Morley interferometry.* – In principle changes in the metric can be measured in a single direction by looking at the light reflected by a remote mass. Wavefronts sent by a reference source would come back at times different than expected, exhibiting a phase difference with respect to the source. The phase difference is observed by looking at the intensity of the beam resulting from the interference between the source

and the reflected beam. The achievable precision is limited by the stability of the basic unit of length, which is the source wavelength.

Space-time deformation by gravitational waves varies with direction, making the comparison between optical paths a much better detection procedure. Sending light beams in two arms of identical length and having them interfere at the end of their return trip allows a much higher precision, as many effects that are common to both arms, like those from the source itself, cancel. The intensity from interference provides the relative phase measurement. This makes the Michelson-Morley interferometer an ideal configuration for detection.

Consider an interferometer where light with power P_{in} impinges on a beam splitter, while the power coming out is P_{out}. Suppose that there is a phase difference between the two beams

$$\Delta\varphi = \pi + \alpha + \phi_{OG},$$

where the static phase $\pi + \alpha$ can be chosen and ϕ_{GW} is the phase difference generated by the gravitational wave. $\alpha = 0$ means destructive interference in transmission, it is usually said that the interferometer is "on the dark fringe". Then

$$P_{\text{out}} = P_{\text{in}} \sin^2\left(\frac{\alpha + \phi_{OG}}{2}\right).$$

For $\phi_{OG} \ll 1$:

$$P_{\text{out}} = P_{\text{in}}\left(\sin^2\frac{\alpha}{2} + \frac{1}{2}\phi_{OG}\sin\alpha\right).$$

The interferometer shows a *sensitivity* of the output power to changes in ϕ_{GW}

$$\frac{\mathrm{d}P_{\text{out}}}{\mathrm{d}\phi_{GW}} = P_{\text{in}}\sin\left(\frac{\alpha + \phi_{GW}}{2}\right)\cos\left(\frac{\alpha + \phi_{GW}}{2}\right) = \sin\left(\alpha + \phi_{GW}\right),$$

which is maximum for $\alpha = \pi/2$. This happens when the instrument is at half fringe: $P_{\text{out}} = P_{\bar{\text{out}}} = P_{\text{in}}/2$, that is when half of the light is reflected toward the light source and the other half reaches the output photodiode. Note: in the field of GW sensitivity denotes the minimum gravitational wave amplitude that can be detected. As will be shown below this is set by the noise present in the detector.

In an ideal interferometer the output intensity is subject to statistical fluctuations of the number of detected photons. This sets the *precision* that can be achieved in the measurement. With a normal laser source the number of photons detected per unit time has a Poisson distribution. If there are on average N pulses per second, it can be shown that the noise spectrum over positive and negative frequencies is

$$S_\gamma(\omega) = 2\pi N^2\delta(\omega) + N.$$

Leaving aside the first term $\delta(\omega)$ corresponding to the average DC current, one has that the variance per Hz is N. In the absence of signal the noise present in the photocurrent is

$$\sigma_{P_{\rm out}} = \sqrt{\frac{\eta P_{\rm in}}{h\nu}} \left|\sin\frac{\alpha}{2}\right|,$$

where η is the quantum efficiency of the photodiode.

The signal-to noise ratio in a 1 Hz band, that is for 1 s measurement time is

$$\mathrm{SNR} = \sqrt{\frac{\eta P_{\rm in}}{2h\nu}} \left|\cos\frac{\alpha}{2}\right| \phi_{OG},$$

which is maximum for $\alpha = 0$: it is more convenient to work on the dark fringe ($\alpha = 0$ however kills the signal, so a little offset is present).

In the GW field the sensitivity is the gravitational wave signal that has the same spectral amplitude of the detector noise, that is, the signal-to-noise ratio is 1. The phase sensitivity measured in 1 second is then

$$\phi_{OG}^{\rm min} = \sqrt{\frac{2h\nu}{\eta P_{\rm in}}}.$$

Consider for example a laser source of infrared light with $\lambda = 1064\,\mathrm{nm}$, incident power of 10 W and an ideal detector ($\eta = 1$). The resulting phase noise is

$$\phi_{OG}^{\rm min} \simeq 2.0 \times 10^{-10}\,\mathrm{rad\,Hz^{-1/2}}.$$

For a wave arriving perpendicularly to the interferometer plane, with + polarization acting with opposite signs on the two arms

$$\phi_{OG} = \frac{2\pi}{\lambda} 2L \frac{h}{2} \times 2.$$

One has

$$h^{\rm min} = \frac{\lambda}{4\pi L}\sqrt{\frac{2h\nu}{\eta P_{\rm in}}}\,.$$

This is, as expected, a spectral density with units $\mathrm{Hz}^{-1/2}$. Moving from the table top instrument to the Virgo interferometer one sees that the km scale interferometer is still off the $10^{-23}\,\mathrm{Hz}^{-1/2}$ target by several orders of magnitude (table V).

TABLE V. – *Table top and Virgo interferometer shot noise.*

	L	$h_{\min}\,\mathrm{Hz}^{-1/2}$
Table top	1 m	1.6×10^{-17}
Virgo	3 km	5.3×10^{-21}

6˙2. *Fabry-Perot cavities.* – The response to gravitational waves increases with the effective optical path between the two test masses. The simplest way is to have multiple reflections, building a delay line. A large interferometer, GEO600 near Hannover, has its arms folded once to double their length, building a Michelson-Morley interferometer with 1.2 km arms. However a larger arm length leads to an increased beam size caused by diffraction. This requires that the multiple reflections occur in well-separated areas, implying large-size mirrors, for a relatively small spot size.

The current approach is to use Fabry-Perot resonant cavities, made by a semitransparent mirror at the input and an essentially totally reflective mirror at the end. This is a device frequently used when one has to have precise wavelength measurements, having a sharp phase response to length variations.

A simplified treatment of the Fabry-Perot Cavity considers the incoming and outgoing electric fields E_{in} and E_{out} at the input mirror, leaving out the temporal dependence given by $\exp(-i\omega t)$, and those inside the cavity E_1, E_2, E_3 e E_4. The relations between the various fields are listed in table VI. It is assumed that optical losses are negligible so that $r_i^2 + t_i^2 = 1$, r_i and t_i real. The $\pi/2$ phase lag between transmitted and reflected amplitude is required for dielectric materials. In gravitational wave interferometers r_2 is usually 1 minus a few parts per million.

Solving for E_1 yields

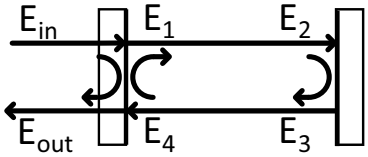
$$E_1 = \frac{t_1}{1 + r_1 r_2 \exp(2ikl)} E_{\mathrm{in}},$$

which, when $\exp(2ikl) = -1$, reaches the maximum value

$$E_1 = \frac{t_1}{1 - r_1 r_2} E_{\mathrm{in}}.$$

TABLE VI. – *Fields in a Fabry-Perot cavity.*

$$\begin{aligned} E_1 &= t_1 E_{\mathrm{in}} + i r_1 E_4 \\ E_2 &= \exp(ikL) E_1 \\ E_3 &= i r_2 E_2 \\ E_4 &= \exp(ikl) E_3 \\ E_{\mathrm{out}} &= i r_1 E_{\mathrm{in}} + t_1 E_4 \end{aligned}$$



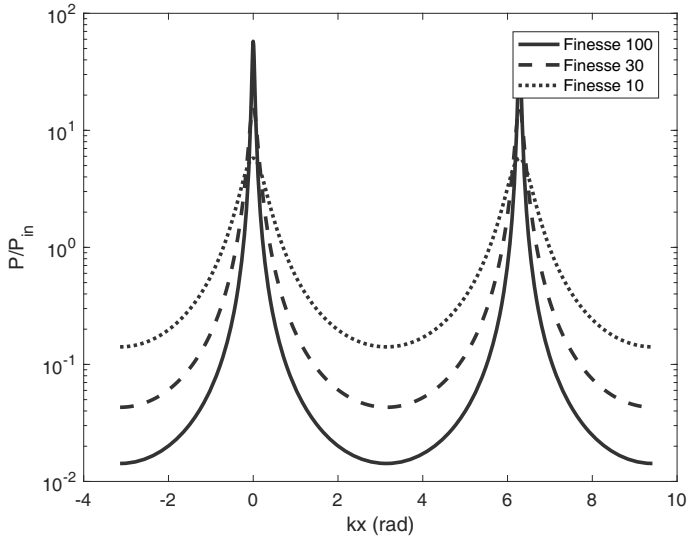


Fig. 6. – Fabry-Perot cavity circulating power with different finesse. Note the slope difference near the maximum.

The various resonances are at optical frequencies

$$\nu_n = \left(n + \frac{1}{2}\right)\frac{c}{2L}$$

or for a cavity length $L_0 = k\lambda/2$. The separation between resonances is called *free spectral range* (FSR) which, for a 3 km cavity, is 50 kHz. The circulating power at resonance is

$$P_1 = \frac{\epsilon_0 c}{2}|E_1|^2 = \left(\frac{t_1}{1 - r_1 r_2}\right)^2 P_{\text{in}},$$

which can be written as

$$P_1 = \left(\frac{t_1}{1 - r_1 r_2}\right)^2 \frac{P_{\text{in}}}{1 + \frac{4\mathcal{F}^2}{\pi^2}\sin^2[x/2]},$$

where $x := 2\pi L/\lambda - k\pi$ is the length detuning from resonance expressed in radians, while $\mathcal{F}$ defines the *finesse* of the cavity

$$\mathcal{F} := \frac{\pi\sqrt{r_1 r_2}}{1 - r_1 r_2}.$$

Figure 6 shows the power circulating in the cavity as function of x.

The reflected field $E_{\rm out}$ is

$$E_{\rm out} = ir_1 E_{\rm in} + ir_2 t_1 \exp(i2kL) E_1 = i\mathcal{R} E_{\rm in}$$

with

$$\mathcal{R} = \frac{r_1 + r_2 \exp(i2kL)}{1 + r_1 r_2 \exp(i2kL)}$$

For $\mathcal{F} \gg 1$, $r_1 r_2 \simeq 1 - \pi/\mathcal{F}$ and

$$r_2 \mathcal{R} = \frac{1 - \pi/\mathcal{F} \exp(i2kL)}{1 - (1 - \pi/\mathcal{F}) \exp(i2kL)} .$$

Near the resonance the sensitivity of the reflected phase ϕ to length variations can be computed, giving

$$\frac{\delta\phi}{\delta L} = \frac{8\mathcal{F}}{\lambda} .$$

This is obtained, assuming negligible absorption, for $(1 - r_2^2)\mathcal{F}/\pi \ll 1$, that is for moderate finesse and low losses.

As in the case of the simple Michelson interferometer, this does not correspond to the maximum precision, as little light is reflected from the cavities, increasing noise in the measurement. Working out the details, that can be found for example in [16, 17], one sees that typically there is a factor 2 reduction in the phase response of the cavity.

The phase difference between two cavities leads to the level

$$h^{\rm min} = \frac{\lambda}{4\mathcal{F}L} \sqrt{\frac{h\nu}{\eta P_{\rm in}}} \sqrt{1 + \frac{4\mathcal{F}^2}{\pi^2} \sin^2\left(\frac{\Omega L}{c}\right)},$$

to be compared with the Michelson case

$$h^{\rm min} = \frac{\lambda}{2\pi L} \sqrt{\frac{h\nu}{\eta P_{\rm in}}}$$

At low frequency $\Omega \ll c/L$ the sensitivity gain is

$$\frac{2\mathcal{F}}{\pi} .$$

For $\mathcal{F} = 450$ the gain is 286, bringing the sensitivity of Advanced Virgo to

$$1.9 \times 10^{-23}\,{\rm Hz}^{-1/2}.$$

6˙3. *Power recycling*. – According to the initial discussion the sensitivity of gravitational wave interferometer increases with the laser power. It is possible to recycle the light coming from the input port of the interferometer, which is intense since we have a dark fringe at the output port. By positioning another semitransparent mirror before the beam splitter one creates a resonant cavity made of the full Michelson-Morley interferometer with Fabry-Perot arms, that reflects light with a well-defined phase. By positioning this recycling mirror suitably one can build another resonant cavity in which light power can build up. The total absorption in the optical system will ultimately limit the available circulating power: however recycling gains $C = 30$–50 are currently achieved. The resulting sensitivity is

$$h^{\min} = \frac{\lambda}{4\mathcal{F}L}\sqrt{\frac{h\nu}{\eta C P_{\rm in}}}\sqrt{1+\frac{4\mathcal{F}^2}{\pi^2}\sin^2\left(\frac{\Omega L}{c}\right)}.$$

6˙4. *Standard quantum limit*. – In the above discussion it appears that having a high power entering the interferometer reduces the shot noise contribution. However this increases the size of the quantum fluctuations in the photon number, causing radiation pressure variations.

One reflected photon per unit time gives an acceleration to a mirror of mass M

$$a = \frac{2h\nu}{Mc}\,.$$

As seen before the one-sided noise spectrum for photon counting is

$$\frac{2P_{\rm in}}{h\nu}\,.$$

Then the acceleration spectrum is

$$S_a(f) = \left(\frac{2h\nu}{Mc}\right)^2\frac{2P_{\rm in}}{h\nu} = \frac{8P_{\rm in}h}{M^2\lambda c}\,.$$

The resulting position noise spectrum is

$$S_x(f) = \frac{P_{\rm in}h}{2\pi^4 M^2\lambda c^2 f^4}\,,$$

giving a f^{-2} behaviour in the linear power spectrum. Taking into account that there are two mirrors per arm and that the arms are anticorrelated, the resulting noise from radiation pressure is

$$h_{RP}(f) = \sqrt{\frac{2P_{\rm in}h}{\lambda c}}\frac{1}{\pi^2 LMf^2}\,.$$

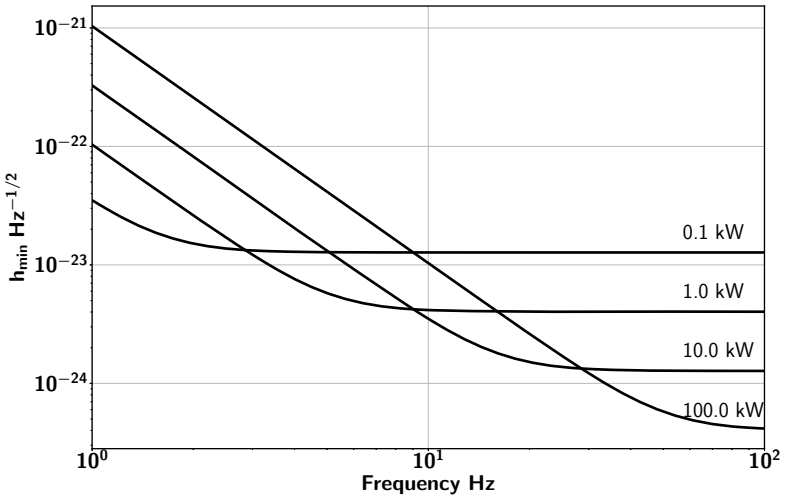


Fig. 7. – Standard quantum limit.

On the other hand it was found that at low frequency the shot noise leads to a spectrum

$$h_{\min} = \frac{\lambda}{4\mathcal{F}L}\sqrt{\frac{h\nu}{P_{\text{in}}}}\,.$$

The sum in quadrature of these two contributions is shown in fig. 7, varying the input power on the beam splitter. This results relies on the photon statistics followed for normal laser light. Nowadays the large interferometers have introduced optical schemes that manipulate the quantum states of light, reducing the uncertainty of the phase measurement at the cost of an increased variance of the intensity. This makes use of *squeezed quantum states* of light, that are a fascinating application of the quantum theory of light using materials with non-linear dielectric response, but is definitely a more advanced subject (see for example [18] and references therein).

7. – Noise curve

In the previous paragraphs the various fundamental contributions to the noise of a gravitational wave interferometer were discussed. A detector able to record gravitational waves over a broad frequency band is the result of many compromises between conflicting requirements. In fig. 8 the fundamental contributions to the noise of Advanced Virgo in its design configuration, including a signal recycling scheme, are presented. In a non-ideal apparatus there are many paths for external disturbances to enter the instrument. Reducing these takes a long experimentation time, that is in contrast with observation time to provide gravitational wave data to the scientific community. A future with a redundant network of interferometers will allow to spend the time necessary to bring each instrument to its design level of performance.

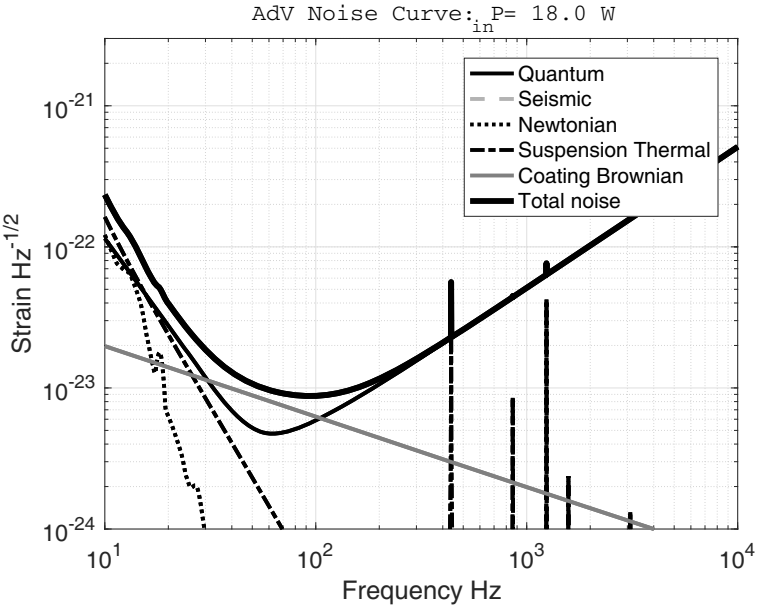


Fig. 8. – Fundamental noise curve.

8. – Ending remarks

These lectures aimed at being a starting point to experimental gravitational wave research. Measuring changes in space-time asks for understanding a variety of physics subjects, from quantum mechanics, to statistical physics, to cosmology. It is the place for curiosity, ingenuity, perseverance, while looking for the unexpected. The wish is that, at the end, new horizons opened and new, fascinating fields of research will attract the reader.

REFERENCES

[1] EINSTEIN A., *Ann. Phys. (Leipzig)*, **49** (1916) 284.
[2] EINSTEIN A., *Preuss. Akad. Wiss. Berlin Sitzber.*, Jan–Juni 1916, p. 688.
[3] FEYNMAN R. P., *An expanded version of the remarks by r.p. feynman on the reality of gravitational waves*, in *The Role of Gravitation in Physics Report from the 1957 Chapel Hill Conference Dean Rickles 1957, Cécile M. DeWitt (eds)*, edited by DEAN RICKLES C. M. D., pp. 279–281.
URL `http://www.edition-open-sources.org/media/sources/5/Sources5.pdf`.
[4] ABBOTT B. P. *et al.*, *Phys. Rev. Lett.*, **116** (2016) 061102.
URL `https://link.aps.org/doi/10.1103/PhysRevLett.116.061102`.
[5] SCHUTZ B. F., *A First Course in General Relativity*, 2nd edition (Cambridge University Press) 2009.

[6] Brazier A. *et al.*, *Mon. Not. R. Astron. Soc.*, **458** (2016) 1267.
URL https://dx.doi.org/10.1093/mnras/stw347.
[7] Królak A., Tinto M. and Vallisneri M., *Phys. Rev. D*, **70** (2004) 022003.
URL https://link.aps.org/doi/10.1103/PhysRevD.70.022003.
[8] Weisberg J. M. and Taylor J. H., *The relativistic binary pulsar b1913+16*, in *Proceedings of "Radio Pulsars" Chania, Crete*, edited by Bailes M., Nice D. J. and Thorsett S. E., ASP. Conf. Series 2003.
[9] Papoulis A. and Pillai S. U., *Probability, Random Variables and Stochastic Processes* (McGraw-Hill) 2001.
[10] Giazotto A., *Phys. Rep.*, **182** (1989) 365.
URL http://www.sciencedirect.com/science/article/pii/0370157389900872.
[11] Accadia T. *et al.*, *J. Instrum.*, **7** (2012) P03012.
URL https://iopscience.iop.org/article/10.1088/1748-0221/7/03/P03012.
[12] Braccini S. *et al.*, *Rev. Sci. Instrum.*, **64** (1993) 310.
URL https://doi.org/10.1063/1.1144249.
[13] Losurdo G. *et al.*, *Rev. Sci. Instrum.*, **70** (1999) 2507.
URL https://doi.org/10.1063/1.1149783.
[14] Callen H. B. and Welton T. A., *Phys. Rev.*, **83** (1951) 34.
URL https://link.aps.org/doi/10.1103/PhysRev.83.34.
[15] Saulson P. R., *Phys. Rev. D*, **42** (1990) 2437.
URL https://link.aps.org/doi/10.1103/PhysRevD.42.2437.
[16] Vinet J.-Y., Meers B., Man C. N. and Brillet A., *Phys. Rev. D*, **38** (1988) 433.
URL https://link.aps.org/doi/10.1103/PhysRevD.38.433.
[17] Vinet J.-Y., *Optical Modeling of Gravitational Wave Interferometers*, in *Proceedings of the International Summer School on Experimental Physics of Gravitational Waves, Urbino, Italy, 1999* (World Scientific Publishing) 2000, pp. 148–257.
[18] Buonanno A. and Chen Y., *Phys. Rev. D*, **64** (2001) 042006.
URL https://link.aps.org/doi/10.1103/PhysRevD.64.042006.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200003

A primer on a real gravitational wave detector

Fulvio Ricci

Dipartimento di Fisica, Sapienza Università di Roma and INFN, Sezione di Roma
Piazza A. Moro 5, 00185 Roma, Italy

Summary. — Despite the simple detection principle, a gravitational wave interferometer is an extremely complex instrument. In these lectures we will try to summarise some of the main features of the detector. We will start from the very basic concepts and gradually advance to a more general treatment. We will focus our discussion on the solutions adopted in the second generation of instruments, by means of which now the era of the gravitational astronomy has started. In the final sections, in view of the development of the future gravitational wave detectors, we will discuss how to reduce two main noise sources: the thermal noise via the use of cryogenic techniques and the readout noise via the strategies conceived to circumvent the quantum limit of the detector.

1. – Introduction

The interferometers for the detection of gravitational waves (GW) as advanced LIGO and Virgo [1, 2] are transducers, which convert the GW space-time strain, a signal at frequency ω in acoustic bandwidth, into an optical signal by modulating the optical field, which oscillates at frequency Ω. It has been pointed out in an elegant paper by Peter Saulson in 1997 [3], that the conversion of the GW strain into an output at the light frequencies $\Omega \pm \omega$ is reminiscent of the class of devices known as parametric amplifiers. The key component of a parametric amplifier is a nonlinear reactance, to which both the

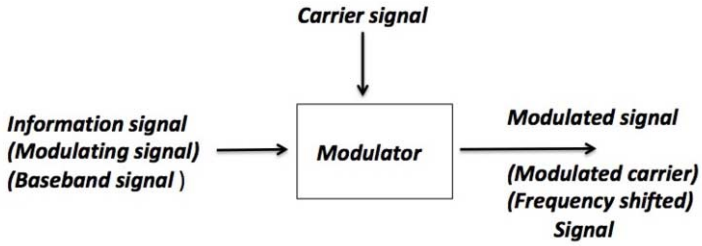


Fig. 1. – The modulation scheme.

input signal and a *pump* signal of a much higher power level are applied. The nonlinear reactance produces other signals at the combination frequencies $m\Omega \pm n\omega$, where m and n are integers. The signal process at the basis of a parametric transducer is the modulation and for this reason we will start our primer on the real GW detector with a *laudatio* of the modulation principle. Then we describe the optical configuration of the GW and its limits in more details. In the second part we will focus on two intrinsic limitations of the instrument sensitivity, which are due to the thermal and readout (optical) noise.

2. – The modulation

The basic method for transmitting information, *i.e.* a signal in the audio bandwidth, is to imprint it in another signal at higher frequency, the *carrier* or *pump signal* (see fig. 1). The modulation, permits to have a *frequency translation*, where the modulation wave at lower frequency, *the baseband signal*, is imprinted in the signal transporting the information, *the carrier*, at higher frequency. To do that some characteristics of the carrier are varied in accordance with the modulation wave.

Different modulation approaches exist: Amplitude Modulation (AM), Frequency Modulation (FM) and Phase modulation (PM). All of those are nonlinear process and among them the AM is the simplest one, whose analytical representation is

$$E(t) = A_c[1 + m(t)\, M]\cos(\Omega t + \theta], \quad \text{with} \quad m(t) = \cos(\omega t + \phi) \quad \text{and} \quad \omega \ll \Omega. \tag{1}$$

The bandwidth associated to the amplitude modulation is limited to 2ω around the carrier frequency Ω and this is the main reason why, in order to transmit high fidelity sound the radio stations are using the FM approach. For FM the carrier's frequency varies with the signal's input, while for PM the carrier's phase varies with the signal's input. As we said, the two modulations differ mainly in the transmission bandwidth. We focus our attention on PM and on the simplest case of a phase modulated by a monochromatic signal, *i.e.*

$$E(t) = A_c \cos[\Omega t + M\, m(t)], \quad \text{with} \quad m(t) = \cos(\omega t + \phi) \quad \text{and} \quad \omega \ll \Omega, \tag{2}$$

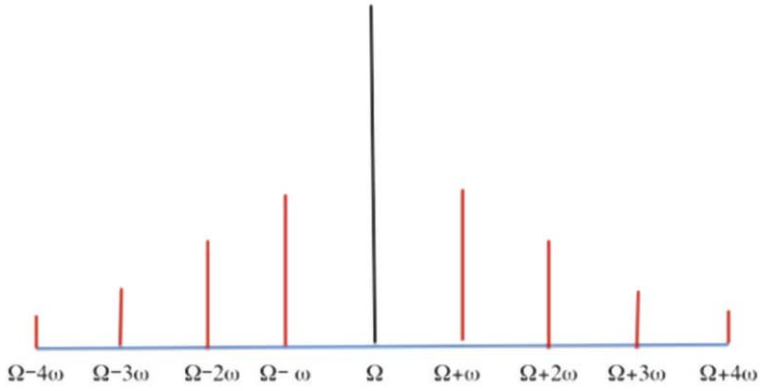


Fig. 2. – The spectrum of a phase modulated signal with a monochromatic line.

where the parameter $M \ll 1$ is the modulation depth. At the first order in M, the amplitude of the PM modulated signal, which is given in the complex plane by the following formula

$$E_o(t) = \frac{A_c}{\sqrt{2}} e^{iM[\omega t+\phi]}, \tag{3}$$

has as quadrature components E_c and E_s,

$$E_c \simeq \frac{A_c}{\sqrt{2}}, \qquad E_s \simeq \frac{A_c}{\sqrt{2}} M \sin(\omega t + \phi)$$

and it is evident that the modulation signal is present just in one of the two quadratures. In the case of a monochromatic modulation the representations of AM and PM signals in the Fourier domain are different. In the first case we are dealing just with three lines at frequency Ω, $\Omega + \omega$ and $\Omega - \omega$, while in the PM case the expansion of eq. (3) in series of Bessel functions,

$$E_o(t) \simeq \frac{A_c}{\sqrt{2}} e^{iM[\omega t+\phi]} = J_o(M) + iJ_1(M)e^{i\omega t+\phi} + iJ_{-1}(M)e^{-i\omega t+\phi} + \ldots, \tag{4}$$

with $J_o(M) = 1$ and $J_1(M) = J_{-1}(M) \simeq M/2$, defines a spectrum with equally spaced lines around the carrier at distance ω (see fig. 2). We note that in the case of PM, sideband fields have $\pi/2$ constant phase shift with respect to the carrier field expressed by the factor i in front of the corresponding terms in the formula (4); therefore its sum is always orthogonal to the carrier field.

Let us now explain why modulation is crucial for the GW detection.

Two optical elements are crucial in an interferometer: the mirror and the 50% beam splitter. Both of them act as four port linear devices, connecting the input to the transmitted and reflected fields on both sides of the optical element. In the mirror case (see

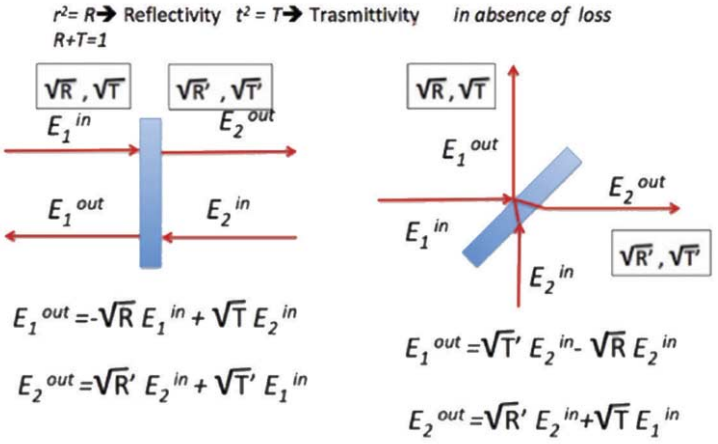


Fig. 3. – Mirror and beam splitter modelled as four port devices of the input-output optical fields.

fig. 3) the quadratures of input light $E^{\rm in}$ impinging on a mirror are related to those of the output beam $E^{\rm out}$ via a matrix whose elements depend on the reflectivity R and the transmittivity T

$$\begin{pmatrix} E_{1c}^{\rm out} \\ E_{1s}^{\rm out} \\ E_{2c}^{\rm out} \\ E_{2s}^{\rm out} \end{pmatrix} = \begin{pmatrix} -\sqrt{R} & 0 & \sqrt{T} & 0 \\ 0 & -\sqrt{R} & 0 & \sqrt{T} \\ \sqrt{T} & 0 & -\sqrt{R} & 0 \\ 0 & \sqrt{T} & 0 & \sqrt{R} \end{pmatrix} \begin{pmatrix} E_{1c}^{\rm in} \\ E_{1s}^{\rm in} \\ E_{2c}^{\rm in} \\ E_{1s}^{2} \end{pmatrix}, \tag{5}$$

while, in the case of the 50% beam splitter, the transfer function matrix between input and output is

$$\begin{pmatrix} -\sqrt{\frac{1}{2}} & 0 & \sqrt{\frac{1}{2}} & 0 \\ 0 & -\sqrt{\frac{1}{2}} & 0 & \sqrt{\frac{1}{2}} \\ \sqrt{\frac{1}{2}} & 0 & -\sqrt{\frac{1}{2}} & 0 \\ 0 & \sqrt{\frac{1}{2}} & 0 & \sqrt{\frac{1}{2}} \end{pmatrix}. \tag{6}$$

When the mirror is displaced by a quantity $x(t) \ll \omega/2\pi c$, the input and out field change: the mirror motion determines a dephasing of the reflected contributions of the

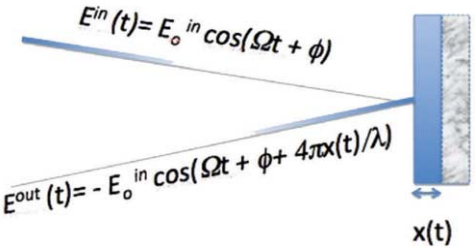


Fig. 4. – The reflected signal modulated by the mirror motion.

input optical fields to the output ones

$$E_1^{\rm out}(t) = -\sqrt{R}\, E_1^{\rm in}(t - 2x(t)/c) + \sqrt{T}\, E_2^{\rm in}(t),$$
$$E_2^{\rm out}(t) = \sqrt{T}\, E_1^{\rm in}(t) + \sqrt{R}\, E_2^{\rm in}(t + 2x(t)/c).$$

In fig. 4 we sketch the output electric fields phase modulated by the mirror motion. Once we analyse the input and output fields in terms of quadrature components, we end up with the same results obtained before for the PM modulation signal: the modulation effect due to the mirror motion is encoded just in the *s-quadrature* of the electric field.

The simple considerations made before were related to the case of a single modulation frequency. The generalisation is made passing to the frequency domain via the Fourier transform of the modulation signal and the corresponding optical field. The complete treatment of this problem ends up by deriving the spectral component $L_s(\omega)$ of the light, which contains the information of the mirror motion: it results that it is proportional to $X(\omega)$, the Fourier transform of $x(t)$, weighted by the amplitude of the optical field

$$L_s(\omega) = E_o X(\omega) \left(\frac{\Omega}{2\pi c}\right).$$

3. – The detection of the modulation component

In the case of a light modulation, if we send directly the beam to a photodiode, we cannot extract information associated to the modulation signal. In fact we have

$$EE^* = \frac{{A_c}^2}{2} e^{i[\Omega t + M\cos(\omega t + \phi)]} e^{-i[\Omega t + M\cos(\omega t + \phi)]} = \text{constant}.$$

The consequence is that we need to consider a more sophisticated approach. In a Michelson interferometer a modulated light is divided in two beams by a 50% beam splitter, then the two beams are recombined on the same optical elements after having travelled along the two arms of the optical device. Let us define the reflectivity of the two end

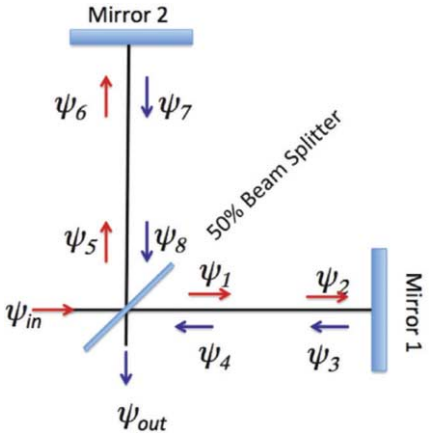


Fig. 5. – The reflected signal modulated by the mirror motion.

mirrors R_1 and R_2 and the two arm lengths l_1 and l_2 (see fig. 5). The optical fields ψ at the input and output of the interferometer are

$$\psi_{\rm in} = K e^{i\chi}, \qquad \psi_1 = \frac{1}{\sqrt{2}}\psi_{\rm in}, \qquad \psi_2 = \psi_1 e^{-ikl_1}, \qquad \psi_3 = i\sqrt{R_1}\psi_2, \qquad \psi_4 = \psi_3 e^{-ikl_1},$$

$$\psi_5 = i\frac{1}{\sqrt{2}}\psi_{\rm in}, \qquad \psi_6 = \psi_5 e^{-ikl_2}, \qquad \psi_7 = i\sqrt{R_2}\psi_6, \qquad \psi_8 = \psi_8 e^{-ikl_2}$$

and finally

$$\psi_{\rm out} = i\frac{1}{\sqrt{2}}\psi_4 + \frac{1}{\sqrt{2}}\psi_8.$$

With a bit of algebra we derive the well known formula for the intensity of the optical field at the output

$$|\psi_{\rm out}|^2 = \frac{1}{4}|\psi_{\rm in}|^2(R_1 + R_2)\left[1 + C\cos(2k\delta l)\right], \tag{7}$$

where $\delta l = l_1 - l_2$ and C is the interferometer contrast,

$$C = \frac{2\sqrt{R_1 R_2}}{R_1 + R_2} = \frac{|\psi_{\rm out}|^2{}_{\rm max} - |\psi_{\rm in}|^2{}_{\rm min}}{|\psi_{\rm out}|^2{}_{\rm max} + |\psi_{\rm in}|^2{}_{\rm min}}\,. \tag{8}$$

The GW interferometer with $l = \frac{l_1+l_2}{2}$ has almost equal arms $l_1 \simeq l_2$ and it set to fulfil the *dark fringe* condition, *i.e.* $\delta l = 0$, since in this way it is insensitive to the noise fluctuations of the injected light. In the absence of a modulated light, the gravitational wave signal changes the difference in the arm length by a quantity $h\,l \ll 1$ and the light intensity at the output of the interferometer given by eq. (7) results to be proportional

to h^2, making the signal detection impossible. We note also that even for the case of an injected modulated light, in the case of a perfect symmetric interferometer ($l_1 = l_2$), the dark condition is verified for all the frequency components of the output light, making again the detection impossible.

To overcome this problem we consider a slightly asymmetric interferometer on which we have injected a modulated light. We call this static difference between the two arms the Schnupp asymmetry $\Delta l_{\rm Sc} = l_1 - l_2$, that we choose of the order of $\Delta l_{\rm Sc} \sim 10^{-3} l$. The asymmetry plays the crucial role of releasing the frequency degeneracy: it makes the output field no longer dark for the sidebands even if the carrier is at dark.

When a GW signal is present and the difference in the arm length is $\delta l = \Delta l_{\rm Sc} + h\, l$, the output of the Michelson interferometer is

$$\psi^m_{\rm out} = i\, \psi_{\rm in}\, e^{i\Omega l/c} \left[J_o\, l\, h\, \frac{\Omega}{c} + 2 J_1 \sin\left(\frac{\omega}{c}\Delta l_{\rm Sc}\right) \cos\left(\omega t + 2\frac{\omega}{c} l\right)\right], \tag{9}$$

with the corresponding intensity

$$|\psi^m_{\rm out}|^2 = |\psi_{\rm in}|^2 \Bigg\{ \left[J_o\, l\, h\, \frac{\Omega}{c}\right]^2 + 2{J_1}^2 \left[1 + 2\cos\left(2\omega t + 4\frac{\omega}{c} l\right)\right] \sin^2\left(\frac{\omega}{c}\Delta l_{\rm Sc}\right) + 2 J_o\, J_1\, l\, h\, \frac{\Omega}{c} \sin\left(\frac{\omega}{c}\Delta l_{\rm Sc}\right) \cos\left(\omega t + 2\frac{\omega}{c} l\right) \Bigg\}. \tag{10}$$

The last term in parentheses of this long formula shows that in a phase component of the modulation we can extract at the interferometer output a signal linearly dependent on the product of the arm length l and the GW strain h.

4. – The readout of the output signal

Now the question is how to extract this signal, which oscillates at the modulation frequency ω. In the present experiments there are two main approaches: heterodyne and DC readout. The general scheme of both systems is sketched in fig. 6. In the case of the first generation of LIGO and Virgo a heterodyne readout scheme was adopted employing radio frequency (RF) modulation-demodulation techniques. The most serious limitation of this system is due to phase noise of the RF modulation and for this reason a new approach has been adopted in the advanced detectors. The DC readout system is a special case of homodyne detection, which makes use of the highly stabilised and filtered carrier light as local oscillator for the readout. We take advantage of the slight detuning of the interferometer arms, so that in practice the role of the local oscillator is played by the fraction of the light that leaks into the signal port due to the residual interferometer asymmetry of the arms.

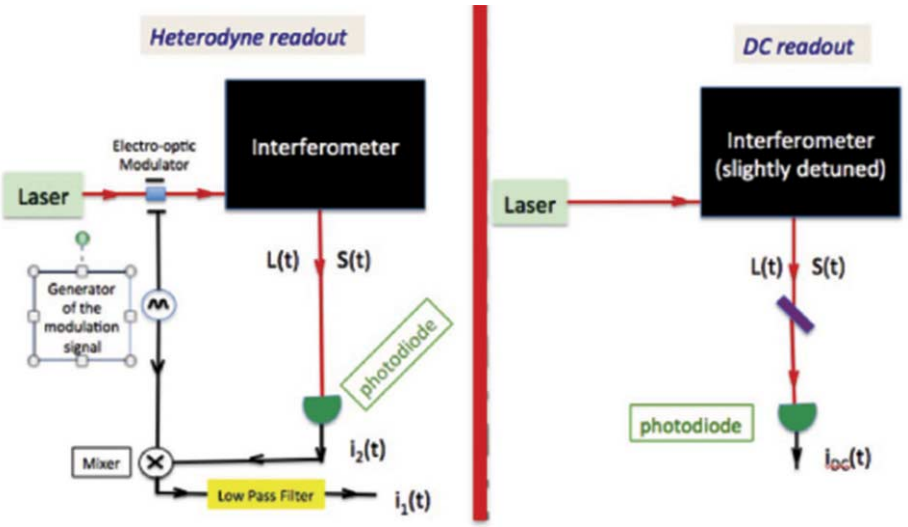


Fig. 6. – On the left we show the sketch of a heterodyne readout, on the right the DC one.

5. – The Fabry-Perot cavities as Michelson arms

The signal detected at the output of the Michelson is due to the light phase shift induced by the GW signal, which is proportional to the arm length l. To increase the phase shift in the GW detectors the light is trapped in each arm into a Fabry-Perot (FP) cavity whose length is $\sim l$. In practice this implies adding two semi-trasparent mirrors near the beam splitter with reflectivity and transmittivity $R_{\rm in} = r_{\rm in}^2$ and $T_{\rm in} = t_{\rm in}^2$. The computation of the optical fields transmitted and reflected by the cavity is done using the same approach applied in the Michelson case. The FP end mirrors are the Michelson ones at very high reflectivity. For simplicity we also indicate the transmittivity and reflectivity of these mirrors as $R_{\rm end} = r_{\rm end}^2$ and $T_{\rm end} = t_{\rm end}^2$ and we relate our consideration to fig. 7. Here we have

$$\psi_{\rm in} = Ke^{i\chi}, \quad \psi_1 = t_{\rm in}\psi_{\rm in} + i\, r_{\rm in}\psi_4, \quad \psi_2 = \psi_1 e^{-ikl}, \quad \psi_3 = i\, r_{\rm end}\psi_2, \quad \psi_4 = \psi_3 e^{-ikl},$$

starting from which, we obtain

$$\psi_r = ir_{\rm in}\psi_{\rm in} + t_{\rm in}\psi_4 = i\psi_{\rm in}\frac{r_{\rm in} + r_{\rm end}\, e^{-i\,2\,k\,l}}{1 + r_{\rm in}r_{\rm end}\, e^{-i\,2\,k\,l}}\,, \tag{11}$$

$$\psi_t = t_{\rm end}\psi_2 = \psi_{\rm in}\frac{t_{\rm in}t_{\rm end}\, e^{-ikl}}{1 + r_{\rm in}r_{\rm end}\, e^{-i2kl}} \tag{12}$$

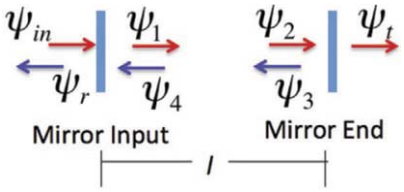


Fig. 7. – The reference scheme for a Fabry-Perot cavity.

and the corresponding intensities are

$$|\psi_r|^2 = \frac{r_{\rm in}^2 r_{\rm end}^2 + 2\, r_{\rm in} r_{\rm end} \cos 2kl}{1 + r_{\rm in}^2 r_{\rm end}^2 + 2\, r_{\rm in} r_{\rm end} \cos 2kl}\,, \tag{13}$$

$$|\psi_t|^2 = \frac{t_{\rm in}^2 t_{\rm end}^2}{1 + r_{\rm in}^2 r_{\rm end}^2 + 2\, r_{\rm in} r_{\rm end} \cos 2kl}\,. \tag{14}$$

Here the resonance effect is evident when the light phase is increased by exactly $2\,\pi$ each two reflections. We are dealing with another interferometric effect and the sharpness of the resonance lines depends on the reflectivity of the mirrors.

To qualify the status of the optical cavity we plot in the complex plane the ratio $\psi_r/\psi_{\rm in}$, derived from formula (11). The resulting curve depends on the relative values of $r_{\rm in}$ and $r_{\rm end}$ and, in the optical jargon originally born in the radio frequency community, the cavity state can be classified as under- or overcoupled (see fig. 8). In the FP cavities installed in the arms of a GW Michelson interferometer the mirror parameters are such that we have $r_{\rm in} < r_{\rm end}$. This implies that the GW FP cavities are overcoupled and almost all the light is reflected back by the cavity.

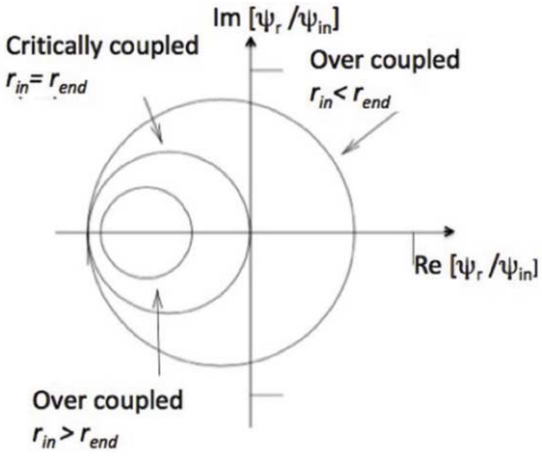


Fig. 8. – The ratio $\psi_r/\psi_{\rm in}$ plotted in the complex plane for different values of $r_{\rm in}$ and $r_{\rm end}$.

Moreover, the main features of the cavity are quantified in terms of few important parameters: FSR, the Free Spectral Range accounting for the distance between two adjacent resonance peaks, the full width at half of the maximum of the resonance curve FWHM, the Finesse $\mathcal{F}$ and the light storage time τ_s in the cavity:

$$\text{FSR} = \frac{c}{2\,l} \qquad \text{FWHM} \simeq \frac{c}{2\pi l}\left(\frac{1 - r_{\text{in}} r_{\text{end}}}{\sqrt{r_{\text{in}} r_{\text{end}}}}\right),$$

$$\mathcal{F} = \frac{\text{FSR}}{\text{FWHM}} \simeq \pi \frac{\sqrt{r_{\text{in}} r_{\text{end}}}}{1 - r_{\text{in}} r_{\text{end}}}\,, \tag{15}$$

$$\tau_s = \frac{1}{\pi}\frac{2l\,\mathcal{F}}{c}\,. \tag{16}$$

Often the phase delay associated to the light trap is interpreted in terms of an effective number of bunches of the light between the two mirror, N_{eff}, or even in terms of and effective optical path of the light, l_{eff}. This approach permits an intuitive calculation of the GW signal amplification factor. Finally, we recall that a more detailed mathematical treatment of the GW signal interaction with the cavity shows that it acts as a low pass filter; the effect is clearly stated by the formula of the effective cavity length expressed as a function of the signal frequency

$$l_{\text{eff}} = \frac{1}{\pi}\frac{2l\,\mathcal{F}}{\sqrt{1 + (\omega\tau_s)^2}}\,. \tag{17}$$

In conclusion, the product $\mathcal{F}\,l$ increases the GW signal amplification of the linear dependence to l_{eff} (eq. (17)), while the existence of the lower angular frequency cut-off of the cavity $\omega_c = 1/\tau_s$ (eq. (16)) reduces the detector bandwidth.

5˙1. *More about the Fabry-Perot cavities.* – In the GW FP cavity the light should bunches back and forth along the same path, but this is just the ideal case if the beam has not a transversal dimension and the two mirrors are perfectly aligned. In a real experiment, the stationary electromagnetic wave in the cavity has an extended spatial profile and bunches between mirrors slightly misaligned. If the mirror surfaces are plane the cavity is far from being stable. In fact, a cavity is stable if a ray launched inside the resonator parallel to the optical axis remains inside the resonator after an infinite number of bounces. If the ray is slightly off axis, we can have still a stable cavity, provided that the ray is reflected in a direction to bring it back toward the cavity center.

It can be demonstrated that the cavity is stable if the product of $g_{\text{in}} = 1 - C_{\text{in}}$ and $g_{\text{end}} = 1 - C_{\text{end}}$ is in the interval $(0, 1)$, where C_{in} and C_{end} are the ratio between the distance between the two mirrors and the curvature radius of the mirror internal surface. For $g_{\text{in}} g_{\text{end}} = 1$ the cavity are marginally stable.

In all the previous consideration we have done the implicit assumption that the fields are infinite plane waves. To give a complete picture of the FP optical properties, we have

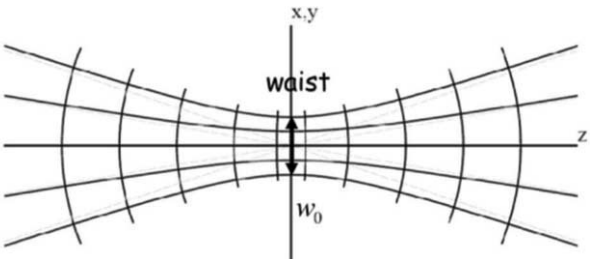


Fig. 9. – The beam profile in a Fabry-Perot cavity.

to treat the stationary light beam in the cavity in the diffraction limit. A resonator generally supports electromagnetic modes having a well-defined and discrete spatial structure and the spatial distribution of the fields of a resonator depends on the shape of its mirrors. In the case of resonators with spherical elements, when we feed the cavity with an almost monochromatic wave along the z-axis, the modes are described by Hermite-Gaussian functions. This result is derived by solving the equation $\nabla^2\psi(x,y,z) + k^2\psi(x,y,z) = 0$, which ends up with the beam profile in the cavity (see fig. 9).

The complete set of solutions of the differential equation for ψ along the propagation z-axis is

$$\psi_{m,n}(x,y,z) = \sqrt{\frac{2}{w^2(z)}}\sqrt{\frac{1}{2^{m+n}m!n!}}\exp\left[-\frac{x^2+y^2}{w^2(z)}\right]\exp\left[ik\frac{x^2+y^2}{R(z)}\right] \times \exp\left[ikz - i(m+n+1)\tan^{-1}\left(\frac{z}{b}\right)\right] H_m\left(\sqrt{(}2)\frac{x}{w(z)}\right) \times H_n\left(\sqrt{(}2)\frac{y}{w(z)}\right), \tag{18}$$

where H_k is the Hermite polynomial of order k.

The first exponential term in (18), $\exp\left[-\frac{x^2+y^2}{w^2(z)}\right]$, defines the Gaussian envelope of the beam profile. It depends on the function $w(z) = w_o\sqrt{1+(\frac{z}{b})^2}$, which is the half-width of the beam, where $b = \pi w_o^2/\lambda$ is the Rayleigh parameter.

The second exponential term, $\exp\left[ik\frac{x^2+y^2}{2\,R(z)}\right]$, is the parabolic wavefront of the beam weighted by the curvature radius function $R(z) = z + \frac{b^2}{z}$.

It follows that the intensity profile of the light beam is characterised by a radius $w(z)$, which is a function of the position along the beam propagation direction. At a radius of $x^2 + y^2 = w^2$ the beam intensity is down by a factor of $1/e^2$ relative to its peak value and $2\,w$ is called the spot size. At $z = 0$ the beam has its minimum radius. This value w_o is the waist radius and the corresponding diameter is called the waist size. At a position $z = b$ the spot size has increased by a factor of $\sqrt{2}$ and continues to increase monotonically. For distances much larger than a Rayleigh parameter the spot size grows linearly with distance and $\theta_o = w_o/b = \lambda/\pi w_o$ is the beam divergence.

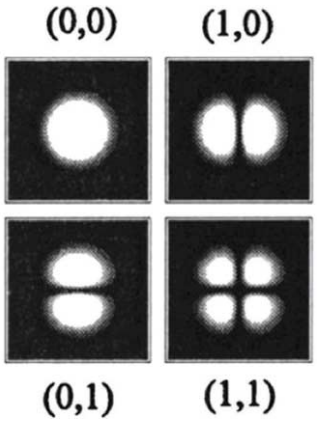


Fig. 10. – The intensity profiles of Hermite-Gauss modes with mode numbers n, m. In this case we present only the non-astigmatic modes ($x = y$).

The other two exponentials of the formula (18) concur to define the *Gouy* phase of the wave

$$\phi(x, y, z) = kz - (m + n + 1) \tan^{-1}\left(\frac{z}{b}\right) + k\frac{x^2 + y^2}{R(z)}, \tag{19}$$

which, after a round trip between the two mirrors located in z_{in} and z_{end}, in the center is changed of the quantity

$$\Delta\phi_{2l} = 2kl - (m + n + 1)\left[\tan^{-1}\left(\frac{z_{\text{end}}}{b}\right) - \tan^{-1}\left(\frac{z_{\text{in}}}{b}\right)\right]. \tag{20}$$

The TEM_{00} with a pure Gaussian profile is selected as the main actor of a GW detector: it is just the basic parent of the mode family. One can easily detect the higher-order transverse modes when the cavity is not perfectly aligned or in general, since the higher-order modes have different resonance frequencies. On resonance the round trip phase change must be $\Delta\phi_{2l} = 2q\pi$ with q arbitrary integer and it can be shown that the Gouy phase contribution can be written in terms of the g factors of the cavity

$$\nu = \nu_{\text{FSR}}\left[(q + 1) + \frac{m + n + 1}{\pi}\cos^{-1}\sqrt{g_{\text{in}}g_{\text{end}}}\right].$$

The first term, $q\,\nu_{\text{FSR}}$, is the planar resonance condition. The second term is an overall shift of the spectrum that depends on the cavity geometry, characterised by g_{in} and g_{end} and mode numbers m and n. In practice, except for confocal or concentric cavities, this implies that different modes with the same frequency will be resonant with a cavity at different lengths and by scanning the resonator length the mode can be seen to appear in sequence (for example, see the mode profiles shown in fig. 10).

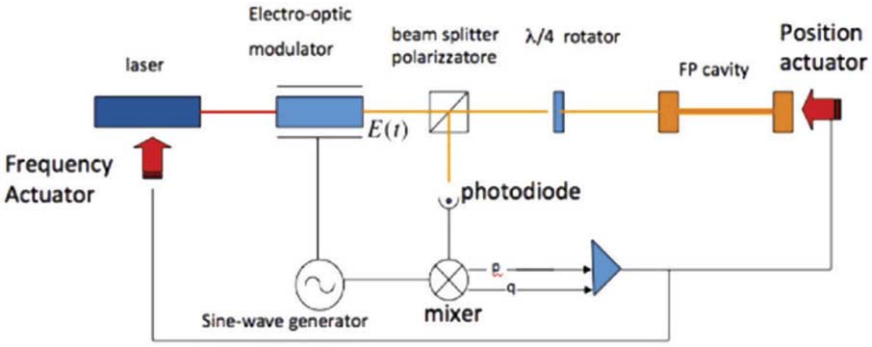


Fig. 11. – The scheme of the Pound-Drever-Hall method to control the FP length.

As we will see later, the detection of the higher modes of the cavity plays a crucial role in the control of the angular degrees of freedom of the GW mirrors.

6. – How to keep the FP cavities in resonance

As we have seen in the previous section the FP resonance condition is obtained when the cavity length, 10^3 m for a GW interferometer, is a multiple of the light wavelength, 10^{-6} m, within an accuracy significantly better than the cavity FWHM $\sim 10^{-8}$–10^{-9} m. This implies to control the cavity length in a stable way with a relative error of the order of 10^{-12}. Again the light modulation provides us with the tool to do it.

In 1946 R. Pound defined the strategy, widely used in radio physics, that today is the method, known as Pound-Drever-Hall, to stabilise the lasers and in general the optical cavities. We refer in particular to the scheme presented in fig. 11. The light is PM modulated at the frequency ω and, as we discussed before, its spectrum is essentially given by a central line, the carrier at Ω, and the two sidebands at $\Omega \pm \omega$. The cavity length is set such that the carrier is resonant $e^{-i2kl} = -1$, while the sidebands are anti-resonants, $e^{\pm 2\omega/c} = e^{\pm\pi}$.

In the presence of a perturbation z of the cavity length the transmitted optical field is

$$\psi_t = i\,\psi_{\rm in}\frac{r_{\rm in} + r_{\rm end}\,e^{i(-2kl-2kz\pm\frac{2\omega l}{c})}}{1 + r_{\rm in}r_{\rm end}\,e^{i(-2kl-2kz\pm\frac{2\omega l}{c})}} \tag{21}$$

and as consequence we have on the detection photodiode of fig. 11 the in-phase component of the transmitted signal proportional to

$$\left\{|\psi_t|^2\right\}_{\text{in phase}} = -4|\psi_{\rm in}|^2 J_0 J_1 t_{\rm in} r_{\rm in} r_{\rm end}(1 + {r_{\rm end}}^2)\frac{\sin 2kz}{1 + (r_{\rm in}r_{\rm end})^4 - 2(r_{\rm in}r_{\rm end})^2\cos 4kz}\,, \tag{22}$$

while the in quadrature component $q = \{|\psi_t|^2\}_{\text{in quadrature}}$ is equal zero.

If we manipulate the previous formula in the case of finesse higher than 1 and $z \ll l$, we derive a simple relation showing what are the parameters affecting the sensitivity of the signal control:

$$\frac{\mathrm{d}}{\mathrm{d}z}\left\{|\psi_t|^2\right\}_{\text{in phase}} \simeq J_o J_1 \frac{8\mathcal{F}}{\lambda}|\psi_{\text{in}}|^2.$$

In practice higher finesse means higher sensitive to length variation; on the other hand higher finesse means also lower FWHM concurring to define the linearity interval of the control feedback loop.

7. – The gravitational wave interferometer

As we said in the previous sections, the FP cavities are used to enhance the optical path of the light, which interacts with the gravitational waves. Thus, in the real configuration of the GW detector we have to deal with several different lengths (see fig. 12): the Michelson arm lengths, l_x and l_y, are in the meter range, while the FP cavity lengths, L_x and L_y, are kilometres. The Schunpp asymmetry, $\Delta l_{xy} = l_x - l_y$ is tens of centimetres. The FP cavities are overcoupled so that the carrier, in resonance in the cavity, is reflected back toward the beam splitter where the two beams are recombined with the carrier on the dark fringe and toward the laser as bright fringe. To gain in terms of light intensity in the interferometer, the light reflected back is recycled by adding the recycling mirror (see fig. 12) between the laser and the beam splitter. We are dealing now with an extra FP cavity where the end mirror is the whole Michelson + FP interferometer and the cavity length is $l_{\text{rc}} = l_{\text{rb}} + \frac{l_x + l_y}{2}$. The mathematical treatment of the optical system is the combination of what it has been presented in the previous sections and, in view of the detection, the computation is focused on the transmittivity of the FP recycling cavity

$$t_{\text{rc}} = \frac{t_{\text{rm}} t_{\text{ifo}} e^{i\Omega_o l_{\text{rc}}}}{1 - r_{\text{rm}} r_{\text{ifo}} e^{i\Omega_o l_{\text{rc}}}}, \tag{23}$$

where t_{rm} and r_{rm} are the usual recycling mirror parameters, while t_{ifo} and r_{ifo} are those of the Michelson interferometer with the FP arms, which assume different expressions for the sidebands

$$r_{\text{ifo}}{}^{\pm} = -e^{\pm i\frac{\omega}{c}(l_x+l_y)}\cos\left(\Delta\frac{\omega}{c}l_{\text{Sc}}\right), \qquad t_{\text{ifo}}{}^{\pm} = \mp e^{\pm i\frac{\omega}{c}(l_x+l_y)}\sin\left(\Delta\frac{\omega}{c}l_{\text{Sc}}\right) \tag{24}$$

and for the carrier

$$r_{\text{ifo}}{}^{c} = e^{i\frac{\Omega}{c}(l_x+l_y)}\left(1 - \frac{\mathcal{F}}{\pi}\epsilon\right), \qquad t_{\text{ifo}}{}^{c} = ie^{i\frac{\Omega}{c}(l_x+l_y)}2\frac{\mathcal{F}\,l\,h}{\pi}\frac{\Omega}{c}\left(1 - \frac{\mathcal{F}}{\pi}\epsilon\right). \tag{25}$$

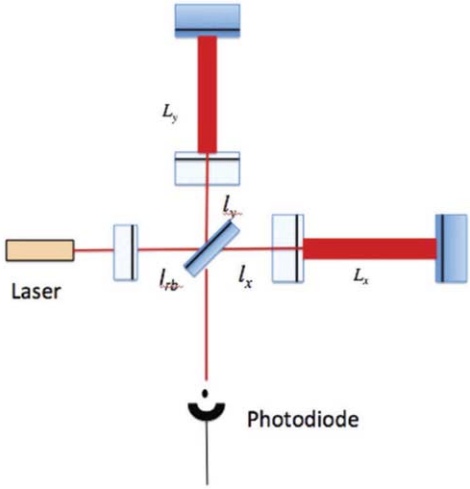


Fig. 12. – The Michelson + the Fabry-Perot cavities + the recycling cavity.

Here we have included in the computation also the optical cavity losses ϵ. Using eqs. (24), (25) in (23) and imposing the resonance conditions

$$e^{i\frac{\Omega}{c}(2l_{\rm rc}+l_x+l_y)} = 1, \qquad e^{i\frac{\omega}{c}(2l_{\rm rc}+l_x+l_y)} = -1,$$

and introducing the finesse $\mathcal{F}_{\rm rc}$ of the recycling cavity, we conclude that at the output we have the signal

$$|\psi_{\rm out}{}^{\rm GW}|^2 = 2J_oJ_1|\psi_{\rm in}|^2\sqrt{\frac{\mathcal{F}_{\rm rc}}{\pi}}\frac{\Omega}{c}\frac{\mathcal{F}}{\pi}l\,h, \tag{26}$$

with a spectral behaviour with respect of the GW signal frequency ν given by

$$|\Psi_{\rm out}{}^{\rm GW}|^2(\nu) = 2J_oJ_1|\psi_{\rm in}|^2\sqrt{\frac{\mathcal{F}_{\rm rc}}{\pi}}\frac{\Omega}{c}\frac{\mathcal{F}}{\pi}l\sqrt{\frac{\nu^2}{\nu^2+{\nu_s}^2}}H(\nu), \tag{27}$$

where $\nu_s = \frac{c}{4l}\mathcal{F}$ is the cut-off frequency of the FP arm cavities (see formula (16) in sect. **5**).

8. – The interferometer control

In sect. **6** we presented the method for keeping a single FP cavity in resonance. To acquire a stable mirror lock there are several constraints that we can express following different points of view. For example we should verify that the time to cross resonance is higher then the light storage time; this condition can be translated into an upper limit

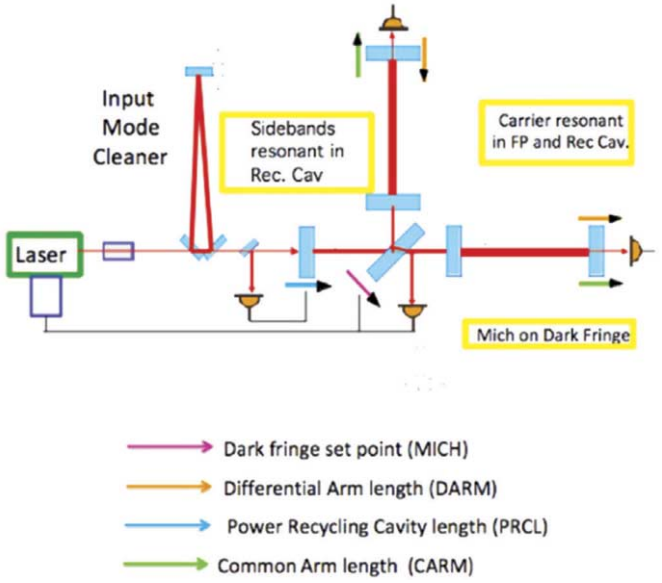


Fig. 13. – The GW interferometer equipped with the photodiode for the longitudinal control.

of the mirror velocity $v_{\max_a}$, which depends on the length L_c and finesse of the cavity $\mathcal{F}_c$, *i.e.*

$$v_{\max_a} = \frac{\pi \lambda c}{4 L_c \mathcal{F}_c{}^2} .$$

It is useful also to look at the maximum oscillation velocities of the mirror that can be compensated by the feedback, $v_{\max_b}$: this is a quantity directly related to the feedback loop bandwidth $\Delta\nu_{fb}$

$$v_{\max_b} = \frac{\pi \lambda \, \Delta\nu_{fb}}{\mathcal{F}_c} .$$

Indeed, the GW optical system must have more cavities in resonance and the Michelson interferometer output has to be kept on the dark fringe. This implies that we need a much more complicated feedback strategy also because we are dealing with suspended optics and we need to achieve a typical lock accuracy in the root mean square displacement of the order of 10^{-12} m. In addition we must impose that the control system has to act on the mirrors without reintroducing noise in the detection bandwidth.

The suspended optical elements are swinging and the first action is to damp their motion with respect to a local reference frame. This action is performed using the so-called *Local Control*: the mirrors are controlled by means of independent ground-based sensors and quasi-inertial actuators. In this way we have the mirror position referred to the laboratory ground (generally affected by the seismic noise).

Then, the operation point of the GW interferometer is locked by using the *Global Control*, *i.e.* the relative position of mirrors is correlated by light itself. To complete this last task we have to control 4 independent lengths:

- $2l_{\mathrm{rc}} + l_x + l_y$, the recycling cavity,
- $l_x + l_y$, the Michelson in the dark fringe,
- L_x, the x FP arm in resonance,
- L_y, the y FP arm in resonance.

To implement the strategy we make use of several photodiodes picking up signals leaking from the interferometer in different points (see fig. 13). The extracted signals, which concur to define the control matrix, are

- PRCL, *i.e.* the length of the power recycling cavity,
- DARM, *i.e.* the differential mode of the interferometer, in practice the length difference of the FP two arms,
- CARM, *i.e.* the common mode of the interferometer or in practice the mean length of the FP two arms,
- MICH, *i.e.* the difference in the length of the Michelson arms.

They are used to actuate on the mirrors and the beam splitter up to a maximum frequency of few tens of hertz. The difficulty to acquire the stable configuration with the longitudinal control feedback engaged depends on the finesse of the cavities. Different procedures have been developed. Here we cite just the one used in Virgo: the variable finesse method [4]. The idea is to take advantage of the fact that the finesse of the recycling cavity changes during the lock acquisition sequence. Two are the main steps of the locking strategy. The first is to lock the interferometer on the half-fringe, so that a large fraction of light escapes through the output port and the power build-up inside the recycling cavity is extremely low. In this way we can operate with all the degrees of freedom weakly coupled, making the control scheme much easier. At the end of this step, the four longitudinal degrees of freedom of the ITF are stably locked in a configuration with a very low recycling gain. Then, the interferometer is adiabatically brought to the operating point with the Michelson on the dark fringe.

Finally, we note that the force needed to acquire the lock is much larger than that to keep it in position, because we need momentum to slow down the mirror oscillations. Strong actuation force implies large electronic noise. For this reasons, once the lock is acquired, in the stable configuration forces are applied mainly on the upper stages of the suspension system, avoiding to actuate directly on the mirror.

All these considerations are applied to control the longitudinal coordinated, the most important one since it concerns the detection of the GW signal. However, other feedback

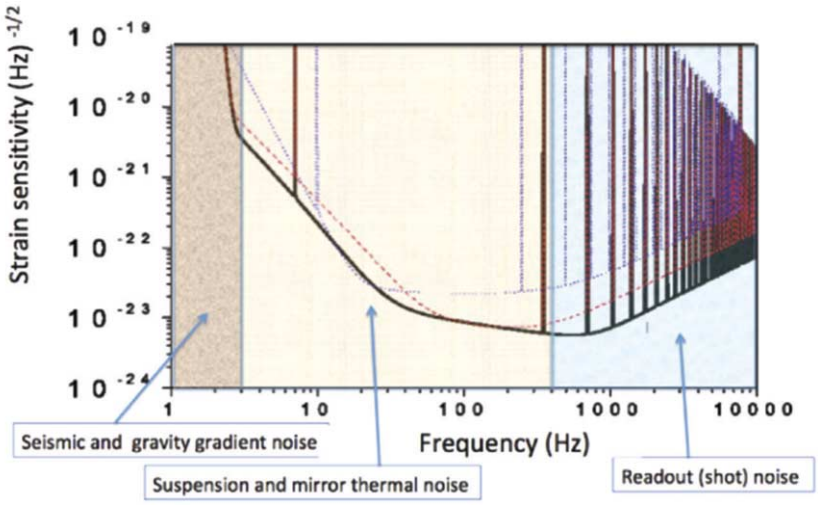


Fig. 14. – A generic sensitivity curve of a GW interferometer, where the frequency range of some of the main noise contributions are shown.

loops are needed to control the angular degrees of freedom of the mirrors. This is done essentially, again, with respect to the light into the interferometer by means of beam wavefront sensing, *i.e.* using quadrant photodiodes sensitive to the higher Hermite-Gauss mode of the light, which are present because of the cavity misalignment.

9. – The sensitivity curve

The sensitivity curve of the GW detector is obtained by considering the noise sources present in the detector and computing their contributions at the input of the system, *i.e.* in terms of the equivalent strain noise of the interferometer. The simplified assumption is that all these different contributions are independent, so that the total spectral density of the noise strain is derived by summing the spectral densities of the different contributions. It comes out that the sensitivity is presented as the square root of this spectral density as a function of the GW signal frequency (see fig. 14).

Here we do not discuss the noise sources, since these have been presented in other contributions to these proceedings. We notice just that the main noise contributions in all the predicted sensitivities of the present and future detectors on the Earth are the readout noise (radiation pressure and shot noise) and the thermal noise (suspension and mirror thermal noise), and in addition, in the lowest part of the detector bandwidth, the residual seismic noise filtered by the suspensions and the gravity gradient noise [5].

This schematic approach permits to discuss the strategies to be pursued for improving the sensitivity and these are the topics discussed shortly in the following sections.

10. – Thermal noise and cryogenics for future gravitational wave detectors

The best sensitivity values of an advanced interferometric detector of gravitational waves are achieved in the frequency range where the dominant source of noise is the thermal one. This is due to the oscillation modes of the suspended mirror; we have to focus our attention mainly on the pendulum mode of the mirror suspended by four wires, which has a characteristic frequency below 1 Hz, and on the thermal fluctuation of the mirror itself contributing to the longitudinal sensitivity of the detector. The mirror is an elastic thick plate whose first symmetric mode of oscillation, moving back and forward the mirror surface, is at around $\sim 5\,\mathrm{kHz}$. Both displacement noise spectra are strongly peaked but in different frequency regions and the GW sensitivity curve is mainly defined by the overlap of the two tails of the noise spectra: for the suspension on the tail on the right side of the 1 Hz peak and for the $\sim 5\,\mathrm{kHz}$ mirror mode on the its left. In both cases, the physics parameters acting on which we can reduce significantly these noise contributions are essentially two: the dissipation mechanisms and the temperature. This is stated by the fluctuation-dissipation theorem, which defines the spectral behaviour of the stochastic force acting on a mechanical linear system.

To define the dynamics of the mechanical system in the frequency domain, we use the force-displacement transfer function $H(\omega)$ and the force-velocity mechanical impedance $Z(\omega)$ as parameters

$$H(\omega) = \frac{F(\omega)}{X(\omega)}\,, \qquad Z(\omega) = \frac{F(\omega)}{v(\omega)}\,.$$

In synthesis the theorem states that the spectral densities of the stochastic force and displacement are

$$S_{ff}(\omega) = 4k_B T\,\mathrm{Re}[Z(\omega)], \qquad S_{xx}(\omega) = 4k_B T\,\mathrm{Im}[H(\omega)],$$

where k_B is the Boltzmann constant and T the equilibrium temperature of the system.

If we apply this theorem to the case of a simple harmonic oscillator, which is a good model for both the pendulum mode of the suspended mirror and the disk vibration mode, we end up to the following results:

$$S_{xx}(\omega) = \frac{4\pi k_B T}{m_p {\omega_o}^4}\frac{\beta_p(\omega)}{m_p} \quad \text{for } \omega \gg \omega_o, \tag{28}$$

$$S_{xx}(\omega) = \frac{4\pi k_B T}{m_d \omega^4}\frac{\beta_d(\omega)}{m_d} \quad \text{for } \omega \ll \omega_o, \tag{29}$$

where m_p and m_d are the equivalent masses of the two oscillation modes of resonance frequency ω_o and $\beta_p(\omega)$, $\beta_d(\omega)$ the dissipation factors.

The advanced detectors LIGO and Virgo operate at room temperature, so that the main effort has been concentrated on the reduction of the dissipation factors $\beta(\omega)$. Here we do not review the results obtained in this context. We notice that the dissipation

associated to the disk vibration is dominated by the mirror coating and an extended research campaign is ongoing to find the good receipt for having both lower optical and acoustic losses [6].

Up to now only the Japanese collaboration KAGRA focused the attention on the second parameter, the temperature. They plan to cool the mirrors at 20 K and this is the new frontier of the GW interferometric detectors. Cooling the mirrors at cryogenic temperatures drastically reduces thermal lensing effect. Below 50 K the material thermal expansion coefficient is so low that the deformation effect of the mirrors can be neglected. In addition, provided a suitable choice of materials compatible with the cryogenic environment, we can obtain a significant reduction of this noise source beyond the simple linear dependence of the noise spectra (28), (29). This was demonstrated in the past by resonant cryogenic detectors, for which at low temperature the acoustic dissipation factor of the selected materials was reduced and the consequent thermal noise contribution was strongly depressed.

Indeed, the cryogenic operation of a GW interferometer represents the main challenge of the new GW interferometer. The use of cryogenic techniques in the interferometer implies to solve several problems:

- use crystal mirrors (instead of amorphous SiO_2) whose optical, elastic and cryogenic properties are compatible with a GW detector requirements;
- define a suitable cooling configuration to reduce the cooling time and, in the stationary condition, to compensate the thermal load due do the mirror absorption of the light;
- design of heat sinks in such a way to bring sufficient refrigeration power without increasing the mechanical dissipations and without compromising the seismic insulation of the suspension;
- protect the system from the vibrations generated by the cooling system.

In the KAGRA detector [7] these items have been analysed and the international collaboration around the project is almost ready to operate the interferometer at 20 K. The mirrors are made of sapphire crystal: the material is harder than SiO_2 but it has been demonstrated that the quality of the mirror surface via polishing can achieve a root mean square below 1 nm as in the case of fused silica substrates. The most critical parameter of the sapphire is the optical absorption: it is ranging from 40 ppm/cm up to 90 ppm/cm in the crystals delivered to KAGRA, values to be compared with few ppm/cm for the SiO_2 case. If the absorption is not decreased the laser power inside the FP arms must be lowered, unless they are able to increase the efficiency to extract the heat absorbed by the mirror. The effort is to get test masses with an absorption of 30 ppm/cm, which implies extracting 0.4 W from each of the arm input mirrors in the low temperature stationary state.

The sapphire mirrors are located in cryostats which include two inner shields: an outer one at 80 K and an inner one at 8 K [8]. Shields and payload are cooled by means

of four pulse tube cryo-coolers. The heat is extracted from the mirror primary via the suspension wires. Thus, the design of the cryogenic suspension is driven by the need to cool down the mirror while keeping its thermal noise as low as possible. The mirror is suspended using four thin rods made of sapphire (35 cm long and 1.6 mm in diameter).

The cooling time is another important parameter since the mirrors will have to be heated up and cooled down once a year to remove the water deposited on the optical surface due to adsorption [9]. Their strategy is to cool the mirror via radiation down to 150 K. The radiation cooling efficiency can be improved by coating the payload (with the exception of the mirror) with diamond-like carbon (DLC). Then, to achieve the mirror temperature of 20 K they rely on appropriate heat links. The heat links are a potential path for the transmission of the cryostat vibrations. This is one of the main challenges whenever cryostats are used in conjunction with an apparatus requiring a very low level of vibrations and it is far from obvious that, as far as the GW sensitivity is improved, the constraints on the mechanical noise transmission of the links became more severe.

At present the strategy for cooling down the payloads, which includes the 23 kg mirror, should permit to reduce the total cooling time from nearly two months to 39 days. The future generation of GW detectors (3G) will host heavier and larger mirrors. A brute force calculation tells us that, if we plan to use the same materials for the payload, a ten times increase of its mass implies a similar multiplication factor in the cooling time. It is far from obvious that 1 year of cooling time is an unacceptable number. Indeed, we have to notice that this is the most pessimistic scenario, because in a 3G configuration the surface of the new payload will be higher than that of KAGRA, contributing to reduce the cooling time. The conclusion is that the cooling efficiency and time is strictly related to the 3G payload design and a robust R&D activity on this domain is needed. We conclude by listing some of the crucial items to be covered by the R&D program focused on a cryogenic suspended mirror:

- develop a suspension last stage prototype working in a cryogenic environment;
- design and test the whole cryogenic system to cool the suspension last stage;
- design and test an attenuation chain of a few mechanical filters able to work in cryogenic environment and to bring sufficient refrigeration power from the top to the bottom;
- develop sensors and actuators needed for the position control of the mirror that are compatible with the cryogenic environment;
- compensate at low temperature the additional noise generated by the refrigeration system;
- low temperature measurements of the heat absorption of the mirror substrates.

11. – Reduction of the readout noise

In the usual classification of the noise sources of a GW detector the readout noise is the sum of the two spectral densities due to the shot noise and radiation pressure noise of the photons impinging on the suspended mirrors of mass M. While the shot noise contribution, $S_{h_{\rm sh}}$ decreases with optical power P, the radiation pressure contribution, $S_{h_{\rm rad}}$ increases. In fact, in the case of a simple Michelson we have

$$S_{h_{\rm sh}}(\nu) = \frac{1}{L^2}\left(\frac{\hbar\, c\, \lambda}{2\pi P}\right), \qquad S_{h_{\rm rad}}(\nu) = \frac{1}{M\,\nu^2\,L^2}\left(\frac{\hbar P}{2\pi^3\, c\,\lambda}\right). \tag{30}$$

Both noises are related to the quantum nature of the e.m. field which interacts with a classical object, the mirror, to monitor continuously its displacement. Thus, it is not surprising that both noises concur to define the quantum limit of the measurement strategy. For a fixed value of P, the crossover of the two curves defines the minimum noise achievable by the readout system, the so-called standard quantum limit (SQL). This is the minimal sum of shot noise and radiation pressure noise. In other words, the light fields enforce Heisenberg uncertainty through complementarity between shot and radiation pressure noise.

The standard quantum limit (SQL) is a concept introduced for the first time in the GW community by V. Braginsky [10,11]. In a semi-classical vision the measurement can be described via a Hamiltonian interaction term where mechanical and e.m. variables are coupled. This implies that the e.m. transducer converts the mechanical information into an electromagnetic signal and, at the same time, *back-acts* to the mechanical system. Now, let us consider again the quadrature components of both the mirror position X_p and X_q and those of the e.m. field E_p and E_q. To circumvent the Heisenberg principle we should be able to develop a strategy scheme playing with the different phases: for example we can try to transfer the signal information embedded in the X_p phase of the mechanical variable to the E_q e.m. phase, while the back-action noise to the E_p phase influences just the X_q (see fig. 15).

In this way it is possible to set transducers by means of which we reduce the variance of X_q counterbalanced by an equivalent increase of the X_p variance, in accordance to the Heisenberg principle [12].

This strategy is called quantum non-demolition technique and it was tested in the case of the low temperature resonant detectors [13]. The first to approach the problem of beating SQL for the optical interferometers was Unruh [14] while, at about the same time, Yuen pointed out that the trick is to correlate the uncertainties in position and momentum in a particular way [15]. This is an important remark on which we will come back later.

Indeed, for a deep comprehension on how to develop a quantum strategy of measurements in a GW interferometer, a quantum optics approach must be followed [16], where the two quadrature components of the e.m. field are written in terms of the creation, $\hat{a}^\dagger$,

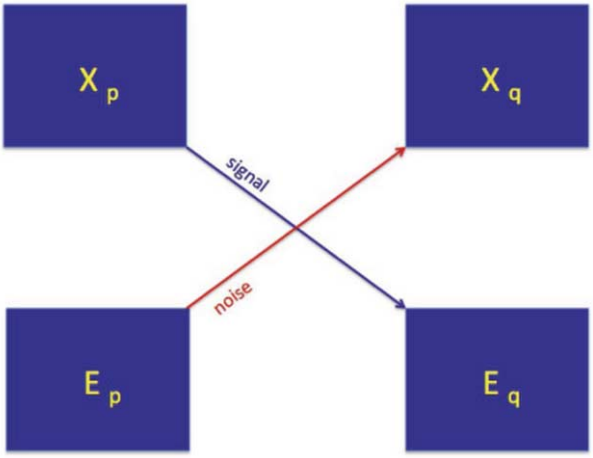


Fig. 15. – The scheme of a back-action evading strategy.

and annihilation, $\hat{a}$, operators

$$\hat{E}_p = E_o \left[\hat{a}^\dagger + \hat{a}\right], \qquad \hat{E}_q = i\, E_o \left[\hat{a}^\dagger - \hat{a}\right]$$

and the product of the standard deviations of two such operators obeys the uncertainty principle.

In quantum optics the coherent state refers to a minimum uncertainty state, with the single free parameter chosen to make the variance equal for the two quadrature components of the field. It means that the uncertainty of a coherent state in the quadrature phase space is a circle (fig. 16). To do a more accurate measurement of one of the two phases means to be able to deforme the circle into an ellipse of equal area, so that the variance of one of the two phases is *squeezed*. Thus, diminishing the quantum noise at a specific quadrature of the wave has as direct consequence an enhancement of the noise of the complementary quadrature.

Depending on the phase angle at which the state's width is reduced, one can distinguish amplitude-squeezed, phase-squeezed, and general quadrature-squeezed states.

Squeezed states of light can be divided into squeezed vacuum and bright squeezed light, depending on the absence or presence of a non-zero mean field. In other words, if the squeezing is applied directly to the electromagnetic vacuum, rather than to a coherent state, the result is called squeezed vacuum.

The vacuum state is present in any region in space where there is no other occupying light state and this is important for GW detector readout. In fact, we recall that the Michelson interferometer of the GW detector is controlled in such way to be its output on the dark fringe. This condition reflects the incident vacuum state back towards the readout photodetector. Thus, quantum noise in a GW detector arises from the vacuum state entering at the dark port of the interferometer.

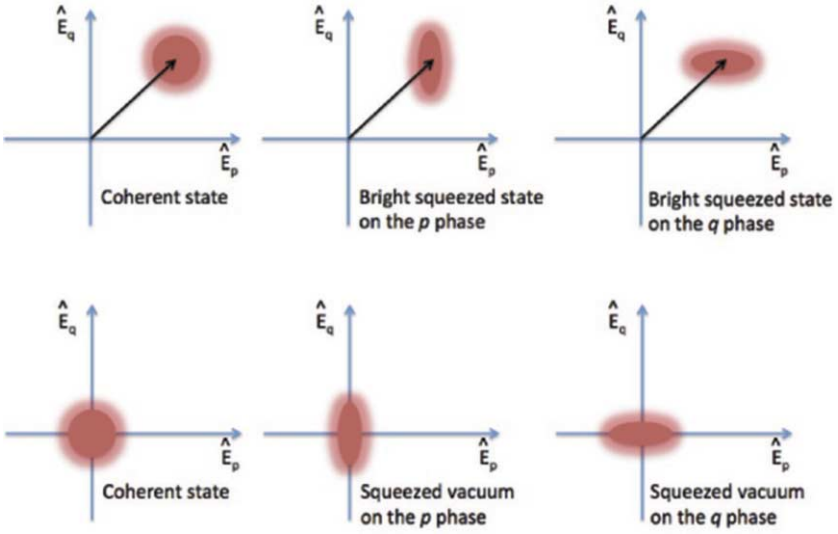


Fig. 16. – We show in the phase plane an example of coherent and bright squeezed states state of the e.m. signal (upper diagrams) and the vacuum and squeezed vacuum states (lower diagrams).

We conclude that a squeezed vacuum state injected into the interferometers dark port modifies the overall quantum noise: a phase squeezed state can reduce the shot noise at the expense of the radiation pressure noise, conversely an amplitude squeezed state can reduce the radiation pressure noise at the expense of the shot one.

Following the line of thinking of Yuen cited above [15], again the modulation concept helps to depict the light fluctuation and the squeezing process. A laser light beam is a photon ensemble characterised by an average frequency and amplitude with individual photons that may have slightly different frequency or amplitude. This implies that quantum noise can be seen as a continuum of sidebands around the carrier Ω, spanning all frequencies $\omega \pm \Omega$ due to the beating between the carrier and the vacuum fluctuations. These sidebands are uncorrelated and have white amplitude and phase noise with the same statistical distribution. The squeezing process introduces a certain level of correlation between randomly fluctuating upper and lower sideband pairs. This concept is present in the full quantum optics treatment of the squeezing problem [16] and the correlation term of the uncertainties allows for improving the precision from the SQL.

There are several methods to produce squeezing states (for a review of this topic see [17,18]), but at present the most advanced technology for producing squeezed states for GW detection is based on the use of optical parametric oscillators (OPOs).

In more details the squeezed vacuum is produced via a process called parametric down-conversion, where a pump photon with frequency ω_p, incident on a dielectric with a χ^2 nonlinearity [19], breaks up into two new photons: a signal photon of frequency ω_s and an idler photon of frequency ω_i, where $\omega_p = \omega_i + \omega_s$. For degenerate parametric down-

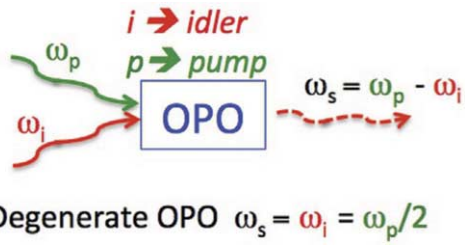


Fig. 17. – The degenerate optical parametric amplifier is pumped by two light beams both controlled in phase with respect to the laser of the GW interferometer. The first beam is at the same frequency as the GW laser, while the second one is at half of this frequency. The OPO mixes the two beams generating a vacuum state with a suitable level of correlation between the two quadratures of the field.

conversion, the idler and signal photons are indistinguishable in frequency ($\omega_i = \omega_s$) and polarisation. A degenerate OPO, where this process takes place inside an optical cavity, has proven to be the most efficient source of quadrature squeezed light: in fact, it acts by introducing correlation between randomly fluctuating sidebands.

Being the signal frequency exactly half the pump frequency, the phase relationship between signal and pump determines whether there is amplification or de-amplification of the signal. We stress that this phase-sensitive amplification occurs only in the case of degenerate amplifiers.

The simplified diagram in fig. 17 shows the principle on the basis of which a vacuum squeezed state is generated.

The squeezed quadrature generated from an OPO squeezers is constant across the frequency spectrum and, since 2012, it is injected successfully in the Anglo-German GW interferometer GEO, improving its shot noise limit [20]. Indeed, an optimised squeezed source for a GW interferometer should have the squeezed ellipse with a rotation angle optimised as a function of the detection frequency. This is called frequency-dependent squeezing, which implies filtering the squeezed vacuum generated by the OPO via an external device with a suitable phase dispersion. An optical cavity, or a sequence of cavities with the carrier frequency locked on the side of the cavity resonance, provide such a dispersion: these are referred to as filter cavities that typically need to be $\sim 100\,\mathrm{m}$ in length to achieve the dispersion over the GW detection bandwidth. The R&D devoted to these filter cavities is an active area of investigation: they need to have extremely low optical loss, to limit the degradation of the produced vacuum squeezing. In addition, their line widths has to be of the order of tens of Hz to be able to rotate the squeezing around the frequency crossover of the spectra of the radiation pressure noise and shot noise.

12. – Conclusion

The era of gravitational wave astronomy is born with the detection of the GW signals emitted by the coalescence of binary systems. After this scientific breakthrough, it is more and more important to improve the performance of the detectors, their sensitivity and reliability for increasing the overall observational time and explore a larger volume of the universe. Indeed, with the advanced detectors we will alternate time periods of data taking to those devoted to the implementation of new experimental solutions. The use of new technologies will bring the advanced detectors to the limit given by the present infrastructures and then a new generation of detectors will get into action. The Einstein Telescope [21] is a European design study project for building a 3G ground-based interferometer. Our wish is that it will be located in the area of the Sos Enattos mine in Sardinia, a site characterised by a very low seismic and environmental noise.

Other ideas for 3G detectors have been presented [22], but all new projects need a robust R&D program, coordinated at international level and carried out by a new generation of researchers. This will permit to progress in the design of a new class of GW instruments, which will make it possible to observe gravitational phenomena up to the border of the universe.

* * *

This work has been supported by *Centro Amaldi di fisica e astrofisica della gravitazione*, INFN and the grant *Ateneo 2017 - Sapienza*.

REFERENCES

[1] AASI J. *et al.*, Advanced LIGO, *Class. Quantum Grav.*, **32** (2015) 074001.
[2] ACERNESE F. *et al.*, Advanced Virgo: a second generation interferometric gravitational wave detector, *Class. Quantum Grav.*, **32** (2015) 024001.
[3] SAULSON P., How an interferometer extracts and amplifies power from a gravitational wave, *Classical Quantum Grav.*, **14** (1997) 2435.
[4] ACERNESE F. *et al.*, The variable finesse locking technique, *Class. Quantum Grav.*, **23** (2006) S85.
[5] HARMS J., Terrestrial Gravity Fluctuations, *Living Rev. Relat.*, **18** (2015) 3.
[6] Optical Coatings and Thermal Noise in Precision Measurement, HARRY G., BODIYA T. P. and DESALVO R. (Editors) (Cambridge University Press) 2012.
[7] SOMIYA K., Detector configuration of KAGRA - the Japanese cryogenic gravitational-wave detector, *Class. Quantum Grav.*, **29** (2012) 124007.
[8] SAKAKIBARA Y. *et al.*, Progress on the cryogenic system for the KAGRA cryogenic interferometric gravitational wave telescope, *Class. Quantum Grav.*, **31** (2014) 224003.
[9] FLAMINIO R., The cryogenic challenge: status of the KAGRA project, *J. Phys.: Conf. Ser.*, **716** (2016) 012034.
[10] *International School of Cosmology and Gravitation, Topics in theoretical and experimental gravitation physics. Proceedings of the School held in Erice, Trapani, Sicily, March 13–25, 1975*, edited by DE SABBATA V. and WEBER J. (Plenum Press, New York) 1977.
[11] *Atti dei convegni Lincei 34: International Meeting on Experimental Gravitation (Pavia, 17–20 September 1976)*, edited by BERTOTTI B. (Accademia dei Lincei) 1977.

[12] Cinquegrana C. *et al.*, Back-action-evading transducing scheme for cryogenic gravitational wave antennas, *Phys. Rev. D*, **8** (1993) 448.
[13] Bonifazi P. *et al.*, Test of a back-action evading scheme on a cryogenic gravitational wave antenna, *Phys. Lett. A*, **215** (1996) 141.
[14] Unruh W. G., in *Quantum Optics, Experimental Gravitation, and Measurement Theory*, edited by Meystre P. and Scully M. (Plenum, New York) 1983, pp. 647–60.
[15] Yuen H., Contractive States and the Standard Quantum Limit for Monitoring Freemass Positions, *Phys. Rev. Lett.*, **51** (1983) 719 and Erratum at *Phys. Rev. Lett.*, **51** (1983) 1603.
[16] Luis A. and Sánchez-Soto L., Breaking the standard quantum limit for interferometric measurements, *Opt. Commun.*, **89** (1992) 140.
[17] Anderson U. L. *et al.*, 30 years of squeezed light generation, *Phys. Scr.*, **91** (2016) 053001.
[18] Chua S. S. Y. *et al.*, Quantum squeezed light in gravitational-wave detectors, *Class. Quantum Grav.*, **31** (2014) 183001.
[19] Sutherland R. L., *Handbook of Nonlinear Optics*, 2nd edition (Marcel Dekker, New York) 2003.
[20] Khalaidovski A. *et al.*, Status of the GEO 600 squeezed-light laser, *J. Phys.: Conf. Ser.*, **363** (2012) 012013.
[21] Punturo M. *et al.*, The Einstein Telescope: a third-generation gravitational wave observatory, *Class. Quantum Grav.*, **27** (2010) 194002.
[22] Dwyer S. E., Sigg D., Ballmer S., Barsotti L., Mavalvala N. and Evans M., A Gravitational Wave Detector with Cosmological Reach, *Phys. Rev. D*, **91** (2015) 082001.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200004

Optical aberrations in gravitational wave detectors and a look at the future

Viviana Fafone

Università di Roma Tor Vergata e INFN, Sezione di Roma Tor Vergata - Roma, Italy

Summary. — In this lecture we will focus our attention on a specific issue connected with the optical properties of the test masses in ground-based interferometers, namely the presence of optical aberrations that can have a relevant impact on the possibility to operate the detector. The second part will present an overview on the medium- and long-term evolution of terrestrial interferometers.

1. – Introduction

The performances of the gravitational wave (GW) detectors are strongly determined by the quality of the materials adopted in the main optical components. The nominal sensitivity of the current and future ground-based detectors is limited by the thermo-mechanical characteristics of the optical coatings which determine the level of the thermal noise in the detector. Several approaches to reducing the mechanical loss are being explored, such as improved processing of conventional amorphous oxide or doped amorphous oxide materials, possibly deposited in nanolayer structures, alternative amorphous materials —typically semiconductors like amorphous silicon or silicon nitride, or crystalline semiconductor mirrors.

The choice of the coatings material has to consider also optical absorptions and uniformity which determine additional effects that can spoil the ultimate performances of the detector. These "thermal effects" have already been observed in the first generation of ground-based interferometers and have become more relevant in the second generation, due to the higher circulating power. Their origin and solutions adopted to mitigate them are described in the two following sections.

2. – Optical aberrations and their effects

The absorption by the high-reflectivity coating of a fraction of the power of the sensing team circulating in the interferometer produces a gradient of temperature inside the substrate. Two different effects originate from this heating:

– Non-uniform optical path length distortions (thermo-optic effect, also termed thermal lensing) due to the temperature dependency of the index of refraction and to the non-zero thermal expansion coefficient of the substrate.

This effect can be understood if we consider a body at uniform temperature T_0, thickness l, refraction index n and non-zero thermo-optic coefficient ($\mathrm{d}n/\mathrm{d}T \neq 0$). The optical path length through the body is $\mathrm{OPL} = n \cdot l$. If we increase uniformly the temperature of the body by a quantity ΔT, since the refraction index depends on the temperature (neglecting at first approximation the thermal expansion), the new OPL through the body becomes:

$$\mathrm{OPL}_1 = n_1 \cdot l, \tag{1}$$

where

$$n_1 = n + \frac{\mathrm{d}n}{\mathrm{d}T}\Delta T, \tag{2}$$

thus

$$\mathrm{OPL}_1 = \left(n + \frac{\mathrm{d}n}{\mathrm{d}T}\Delta T\right) \cdot l = n \cdot l + \frac{\mathrm{d}n}{\mathrm{d}T}\Delta T \cdot l = \mathrm{OPL} + \Delta\mathrm{OPL}, \tag{3}$$

and we can define the quantity *optical path length increase* as

$$\Delta\mathrm{OPL} = \frac{\mathrm{d}n}{\mathrm{d}T}\Delta T \cdot l. \tag{4}$$

If we include the thermal expansion, eq. (4) becomes [1]

$$\Delta\mathrm{OPL} = \frac{\mathrm{d}n}{\mathrm{d}T}\Delta T \cdot l + \alpha(1+\sigma)(n-1)\Delta T \cdot l \tag{5}$$

where α is the thermal-expansion coefficient and σ the Poisson's ratio. In the most general case of a non-uniform heating, the expression for the optical path length increase reads

$$\Delta\mathrm{OPL}(r,\theta) = \frac{\mathrm{d}n}{\mathrm{d}T}\int_0^l \Delta T(r,\theta,z)\mathrm{d}z + \alpha(1+\sigma)(n-1)\int_0^l \Delta T(r,\theta,z)\mathrm{d}z. \tag{6}$$

– Change of the profile of the optics surface, due to thermal expansion (thermo-elastic deformation).

A fraction of the power is also absorbed in the bulk of the optics thus generating the same effects. However this absorption is typically smaller than that occurring in the coatings. As a reference, the power absorbed in the coatings of the Advanced Virgo test masses at the design input power (125 W) will be about 0.35 W to be compared with absorptions of the order of 0.03 W in the bulk.

Thermal lensing can be computed analytically in simple cases, for example when we consider uniform absorptions and an impinging Gaussian beam. In the following the computation is reported. References for more comprehensive studies can be found in [1] and [2].

For a Gaussian profiled beam (fundamental TEM_{00} mode), the intensity as a function of the radial coordinate can be expressed as

$$I(r) = \frac{2P}{\pi w^2} e^{-\frac{2r^2}{w^2}}. \tag{7}$$

It does not depends on any angular coordinate nor on z. Thus we can exploit the symmetry of the problem and use cylindrical coordinates (r, z, θ), with $0 < r < a$ and $-h/2 < z < h/2$, with a the radius and h the thickness of the mirror.

In the most general case, the time evolution of the temperature field is found by solving the Fourier equation:

$$\left[\rho C\partial_t - K_c\nabla^2\right] T(r,z,t) = p(r,z,t), \tag{8}$$

where T is the temperature increase (it was ΔT), ρ the density, C the specific heat, K_c the thermal conductivity, p is the possible heat source inside the mirror and $\partial_t = \partial/\partial t$.

We want to determine the steady-state ($\partial_t \to 0$) temperature field, when the heat lost by radiation balances the power absorbed by the mirror. If we limit ourselves only to the heat absorbed by the mirror surface, then there are no internal heat sources and the Fourier equation becomes the Laplace one

$$\nabla^2 T(r,z) = 0, \tag{9}$$

which in cylindrical coordinates can be rewritten as

$$\left(\partial_r^2 + \frac{1}{r}\partial_r + \partial_z^2\right) T(r,z) = 0. \tag{10}$$

The general solution to eq. (10) is a harmonic function of the kind

$$T(r,z) = J_0(kr)\left(Ae^{kz} + Be^{-kz}\right), \tag{11}$$

where $J_0(x)$ are the Bessel functions of the first kind, A, B e k the constants determined by the boundary conditions.

On the mirror surface hit by the laser beam ($z = -h/2$), the boundary condition will be

$$-K_c\left[\frac{\partial T(r,z)}{\partial z}\right]_{z=-h/2} = -\sigma_e\left([T_0 + T(r,-h/2)]^4 - T_0^4\right) + \epsilon I(r), \tag{12}$$

ϵ is the absorption coefficient of the mirror high-reflectivity coating and σ_e is the emissivity of fused silica. Since $T \ll T_0$, because T represents the temperature increase and T_0 the room temperature, it is possible to linearize the term relative to irradiance and rewrite eq. (12) as

$$-K_c\left[\frac{\partial T(r,z)}{\partial z}\right]_{z=-h/2} = -4\sigma_e T_0^3 T(r,-h/2) + \epsilon I(r). \tag{13}$$

On the opposite surface ($z = h/2$), it will be

$$-K_c\left[\frac{\partial T(r,z)}{\partial z}\right]_{z=h/2} = 4\sigma_e T_0^3 T(r,h/2) + \epsilon I(r). \tag{14}$$

And finally, the irradiance boundary condition on the barrel of the mirror ($r = a$) reads

$$-K_c\left[\frac{\partial T(r,z)}{\partial r}\right]_{r=a} = 4\sigma_e T_0^3 T(a,z). \tag{15}$$

Let us consider eq. (15): we know that the general solution is described by eq. (11) and that a Bessel function has the following property:

$$\partial_x J_0(x) = -J_1(x). \tag{16}$$

By inserting eq. (11) into eq. (15) and using the property (16), it is straightforward to get

$$K_c k J_1(ka) = 4\sigma_e T_0^3 J_0(ka). \tag{17}$$

If we define $\chi \equiv \frac{4\sigma_e T_0^3 a}{K_c}$ and $\zeta = ka$, the previous equation becomes

$$\zeta J_1(\zeta) - \chi J_0(\zeta) = 0, \tag{18}$$

which has a discrete and infinite set of solutions $\{\zeta_n,\ n = 1, 2\ldots\}$. Thus the constant k gets values $k_n = \zeta_n/a$ and the general solution to eq. (10) can be written as

$$T(r,z) = \sum_n \left(A_n e^{k_n z} + B e^{-k_n z}\right) J_0(\zeta_n r/a). \tag{19}$$

It can be shown from the Sturm-Liouville theorem that the functions $\{J_0(\zeta_n r/a),\ n = 1, 2\ldots\}$ form a complete orthogonal basis for functions defined in the interval $[0, a]$. Thus, any function $I(r)$ can be expanded in series of J_0 in $[0, a]$ (Fourier-Bessel series).

$$I(r) = \sum_{n=0}^{\infty} I_n J_0(\zeta_n r/a) \tag{20}$$

and

$$I_n = \frac{2\zeta_n^2}{a^2\left(\chi^2 + \zeta_n^2\right) J_0(\zeta_n)^2} \int_0^a I(r) J_0(\zeta_n r/a) r \,\mathrm{d}r. \tag{21}$$

In the case of interferometric detectors, $I(r)$ is given by eq. (7), so that

$$I_n = \frac{2\zeta_n^2}{a^2\left(\chi^2 + \zeta_n^2\right) J_0(\zeta_n)^2} \int_0^a \frac{2P}{\pi w^2} e^{-\frac{2r^2}{w^2}} J_0(\zeta_n r/a) r \,\mathrm{d}r. \tag{22}$$

Moreover, to minimize clipping losses the radius of the mirrors is chosen to be much larger than the size of the Gaussian laser beam. So that in eq. (22) it is possible to replace the upper integration limit with ∞, getting

$$I_n = \frac{P}{\pi w^2} \frac{2\zeta_n^2}{a^2\left(\chi^2 + \zeta_n^2\right) J_0(\zeta_n)^2} e^{-\frac{w^2\zeta_n^2}{8a^2}} \tag{23}$$

and finally, the intensity profile of the laser Gaussian beam becomes

$$I(r) = \frac{P}{\pi w^2} \sum_{n=0}^{\infty} \frac{2\zeta_n^2}{a^2\left(\chi^2 + \zeta_n^2\right) J_0(\zeta_n)^2} e^{-\frac{w^2\zeta_n^2}{8a^2}} J_0(\zeta_n r/a). \tag{24}$$

Using eqs. (19) and (24), the boundary conditions (13) and (14) reduce to a linear system of equations, where A_n and B_n are the unknowns:

$$\left(\zeta_n - \chi\right)\Gamma_n^2 A_n - \left(\zeta_n + \chi\right) B_n = \frac{-\epsilon I_n a \Gamma_n}{K_c}, \tag{25}$$

$$\left(\zeta_n + \chi\right) A_n - \left(\zeta_n - \chi\right)\Gamma_n^2 B_n = 0, \tag{26}$$

having defined, for sake of simplicity, $\Gamma_n = e^{-\frac{\zeta_n h}{2a}}$. Solving for A_n and B_n is straightforward and we are finally able to write the equation describing the temperature field into

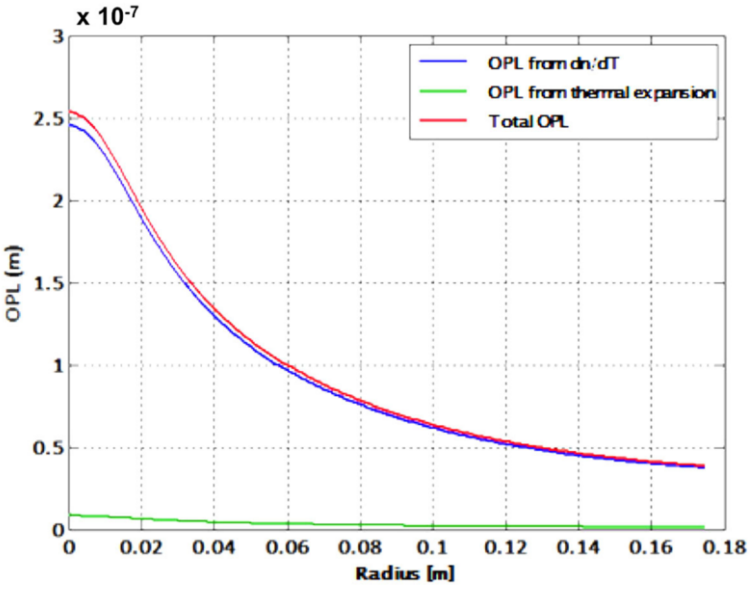


Fig. 1. – Optical path length increase inside a Virgo mirror for 0.2 W total absorbed power as a function of the radial coordinate (0 is the center of the mirror).

a cylindrical mirror that absorbs a fraction ϵ of the power stored in a Gaussian-profiled laser beam:

$$T(r,z) = \sum_n \frac{\epsilon I_n a}{K_c} e^{-\frac{\zeta_n h}{2a}} \frac{(\zeta_n - \chi)e^{-\frac{\zeta_n (h-z)}{a}} + (\zeta_n + \chi)e^{-\frac{\zeta_n z}{a}}}{(\zeta_n + \chi)^2 - (\zeta_n - \chi)^2 e^{-\frac{2\zeta_n h}{a}}} J_0(\zeta_n r/a). \tag{27}$$

In a very similar manner, it is possible to determine the temperature field when the heating occurs inside the bulk of the mirror [1].

We are now able to evaluate the optical path length increase. Inserting eq. (27) into (6) and performing the integration along the z direction, that represents the thickness of the mirror, one gets

$$\Delta\mathrm{OPL}(r) = \left(\frac{\mathrm{d}n}{\mathrm{d}T} + \alpha(1+\sigma)(n-1)\right) \cdot \sum_n \frac{\epsilon I_n a^2}{K_c \zeta_n} \frac{1 - e^{-\frac{\zeta_n h}{a}}}{\zeta_n + \chi - (\zeta_n - \chi)\, e^{-\frac{\zeta_n h}{a}}} J_0(\zeta_n r/a) \tag{28}$$

An analogue expression can be found for the optical path length increase due to substrate absorption [1]. The total ΔOPL is the sum of these two contributions.

Figure 1 shows the optical path length increase due to coating absorption in a Virgo mirror for 0.2 W of absorbed power.

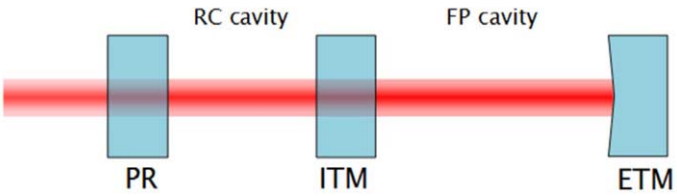


Fig. 2. – Linear scheme of a power recycled Michelson interferometer with Fabry-Perot arm cavities.

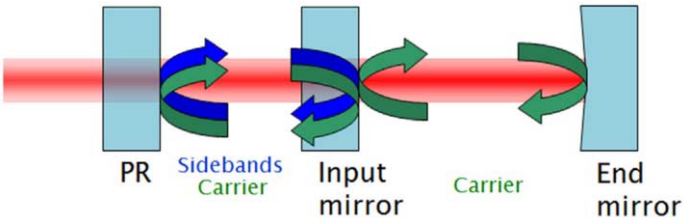


Fig. 3. – Resonant conditions of the fields circulating in the interferometer.

One important thing to notice is that thermal lensing does not create a spherical lens.

Another important point to clarify is where thermal effects arise in a GW interferometric detector.

Concerning the thermal lensing, let us consider the scheme shown in fig. 2, where the interferometer has been "linearized", *i.e.* only two coupled cavities, the power recycling and one Fabry-Perot, are drawn. The Power Recycling (PR) mirror will absorb for sure some power on its high reflectivity surface and in its bulk. So, the refraction index inside the substrate will change due to the temperature increase and some thermal lensing will appear. However, the PR bulk lies outside the interferometer. In fact, the Power Recycling cavity is defined by the High-Reflectivity (HR) surfaces of the PR mirror and input test mass. We can thus conclude that the thermal lensing in the PR mirror does not affect the interferometer.

Similar considerations apply to the End Test Masses (ETM): their substrate is out of the Fabry-Perot cavity, thus out of the interferometer and is not crossed by the beam.

Input Test Masses (ITM) absorb some of the power stored in the recycling cavity in their substrates, plus they absorb some of the Fabry-Perot cavity power on the high-reflectivity coating. So, the ITM substrate is where the thermal lensing occurs, as it is inside the power recycling cavity.

As explained in other lectures, three fields circulate inside the interferometer: the carrier and the lower and upper radio-frequency sidebands. Thermal lensing will affect only the control sidebands. To help understand this sentence, we can give a look at fig. 3. The carrier resonates both in the recycling and in the Fabry-Perot cavities. It will sense the thermal lensing in the ITM substrate, but when it enters the FP cavity it will be cleaned by the Finesse of the cavity itself that acts as a filter.

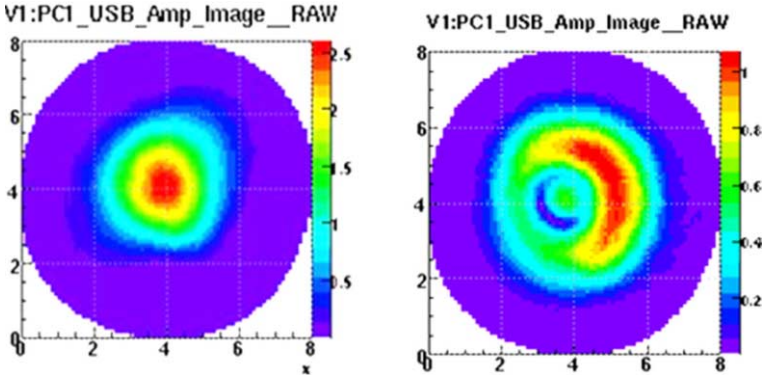


Fig. 4. – Amplitude images of a sideband field. Left: with no thermal lensing. Right: in the presence of strong thermal lensing. Lack of cylindrical symmetry may be due to some tilted optic.

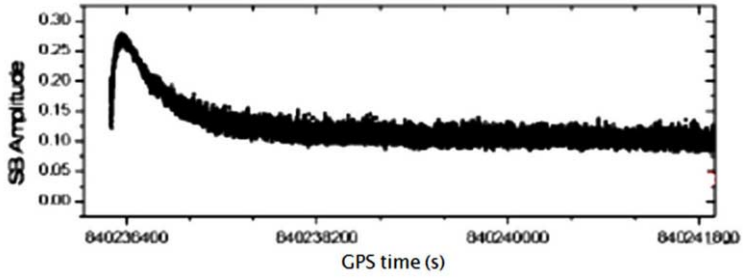


Fig. 5. – Typical transient decrease of the sidebands amplitude as observed during a Virgo lock acquisition.

On the contrary, the sidebands only resonate in the recycling cavity and, so, they do not take advantage of the filtering effect of the FP cavity. Radio-frequency sidebands fully sense the thermal lens in the ITM substrate. Since thermal effects give raise to a non-spherical lens, the sidebands fields in the recycling cavity are strongly aberrated (see fig. 4).

The cavity, seen by the sidebands, is "less" resonant and so the sidebands power decreases (as shown in fig. 5). At certain level, the sidebands amplitude becomes too low to keep the interferometer locked. In this situation, the only solution is to decrease the power at the input of the interferometer, to reduce the strength of the thermal lens. However, in this way, the sensitivity of the detector is degraded because of the increase of the shot noise.

Concerning the thermo-elastic deformation, the main effect seen by the interferometer is the change of the radius of curvature (RoC) of the ITM and ETM. This deformation was not relevant at the time of Virgo, but it increases the radius of curvature by about 40 m in Advanced Virgo at nominal design input power. As a consequence, the Fabry-Perot cavity will become less concentric, and the spot sizes at the mirrors will shrink, leading

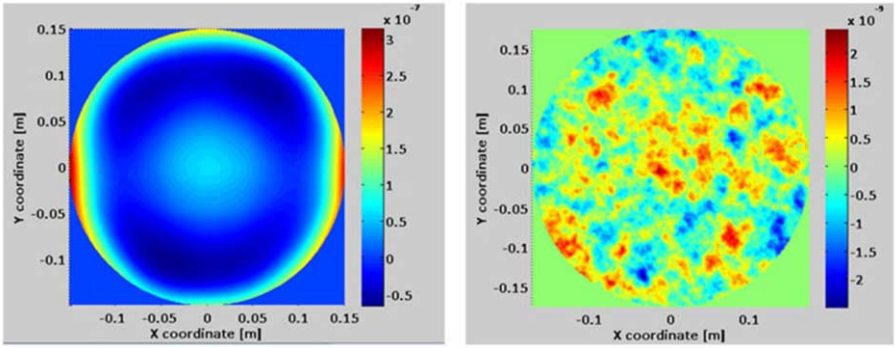


Fig. 6. – Left: substrate transmission map measured at LMA on an Advanced-LIGO test mass, the color scale is 10^{-7} m. Right: simulated surface map. Color scale in nanometers.

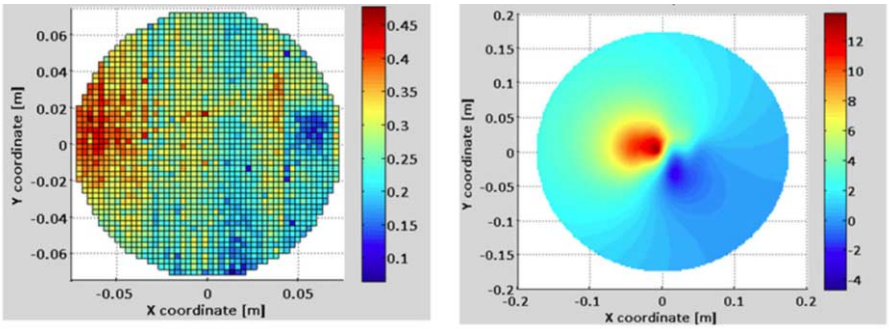


Fig. 7. – Left: coating absorption map measured at LMA on an Advanced-LIGO test mass. Right: resulting non-uniformity of the OPL for 125 W of ITF input power. Color scale in nanometers.

to an increase of thermal noise. Moreover, higher order modes can become resonant in the arm cavities, considerably increasing round trip losses.

So far, we have focused our attention only on thermal effects, both for the optical path length increase in the recycling cavity and for the increase of the test masses RoCs in the Fabry-Perot cavity. Moreover, up to now we have considered only effects arising from uniform absorption, *i.e.* with cylindrical symmetry.

However, in an ITF there are several sources of optical defects, often without any symmetry, arising from imperfections in the production and polishing of the glass used for the various substrates in the recycling cavity. Surface figure errors on reflective and transmissive surfaces do contribute to the aberrations as well as spatial variations in the index of refraction of the substrates (see fig. 6).

Finally, we also have to consider the non-homogeneity of the high-reflectivity coating absorptions, that generates a non-symmetric optical path length increase, as shown in fig. 7.

All these effects have to be considered in the aberration budget of the interferometer and their impact properly evaluated in order to understand the behaviour of the detector and to design a suitable system to minimise them.

3. – Correction of optical aberrations

The general guideline to design a system able to correct optical aberrations is to introduce in the interferometer a suitable "optical power" that can compensate the defects due to thermal effects or structural imperfections. This system, named Thermal Compensation System (TCS) must be adaptive, in order to be able to change at will the strength and shape of the corrective optics, to follow the different interferometer operating conditions.

In order to correct the thermal lensing, we must somehow induce in the input mirrors a "lens" equal but opposite to the one coming from aberrations, reminding that thermal lensing generates a lens that is not spherical. Another important point to establish is to what extent the optical defects must be compensated. The input mirrors are cylinders with a radius of 175 mm, is it necessary to correct the optical path length up to the edge? The answer to this question comes from optical simulations. In fact, we must recall that the size of the interferometer beam is much smaller than that of the mirror. So, the beam will sense optical path length distortions only where it exists. In practice, it turns out that the extent over which the aberration effects must be corrected is approximately 1.5 times the size of the beam on the ITMs. For instance, in Virgo the YAG laser radius on the ITM was about 22 mm, thus the area to be compensated had a radius of about 33 mm. In advanced Virgo the laser spot on the mirror has a radius of about 50 mm and the compensated area must increase accordingly.

The correction chosen in Virgo and advanced Virgo to compensate for the axisymmetric aberrations (as those induced by the adsorption of the Gaussian sensing beam) is an annular heating pattern generated by a CO_2 laser ($\lambda = 10.6\,\mu$m).

The CO_2 beam was impinging directly on the input test masses [3] in Virgo, while in advanced Virgo the CO_2 beam shines an additional transmissive optics, named Compensation Plate (CP) that has been added in the recycling cavity, in order to avoid that the noise introduced by the CO_2 laser beam is amplified by the factor $2F/\pi$ (where F if the finesse of the Fabry-Perot cavity) and to make it compliant with the higher sensitivity of the interferometer. The thickness of the CP has been optimized by minimizing the heat escaping from its barrel and taking into account the need to accumulate enough optical path length. The distance between CP and ITM is 20 cm, this allows to minimize the radiative coupling between the two optics. In fact, the heated CP radiates heat towards the test mass. The heating of the TM is uniform, but since the barrel of the input mirror radiates a part of the heat away, a radial temperature gradient is established. This gives rise to an increase of optical path length that adds to the thermal lensing.

The corrective annular pattern is generated with an axicon lens (an optic with a

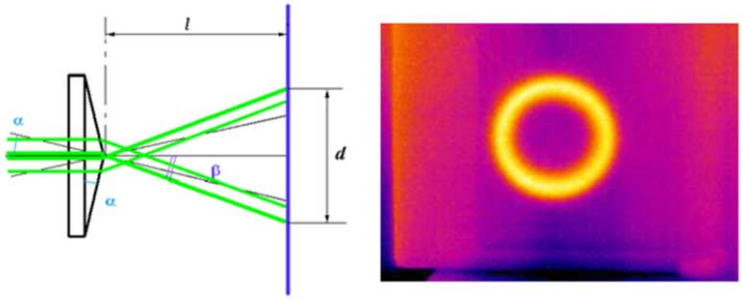


Fig. 8. – Left: working principle of an axicon lens. Right: intensity profile of a laser beam after passing through an axicon.

conical surface) that converts the laser Gaussian beam into an annular beam([1]).

The left image of fig. 8 shows the working principle of an axicon lens, while on the right a measured intensity profile of a CO_2 laser after going through an axicon is reported. An acousto-optic modulator([2]) is used to reduce the intensity noise of the laser, while DC control of the power content of the beam is accomplished using a half-wave plate and a fixed polarizer. Such a system allows to significantly reduce the deviation of the beam impinging on the axicon, which would otherwise result in a non-homogeneous intensity distribution inside the annular pattern.

The efficiency of the TCS has been tested already at the time of Virgo [4] by progressively increasing the CO_2 power, while looking at the behaviour of the interferometer and at the optical phase and amplitude of the carrier and modulation sidebands recorded by a phase camera [5]. The TCS power was increased until the thermal effects in the input mirrors were totally compensated [3, 6], as if the interferometer was operating at low power. In the upper left plot of fig. 9 it is possible to see that when no correction is applied the sideband field is strongly aberrated. The other plots in the same figure show that, as the TCS power is increased, the sidebands amplitude recovers the expected Gaussian shape.

The performances of the TCS allowed the interferometer to run VSR2 [7] with 17 W input power compared to 8 W during VSR1 [8].

([1]) Two axicon lenses are actually used in Advanced Virgo to better tune the shape of the heating pattern.

([2]) An acousto-optic modulator (AOM), also called a Bragg cell, uses the acousto-optic effect to diffract and shift the frequency of light using sound waves (usually at radio-frequency). A piezoelectric transducer is attached to a material such as glass. An oscillating electric signal drives the transducer to vibrate, which creates sound waves in the glass. These can be thought of as moving periodic planes of expansion and compression that change the index of refraction. Incoming light scatters off the resulting periodic index modulation and interference occurs similar to Bragg diffraction. The interaction can be thought of as four-wave mixing between phonons and photons.

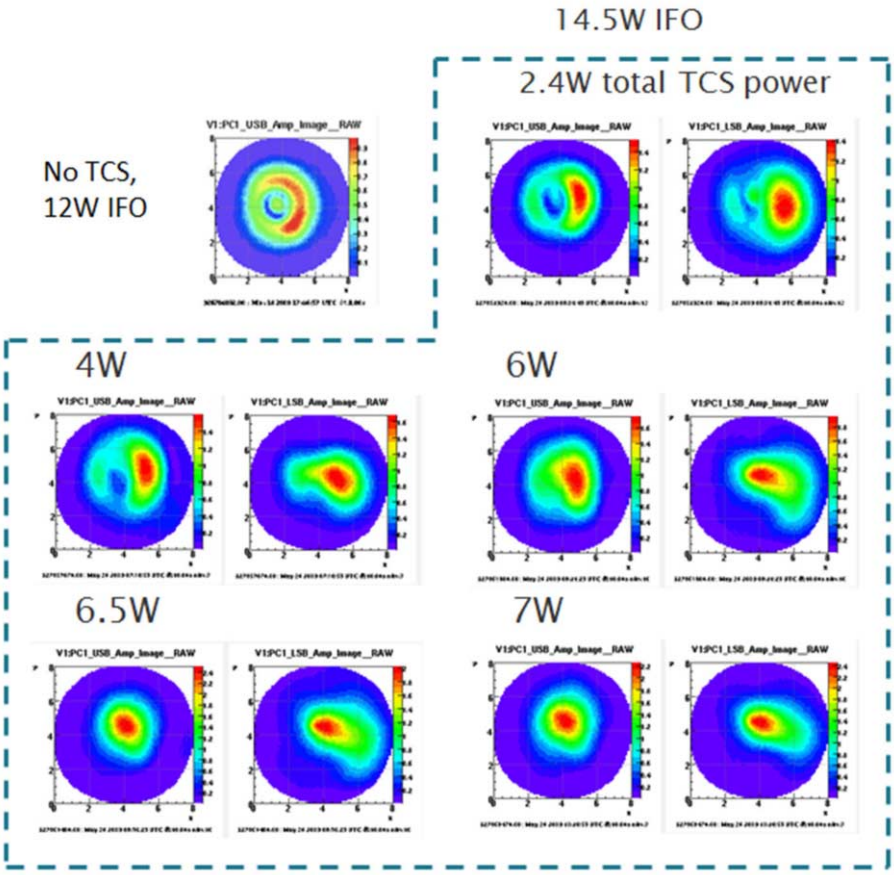


Fig. 9. – Amplitude images of the sidebands recorded with a phase camera. Upper left: no TCS (one only is shown for simplicity). The other plots show the change in shape as a function of the TCS power.

The actuator chosen to control the radius of curvature of the test mass follows from the same concept already used in the past in GW interferometric detectors [9]: a ring heater (RH) that radiates power on the test mass. Four ring heaters are installed in Advanced Virgo, one around each test mass. The position of the RH along the barrel of the TM is such as to maximize its efficiency. The input mirror RH also provides limited compensation of thermo-optic effect in the recycling cavities. The Advanced Virgo RH is equipped with a reflecting shield to maximize the amount of power reaching the test mass.

The conceptual actuation scheme of the compensation system of Advanced Virgo [10, 11] is shown in fig. 10, left. A pictorial view of the input payload, with the ITM, RH and CP, is shown in fig. 10, right.

This scheme also allows to reduce the coupling between the two degrees of freedom (lensing and RoC), so to have a control matrix as diagonal as possible.

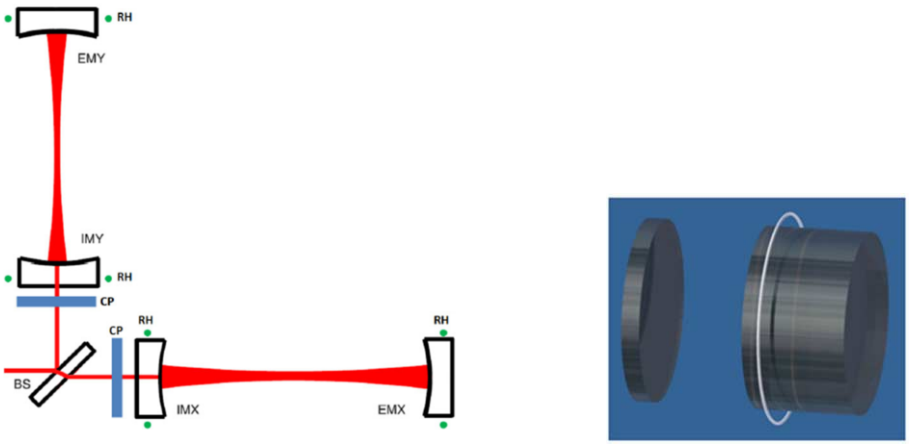


Fig. 10. – Left: Actuation scheme of the Advanced-Virgo TCS: blue rectangles represent the CPs (heated by the CO_2 lasers) while the green dots around the test masses are the ring heaters. Right: Picture of the input payload, comprising the ITM with the RH and the CP. Figure from [11].

The aberrations in the recycling cavity optics are sensed by several complementary techniques. The amplitude of the optical path length variation appears in some ITF channels, such as the power stored in the radio frequency sidebands. These are scalar quantities that can only give a measurement of the amount of power scattered into higher-order modes. Moreover, phase cameras [5] sense the intensity distribution and phase of the fields in the recycling cavity (carrier and sidebands). In addition, each optic with a significant thermal load is independently monitored. The HR face of each test mass is monitored in reflection for deformation. The input test mass/compensation plate phase profile are monitored on reflection on-axis from the recycling cavity side.

The TCS sensors, dedicated to the measurement of thermally induced distortions, consist of a Hartmann Wavefront Sensor (HWS) and a probe beam (at a different wavelength than the ITF beam) whose wavefront contains the thermal aberration information to be sensed. The working principle of the HWS is shown in the left panel of fig. 11. An aberrated wavefront W' is incident on a Hartmann plate (essentially a plate containing a series of apertures, see the right picture of fig. 11). The resulting rays propagate a distance L, normal to the wavefront, and are incident on a CCD. The new spot position, x_i', is measured and compared to a reference spot positions, x_i, determined using a non-aberrated wavefront W. The wavefront gradient in the i-th position is given by

$$\frac{\partial \Delta W}{\partial x} = \frac{\Delta x_i}{L}\,. \tag{29}$$

The Hartmann sensor selected for Advanced Virgo is that already developed and characterized on test bench experiments and in the Gingin High-Optical-Power Test Facility for the measurement of wavefront distortion [12].

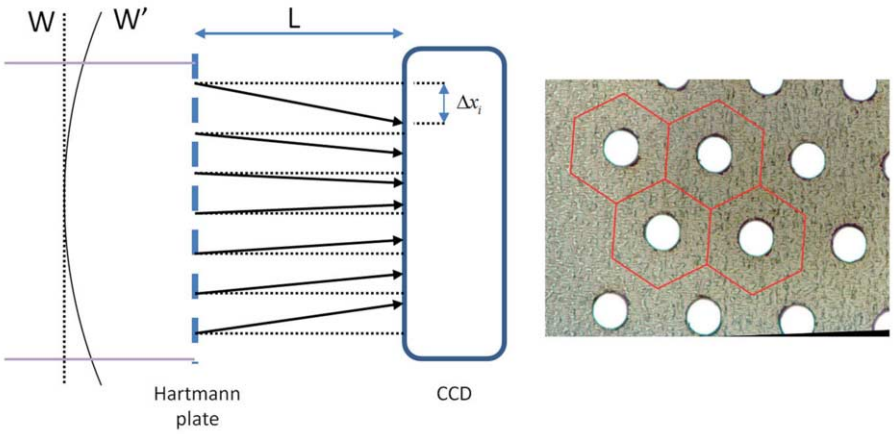


Fig. 11. – Left: working principle of the Hartmann Wavefront Sensor. Right: close view of a Hartmann plate.

4. – Mid and longer term perspective for ground-based detectors

At the time of writing this report, two observing runs have been completed. The first observing run (O1) of Advanced LIGO, which took place from September 12th, 2015 until January 19th, 2016 realized the first detections of gravitational waves from stellar-mass binary black holes (BBHs) [13-16]. After an upgrade and commissioning period, the second observing run (O2) of the Advanced LIGO detectors [17] started on November 30th, 2016, and ended on August 25th, 2017. On August 1st, 2017 Advanced Virgo [18] joined the observing run, enabling the first three-detector observations of GWs [19]. The joint LIGO-Virgo O3 run is scheduled to start in April 2019 with an approximate duration of one year.

At present, all data collected in the two observing runs have been deeply reanalyzed with improved search algorithms in order to reevaluate the significance of previously identified GW events and to potentially discover new ones [20]. The searches identified a total of ten BBH mergers and one binary neutron star (BNS) signal. The GW events are identified as follows: GW150914 (the first binary black hole merger detected), GW151012, GW151226, GW170104, GW170608, GW170729, GW170809, GW170814 (the first binary black hole merger detected by three interferometers), GW170817 (the first — and only, for the time being — binary neutron star detected by LIGO and Virgo), GW170818 and GW170823.

Apart from the three mot famous events, GW170729's mass has been estimated to be $85.1^{+15.6}_{-10.9} M_{\odot}$, making it the highest-mass BBH observed to date. GW170818 is the second BBH observed in triple-coincidence between the two LIGO observatories and Virgo after GW170814. As the sky location is primarily determined by the differences in the times of arrival of the GW signal at the different detector sites, LIGO-Virgo coincident events

have a vastly improved sky localization, which is crucial for electromagnetic follow-up campaigns.

Gravitational waves from compact binaries carry information about the properties of the source such as the masses and spins. The observation of these GW events allows us to place constraints on the rates of stellar mass BBH and BNS mergers in the Universe and probe their mass and spin distributions, putting them into astrophysical context. The non-observation of GWs from a neutron star-black hole binary (NSBH) yields a stronger 90% upper limit on the rate.

From the signals observed till now it is evident how relevant is the presence of a fully operating network: the presence of at least three detectors is in fact crucial to maximise the scientific outcome. For this reason the GW community is working toward an internationally coordinated effort for the upgrade of the present infrastructures and for the design and realization of future GW observatories, the so-called third generation (3G).

In the mid term timeline, the "plus" versions of Advanced Virgo and Advanced LIGO will exploit at best the present infrastructures, by implementing solutions to decrease the most offending noise sources: thermal noise in the mirrors, quantum noise, newtonian noise. The Advanced Virgo plus (AdV+) program will take the next 10 years, and will develop, in parallel with LIGO A+ and KAGRA, in two phases. The first-phase upgrades will include frequency-dependent squeezing, Newtonian noise cancellation, and signal recycling and will conclude with an observing run with a projected BNS range above 160 Mpc (roughly three times O2 sensitivity) in the mid-2020s. Then, phase 2 will demand better mirror coatings as well as larger beams and test masses which are projected to yield a BNS range around 260 Mpc (four times O2 sensitivities), in the timeline 2026-2027.

A further increase of the sensitivity of the ground-based interferometers requires changes in the infrastructures pushing toward a third generation detectors. The European project, named Einstein Telescope (ET) is in the middle of its second decade since its conception. The idea was formulated in 2004 and was studied in detail in an EU funded Design Study (2008–2011) [21].

From the US side, the 3G solution is focusing toward Cosmic Explorer [22], a 40 km-arm-length interferometer to be housed in a new observatory facility. An intermediate step is also being considered: LIGO Voyager [23] which is intended for the existing LIGO Observatories, with the aim to reach the existing facility sensitivity limits. It will take advantage of new technologies currently being researched, including silicon mirrors operated at 123 K temperatures, and high power 1.5–2 mm wavelength lasers with squeezed light injection.

The ET concept relies on a triangular-shaped facility (which allows to resolve the polarization of gravitational waves without the need of additional detectors), with 10 km long arms and three co-located instruments. Each instrument comprises of two detectors each, in a xylophone configuration in which one detector maximizes the sensitivity at low frequency, while the other detector maximizes the high-frequency performance. The output of the two detectors are then combined to provide a broadband sensitivity.

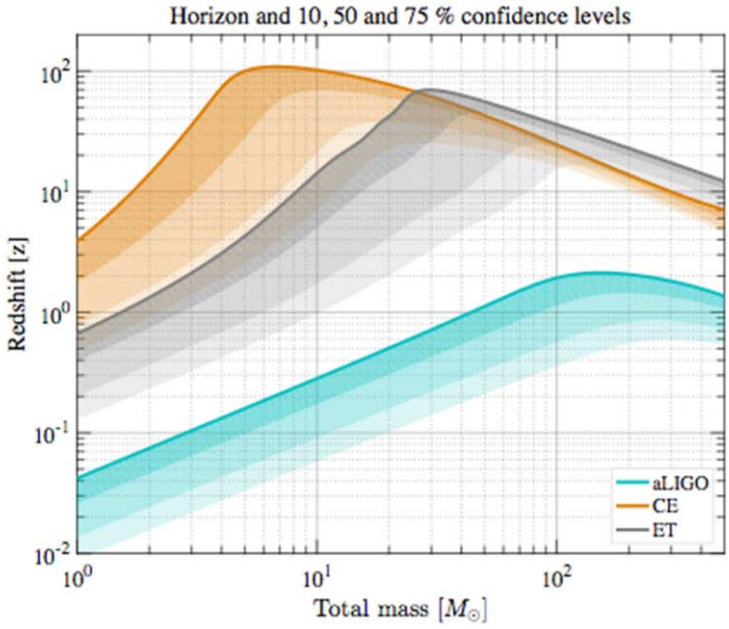


Fig. 12. – Astrophysical reach of Cosmic Explorer and ET for equal-mass (non-spinning) compact binary inspiral systems. The maximum observable distance is shown as a function of the total intrinsic mass of the system. Baring a large number of primordial black holes, at redshifts larger than $z \simeq 10$ there will be few sources. Thus a horizon of $z > 20$ for a given mass should be taken to indicate that essentially all compact binary coalescence in the universe will be observable by a network of similar detectors, many with a high signal-to-noise ratio. Similar curves for a second-generation interferometer (aLIGO) are shown for comparison. A Hubble constant of 67.9 km/s/Mpc and a ΛCDM model of expansion was assumed. Figure from [23].

The main advantage of the xylophone configuration is to decouple technologies that are difficult to coexist in a single detector, like, for example, cryogenics (about 20 K) and high power operations. The ET xylophone concept calls for a low frequency cryogenic detector, with low circulating power, and a high-frequency room temperature detector with high circulating power. The other main features envisaged for ET are: underground operation to limit the effect of the seismic noise and mirrors cooled at low temperature to directly reduce the thermal vibration of the test masses.

The Cosmic Explorer concept relies on an L-shaped configuration with 40 km arms. Increasing the arm length beyond the existing 4 km facilities is crucial to take advantage of the scaling of fundamental noise with length [22].

The 3G detectors aim to improve the sensitivity over the present 2G network by more than a factor of 10, with their astrophysical reach targeting cosmological distances. The astrophysical reach for binary systems can be represented by redshift *vs.* source mass, as illustrated in fig. 12, showing how ET and Cosmic Explorer cover a wide range of binary sources.

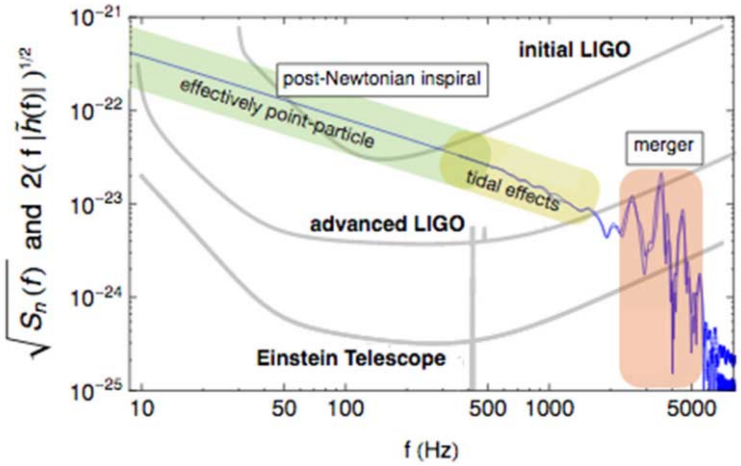


Fig. 13. – A schematic illustration of the gravitational-wave signal emitted during the late stages of a binary neutron star inspiral at 100 Mpc. The effective signal strain is compared to the sensitivity of different generations of detectors at their design sensitivity. Above 100 Hz or so the tidal compressibility is expected to leave a secular imprint on the signal. The eventual merger involves violent dynamics, which also encodes the matter equation of state. The merger signal is expected at a few kHz, making it difficult to observe with the current generation of detectors, but it should be within reach of third generation detectors like the Einstein Telescope. Figure from [24].

The 3G detectors will be able to shed light on hot sectors of physics and astrophysics and will be a fundamental tool to:

- investigate nature of gravity, compact objects and dark matter;
- assess the physics of densest, hottest matter in nature;
- provide the first ever cosmic census of stellar mass binary black holes, star remnants and seeds to supermassive black holes;
- understand the physics of the early Universe, probe its dark sectors and relevant particle physics at the highest energy scales.

All these topics will greatly benefit from the improvement in sensitivity of 3G detectors. As an example, fig. 13 shows the GW signal from a BNS merger compared with the sensitivity curves of first-, second- and third-generation interferometers. The goal of getting a 10× better sensitivities with respect to the advanced detectors, relies on an effective research and development program focused on fundamental technological aspects: development of high-quality, massive-mirror substrates (suitable for operation at cryogenic temperatures, with sensing laser wavelengths different from the present 1 micron); development of low optical and mechanical loss coatings deposited on large areas; suitable light sources (lasers and squeezers) and low-noise suspensions and cryogenic systems.

REFERENCES

[1] VINET J. Y., *Living Rev. Relativ.*, **12** (2009) 5.
[2] Virgo Physics Book, `https://wwwcascina.virgo.infn.it/vpb/vpb2.ps.gz`.
[3] ACCADIA T. *et al.*, *Proceedings of the 12th Marcell Grossmann Meeting*, edited by DAMOUR T., JANTZEN R. T. and RUFFINI R. (World Scientific, Singapore) 2011.
[4] ACCADIA T. *et al.* (VIRGO COLLABORATION), *Journal of Instrumentation*, **7** (2012) P03012.
[5] GODA K., OTTAWAY D., CONNELLY B., ADHIKARI R., MAVALVALA N. and GRETARSSON A., *Opt. Lett.*, **29** (2004) 1452.
[6] ACCADIA T. *et al.*, *Class. Quantum Grav.*, **27** (2010) 084002.
[7] ACCADIA T. *et al.*, *Class. Quantum Grav.*, **28** (2011) 114002.
[8] ACERNESE F. *et al.*, *Class. Quantum Grav.*, **25** (2008) 184001.
[9] LUEK H. *et al.*, *Class. Quantum Grav.*, **21** (2004) S985-S989.
[10] ROCCHI A. *et al.*, *Proceedings of the 46th Rencontres de Moriond*, edited by AUGE E., DUMARCHEZ J. and VAN J. T. T. (The Gioi Publishers, Vietnam) 2011.
[11] ROCCHI A. *et al.*, *J. Phys. Conf. Series*, **363** (2012) 012016.
[12] FAN Y. *et al.*, *Rev. Sci. Instrum.*, **79** (2008) 104501.
[13] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. Lett.*, **116** (2016) 061102.
[14] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. Lett.*, **116** (2016) 241103.
[15] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. D*, **93** (2016) 122003.
[16] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. X*, **6** (2016) 041015.
[17] AASI J. *et al.* (LIGO SCIENTIFIC COLLABORATION), *Class. Quantum. Grav.*, **32** (2015) 074001.
[18] ACERNESE F. *et al.* (VIRGO COLLABORATION), *Class. Quantum. Grav.*, **32** (2015) 024001.
[19] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. Lett.*, **119** (2017) 141101.
[20] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. X*, **9** (2019) 131040.
[21] PUNTURO M. *et al.*, *Class. Quantum Grav.*, **27** (2010) 194002.
[22] ABBOTT B. P. *et al.* (LIGO SCIENTIFIC COLLABORATION, VIRGO COLLABORATION), *Phys. Rev. Lett.*, **119** (2017) 141101.
[23] LIGO Scientific Collaboration, Instrument Science White Paper - LIGO-T1600119-v4 (2016).
[24] ANDERSSON NILS, *J. Astrophys. Astron.*, **38** (2018) 58, DOI 10.1007/s12036-017-9463-8.

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200005

Astrophysics of stellar black holes

Michela Mapelli

INAF-Osservatorio Astronomico di Padova - Vicolo dell'Osservatorio 5, I-35142 Padova, Italy

INFN-Sezione Milano Bicocca - Piazza della Scienza 3, I-20126 Milano, Italy

Institute for Astrophysics and Particle Physics, University of Innsbruck
Technikerstrasse 25/8, A-6020 Innsbruck, Austria

Summary. — On September 14,2015, the LIGO interferometers captured a gravitational wave (GW) signal from two merging black holes (BHs), opening the era of GW astrophysics. Five BH mergers have been reported so far, three of them involving massive BHs ($> 30M_{\odot}$). According to stellar evolution models, such massive BHs can originate from massive relatively metal-poor stars. Alternatively, gravitational instabilities in the early Universe were claimed to produce BHs in this mass range. The formation channels of merging BH binaries are still an open question: a plethora of uncertainties affect the evolution of massive stellar binaries (*e.g.* the process of common envelope) and their dynamics. This review is intended to discuss the open questions about BH binaries, and to present the state-of-the-art knowledge about the astrophysics of black holes for non-specialists, in light of the first LIGO detections.

1. – Lesson learned from the first direct gravitational wave detections

On September 14,2015, the LIGO interferometers captured a gravitational wave (GW) signal from two merging black holes (BHs [1]). This event, named GW150914, is the first direct detection of GWs, about hundred years after Einstein's prediction. To date,

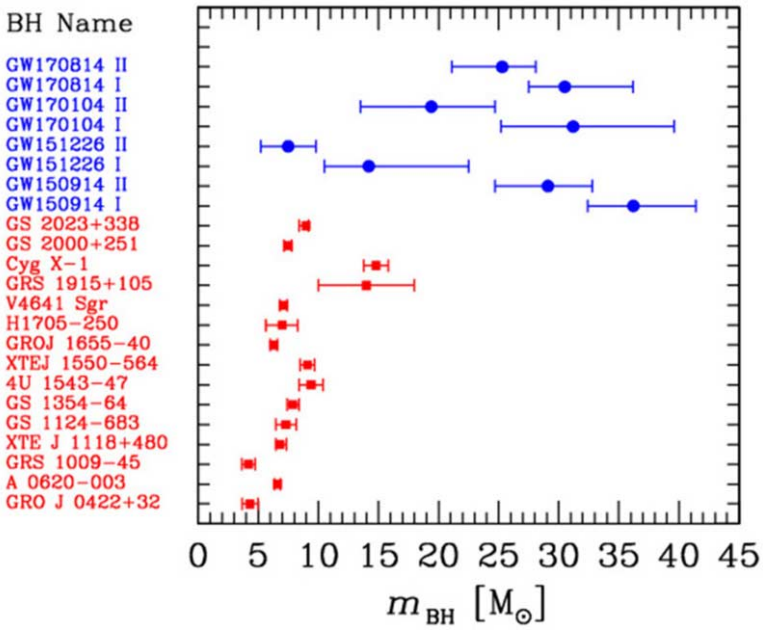


Fig. 1. – A compilation of BH masses $m_{\rm BH}$ from observations. Red squares: BHs with dynamical mass measurement in X-ray binaries [20,21]. This selected sample is quite conservative, because uncertain and debated results are not being shown (*e.g.* IC10 X-1 [22]). Blue circles: BHs in the first four published GW events [3,5,6].

four more BH mergers have been observed (GW151226, GW170104, GW170608 and GW170814 [2-5]), plus a sixth non-confirmed signal (LVT151012 [6]).

Astrophysicists have learned a bunch of breakthrough concepts from GW detections [7]. We now know that double BH binaries (BHBs, *i.e.* binaries composed of two BHs) exist. They have been studied and modelled for a long time (*e.g.* [8-15]), but their observational confirmation was still missing. Moreover, GW detections imply that some BHBs are able to merge within a Hubble time. Finally, three out of five merging BHBs detected so far (GW150914, GW170104 and GW170814) host BHs with mass in excess of $20M_\odot$.

This result is surprising under many respects. First, the only BHs for which we have a dynamical mass measurement, *i.e.* about a dozen of BHs in X-ray binaries, have all mass well below $20M_\odot$ (see fig. 1 for a compilation of measured BH masses). Second, most population-synthesis codes did not predict the existence of BHs with mass $m_{\rm BH} > 30M_\odot$. Thus, the first GW detections have urged the astrophysical community to deeply revise the models of BH formation and evolution.

This review discusses the formation channels of BHs and BHBs in light of the challenges posed by recent GW detections. It is aimed at students and non-expert of the field, being the proceeding of the lecture held for the 200 Course on "Gravitational Waves and Cosmology" at the International School of Physics "Enrico Fermi".

2. – The formation of compact remnants from stellar evolution and supernova explosions

BHs and neutron stars (NSs) are expected to form as remnants of massive ($\gtrsim 8 M_\odot$) stars. An alternative theory predicts that BHs can form also from gravitational collapse in the early Universe (the so called primordial BHs, *e.g.* [16-18]). In this review, we will focus on BHs of stellar origin.

The mass function of BHs is highly uncertain, because it may be affected by a number of barely understood processes. In particular, stellar winds and supernova (SN) explosions both play a major role for the formation of compact remnants. Processes occuring in close binary systems (*e.g.* mass transfer and common envelope) are a further complication and will be discussed in the next section.

2˙1. *Stellar winds and stellar evolution.* – Stellar winds are outflows of gas from the atmosphere of a star. In cold stars (*e.g.* red giants and asymptotic giant branch stars) they are mainly induced by radiation pressure on dust, which forms in the cold outer layers (*e.g.* [19]). In massive hot stars (O and B main-sequence stars, luminous blue variables and Wolf-Rayet stars), stellar winds are powered by the coupling between the momentum of photons and that of metal ions present in the stellar photosphere. A large number of strong and weak resonant metal lines are responsible for this coupling (see *e.g.* [23] for a review).

Understanding stellar winds is tremendously important for the study of compact objects, because mass loss determines the pre-SN mass of a star (both its total mass and its core mass), which in turn affects the outcome of a SN explosion [24-27].

Early work on stellar winds (*e.g.* [28-30]) highlighted that the mass loss of O and B stars depends on metallicity as $\dot{M} \propto Z^\alpha$ (with $\alpha \sim 0.5$–1.0, depending on the model). However, such early work did not account for multiple scattering, *i.e.* for the possibility that a photon interacts several times before being absorbed or leaving the photosphere. Vink *et al.* [31] accounted for multiple scatterings and found a universal metallicity dependence $\dot{M} \propto Z^{0.85} v_\infty^p$, where v_∞ is the therminal velocity and $p = -1.23$ ($p = -1.60$) for stars with effective temperature $T_{\rm eff} \gtrsim 25000\,{\rm K}$ ($12000\,{\rm K} \lesssim T_{\rm eff} \lesssim 25000\,{\rm K}$).

The situation is more uncertain for post-main sequence stars. For Wolf-Rayet (WR) stars, *i.e.* naked helium cores, [32] predict a similar trend with metallicity $\dot{M} \propto Z^{0.86}$. With a different numerical approach (which accounts also for wind clumping), [33] find a strong dependence of WR mass loss on metallicity but also on the electron-scattering Eddington factor $\Gamma_e = \kappa_e L_*/(4\pi c G M_*)$, where κ_e is the cross section for electron scattering, L_* is the stellar luminosity, c is the speed of light, G is the gravity constant, and M_* is the stellar mass. The importance of Γ_e has become increasingly clear in the last few years [34-36], but, unfortunately, only few stellar evolution models include this effect.

For example, [37, 38] adopt a mass loss prescriptions $\dot{M} \propto Z^\alpha$, where $\alpha = 0.85$ if $\Gamma_e < 2/3$ and $\alpha = 2.45 - 2.4\,\Gamma_e$ if $2/3 \leq \Gamma_e \leq 1$. This simple formula accounts for the fact that metallicity dependence tends to vanish when the star is close to be

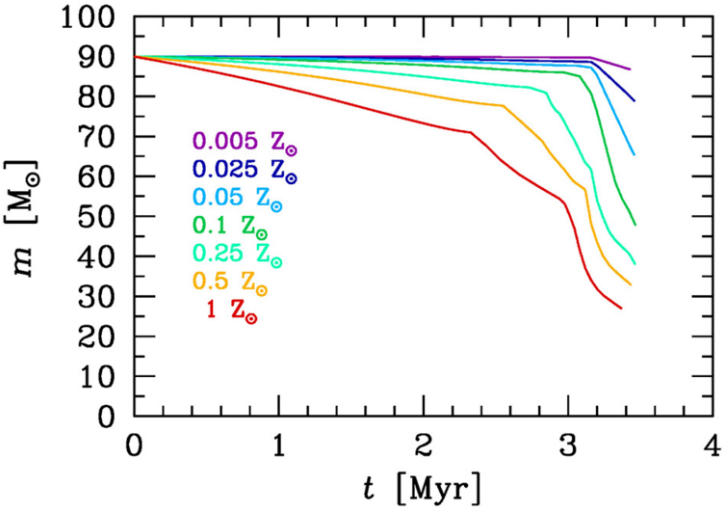


Fig. 2. – Evolution of stellar mass as a function of time for a star with ZAMS mass $m_{\rm ZAMS} = 90M_\odot$ and seven different metallicities, ranging from $0.005Z_\odot$ up to $Z_\odot$ (we assumed $Z_\odot = 0.02$). These curves were obtained with the SEVN population-synthesis code [39], adopting PARSEC stellar evolution tracks.

radiation pressure dominated, as clearly shown by fig. 10 of [33]. Figure 2 shows the mass evolution of a star with zero-age main sequence (ZAMS) mass $m_{\rm ZAMS} = 90M_\odot$ for seven different metallicities, as obtained with the SEVN code [39]. At the end of its life, a solar-metallicity star (here we assume $Z_\odot = 0.02$) has lost more than 2/3 of its initial mass, while the most metal-poor star in the figure ($Z = 0.005Z_\odot$) has retained almost all its initial mass.

Other aspects of massive star evolution also affect the pre-SN mass of a star. For example, surface magnetic fields appear to strongly quench stellar winds by magnetic confinement [40-42]. In particular, [42] show that a non-magnetic star model with metallicity $\sim 0.1Z_\odot$ and a magnetic star model with solar metallicity and Alfvén radius $R_A \sim 4\,R_\odot$ undergo approximately the same mass loss according to this model. This cannot be neglected because surface magnetic fields are detected in ~ 10 per cent hot stars [43], but is currently not included in models of compact-object formation.

Finally, rotation affects the evolution of a massive star in several ways (*e.g.* [44-46]). Describing the effects of rotation in detail is beyond the aims of this review. For this, we refer to a recent review of Marco Limongi [46]. As a general rule of thumb, rotation increases the stellar luminosity. This implies that mass loss is generally enhanced if rotation is accounted for. On the other hand, rotation also induces chemical mixing, which leads to the formation of larger Helium and Carbon-Oxygen cores. While enhanced mass loss implies smaller pre-SN masses, the formation of bigger cores has also strong implication for the final fate of a massive star, as we discuss in the following section.

2˙2. *Supernovae (SNe).* – The mechanisms triggering core-collapse SNe are still highly uncertain. The basic framework and open issues are the following. As the mass of the central degenerate core reaches the Chandrasekhar mass [47], the degeneracy pressure of relativistic electrons becomes insufficient to support it against collapse. Moreover, electrons are increasingly removed, because protons capture them producing neutrons and neutrinos. This transforms the core into a new state, where matter is essentially composed of neutrons, which support the core against collapse by their degeneracy pressure. To reach this new equilibrium, the core collapses from a radius of few thousand km down to a radius of few ten km in less than a second. The gravitational energy gained from the collapse is $W \sim 5 \times 10^{53}\,\mathrm{erg}\,(M_{\mathrm{PNS}}/1.4\,M_\odot)^2\,(10\,\mathrm{km}/R_{\mathrm{PNS}})\,\mathrm{erg}$, where M_{PNS} and R_{PNS} are the mass and radius of the proto-neutron star (PNS).

The main problem is to explain how this gravitational energy can be (at least partially) transferred to the stellar envelope triggering the SN explosion [48, 49]. Several mechanisms have been proposed, including rotationally-driven SNe and/or magnetically-driven SNe (see, *e.g.*, [50-52] and references therein). The most commonly investigated mechanism is the convective SN engine. According to this model, the collapsing core bounces, driving a shock. For the SN explosion to occur, this shock must reverse the supersonic infall of matter from the outer layers of the star. Most of the energy in the shock consists in a flux of neutrinos. As soon as neutrinos are free to leak out (because the shock has become diffuse enough), their energy is lost and the shock stalls. The SN occurs only if the shock is revived by some mechanism. If the region between the PNS surface and the shock stalling radius becomes convectively unstable (*e.g.* because of an entropy gradient), neutrinos can be injected more efficiently and deposit more energy into the shock, possibly reviving it. In other words, if the convective region overcomes the ram pressure of the infalling material, an explosion is launched. If not, the SN fails.

While this is the general idea of the convective engine, the truth is that fully self-consistent simulations of core collapse with state-of-the-art treatment of neutrino transport do not lead to explosions in spherical symmetry except for the lighter SN progenitors ($\lesssim 10 M_\odot$ [52,53]). Simulations which do not require the assumption of spherical symmetry (*i.e.* run at least in 2D) appear to produce successful explosions from first principles for a larger range of progenitor masses (see, *e.g.*, [54, 55]). However, 2D and 3D simulations are still computationally challenging and cannot be used to make a study of the mass distribution of compact remnants.

Thus, in order to study compact-object masses, SN explosions are artificially induced by injecting in the pre-SN model some amount of kinetic energy (kinetic bomb) or thermal energy (thermal bomb) at an arbitrary mass location. The evolution of the shock is then followed by means of 1D hydrodynamical simulations with some relatively simplified treatment for neutrinos. This allows to simulate hundreds of stellar models.

Following this approach, O'Connor and Ott [56] propose a criterion to decide whether a SN is successful or not, based on the compactness parameter:

$$\xi_M = \frac{M/M_\odot}{R(M)/1000\,\mathrm{km}}\,, \tag{1}$$

where $R(M)$ is the radius which encloses a given mass M. Usually, the compactness is defined for $M = 2.5 M_\odot$ ($\xi_{2.5}$). Authors of [56] measure the compactness at core bounce([1]) in their simulations and find that the larger $\xi_{2.5}$, the shorter the time to form a BH (as shown in their fig. 6). This means that stars with a larger value of $\xi_{2.5}$ are more likely to collapse to a BH without SN explosion.

Rotating stellar models have lower values of $\xi_{2.5}$ (because of the centrifugal force) but produce lower neutrino luminosity and thus are more likely to form BHs without a SN explosion than non-rotating models [56].

The work by Ugliano *et al.* [57] and Horiuchi *et al.* [58] indicate that the best threshold between exploding and non-exploding models is $\xi_{2.5} \sim 0.2$.

Finally [53] stress that a single criterion (*e.g.* the compactness) cannot capture the complex physics of core-collapse SN explosion. They introduce a two-parameter criterion based on

$$M_4 = \frac{m(s=4)}{M_\odot} \quad \text{and} \quad \mu_4 = \left[\frac{\mathrm{d}m/M_\odot}{\mathrm{d}R/1000\,\mathrm{km}}\right]_{s=4}, \tag{2}$$

where M_4 is the mass (at the onset of collapse) where the dimensionless entropy per baryon is $s = 4$, and μ_4 is the spatial derivative at the location of M_4.

This choice is motivated by the fact that, in their 1D simulations, the explosion sets shortly after M_4 has fallen through the shock and well before the shell enclosing $M_4 + 0.3 M_\odot$ has collapsed. They show that exploding models can be distinguished from non-exploding models in the μ_4 *versus* $M_4\,\mu_4$ plane (see their fig. 6) by a linear fit

$$y(x) = k_1\, x + k_2, \tag{3}$$

where $y(x) = \mu_4$, $x = M_4\,\mu_4$, and k_1 and k_2 are numerical coefficients which depend on the model (see table 2 of [53]). The reason of this behaviour is that μ_4 scales with the rate of mass infall from the outer layers (thus the larger μ_4, the lower the chance of the SN to occur), while $M_4\,\mu_4$ scales with the neutrino luminosity (thus the larger $M_4\,\mu_4$, the higher the chance of a SN explosion). Finally, [53] stress that the fallback is quite inefficient ($< 0.05 M_\odot$) when the SN occurs.

The models discussed so far depend on quantities ($\xi_{2.5}$, M_4, μ_4) which can be evaluated no earlier than the onset of core collapse. Thus, stellar evolution models are required which integrate a massive star till the iron core has formed. This is prohibitive for most stellar evolution models (with few remarkable exceptions, *e.g.* FRANEC [45] and MESA [59]). Reference [60] propose a simplified approach. Based on their simulations (see also [24,25,61]), they propose that the mass of the compact remnants depends on two quantities: the carbon-oxygen core mass $m_{\rm CO}$ and the total final mass of the star $m_{\rm fin}$. In particular, $m_{\rm CO}$ determines whether the star will undergo a core-collapse

([1]) Authors of [57] show that $\xi_{2.5}$ is not significantly different at core bounce or at the onset of collapse.

SN or will collapse to a BH directly (namely, stars with $m_{\rm CO} > 7.6 M_\odot$ collapse to a BH directly), whereas $m_{\rm fin}$ determines the amount of fallback on the proto NS. In this formalism, the only free parameter is the time to launch the shock. The explosion energy is significantly reduced if the shock is launched $\gg 250$ ms after the onset of the collapse (*delayed* SN explosion) with respect to an explosion launched in the first ~ 250 ms (*rapid* SN explosion [60]).

While this approach is quite simplified with respect to other prescriptions, [46] show that there is a strong correlation between the final carbon-oxygen mass and the compactness parameter $\xi_{2.5}$ at the onset of collapse, regardless of the rotation velocity of the progenitor star (see fig. 21 of [46]).

Finally, it is important to recall pair-instability and pulsational pair-instability SNe [62-65]. If the helium core of a star grows above $\sim 30 M_\odot$ and the core temperature is $\gtrsim 7 \times 10^8$ K, the process of electron-positron pair production becomes effective. It removes photon pressure from the core producing a sudden collapse before the iron core is formed. For $M_{\rm He} > 135 M_\odot$, the collapse cannot be reversed and the star collapses directly in to a BH [65]. If $135 \gtrsim M_{\rm He} \gtrsim 64 M_\odot$, the collapse triggers an explosive burning of heavier elements, which has disruptive effects. This leads to a complete disruption of the star, leaving no remnant (the so-called pair-instability SN [66]). For $64 \gtrsim M_{\rm He} \gtrsim 32 M_\odot$, pair production induces a series of pulsations of the core (pulsational pair instability SNe), which trigger an enhanced mass loss [65]. At the end of this instability phase a remnant with non-zero mass is produced, significantly lighter than in case of a direct collapse.

2˙3. *The mass of compact remnants.* – The previous sections suggest that our knowledge of remnant mass is hampered by severe uncertainties, connected with both stellar winds and core-collapse SNe. Thus, models of the mass spectrum of compact remnants must be taken with a grain of salt. However, few robust features can be drawn.

Figure 3 is a simplified version of figs. 2 and 3 of Heger *et al.* [67]. The final mass of a star and the mass of the compact remnant are shown as a function of the ZAMS mass. The left and the right-hand panels show the case of a solar metallicity star and of a metal-free star, respectively. In the case of the solar metallicity star, the final mass of the star is much lower than the initial one, because stellar winds are extremely efficient. The mass of the compact remnant is also much lower than the final mass of the star because a core-collapse SN always take place.

In contrast, a metal-free star (*i.e.* a Population III star) loses a negligible fraction of its mass by stellar winds (the blue and the black line in fig. 3 are superimposed). As for the mass of the compact remnant, fig. 3 shows that there are two regimes: below a given threshold (≈ 30–$40 M_\odot$) the SN explosion succeeds even at zero metallicity and the mass of the compact remnant is relatively small. Above this threshold, the mass of the star (in terms of both core mass and envelope mass) is sufficiently large that the SN fails. Most of the final stellar mass collapses to a BH, whose mass is significantly larger than in the case of a SN explosion.

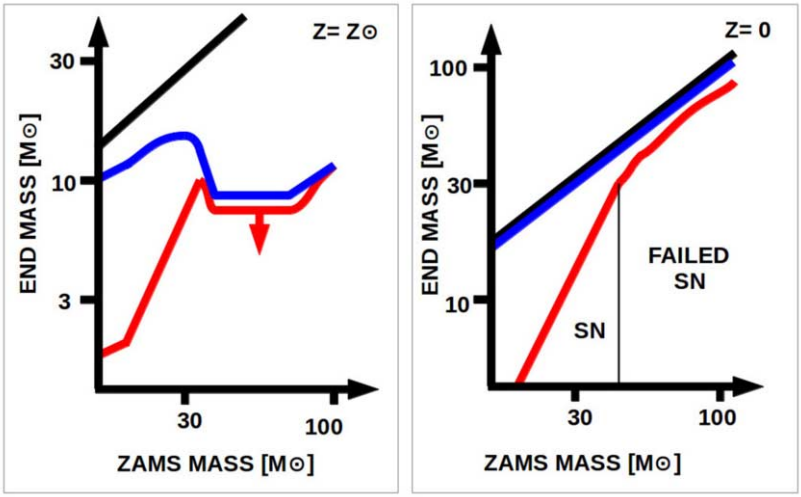


Fig. 3. – Final mass of a star ($m_{\rm fin}$, blue lines) and mass of the compact remnant (m_{rem}, red lines) as a function of the ZAMS mass of the star. The thick black line marks the region where $m_{\rm fin} = m_{\rm ZAMS}$. Left-hand panel: solar metallicity star. Right-hand panel: metal-free star. The red arrow on the left-hand panel is an upper limit for the remnant mass. Vertical thin black line in the right-hand panel: approximate separation between successful and failed SNe at $Z = 0$. This cartoon was inspired by figs. 2 and 3 of Heger *et al.* [67].

What happens at intermediate metallicity between solar and zero, *i.e.* in the vast majority of the Universe we know? Predicting what happens to a metal-free star is relatively simple, because its evolution does not depend on the interplay between metals and stellar winds. The fate of a solar metallicity star is more problematic, because we must account for line-driven stellar winds, but most data we have about massive star winds are for nearly solar metallicity stars, which makes models easier to calibrate. Instead, modelling intermediate metallicities is significantly more complicated, because the details depend on the interplay between metals and stellar winds and only limited data are available for calibration (mostly data for the Large and Small Magellanic Clouds).

As a rule of thumb (see, *e.g.*, [39, 60]), we can draw the following considerations. If the zero-age main-sequence (ZAMS) mass of a star is large ($m_{\rm ZAMS} \gtrsim 30 M_\odot$), then the amount of mass lost by stellar winds is the main effect which determines the mass of the compact remnant. At low metallicity ($\lesssim 0.1 Z_\odot$) and for a low Eddington factor ($\Gamma_e < 0.6$), mass loss by stellar winds is not particularly large. Thus, the final mass $m_{\rm fin}$ and the carbon-oxygen mass $m_{\rm CO}$ of the star may be sufficiently large to avoid a core-collapse SN explosion: this leads the star to form a massive BH ($\gtrsim 20 M_\odot$) by direct collapse, unless pair-instability and pulsational-pair instability SNe occur. At high metallicity ($\approx Z_\odot$) or large Eddington factor ($\Gamma_e > 0.6$), mass loss by stellar winds is particularly efficient and may lead to a small $m_{\rm fin}$ and $m_{\rm CO}$: the star is expected to undergo a core-collapse SN and to leave a relatively small remnant.

If the ZAMS mass of a star is relatively low ($7 < m_{\rm ZAMS} < 30 M_\odot$), then stellar winds are not important (with the possible exception of super asymptotic giant branch stars), regardless of the metallicity. In this case, the details of the SN explosion (*e.g.* energy of the explosion and amount of fallback) are crucial to determine the final mass of the remnant.

This general sketch may be affected by several factors, such as pair-instability SNe, pulsational pair-instability SNe (*e.g.* [65]) and an *island scenario* for core-collapse SNe (*e.g.* [53]).

Taking these effects into account leads to a prediction for the mass spectrum of compact remnants similar to the one shown in fig. 4. This figure shows that the mass of the compact remnant strongly depends on the metallicity of the progenitor star if $m_{\rm ZAMS} \gtrsim 30 M_\odot$. In most cases, the lower the metallicity of the progenitor, the larger the maximum mass of the compact remnant [26, 27, 39, 67-69]. However, for metal-poor stars ($Z < 10^{-3}$) with ZAMS mass $230 > m_{\rm ZAMS} > 110 M_\odot$ pair instability SNe lead to the complete disruption of the star and no remnant is left. Only very massive ($m_{\rm ZAMS} > 230 M_\odot$) metal-poor ($Z < 10^{-3}$) stars can collapse to a BH directly, producing intermediate-mass BHs (*i.e.* BHs with mass $\gtrsim 100 M_\odot$).

The mass spectrum of relatively low-mass stars ($8 < M_{\rm ZAMS} < 30 M_\odot$) is not significantly affected by metallicity. The assumed core-collapse SN model is the most important factor in this mass range [60].

3. – Binaries of stellar black holes

Naively, one could think that if two massive stars are members of a binary system, they will eventually become a double BH binary and the mass of each BH will be the same as if its progenitor star was a single star. This is true only if the binary system is sufficiently wide (detached binary) for its entire evolution. If the binary is close enough, it will evolve through several processes which might significantly change its final fate. Here, we will mention some of the most important ones.

3˙1. *Mass transfer*. – If two stars exchange matter to each other, it means they undergo a mass transfer episode. This might be driven either by stellar winds or by an episode of Roche-lobe filling.

When a massive star loses mass by stellar winds, its companion might be able to capture some of this mass. This will depend on the amount of mass which is lost and on the relative velocity of the wind with respect to the companion star. Based on the Bondi and Hoyle [71] formalism, Hurley *et al.* [72] describe the mean mass accretion rate by stellar winds as

$$\dot{M}_{\rm A} = \frac{1}{\sqrt{1-e^2}} \left(\frac{G\, M_{\rm A}}{v_{\rm w}^2} \right)^2 \frac{\alpha_{\rm w}}{2\, a^2} \frac{1}{[1 + (v_{\rm orb}/v_{\rm w})^2]^{3/2}} \dot{M}_{\rm D}, \tag{4}$$

where e is the binary eccentricity, G is the gravitational constant, $M_{\rm A}$ is the mass of the accreting star, $v_{\rm w}$ is the velocity of the wind, $\alpha_{\rm w} \sim 3/2$ is an efficiency constant, a is the semi-major axis of the binary, $v_{\rm orb} = \sqrt{G\,(M_{\rm A} + M_{\rm D})/a}$ is the orbital velocity of the

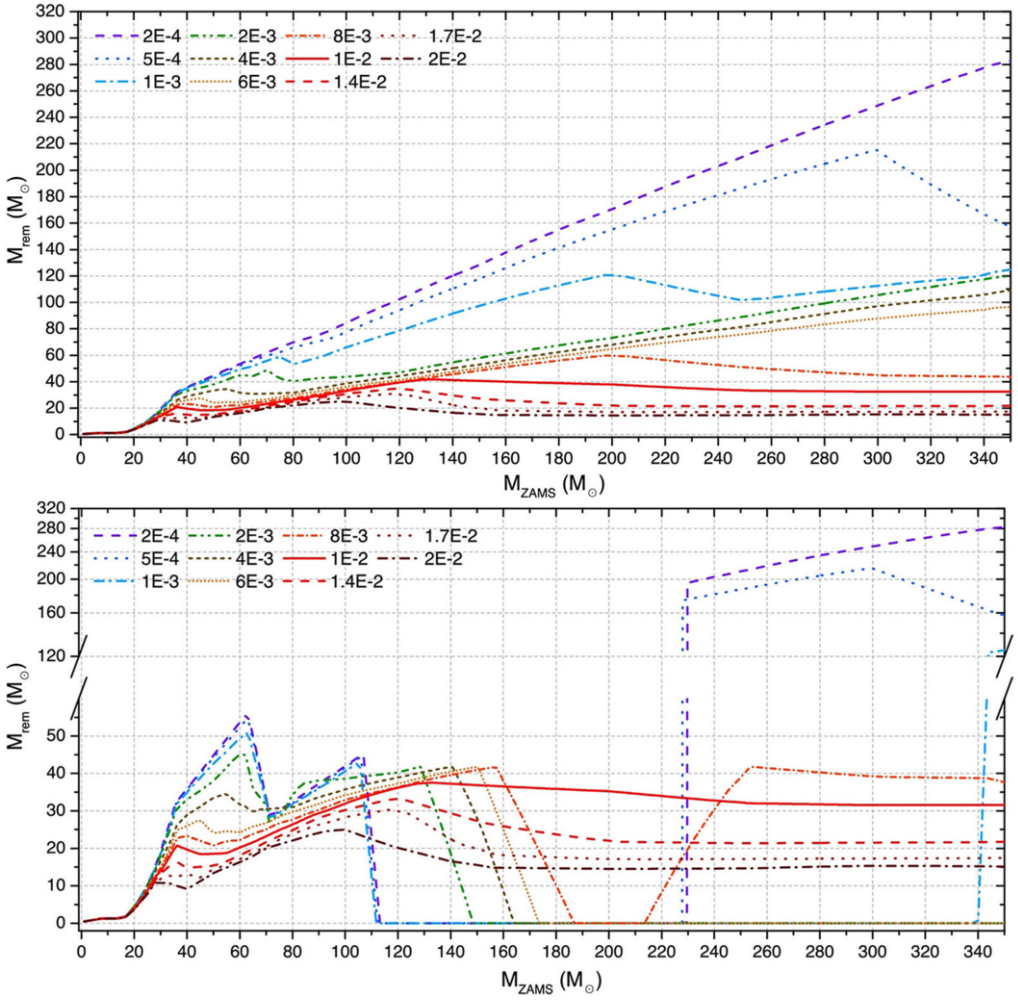


Fig. 4. – Mass of the compact remnant ($M_{\rm rem}$) as a function of the ZAMS mass of the star ($M_{\rm ZAMS}$) from Spera and Mapelli [70]. Lower (upper) panel: pulsational pair-instability and pair-instability SNe are (not) included. In both panels: dash-dotted brown line: $Z = 2.0 \times 10^{-2}$; dotted dark orange line: $Z = 1.7 \times 10^{-2}$; dashed red line: $Z = 1.4 \times 10^{-2}$; solid red line: $Z = 1.0 \times 10^{-2}$; short-dash-dotted orange line: $Z = 8.0 \times 10^{-3}$; short dotted light orange line: $Z = 6.0 \times 10^{-3}$; short-dashed green line: $Z = 4.0 \times 10^{-3}$; dash-double-dotted green line: $Z = 2.0 \times 10^{-3}$; dash-dotted light blue line: $Z = 1.0 \times 10^{-3}$; dotted blue line: $Z = 5.0 \times 10^{-4}$; dashed violet line: $Z = 2.0 \times 10^{-4}$. A delayed core-collapse SN mechanism has been assumed, following the prescriptions of [60]. From figs. 1 and 2 of [70].

binary ($M_{\rm D}$ being the mass of the donor), and $\dot{M}_{\rm D}$ is the mass loss rate by the donor in a given time span. Since $\dot{M}_{\rm D}$ is usually quite low ($< 10^{-3} M_\odot\,{\rm yr}^{-1}$) and $v_{\rm w}$ is usually quite high ($> 1000\,{\rm km\,s}^{-1}$ for a line-driven wind) with respect to the orbital velocity, this kind of mass transfer is usually rather inefficient.

Mass transfer by Roche lobe overflow is usually more efficient. The Roche lobe of a star in a binary system is the maximum equipotential surface around the star within which matter is bound to the star. While the exact shape of the Roche lobe should be calculated numerically, a widely used approximate formula [73] is

$$r_{\mathrm{L},1} = a\,\frac{0.49\,q^{2/3}}{0.6\,q^{2/3} + \ln(1+q^{1/3})}\,, \tag{5}$$

where a is the semi-major axis of the binary and $q = m_1/m_2$ (m_1 and m_2 are the masses of the two stars in the binary).

The Roche lobes of the two stars in a binary are thus connected by the L1 Lagrangian point. Since the Roche lobes are equipotential surfaces, matter orbiting at or beyond the Roche lobe can flow freely from one star to the other. We say that a star overfills (underfills) its Roche lobe when its radius is larger (smaller) than the Roche lobe. If a star overfills its Roche lobe, a part of its mass flows toward the companion star which can accrete (a part of) it. The former and the latter are thus called donor and accretor star, respectively.

Mass transfer obviously changes the mass of the two stars in a binary, and thus the final mass of the compact remnants of such stars, but also the orbital properties of the binary. If mass transfer is non-conservative (which is the most realistic case in both mass transfer by stellar winds and Roche lobe overflow), it leads to an angular-momentum loss, which in turn affects the semi-major axis.

An important point about Roche lobe overflow is to estimate whether it is (un)stable an on which timescale. A commonly used approach consists in comparing the following quantities [72, 74-76]:

$$\begin{aligned}
\zeta_{\mathrm{ad}} &= \left(\frac{\mathrm{d}\ln R_{\mathrm{D}}}{\mathrm{d}\ln M_{\mathrm{D}}}\right)_{\mathrm{ad}}\,,\\
\zeta_{\mathrm{th}} &= \left(\frac{\mathrm{d}\ln R_{\mathrm{D}}}{\mathrm{d}\ln M_{\mathrm{D}}}\right)_{\mathrm{th}}\,,\\
\zeta_{\mathrm{L}} &= \left(\frac{\mathrm{d}\ln r_{\mathrm{L,D}}}{\mathrm{d}\ln M_{\mathrm{D}}}\right),
\end{aligned} \tag{6}$$

where ζ_{ad} is the change of radius of the donor (induced by the mass loss) needed to adiabatically adjust the star to a new hydrostatic equilibrium, ζ_{th} is the change of the radius of the donor (induced by the mass loss) needed to adjust the star to a new thermal equilibrium, and ζ_{L} is the change of the Roche lobe of the donor (induced by the mass loss).

If $\zeta_{\mathrm{L}} > \zeta_{\mathrm{ad}}$, then the star expands faster than the Roche lobe and mass transfer is dynamically unstable. If $\zeta_{\mathrm{ad}} > \zeta_{\mathrm{L}} > \zeta_{\mathrm{th}}$, then mass transfer becomes unstable over a Kelvin-Helmholtz timescale. Finally, if $\zeta_{\mathrm{ad}} > \zeta_{\mathrm{th}} > \zeta_{\mathrm{L}}$, mass transfer is stable until stellar evolution causes a further expansion of the radius.

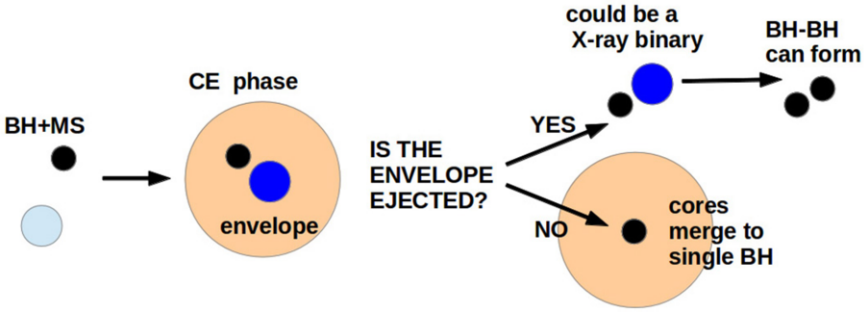


Fig. 5. – Cartoon of the evolution of a BH binary which evolves through a CE phase. The companion of the BH is initially in the main sequence (MS). When the companion evolves off the MS, becoming a giant star, it overfills its Roche lobe. The BH and the giant star enter a CE. The core of the giant and the BH spiral in because of the gas drag exerted by the envelope. If the envelope is ejected, we are left with a new binary, composed of the BH and the naked helium core of the giant. The new binary has a much smaller orbital separation than the initial binary. If the naked helium core becomes a BH and its natal kick does not disrupt the binary, then a double BH binary is born, possibly with a small semi-major axis. In contrast, if the envelope is not ejected, the BH and the helium core spiral in, till they merge together. A single BH is left, which will not be a source of GWs.

If mass transfer is dynamically unstable ($\zeta_{\rm L} > \zeta_{\rm ad}$) or both stars overfill their Roche lobe, then the binary is expected to merge (if the donor lacks a steep density gradient between the core and the envelope) or to enter common envelope (CE).

3˙2. *Common envelope (CE).* – If two stars enter in CE, their envelope(s) stop co-rotating with their cores. The two stellar cores (or the compact remnant and the core of the star, if the binary is already single degenerate) are embedded in the same non-corotating envelope and start spiralling in as an effect of gas drag exerted by the envelope. Part of the orbital energy lost by the cores as an effect of this drag is likely converted into heating of the envelope, making it more loosely bound. If this process leads to the ejection of the envelope, then the binary survives, but the post-CE binary is composed of two naked stellar cores (or a compact remnant and a naked stellar core). Moreover, the orbital separation of the two cores (or the orbital separation of the compact remnant and the core) is considerably smaller than the initial orbital separation of the binary, as an effect of the spiral in. This circumstance is crucial for the fate of a BH binary. In fact, if the binary which survives a CE phase evolves into a double BH binary, this double BH binary will have a very short semi-major axis, much shorter than the sum of the maximum radii of the progenitor stars, and may be able to merge by GW emission within a Hubble time.

In contrast, if the envelope is not ejected, the two cores (or the compact remnant and the core) spiral in till they eventually merge. This premature merger of a binary during a CE phase prevents the binary from evolving into a double BH binary. The cartoon in fig. 5 summarizes these possible outcomes.

The $\alpha\lambda$ formalism [77] is the most common formalism adopted to describe a common

envelope. The basic idea of this formalism is that the energy needed to unbind the envelope comes uniquely from the loss of orbital energy of the two cores during the spiral-in.

The fraction of the orbital energy of the two cores which goes into unbinding the envelope can be expressed as

$$\Delta E = \alpha \left(E_{\rm b,f} - E_{\rm b,i}\right) = \alpha \, \frac{G \, m_{\rm c1} \, m_{\rm c2}}{2} \, \left(\frac{1}{a_{\rm f}} - \frac{1}{a_{\rm i}}\right), \tag{7}$$

where $E_{\rm b,i}$ ($E_{\rm b,f}$) is the orbital binding energy of the two cores before (after) the CE phase, $a_{\rm i}$ ($a_{\rm f}$) is the semi-major axis before (after) the CE phase, $m_{\rm c1}$ and $m_{\rm c2}$ are the masses of the two cores, and α is a dimensionless parameter which measures which fraction of the removed orbital energy is transferred into the envelope. If the primary is already a compact object (as in fig. 5), $m_{\rm c1}$ is the mass of the compact object.

The binding energy of the envelope is

$$E_{\rm env} = \frac{G}{\lambda} \, \left[\frac{m_{\rm env,1} \, m_1}{R_1} + \frac{m_{\rm env,2} \, m_2}{R_2}\right], \tag{8}$$

where m_1 and m_2 are the masses of the primary and the secondary member of the binary, $m_{\rm env,1}$ and $m_{\rm env,2}$ are the masses of the envelope of the primary and the secondary member of the binary, R_1 and R_2 are the radii of the primary and the secondary member of the binary, and λ is the parameter which measures the concentration of the envelope (the smaller λ, the more concentrated is the envelope).

By imposing $\Delta E = E_{\rm env}$ we can derive which is the value of the final semi-major axis $a_{\rm f}$ for which the envelope is ejected:

$$\frac{1}{a_{\rm f}} = \frac{1}{\alpha \, \lambda} \, \frac{2}{m_{\rm c1} \, m_{\rm c2}} \, \left[\frac{m_{\rm env,1} \, m_1}{R_1} + \frac{m_{\rm env,2} \, m_2}{R_2}\right] + \frac{1}{a_{\rm i}} \, . \tag{9}$$

If $a_{\rm f}$ is lower than the sum of the radii of the two cores (or than the sum of the Roche lobe radii of the cores), then the binary will merge during CE, otherwise the binary survives and eq. (9) tells us the final orbital separation. This means that the larger (smaller) $\alpha \, \lambda$, the larger (smaller) the final orbital separation.

Actually, we have known for a long time (see [78] for a review) that this simple formalism is a poor description of the physics of CE, which is considerably more complicated. For example, there is a number of observed systems for which an $\alpha > 1$ is required, which is obviously unphysical. Moreover, λ cannot be the same for all stars. It is expected to vary wildly not only from star to star but also during different evolutionary stages of the same star. Several authors [79,80] have estimated $E_{\rm env}$ directly from their stellar models, which removes the λ parameter from eq. (9) and significantly improves this formalism. However, even in this case, we cannot get rid of the α parameter.

Figure 6 shows the distribution of total masses of merging BHs obtained with the same code (MOBSE [81]) by changing solely the value of $\alpha\lambda$. The difference between the

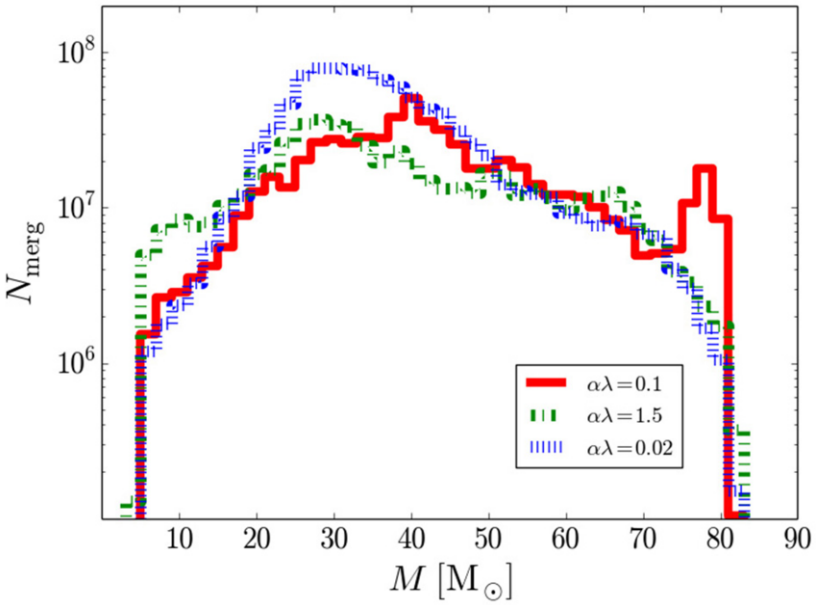


Fig. 6. – Distribution of total masses ($M = m_1 + m_2$) of merging BH binaries in the LIGO-Virgo instrumental horizon, obtained with the MOBSE code [81]. Merging binaries come from progenitors with different metallicity and were sampled as described in [86]. The only difference between the three histograms is the value of $\alpha\lambda$ in the CE formalism. Red solid line: $\alpha\lambda = 0.1$; green dot-dashed line: $\alpha\lambda = 1.5$; blue dotted line: $\alpha\lambda = 0.02$.

three mass distributions is a clear example of how important is CE for the demography of BH binaries.

Thus, it would be extremely important to model the CE in detail, for example with numerical simulations. A lot of effort has been put on this in the last few years, but there are still many open questions (see the review by [78]). For example, we do not have self-consistent models of the onset of CE, when an unstable mass transfer prevents the envelope from co-rotating with the core. Usually, hydrodynamical simulations of CE start when the core of the companion is already at the surface of the envelope.

The only part of CE which has been successfully modelled by several authors (*e.g.*, [82-85]) is the initial spiral in phase, when the two cores spiral in on a dynamical time scale (≈ 100 days).

However, at the end of this dynamical spiral in only a small fraction of the envelope (~ 25 per cent [85]) appears to be ejected in most simulations. When the two cores are sufficiently close that they are separated only by a small gas mass, the spiral-in slows down and the system evolves on the Kelvin-Helmholtz timescale of the envelope ($\approx 10^{3-5}$ years). Simulating the system for a Kelvin-Helmholtz timescale is prohibitive for current simulations. Thus, the description of CE remains the conundrum of massive binary evolution.

3˙3. *Alternative evolution to CE.* – Massive fast-rotating stars can have a chemically homogeneous evolution (CHE): they do not develop a chemical composition gradient because of the mixing induced by rotation. This is particularly true if the star is metal poor, because stellar winds are not efficient in removing angular momentum. If a binary is very close, the spins of its members are even increased during stellar life, because of tidal synchronisation. The radii of stars following CHE are usually much smaller than the radii of stars developing a chemical composition gradient [87, 88]. This implies that even very close binaries (few tens of solar radii) can avoid CE.

Authors of ref. [89] simulate very close binaries whose components are fast-rotating massive stars. A number of their simulated binaries evolve into contact binaries where both binary components fill and even overfill their Roche volumes. If metallicity is sufficiently low and rotation sufficiently fast, these binaries may evolve as "over-contact" binaries: the over-contact phase differs from a classical CE phase because co-rotation can, in principle, be maintained as long as material does not overflow the L2 point. This means that a spiral-in that is due to viscous drag can be avoided, resulting in a stable system evolving on a nuclear timescale.

Such over-contact binaries maintain relatively small stellar radii during their evolution (few tens of the solar radius) and may evolve into a double BH binary with a very short orbital period. This scenario predicts the formation of merging BHs with relatively large masses ($> 20 M_\odot$), nearly equal mass ($q = 1$), and with large aligned spins. The latter prediction is quite at odds with the effective spins of GW150914, GW170104, and GW170814.

4. – The dynamics of black hole binaries

In the previous sections of this review, we discussed the formation of BH binaries as isolated binaries. There is an alternative channel for BH binary formation: the dynamical evolution scenario.

4˙1. *Dynamically active environments.* – Collisional dynamics is important for the evolution of binaries only if they are in a dense environment ($\gtrsim 10^3$ stars pc^{-3}), such as a star cluster. On the other hand, astrophysicists believe that the vast majority of massive stars (which are BH progenitors) form in star clusters [90-93].

Most studies of dynamical formation of BH binaries focus on globular clusters (*e.g.*, [12, 13, 94-100]). *Globular clusters* are old stellar systems ($\sim 12\,\mathrm{Gyr}$), mostly very massive ($> 10^4 M_\odot$) and dense ($> 10^4 M_\odot\,\mathrm{pc}^{-3}$). They are sites of intense dynamical processes (such as the gravothermal catastrophe). However, globular clusters represent a tiny fraction of the baryonic mass in the Universe ($\lesssim 1$ per cent [101]).

In contrast, only few studies of BH binaries (*e.g.* [102-107]) focus on *young star clusters*. These young ($\lesssim 100\,\mathrm{Myr}$), relatively dense ($> 10^3 M_\odot\,\mathrm{pc}^{-3}$) stellar systems are thought to be the most common birthplace of massive stars. When they evaporate (by gas loss) or are disrupted by the tidal field of their host galaxy, their stellar content is released into the field. Thus, it is reasonable to expect that a large fraction of BH

binaries which are now in the field may have formed in young star clusters, where they participated in the dynamics of the cluster. The reason why young star clusters have been neglected in the past is exquisitely numerical: the dynamics of young star clusters needs to be studied with direct N-body simulations, which are rather expensive (they scale as N^2), combined with population-synthesis simulations. Moreover, their dynamical evolution may be significantly affected by the presence of gas. Including gas would require a challenging interface between direct N-body simulations and hydrodynamical simulations, which has been done in very few cases [108-115] and has been never used to study BH binaries.

A third flavour of star cluster where BH binaries might form and evolve dynamically are *nuclear star clusters*, *i.e.* star clusters which lie in the nuclei of galaxies. Nuclear star clusters are rather common in galaxies (*e.g.* [116-118]), are usually more massive and denser than globular clusters, and may co-exist with super-massive black holes. These features make them unique among star clusters, for the effects that we shall describe in the next sections.

4˙2. *Three-body encounters*. – We now review what are the main dynamical effects which can affect a BH binary, starting from three-body encounters.

Binaries have a energy reservoir, their internal energy:

$$E_{\rm int} = \frac{1}{2}\,\mu\,v^2 - \frac{G\,m_1\,m_2}{r}\,, \tag{10}$$

where $\mu = m_1\,m_2/(m_1+m_2)$ is the reduced mass of the binary (whose components have mass m_1 and m_2), v is the relative velocity between the two members of the binary, and r is the distance between the two members of the binary. As shown by Kepler's laws, $E_{\rm int} = -E_{\rm b} = -G\,m_1\,m_2/(2\,a)$, where $E_{\rm b}$ is the binding energy of the binary (a being the semi-major axis of the binary).

The internal energy of a binary can be exchanged with other stars only if the binary undergoes a close encounter with a star, so that its orbital parameters are perturbed by the intruder. This happens only if a single star approaches the binary by few times its orbital separation. We define this close encounter between a binary and a single star as a *three-body encounter*. For this to happen with a non-negligible frequency, the binary must be in a dense environment, because the rate of three-body encounters scales with the local density of stars.

Three-body encounters have crucial effects on BH binaries, such as *exchanges, hardening*, and *ejections*.

4˙3. *Exchanges*. – Dynamical exchanges are three-body encounters during which one of the former members of the binary is replaced by the intruder (see fig. 7).

Exchanges may lead to the formation of new double BH binaries. As is shown in fig. 7, if a binary composed of a BH and a low-mass star undergoes an exchange with a single BH, this leads to the formation of a new double BH binary. This is a very important difference between BHs in the field and in star clusters: a BH which forms as

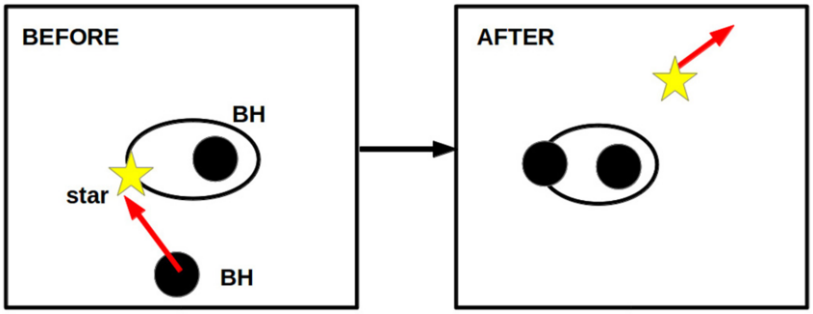


Fig. 7. – Cartoon of a dynamical exchange. A binary composed of a BH and a star interacts with a BH, which replaces the star.

a single object in the field has negligible chances to become member of a binary system, while a single BH in the core of a star cluster has good chances of becoming member of a binary by exchanges.

Exchanges are expected to lead to the formation of many more double BH binaries than they can destroy, because the probability for an intruder to replace one of the members of a binary is ≈ 0 if the intruder is less massive than both binary members, while it suddenly jumps to ~ 1 if the intruder is more massive than one of the members of the binary [119]. Since BHs are among the most massive bodies in a star cluster (after their massive progenitors transform into them), they are very efficient in acquiring companions through dynamical exchanges.

Thus, exchanges are a crucial mechanism to form BH binaries dynamically. By means of direct N-body simulations, Ziosi *et al.* [102] show that > 90 per cent of double BH binaries in young star clusters form by dynamical exchange.

Moreover, BH binaries formed via dynamical exchange will have some distinctive features with respect to field BH binaries (see, *e.g.*, [102]):

– double BH binaries formed by exchanges will be (on average) more massive than isolated double BH binaries, because more massive intruders have higher chances to acquire companions;

– exchanges trigger the formation of highly eccentric double BH binaries (eccentricity is then significantly reduced by circularisation induced by GW emission, if the binary enters the regime where GW emission is effective);

– double BH binaries born by exchange will likely have misaligned spins, because exchanges tend to randomize the spins.

Zevin *et al.* [120] compare a set of simulations of field binaries with a set of simulations of globular cluster binaries, run with the same population-synthesis code. The most striking difference between merging BH binaries in their globular cluster simulations and in their population-synthesis simulations is the dearth of merging BHs with mass

$< 10 M_\odot$ in the globular cluster simulations. This is due to the fact that exchanges tend to destroy binaries composed of light BHs.

Spin misalignments are another possible feature to discriminate between field binaries and star cluster binaries (*e.g.*, [121, 122]). Unfortunately, there is no robust theory to predict the magnitude of the spin of a BH given the spin of its parent star [123]. In principle, we do not know whether BH spins are low (~ 0) or high (~ 1), given the spin of the parent star. However, we can reasonably affirm that the orientation of the spin of a BH matches the orientation of the spin of its progenitor star, if the latter evolved in isolation and directly collapsed to a BH.

Thus, we expect that an isolated binary in which the secondary becomes a BH by direct collapse results in a double BH binary with aligned spins (*i.e.* the spins of the two BHs have the same orientation, which is approximately the same as the orbital angular-momentum direction of the binary), because tidal evolution and mass transfer in a binary tend to synchronise the spins [72]. On the other hand, if the secondary undergoes a SN explosion, the natal kick may reshuffle spins. For dynamically formed BH binaries (through exchange) we expect misaligned, or even nearly isotropic spins, because any original spin alignment is completely reset by three-body encounters.

Currently, we have only poor constraints on BH spins for the first three GW detections. LIGO can only provide a measurement for the effective spin $\chi_{\rm eff}$ which is the sum of the components of the spins of the two BHs along the orbital angular-momentum vector of the binary:

$$\chi_{\rm eff} = \frac{c}{G\,(m_1+m_2)}\left(\frac{\vec{S}_1}{m_1}+\frac{\vec{S}_2}{m_2}\right)\cdot\frac{\vec{L}}{|\vec{L}|} \equiv \frac{1}{m_1+m_2}\,(m_1\,\chi_1+m_2\,\chi_2), \tag{11}$$

where $\vec{S}_1$ and $\vec{S}_2$ are the spin angular-momentum vectors of the BHs, $\vec{L}$ is the orbital angular momentum of the binary, χ_1 and χ_2 are the dimensionless projections of the individual BH spins, respectively. By construction, $-1 \leq \chi_{\rm eff} \leq 1$. A positive (negative) value of $\chi_{\rm eff}$ means that the spins of the two BHs are aligned (counter-aligned).

Only for GW151226, the measured value of $\chi_{\rm eff}$ is significantly larger than zero, indicating (at least partial) alignment. For GW150914, LVT151012, GW170104 and GW170814, the effective spin is consistent with zero and may be either negative or positive. With a Bayesian approach, [121] estimate that an isotropic distribution of spins can be preferred with respect to an aligned distribution of spins at 2.4σ confidence level, based on the first three detections and on the candidate event LVT151012.

4˙4. *Hardening.* – If a double BH binary undergoes a number of three-body encounters during its life, we expect that its semi-major axis will shrink as an effect of the encounters. This process is called dynamical *hardening*.

Following [124], we call hard binaries (soft binaries) those binaries with binding energy larger (smaller) than the average kinetic energy of a star in the star cluster. According to Heggie's law [124], hard binaries tend to harden (*i.e.* to become more and more bound) via three-body encounters. In other words, a fraction of the internal energy of a hard

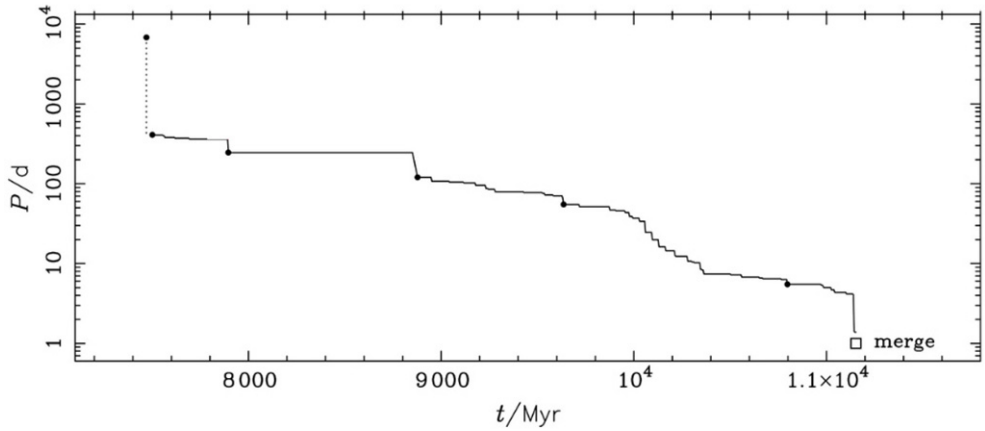


Fig. 8. – Time evolution of the orbital period of a simulated BH binary (from an N-body simulation of a globular cluster). Adapted from fig. 3 of Hurley *et al.* [125].

binary can be transferred into kinetic energy of the intruders and of the centre-of-mass of the binary during three-body encounters. This means that the binary loses internal energy and its semi-major axis shrinks.

Most double BH binaries are expected to be hard binaries, because BHs are among the most massive bodies in star clusters. Thus, double BH binaries are expected to harden as a consequence of three-body encounters. The hardening process may be sufficiently effective to shrink a BH binary till it enters the regime where GW emission is efficient: a BH binary which is initially too loose to merge may then become a GW source thanks to dynamical hardening.

Figure 8 is a clear example of this process. It shows the period evolution of a simulated double BH binary in a globular cluster. The orbital period decreases till the binary eventually merges. The decrease of the orbital period is not smooth, but proceeds by smaller and larger steps, which indicate hardening due to three-body encounters and even few exchanges.

It is even possible to make a simple analytic estimate of the evolution of the semi-major axis of a double BH binary which is affected by three-body encounters and by GW emission (eq. (9) of [14]):

$$\frac{\mathrm{d}a}{\mathrm{d}t} = -2\,\pi\,\xi\,\frac{G\,\rho}{\sigma}\,a^2 - \frac{64}{5}\,\frac{G^3\,m_1\,m_2\,(m_1+m_2)}{c^5\,(1-e^2)^{7/2}}\,a^{-3}, \tag{12}$$

where $\xi \sim 0.2$–1 is a dimensionless parameter [126], ρ is the local mass density of stars, σ is the local velocity dispersion, c is the light speed, e is the eccentricity. The first part of the right-hand term of eq. (12) accounts for the effect of three-body hardening on the semi-major axis. It scales as $\mathrm{d}a/\mathrm{d}t \propto a^2$, indicating that the larger the binary, the more effective the hardening. This can be easily understood considering that the geometric cross section for three-body interactions with a binary scales as a^2.

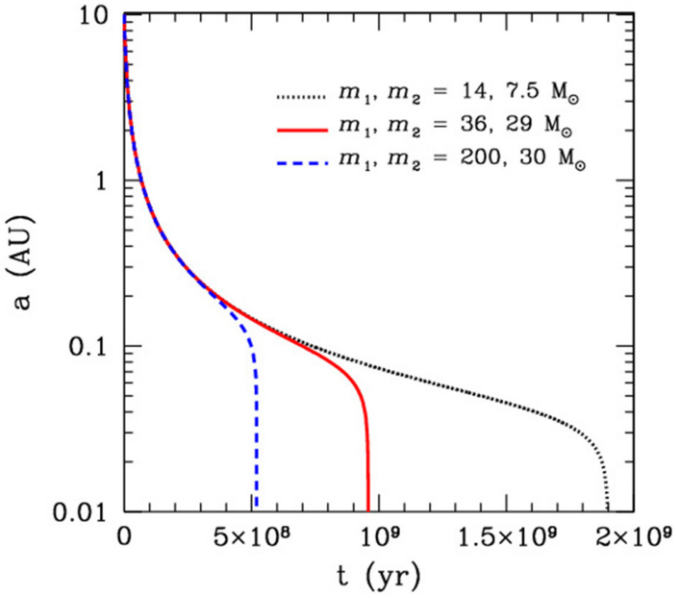


Fig. 9. – Time evolution of the semi-major axis of three BH binaries estimated from eq. (12). Blue dashed line: BH binary with masses $m_1 = 200M_\odot$, $m_2 = 30M_\odot$; red solid line: $m_1 = 36M_\odot$, $m_2 = 29M_\odot$; black dotted line: $m_1 = 14M_\odot$, $m_2 = 7.5M_\odot$. For all BH binaries: $\xi = 1$, $\rho = 10^5 M_\odot\,\mathrm{pc}^{-3}$, $\sigma = 10\,\mathrm{km\ s}^{-1}$, initial semi-major axis of the BH binary $a_\mathrm{i} = 10\,\mathrm{AU}$.

The second part of the right-hand term of eq. (12) accounts for energy loss by GW emission. It is the first-order approximation of the calculation by [127]. It scales as $\mathrm{d}a/\mathrm{d}t \propto a^{-3}$ indicating that GW emission becomes efficient only when the two BHs are very close to each other.

In fig. 9 we solve eq. (12) numerically for three double BH binaries with different mass. All binaries evolve through i) a first phase in which hardening by three-body encounters dominates the evolution of the binary, ii) a second phase in which the semi-major axis stalls because three-body encounters become less efficient as the semi-major axis shrink, but the binary is still too large for GW emission to become efficient, and iii) a third phase in which the semi-major axis drops because the binary enters the regime where GW emission is efficient.

4·5. *Dynamical ejections*. – During three-body encounters, a fraction of the internal energy of a hard binary is transferred into kinetic energy of the intruders and of the centre-of-mass of the binary. As a consequence, the binary recoils. The recoil velocity is generally of the order of few km s^{-1}, but can be up to several hundred km s^{-1}.

Since the escape velocity from a globular cluster is $\sim 30\,\mathrm{km\,s}^{-1}$ and the escape velocity from a young star cluster or an open cluster is even lower, both the recoiling binary and the intruder can be ejected from the parent star cluster. If the binary and/or the intruder

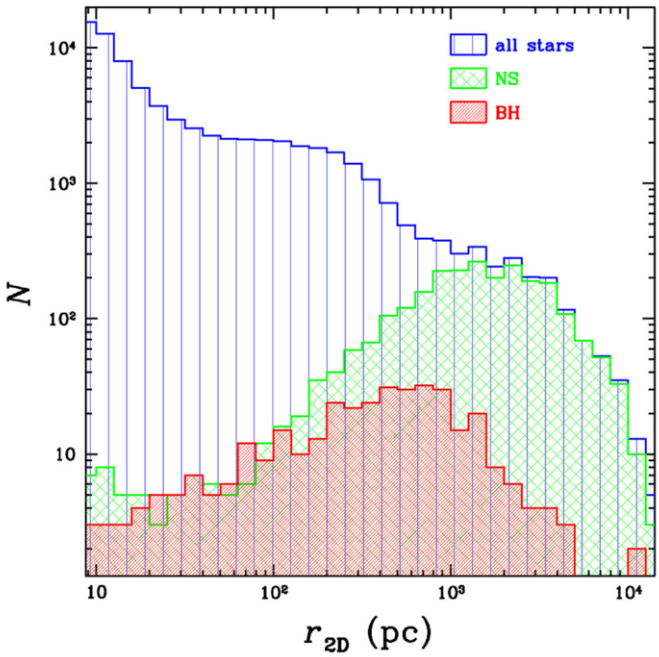


Fig. 10. – Distribution of 2-dimensional distances (r_{2D}) of stars from the centre of the parent star cluster 100 Myr after the beginning of the simulation. This plot combines 300 simulations of young star clusters discussed in [69]. All stars shown in the figure have distance > 10 pc from the centre of the star cluster, indicating that they were ejected. Blue histogram: all ejected stars; green histogram: ejected NSs; red histogram: ejected BHs. NSs are ejected both by SN kick (50 per cent) and by dynamical recoil (50 per cent), while most BHs are ejected by dynamical recoil.

are ejected, they become field objects and cannot participate in the dynamics of the star cluster anymore. Thus, not only the double BH binary stops hardening, but also the intruder, if it is another compact object, loses any chance of entering a new binary by dynamical exchange.

Figure 10 shows that this process is particularly efficient in young star clusters, especially for NSs (which are lighter than BHs). In 100 Myr, about 90 per cent of NSs and 40 per cent of BHs are ejected from the parent cluster, either by dynamical recoil or by SN kick [69, 95].

Dynamical ejections of double NSs and of BH-NS binaries were proposed to be one of the possible explanations for the hostless short gamma-ray bursts, *i.e.* gamma-ray bursts whose position in the sky appears to be outside any observed galaxy [128]. Hostless bursts may be ~ 25 per cent of all short gamma-ray bursts.

In general, ejections of compact objects and compact-object binaries from their parent star cluster can be the result of at least three different processes:

– dynamical ejections;

– SN kicks [60, 129];

– GW recoil [130-132].

GW recoil is a relativistic kick occurring when a BH binary merges. It results in kick velocities up to thousands of km s^{-1} and usually of the order of hundreds of km s^{-1}.

Ejections (by dynamics, SN kick or GW recoil) may be the main process at work against mergers of second-generation BHs, where for second-generation BHs we mean BHs which were born from the merger of two BHs rather than from the collapse of a star [133]. In globular clusters, open clusters and young star clusters, a BH binary has good chances of being ejected by three-body encounters before it merges (see [134] for a detailed calculation) and a very high chance of being ejected by GW recoil after it merges. The only place where merging BHs can easily avoid ejection by GW recoil are nuclear-star clusters, whose escape velocity is hundreds of km s^{-1}.

4˙6. *Formation of intermediate-mass black holes by runaway collisions.* – In sect. **2˙3**, we have mentioned that intermediate-mass black holes (IMBHs, *i.e.* BHs with mass $100 \lesssim m_{\rm BH} \lesssim 10^4 M_\odot$) form from the direct collapse of metal-poor extremely massive stars [70]. Other formation channels have been proposed for IMBHs and most of them involve dynamics of star clusters. The formation of massive BHs by runaway collisions has been originally proposed about half a century ago [135, 136] and was then elaborated by several authors (*e.g.* [103, 137-142]).

The basic idea (see fig. 11 for a cartoon) is that if the most massive stars in a dense young star cluster sink to the centre of the cluster by dynamical friction on a time shorter than their lifetime (*i.e.* before core-collapse SNe take place, removing a large fraction of their mass), then the density of massive stars in the cluster core becomes extremely high. This makes collisions between massive stars extremely likely. Actually, direct N-body simulations show that collisions between massive stars proceed in a runaway sense, leading to the formation of a very massive ($\gg 100 M_\odot$) star [138]. The main open question is: "What is the final mass of the collision product? Is the collision product going to collapse to an IMBH?".

There are essentially two critical issues: i) how much mass is lost during the collisions? ii) how much mass does the very-massive star lose by stellar winds?

Hydrodynamical simulations of colliding stars [143, 144] show that massive star can lose ≈ 25 per cent of their mass during collisions. Even if we optimistically assume that no mass is lost during and immediately after the collision (when the collision product relaxes to a new equilibrium), the resulting very massive star will be strongly radiation pressure dominated and is expected to lose a significant fraction of its mass by stellar winds. Recent studies including the effect of the Eddington factor on mass loss [70, 103] show that IMBHs cannot form from runaway collisions at solar metallicity. At lower metallicity ($Z \lesssim 0.1 Z_\odot$) approximately 10–30 per cent of runaway collision products in

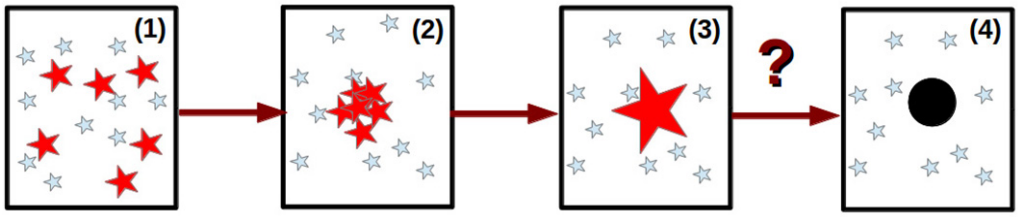


Fig. 11. – Cartoon of the runaway collision scenario in dense young star clusters (see, *e.g.*, [138]). From left to right: (1) the massive stars (red big stars) and the low-mass stars (blue small stars) follow the same initial spatial distribution; (2) dynamical friction leads the massive stars to sink to the core of the cluster, where they start colliding between each other; (3) a very massive star ($\gg 100 M_\odot$) forms as a consequence of the runaway collisions; (4) this massive star might be able to directly collapse into a BH.

young dense star clusters can become IMBHs by direct collapse (they also avoid being disrupted by pair-instability SNe).

The majority of runaway collision products do not become IMBHs but they end up as relatively massive BHs ($\sim$ 20–90$M_\odot$ [103]). If they remain inside their parent star cluster, such massive BHs are extremely efficient in acquiring companions by dynamical exchanges. Mapelli [103] find that all stable binaries formed by the runaway collision product are double compact object binaries and thus are possibly important sources of GWs in the LIGO-Virgo range.

4˙7. *Formation of intermediate-mass black holes by repeated mergers.* – The runaway collision scenario occurs only in the early stages of the evolution of a star cluster. However, it has been proposed that IMBHs form even in old clusters (*e.g.*, globular clusters) by repeated mergers of smaller BHs (*e.g.*, [134, 142]).

The simple idea is illustrated in fig. 12. A stellar BH binary in a star cluster is usually a rather hard binary. Thus, it shrinks by dynamical hardening till it may enter the regime where GW emission is effective. In this case, the BH binary merges leaving a single more massive BH. Given its relatively large mass, the new BH has good chances to acquire a new companion by exchange. Then, the new BH binary starts hardening again by three-body encounters and the story may repeat several times, till the main BH becomes an IMBH.

This scenario has one big advantage: it does not depend on stellar evolution, so we are confident that the BH will grow in mass by mergers, if it remains inside the cluster. However, there are several issues. First, the BH binary may be ejected by dynamical recoils, received as an effect of three-body encounters. Recoils get stronger and stronger, as the orbital separation decreases [134]. The BH binary will avoid ejection by dynamical recoil only if it is sufficiently massive ($\gtrsim 50 M_\odot$ for a dense globular cluster [14]). If the BH binary is ejected, the loop breaks and no IMBH is formed.

Second and even more important, the merger of two BHs involves a relativistic kick. This kick may be as large as hundreds of km s^{-1} [130], leading to the ejection of the BH

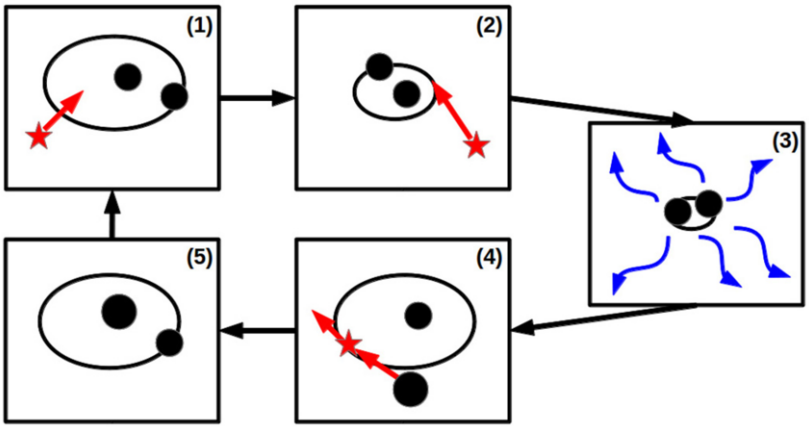


Fig. 12. – Cartoon of the repeated merger scenario in old-star clusters (see, *e.g.*, [134, 142]). From top to bottom and from left to right: (1) a BH binary undergoes three-body encounters in a star cluster; (2) three-body encounters harden the BH binary, shrinking its semi-major axis; (3) the BH binary hardens by three-body encounters till it enters the regime where GW emission is efficient: the binary semi-major axis decays by GW emission and the binary merges; (4) a single bigger BH forms as result of the merger, which may acquire a new companion by dynamical exchange (if it is not ejected by GW recoil); (5) the new binary containing the bigger BH starts shrinking again by three body encounters (1). This loop may be repeated several times till the main BH becomes an IMBH.

from the parent star cluster (unless it is a nuclear star cluster). Also in this case, the loop breaks and no IMBH is formed.

Finally, even if the BH binary is not ejected, this scenario is relatively inefficient: if the seed BH is $\sim 50 M_\odot$, several Gyr are required to form an IMBH with mass $\sim 500 M_\odot$ [134].

Monte Carlo simulations by [142] show that both the runaway collision scenario and the repeated-merger scenario can be at work in star clusters. Figure 13 clearly shows that two classes of IMBHs form in simulations: i) runaway collision IMBHs form in the first few Myr of the life of a star cluster and grow in mass very efficiently; ii) repeated-merger IMBHs start forming much later ($\gtrsim 5\,\mathrm{Gyr}$) and their growth is less efficient. It is important to note that these simulations do not include prescriptions for GW recoils, which might dramatically suppress the formation of IMBHs by repeated mergers [145].

4˙8. *Kozai-Lidov resonance.* – Unlike the other dynamical processes discussed so far, Kozai-Lidov (KL) resonance [146, 147] can occur both in the field and in star clusters. KL resonance appears whenever we have a stable hierarchical triple system (*i.e.* a triple composed of an inner binary and an outer body orbiting the inner binary), in which the orbital plane of the outer body is inclined with respect to the orbital plane of the inner binary. Periodic perturbations induced by the outer body on the inner binary cause i) the eccentricity of the inner binary and ii) the inclination between the orbital plane of the inner binary and that of the outer body to oscillate.

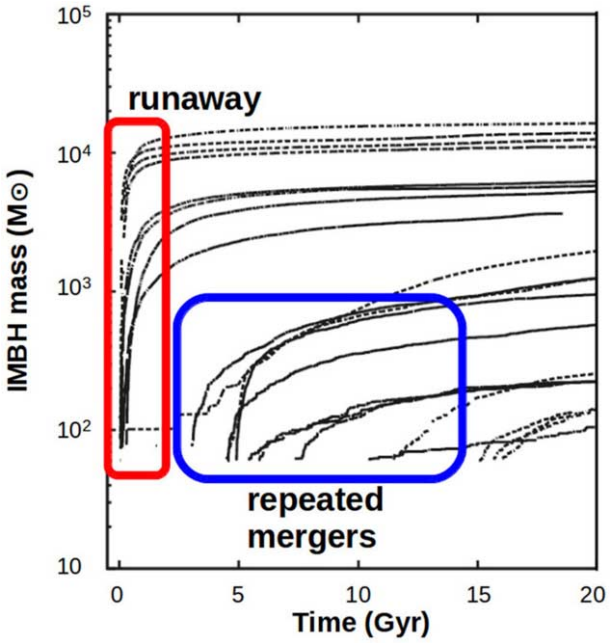


Fig. 13. – IMBH mass growth as a function of time in Monte Carlo simulations of globular clusters [142]. The red (blue) box highlights IMBHs formed by runaway collisions (repeated mergers). Dashed lines: models with reduced mass accretion on to the IMBH and reduced star expansion after merger events. Solid lines: models with the standard prescription for IMBH mass accretion rates and post-merger expansion. See [142] for further details. Adapted from fig. 6 of [142].

Figure 14 is an example of KL oscillations in three-body simulations. It is worth noting that the semi-major axis is not affected, because KL resonance does not imply an energy exchange between inner and outer binary.

KL oscillations may enhance BH binary mergers because the timescale for merger by GW emission strongly depends on the eccentricity e of the binary [127]:

$$t_{\rm GW} = \frac{5}{256}\,\frac{c^5\,a^4\,(1-e^2)^{7/2}}{G^3\,m_1\,m_2\,(m_1+m_2)}\,. \tag{13}$$

It might seem that hierarchical triples are rather exotic systems. This is not the case. In fact, $\sim$ 10 per cent of low-mass stars are in triple systems [148-150]. This fraction gradually increases for more massive stars [151], up to $\sim$ 50 per cent for B-type stars [152-155]. In star clusters, stable hierarchical triple systems may form dynamically, via four-body or multi-body encounters.

Kimpson *et al.* [104] find that KL resonance may enhance the BH merger rate by $\approx$ 40 per cent in young star clusters and open clusters. On the other hand, Antonini *et al.* [156] find that KL resonance in field triples can account for $\lesssim 3$ mergers Gpc^{-3} yr^{-1}.

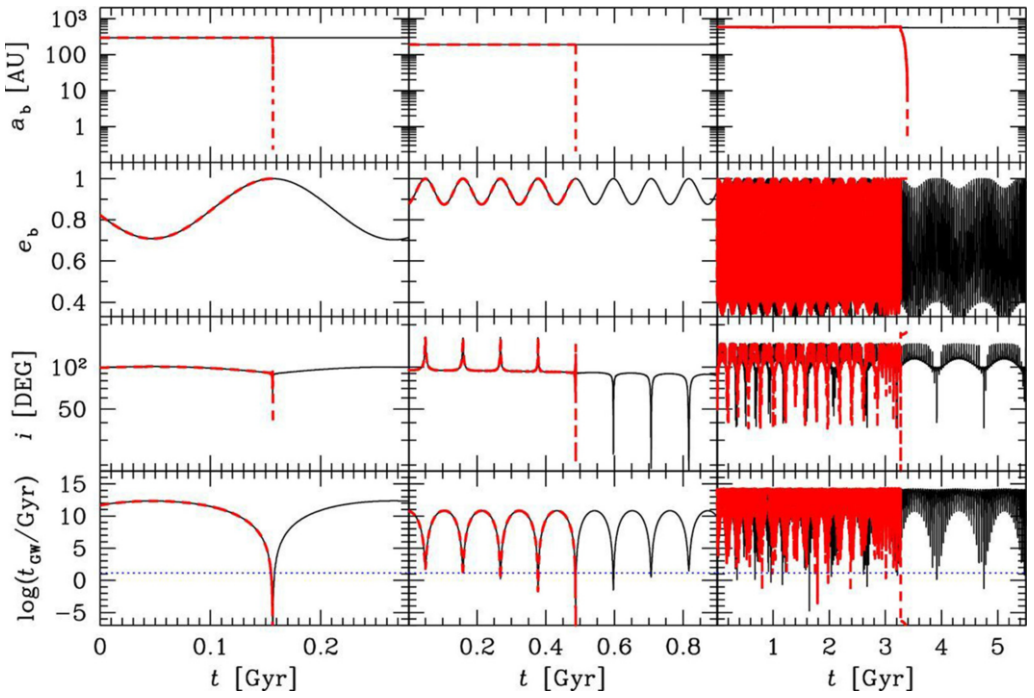


Fig. 14. – Simulation of three hierarchical triples undergoing Kozai-Lidov (KL) resonance. Black solid line: simulation without post-Newtonian terms. Red dashed line: simulation with 2.5 Post-Newtonian term. From top to bottom: time evolution of semi-major axis; time evolution of the eccentricity of the inner binary; time evolution of the inclination between the orbital plane of the inner and outer binary; time for coalescence by GW emission as a function of time, expressed as in [127]. From fig. 5 of [104].

The main signature of the merger of a KL system is the non-zero eccentricity until very few seconds before the merger. Eccentricity might be significantly non-zero even when the system enters the LIGO-Virgo frequency range, as shown in fig. 15.

KL resonances have an intriguing application in nuclear-star clusters. If the stellar BH binary is gravitationally bound to the super-massive BH (SMBH) at the centre of the galaxy, then we have a peculiar triple system where the inner binary is composed of the stellar BH binary and the outer body is the SMBH [157]. Also in this case, the merging BH has good chances of retaining a non-zero eccentricity till it emits GWs in the LIGO-Virgo frequency range.

4˙9. *Summary of dynamics and open issues.* – In this section, we have seen that dynamics is a crucial ingredient to understand BH demography. Dynamical interactions (three- and few-body close encounters) can favour the coalescence of BH binaries through dynamical hardening. New BH binaries can form via dynamical exchanges. Both processes suggest a boost of the BH binary merger rate in a dynamically active environment.

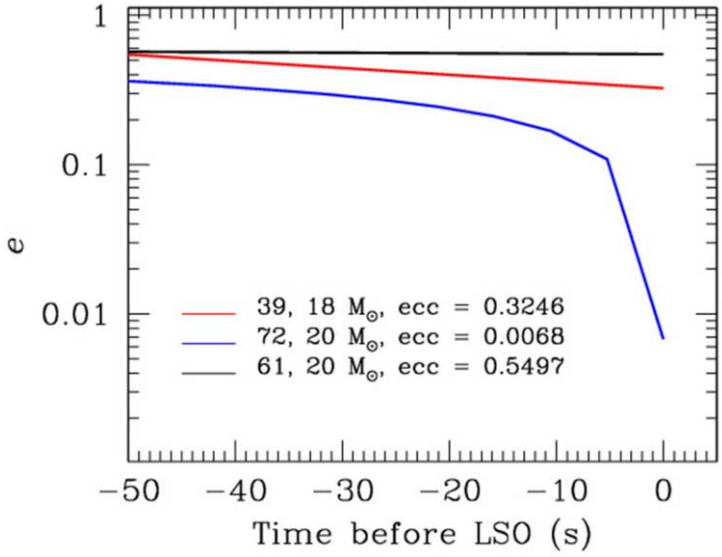


Fig. 15. – Evolution of the eccentricity in the last few seconds before the last stable orbit (LSO) of the three hierarchical triples already shown in fig. 14. The LSO was estimated as $a_{\rm LSO} = 6\,G\,(m_1 + m_2)/c^2$. The eccentricity at the LSO is $e = 0.3246$ for the system with m_1, $m_2 = 39$, $18 M_\odot$, $e = 0.0068$ for m_1, $m_2 = 72$, $20 M_\odot$, and $e = 0.5497$ for m_1, $m_2 = 61$, $20 M_\odot$.

Moreover, exchanges favour the formation of more massive binaries, with higher initial eccentricity and with misaligned spins. Also KL resonances favour the coalescence of more massive binaries and with higher eccentricity, even close to the last stable orbit.

On the other hand, three-body encounters might trigger the ejection of compact-object binaries from their natal environment, inducing a significant displacement between the birth place of the binary and the location of its merger.

Finally, dynamics can lead to the formation of IMBHs, with mass of few hundreds solar masses.

We expect dynamics to be important for BH binaries also because massive stars (which are the progenitors of BHs) form preferentially in young star clusters [93], which are dynamically active places. Despite this, the vast majority of studies of BH demography either do not include dynamics or focus only on old globular clusters. Globular clusters are indeed systems where dynamics is tremendously important for BH binaries, but they represent a small fraction of the stellar mass in the Universe (less than $\sim$ few per cent [101]).

In contrast, young star clusters are the most common birthplace of massive stars, but only few works focus on BH binaries in young star clusters [102,103,105-107]. Even these studies assume overly simplified initial conditions for their simulations. For example, they do not include gas, which is a fundamental component of young star clusters. The main reason for this "omission" is purely numerical. Old massive globular clusters are relaxed

(nearly) spherically symmetric structures with no gas. For most of their life, they host only low-mass stars and compact remnants. They can be evolved with Monte Carlo codes (*e.g.* MOCCA [158]), which are considerably faster than direct N-body codes. In contrast, young star clusters are asymmetric, not yet relaxed, rich of gas and of substructures. Direct N-body codes, which scale as N^2, are required to model young star clusters. Adding gas to the general picture requires to combine expensive direct N-body simulations with hydrodynamical simulations. Only few authors attempted to model young star clusters with both direct N-body and hydrodynamics, in most cases simply by dividing the integration in the two separate parts: first the hydrodynamics and then the direct N-body dynamics, after most gas has been converted into stars [108-115, 159]. Last but not least, young star clusters host short-lived massive stars. Thus, stellar evolution processes are very fast in young star clusters and they have a dramatic impact on the (hydro)dynamics of the system. A heroic theoretical effort is needed to properly model young star clusters and their impact on the demography of BH binaries.

5. – Black hole binaries in cosmological context

Population synthesis codes model BH binaries as isolated systems, with essentially no information on the environment. Even dynamical simulations are restricted to a limited environment: the parent star cluster. On the other hand, BH binaries merging within the LIGO-Virgo instrumental horizon (typically at redshift $z \sim 0.1$–0.2) might have formed at much higher redshift. Moreover, third-generation ground-based gravitational-wave detectors (*e.g.*, Einstein Telescope [160]) will be able to observe merging BH binaries up to redshift $z \approx 10$. Thus, the cosmological framework in which BH binaries and their progenitors evolve cannot be neglected.

Accounting for cosmology in models of stellar-size BH binaries appears as a desperate challenge, because of the humongous dynamical range: orbital separation of BH binaries are of the order of few astronomical units, while cosmic structures are several hundreds of Mpc. Few theoretical studies have faced this challenge, adopting two different procedures.

5˙1. *Analytic prescriptions*. – Some authors (*e.g.*, [161-164]) combine the outputs of population synthesis codes with analytic prescriptions. The main ingredients are the cosmic star formation rate density and the average evolution of metallicity with redshift [165]. In some previous work (*e.g.*, [161, 164]) a Press-Schechter–like formalism is adopted, to include the mass of the host galaxy in the general picture. Lamberts *et al.* [164] even include a redshift-dependent description for the mass-metallicity relation (hereafter MZR), to account for the fact that the mass of a galaxy and its observed metallicity are deeply connected. The main advantage of this procedure is that the star formation rate and the metallicity evolution can be derived more straightforwardly from the data. The main drawback is that it is extremely difficult, if not impossible, to trace the evolution of the host galaxy of the BH binary, through its galaxy merger tree.

5˙2. *Cosmological simulations*. – The alternative approach feeds the outputs of population synthesis simulations into cosmological simulations [86, 166, 167], through a Monte

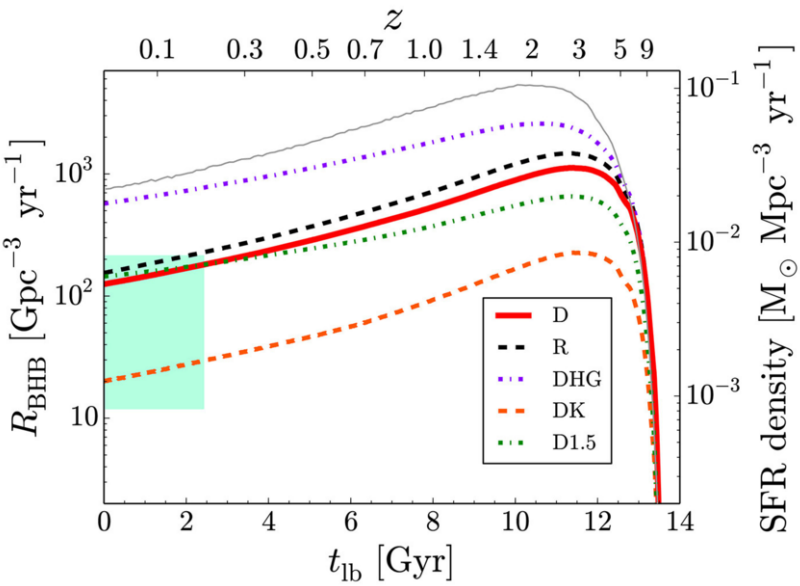


Fig. 16. – Cosmic merger rate density of BH binaries in the comoving frame ($R_{\rm BHB}$) as a function of the look-back time ($t_{\rm lb}$, bottom x-axis) and of the redshift (z, top x-axis) for several population synthesis models [81] interfaced with the Illustris simulation. Red solid line: model D (delayed SN model [60], CE parameters $\alpha\lambda = 0.1$, fiducial distribution for the natal kicks of the BHs). Black dashed line: model R, same as model D but for a rapid SN model [60]. Green dot-dashed line: model D1.5 (same as model D but for $\alpha\lambda = 1.5$). Violet dot-dashed line: model DHG (same as model D but assuming that Hertzsprung gap donors can survive a CE phase). Orange dashed line: model DK (same as model D but with maximum BH natal kicks). Right-hand y-axis: cosmic star formation rate density in the Illustris, shown as a thin grey line. From fig. 1 of [86].

Carlo approach. This has the clear advantage that the properties of the host galaxies can be easily reconstructed across cosmic time. However, the ideal thing would be to have a high-resolution cosmological simulation (sufficient to resolve also small dwarf galaxies) with a box as large as the instrumental horizon of the GW detectors. This is obviously impossible. High-resolution simulations have usually a box of few comoving Mpc^3, while simulations with a larger box cannot resolve dwarf galaxies. Moreover, this procedure requires to use the cosmic star formation rate density and the redshift-dependent MZR which are intrinsic to the cosmological simulations. While most state-of-the-art cosmological simulations reproduce the cosmic star formation rate density reasonably well, the MZR is an elusive feature, creating more than a trouble even in the most advanced cosmological simulations.

Recently, [86] combined their population-synthesis simulations with the Illustris cosmological box [168-170]. The size of the Illustris (length = 106.5 comoving Mpc) is sufficient to satisfy the cosmological principle, but galaxies with stellar mass $\lesssim 10^8 M_\odot$

are heavily under-resolved. The Illustris matches the cosmic star formation rate density and recover a MZR. However, the simulated MZR is sensibly different from the observed relation: the curve is relatively steeper at low metallicity and there is no flattening at high mass [171, 172]. Uncertainties connected with the MZR are estimated to affect the BH merger rate by $\sim$ 20 per cent. Keeping these *caveats* in mind [115] are able to reconstruct the BH merger rate across cosmic time.

Figure 16 shows the cosmic merger rate of BH binaries in the comoving frame obtained by [86] for several different population-synthesis models (the assumptions for CE treatment, natal kick distribution and SN recipes are changed from model to model). In all considered models, the cosmic BH merger rate has a peak at relatively large redshift ($z \sim 2$–4) and then decreases slowly down to the current time. The trend of the cosmic merger rate is quite similar to the trend of the cosmic star formation rate density curve.

Authors of ref. [167] adopt a complementary approach to study the importance of dwarf galaxies for GW detections. They use the GAMESH pipeline to produce a high-resolution simulation of the Local Group (length = 4 Mpc comoving). This means that the considered portion of the Universe is strongly biased, but the resolution is sufficient to investigate BH binaries in small ($> 10^6 M_\odot$) dwarf galaxies. One of their main conclusions is that GW150914-like events originate mostly from small metal-poor galaxies.

The approach of comparing the results of large-box cosmological simulations with small-box high-resolution simulations is surely crucial to understand how BH binaries populate the galaxies across cosmic time.

6. – Summary and outlook

We reviewed our current understanding of the astrophysics of stellar-mass BHs. The era of gravitational-wave astrophysics has just begun and has already produced two formidable results: BH binaries exist and can host BHs with mass $> 30 M_\odot$ [1, 7].

According to nowadays stellar evolution and supernova theories, such massive BHs can form only from massive relatively metal-poor stars. At low metallicity, stellar winds are quenched and stars end their life with a larger mass than their metal-rich analogues. If the final mass and the final core mass are sufficiently large, these stars can directly collapse into BHs, producing BHs with mass $\gtrsim 30 M_\odot$ [26, 27]. An alternative scenario predicts that $\sim$ 30–40$M_\odot$ BHs are the result of gravitational instabilities in the very early Universe (primordial BHs, *e.g.* [17]).

The formation channels of merging BH binaries are still an open question. All proposed scenarios have several drawbacks and uncertainties. While mass transfer and common envelope are a major issue in the isolated binary evolution scenario, even the dynamical evolution is still effected by major issues (*e.g.* the small statistics about BHs in young star clusters, and the major simplifications adopted in dynamical simulations).

Finally, a global picture is missing, which combines stellar and binary evolution with dynamics and cosmology, aimed at reconstructing the BH merger history across cosmic time. This is crucial for the astrophysical interpretation of LIGO-Virgo data and for meeting the challenge of third-generation ground-based GW detectors.

* * *

I thank the organizers and the participants of the International School of Physics "Enrico Fermi" (Course 200 - Gravitational Waves and Cosmology) for the enlightening discussions and for giving me the opportunity to think about these lectures. Numerical calculations have been performed through a CINECA-INFN agreement and through a CINECA-INAF agreement (Accordo Quadro INAF-CINECA 2017), providing access to resources on GALILEO and MARCONI at CINECA. MM acknowledges financial support from the Italian Ministry of Education, University and Research (MIUR) through grant FIRB 2012 RBFR12PM1F, from INAF through grant PRIN-2014-14, and from the MERAC Foundation.

REFERENCES

[1] ABBOTT B. P., ABBOTT R., ABBOTT T. D. *et al.*, *Phys. Rev. Lett.*, **116** (2016) 061102.
[2] ABBOTT B. P., ABBOTT R., ABBOTT T. D. *et al.*, *Phys. Rev. Lett.*, **116** (2016) 241103.
[3] ABBOTT B. P., ABBOTT R., ABBOTT T. D. *et al.*, *Phys. Rev. Lett.*, **118** (2017) 221101.
[4] The LIGO Scientific Collaboration, the Virgo Collaboration, ABBOTT B. P. *et al.*, arXiv:1711.05578 (2017).
[5] The LIGO Scientific Collaboration, the Virgo Collaboration, ABBOTT B. P. *et al.*,, arXiv:1709.09660 (2017).
[6] ABBOTT B. P., ABBOTT R., ABBOTT T. D. *et al.*, *Phys. Rev. X*, **6** (2016) 041015.
[7] ABBOTT B. P., ABBOTT R., ABBOTT T. D. *et al.*, *Astrophys. J. Lett.*, **818** (2016) L22.
[8] TUTUKOV A. and YUNGELSON L., *Nauch. Inform.*, **27** (1973) 70.
[9] THORNE K. S., in *Three Hundred Years of Gravitation*, edited by HAWKING S. and ISRAEL W. (Cambridge University Press, UK) 1987, p. 330.
[10] SCHUTZ B. F., *Class. Quantum Grav.*, **6** (1989) 1761.
[11] KULKARNI S. R., HUT P. and MCMILLAN S., *Nature*, **364** (1993) 421.
[12] SIGURDSSON S. and HERNQUIST L., *Nature*, **364** (1993) 423.
[13] PORTEGIES ZWART S. F. and MCMILLAN S. L. W., *Astrophys. J. Lett.*, **528** (2000) L17.
[14] COLPI M., MAPELLI M. and POSSENTI A., *Astrophys. J.*, **599** (2003) 1260.
[15] BELCZYNSKI K., SADOWSKI A. and RASIO F. A., *Astrophys. J.*, **611** (2004) 1068.
[16] BIRD S., CHOLIS I., MUÑOZ J. B. *et al.*, *Phys. Rev. Lett.*, **116** (2016) 201301.
[17] CARR B., KÜHNEL F. and SANDSTAD M., *Phys. Rev. Doc.*, **94** (2016) 083504.
[18] INOMATA K., KAWASAKI M., MUKAIDA K., TADA Y. and YANAGIDA T. T., *Phys. Rev. Doc.*, **95** (2017) 123510.
[19] VAN LOON J. T., CIONI M.-R. L., ZIJLSTRA A. A. and LOUP C., *Astron. Astrophys.*, **438** (2005) 273.
[20] OROSZ J. A., "A Massive Star Odyssey: From Main Sequence to Supernova", in *Proceedings of the IAU Symposium*, Vol. **212** (Astronomical Society of the Pacific) 2003, p. 365.
[21] ÖZEL F., PSALTIS D., NARAYAN R. and MCCLINTOCK J. E., *Astrophys. J.*, **725** (2010) 1918.
[22] LAYCOCK S. G. T., MACCARONE T. J. and CHRISTODOULOU D. M., *Mon. Not. R. Astron. Soc.*, **452** (2015) L31.
[23] BRESOLIN F. and KUDRITZKI R. P., *Origin and Evolution of the Elements*, **283** (2004) 283.
[24] FRYER C. L., *Astrophys. J.*, **522** (1999) 413.

[25] FRYER C. L. and KALOGERA V., *Astrophys. J.*, **554** (2001) 548.
[26] MAPELLI M., COLPI M. and ZAMPIERI L., *Mon. Not. R. Astron. Soc.*, **395** (2009) L71.
[27] BELCZYNSKI K., BULIK T., FRYER C. L. *et al.*, *Astrophys. J.*, **714** (2010) 1217.
[28] ABBOTT D. C., *Astrophys. J.*, **259** (1982) 282.
[29] KUDRITZKI R. P., PAULDRACH A. and PULS J., *Astron. Astrophys.*, **173** (1987) 293.
[30] LEITHERER C., ROBERT C. and DRISSEN L., *Astrophys. J.*, **401** (1992) 596.
[31] VINK J. S., DE KOTER A. and LAMERS H. J. G. L. M., *Astron. Astrophys.*, **369** (2001) 574.
[32] VINK J. S. and DE KOTER A., *Astron. Astrophys.*, **442** (2005) 587.
[33] GRÄFENER G. and HAMANN W.-R., *Astron. Astrophys.*, **482** (2008) 945.
[34] GRÄFENER G., VINK J. S., DE KOTER A. and LANGER N., *Astron. Astrophys.*, **535** (2011) A56.
[35] VINK J. S., MUIJRES L. E., ANTHONISSE B. *et al.*, *Astron. Astrophys.*, **531** (2011) A132.
[36] VINK J. S., *Philos. Trans. R. Soc. London Ser. A*, **375** (2017) 20160269.
[37] TANG J., BRESSAN A., ROSENFIELD P. *et al.*, *Mon. Not. R. Astron. Soc.*, **445** (2014) 4287.
[38] CHEN Y., BRESSAN A., GIRARDI L. *et al.*, *Mon. Not. R. Astron. Soc.*, **452** (2015) 1068.
[39] SPERA M., MAPELLI M. and BRESSAN A., *Mon. Not. R. Astron. Soc.*, **451** (2015) 4086.
[40] GEORGY C., MEYNET G., EKSTRÖM S. *et al.*, *Astron. Astrophys.*, **599** (2017) L5.
[41] KESZTHELYI Z., PULS J. and WADE G. A., *Astron. Astrophys.*, **598** (2017) A4.
[42] PETIT V., KESZTHELYI Z., MACINNIS R. *et al.*, *Mon. Not. R. Astron. Soc.*, **466** (2017) 1052.
[43] WADE G. A., NEINER C., ALECIAN E. *et al.*, *Mon. Not. R. Astron. Soc.*, **456** (2016) 2.
[44] MAEDER A., *Physics, Formation and Evolution of Rotating Stars, Astronomy and Astrophysics Library* (Springer Berlin Heidelberg) 2009.
[45] CHIEFFI A. and LIMONGI M., *Astrophys. J.*, **764** (2013) 21.
[46] LIMONGI M., arXiv:1706.01913 (2017).
[47] CHANDRASEKHAR S., *Astrophys. J.*, **74** (1931) 81.
[48] COLGATE S. A. and WHITE R. H., *Astrophys. J.*, **143** (1966) 626.
[49] BETHE H. A. and WILSON J. R., *Astrophys. J.*, **295** (1985) 14.
[50] JANKA H.-T., *Annu. Rev. Nucl. Part. Sci.*, **62** (2012) 407.
[51] FRYER C., 2014, *Proceedings of Frontier Research in Astrophysics (FRAPWS2014)* held 26-31 May, 2014 in Mondello (Palermo), Italy. Online at `<A href=‘‘http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=237’’>http://pos.sissa.it/cgi-bin/reader/conf.cgi?confid=237</A>, id.4, 4`.
[52] FOGLIZZO T., KAZERONI R., GUILET J. *et al.*, *Publ. Astron. Soc. Australia*, **32** (2015) e009.
[53] ERTL T., JANKA H.-T., WOOSLEY S. E., SUKHBOLD T. and UGLIANO M., *Astrophys. J.*, **818** (2016) 124.
[54] MÜLLER B., JANKA H.-T. and MAREK A., *Astrophys. J.*, **756** (2012) 84.
[55] MÜLLER B., JANKA H.-T. and HEGER A., *Astrophys. J.*, **761** (2012) 72.
[56] O'CONNOR E. and OTT C. D., *Astrophys. J.*, **730** (2011) 70.
[57] UGLIANO M., JANKA H.-T., MAREK A. and ARCONES A., *Astrophys. J.*, **757** (2012) 69.
[58] HORIUCHI S., NAKAMURA K., TAKIWAKI T., KOTAKE K. and TANAKA M., *Mon. Not. R. Astron. Soc.*, **445** (2014) L99.
[59] PAXTON B., MARCHANT P., SCHWAB J. *et al.*, *Astrophys. J. Suppl.*, **220** (2015) 15.
[60] FRYER C. L., BELCZYNSKI K., WIKTOROWICZ G. *et al.*, *Astrophys. J.*, **749** (2012) 91.
[61] FRYER C. L., *New Astron.*, **50** (2006) 492.
[62] FOWLER W. A. and HOYLE F., *Astrophys. J. Suppl.*, **9** (1964) 201.
[63] BARKAT Z., RAKAVY G. and SACK N., *Phys. Rev. Lett.*, **18** (1967) 379.

[64] RAKAVY G. and SHAVIV G., *Astrophys. J.*, **148** (1967) 803.
[65] WOOSLEY S. E., *Astrophys. J.*, **836** (2017) 244.
[66] HEGER A. and WOOSLEY S. E., *Astrophys. J.*, **567** (2002) 532.
[67] HEGER A., WOOSLEY S. E., FRYER C. L. and LANGER N., *From Twilight to Highlight: The Physics of Supernovae, ESO Astrophysics Symposia*, edited by W. HILLEBRANDT and B. LEIBUNDGUT (Springer-Verlag) 2003, p. 3.
[68] MAPELLI M., RIPAMONTI E., ZAMPIERI L., COLPI M. and BRESSAN A., *Mon. Not. R. Astron. Soc.*, **408** (2010) 234.
[69] MAPELLI M., ZAMPIERI L., RIPAMONTI E. and BRESSAN A., *Mon. Not. R. Astron. Soc.*, **429** (2013) 2298.
[70] SPERA M. and MAPELLI M., *Mon. Not. R. Astron. Soc.*, **470** (2017) 4739.
[71] BONDI H. and HOYLE F., *Mon. Not. R. Astron. Soc.*, **104** (1944) 273.
[72] HURLEY J. R., TOUT C. A. and POLS O. R., *Mon. Not. R. Astron. Soc.*, **329** (2002) 897.
[73] EGGLETON P. P., *Astrophys. J.*, **268** (1983) 368.
[74] WEBBINK R. F., *Interact. Binary Stars*, **39** (1985) 39.
[75] PORTEGIES ZWART S. F. and VERBUNT F., *Astron. Astrophys.*, **309** (1996) 179.
[76] TOUT C. A., AARSETH S. J., POLS O. R. and EGGLETON P. P., *Mon. Not. R. Astron. Soc.*, **291** (1997) 732.
[77] WEBBINK R. F., *Astrophys. J.*, **277** (1984) 355.
[78] IVANOVA N., JUSTHAM S., CHEN X. *et al.*, *Astron. Astrophys. Rev.*, **21** (2013) 59.
[79] XU X.-J. and LI X.-D., *Astrophys. J.*, **716** (2010) 114.
[80] LOVERIDGE A. J., VAN DER SLUYS M. V. and KALOGERA V., *Astrophys. J.*, **743** (2011) 49.
[81] GIACOBBO N., MAPELLI M. and SPERA M., *Mon. Not. R. Astron. Soc.*, **474** (2018) 2959.
[82] RICKER P. M. and TAAM R. E., *Astrophys. J. Lett.*, **672** (2008) L41.
[83] RICKER P. M. and TAAM R. E., *Astrophys. J.*, **746** (2012) 74.
[84] PASSY J.-C., DE MARCO O., FRYER C. L. *et al.*, *Astrophys. J.*, **744** (2012) 52.
[85] OHLMANN S. T., RÖPKE F. K., PAKMOR R. and SPRINGEL V., *Astrophys. J. Lett.*, **816** (2016) L9.
[86] MAPELLI M., GIACOBBO N., RIPAMONTI E. and SPERA M., *Mon. Not. R. Astron. Soc.*, **472** (2017) 2422.
[87] DE MINK S. E. and MANDEL I., *Mon. Not. R. Astron. Soc.*, **460** (2016) 3545.
[88] MANDEL I. and DE MINK S. E., *Mon. Not. R. Astron. Soc.*, **458** (2016) 2634.
[89] MARCHANT P., LANGER N., PODSIADLOWSKI P., TAURIS T. M. and MORIYA T. J., *Astron. Astrophys.*, **588** (2016) A50.
[90] LADA C. J. and LADA E. A., *Annu. Rev. Astron. Astrophys.*, **41** (2003) 57.
[91] WEIDNER C. and KROUPA P., *Mon. Not. R. Astron. Soc.*, **365** (2006) 1333.
[92] WEIDNER C., KROUPA P. and BONNELL I. A. D., *Mon. Not. R. Astron. Soc.*, **401** (2010) 275.
[93] PORTEGIES ZWART S. F., MCMILLAN S. L. W. and GIELES M., *Annu. Rev. Astron. Astrophys.*, **48** (2010) 431.
[94] DOWNING J. M. B., BENACQUISTA M. J., GIERSZ M. and SPURZEM R., *Mon. Not. R. Astron. Soc.*, **407** (2010) 1946.
[95] DOWNING J. M. B., BENACQUISTA M. J., GIERSZ M. and SPURZEM R., *Mon. Not. R. Astron. Soc.*, **416** (2011) 133.
[96] BENACQUISTA M. J. and DOWNING J. M. B., *Living Rev. Relativ.*, **16** (2013) 4.
[97] TANIKAWA A., *Mon. Not. R. Astron. Soc.*, **435** (2013) 1358.
[98] RODRIGUEZ C. L., MORSCHER M., PATTABIRAMAN B. *et al.*, *Phys. Rev. Lett.*, **115** (2015) 051101.

[99] RODRIGUEZ C. L., CHATTERJEE S. and RASIO F. A., *Phys. Rev. Doc.*, **93** (2016) 084029.
[100] ASKAR A., SZKUDLAREK M., GONDEK-ROSIŃSKA D., GIERSZ M. and BULIK T., *Mon. Not. R. Astron. Soc.*, **464** (2017) L36.
[101] HARRIS W. E., HARRIS G. L. H. and ALESSI M., *Astrophys. J.*, **772** (2013) 82.
[102] ZIOSI B. M., MAPELLI M., BRANCHESI M. and TORMEN G., *Mon. Not. R. Astron. Soc.*, **441** (2014) 3703.
[103] MAPELLI M., *Mon. Not. R. Astron. Soc.*, **459** (2016) 3432.
[104] KIMPSON T. O., SPERA M., MAPELLI M. and ZIOSI B. M., *Mon. Not. R. Astron. Soc.*, **463** (2016) 2443.
[105] BANERJEE S., *Mon. Not. R. Astron. Soc.*, **467** (2017) 524.
[106] BANERJEE S., arXiv:1707.00922 (2017).
[107] FUJII M., TANIKAWA A. and MAKINO J., arXiv:1709.02058 (2017).
[108] MOECKEL N. and BATE M. R., *Mon. Not. R. Astron. Soc.*, **404** (2010) 721.
[109] FUJII M. S. and PORTEGIES ZWART S., *Mon. Not. R. Astron. Soc.*, **449** (2015) 726.
[110] FUJII M. S. and PORTEGIES ZWART S., *Astrophys. J.*, **817** (2016) 4.
[111] PARKER R. J. and DALE J. E., *Mon. Not. R. Astron. Soc.*, **432** (2013) 986.
[112] PARKER R. J., DALE J. E. and ERCOLANO B., *Mon. Not. R. Astron. Soc.*, **446** (2015) 4278.
[113] PARKER R. J. and DALE J. E., *Mon. Not. R. Astron. Soc.*, **451** (2015) 3664.
[114] PARKER R. J. and DALE J. E., *Mon. Not. R. Astron. Soc.*, **470** (2017) 390.
[115] MAPELLI M., *Mon. Not. R. Astron. Soc.*, **467** (2017) 3255.
[116] BÖKER T., LAINE S., VAN DER MAREL R. P. *et al.*, *Astronom. J.*, **123** (2002) 1389.
[117] FERRARESE L., CÔTÉ P., DALLA BONTÀ E. *et al.*, *Astrophys. J. Lett.*, **644** (2006) L21.
[118] GRAHAM A. W. and SPITLER L. R., *Mon. Not. R. Astron. Soc.*, **397** (2009) 2148.
[119] HILLS J. G. and FULLERTON L. W., *Astronom. J.*, **85** (1980) 1281.
[120] ZEVIN M., PANKOW C., RODRIGUEZ C. L. *et al.*, *Astrophys. J.*, **846** (2017) 82.
[121] FARR W. M., STEVENSON S., MILLER M. C. *et al.*, *Nature*, **548** (2017) 426.
[122] FARR B., HOLZ D. E. and FARR W. M., arXiv:1709.07896 (2017).
[123] MILLER M. C. and MILLER J. M., *Phys. Rep.*, **548** (2015) 1.
[124] HEGGIE D. C., *Mon. Not. R. Astron. Soc.*, **173** (1975) 729.
[125] HURLEY J. R., SIPPEL A. C., TOUT C. A. and AARSETH S. J., *Publ. Astron. Soc. Australia*, **33** (2016) e036.
[126] HILLS J. G., *Astronom. J.*, **88** (1983) 1269.
[127] PETERS P. C., *Phys. Rev.*, **136** (1964) 1224.
[128] FONG W., BERGER E., CHORNOCK R. *et al.*, *Astrophys. J.*, **769** (2013) 56.
[129] HOBBS G., LORIMER D. R., LYNE A. G. and KRAMER M., *Mon. Not. R. Astron. Soc.*, **360** (2005) 974.
[130] LOUSTO C. O. and ZLOCHOWER Y., *Phys. Rev. Doc.*, **79** (2009) 064018.
[131] CAMPANELLI M., LOUSTO C., ZLOCHOWER Y. and MERRITT D., *Astrophys. J. Lett.*, **659** (2007) L5.
[132] GONZÁLEZ J. A., SPERHAKE U., BRÜGMANN B., HANNAM M. and HUSA S., *Phys. Rev. Lett.*, **98** (2007) 091101.
[133] GEROSA D. and BERTI E., *Phys. Rev. Doc.*, **95** (2017) 124046.
[134] MILLER M. C. and HAMILTON D. P., *Mon. Not. R. Astron. Soc.*, **330** (2002) 232.
[135] COLGATE S. A., *Astrophys. J.*, **150** (1967) 163.
[136] SANDERS R. H., *Astrophys. J.*, **162** (1970) 791.
[137] PORTEGIES ZWART S. F., MAKINO J., MCMILLAN S. L. W. and HUT P., *Astron. Astrophys.*, **348** (1999) 117.
[138] PORTEGIES ZWART S. F. and MCMILLAN S. L. W., *Astrophys. J.*, **576** (2002) 899.

[139] PORTEGIES ZWART S. F., BAUMGARDT H., HUT P., MAKINO J. and MCMILLAN S. L. W., *Nature*, **428** (2004) 724.
[140] GÜRKAN M. A., FREGEAU J. M. and RASIO F. A., *Astrophys. J. Lett.*, **640** (2006) L39.
[141] FREITAG M., GÜRKAN M. A. and RASIO F. A., *Mon. Not. R. Astron. Soc.*, **368** (2006) 141.
[142] GIERSZ M., LEIGH N., HYPKI A., LÜTZGENDORF N. and ASKAR A., *Mon. Not. R. Astron. Soc.*, **454** (2015) 3150.
[143] GABUROV E., LOMBARDI J. C. and PORTEGIES ZWART S., *Mon. Not. R. Astron. Soc.*, **383** (2008) L5.
[144] GABUROV E., LOMBARDI J. C., JR. and PORTEGIES ZWART S., *Mon. Not. R. Astron. Soc.*, **402** (2010) 105.
[145] HOLLEY-BOCKELMANN K., GÜLTEKIN K., SHOEMAKER D. and YUNES N., *Astrophys. J.*, **686** (2008) 829.
[146] KOZAI Y., *Astronom. J.*, **67** (1962) 591.
[147] LIDOV M. L., *Plan. Space Sci.*, **9** (1962) 719.
[148] TOKOVININ A., *Mon. Not. R. Astron. Soc.*, **389** (2008) 925.
[149] TOKOVININ A., *Astronom. J.*, **147** (2014) 87.
[150] RAGHAVAN D., MCALISTER H. A., HENRY T. J. *et al.*, *Astrophys. J. Suppl.*, **190** (2010) 1.
[151] DUCHÊNE G. and KRAUS A., *Annu. Rev. Astron. Astrophys.*, **51** (2013) 269.
[152] REMAGE EVANS N., *Bull. Soc. R. Sci. Liege*, **80** (2011) 663.
[153] SANA H., LE BOUQUIN J.-B., LACOUR S. *et al.*, *Astrophys. J. Suppl.*, **215** (2014) 15.
[154] MOE M. and DI STEFANO R., *Astrophys. J. Suppl.*, **230** (2017) 15.
[155] TOONEN S., HAMERS A. and PORTEGIES ZWART S., *Comput. Astrophys. Cosmol.*, **3** (2016) 6.
[156] ANTONINI F., TOONEN S. and HAMERS A. S., *Astrophys. J.*, **841** (2017) 77.
[157] ANTONINI F. and PERETS H. B., *Astrophys. J.*, **757** (2012) 27.
[158] HYPKI A. and GIERSZ M., *Mon. Not. R. Astron. Soc.*, **429** (2013) 1221.
[159] PELUPESSY F. I. and PORTEGIES ZWART S., *Mon. Not. R. Astron. Soc.*, **420** (2012) 1503.
[160] PUNTURO M., ABERNATHY M., ACERNESE F. *et al.*, *Class. Quantum Grav.*, **27** (2010) 194002.
[161] DOMINIK M., BELCZYNSKI K., FRYER C. *et al.*, *Astrophys. J.*, **779** (2013) 72.
[162] DOMINIK M., BERTI E., O'SHAUGHNESSY R. *et al.*, *Astrophys. J.*, **806** (2015) 263.
[163] BELCZYNSKI K., HOLZ D. E., BULIK T. and O'SHAUGHNESSY R., *Nature*, **534** (2016) 512.
[164] LAMBERTS A., GARRISON-KIMMEL S., CLAUSEN D. R. and HOPKINS P. F., *Mon. Not. R. Astron. Soc.*, **463** (2016) L31.
[165] MADAU P. and DICKINSON M., *Annu. Rev. Astron. Astrophys.*, **52** (2014) 415.
[166] O'SHAUGHNESSY R., BELLOVARY J. M., BROOKS A. *et al.*, *Mon. Not. R. Astron. Soc.*, **464** (2017) 2831.
[167] SCHNEIDER R., GRAZIANI L., MARASSI S. *et al.*, *Mon. Not. R. Astron. Soc.*, **471** (2017) L105.
[168] VOGELSBERGER M., GENEL S., SPRINGEL V. *et al.*, *Nature*, **509** (2014) 177.
[169] VOGELSBERGER M., GENEL S., SPRINGEL V. *et al.*, *Mon. Not. R. Astron. Soc.*, **444** (2014) 1518.
[170] NELSON D., PILLEPICH A., GENEL S. *et al.*, *Astron. Comput.*, **13** (2015) 12.
[171] VOGELSBERGER M., GENEL S., SIJACKI D. *et al.*, *Mon. Not. R. Astron. Soc.*, **436** (2013) 3031.
[172] TORREY P., VOGELSBERGER M., GENEL S. *et al.*, *Mon. Not. R. Astron. Soc.*, **438** (2014) 1985.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200006

GW170817: the dawn of multi-messenger astronomy inlcuding gravitational waves(*)

Marica Branchesi(**)

Gran Sasso Science Institute - viale F. Crispi 7, L'Aquila, Italy

INFN-Laboratori Nazionali del Gran Sasso - Assergi (AQ), Italy

Summary. — On 2017 August 17 the Advanced LIGO and Advanced Virgo detectors detected for the first time the signal, GW170817, from the coalescence of a binary system of neutron stars (Abbott B. P. *et al.*, *Phys. Rev. Lett.*, **119** (2017) 161101). Exactly 1.7 s after the merger time (12:41:04 UTC) the Fermi Gamma-ray Burst Monitor independently detected a gamma-ray burst, GRB170817A (Abbott B. P. *et al.*, *Astrophys. J. Lett.*, **848** (2017) L13; L14). An extensive observing campaign involving more than 70 world-wide ground and space observatories was performed leading to the discovery of the counterpart signals across all the electromagnetic spectrum (Abbott B. P. *et al.*, *Astrophys. J. Lett.*, **848** (2017) L12). This observational campaign marks the birth of multi-messenger astronomy, which uses gravitational waves and electromagnetic emission. The collected multi-messenger data confirmed ten-year-old theoretical models. And at the same time, the richness of details of the taken data will require to develop new theory and to make other observations in the coming years to be interpreted.

(*) Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer Nature, *Multiple Messengers and Challenges in Astroparticle Physics*, edited by Roberto Aloisio, Eugenio Coccia and Francesco Vissani, © 2018, pp. 489-497.
(**) E-mail: marica.branchesi@gssi.infn.it

1. – The first gravitational-wave observation of the coalescence of a binary system of neutron stars

This gravitational-wave signal GW170817 (about 100 s starting from 24 Hz) was the loudest yet observed, with a combined signal-to-noise ratio (SNR) of 32.4 and a false-alarm-rate estimate of less than one per 8.0×10^4 years [1]. The signal carried information about the astrophysical source: masses, tidal deformability, distance, inclination, and sky localization. The component masses were inferred to be between 0.86 and $2.26 M_\odot$, and when restricting the component spins to a range consistent with the BNS observed population, the component masses result in the range 1.17–$1.60 M_\odot$. These masses are in agreement with the dynamically measured masses of known neutron stars in binaries [5], and their inconsistency with the masses of known black holes in galactic binary systems [6], suggesting that the GW170817 source was composed of two neutron stars.

However, the masses and lower limit on the compactness estimated from the gravitational-wave observations does not allow us to exclude objects more compact than neutron stars such as quark stars, black holes, or more exotic objects. The detection of the electromagnetic counterpart demonstrates the presence of matter confirming at least the presence of a neutron star in the system.

Details of the objects' internal structure become important as the orbital separation approaches the size of the bodies. As the gravitational-wave frequency increases, tidal effects in binary neutron stars increasingly affect the phase and become significant above $f_{\rm GW} \simeq 600$ Hz. Unfortunately, the noise of the LIGO and Virgo detectors at these frequencies made it difficult to extract signal information and only a lower bound on the energy emitted $E_{\rm rad} > 0.025 M_\odot c^2$ was set before the onset of strong tidal effects. Estimates of the tidal deformability (or polarizability) disfavor equations of state that predict less compact stars.

The luminosity distance estimate of 40^{+8}_{-14} Mpc set this GW signal as the closest ever observed and, by association, the closest short gamma-ray burst with a distance measurement. The viewing angle is constrained to be less than 56 degrees. The source was localized by the Advanced LIGO and Virgo network within a sky region of 28 deg^2 (90% probability) near the southern end of the constellation Hydra. The advanced Virgo interferometer was essential for the sky localization. The low signal amplitude observed (SNR=2) in Virgo with respect to the signal detected by the LIGO detectors indicated that the source was close to a blind spot of the interferometer significantly constraining the sky position. Distance and sky position located the source to within a volume of 380 Mpc^3. A region of the local Universe where properties and distribution of the galaxies are known. This played a key role to drive the successful optical search, which identified the counterpart and the host galaxy.

Binary neutron star mergers may result in a black hole, or short- or long-lived neutron star. In all these cases the remnant is expected to emit gravitational waves. However, the ringdown of a black hole produces gravitational waves at frequencies around 6 kHz, where the interferometer response is strongly reduced making its observation unfeasible. Searches have been made for short (tens of ms) and intermediate duration ($\leq$ 500 s)

gravitational-wave signals from a neutron star remnant at frequencies up to 4 kHz without evidence of a postmerger signal of astrophysical origin. The upper limits placed on the strength of the signal cannot definitively rule out the existence of a short- or long-lived postmerger neutron star.

The detection of GW170817 enables to estimate the local coalescence rate density of binary systems of neutron stars to $1540^{3200}_{-1220}\mathrm{Gpc}^{-3}\mathrm{yr}^{-1}$ [1].

2. – Discovery of the high-energy counterpart

A hard-X-ray electromagnetic (EM) counterpart, short-duration, sub-luminous Gamma Ray Burst, GRB170817A, was detected by the Fermi-GBM at 12:41:06 UTC, and announced automatically just 14 s after. It was also detected by the International Gamma-Ray Astrophysics Laboratory (INTEGRAL) through an offline search [7] initiated by the LIGO-Virgo and Fermi-GBM notice.

The probability of the near-simultaneous temporal and spatial observation of GRB170817A and GW170817 occurring by chance is 5.0×10^{-8}. The association of GW170817 and GRB170817A provides insight into the origin of short gamma-ray bursts and fundamental physics. These observations are the first firm evidence that at least a fraction of short-duration GRBs are associated with compact star mergers [2].

The arrival delay of only $(+1.74 \pm 0.05)$ s between high-energy and GW signal and travelled distance places stringent constraints on the difference between the speed of gravity and the speed of light to be between $+3 \times 10^{-15}$ and $+7 \times 10^{-16}$ times the speed of light. The measurement is consistent with GWs propagating at the speed of light as predicted by general relativity. The strong constraint on the speed of GWs ruled out many classes of modified gravity models (see, *e.g.* [8, 9]).

GRB170817A is the closest short GRB with known distance. Its fluence of $2.2 \times 10^{-7}\mathrm{erg\,cm}^{-2}$ in the 10–1000 keV energy range corresponds to a gamma-ray isotropic equivalent energy E_{iso} 4.3×10^{46} erg, which makes it three to four orders of magnitude smaller than the average energy of short GRBs with known redshift. The low luminosity has been interpreted as:

- uniform jet observed at large angles, provided that the jet bulk Lorentz factor is significantly smaller than usually assumed (see, *e.g.* [10]);
- emission from a structured jet, with a fast and energetic inner part surrounded by a slower component [11]. The slower component can be a cocoon, formed by the deposit of the energy of the jet when it excavates its way through the merger (dynamical) and post-merger (disk-viscosity-driven and neutrino-driven) ejecta. This shocked jet material may produce a quasi-isotropic prompt, high-energy EM signal (see, *e.g.* [12, 13]);
- jet-less scenario where an isotropic fireball expands ahead of the kilonova ejecta [14].

3. – The multi-wavelength electromagnetic follow-up campaign

The LIGO-Virgo sky localization of GW170817 (placing the source in a region of about 30 sq. degrees and at a distance of 40 Mpc) were sent to the astronomers at 17:54:51 UTC and became observable to telescopes in Chile about 10 hr after the merger. Immediately, the One-Meter, Two-Hemisphere (1M2H) team using the 1 m Swope Telescope started to observe the galaxies in the volume corresponding to the GW170817 signal. The galaxies were ranked based on stellar mass and star formation rate. In one of them, NGC4993, an early-type galaxy at the same distance as the gravitational-wave signal, they discovered a new and bright blue optical transient (SSS17a, now with the International Astronomical Union identification of AT 2017gfo) at 23:33 UTC. The optical transient was independently detected by multiple teams during the same night. Then, all the telescopes and satellites around the world started the most extensive observing campaign ever across the electromagnetic spectrum pointing at AT 2017gfo for weeks [4]. The observations taken by a network of ground-based telescopes, from 40 cm to 10 m, and space-based observatories in the ultraviolet (UV), optical, and near-infrared showed a peculiar spectral energy evolution: a rapid dimming of initial UV-blue emission (see, *e.g.* [15]), brightening of the red and near-infrared emission, which began to fade after a few days to a week (see, *e.g.* [16]). The telescopes involved in the photometric observation include: CTIO1.3 m, DECam, IRSF, the Gemini-South FLAMINGO2, Gemini-South GMOS, GROND, HST, iTelescope.Net telescopes, the Korea Microlensing Telescope Network, Las Cumbres Observatory, the Lee Sang Gak Telescope, the Magellan-Baade and Magellan-Clay 6.5 m telescopes, the Nordic Optical Telescope, Pan-STARRS1, REM/ROS2 and REM/REMIR, SkyMapper, Subaru Hyper Suprime-Cam, ESO-VISTA, ESO-VST/OmegaCAM, and ESO-VLT/FORS2.

The first spectrum of AT 2017gfo was taken by the Magellan-Clay telescope. It showed a blue and featureless continuum between 4000 and 10000 Å. It appeared as an unusual spectrum, but common to cataclysmic-variable stars and young core-collapse supernovae. The following spectra taken 24–48 hrs later (by the SALT-RSS, ePESSTO with the EFOSC2 instrument in spectroscopic mode at the ESO New Technology Telescope, the X-shooter spectrograph on the ESO Very Large Telescope, and the Goodman Spectrograph on the 4 m SOAR telescope) were crucial to exclude any transient contaminant and firmly associate AT 2017gfo with GW170817. The spectra showed an exceptionally fast spectral evolution in the blue part and absence of any features identifiable with absorption lines common in supernova-like transients. The rapid fading in the blue, the spectral energy evolution and appearance of broad spectral features in the red near-IR reproduced the expectations of disk wind outflow and dynamical ejecta kilonovae models (see, *e.g.* [17-20]). During the coalescence of two neutron stars, the mass ejected at extremely high velocity in the interstellar medium is the ideal site for the formation of heavy elements for rapid neutron capture. While the radioactive decay of such elements determines the emission called kilonova, the Universe is enriched with elements heavier than iron, the r-process nucleosynthesis elements. The ESO-VLT/X-shooter spectra, which simultaneously cover the wavelength range 3200–24800 Å and were taken over 2

weeks with a close to daily sampling revealed signatures of the radioactive decay of these elements [10, 21]. Coalescences of two neutron stars are very rare events; in Milky Way equivalent galaxies, they can happen with a rate of 32–474 Myr^{-1} [22], but the quantity of r-process materials that can be formed are enough to explain the abundance of heavy elements in the Universe (see, *e.g.* [25, 26, 23, 24]).

The UV/optical/infrared observations of GW170817 compared with kilonova models and numerical relativity results were used to estimate a lower bound on the tidal deformability parameter. Together with the upper bound from the GW signal, both extremely stiff and soft NS equations of state, can be ruled out [27].

After a significant upper limit on the emission associated to no detections by Swift-XRT and the Chandra X-ray Observatory, on 26 Aug 2017 Chandra detected X-ray emission at the position of AT 2017gfo.

The observed X-ray flux implies an isotropic luminosity of $9 \times 10^{38} \mathrm{erg\,s}^{-1}$ if located in NGC 4993 [28]. Further Chandra observations, performed between 1 and 2 Sept. 2017, confirmed the presence of continued X-ray activity, and suggested a slight increase in luminosity.

On September 2 and September 3 (16 days after the merger) a radio emission appeared in the Karl G. Jansky Very Large Array (VLA) data at a frequency of 3 GHz and at a frequency of 6 GHz. The Australia Telescope Compact Array (ATCA) also detected the source on September 5 in the 5.5–9 GHz band. In observations at 3 GHz with the VLA, the source shows evidence of an increase in flux density over a timescale of 2 weeks [29]. The observed X-ray radio emissions are consistent with the emission from a relativistic collimated jet viewed off-axis. The observer not being aligned with the opening angle of the jetted outflow is expected to observe the afterglow when the jet has spread and decelerated enough so that the beaming cone of the emission includes the observer's line of sight; the onset of the afterglow is delayed by several days or weeks and the emission weaker with respect to an on-axis observer. The X-ray and radio detection were interpreted as the first observations of an off-axis GRB (see, *e.g.* [28, 30]), a scenario theorized for about two decades but never observed before. Both the X-ray and radio emission were also consistent with a cocoon that breaks out when the jet transfers a large fraction of its energy into the surrounding ejecta. The cocoon expands over a wide angle at mildly relativistic velocities [31].

No ultra-high-energy gamma rays and no neutrino candidates consistent with the source were found in follow-up searches.

The proximity to the Sun did not allow the optical and X-ray observations of the GW source for about 3 months. Observations of GW170817 continued in the radio band (0.6–18 GHz frequency range) with the VLA, the ATCA and the upgraded Giant Metrewave Radio Telescope, showing a steady rise of the light curve and a spectrum consistent with optically thin synchrotron emission [32]. These observations completely exclude the collimated ultra-relativistic uniform jet viewed off-axis and an isotropic thin and uniform fireball, and requires a mildly relativistic wide-angle outflow moving towards the observer. This outflow could be the high-velocity tail of the neutron-rich material dynamically ejected during the merger (isotropic outflow with velocity profile) but the

most likely scenario remains the cocoon able to explain also the prompt gamma-rays. Chandra observations during the first days of December 2017 showed brightening X-ray emission [33, 30, 28]. The similar, slow rising of both the X-ray and radio emissions indicate a common (synchrotron) source as origin. On December 6, AT 2017gfo was also detected in the optical band by the HST [34].

The multi-messenger campaign of GW170817 marked also the beginning of cosmology with gravitational waves [35]. Combining the distance to the source inferred from the gravitational-wave signal with the recession velocity of the galaxy inferred from measurements of the redshift using electromagnetic data, the present-day expansion rate of the Universe has been determined as given by the Hubble constant of $70.0^{+12.0}_{-8.0}\mathrm{km\,s^{-1}Mpc^{-1}}$ (maximum a posteriori and 68% credible interval). The Hubble constant estimate is consistent with existing measurements, including the cosmic microwave background measurements from Planck [36], the Type Ia supernova measurements from SHoES [37], the baryon acoustic oscillations measurements from SDSS [38], the strong lensing measurements from H0LiCOW [39], while being completely independent of them. The large uncertainty in the H0 measurement with GWs due to the degeneracy between distance and inclination in the GW measurements could improve with precise polarization measurements, and additional standard-siren measurements from future gravitational-wave sources will provide more precise constraints of this cosmological parameter.

As described above, the first multi-messenger campaign had and will have strong impact in many astrophysics fields starting from emission process physics and nuclear astrophsyics, to fundamental physics and cosmology. This campaign demonstrated the importance of the development, coordination and collaboration of a network of gravitational-wave, electromagnetic, and neutrino observatories, the importance of joint analysis and interpretation of the multi-messenger data. We are in an exciting new era of multi-messenger, time-domain astronomy.

REFERENCES

[1] ABBOTT B. P., ABBOTT R., ABBOTT T. D., ACERNESE F., ACKLEY K., ADAMS C., ADAMS T., ADDESSO P., ADHIKARI R. X., ADYA V. B. *et al.*., *Phys. Rev. Lett.*, **119** (2017) 161101.

[2] ABBOTT B. P., ABBOTT R., ABBOTT T. D., ACERNESE F., ACKLEY K., ADAMS C., ADAMS T., ADDESSO P., ADHIKARI R. X., ADYA V. B. *et al.*, *Astrophys. J. Lett.*, **848** (2017) L13.

[3] GOLDSTEIN A., VERES P., BURNS E., BRIGGS M. S., HAMBURG R., KOCEVSKI D., WILSON-HODGE C. A., PREECE R. D., POOLAKKIL S., ROBERTS O. J., HUI C. M., CONNAUGHTON V., RACUSIN J., VON KIENLIN A., DAL CANTON T., CHRISTENSEN N., LITTENBERG T., SIELLEZ K., BLACKBURN L., BROIDA J., BISSALDI E., CLEVELAND W. H., GIBBY M. H., GILES M. M., KIPPEN R. M., MCBREEN S., MCENERY J., MEEGAN C. A., PACIESAS W. S. and STANBRO M., *Astrophys. J. Lett.*, **848** (2017) L14.

[4] ABBOTT B. P., ABBOTT R., ABBOTT T. D., ACERNESE F., ACKLEY K., ADAMS C., ADAMS T., ADDESSO P., ADHIKARI R. X., ADYA V. B. *et al.*, *Astrophys. J. Lett.*, **848** (2017) L12.

[5] Antoniadis J., Freire P. C. C., Wex N., Tauris T. M., Lynch R. S., van Kerkwijk M. H., Kramer M., Bassa C., Dhillon V. S., Driebe T., Hessels J. W. T., Kaspi V. M., Kondratiev V. I., Langer N., Marsh T. R., McLaughlin M. A., Pennucci T. T., Ransom S. M., Stairs I. H., van Leeuwen J., Verbiest J. P. W. and Whelan D. G., *Science*, **340** (2013) 448.

[6] Kreidberg L., Bailyn C. D., Farr W. M. and Kalogera V., *Astrophys. J.*, **757** (2012) 36.

[7] Savchenko V., Ferrigno C., Kuulkers E., Bazzano A., Bozzo E., Brandt S., Chenevez J., Courvoisier T. J.-L., Diehl R., Domingo A., Hanlon L., Jourdain E., von Kienlin A., Laurent P., Lebrun F., Lutovinov A., Martin-Carrillo A., Mereghetti S., Natalucci L., Rodi J., Roques J.-P., Sunyaev R. and Ubertini P., *Astrophys. J. Lett.*, **848** (2017) L15.

[8] Baker T., Bellini E., Ferreira P. G., Lagos M., Noller J. and Sawicki I., *Phys. Rev. Lett.*, **119** (2017) 251301.

[9] Creminelli P. and Vernizzi F., *Phys. Rev. Lett.*, **119** (2017) 251302.

[10] Pian E., D'Avanzo P., Benetti S., Branchesi M., Brocato E., Campana S., Cappellaro E., Covino S., D'Elia V., Fynbo J. P. U., Getman F., Ghirlanda G., Ghisellini G., Grado A., Greco G., Hjorth J., Kouveliotou C., Levan A., Limatola L., Malesani D., Mazzali P. A., Melandri A., Møller P., Nicastro L., Palazzi E., Piranomonte S., Rossi A., Salafia O. S., Selsing J., Stratta G., Tanaka M., Tanvir N. R., Tomasella L., Watson D., Yang S., Amati L., Antonelli L. A., Ascenzi S., Bernardini M. G., Boër M., Bufano F., Bulgarelli A., Capaccioli M., Casella P., Castro-Tirado A. J., Chassande-Mottin E., Ciolfi R., Copperwheat C. M., Dadina M., De Cesare G., di Paola A., Fan Y. Z., Gendre B., Giuffrida G., Giunta A., Hunt L. K., Israel G. L., Jin Z.-P., Kasliwal M. M., Klose S., Lisi M., Longo F., Maiorano E., Mapelli M., Masetti N., Nava L., Patricelli B., Perley D., Pescalli A., Piran T., Possenti A., Pulone L., Razzano M., Salvaterra R., Schipani P., Spera M., Stamerra A., Stella L., Tagliaferri G., Testa V., Troja E., Turatto M., Vergani S. D. and Vergani D., *Nature*, **551** (2017) 67.

[11] Kathirgamaraju A., Barniol Duran R. and Giannios D., *Mon. Not. R. Astron. Soc.*, **473** (2018) L121.

[12] Lazzati D., Deich A., Morsony B. J. and Workman J. C., *Mon. Not. R. Astron. Soc.*, **471** (2017) 1652.

[13] Lazzati D., López-Cámara D., Cantiello M., Morsony B. J., Perna R. and Workman J. C., *Astrophys. J. Lett.*, **848** (2017) L6.

[14] Salafia, O. S., Ghisellini G. and Ghirlanda G., *Mon. Not. R. Astron. Soc.*, **474** (2018) L7.

[15] Evans P. A., Cenko S. B., Kennea J. A., Emery S. W. K., Kuin N. P. M., Korobkin O., Wollaeger R. T., Fryer C. L., Madsen K. K., Harrison F. A., Xu Y., Nakar E., Hotokezaka K., Lien A., Campana S., Oates S. R., Troja E., Breeveld A. A., Marshall F. E., Barthelmy S. D., Beardmore A. P., Burrows D. N., Cusumano G., D'Ai A., D'Avanzo P., D'Elia V., de Pasquale M., Even W. P., Fontes C. J., Forster K., Garcia J., Giommi P., Grefenstette B., Gronwall C., Hartmann D. H., Heida M., Hungerford A. L., Kasliwal M. M., Krimm H. A., Levan A. J., Malesani D., Melandri A., Miyasaka H., Nousek J. A., O'Brien P. T., Osborne J. P., Pagani C., Page K. L., Palmer D. M., Perri M., Pike S., Racusin J. L., Rosswog S., Siegel M. H., Sakamoto T., Sbarufatti B., Tagliaferri G., Tanvir N. R. and Tohuvavohu A., *Science*, **358** (2017) 1565.

[16] VILLAR V. A., GUILLOCHON J., BERGER E., METZGER B. D., COWPERTHWAITE P. S., NICHOLL M., ALEXANDER K. D., BLANCHARD P. K., CHORNOCK R., EFTEKHARI T., FONG W., MARGUTTI R. and WILLIAMS P. K. G., *Astrophys. J. Lett.*, **851** (2017) L21.
[17] KASEN D., BADNELL N. R. and BARNES J., *Astrophys. J. Lett.*, **774** (2013) 25.
[18] METZGER B. D., *Living Rev. Relativ.*, **20** (2017) 3.
[19] KASEN D., METZGER B., BARNES J., QUATAERT E. and RAMIREZ-RUIZ E., *Nature*, **551** (2017) 80.
[20] TANAKA M., UTSUMI Y., MAZZALI P. A., TOMINAGA N., YOSHIDA M., SEKIGUCHI Y., MOROKUMA T., MOTOHARA K., OHTA K., KAWABATA K. S., ABE F., AOKI K., ASAKURA Y., BAAR S., BARWAY S., BOND I. A., DOI M., FUJIYOSHI T., FURUSAWA H., HONDA S., ITOH Y., KAWABATA M., KAWAI N., KIM J. H., LEE C.-H., MIYAZAKI S., MORIHANA K., NAGASHIMA H., NAGAYAMA T., NAKAOKA T., NAKATA F., OHSAWA R., OHSHIMA T., OKITA H., SAITO T., SUMI T., TAJITSU A., TAKAHASHI J., TAKAYAMA M., TAMURA Y., TANAKA I., TERAI T., TRISTRAM P. J., YASUDA N. and ZENKO T., *PASJ*, **69** (2017) 102.
[21] SMARTT S. J., CHEN T.-W., JERKSTRAND A., COUGHLIN M., KANKARE E., SIM S. A., FRASER M., INSERRA C., MAGUIRE K., CHAMBERS K. C., HUBER M. E., KRÜHLER T., LELOUDAS G., MAGEE M., SHINGLES L. J., SMITH K. W., YOUNG D. R., TONRY J., KOTAK R., GAL-YAM A., LYMAN J. D., HOMAN D. S., AGLIOZZO C., ANDERSON J. P., ANGUS C. R., ASHALL C., BARBARINO C., BAUER F. E., BERTON M., BOTTICELLA M. T., BULLA M., BULGER J., CANNIZZARO G., CANO Z., CARTIER R., CIKOTA A., CLARK P., DE CIA A., DELLA VALLE M., DENNEAU L., DENNEFELD M., DESSART L., DIMITRIADIS G., ELIAS-ROSA N., FIRTH R. E., FLEWELLING H., FLÖRS A., FRANCKOWIAK A., FROHMAIER C., GALBANY L., GONZÁLEZ-GAITÁN S., GREINER J., GROMADZKI M., GUELBENZU A. N., GUTIÉRREZ C. P., HAMANOWICZ A., HANLON L., HARMANEN J., HEINTZ K. E., HEINZE A., HERNANDEZ M.-S., HODGKIN S. T., HOOK I. M., IZZO L., JAMES P. A., JONKER P. G., KERZENDORF W. E., KLOSE S., KOSTRZEWA-RUTKOWSKA Z., KOWALSKI M., KROMER M., KUNCARAYAKTI H., LAWRENCE A., LOWE T. B., MAGNIER E. A., MANULIS I., MARTIN-CARRILLO A., MATTILA S., MCBRIEN O., MÜLLER A., NORDIN J., O'NEILL D., ONORI F., PALMERIO J. T., PASTORELLO A., PATAT F., PIGNATA G., PODSIADLOWSKI P., PUMO M. L., PRENTICE S. J., RAU A., RAZZA A., REST A., REYNOLDS T., ROY R., RUITER A. J., RYBICKI K. A., SALMON L., SCHADY P., SCHULTZ A. S. B., SCHWEYER T., SEITENZAHL I. R., SMITH M., SOLLERMAN J., STALDER B., STUBBS C. W., SULLIVAN M., SZEGEDI H., TADDIA F., TAUBENBERGER S., TERRERAN G., VAN SOELEN B., VOS J., WAINSCOAT R. J., WALTON N. A., WATERS C., WEILAND H., WILLMAN M., WISEMAN P., WRIGHT D. E., WYRZYKOWSKI Ł. and YARON O., *Nature*, **551** (2017) 75.
[22] ABBOTT B. P., ABBOTT R., ABBOTT T. D., ACERNESE F., ACKLEY K., ADAMS C., ADAMS T., ADDESSO P., ADHIKARI R. X., ADYA V. B. *et al.*, *Astrophys. J. Lett.*, **850** (2017) L40.
[23] ABBOTT B. P., ABBOTT R., ABBOTT T. D., ACERNESE F., ACKLEY K., ADAMS C., ADAMS T., ADDESSO P., ADHIKARI R. X., ADYA V. B. *et al.*, *Astrophys. J. Lett.*, **850** (2017) L39.
[24] CÔTÉ B., FRYER C. L., BELCZYNSKI K., KOROBKIN O., CHRUŚLIŃSKA M., VASSH N., MUMPOWER M. R., LIPPUNER J., SPROUSE T. M., SURMAN R. and WOLLAEGER R., ArXiv e-print 1710.05875 (2017).
[25] DROUT M. R., PIRO A. L., SHAPPEE B. J., KILPATRICK C. D., SIMON J. D., CONTRERAS C., COULTER D. A., FOLEY R. J., SIEBERT M. R., MORRELL N., BOUTSIA K., DI MILLE F., HOLOIEN T. W.-S., KASEN D., KOLLMEIER J. A., MADORE B. F., MONSON A. J., MURGUIA-BERTHIER A., PAN Y.-C., PROCHASKA J. X., RAMIREZ-RUIZ E., REST A.,

ADAMS C., ALATALO K., BAÑADOS E., BAUGHMAN J., BEERS T. C., BERNSTEIN R. A., BITSAKIS T., CAMPILLAY A., HANSEN T. T., HIGGS C. R., JI A. P., MARAVELIAS G., MARSHALL J. L., MONI BIDIN C., PRIETO J. L., RASMUSSEN K. C., ROJAS-BRAVO C., STROM A. L., ULLOA N., VARGAS-GONZÁLEZ J., WAN Z. and WHITTEN D. D., *Science*, **358** (2017) 1560.

[26] KASLIWAL M. M., NAKAR E., SINGER L. P., KAPLAN D. L., COOK D. O., VAN SISTINE A., LAU R. M., FREMLING C., GOTTLIEB O., JENCSON J. E., ADAMS S. M., FEINDT U., HOTOKEZAKA K., GHOSH S., PERLEY D. A., YU P.-C., PIRAN T., ALLISON J. R., ANUPAMA G. C., BALASUBRAMANIAN A., BANNISTER K. W, BALLY J., BARNES J., BARWAY S., BELLM E., BHALERAO V., BHATTACHARYA D., BLAGORODNOVA N., BLOOM J. S., BRADY P. R., CANNELLA C., CHATTERJEE D., CENKO S. B., COBB B. E., COPPERWHEAT C., CORSI A., DE K., DOBIE D., EMERY S. W. K., EVANS P. A., FOX O. D., FRAIL D. A., FROHMAIER C., GOOBAR A., HALLINAN G., HARRISON F., HELOU G., HINDERER T., HO A. Y. Q., HORESH A., IP W.-H., ITOH R., KASEN D., KIM H., KUIN N. P. M., KUPFER T., LYNCH C., MADSEN K., MAZZALI P. A., MILLER A. A., MOOLEY K., MURPHY T., NGEOW C.-C., NICHOLS D., NISSANKE S., NUGENT P., OFEK E. O., QI H., QUIMBY R. M., ROSSWOG S., RUSU F., SADLER E. M., SCHMIDT P., SOLLERMAN J., STEELE I., WILLIAMSON A. R., XU Y., YAN L., YATSU Y., ZHANG C. and ZHAO W., ArXiv e-print 1710.05436 (2017).

[27] RADICE D., PEREGO A., ZAPPA F. and BERNUZZI S., ArXiv e-print 1711.03647 (2017).

[28] TROJA E., PIRO L., VAN EERTEN H., WOLLAEGER R. T., IM M., FOX O. D., BUTLER N. R., CENKO S. B., SAKAMOTO T., FRYER C. L., RICCI R., LIEN A., RYAN R. E., KOROBKIN O., LEE S.-K., BURGESS J. M., LEE W. H., WATSON A. M., CHOI C., COVINO S., D'AVANZO P., FONTES C. J., GONZÁLEZ J. B., KHANDRIKA H. G., KIM J., KIM S.-L., LEE C.-U., LEE H. M., KUTYREV A., LIM G., SÁNCHEZ-RAMÍREZ R., VEILLEUX S., WIERINGA M. H. and YOON Y., *Nature*, **551** (2017) 71.

[29] HALLINAN G., CORSI A., P. MOOLEY K., HOTOKEZAKA KENTA, NAKAR E., KASLIWAL M., KAPLAN DAVID, A. FRAIL D., T. MYERS S., MURPHY T., DE K., DOBIE D., ALLISON JAMES, W. BANNISTER K., BHALERAO VARUN, CHANDRA P., E. CLARKE T., GIACINTUCCI S., Y. Q. HO A. and P. SINGER L., *Science*, **358** (2017) 1579.

[30] MARGUTTI R., BERGER E., FONG W., GUIDORZI C., ALEXANDER K. D., METZGER B. D., BLANCHARD P. K., COWPERTHWAITE P. S., CHORNOCK R., EFTEKHARI T., NICHOLL M., VILLAR V. A., WILLIAMS P. K. G., ANNIS J., BROWN D. A., CHEN H., DOCTOR Z., FRIEMAN J. A., HOLZ D. E., SAKO M. and SOARES-SANTOS M., *Astrophys. J. Lett.*, **848** (2017) L20.

[31] GOTTLIEB O., NAKAR E., PIRAN T. and HOTOKEZAKA K., ArXiv e-print 1710.05896 (2017).

[32] MOOLEY K. P., NAKAR E., HOTOKEZAKA K., HALLINAN G., CORSI A., FRAIL D. A., HORESH A., MURPHY T., LENC E., KAPLAN D. L., DE K., DOBIE D., CHANDRA P., DELLER A., GOTTLIEB O., KASLIWAL M. M., KULKARNI S. R., MYERS S. T., NISSANKE S., PIRAN T., LYNCH C., BHALERAO V., BOURKE S., BANNISTER K. W. and SINGER L. P., *Nature*, **554** (2018) 207.

[33] HAGGARD D., RUAN J. J., NYNKA M., KALOGERA V. and EVANS P., "LIGO/Virgo GW170817: Brightening X-ray Emission from GW170817/GRB170817A/SSS17a", GCN Circular 22206 (2017).

[34] LEVAN A. J., LYMAN J. D., TANVIR N. R., MANDEL I., HJORTH J., FRUCHTER A. S., KANGAS T., GOMPERTZ B., WIERSEMA K. D., STEEGHS D. and ROSSWOG S., "LIGO/Virgo GW170817: Further Hubble Space Telescope observations", GCN Circular 22207 (2017).

[35] ABBOTT B. P., ABBOTT R., ABBOTT T. D., ACERNESE F., ACKLEY K., ADAMS C., ADAMS T., ADDESSO P., ADHIKARI R. X., ADYA V. B. *et al.*, *Nature*, **551** (2017) 85.
[36] PLANCK COLLABORATION (ADE P. A. R., AGHANIM N., ARNAUD M., ASHDOWN M., AUMONT J., BACCIGALUPI C., BANDAY A. J., BARREIRO R. B., BARTLETT J. G. *et al.*), *Astron. Astrophys.*, **594** (2016) A13.
[37] RIESS A. G., MACRI L. M., HOFFMANN S. L., SCOLNIC D., CASERTANO S., FILIPPENKO A. V., TUCKER B. E., REID M. J., JONES D. O., SILVERMAN J. M., CHORNOCK R., CHALLIS P., YUAN W., BROWN P. J. Y FOLEY R. J., *Astrophys. J.*, **826** (2016) 56.
[38] BOSS COLLABORATION (AUBOURG É., BAILEY S., BAUTISTA J. E., BEUTLER F., BHARDWAJ V., BIZYAEV D., BLANTON M., BLOMQVIST M., BOLTON A. S., BOVY J., BREWINGTON H., BRINKMANN J., BROWNSTEIN J. R., BURDEN A., BUSCA N. G., CARITHERS W., CHUANG C.-H., COMPARAT J., CROFT R. A. C., CUESTA A. J., DAWSON K. S., DELUBAC T., EISENSTEIN D. J., FONT-RIBERA A., GE J., LE GOFF J.-M., GONTCHO S. G. A., GOTT J. R., GUNN J. E., GUO H., GUY J., HAMILTON J.-C., HO S., HONSCHEID K., HOWLETT C., KIRKBY D., KITAURA F. S., KNEIB J.-P., LEE K.-G., LONG D., LUPTON R. H., MAGAÑA M. V., MALANUSHENKO V., MALANUSHENKO E., MANERA M., MARASTON C., MARGALA D., MCBRIDE C. K., MIRALDA-ESCUDÉ J., MYERS A. D., NICHOL R. C., NOTERDAEME P., NUZA S. E., OLMSTEAD M. D., ORAVETZ D., PÂRIS I., PADMANABHAN N., PALANQUE-DELABROUILLE N., PAN K., PELLEJERO-IBANEZ M., PERCIVAL W. J., PETITJEAN P., PIERI M. M., PRADA F., REID B., RICH J., ROE N. A., ROSS A. J., ROSS N. P., ROSSI G., RUBIÑO-MARTÍN J. A., SÁNCHEZ A. G., SAMUSHIA L., GÉNOVA-SANTOS R. T., SCÓCCOLA C. G., SCHLEGEL D. J., SCHNEIDER D. P., SEO H.-J., SHELDON E., SIMMONS A., SKIBBA R. A., SLOSAR A., STRAUSS M. A., THOMAS D., TINKER J. L., TOJEIRO R., VAZQUEZ J. A., VIEL M., WAKE D. A., WEAVER B. A., WEINBERG D. H., WOOD-VASEY W. M., YÈCHE C., ZEHAVI I. and ZHAO G.-B.), *Phys. Rev. D*, **92** (2015) 123516.
[39] BONVIN V., COURBIN F., SUYU S. H., MARSHALL P. J., RUSU C. E., SLUSE D., TEWES M., WONG K. C., COLLETT T., FASSNACHT C. D., TREU T., AUGER M. W., HILBERT S., KOOPMANS L. V. E., MEYLAN G., RUMBAUGH N., SONNENFELD A. and SPINIELLO C., *Mon. Not. R. Astron. Soc.*, **465** (2017) 4914.

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200007

The standard model of cosmology: A skeptic's guide

Douglas Scott(*)

Department of Physics & Astronomy, University of British Columbia - Vancouver, Canada

Summary. — The status of the standard cosmological model, also known as "ΛCDM" is described. With some simple assumptions, this model fits a wide range of data, with just six (or seven) free parameters. One should be skeptical about this claim, since it implies that we now have an astonishingly good picture of the statistical properties of the large-scale Universe. However, the successes of the model cannot be denied, including more than $1000\,\sigma$ worth of detection of CMB anisotropy power. The model is older than most modern astrophysicists seem to appreciate, and has not fundamentally changed for more than a quarter of a century. Tensions and anomalies are often discussed, and while we should of course be open to the possibility of new physics, we should also be skeptical about the importance of 2–$3\,\sigma$ differences between data sets until they become more significant. Still, today's SMC is surely not the full story and we should be looking for extensions or new ingredients to the model, guided throughout by a skeptical outlook.

1. – What is the standard model of cosmology?

The currently best-fitting picture for describing the statistics of the Universe on large scales, the standard model of cosmology (or SMC), is often known as ΛCDM, since it is a model in which the matter is mostly cold and dark (*i.e.* effectively collisionless and

(*) E-mail: dscott@phas.ubc.ca

with no electromagnetic interactions, CDM), with the bulk of the energy density of the Universe behaving like vacuum energy (*i.e.* like the cosmological constant of general relativity, Λ). But things are even more specific than that, with the values of only about half-a-dozen free parameters being enough to make a Universe that looks statistically just like the one we live in — and several of those parameters are now known to a really impressive level of precision. So the "SMC" is now quite precisely prescribed.

It is an astonishing achievement of modern cosmology that we have come to have such a successful model, especially when one considers that there is no *a posteriori* reason to expect things to be this simple. In physics we are driven to accept a model for several reasons — certainly that it fits the data, but also because of some less well defined notion of æsthetics. The simple group-theoretical underpinnings of the standard model of particle physics and the elegance of the field equations of general relativity are obvious examples of this. Sure, they fit lots and lots of experimental data, but they are also really *nice*! But for cosmology, no one would claim that the SMC is beautiful, or even that it has to be correct because all alternatives are uglier. Certainly the SMC has some degree of simplicity (since it does not need many free parameters), but why do those parameters have the values that they do (see sects. **2** and **3**)? And why are there not lots of other parameters required (see sect. **4**)? Despite the fact that nothing in the basic cosmological picture has changed since the early-to-mid-1990s (see sect. **5**), most cosmologists are expecting something else to be just around the corner. After all, surely the SMC can not be all there is?

In these notes I would like to bring some attention to the idea that we should be skeptical([1]) here, since we are dealing with very large themes. A model that purports to describe the whole of the observable Universe should be met with a decent dose of incredulity! It is important that we retain a healthy level of skepticism when discussing any such claims. But at the same time we should also remember to be skeptical about *counter*-claims (sects. **6** and **7**) that have not passed the same level of scrutiny. And we should keep in mind criteria that define what skepticism is (sect. **8**), so that we can isolate the successes of the SMC, while remembering that parts of modern cosmology's lore remain quite speculative (sect. **9**).

2. – The parameters and assumptions of the SMC

Let us be explicit about the standard model by giving the modern values of its basic parameters. Right now the determination of these quantities is driven by cosmic microwave background (CMB) anisotropy experiments, and in particular by results from the *Planck* satellite [1] (supported by many other kinds of data) on the CMB power spectra (which are discussed further in sect. **4**). Because of this, the basic parameter set is currently given in terms of the quantities that are most directly measured by CMB

([1]) I am going to use the American English spelling of this word, since to many non-British people, "sceptic" looks too much like "septic".

TABLE I. – *Basic cosmological parameters, from a combination of Planck* 2015 *data and other constraints from BAO, SNe and* H_0 *data (see ref.* [2]*). The CMB temperature comes from an analysis of the monopole spectral data by Fixsen* [3].

Physical baryon density	$\Omega_\mathrm{b}h^2$	0.02227 ± 0.00020
Physical CDM density	$\Omega_\mathrm{c}h^2$	0.1184 ± 0.0012
Angular parameter	$100\theta_*$	1.04106 ± 0.00041
Reionisation optical depth	τ	0.067 ± 0.013
Power spectrum amplitude	$\ln(10^{10}A_\mathrm{s})$	3.064 ± 0.024
Power spectrum slope	n_s	0.9681 ± 0.0044
CMB temperature	T_0 [K]	2.7255 ± 0.0006

experiments. This means that the parameters most often discussed in relation to observational constraints are not necessarily the ones that are simplest to explain to the general public, or that are the focus of non-CMB cosmologists. These parameters are listed in table I. The set consists of: two densities, $\Omega_\mathrm{b}h^2$ and $\Omega_\mathrm{c}h^2$ (for baryonic matter and cold dark matter separately, since they have distinct effects on the CMB power spectra), including a scaling of physical density with the dimensionless Hubble parameter, $h \equiv H_0/100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; a parameter θ_* that corresponds to the sound horizon divided by the angular diameter distance to last scattering, which quantifies sliding the CMB power spectra left and right; the amplitude A_s of the initial power spectrum of density perturbations, defined at a particular scale, and often given as a logarithm; the slope n of the initial power spectrum as a function of wavenumber; and a parameter τ describing how much the primary CMB anisotropies are scattered by the reionised medium at low redshifts.

By far the best determined of these parameters is θ_*, with a signal-to-noise ratio (S/N) of about 2500 (from table I, or about 2300 from the CMB alone). Then follows A_s, $\Omega_\mathrm{b}h^2$ and $\Omega_\mathrm{c}h^2$, with $S/N \simeq 100$, while n_s and τ only differ from their default values (of 1 and 0, respectively) at $S/N \simeq 5$. Other cosmological parameters that are often discussed include H_0, t_0, Ω_m, Ω_Λ, z_reion, etc., which are not independent, but can be determined from the six parameters in the context of the SMC.

Although it is often stated that there are six basic parameters, there is a seventh that is often ignored. This is the temperature of the CMB today (or equivalently the radiation density), which is constrained using data from the *COBE*-FIRAS instrument [4], as well as from several other experiments (see ref. [3]). The determination is systematics dominated, with $S/N \simeq 5000$. It is hence more precise than other parameters, and dramatically better determined than other densities. For that reason it is usually considered to be fixed, and not a free parameter at all. However, the precision is starting to approach the cosmic-variance limit, and so if T_0 was measured with much smaller errors, we should have to consider the fact that we can only measure parameters within our Hubble patch and not actually "background" parameter values (see ref. [5] for discussion).

TABLE II. – *Some assumptions of the SMC. Note that several of these apply to our observable volume (which is the only part of the Universe that we can test) only.*

Understanding the Cosmos is possible for human beings
Physics is the same everywhere and at all times
General relativity is the correct theory of gravity on cosmological scales
The Universe is approximately statistically homogeneous and isotropic
The Universe is spatially flat on large scales
The dark energy behaves like a cosmological constant, with $w = -1$
The dark matter is collisionless and cold for the purposes of cosmology
There are three species of nearly massless neutrinos
There are no additional light particles contributing to the background
Density perturbations are adiabatic in nature
The initial conditions were Gaussian
The running of the primordial power spectrum is negligible
The contribution of gravitational waves is negligible
Topological defects were unimportant for structure formation
The physics of recombination is fully understood
One parameter is sufficient to describe the effects of reionisation

But (to be skeptical about this), we might wonder whether there are other hidden parameters. There definitely are, to some extent, but mostly any additions to the SMC are better cast as assumptions. In fact there are many of these, and it is important to be clear that the six (or seven) parameters of the SMC are only descriptive of the Universe within a specific framework. A list of these assumptions is given in table II (and the reader can probably think of more).

All of these assumptions are testable, and they all *have* been investigated. Many of them are tested through putting limits on extensions to the SMC, *e.g.* checking whether the curvature is consistent with flat space, whether there is evidence for modified gravity, non-trivial dark energy (*i.e.* $w \neq -1$), or non-Gaussianity, or whether there are signs of the effects of massive neutrinos or cosmic strings (*e.g.* see refs. [6, 2]).

Nevertheless, this is definitely a place where we need to exercise caution. The confidence with which we know the values of the basic set of six (or seven) parameters depends on this being the full parameters space. If there are more ingredients in the actual model, then the parameters in the basic set will have larger uncertainties. For example, if we consider models that allow curvature then the constraints on w are very much weakened. Hence we need to look carefully at these tests. Right now there is no strong evidence for any additional parameter, but we fully expect that there will be more ingredients required as the data improve, *e.g.* that the effects of massive neutrinos or primordial

gravitational waves will eventually be measured. And there may be genuine surprises of course, like multiple kinds of dark matter or dark energy, or important extra components, such as magnetic fields or isocurvature modes.

Nevertheless, there *has* been caution exercised, and despite attempts to find evidence for additional parameters, the basic set continues to fit very high signal-to-noise data — particularly the CMB power spectra.

3. – The numbers that describe the Universe

Since it appears that the set of numbers required to statistically describe the cosmological model has just seven elements, then these values become important quantities that should be better known, among astronomers and non-astronomers alike. Many people follow the detailed statistics of their favourite sports teams, or can name the capital cities of various countries, or give the sequence of colours of the rainbow, or list the wives of Henry VIII in order, or name the actors who have played their favourite time-travelling alien. Almost everyone learns the list of planets in the Solar System, through the mnemonic about pizzas (that no longer includes pizza!). So why do most humans not know the numbers that describe the Universe that we live in?

Perhaps one of the problems is that the usual six parameters coming from CMB anisotropies are quite esoteric. This becomes apparent as soon as one tries to explain the values in table I to the general public. However, these six arcane numbers (together with the assumptions that we have already discussed) span the space of all parameters, and hence it is easy to present versions that are simpler (like the age of the Universe, t_0, or the density of some component, like $\rho_{\rm M}$) in more familiar units. Let us highlight a few variants of quantities that are useful in describing our Universe (table III), in the hope that some of them may catch on! Further examples along these lines can be found in the paper "Cosmic Mnemonics" [7].

With enough effort, it is easy to find numerological coincidences. One should obviously be skeptical about claims of significance for such things though! For example, from the table we see that the number of particles in the observable Universe (mostly photons) is about α^{-42} (where α is the fine-structure constant), and additionally in the standard model, the Earth forms at a redshift corresponding to $z = 0.42$. These facts could be used to suggest a link with Douglas Adams' universal answer.

4. – Information in the SMC

Cosmological information comes from many sources. However, at the present time, the CMB dominates the constraints on the SMC. This "era of precision cosmology" can be seen through plots of the current status of the power spectra coming from the CMB. Figure 1 shows the TT (which dominates the information) and TE (which is catching up) spectra, while fig. 2 shows the EE (now also impressive) and BB (still in its infancy) spectra. The grey line is the 6-parameter ΛCDM model fit to the *Planck* TT data, and one can see how well it matches the other power spectra.

TABLE III. – *Variants on the numbers that describe our Universe.*

Characteristic scale on the CMB sky, $\theta_* \simeq 0.6^\circ$ (think eclipse!)
Radius of observable Universe $\simeq 400\,\mathrm{Ym}$
Age of the Universe $t_0 \simeq 5$ trillion days $\simeq 5 \times 2^{200} t_{\mathrm{Pl}}$
Age of the Universe is triple the age of the Earth, $t_0 \simeq 3\,t_\oplus$
$H_0\,t_0$ is slightly less than 1, and $H\,t$ will be unity in about 1 billion years
H_0 will asymptote to the value $56\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ in the far future
Cosmological constant, $\Lambda \simeq 10^{-35}\,\mathrm{s}^{-2}$ ("ten square attohertz")
Critical density, ρ_{crit}, corresponds to 5 proton masses per cubic metre
Density ratios, $\Omega_{\mathrm{c}}/\Omega_{\mathrm{b}} \simeq 2\,\Omega_\Lambda/\Omega_{\mathrm{m}} \simeq 5.3$
Density parameter for photons, $\Omega_\gamma \simeq \alpha^2$
Variance of density in spheres is unity at about 9 Mpc (no h^{-1})
Amplitude of position-space density perturbations on Hubble scale, $\sigma \simeq 6 \times 10^{-6}$
Temperature at last scattering epoch $T_{\mathrm{CMB}} \simeq 3000\,\mathrm{K}$ (think M giant!)
Age at last-scattering epoch, $t_{\mathrm{rec}} \simeq 370\,\mathrm{kyr}$
Age at reionisation, $t_{\mathrm{reion}} \simeq 600\,\mathrm{Myr}$
Number of particles in observable Universe $\simeq \alpha^{-42}$

When we add up the total $(S/N)^2$ from the *Planck* CMB power spectra, over the part of the sky conservatively believed to be free of foreground emission, we find that the *Planck* TT, TE and EE measurements together correspond to about $900\,\sigma$. Adding the higher multipole measurements from ACTPol and SPTPol means that today's CMB power spectrum determinations together represent more than $1000\,\sigma$ of detection. If we were skeptical about the success of the SMC, then we should take note that it requires just a few simplifying assumptions and seven free parameters to fit this huge amount of information — quite a remarkable achievement.

How does the constraining power of the CMB work? The simple answer is that it just depends on the number of modes that are measured, where the mode amplitudes $a_{\ell m}$ come from expanding the sky as $T(\theta,\phi) = \sum_{\ell m} a_{\ell m} Y_{\ell m}$. Since the anisotropies are Gaussian, then each $a_{\ell m}$ gives just a little bit of information about the expectation value of the power in the $a_{\ell m}$s, C_ℓ (or equivalently, the variance), and the total constraining power is just about counting the number of modes. In more detail, if we have a cosmic-variance-limited experiment, with $\Delta C_\ell = \sqrt{2/(2\ell+1)}C_\ell$, then the total signal-to-noise ratio in the power spectrum is

$$(1) \qquad (S/N)^2 \equiv \sum_{\ell=2}^{\ell_{\max}} (C_\ell/\Delta C_\ell)^2 = \frac{1}{2}\sum_{\ell=2}^{\ell_{\max}} (2\ell+1) = \frac{1}{2}\left[\ell_{\max}(\ell_{\max}+2) - 3\right] \simeq \ell_{\max}^2.$$

But since the number of modes is just $\sum_{\ell=2}^{\ell_{\max}} \sum_{m=-\ell}^{+\ell}$, then this means that the total $(S/N)^2$ is just half the number of modes.

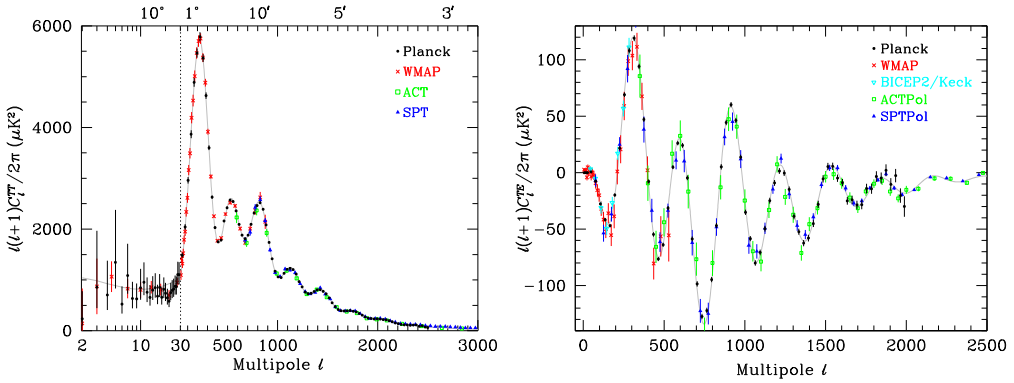


Fig. 1. – CMB temperature anisotropy power spectrum (left) and temperature-polarisation cross-power spectrum (right), from *Planck, WMAP*, BICEP/Keck, ACT and SPT (see ref. [8] for full references). This demonstrates the current precision with which these power spectra have been measured.

To the extent that through *Planck*, we have measured all the modes out to $\ell \simeq 1500$ (over a large fraction of the sky), and the damping plus foregrounds means that we can not go far beyond $\ell \simeq 3000$ (say) for primary CMB anisotropy measurements, then we are a good way through gathering all the information we can get from C_ℓ^{TT}. But what about polarisation? Assuming for the moment that C_ℓ^{BB} is negligible, then the existence of both C_ℓ^{TE} and C_ℓ^{EE} would seem to confuse matters. But really the situation is simply that we can measure the scalar field E, in addition to T, for each pixel. And hence, provided that we measure both C_ℓ^{TE} and C_ℓ^{EE} out to some $\ell_{\rm max}$, then we have exactly twice as much information as we would obtain from C_ℓ^{TT} alone. This means that the

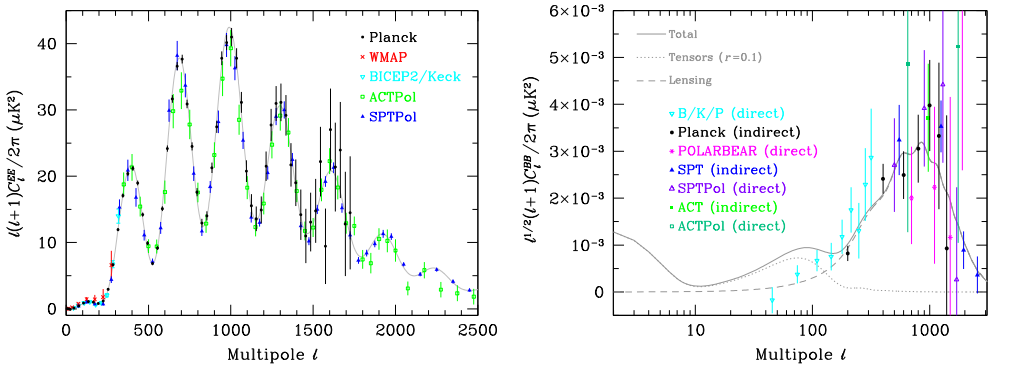


Fig. 2. – Polarisation EE (left) and BB (right) power spectra from several experiments (see ref. [8] for references). The BB spectrum here is scaled by a power of ℓ that makes it possible to see all three of the expected peaks (from reionisation, recombination and lensing).

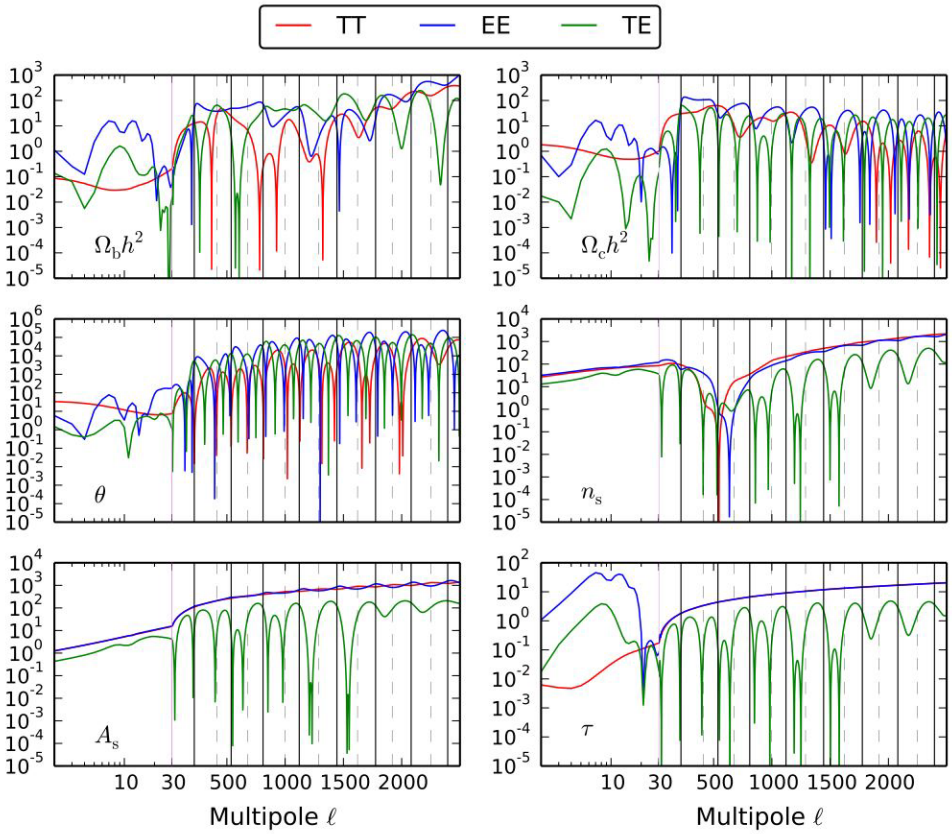


Fig. 3. – Square of signal-to-noise ratio for each multipole, for a cosmic-variance-limited experiment covering half the sky. The results for the TT, TE and EE power spectra are shown on panels for each of the six parameters of the SMC, with a logarithmic horizontal axis for the first 30 multipoles. Vertical lines mark peaks and troughs in the TT power spectrum. These panels show the complex mapping of power spectrum constraints onto parameters over different multipole ranges.

total "information content" can be defined to be just a count of the number of modes probed, in any of the CMB fields.

Of course not all information is equal. For example, *any* large-angle BB measurement would provide us with an entirely new kind of information, enabling us to determine an additional parameter (r, the tensor-to-scalar ratio) that is otherwise hard to constrain. Moreover, it is well known that adding polarisation data helps to break some parameter degeneracies. So we should like to know how the parameter constraints map onto the power spectrum modes. The proper way to discuss this is through the Fisher matrix, which includes derivatives of the power spectra with respect to the cosmological parameters; this is demonstrated in fig. 3 and discussed fully in ref. [9]. One can see that some multipole ranges are particularly important for some parameters, and as we go up in ℓ, so

that new peaks or troughs are included, the constraining power can change dramatically.

To focus on one example, the behaviour of the A_s panel is simple — if there was only an overall normalisation to measure, then the constraints would just come from the S/N (and mode counting) expression given in eq. (1), with polarisation giving equal constraining power to temperature. In that sense, A_s is a "linear" parameter, since there is a simple relationship between its total S/N and the parameter constraint. However, if the dependence is less trivial, then the relationship is "non-linear", and hence the way that the total $(S/N)^2$ is shared out among the parameters is more complicated. A good example is θ_*, which determines the amount by which the power spectra can be slid left and right in multipole — this can be determined to a great precision (because of the relative sharpness of the acoustic peak structure), which is why this is the best determined parameter of the SMC today. In fact the total S/N in θ_* from *Planck* is around 2500, which is considerably more than the *total* S/N in the power spectra!

CMB polarisation has yet to become particularly constraining for the parameters of the SMC. But that situation will change as new experiments add modes, doubling at least (or more, since high-ℓ E-mode measurements are not as limited by foregrounds) the information achievable from temperature alone, and providing specifically useful degeneracy-breaking capability. The discussion above can be extended to the BB spectrum, as well as the lensing spectrum $C_\ell^{\phi\phi}$ (which comes from the temperature trispectrum, and gives an additional field, ϕ). This approach is useful for discussing future experimental constraints, and how they map onto parameters. But one thing it tells us is that eventually we shall run out of CMB information. This is essentially because the CMB information is almost entirely restricted to two dimensions. The same thinking can also inform discussions of more ambitious attempts to extract the much larger amount of information contained in 3d surveys — around $(c\,k_{\rm max}/H_0)^3$, if we can get all the k modes down to scales $k_{\rm max}$ [10]. This means that in principle we could one day measure enough modes to give $\gg 10^6\,\sigma$ of power detection.

5. – The venerableness of the SMC

It is clear that the standard model of cosmology is now well established. So well established in fact that a great deal of effort in modern cosmology is directed towards trying to find extensions to the model. For example, searching for evidence that the dark energy is evolving or that more than two parameters are needed to characterise the perturbations. Such searches for "physics beyond the standard model" makes one think of similar endeavours to find evidence to extend the Standard Model of particle physics. When this sort of thing comes up, it has been traditional for cosmologists to claim that the SMC is relatively young and still in an exploratory stage — so it is *nothing like* the chasing of $3\,\sigma$ effects that appears to have motivated much experimental particle physics for decades. But in fact the SMC is actually quite long in the tooth itself by now!

So how old is the SMC? Certainly if one goes back a dozen years to a previous overview by this same author [11], one finds little that has changed. The model is much more precisely determined of course, but all the ingredients are already in place. Indeed,

one can go back earlier, *e.g.* to the paper "What Have We Already Learned from the Cosmic Microwave Background?" (also known as "What Has the CMB Ever Done For Us?") written in 1998 [12], and find that the basic picture is just the same. In fact many expositions of the history of cosmology state that the model became established with the detection of cosmic acceleration in 1998. Of course that was an important part of the story, but I think it is clear that what we now call ΛCDM was *already* the best-fitting model when the supernova data came in and confirmed it — to the extent that even the most skeptical cosmologists had to take Λ seriously. Many papers had already pointed out before 1998 that a collection of results pointed to a flat Λ model being the best way of extending what had previously been called "standard CDM" (or sCDM), *i.e.* a CDM-dominated model with $\Omega_{\rm M} = 1$ and an initial power spectrum of exactly the Harrison-Zeldovich-Peebles form ($n = 1$). Among these results were: the need for more power on large scales (to match galaxy clustering data); the fact that most measurements of the density parameter tended to give $\Omega \lesssim 0.3$; that few measurements of H_0 gave values in the $\simeq 50\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ range that were needed to make the age of the Universe old enough for the stars within it; that the amplitude of density perturbations from the *COBE* satellite's CMB anisotropies pointed towards adiabatic perturbations in a Λ-dominated (rather than open) model; and that indications from smaller-scale CMB anisotropies were suggesting an acoustic scale consistent with flat geometry [13, 14].

Several of these arguments were compiled in two essays in 1995, one by Ostriker and Steinhardt [15] and the other by Krauss and Turner [16]. Not everyone was convinced of course, and some nostalgic theorists still tried to cling to the Einstein-de Sitter elegance of sCDM [17, 18]. However, the writing was on the wall, and of all the flavours to add to sCDM, it was apparent by the mid-90s that ΛCDM gave the best fit (even if you did not necessarily like it!). Indeed it is possible to find earlier papers pointing to this model being preferred by a combination of data — and here the 1990 Nature paper by Efstathiou, Maddox and Sutherland [19] is a particular standout. That is not to say that there were not papers proposing quite different models at the same time, but just that the currently understood SMC was already there in the early 90s, with reasons to believe that it provided the "best-buy" cosmology.

What this means is that the SMC is older than most people appreciate — something like a quarter of century old, making it more than half the age of the SM of particle physics! As an indication of just how long ago that was, in the early 90s we were using dial-up modems to connect with the internet, the main browser was (the pre-Netscape) Mosaic and the world's first text message was being sent!

6. – Tensions

The idea of "tensions" has already come up. So let us take that particular bull by the horns right now. There are several different minor chinks in the armour of the SMC that are pointed to by various researchers. A list of some of them is given in table IV. None are sufficiently significant to call them an actual discrepancy, hence the use of the

TABLE IV. – *List of claimed tensions (not complete!). Are any of these of consequence?*

The amplitude σ_8 between CMB and cluster abundance
Galaxy cosmic shear *versus* CMB constraints on $\sigma_8\Omega_{\rm m}^{0.5}$
H_0 between traditional direct estimates and indirect CMB estimates
Planck versus WMAP TT power spectra
Planck high-ℓ *versus* low-ℓ data
Preference for $A_{\rm L} > 1$ (apparent lensing effect) in *Planck* data
Small-scale galaxy formation controversies

word "tension"([2]). What should a skeptic make of an apparent disagreement between different data sets at the 2–3 σ level?

Well, let us remember that today's CMB data contains more than 1000 σ worth of detection. We can ask how many 3 σ results there are in 1000 σ. Since signal-to-noise ratios add quadratically, the answer is the number of times that 3^2 goes into 1000^2, and the answer is *more than 100000*. So why is there so much focus on specific issues that are barely at the 3 σ level?

Of course part of the answer is that we should not accept that the SMC is the last word, but should keep an open mind to other possibilities. There is also a strong motivation to look for flaws that require revisions, since we are all hoping to find fundamentally new physics by further confronting the SMC with cosmological data. And there *may* be evidence of such things lurking in low S/N disparities. The trick of course is to find the hints of disparity that grow from mere "tensions" into genuinely significant differences.

However, concentrating on a few of the > 100000 potential 3-ish σ effects seems like a misplaced kind of skepticism. Trying to find minor deficiencies in conventional wisdom seems to me to be a bit like chasing conspiracy theories. In any situation, you can always find *something* that does not seem to make sense — but you should be assessing the evidence carefully, bearing in mind the context. Here the context is: 1) that the model (the SMC) fits a number of observational phenomena very well; 2) that some of the uncertainties are of a systematic rather than statistical nature; and 3) that there are a very large number of potential tensions that could be selected from the $> 1000\,\sigma$ of measured information.

Let us look at a couple of aspects of the history of the development of the SMC in order to see if there are any lessons we can learn. Although the SMC was already in place by the early 1990s, there are some observations that have changed considerably since that time. In particular, determinations of the age of the Universe (from estimating the ages of the oldest globular clusters) and determinations of the baryon abundance (coming from Big

([2]) I am reminded of the skeptics' joke that when "alternative medicine" works it is just called "medicine". In the same way, if a "tension" was actually big enough to be significant, then it would be called a "difference".

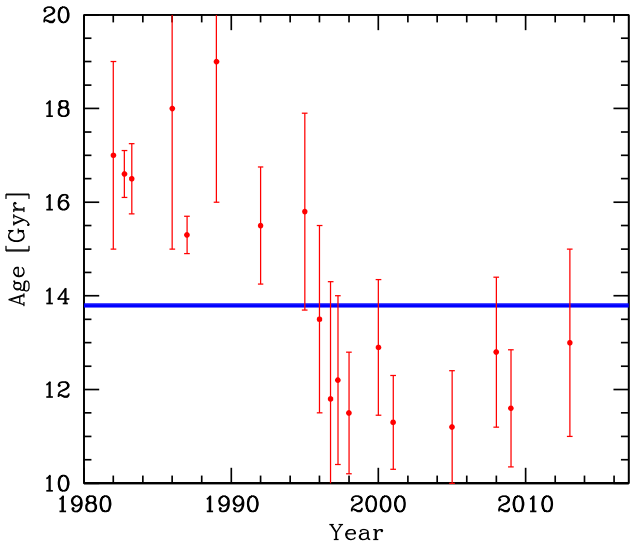


Fig. 4. – Estimated ages of the oldest globular clusters, taken from a representative set of review-like articles. The preferred value dropped dramatically in the mid-1990s. The blue band is the current best-fitting value for the age of the Universe, based on the 2015 *Planck* data (including uncertainty).

Bang nucleosynthesis) changed in value in the mid-1990s. Figure 4 shows the situation for cluster ages. The values plotted come from a representative selection of papers by several of the groups working on this problem at that time. In the early 90s the oldest clusters were stated to be perhaps 17 billion years old, with lower limits at around the 15 billion year level. However, in hindsight it is clear that those estimates were incorrect because they were dominated by systematic uncertainties. As other cosmological measurements improved, and it became clear that the Universe had an age that was probably no more than 14 billion years, the cluster ages were revised to become consistent. One might imagine that the change could be traced to one particular effect that was fixed — but that really is not the case. Instead there were several tweaks made over the years, most of which had the same sign and resulted in the ages of the oldest clusters coming down to around 12 billion years. The situation is really not very satisfying! But I suspect this is often the way things are when the uncertainties are to do with assumptions and approximations in the analysis, rather than just being statistical.

The situation with baryon abundance is fairly similar. But here it is harder to make a plot of the values, since often there were no clear errors given! Instead it was common to write down some feasible range for the baryon-to-photon ratio, which was bracketed by different light-element abundances (with little effort made at that time to designate 95% confidence ranges, or give $\pm 1\,\sigma$ values, or the equivalent). Despite this difficulty in interpreting the uncertainties in the old results, what is clear is that the preferred value

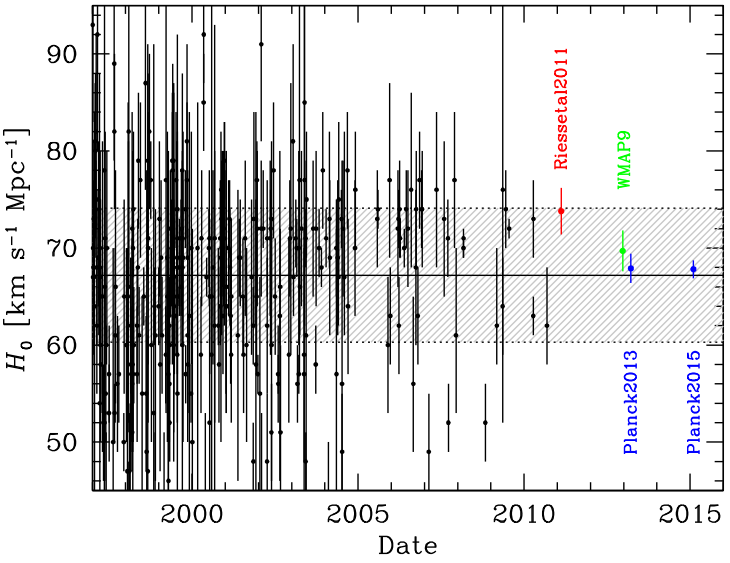


Fig. 5. – Estimates of the Hubble constant between 1997 and 2010, compiled by John Huchra (black points [20]). These values were combined into a best estimate (using practices for combining nuclear physics data) by Pritychenko [21], with this value and uncertainty represented by the hatched region. The new value published by Riess *et al.* in 2011 [22] is shown by the red point. The estimate derived from the full *WMAP* data set is shown in green [23]. The values obtained from the 2013 [6] and 2015 [2] *Planck* data releases are the two blue points.

around 1990 was something like $\Omega_b h^2 \simeq 0.012$, while 10 years later it was around 0.022. The change corresponded to a lot of σ (with whatever value of uncertainty you used). Again, there was not one reason for this change, but probably a list of things contributed to the increase (the availability of damped Ly α systems for deuterium abundance measurements being part of it).

The tension that perhaps attracts the most current attention is that some of the most recent and precise values for the Hubble constant, determined using standard distance-estimation techniques, appear higher than the value determined from the SMC parameters that best fit CMB data. Figure 5 shows a large collection of published H_0 determinations from the period 1997–2010, including the error bars. The horizontal band gives an average of these data (using procedures developed to deal with apparently disparate nuclear and particle physics data [21]). The newer value from Riess *et al.* (2011) [22] is indicated, together with values coming from *WMAP* and *Planck*. Other recent values could be added, but would not change the basic picture. Placed in this context we can see the systematics-dominated history of attempts to directly determine H_0 (which certainly goes back earlier than 1997!). We have to assume that most sets of authors of historical values believed that their published uncertainties were a fair representation of their confidence in the data. And yet it is clear that at any given time the errors were

being underestimated by most groups. Perhaps the situation is genuinely different now, and the new (smaller) error bars are correct. But it is hard not to be skeptical.

In any case, we know that measurement techniques are continuing to improve, and that if this CMB-*versus*-direct-determination tension in H_0 arises from some genuinely new physics, then the statistical confidence in the differences will grow. Time will tell.

Turning to the example of *Planck*'s large-scale *versus* small-scale constraints, there are some additional issues for the skeptic to keep in mind. As discussed in the Planck Collaboration Intermediate LI paper [24], this situation is not as simple as it might at first seem. There are certainly parameter shifts when one considers *Planck* low-ℓ *versus Planck* high-ℓ data, and these shifts may seem to be at the 2–3 σ level. However, the parameter space has six (or maybe five, considering that τ is hardly measured) dimensions, with many other directions in this space corresponding to particular parameter combinations. So one will see a $2\,\sigma$ deviation in *some* direction more than 5% of the time, and hence it is necessary to take into account the whole parameter set when assessing this kind of tension. When that is done, the differences of low-ℓ *versus* high-ℓ parameters have a probability to exceed above 10% (*i.e.* nothing to write home about). On top of that, it is unclear how one is choosing the angular scale to look for a split. And there is also a difficulty in assessing how unlikely it might be for a data excursion to map onto a parameter shift. The conclusion is that *Planck* and *WMAP* are in spectacular agreement where they overlap, and the shifts seen at higher multipoles are just about as big as you would expect the shifts to be. That is not to say that there might not also be problems with some of the foreground modelling, or indeed some physics missing from the SMC — it is just that the data do not *require* these things at the moment.

7. – Anomalies

The other word that is much heard when discussing the CMB is "anomalies". What is meant here is a feature (at a low level of significance) that appears to be unexpected in the SMC, pointing to perhaps some kind of non-Gaussianity or breaking of statistical isotropy. There are several examples that have been suggested over the years, with different researchers claiming importance for one or other. Table V gives a partial list. It seems extremely hard to believe that *all* of these are pointing to deficiencies in the conventional picture. One should be skeptical about each of them, particularly because of the issue of *a posteriori* statistics. This issue is one that causes enough debate among cosmologists, that I am going to discuss it in some detail.

The problem is that all of these "anomalies" had their statistical significance assessed *after* they were discovered. Hence, in order to fairly determine how unlikely they are, it is necessary to consider other anomalies that may have been discovered instead. Statisticians call this the "multiplicity of tests" issue, which I think is the most helpful way to think about it.

Let us take the so-called CMB Cold Spot as an example (as shown in the left-hand panel of fig. 6). The probability of finding a cold region of exactly this size and shape in exactly this particular direction is obviously vanishingly small. No one would consider

TABLE V. – *List of claimed temperature anomalies. This is not intended to be a complete list (and several of these anomalies are related to each other).*

Low quadrupole and other low-ℓ modes
Deficit in power at $\ell \simeq 20$–30
Low variance
Lack of correlation at large angular scales
The Cold Spot
Other features on the CMB sky
Hemispheric asymmetry
Dipole modulation
Alignment of low-order multipoles
Odd-even multipole asymmetry
Other features in the power spectrum

such a calculation to be useful, and at the very least would appreciate that the spot could have been found in any direction — hence to assess the significance one could look in simulated skies for similarly extreme cold spots that occur anywhere. The probability determined in this way then becomes of order 0.1%. However, a specific scale was chosen for the spot, or to be more explicit, a filter of a particular shape and scale was chosen. It turns out that the Cold Spot is not very extreme if a purely Gaussian filter is used, but is pulled out at higher amplitude by a "compensated" filter (*e.g.* what is often called a "Mexican hat") with a scale of about 5°. This means that one should marginalise over the scale (within some reasonable bounds) and over a set of potential filter shapes (that one might have chosen) as well. On top of all this it is obviously clear that one needs to consider *hot* spots as well as cold spots (this may seem so self-evident that it does not need to be stated, but in fact several papers have *only* assessed the significance of cold spots). And the situation is more complicated than that, since if there had been a fairly conspicuous pair of neighbouring spots, or a hot spot diametrically opposite a cold spot, or even a triangle of spots, then one might equally well have been writing papers about the anomalous feature that was discovered.

The point is that in each Hubble patch (with the CMB sky being an independent realisation of the underlying power spectrum), there will be features on the sky or in the power spectrum, that appear anomalous. One has to consider the set of potential anomalies in each patch in order to assess whether a feature is extreme enough to get excited about. In practice 2–$3\,\sigma$ anomalies go away when you marginalise over these possibilities, but $\gtrsim 5\,\sigma$ anomalies would remain anomalous after maginalisation.

A criticism of this way of thinking is that it is just *too* skeptical! The argument is that if you try hard enough to marginalise over possible tests then you can make *anything* appear to be insignificant. I do not think this is true, since you have to be

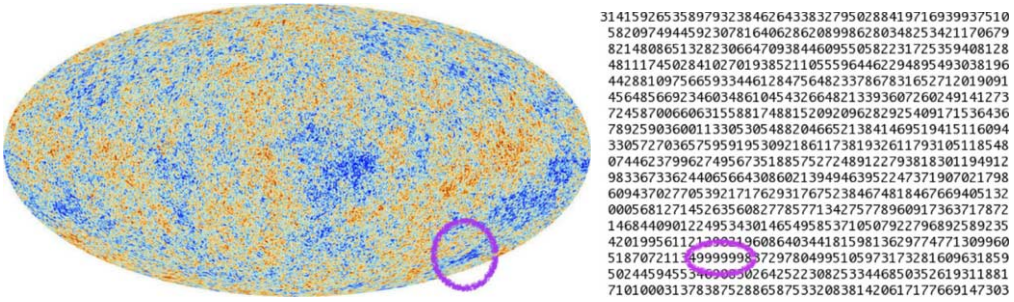


Fig. 6. – Left: map of the CMB sky from *Planck*, with the position of the so-called "Cold Spot" indicated. Right: the first 900 digits of π, showing the "hot spot" of six 9s (also known as the Feynman point).

reasonable here (as in any assessment of statistical evidence, where there is always some subjectivity). And I stand by the claim that it is hard to make $5\,\sigma$ effects go away, while 2–$3\,\sigma$ effects that are subject to a posteriori statistics should always be viewed with extreme skepticism.

Another way to look at this is to make an analogous study of something that you are confident is genuinely random. This was done in the paper with Dr. Frolop [25], comparing the CMB anomalies with patterns in the digits of π. Several examples are given there, but let us just pick one. As illustrated in the right panel of fig. 6, there are six consecutive occurrences of the digit "9" at the 762nd digit of π. Assuming that the digits are random, a simple calculation (considering the number of ways of placing the run of 9s in the first 762 digits, and the number of ways of picking the other 756 digits) gives a probability of $756/10^6$. Of course the run of six numbers did not have to be 9 (even although you could consider 9 to be special because it is the highest), and hence the probability for any run of six is 10 times greater — but that is still less than 1%! So why does this not shake our faith in the digits of π being random? The answer is that a posteriori statistical effects can be subtle. In this particular case, we would have found an equal probability for obtaining a run of *five* 9s at an earlier digit, or *four* 9s even earlier, and so on. When including this hidden "multiplicity of tests", the probability becomes of order 10%, *i.e.* not small enough to decide that there are messages written in the digits of π! And in addition to all of that, there are other patterns that might also have been remarked upon if they had been found, a run of "123456" for example, or an alternating series like "9090909". Perhaps these are not quite *as* striking as six 9s, but they should get *some* weight in considering the set of tests, and hence in the assessment of the significance of the anomaly. I am convinced that if one went to the trouble of looking for such things, there would appear to be something conspicuous in most chunks of π (in every 1000 digits, say) — just like there are some apparent "anomalies" in every Hubble patch's CMB sky.

Despite all these words of caution, let me add that one should *still* continue the search for anomalies, since any genuinely significant large-scale oddities could be signs of exciting new physics (*e.g.* see discussion in ref. [26]). And of course sometimes $3\,\sigma$

things will become $5\,\sigma$ when more data are included — so it is worth continuing these investigations. A problem is that the large-angle temperature field has already been well mapped, and those data are now limited by cosmic variance. So the only way to make progress is to include new data, such as from CMB polarisation [27]. What would be particularly good would be to find some kind of natural explanation for an anomaly, with no (or very few) free parameters, which also makes a clear prediction for some *new* observable, such as polarisation, lensing, or 21 cm observations. If there is such a prediction, and a $3\,\sigma$ result is found, then that really would mean 99.7% confidence.

8. – The nature of skepticism

I have said a lot about being skeptical — but what does that really mean? What I am talking about here is the concept philosophers might call "scientific skepticism", which involves questioning assertions that lack empirical evidence. I believe this to be a fundamental part of scientific inquiry. It can be summed up through the phrase "extraordinary claims require extraordinary evidence" (popularised by Carl Sagan) — and obviously that applies well to cosmology, through its grandiose themes, just as it does to pseudo-science. Science is not completely mechanical and dispassionate, since it includes speculation and creativity as part of the process of development — but that is not the same as accepting every new idea that comes along. At the other end of the spectrum, it is also important not to fall into the trap of "denialism", *i.e.* adopting a position that rejects every claim *even if* there is good evidence to support it (like climate change, or dark matter perhaps).

To be a bit more explicit about skepticism, let me pick the writings of a particular modern philosopher, namely Mario Bunge, who has written extensively on the topic of scientific epistemology. Among other definitions, he describes how any authentic science must include "changeability, compatibility with the bulk of the antecedent knowledge, partial intersection with at least one other science, and control by the scientific community" [28]. These ideas give a little more content to the notions of hypothesis testing, falsifiability, parsimony, etc., that we learn about in school. And they make clear that the skeptical approach is central to the establishment and evolution of scientific ideas.

As examples of topics that fail to meet these criteria and land up in the pseudo-science category, Bunge lists "astrology, alchemy, parapsychology, characterology, graphology, creation 'science', 'intelligent design', Christian 'Science', dowsing, homeopathy, and memetics". However, Bunge also states that "cosmology is still rife with speculations that contradict solid principles of physics"! He says that for good reason — the SMC lives within the domain of "physical cosmology" and has passed a wide array of tests, but, on the other hand, the most theoretical aspects of cosmology are indeed in an entirely different conjectural realm. Hence it is important to separate the concrete parts of modern cosmology (the answers to the "what" questions) from the areas where we are still speculating wildly (and trying to find answers to the "why" questions).

9. – Beyond the SMC

We should all like to understand where the whole Universe comes from, or explain away the dark matter and dark energy. Speculation is certainly good, but *believing* your speculation (before it has passed any tests) is bad science! The correct approach should be to investigate the consequences of your idea and try to determine if there are definitive predictions that can be confronted with data.

I feel that there is a kind of malady that infects some cosmologists, where pretty much any outlandish and unorthodox idea is considered at the same level as the conventional picture — rather than giving it a higher degree of skepticism, like all extraordinary claims deserve. Perhaps part of the blame here is that modern physics in general, and the SMC in particular, contain some fairly bizarre-sounding concepts. We teach students about quantum mechanics and black holes, that we can build a model for the whole observable Universe, that there are hypothetical particles that dominate all matter and that a negative-pressure fluid is driving the cosmic acceleration. So perhaps students start to think that any hare-brained scheme is equally worth pursuing?

I can not shake the feeling that a good dose of skepticism would help keep things in perspective.

An example of this is inflation (see [29] and [30] for discussions). It is undoubtedly an appealing idea, and there is a great deal of circumstantial evidence to support it — so I think it is entirely reasonable to be a *fan* of inflationary cosmology. But since inflation is really a framework rather than a model, we can not assert that any of the observations actually prove that inflation is correct([3]). It seems reasonable to assume that whatever picture turns out to describe the early Universe, and generates the perturbations, it will contain some of the features of the current inflationary paradigm. But I do not think we can proclaim that we know that it will include *all* the ingredients — not until we have some more direct evidence.

However, one of the problems with assessing the merits of inflation is that there is not a good alternative. Sure, there might be some ideas suggested as counter-proposals, but they tend to seem much more *ad hoc*, or create more problems than they solve, or have predictions that are less well developed. And the same issue applies more broadly across other "alternative" theories. The SMC has been developed over decades and the calculations are relatively straightforward (involving Gaussian perturbations, linear theory, well-understood physics, etc.) — but there is no reason to expect the same to be true for some unconventional new idea. So if an alternative is proposed, then it is not trivial to determine whether it can match the precision tests of the SMC. We just have to be a little patient until the calculations can be done accurately enough.

Despite the need to be open to alternatives, when there are clear predictions, it is

([3]) I have spent a long time trying to determine just what we have learned regarding inflation and hence where we are on the spectrum of proof for this concept, and I have come up with the following statement (which I challenge anyone to disagree with): "*something like inflation is something like proven*".

still important to be skeptical if they just do not fit the data. As an example, we call the dominant form of energy in the Universe "dark energy", as though its properties were mysterious and unknown — and a huge amount of effort is going into measuring its equation of state (w as a function of redshift) with increasing precision. But the reality is that all measurements so far are consistent with this component being simply vacuum energy with $w = -1$. I have heard people say that it is much more likely to be a model with $w \neq -1$, since $w = -1$ has zero probability! But really, there is no sensible model that gives a definite prediction other than pure vacuum, and so we are left with the notion that there is just a universal constant, Λ, that gives a small (but non-zero) energy density to empty space.

Another example is dark matter. It is obvious that an alternative explanation for galaxy rotation curves *might be* that we can modify our theory of gravity. And there have been several suggestions along those lines (see, *e.g.*, [31]). However, the evidence for dark matter comes from a lot more than rotation curves of galaxies, *e.g.* the depth of cluster potential wells and measurements of gravitational lensing. But in fact the most robust evidence for dark matter comes from the CMB anisotropies — there is *no* model for fitting the power spectra that does not include a lot more CDM than baryonic matter. Here we have a choice between abandoning GR (or even Newtonian gravity) or just imagining that there is a component of matter that is not very shiny! Even without guidance from data, it seems fairly clear that the parsimonious explanation is to have a particle that is like a heavier version of the neutrino. But the skeptic should come down more heavily on the side of CDM when comparing with clustering, lensing and (particularly) CMB data.

A related issue is the evaluation of some of the small-scale puzzles associated with galaxies. It has become common to propose models ascribing these to some property of the dark matter (just strong enough to detect, without messing up the SMC predictions entirely). However, galaxy formation is a complicated business [32], involving non-linear complexity, hydrodynamics, feedback processes, etc. Since we *know* that we do not fully understand baryonic physics, we should be skeptical about assertions that some new property of dark matter has been discovered because of indications coming from non-linear scales.

Despite the examples given here, the SMC is in no sense a complete model, and there will surely be several additions eventually. Table VI lists some potential questions relating to physics *beyond* the SMC. Will one of these lead to the next breakthrough? Right now the path to progress is not at all clear. Maybe it will turn out to be something else entirely, something unexpected and outlandish — but only if the evidence strongly supports that.

10. – Conclusions

I have given an overview of the current status of the Standard Model of Cosmology, the SMC, and stressed how important it is to maintain a healthy level of skepticism when assessing the successes of this model, and in evaluating the merits of extensions to it.

So when should one be skeptical and when not? That is the trick of course! Obviously

TABLE VI. – *Physics beyond the SMC. Which of these questions will turn out to be fruitful?*

Where did the parameters come from?
Did inflation happen?
Can we explain the value of Λ?
Why is $\Omega_{\rm c}/\Omega_{\rm b} \simeq 5.3$?
Are any anomalies or tensions worthy of attention?
Can we detect primordial gravitational waves?
Can we detect primordial non-Gausianity?
Are there missing ingredients to the SMC?
Will neutrino properties be measurable?
Can we predict reionisation from first principles?

the aim is to be *right*, and it is never clear how to forecast the future. There was a time when hardly anyone believed that the solution to the solar neutrino problem lay in the properties of neutrinos — but a small number of people got it right before the rest of us. Similarly, some people saw that ΛCDM fit most of the data while many others in cosmology were working on things like "open CDM" or "mixed dark matter". Since no practicing cosmologist believes that the current SMC will be the last word on a statistical description of the Universe, then there are surely developments that are yet to come. The goal is (somehow) to pick the 2–3 σ effects that grow to be important parts of the model — and by implication, part of this process involves ignoring most of the other claims for chinks in the SMC's armour.

There were times in the history of cosmology when it was fairly clear what directions were going to be fruitful for pursuing calculations or observations. I think it is not just that we have the benefit of hindsight — it really was the case that at one time studying hot *versus* cold dark matter was obviously a good idea, and at some other time developing the theory of CMB anisotropies or building experiments to probe degree-scale anisotropies were clearly worthwhile. However, right now it is not at all obvious where cosmology is going next.

This means that this is either the worst time or the best time to be a cosmologist! If you have a good idea (and it turns out to be right) you could find yourself on your own making the next major contribution to our understanding of the whole of the Cosmos.

* * *

I acknowledge many enjoyable discussions about some of the issues contained in this contribution with members of the Planck Collaboration, including those who were at UBC, particularly Dago Contreras, Ali Narimani and Jim Zibin.

REFERENCES

[1] PLANCK COLLABORATION I, *Astron. Astrophys.*, **594** (2016) A1.
[2] PLANCK COLLABORATION XIII, *Astron. Astrophys.*, **594** (2016) A13.
[3] FIXSEN D. J., *Astrophys. J.*, **707** (2009) 916.
[4] FIXSEN D. J., CHENG E. S., GALES J. M., MATHER J. C., SHAFER R. A. and WRIGHT E. L., *Astrophys. J.*, **473** (1996) 576.
[5] ZIBIN J. P. and SCOTT D., *Phys. Rev. D*, **78** (2008) 123529.
[6] PLANCK COLLABORATION XVI, *Astron. Astrophys.*, **571** (2014) A16.
[7] SCOTT D., NARIMANI A. and PAGE D. N., *Phys. Canada*, **70** (2013) 258.
[8] SCOTT D. and SMOOT G., *Chin. Phys. C*, **40** (2016) 100001.
[9] SCOTT D., CONTRERAS D., NARIMANI A. and MA Y.-Z., *JCAP*, **6** (2016) 046.
[10] MA Y.-Z. and SCOTT D., *Phys. Rev. D*, **93** (2016) 083510.
[11] SCOTT D., *Canadian J. Phys.*, **84** (2006) 419.
[12] LAWRENCE C. R., SCOTT D. and WHITE M., *Proc. Astron. Soc. Pacific.*, **111** (1999) 525.
[13] SCOTT D. and WHITE M., *The Existence of Baryons at $z = 1000$*, in *Proceedings of the CMB Anisotropies Two Years after COBE: Observations, Theory and the Future*, edited by KRAUSS L. M. 1994, p. 214.
[14] PIERPAOLI E., SCOTT D. and WHITE M., *Mod. Phys. Lett. A*, **15** (2000) 1357.
[15] OSTRIKER J. P. and STEINHARDT P. J., *Nature*, **377** (1995) 600, astro-ph/9505066.
[16] KRAUSS L. and TURNER M. S., *Gen. Rel. Grav.*, **27** (1995) 1137, astro-ph/9504003.
[17] WHITE M., SCOTT D., SILK J. and DAVIS M., *Mon. Not. R. Astron. Soc.*, **276** (1995) L69.
[18] WHITE M., VIANA P. T. P., LIDDLE A. R. and SCOTT D., *Mon. Not. R. Astron. Soc.*, **283** (1996) 107.
[19] EFSTATHIOU G., SUTHERLAND W. J. and MADDOX S. J., *Nature*, **348** (1990) 705.
[20] `Url: https://www.cfa.harvard.edu/~dfabricant/huchra/hubble.plot.dat`.
[21] PRITYCHENKO B., *A nuclear data approach for the Hubble constant measurements*, *Eur. Phys. J. Web of Conf.*, **146** (2017) 01006.
[22] RIESS A. G., MACRI L., CASERTANO S., LAMPEITL H., FERGUSON H. C., FILIPPENKO A. V., JHA S. W., LI W. and CHORNOCK R., *Astrophys. J.*, **730** (2011) 119.
[23] BENNETT C. L. *et al.*, *Astrophys. J. Suppl.*, **208** (2013) 20.
[24] PLANCK COLLABORATION INTERM. LI, *Astron. Astrophys.*, **607** (2017) A95.
[25] FROLOP A. and SCOTT D., ArXiv e-prints, (2016), 1603.09703.
[26] CONTRERAS D., HUTCHINSON J., MOSS A., SCOTT D. and ZIBIN J. P., *Phys. Rev. D*, **97** (2018) 063504.
[27] CONTRERAS D., ZIBIN J. P., SCOTT D., BANDAY A. J. and GÓRSKI K. M., *Phys. Rev. D*, **96** (2017) 123522.
[28] BUNGE M., *Skept. Inqui.*, **30 (4)** (2006) 29.
[29] PLANCK COLLABORATION XXII, *Astron. Astrophys.*, **571** (2014) A22.
[30] PLANCK COLLABORATION XX, *Astron. Astrophys.*, **594** (2016) A20.
[31] SCOTT D., WHITE M., COHN J. D. and PIERPAOLI E., ArXiv *Astrophysics e-prints*, (2001), astro-ph/0104435.
[32] SCOTT D., *Astron. Geophys.*, **52** (2011) 6.31, arXiv:1112.0285.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200008

The theory of inflation

J. Martin(*)

Institut d'Astrophysique de Paris, UMR 7096-CNRS, Université Pierre et Marie Curie
98 bis boulevard Arago, 75014 Paris, France

Summary. — This article contains a concise review of the theory of inflation. We discuss its main theoretical aspects as well as its observational predictions. We also explain how the most recent astrophysical observations constrain the inflationary scenario.

1. – Introduction

The theory of inflation was invented at the end of the 70's and beginning of the 80's in order to improve the hot Big Bang model [1-6]. It consists in a phase of accelerated expansion taking place in the early Universe, at very high-energy scales, possibly as high as 10^{15} GeV. Not only inflation solves the puzzles of the standard model but it also provides a convincing mechanism for structure formation [7-12] (for reviews, see, *e.g.*, refs. [13, 14]) which, interestingly enough, is based on General Relativity (GR) and Quantum Mechanics (QM), two theories notoriously difficult to combine.

(*) E-mail: jmartin@iap.fr

On the observational front, the progresses have also been enormous, culminating recently with the publication of the high-accuracy measurement of the Cosmic Microwave Background (CMB) anisotropies by the European Space Agency (ESA) satellite Planck [15-18]. For the first time, this satellite has been able to show that the spectral index of the scalar power spectrum is close to one (exact scale invariance) but not exactly one, the deviation from one being detected at a statistically significant level, namely at more than 5σ. This is a crucial landmark because this was a prediction of inflation (and not a post-diction). This is the reason why inflation is now viewed as the front-runner candidate for describing the physical conditions that prevailed in the early Universe [19].

The aim of these lectures is to give a brief introduction to the theory of inflation. It is organized as follows. In the next section, sect. **2**, we discuss the motivations for inflation. In sect. **2**˙1, we first present the standard model of Cosmology, the hot Big Bang phase, as it was prior to the invention of inflation. Then, in sect. **2**˙2, we discuss the puzzles of the hot Big Bang phase and why a phase of accelerated expansion can solve them. In sect. **2**˙3, we discuss how inflation can be realized in practice and how it comes to an end (the theory of reheating). In sect. **3**, we discuss the theory of inflationary cosmological perturbations of quantum-mechanical origin. We first show that the quantum state of the perturbations at the end of inflation is peculiar (a two-mode squeezed state) and then we calculate the power spectrum in the slow-roll approximation. In sect. **4**, we briefly describe more complicated ways to realize inflation, in particular multiple field inflation. In sect. **5**, we discuss the observational status of inflation. We argue that the simplest class of scenarios is the preferred one and present observational constraints on the shape of the potential and on the reheating phase. Finally, in sect. **6**, we recap the main points and briefly discuss the future of inflation.

2. – Why inflation?

2˙1. *The pre-inflationary standard model.* – Among the four fundamental interactions that have been identified in Nature, gravity is the important one when it comes to Cosmology. Indeed, the Universe being neutral, this is the only force left with an infinite range and, therefore, the only one which can shape the Universe on astrophysical scales. The gravitational interaction being described by GR, any attempt to construct a model of the cosmos must be based on this theory. In addition, the standard model of Cosmology, the so-called hot Big Bang model, is based on a second fundamental assumption, namely the cosmological principle which states that, on large scales, the Universe is homogeneous and isotropic. This means that the general relativistic metric describing our Universe can be taken to be the Friedman-Lemaitre-Robertson-Walker (FLRW) one, namely

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left(\frac{\mathrm{d}r^2}{1-\mathcal{K}r^2} + r^2\mathrm{d}\theta^2 + r^2\sin^2\theta\mathrm{d}\varphi^2\right), \tag{1}$$

where $a(t)$ is the scale factor and $\mathcal{K}$ is a constant related to the curvature radius of space $r_{\mathrm{curv}} = a(t)/\sqrt{|\mathcal{K}|}$. Assuming that matter is described by perfect fluids, the

corresponding Einstein equations read

$$\frac{\dot{a}^2}{a^2} + \frac{\mathcal{K}}{a^2} = \frac{1}{3M_{\rm Pl}^2}\sum_{i=1}^{N}\rho_i + \frac{\Lambda_{\rm B}}{3}\,, \tag{2}$$

$$-\left(2\frac{\ddot{a}}{a} + \frac{\dot{a}^2}{a^2} + \frac{\mathcal{K}}{a^2}\right) = \frac{1}{M_{\rm Pl}^2}\sum_{i=1}^{N}p_i - \Lambda_{\rm B}, \tag{3}$$

where $M_{\rm Pl}$ is the Planck mass and $\Lambda_{\rm B}$ is the bare cosmological constant. The quantities ρ_i and p_i are respectively the energy density and pressure of the fluid "i". In the hot Big Bang model, one has five species, photons, neutrinos (which form radiation) and cold dark matter (cdm) and baryons (which form pressure-less matter) plus dark energy (given by the cosmological constant). Photons and neutrinos have a constant equation of state equal to 1/3, which means that $p_\gamma = \rho_\gamma/3$ and $p_\nu = \rho_\nu/3$. As already mentioned cdm and baryons have vanishing pressure. Finally dark energy (de) has a vacuum equation of state, meaning that $p_{\rm de} = -\rho_{\rm de}$. The standard model is also such that the spatial curvature vanishes, $\mathcal{K} = 0$. The free parameters are $H_0 \equiv \dot{a}/a|_{\rm now}$ (a dot denotes a derivative with respect to cosmic time), $\Lambda_{\rm B}$, ρ_γ, ρ_ν, $\rho_{\rm cdm}$, $\rho_{\rm b}$ and τ the optical depth that describes how the universe re-ionizes. We also have two extra parameters describing the perturbations, $A_{\rm S}$ and $n_{\rm S}$ that will be introduced later on. This means a total of nine parameters. However, introducing the critical energy density $\rho_{\rm cri} = 3H^2M_{\rm Pl}^2$ and defining $\Omega_i \equiv \rho_i/\rho_{\rm cri}$, the fact that $\mathcal{K} = 0$ means that the Friedman equation (2) can be rewritten as a constraint, $\Omega_\gamma + \Omega_\nu + \Omega_{\rm cdm} + \Omega_{\rm b} + \Omega_{\rm de} \equiv \Omega_{\rm tot} = 1$. So, in fact, we have eight free parameters (often, ρ_γ and ρ_ν are not viewed as free parameters because they are precisely determined by the CMB measurement and the number of neutrinos family; in that case we have a six-parameter model). These free parameters have now been measured with good precision (at the percentage level) [15,17]. For the expansion rate, one has $H_0 = 100h\,{\rm km}\times{\rm s}^{-1}\times{\rm Mpc}^{-1}$ with $h \simeq 0.67$, and for the matter content in the present day Universe, $\Omega_\gamma h^2 \simeq 2.47\times10^{-5}$, $\Omega_\nu h^2 \simeq 1.68\times10^{-5}$ (assuming three families of neutrinos), $\Omega_{\rm cdm}h^2 \simeq 0.1198$, $\Omega_{\rm b}h^2 \simeq 0.02255$ and $\Omega_{\rm de}h^2 \simeq 0.306$.

Knowing the matter content, by integrating the Einstein equations, we can infer the history of the Universe. The early Universe was dominated by radiation, with a scale factor given by $a(t) \propto t^{1/2}$ from the initial singularity until a redshift $z_{\rm eq} \simeq 3400$. Then, pressure-less matter took over with a scale factor $a(t) \propto t^{2/3}$ until a redshift of order one. Then, dark energy started to dominate and we still live in this epoch. The history of the Universe is thus made of three successive eras.

This simple model, except for the presence of dark energy, was already known before the 80's (although, at that time, the parameters were not measured with today accuracy) and has a great explanatory power. As mentioned before, it is known as the hot Big Bang model or the ΛCDM model in its most modern incarnation and is considered as the most convincing model for Cosmology. Why, then, the simple version presented above is nevertheless considered as not fully satisfactory thus motivating the introduction of inflation? We now turn to this question.

2˙2. *The puzzles of the standard model.* – With a few parameters, the pre-inflationary standard model of Cosmology was (is) able to explain a very large number of observational facts. Therefore it may seem strange to view it as not totally satisfactory. In fact, the difficulties of the hot Big Bang model are all related to the initial conditions. For instance, it is difficult to understand why spatial curvature is so small today. Indeed, the expansion during the hot Big Bang phase is decelerated and this means that $\Omega_{\rm tot}-1$ is growing. Therefore, since $\Omega_{\rm tot}-1$ is, today, very close to zero, this implies that it was in fact incredibly small in the early Universe (say, at BBN). Of course, it is always possible to postulate that the initial conditions were just such that it was the case. However, there is another explanation which consists in assuming that there was an accelerated phase of expansion, $\ddot{a}>0$, prior to the hot Big Bang epoch. This new phase of accelerated expansion is called "inflation". Then, the initial conditions at the beginning of the hot Big Bang epoch are now viewed as the "final conditions" at the end of inflation. Moreover, during a phase of accelerated expansion $\Omega_{\rm tot}-1$ is decreasing. Therefore, if $\Omega_{\rm tot}-1$ sufficiently decreases during inflation, it can entirely compensate the subsequent growth during the hot Big Bang phase and we understand why it is still small today. One can show that the compensation occurs if we have more than 60 e-folds of inflation. In some sense, inflation is a physical mechanism which puts the hot Big Bang phase on the "right tracks" by automatically single outing the right initial conditions.

Quite remarkably, one can show that all the puzzles of the standard model can be solved by the same mechanism [3]. For instance, this is the case of the so-called horizon problem. According to the hot Big Bang model, the angular scale of the horizon on the last scattering surface (where the CMB radiation was emitted) is $\simeq 1^\circ$. This means that we should expect the temperature to be strongly inhomogeneous on this scale all over the sky. As is well known, this is not the case since the CMB is, on the contrary, extremely homogeneous and isotropic. However, if one has 60 e-folds of inflation before the hot Big Bang phase, then the horizon at the last scattering surface covers the entire celestial sphere today and the problem is gone. We stress again that the number of e-folds needed to solve the problem turns out to be the same as for the flatness problem, namely 60.

Of course, postulating an accelerating phase is not sufficient. One must also identify a physical mechanism that could be responsible for it. In the next section, we discuss this question.

2˙3. *Basics of inflation.* – We have seen before that, if there is a phase of accelerated expansion in the early Universe, then the puzzles of the hot Big Bang model can be explained. As long as the gravitational field is described by GR and the cosmological principle valid, the acceleration of the scale factor can be expressed as

$$\frac{\ddot{a}}{a}=-\frac{1}{6M_{\rm Pl}^2}\sum_{i=1}^{N}(\rho_i+3p_i)+\frac{1}{3}\Lambda_{\rm B}. \tag{4}$$

Assuming that the cosmological constant does not play a role in the early Universe (given its present-day value), the condition for having $\ddot{a}>0$ reads

$$\rho_{\rm T} + 3p_{\rm T} < 0, \tag{5}$$

where $\rho_{\rm T} = \sum_{i=1}^{N} \rho_i$ and $p_{\rm T} = \sum_{i=1}^{N} p_i$ denote the total energy density and pressure. Given that the energy density must be positive, we are left with the condition that the pressure must be negative.

In usual situations, the pressure of a fluid is positive. This is for instance the case of radiation. However, inflation is supposed to take place in the very early Universe, at extremely high redshifts, and at those energies, hydrodynamics is clearly not the appropriate framework to describe matter. We should rather use field theory. The simplest type of field, compatible with the cosmological principle and the FLRW symmetries is a scalar field. We therefore assume that the matter content of the early Universe was dominated by a homogeneous scalar field $\phi(t)$ called, for obvious reasons, the "inflaton". The corresponding action is given by

$$\mathcal{L} = -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi - V(\phi) + \mathcal{L}_{\rm int}(\phi, A_\mu, \Psi), \tag{6}$$

where $V(\phi)$ is the inflaton potential and $\mathcal{L}_{\rm int}$ describes the interaction of the inflaton field with the other fields present such as gauge bosons A_μ or fermions Ψ. Then, by varying this action with respect to the metric tensor, one can calculate the energy momentum tensor and, therefore, the energy density and the pressure of the system. Ignoring for the moment the interaction term, this leads to

$$\rho = \frac{\dot{\phi}^2}{2} + V(\phi), \qquad p = \frac{\dot{\phi}^2}{2} - V(\phi). \tag{7}$$

We see that energy density is positive definite as it should (of course, $V(\phi) > 0$) but this is not the case of pressure. If the potential energy dominates over the kinetic energy, then $p < 0$. This will be the case if the kinetic energy is small or, in other words, if the inflaton moves slowly along its potential. And this will happen if the potential is nearly flat. We conclude that, if the inflaton dominates the energy budget at early times and if its potential is almost flat, then a phase of inflation can occur. This is the basics idea that underlies the theory of inflation.

At the technical level, the evolution of the system is controlled by the Friedmann and Klein-Gordon equations, namely

$$H^2 = \frac{1}{3M_{\rm Pl}^2} \left[\frac{\dot{\phi}^2}{2} + V(\phi) \right], \qquad \ddot{\phi} + 3H\dot{\phi} + V_\phi = 0, \tag{8}$$

where the subscript ϕ means the derivative with respect to the inflaton field. Unfortunately, this system of equations cannot be solved analytically unless the potential has a very specific form (for instance, $V(\phi) \propto e^{-\alpha\phi}$, a model called power law inflation). Therefore, we have to use either numerical calculations or a perturbative method. In

general, a perturbative method is based on an expansion of the relevant physical quantities in terms of a small parameter (or several) naturally present in the problem (for instance a coupling constant in field theory). Here, one can use the fact that the potential is nearly flat. If it is exactly flat, then the scalar field acts as a cosmological constant and the corresponding solution is de Sitter. One can then expand the solution of the system (8) around de Sitter. Since the de Sitter solution corresponds to a constant Hubble parameter, one can define small parameters by considering the derivatives of H and, then, expand the solution in these parameters. They are called horizon flow parameters or slow-roll parameters and are defined by [20, 21]

$$\epsilon_{n+1} \equiv \frac{\mathrm{d}\ln|\epsilon_n|}{\mathrm{d}N}, \quad n \geq 0, \tag{9}$$

where $\epsilon_0 \equiv H_{\rm ini}/H$ stands at the top of the hierarchy and $N \equiv \ln(a/a_{\rm ini})$ is the number of e-folds. The first Hubble flow parameter can be expressed as

$$\epsilon_1 = -\frac{\dot{H}}{H^2} = 1 - \frac{\ddot{a}}{aH^2} = \frac{3\dot{\phi}^2}{2}\frac{1}{\dot{\phi}^2/2 + V(\phi)}. \tag{10}$$

As mentioned above, it is related to the first derivative of the Hubble parameter. The second Hubble flow parameter, ϵ_2, would be related to $\ddot{H}$ and so on. We also see on the second expression of ϵ_1 that $\epsilon_1 < 1$ when $\ddot{a} > 0$, that is to say when inflation occurs. Of course, $\epsilon_1 \ll 1$ when the inflationary expansion is close to that of de Sitter. Finally, the third expression of ϵ_1 makes clear that it is a very small quantity when the kinetic energy is small compared to the total energy and, therefore, compared to the potential energy. In fact, there is yet another way to express the Hubble flow parameters. If one assumes that $\epsilon_n \ll 1$ (the following expressions are therefore approximate contrary to eqs. (10) which are exact), then the first three Hubble flow parameters can be written as [22]

$$\epsilon_1 \simeq \frac{M_{\rm Pl}^2}{2}\left(\frac{V_\phi}{V}\right)^2, \tag{11}$$

$$\epsilon_2 \simeq 2M_{\rm Pl}^2\left[\left(\frac{V_\phi}{V}\right)^2 - \frac{V_{\phi\phi}}{V}\right], \tag{12}$$

$$\epsilon_2\epsilon_3 \simeq 2M_{\rm Pl}^4\left[\frac{V_{\phi\phi\phi}V_\phi}{V^2} - 3\frac{V_{\phi\phi}}{V}\left(\frac{V_\phi}{V}\right)^2 + 2\left(\frac{V_\phi}{V}\right)^4\right]. \tag{13}$$

It is then clear that, when the inflaton potential is nearly flat, one has $\epsilon_n \ll 1$. The Hubble flow parameters are in fact nothing but a measure of the flatness of the inflaton potential.

Having identified the small parameters of the problem, one can now use them and design a method of approximation based on an expansion in terms of the ϵ_n's. This is

called the slow-roll approximation. The first step consists in re-writing the Friedman and Klein-Gordon equations (8) in terms of the ϵ_n's. This leads to

$$H^2 = \frac{V}{M_{\rm Pl}^2(3-\epsilon_1)}\,, \tag{14}$$

$$\left(1+\frac{\epsilon_2}{6-2\epsilon_1}\right)\frac{\mathrm{d}\phi}{\mathrm{d}N} = -M_{\rm Pl}^2\frac{\mathrm{d}\ln V}{\mathrm{d}\phi}\,. \tag{15}$$

At this stage, these expressions are exact. Then, we expand them at leading order in the Hubble flow parameters. This gives

$$H^2 \simeq \frac{V}{3M_{\rm Pl}^2}, \qquad \frac{\mathrm{d}\phi}{\mathrm{d}N} \simeq -M_{\rm Pl}^2\frac{\mathrm{d}\ln V}{\mathrm{d}\phi}\,. \tag{16}$$

Unsurprisingly, we now see that the expansion rate of the Universe is solely controlled by the potential energy. One great advantage of the above equations is that they can be integrated exactly. The solution reads

$$N - N_{\rm ini} = -\frac{1}{M_{\rm Pl}^2}\int_{\phi_{\rm ini}}^{\phi}\frac{V(\chi)}{V_\chi(\chi)}\,\mathrm{d}\chi, \tag{17}$$

$\phi_{\rm ini}$ being the initial value of the inflaton. If the above integral can be performed, then one obtains $N = N(\phi)$ and by inverting it, one arrives at the trajectory, $\phi = \phi(N)$. If one assumes a potential $V(\phi)$, this solution can be compared with the exact solution obtained by a numerical integration. In practice, as long as $\epsilon_n \ll 1$, eq. (17) turns out to be an excellent approximation.

We now turn to another crucial question of the inflationary scenario, namely how it comes to an end [23-26]. At this stage, let us recall that inflation is not an alternative to the ΛCDM model but just an additional ingredient. A phase of inflation is supposed to take place in the early Universe for the reasons explained in sect. **2**˙2 but, then, it must be smoothly connected to the standard ΛCDM phase. On a more practical side, it is known that the expansion of the Universe was radiation dominated during the Big Bang Nucleosynthesis (BBN) (otherwise the production of light elements, which is known to be in good agreement with the data, would be drastically modified) and, therefore, inflation must have stopped by that time.

There exists different mechanisms to stop inflation but the simplest one is just that, at some point, the potential is no longer flat enough to support inflation. Usually this happens in the vicinity of the minimum of the potential. Technically, this means that the slow-roll approximation is no longer valid. In fact, from eq. (10), one sees that the expansion is no longer accelerated when $\epsilon_1 = 1$ which, therefore, defines the time at which inflation comes to an end. Then, the field starts oscillating at the bottom of its potential. If $m^2 = \mathrm{d}^2V/\mathrm{d}\phi^2$ is the mass around the local minimum, the field behaves

as [23]

$$\phi(t) = \phi_{\rm end}\left(\frac{a_{\rm end}}{a}\right)^{3/2}\sin(mt), \tag{18}$$

namely the field oscillates with a frequency given by its mass. Of course, in this regime, the kinetic energy is no longer sub-dominant compared to the potential energy. In fact, there is now equipartition between them which means that $\langle p\rangle_t = 0$. This implies that the averaged energy density behaves as dust as also revealed by the fact that the overall amplitude of the inflaton is proportional to $a^{-3/2}$.

The above behavior is valid if one neglects the interaction of the inflaton with the other fields or, in other words, for times much smaller than the inflaton life time Γ^{-1}, where Γ is the total inflaton decay rate. If this is taken into account, then eq. (18) becomes

$$\phi(t) = \phi_{\rm end}e^{-\Gamma t}\left(\frac{a_{\rm end}}{a}\right)^{3/2}\sin(mt), \tag{19}$$

which shows that the total energy density stored in the inflaton field quickly goes to zero. This energy is transferred to the inflaton decay products. Then, these decay products thermalize and the radiation-dominated epoch starts at a temperature which is known as the reheating temperature $T_{\rm rh}$. This is the first time that a temperature can be defined in the history of the Universe. Equivalently, this also determines the reheating energy density, $\rho_{\rm reh}$, that is to say the energy density at which one starts the ΛCDM model. It is given by

$$\rho_{\rm reh} = g_*\frac{\pi^2}{30}T_{\rm reh}^4, \tag{20}$$

where g_* encodes the number of relativistic degrees of freedom.

It is also interesting to study the evolution of the equation of state during the reheating. We know it must transit between -1 and $1/3$. In fact, observationally speaking, the mean equation of state is easier to probe. It is defined by [27-30]

$$\overline{w}_{\rm reh} \equiv \frac{1}{\Delta N}\int_{N_{\rm end}}^{N_{\rm reh}} w_{\rm reh}(n){\rm d}n, \tag{21}$$

where $\Delta N \equiv N_{\rm reh} - N_{\rm end}$ is the total number of e-folds during reheating and $w_{\rm reh} \equiv p_{\rm T}/\rho_{\rm T}$ is the instantaneous equation of state. The quantity $\overline{w}_{\rm reh}$ controls the evolution of the total energy density since one has

$$\rho_{\rm reh} = \rho_{\rm end}\, e^{-3(1+\overline{w}_{\rm reh})\Delta N}, \tag{22}$$

where $\rho_{\rm end}$ is the energy density at the end of inflation, namely when $\epsilon_1 = 1$. If one is given a model of inflation, then this quantity can be easily calculated.

It is also relevant to introduce the reheating parameter which is a quantity depending on $\rho_{\rm reh}$ and $\overline{w}_{\rm reh}$. Explicitly, it reads [27-30]

$$R_{\rm rad} \equiv \left(\frac{\rho_{\rm reh}}{\rho_{\rm end}}\right)^{(1-3\overline{w}_{\rm reh})/(12+12\overline{w}_{\rm reh})}. \tag{23}$$

The reason why this parameter is important can be found in refs. [27-29, 19, 30]. It turns out that, when one tries to constrain reheating with the CMB, we end up constraining this parameter. As simple check allows us to understand why. Observationally speaking there should not be any difference between a model where reheating proceeds instantaneously and a model where reheating proceeds with an equation of state 1/3. If $R_{\rm rad}$ is the only combination of parameters we can access to, it should therefore have the same value for those two situations. And, indeed, it is easy to check that $R_{\rm rad} = 1$ if $\rho_{\rm reh} = \rho_{\rm end}$ (instantaneous reheating) or $\overline{w}_{\rm reh} = 1/3$ (radiative reheating).

Let us now illustrate the previous considerations on a simple example. Suppose the inflaton potential is given by $V(\phi) = m^2\phi^2/2$. Then, it is easy to perform the integral in eq. (17) and the corresponding trajectory reads

$$\phi(N) = \sqrt{\phi_{\rm ini}^2 - 4M_{\rm Pl}^2(N - N_{\rm ini})}. \tag{24}$$

As explained before, inflation stops when $\epsilon_1 = 1$ which, in this case, means $\phi_{\rm end} = \sqrt{2}M_{\rm Pl}$. From this result, one can also compute the total number of e-folds. One finds

$$N_{\rm T} \equiv N_{\rm end} - N_{\rm ini} = \frac{1}{4}\frac{\phi_{\rm ini}^2}{M_{\rm Pl}^2} - \frac{1}{2}. \tag{25}$$

This relation means that, in order to have more than 60 e-folds, one should start from $\phi_{\rm ini} \gtrsim 15M_{\rm Pl}$. Finally, the reheating will be completed when $H \simeq \Gamma$, namely $g_*\pi^2 T_{\rm reh}^4/30 \simeq M_{\rm Pl}^2\Gamma^2$ or

$$T_{\rm reh} \simeq \left(\frac{30}{g_*\pi^2}\right)^{1/4} M_{\rm Pl}^{1/2}\Gamma^{1/2}. \tag{26}$$

We see that the reheating temperature scales as the square root of the decay rate.

3. – Inflationary cosmological perturbations

We now turn to the theory of cosmological perturbations of quantum-mechanical origin. This part of the inflationary scenario makes use of GR and QM and as such is particularly interesting. Moreover, it allows us to build a bridge between theoretical considerations and actual astrophysical measurements. Therefore, it plays a crucial role in our attempts to observationally probe inflation.

So far, we have considered that the Universe was homogeneous and isotropic. Clearly, in the real world, this is not the case. Going beyond the cosmological principle is *a priori*

technically challenging since this means solving Einstein equations in an inhomogeneous and anisotropic situation. Fortunately, we know that the amplitude of these inhomogeneities were small in the early Universe as revealed by the fact that $\delta T/T \simeq 10^{-5}$ on the last scattering surface located at a redshift of $z_{\rm lss} \simeq 1100$. Since the amplification of the fluctuations proceeds by gravitational collapse, the amplitude of the inhomogeneities were even smaller during inflation. As a consequence, one can study their behavior perturbatively. Moreover, restricting ourselves to linear perturbations (leading order) is sufficient. Based on the previous considerations, we can then write [31]

$$g_{\mu\nu}(\eta,\boldsymbol{x}) = g_{\mu\nu}^{\rm FLRW}(t) + \delta g_{\mu\nu}\left(\eta,\boldsymbol{x}\right) + \ldots, \tag{27}$$

with the assumption that $|\delta g_{\mu\nu}(\eta,\boldsymbol{x})| \ll |g_{\mu\nu}^{\rm FLRW}(t)|$. The tensor $\delta g_{\mu\nu}$ can be decomposed into three types of fluctuations, scalar, vector and tensor or gravitational waves. The study of scalar perturbations can be reduced to the study of a single quantity, the curvature perturbation $\zeta(\eta,\boldsymbol{x})$ and the primordial gravitational waves can be described by a transverse and traceless two rank tensor $h_{ij}(\eta,\boldsymbol{x})$, $h^i{}_i = \partial_i h^i{}_j = 0$. Vector perturbations do not play a role during inflation. As was already mentioned, the evolution of the Universe is controlled by the Einstein equations, $G_{\mu\nu} = T_{\mu\nu}$. Since we expand the metric tensor in terms of the perturbations, one must do the same for the Einstein tensor, $G_{\mu\nu} = G_{\mu\nu}^{\rm FLRW} + \delta G_{\mu\nu}$ and for the stress energy tensor, $T_{\mu\nu} = T_{\mu\nu}^{\rm FLRW} + \delta T_{\mu\nu}$. Then, the equations describing the behavior of the perturbations are

$$\delta G_{\mu\nu} = \delta T_{\mu\nu}. \tag{28}$$

Of course, these equations are now partial differential equations since the perturbations are supposed to describe the early inhomogeneous and anisotropic Universe. But since these equations are linear, they can be solved by going to Fourier space.

Then, the idea is to quantize the system. The motivation is that this will provide a source for the cosmological perturbations (in other words, this will fix the initial conditions). This source will be the unavoidable quantum fluctuations of the inflaton and gravitational fields at the beginning of inflation. On the technical front, this means that $\delta g_{\mu\nu}$ will be promoted to a quantum operator, $\delta g_{\mu\nu} \to \delta\hat{g}_{\mu\nu}$. As a consequence, curvature perturbations and gravitational waves also become quantum operators, $\hat{\zeta}$ and $\hat{h}_{ij}$.

One fundamental assumption of inflation is that, initially, the quantum perturbations are placed in the vacuum state. Then, this state will evolve as the Universe expands. At the end of inflation, the system will be placed into a strongly two-mode squeezed state. This state is a very peculiar state and is defined as follows (here, we follow the presentation of ref. [32]). Let us consider a one-dimensional quantum oscillator. As is well known, its vacuum state is a Gaussian state whose wave function is given by

$$\Psi_0(x) = \frac{1}{\pi^{1/4}} e^{-x^2/2}, \tag{29}$$

where x is the position of the oscillator. This state, written in the momentum basis, reads

$$\tilde{\Psi}_0(p) = \frac{1}{\pi^{1/4}} e^{-p^2/2}, \tag{30}$$

where p is the conjugate momentum of x. An interesting feature of the vacuum state is that the dispersion in position and momentum are equal, namely

$$\langle \Delta\hat{x}^2 \rangle = \langle \Delta\hat{p}^2 \rangle = \frac{1}{2} \tag{31}$$

and saturates the Heisenberg inequality $\langle \Delta\hat{x}^2 \rangle \langle \Delta\hat{p}^2 \rangle = \frac{1}{4}$. A one-mode squeezed state is also a Gaussian state but, in position basis and momentum basis, its wave function is given by

$$\Psi_R(p) = \frac{\sqrt{R}}{\pi^{1/4}} e^{-R^2 x^2/2}, \qquad \tilde{\Psi}_R(p) = \frac{1}{\pi^{1/4}\sqrt{R}} e^{-p^2/(2R^2)}. \tag{32}$$

We see that the wave function now depends on an additional parameter, R. As a consequence, the dispersion in position and momentum are no longer equal,

$$\langle \Delta\hat{x}^2 \rangle = \frac{1}{2R^2}\,, \qquad \langle \Delta\hat{p}^2 \rangle = \frac{R^2}{2}\,, \tag{33}$$

although they still saturates the Heisenberg inequality. If $R > 1$, then the dispersion in position is smaller than that of the vacuum. We say that the state is squeezed in position, hence its name. Of course, since one has to satisfy the Heisenberg inequality, the price to pay is that the dispersion in momentum is larger. If $R < 1$, we have the opposite situation and the state is squeezed in momentum.

Then, let us consider two oscillators. The vacuum state of this system in position basis (namely the position of the first oscillator also referred to as the position of Alice and the position of the second oscillator also referred as to the position of Bob) can be written as

$$\Psi_0(x_1, x_2) = \frac{1}{\sqrt{\pi}} e^{-x_1^2/2 - x_2^2/2} = \frac{1}{\sqrt{\pi}} e^{-(x_1 - x_2)^2/4} e^{-(x_1 + x_2)^2/4}. \tag{34}$$

We see that the position of Alice and Bob are uncorrelated. From this expression, we are now in a position to introduce the two-mode squeezed state which is given by

$$\Psi_R(x_1, x_2) = \frac{1}{\sqrt{\pi}} e^{-R^2(x_1 - x_2)^2/4} e^{-(x_1 + x_2)^2/(4R^2)}, \tag{35}$$

where the squeezing factor R appears again and is related to the squeezing parameter r by $R = \ln r$. We see that the position of Alice and Bob are now correlated. It is also

interesting to notice that the two-mode squeezed state does not imply squeezing for Alice or Bob. Indeed, it is easy to check that

$$\langle \Delta\hat{x}_1^2 \rangle = \langle \Delta\hat{x}_2^2 \rangle = \frac{1+R^4}{4R^2}\,. \tag{36}$$

These dispersions are always larger than those one would obtain from the vacuum state. This is related to the fact that, if one traces out, say, Alice's degree of freedom, the obtained state of Bob is not a one-mode squeezed state but a thermal state.

The quantum-mechanical properties of inflation discussed above are clearly fascinating. Based on this aspect of the theory, one can wonder whether it would be possible to exhibit quantum effects in the sky. This was first discussed in refs. [33, 34] and, more recently, in refs. [35-40].

Let us now turn to a quantitative characterization of the cosmological fluctuations originating from inflation. As usual this will be done by computing the various correlation functions of scalar and tensor perturbations (in the following, we mainly focus on the scalar sector). The simplest correlation function is evidently the two-point correlation function which is given by

$$\langle \zeta^2(\eta, \boldsymbol{x}) \rangle = \int_0^{+\infty} \frac{\mathrm{d}k}{k} \mathcal{P}_\zeta(k), \tag{37}$$

where brackets mean quantum averages in the two mode squeezed state described above and where $\mathcal{P}_\zeta(k) = k^3|\zeta_{\boldsymbol{k}}|^2/(2\pi^2)$ is, by definition, the power spectrum of scalar perturbations. This scalar power spectrum is a very important quantity because it can be probed observationally by measuring the CMB anisotropies or by measuring the distributions of galaxies across our Universe. Using the slow-roll approximation introduced above, it can also be calculated for an arbitrary potential $V(\phi)$ and the result reads

$$\mathcal{P}_\zeta(k) = \mathcal{P}_{\zeta 0}(k_\mathrm{P}) \left[a_0^{(\mathrm{S})} + a_1^{(\mathrm{S})} \ln\left(\frac{k}{k_\mathrm{P}}\right) + \frac{a_2^{(\mathrm{S})}}{2} \ln^2\left(\frac{k}{k_\mathrm{P}}\right) + \ldots \right], \tag{38}$$

where k_P is a pivot scale and the global amplitude can be expressed as

$$\mathcal{P}_{\zeta_0} = \frac{H_*^2}{8\pi^2 \epsilon_{1*} M_\mathrm{Pl}^2}. \tag{39}$$

In the above formula, a star means that the corresponding quantity has been calculated at the time at which the pivot scale crossed out the Hubble radius during inflation. We notice that the amplitude of the correlation function depends on the square of the Hubble rate during inflation (measured in Planck units) and is inversely proportional to the first slow-roll parameter. All these quantities are scale independent and so is the global amplitude. This result is viewed as one of the most important success of inflation. Indeed, before the invention of inflation, it was already known that a scale-invariant

power spectrum (or Harrisson-Zeldovitch power spectrum) is a good fit to the data. But its origin was mysterious and there was no convincing physical mechanism to produce it. Inflation, on the contrary, naturally implies this property. In fact, generically, exact scale invariance is not a prediction of inflation because, as can be seen in eq. (38), the overall amplitude receives small, scale dependent, logarithmic corrections. The amplitudes of those corrections is determined by the Hubble flow parameters, namely [20, 41-45, 21, 46, 47],

$$
\begin{aligned}
(40)\quad a_0^{(\mathrm{S})} &= 1 - 2\,(C+1)\epsilon_{1*} - C\epsilon_{2*} + \left(2C^2 + 2C + \frac{\pi^2}{2} - 5\right)\epsilon_{1*}^2 \\
&\quad + \left(C^2 - C + \frac{7\pi^2}{12} - 7\right)\epsilon_{1*}\epsilon_{2*} + \left(\frac{1}{2}C^2 + \frac{\pi^2}{8} - 1\right)\epsilon_{2*}^2 \\
&\quad + \left(-\frac{1}{2}C^2 + \frac{\pi^2}{24}\right)\epsilon_{2*}\epsilon_{3*} + \ldots,
\end{aligned}
$$

$$
(41)\quad a_1^{(\mathrm{S})} = -2\epsilon_{1*} - \epsilon_{2*} + 2(2C+1)\epsilon_{1*}^2 + (2C-1)\epsilon_{1*}\epsilon_{2*} + C\epsilon_{2*}^2 - C\epsilon_{2*}\epsilon_{3*} + \ldots,
$$

$$
(42)\quad a_2^{(\mathrm{S})} = 4\epsilon_{1*}^2 + 2\epsilon_{1*}\epsilon_{2*} + \epsilon_{2*}^2 - \epsilon_{2*}\epsilon_{3*} + \ldots,
$$

$$
(43)\quad a_3^{(\mathrm{S})} = \mathcal{O}(\epsilon_{n*}^3),
$$

where $C \equiv \gamma_{\mathrm{E}} + \ln 2 - 2 \approx -0.7296$, γ_{E} being the Euler constant. Therefore, the exact prediction of inflation (really a prediction since it was made before it was checked) is that the power spectrum should be almost scale invariant but not exactly scale invariant. This prediction has been recently confirmed for the first time by the Planck data. Technically, one defines the spectral index, which is the logarithmic derivative of $\ln \mathcal{P}_\zeta(k)$, namely

$$
(44)\qquad n_{\mathrm{S}} = 1 - 2\epsilon_{1*} - \epsilon_{2*},
$$

where $n_{\mathrm{S}} = 1$ corresponds to exact scale invariance. As will be discussed in more details in the following, Planck has measured $n_{\mathrm{S}} \simeq 0.96$ and $n_{\mathrm{S}} = 1$ is now excluded at more than 5σ. We also see that the spectral index depends on the two first Hubble flow parameters. As a consequence, a measurement of n_{S} is also a measurements of ϵ_{1*} and ϵ_{2*}, that is to say of the first and second derivative of the inflaton potential. This explains how astrophysical measurements can constrain the theory of inflation.

The treatment of tensor modes (primordial gravitational waves) proceeds in the very same way. One can compute the two-point correlation and the power spectrum using the slow-roll approximation. One then arrives at the following expression:

$$
(45)\qquad \mathcal{P}_h(k) = \mathcal{P}_{h0}(k_{\mathrm{P}}) \left[a_0^{(\mathrm{T})} + a_1^{(\mathrm{T})} \ln\left(\frac{k}{k_{\mathrm{P}}}\right) + \frac{a_2^{(\mathrm{T})}}{2} \ln^2\left(\frac{k}{k_{\mathrm{P}}}\right) + \ldots\right],
$$

where the amplitude $\mathcal{P}_{h0}(k_{\rm P})$ is given by

$$\mathcal{P}_{h_0} = \frac{2H_*^2}{\pi^2 M_{\rm Pl}^2}\,. \tag{46}$$

As was the case for scalar perturbations, the overall amplitude is also given by the square of the expansion rate during inflation measured in Planck units. Of course the big difference is that the first slow-roll parameter ϵ_{1*} is now absent. This means that a measurement of the tensor modes would immediately provide the energy scale of inflation. Notice that $\mathcal{P}_{h0}(k_{\rm P})$ is also scale independent and, at leading order, the tensor power spectrum is therefore scale invariant. However, as was also the case for scalar modes, this scale-invariant amplitude receives small, scale-dependent, logarithmic corrections the amplitude of which can be expressed as [21]

$$\begin{aligned} a_0^{(\rm T)} &= 1 - 2\,(C+1)\epsilon_{1*} + \left(2C^2 + 2C + \frac{\pi^2}{2} - 5\right)\epsilon_{1*}^2 \\ &\quad + \left(-C^2 - 2C + \frac{\pi^2}{12} - 2\right)\epsilon_{1*}\epsilon_{2*} + \ldots, \end{aligned} \tag{47}$$

$$a_1^{(\rm T)} = -2\epsilon_{1*} + 2(2C+1)\epsilon_{1*}^2 - 2(C+1)\epsilon_{1*}\epsilon_{2*} + \ldots, \tag{48}$$

$$a_2^{(\rm T)} = 4\epsilon_{1*}^2 - 2\epsilon_{1*}\epsilon_{2*} + \ldots, \tag{49}$$

$$a_3^{(\rm T)} = \mathcal{O}(\epsilon_{n*}^3). \tag{50}$$

From the coefficient $a_1^{(\rm T)}$, one can read the tensor spectral index (at first order in slow-roll). One obtains

$$n_{\rm T} = -2\epsilon_1. \tag{51}$$

Exact scale invariance corresponds to $n_{\rm T} = 0$ (for historical reasons, the convention differs from that of scalars). Another difference is that $n_{\rm T}$ depends on ϵ_{1*} only while $n_{\rm S}$ depends on ϵ_{1*} and ϵ_{2*}. Given that ϵ_{1*} is always positive, this implies that $n_{\rm T}$ is always negative (or red).

Finally, one can also calculate the tensor amplitude to scalar amplitude r. Using the previous expressions, one obtains

$$r \equiv \frac{\mathcal{P}_h}{\mathcal{P}_\zeta} = 16\epsilon_{1*}. \tag{52}$$

Since, by definition, $\epsilon_{1*} \ll 1$, this means that gravitational waves are sub-dominant (which explains why they have not yet been detected [48, 49]). Notice that there is *a priori* no lower bound on r. Therefore, if r turns out to be very small, primordial gravitational waves will probably never been detected but this would be in no way in

contradiction with the predictions of inflation. At the time of writing, it is believed that the next generations of telescope and satellites will be able to reach the level $r \sim 10^{-3}$ maybe a bit smaller. Let us hope that Nature has produced an r larger than this limit!

To conclude this section, let us mention Non-Gaussianities (NG). So far, we have restricted our considerations to two-point correlation functions. Of course, higher correlation functions are also of great interest. Usually, the three-point correlation function (bispectrum) and the four-point correlation function (trispectrum) are considered. For the models described previously, NG are very small (of the order of the slow-roll parameters) [50-53]. The reason is easy to understand. We have started from a Gaussian state and the evolution of the perturbations is linear. As a consequence, the appearance of any NG is necessarily related to non-linearities, which are very small.

4. – Extensions

So far, we have described the simplest way to realize inflation. However, since the invention of inflation in the 80's, more complicated scenarios have been imagined. In this section, we say a few words about them.

The most generic extension is probably to consider models where, instead of having one scalar field, one has several ones playing an active role during inflation [54]. This appears to be a natural approach given that inflation can occur at energy scales as high as $10^{15}\,\mathrm{GeV}$. At those scales, it is believed that particle physics is no longer described by the standard model but by its extensions (SUSY, SUGRA, string theory, etc., ...). And, usually, in these alternative frameworks, there are plethora of scalar fields.

Clearly, multiple-field inflation scenarios are more complicated and it is more difficult to make generic predictions. However, one can list three main modifications. Firstly, there is the possibility of having non-adiabatic perturbations, which is impossible for single field models. The reason is that, if several scalar fields are present during inflation, then the corresponding decay products can have different origin resulting in the possible presence of non-adiabatic perturbations. Secondly, non-adiabatic perturbations can source the evolution of curvature perturbations. As a result, if they are present during inflation and reheating, $\zeta(\eta, \boldsymbol{x})$ on large scales is no longer a conserved quantity. This has drastic consequences, especially for reheating, which then becomes potentially dependent on the details of physical processes going on on scales smaller than the Hubble radius. Thirdly, it is possible to produce non-negligible NG. As already mentioned, these modifications are not mandatory and must be analyzed on a model by model basis.

Yet other extensions are also possible such that having a non-canonical kinetic term for the scalar field. They are called K-inflation models [55, 56] (for the observational status of this class of models, see refs. [46, 57, 47]). It is also possible to have models with features [58, 59]. This means a model of inflation where, in some limited region, the potential is not flat. This usually causes a transitory violation of the slow-roll approximation which can result in oscillations in the power spectrum and non-negligible NG [60-62]. More complicated models are possible, for instance by combining the var-

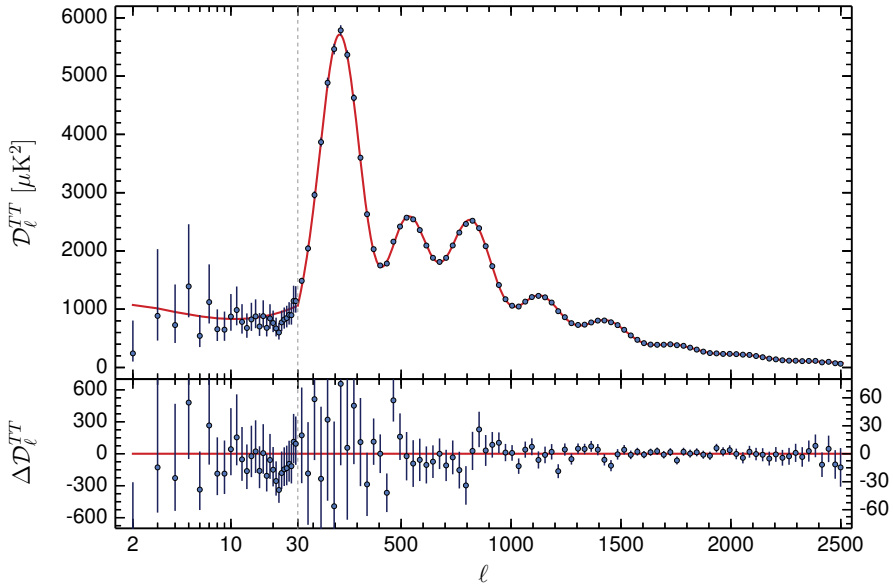


Fig. 1. – Multipole moments *versus* angular scale from Planck 2015 data. The multipole moments are obtained from the CMB map by Fourier transforming it according to: $\langle \delta T/T(\boldsymbol{e}_1)\delta T/T(\boldsymbol{e}_2)\rangle = (4\pi)^{-1}\sum_\ell (2\ell+1)C_\ell P_\ell(\cos\theta)$, where θ is the angle between two directions $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ and P_ℓ is a Legendre polynomial. The multipole moments C_ℓ are interpreted as the power of the signal at a given angle θ. Notice that $\mathcal{D}_\ell$ is related to C_ℓ by $\mathcal{D}_\ell = \ell(\ell+1)C_\ell/(2\pi)$. The red curve corresponds to the best fit and is consistent with the predictions of single field, slow-roll, inflation. Figure taken from ref. [17].

ious ingredients discussed above [63], but we will not discuss them here. We now turn to another question, namely how the observations can discriminate among these various possibilities.

5. – Inflation and CMB observations

The Planck satellite has recently measured the CMB temperature, see fig. 1, and polarization, see figs. 2 and 3, anisotropies with unprecedented accuracy. These new data allow us to constrain inflation and to learn which was the version of inflation realized in the early Universe.

In brief, Planck has shown that the Universe is so spatially flat, that the perturbations are adiabatic and Gaussian [18]. These results are all consistent with single field (with minimal kinetic term), slow-roll, inflation which, therefore, appears to be the preferred class of models. This does not mean that the more complicated versions discussed in sect. **4** are ruled out but just that, at the moment, they are not needed in order to explain the data.

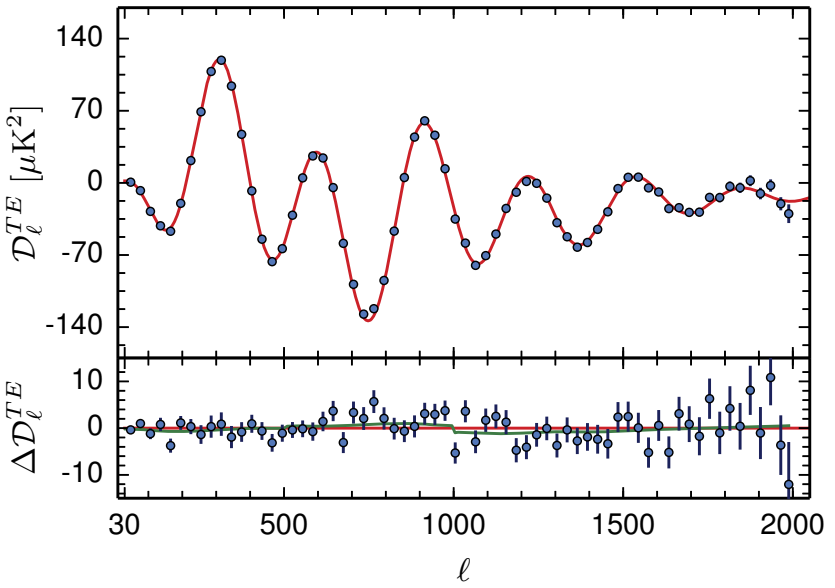


Fig. 2. – Multipole moments corresponding to the correlation between temperature and E-mode polarization anisotropies. The red solid line is obtained from temperature measurements only, see fig. 1. The lower panel shows the residual with respect to this best fit. Figure taken from ref. [17].

With regards to inflation, probably the most important discovery made by the Planck satellite is the measurement of the scalar spectral index [18]

$$n_{\rm S} = 0.969 \pm 0.005. \tag{53}$$

For the first time, the value $n_{\rm S} = 1$ is excluded at more than 5σ. As was already discussed above, the fact that the power spectrum must be scale invariant (the so-called Harrisson-Zeldovitch power spectrum) was known long ago (before the invention of inflation). But the non-trivial prediction of inflation was that $n_{\rm S}$ should be close to one but not exactly one. And this is exactly what has been observed for the first time by the Planck satellite.

Another important of piece of information is that, unfortunately, so far, no gravitational wave has been detected. This means the following upper bound on the tensor to scalar ratio r [48]

$$r \lesssim 0.08. \tag{54}$$

From the measurements of those quantities, one can also infer constraints on the Hubble flow parameters, see fig. 4 and refs. [64,65,19]. We see that $P_* \equiv \mathcal{P}_{\zeta 0} a_0^{(\rm S)}$ and ϵ_{2*} are constrained while there only exists an upper bound on ϵ_{1*}. Of course, P_* is determined because one knows the amplitude of CMB fluctuations (namely $\delta T/T \simeq 10^{-5}$). On the

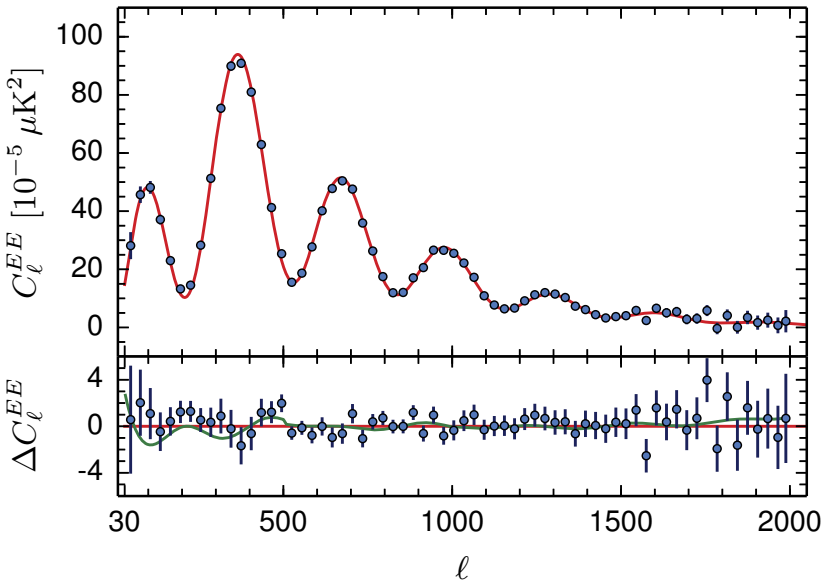


Fig. 3. – Same as in fig. 2 but for the E-mode power spectrum obtained from Planck 2015. Figure taken from ref. [17].

other hand, the upper bound on ϵ_{1*} originates from eq. (52) and the fact that we only have an upper bound on r. Given that $H_*^2/M_{\rm Pl}^2 \simeq 8\pi^2\epsilon_{1*}P_*$, this means that we only have an upper bound on the energy scale of inflation, namely

$$H_* \lesssim 1.2 \times 10^{14}\,\text{GeV}, \tag{55}$$

or $\rho_*^{1/4} \lesssim 2.2{\times}10^{16}\,\text{GeV}$. Finally, the third slow-roll parameter, ϵ_{3*} is not well constrained which means that we do not have yet a detection of a running.

We have seen before that the slow-roll parameters carry information about the shape of the inflaton potential. Since we have obtained constraints on these parameters, we must be able to say something about the shape of the inflaton potential itself [64, 65, 19]. In order to answer this question, one can calculate the Bayesian evidence of the various models of inflation. The Bayesian evidence is the integral of the likelihood function over the prior space. It characterizes the performance of a model and its ability to fit the data [66]. The larger the evidence, the better the model. In refs. [64, 65, 19], the Bayesian evidence of nearly two hundred models were computed. The result of this computation is displayed in figs. 5 where the number of unconstrained parameters is also indicated. A detailed analysis of those results has been published in refs. [64, 65, 19], but the bottom line is that plateau inflationary models are the "best" models according to the Planck data. A plateau potential is a potential which flattens out at infinity. The prototype of

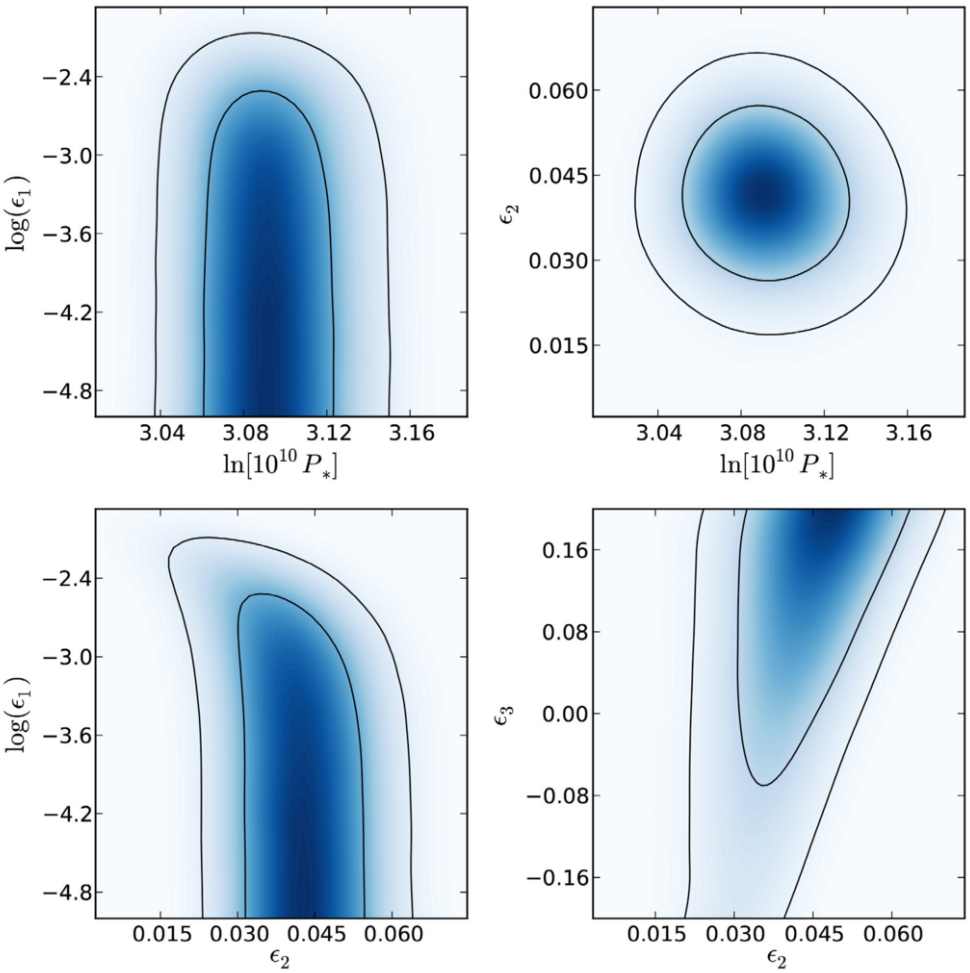


Fig. 4. – Posterior distributions of the parameters ϵ_{1*}, ϵ_{2*}, ϵ_{3*} and $P_* \equiv a_0^{(\mathrm{S})}\mathcal{P}_{\zeta 0}$. The posteriors are taken to be Jeffreys's priors for P_* and ϵ_{1*} and flat priors for ϵ_{2*} and ϵ_{3*}.

this class of models is the so-called Starobinsky model given by

$$V(\phi) = M^4 \left(1 - e^{-\sqrt{2/3}\phi/M_{\mathrm{Pl}}}\right)^2. \tag{56}$$

This conclusion is non-trivial since models that were historically considered as leading candidates, such as $V(\phi) = m^2\phi^2/2$, are now strongly disfavored compared to plateau models.

Let us also notice another interesting point. The prediction of plateau models for r is, roughly speaking, $r \simeq 10^{-3}$. As indicated before, this value is in principle reachable

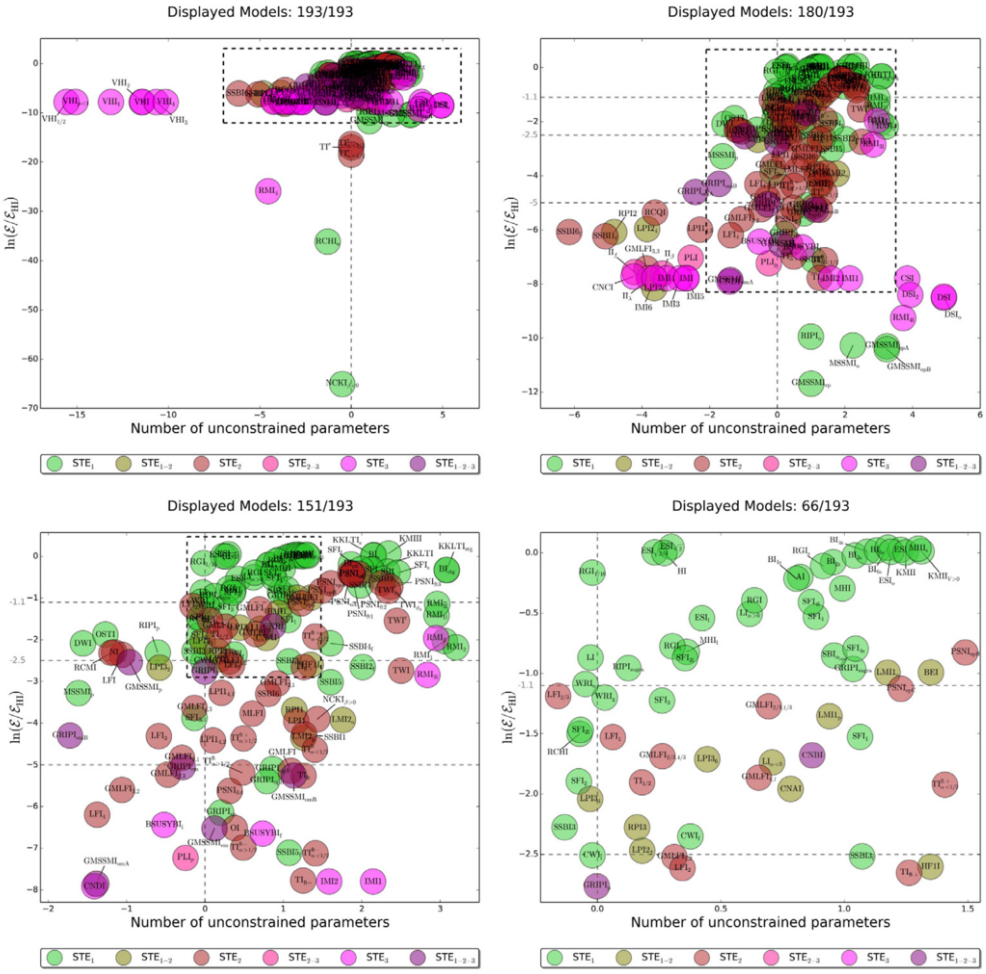


Fig. 5. – Bayesian evidence *versus* number of unconstrained parameters for different models of inflation. Each circle represents a given inflationary scenario (the size of the circle has no meaning). The upper right panel is a zoom on the "best" region (the square delimited by the dashed black line) of the upper left panel. In the same way, the bottom left plot is a zoom on the "best" region of the upper right. Finally, the bottom right is a zoom on the "best" region of the bottom left figure.

by the next generation of instruments. This means that there is maybe a good chance to detect primordial gravitational in a non too distant future (say, a decade).

Finally, let us discuss what the Planck data imply for reheating. As was discussed before, constraints on reheating are expressed through constraints on the reheating parameter $R_{\rm rad}$ defined in eq. (23). In refs. [27-30], the posterior distributions was derived for the nearly two hundred models already considered before for the calculation of the

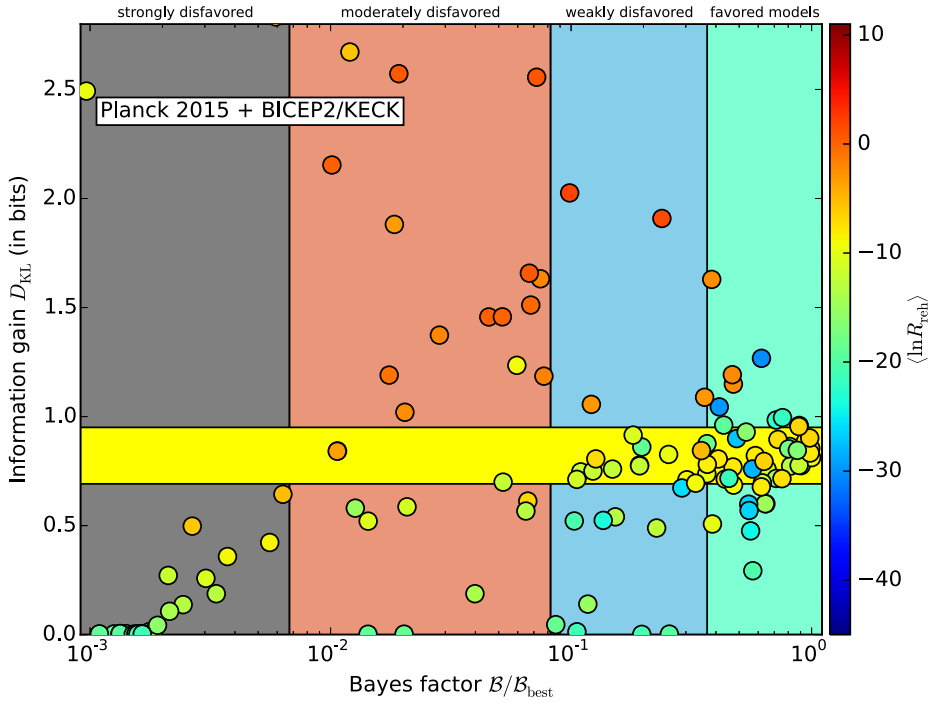


Fig. 6. – Kullback-Leibler $D_{\rm KL}$ divergence *versus* Bayesian evidence for various models of inflation. The mean value is given by $\langle D_{\rm KL} \rangle = 0.82 \pm 0.13$ and the yellow band represents the one-sigma deviation around this mean value.

Bayesian evidence. The situation is summarized in fig. 6. It represents the Kullback-Leibler divergence between the prior distribution and the posterior *versus* the Bayesian evidence for different models of inflation (represented by circles). The Kullback-Leibler divergence is defined by

$$D_{\rm KL} = \int P\left(\ln R_{\rm reh}|D\right) \ln\left[\frac{P\left(\ln R_{\rm reh}|D\right)}{\pi\left(\ln R_{\rm reh}\right)}\right] {\rm d}\ln R_{\rm reh}, \tag{57}$$

where $R_{\rm reh}$ is given by $\ln R_{\rm reh} = \ln R_{\rm rad} + \ln(\rho_{\rm end}/M_{\rm Pl}^4)/4$ and is therefore, for a given model of inflation, in a one-to-one correspondence with $R_{\rm rad}$. The quantity π represents the prior on $R_{\rm reh}$ and P the posterior. The Kullback-Leibler divergence measures the "distance" between the prior and the posterior and, as a consequence, also represents the amount of information provided by the data D (of course, here, the Planck data) about $\ln R_{\rm reh}$. The constraints are model dependent and one has a posterior distribution per model of inflation, an amount of information which, given the number of scenarios analyzed, is difficult to deal with. The value of $D_{\rm KL}$ is one way to summarize the information about reheating for a given model to one number. In this sense, fig. 6

completely describes what, for each known model of inflation, the Planck data implies with regards to the ability to fit the data and to reheating. Let us also notice that one can calculate the mean value of D_{KL}. One finds $\langle D_{\mathrm{KL}} \rangle = 0.82 \pm 0.13$, which expresses the fact that reheating is globally constrained by the Planck data.

6. – Conclusions

In this short review, we have discussed the theory of inflation. Over the years, the inflationary scenario has become a crucial ingredient in our understanding of Cosmology. It is important to stress that inflation is not an alternative to the standard model of Cosmology, it is rather a new part of it.

Invented in the 80's, inflation has recently witnessed new developments with the publication of the high accuracy Planck data. Clearly, these data have boosted our confidence in inflation. In particular, the measurement of the spectral index to be close but not equal to one is an important confirmation of an inflationary prediction. Admittedly, it is probably not the final proof that inflation actually occurred in the early Universe but it nevertheless represents a very strong argument in its favor. From the Planck data, we have also learned that inflation is probably realized in its simplest version (single field, slow-roll, with minimal kinetic term) and that the best scenario is a plateau model for which the potential flattens out at very large values of the field.

What is then the next step? Clearly, the detection of primordial gravitational waves will play a crucial role. It is an unambiguous prediction of inflation that has not yet been confirmed. Future missions will be able to reach $r \sim 10^{-3}$. Unfortunately, inflation, as a paradigm, does not predict the value of r even if r is predicted if a precise scenario is given. However, the best model of inflation, the Starobinsky model, predicts a value of r which, in principle, could be detected in the future.

Let us also add that the detection of NG will also certainly play an important role in the future. Given that we deal with the simplest class of models, the expected signal is very small and its detection will be challenging (if possible). But, obviously, this would be of crucial importance.

Of course, inflation is not a perfect scenario and some of its aspects remain unclear. But, as an effective model of the early Universe, it scores pretty well. Let us see whether its performances remain so efficient in the future.

* * *

I would like to thank the organizers, especially Profs. J. Silk and N. Vittorio, and the Italian Physical Society (SIF) for having invited me to lecture at the International School of Physics Enrico Fermi (from 26th June to 19th July 2017) at the beautiful Villa Monastero located in Varenna, Lake of Como (Italy).

REFERENCES

[1] STAROBINSKY A. A., *Phys. Lett. B*, **91** (1980) 99.
[2] STAROBINSKY A. A., *Phys. Lett. B*, **117** (1982) 175.
[3] GUTH A. H., *Phys. Rev. D*, **23** (1981) 347.
[4] LINDE A. D., *Phys. Lett. B*, **108** (1982) 389.
[5] ALBRECHT A. and STEINHARDT P. J., *Phys. Rev. Lett.*, **48** (1982) 1220.
[6] LINDE A. D., *Phys. Lett. B*, **129** (1983) 177.
[7] STAROBINSKY A. A., *JETP Lett.*, **30** (1979) 682.
[8] MUKHANOV V. F. and CHIBISOV G., *JETP Lett.*, **33** (1981) 532.
[9] MUKHANOV V. F. and CHIBISOV G., *Sov. Phys. JETP*, **56** (1982) 258.
[10] GUTH A. H. and PI S., *Phys. Rev. Lett.*, **49** (1982) 1110.
[11] HAWKING S., *Phys. Lett. B*, **115** (1982) 295.
[12] BARDEEN J. M., STEINHARDT P. J. and TURNER M. S., *Phys. Rev. D*, **28** (1983) 679.
[13] MARTIN J., *Lect. Notes Phys.*, **669** (2005) 199.
[14] MARTIN J., *Lect. Notes Phys.*, **738** (2008) 193.
[15] ADE P. *et al.*, *Astron. Astrophys.*, **571** (2014) A16.
[16] ADE P. A. R. *et al.*, *Astron. Astrophys.*, **571** (2014) A22.
[17] ADE P. A. R. *et al.*, *Astron. Astrophys.*, **594** (2016) A13.
[18] ADE P. A. R. *et al.*, *Astron. Astrophys.*, **594** (2016) A20.
[19] MARTIN J., *Astrophys. Space Sci. Proc.*, **45** (2016) 41.
[20] SCHWARZ D. J., TERRERO-ESCALANTE C. A. and GARCIA A. A., *Phys. Lett. B*, **517** (2001) 243.
[21] LEACH S. M., LIDDLE A. R., MARTIN J. and SCHWARZ D. J., *Phys. Rev. D*, **66** (2002) 023515.
[22] LIDDLE A. R., PARSONS P. and BARROW J. D., *Phys. Rev. D*, **50** (1994) 7222.
[23] TURNER M. S., *Phys. Rev. D*, **28** (1983) 1243.
[24] TRASCHEN J. H. and BRANDENBERGER R. H., *Phys. Rev. D*, **42** (1990) 2491.
[25] KOFMAN L., LINDE A. D. and STAROBINSKY A. A., *Phys. Rev. D*, **56** (1997) 3258.
[26] AMIN M. A., HERTZBERG M. P., KAISER D. I. and KAROUBY J., *Int. J. Mod. Phys. D*, **24** (2014) 1530003.
[27] MARTIN J. and RINGEVAL C., *JCAP*, **0608** (2006) 009.
[28] MARTIN J. and RINGEVAL C., *Phys. Rev. D*, **82** (2010) 023511.
[29] MARTIN J., RINGEVAL C. and VENNIN V., *Phys. Rev. Lett.*, **114** (2015) 081303.
[30] MARTIN J., RINGEVAL C. and VENNIN V., *Phys. Rev. D*, **93** (2016) 103532.
[31] MUKHANOV V. F., FELDMAN H. and BRANDENBERGER R. H., *Phys. Rept.*, **215** (1992) 203.
[32] LVOVSKY A. I. (2014).
[33] GRISHCHUK L. and SIDOROV Y., *Phys. Rev. D*, **42** (1990) 3413.
[34] GRISHCHUK L., HAUS H. and BERGMAN K., *Phys. Rev. D*, **46** (1992) 1440.
[35] MARTIN J., VENNIN V. and PETER P., *Phys. Rev. D*, **86** (2012) 103524.
[36] MARTIN J. and VENNIN V., *Phys. Rev. D*, **93** (2016) 023505.
[37] MARTIN J. and VENNIN V., *Phys. Rev. A*, **93** (2016) 062117.
[38] MARTIN J. and VENNIN V., *Phys. Rev. A*, **94** (2016) 052135.
[39] MARTIN J. and VENNIN V., *Phys. Rev. D*, **96** (2017) 063501.
[40] MARTIN J. and VENNIN V., *JCAP*, **1805** (2018) 063.
[41] CASADIO R., FINELLI F., LUZZI M. and VENTURI G., *Phys. Rev. D*, **71** (2005) 043517.
[42] CASADIO R., FINELLI F., LUZZI M. and VENTURI G., *Phys. Lett. B*, **625** (2005) 1.
[43] CASADIO R., FINELLI F., LUZZI M. and VENTURI G., *Phys. Rev. D*, **72** (2005) 103516.
[44] GONG J.-O. and STEWART E. D., *Phys. Lett. B*, **510** (2001) 1.

[45] Choe J., Gong J.-O. and Stewart E. D., *JCAP*, **0407** (2004) 012.
[46] Lorenz L., Martin J. and Ringeval C., *Phys. Rev. D*, **78** (2008) 083513.
[47] Martin J., Ringeval C. and Vennin V., *JCAP*, **1306** (2013) 021.
[48] Ade P. A. R. *et al.*, *Phys. Rev. Lett.*, **114** (2015) 101301.
[49] Martin J., Ringeval C., Trotta R. and Vennin V., *Phys. Rev. D*, **90** (2014) 063501.
[50] Gangui A., Lucchin F., Matarrese S. and Mollerach S., *Astrophys. J.*, **430** (1994) 447.
[51] Gangui A., *Phys. Rev. D*, **50** (1994) 3684.
[52] Gangui A. and Martin J., *Mon. Not. Roy. Astron. Soc.*, **313** (2000) 323.
[53] Maldacena J. M., *JHEP*, **05** (2003) 013.
[54] Wands D., *Lect. Notes Phys.*, **738** (2008) 275.
[55] Armendariz-Picon C., Damour T. and Mukhanov V. F., *Phys. Lett. B*, **458** (1999) 209.
[56] Garriga J. and Mukhanov V. F., *Phys. Lett. B*, **458** (1999) 219.
[57] Lorenz L., Martin J. and Ringeval C., *Phys. Rev. D*, **78** (2008) 063543.
[58] Starobinsky A. A., *JETP Lett.*, **55** (1992) 489, (*Pisma Zh. Eksp. Teor. Fiz.*, **55** (1992) 477).
[59] Hazra D. K., Aich M., Jain R. K., Sriramkumar L. and Souradeep T., *JCAP*, **1010** (2010) 008.
[60] Martin J. and Sriramkumar L., *JCAP*, **1201** (2012) 008.
[61] Hazra D. K., Sriramkumar L. and Martin J., *JCAP*, **1305** (2013) 026.
[62] Martin J., Sriramkumar L. and Hazra D. K., *JCAP*, **1409** (2014) 039.
[63] Vila S., Martin J. and Steer D., *JCAP*, **1408** (2014) 032.
[64] Martin J., Ringeval C. and Vennin V., *Phys. Dark Univ.*, **5-6** (2014) 75.
[65] Martin J., Ringeval C., Trotta R. and Vennin V., *JCAP*, **1403** (2014) 039.
[66] Trotta R., *Contemp. Phys.*, **49** (2008) 71.

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200009

Primordial Non-Gaussianity

M. Celoria

Gran Sasso Science Institute - via F. Crispi 7, I-67100, L'Aquila, Italy

ICTP, International Centre for Theoretical Physics - Strada Costiera 11, 34151, Trieste, Italy

S. Matarrese

Dipartimento di Fisica e Astronomia "Galileo Galilei", Università di Padova via Marzolo 8, I-35131, Padova, Italy

INFN, Sezione di Padova - via Marzolo 8, I-35131, Padova, Italy

INAF-Osservatorio Astronomico di Padova - vicolo dell Osservatorio 5, I-35122 Padova, Italy

Gran Sasso Science Institute - via F. Crispi 7, I-67100, L'Aquila, Italy

Summary. — Here we review the present status of modelling of and searching for primordial non-Gaussianity of cosmological perturbations. After introducing the models for non-Gaussianity generation during inflation, we discuss the search for non-Gaussian signatures in the Cosmic Microwave Background and in the Large-Scale Structure of the Universe.

1. – Introduction

According to *Planck* 2018, "the 6-parameter ΛCDM model provides an astonishingly accurate description of the Universe form times prior to 380000 years after the Big Bang, defining the last-scattering surface observed via the Cosmic Microwave Background (CMB) radiation, to the present day at an age of 13.8 billion years" [1].

Actually, the concordance model describes the evolution of tiny fluctuations on top of a homogeneous and isotropic background from the early Universe to the time of observation. Specifically, we can study how these modes looked like at the last-scattering surface, by analyzing the inhomogeneities in the CMB, and, in recent times, by observing the Large-Scale Structure (LSS) of the Universe.

In the standard model of cosmology, the primordial perturbations, corresponding to the seeds for the LSS, are chosen from a Gaussian distribution with random phases. This assumption is justified from experimental evidences as deviations from this Gaussian hypothesis, *i.e.* Primordial Non-Gaussianity (PNG), have not been observed yet.

Note that from the theoretical point of view, it is not surprising that "a Gaussian random field may provide a good description of the properties of density fluctuations" [2]. Actually, "the central limit theorem implies that a Gaussian distribution arises whenever one has a variable [...] which is a linear superposition of a large number of independent random variables [...] which are all drawn from the same distribution" [2].

For this reason, deviations from perfect Gaussianity can provide relevant information on the early Universe and research on PNG is particularly important, especially if these initial conditions were generated by some dynamical process, such as, for example, inflation in the Early Universe.

Actually, while "small-amplitude curvature perturbations generated by quantum fluctuations in an inflationary phase [...] would yield a nearly Gaussian random density field" [2], direct measurements of non-Gaussianity would allow us to go beyond the free-field limit, providing information concerning the degrees of freedom, the possible symmetries and the interactions characterizing the inflationary action.

1˙1. *Historical outline*. – To be, or not to be Gaussian? The quest for Non-Gaussianity (NG) has a long story, already during the late seventies observations indicated that the patterns in the LSS could not be related to a Gaussian distribution. More precisely, NG in the LSS was measured in 1977 by Groth and Peebles [3] who computed the 3-point function of galaxies, raising the question whether this feature was only associated to non-linear gravitational clustering or it also included some signature of primordial NG.

In the subsequent years, and especially during the late eighties, the consequences of strongly non-Gaussian initial conditions were investigated in order to explore alternative structure formation models. However, these extreme possibilities were later excluded by CMB and LSS observations with increased accuracy.

Remarkably, the early nineties featured the beginning of a new era of non-Gaussian models from inflation, characterized by a small $f_{\rm NL}$, compatible with observations [4-8]. At the same time, N-body simulations started to play a crucial role determining the LSS of the Universe arising from the non-linear gravitational clustering of non-Gaussian Cold Dark Matter perturbations. The view on NG using N-body simulations around 1990 can be depicted in fig. 1. From the theoretical point of view, this line of research continued until the new millennium, when PNG finally emerged as a new "smoking gun" of (non-standard) inflation models [10,11], probing interactions among fields at the

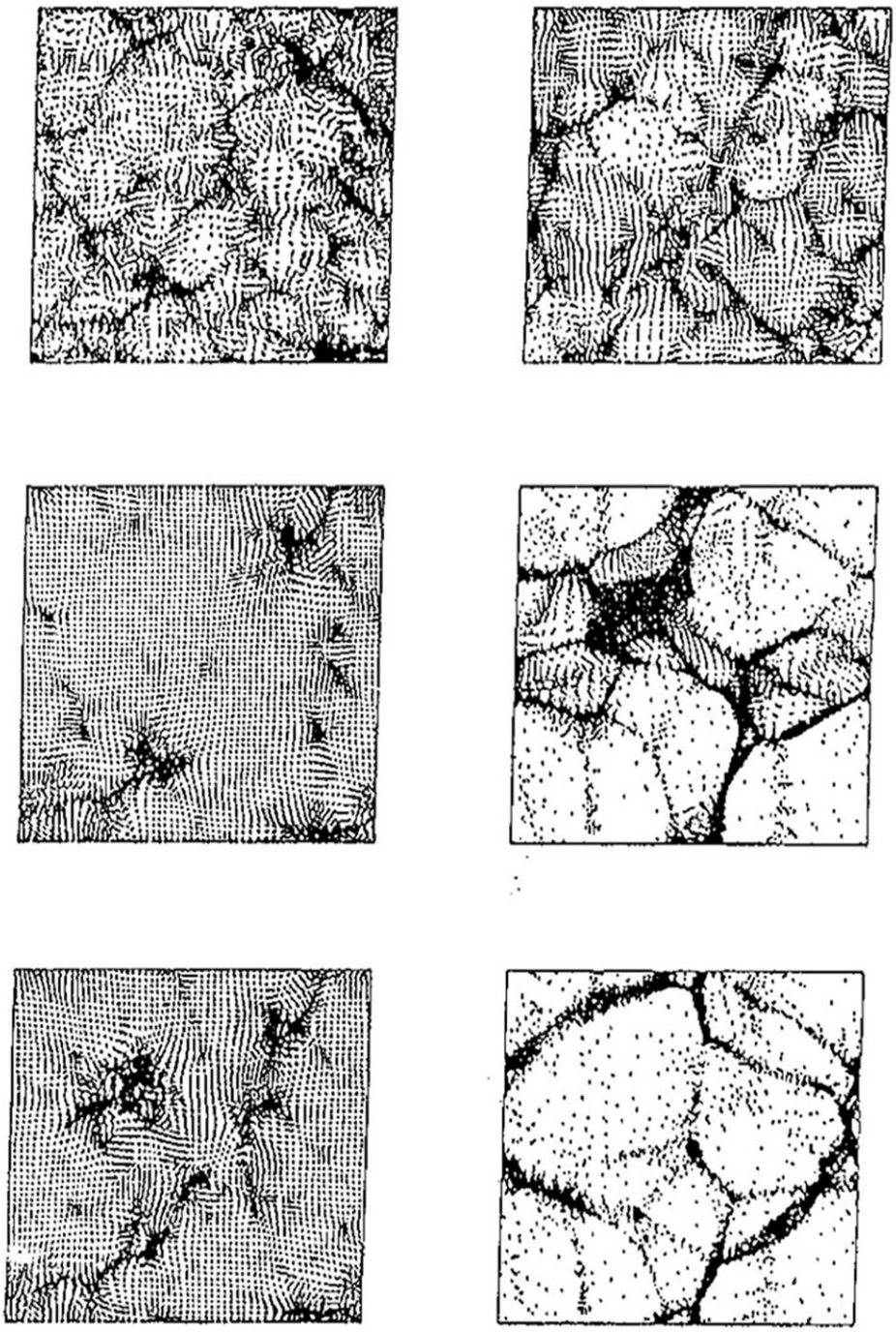


Fig. 1. – Projected particle positions in slices of depth one sixteenth of the computational box-size at the present time t_0. The slices refer to different models. From [9].

highest energy scales, which complements the search for primordial gravitational waves (PGW).

As for the experimental and observational side, the bispectra for PSCz [12] and the IRAS [13] redshift catalogues were determined in 2001, and for 2dF galaxies in 2002 [14]. More recently, the three-point correlation functions for the WiggleZ spectroscopic galaxy survey and the Baryon Oscillation Spectroscopic Survey were determined in [15] and [16, 17], respectively([1]).

Finally, the very stringent *Planck* constraints on PNG [26, 27] rose the question whether this route is still viable or not, and the present-day challenge is to detect (or constrain) mild or weak deviations from primordial Gaussian initial conditions.

2. – Non-Gaussianity in the initial conditions

The inflationary paradigm, a phase of accelerated expansion in the early Universe, was originally proposed at the end of the seventies in order to overcome some inconsistencies of the hot Big Bang model, which was plagued by the so-called flatness and horizon problems. At the same time, inflation suggested a quantum origin for the density fluctuations in the Universe, thereby providing a convincing dynamical mechanism for structure formation. Generally, the testable predictions of inflationary models are

- a critical value for the total energy density;
- almost, but not exact, scale-invariant and nearly Gaussian adiabatic density fluctuations;
- almost, but not exact, scale-invariant stochastic background of relic gravitational waves.

Note that *Planck* data have confirmed these predictions, for example the measured spectral index of the scalar power spectrum is $n_s = 0.9649 \pm 0.0042$ at 68% CL and no evidence for a scale dependence of n_s has been found [28].

Also spatial flatness is confirmed at a precision of 0.4% at 95% CL with the combination of BAO data [28]. Prospects for further improving measurements of spatial curvature are discussed in [29].

While primordial gravitational waves have not been yet detected, the upper limit on the tensor-to-scalar ratio from the BICEP2/Keck CMB polarization experiments is $r_{0.05} < 0.07$ at 95% confidence, which tightens to $r_{0.05} < 0.06$ in conjunction with *Planck* temperature measurements and other data [28, 30].

([1]) Note that even if the sensitivity is not competitive with CMB data [7], interesting bounds on local $f_{\rm NL}$ from current power spectrum constraints can be found in [18-25].

In order to reconstruct the inflationary action, we need two ingredients:

– the stochastic GW background, providing information on the inflationary energy scales;

– deviations from Gaussian initial conditions, providing information on the possible interactions. Moreover, PNG features can help us to distinguish inflation models which would yield the same predictions for n_s and r.

Many primordial (inflationary) models of non-Gaussianity can be represented in configuration space by the simple formula [4-8]

$$\Phi = \varphi_L + f_{\rm NL}\left(\varphi_L^2 - \left\langle\varphi_L^2\right\rangle\right) + g_{\rm NL}\left(\varphi_L^3 - \left\langle\varphi_L^2\right\rangle\varphi_L\right) + \dots, \tag{1}$$

where Φ is the large-scale gravitational potential (or equivalently in terms of the gauge-invariant comoving curvature perturbation ζ which on super-horizon scales satisfies the relation $\Phi = 3\,\zeta/5$), φ_L its linear Gaussian contribution and $f_{\rm NL}$ the dimensionless non-linearity parameter (or more generally non-linearity function).

2˙1. *Non-Gaussianity and higher-order statistics.* – The simplest statistics measuring NG is the 3-point function or its Fourier transform, the "bispectrum":

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\Phi(\mathbf{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B_\Phi(k_1,k_2,k_3), \tag{2}$$

which carries shape information. In the simple linear and quadratic model, parametrized by φ_L and $f_{\rm NL}$, the bispectrum of the gravitational potential reads

$$B_\Phi(k_1,k_2,k_3) = 2f_{\rm NL}[P_\Phi(k_1)P_\Phi(k_1) + \text{cyclic terms}] \tag{3}$$

where we applied Wick's theorem and

$$\langle\Phi(\mathbf{k}_1)\Phi(\mathbf{k}_2)\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2)P_\Phi(k_1). \tag{4}$$

In order to evaluate NG from the early Universe to the present time, we need, first of all, to calculate non-Gaussianity during inflation using a self-consistent method.

Then, we need to evolve scalar (vector) and tensor perturbations to second order outside the horizon, matching conserved second-order gauge-invariant variables, such as the comoving curvature perturbation $\zeta^{(2)}$ (or non-linear generalizations of it), to its value at the end of inflation (accurately accounting for reheating).

Finally, we consistently study the evolution of the perturbations after they re-entered the Hubble radius, by computing the second-order radiation transfer function for the CMB and the second-order matter transfer function for the LSS.

Although this procedure is involved, PNG represents a fundamental tool to probe fundamental physics (*e.g.* UV completion of the standard model of particle physics or general relativity such as string theory) during inflation at energies from the Grand

Unified Theories (GUT) scale $\sim 10^{15}\,\mathrm{GeV}$ to the Planck scale $\sim 10^{19}\,\mathrm{GeV}$, as different inflationary models predict different amplitudes and shapes of the bispectrum. For example, even tough standard models of slow-roll inflation predict tiny deviations from Gaussianity [4-8, 10, 11], consistent with the 2013 and 2015 *Planck* results, specific oscillatory PNG features can point to particular string-theory models as shown in [31, 32].

In conclusion, searching for PNG is interesting per-se for theoretically well-motivated models of inflation and, as shown in *Planck* 2015 results [27] (see also [28]), can severely limit various classes of inflationary models beyond the simplest paradigm.

2˙2. *Bispectrum of a self-interacting scalar field in de Sitter space.* – Consider a scalar field χ with cubic self-interactions, *i.e.* with an interaction term of the form $\lambda\chi^3/6$ in the Lagrangian.

Writing the field as $\chi = \chi_0 + \delta\chi$, where $\delta\chi$ represent the fluctuations around its vacuum expectation value $\chi_0 = \langle 0|\chi|0\rangle$, the two- and three-point functions in Fourier space [33, 6, 34], after the rescaling $\delta\chi = \delta\hat{\chi}/a$ are given by

$$\begin{aligned}(5)\qquad &\langle 0|\delta\hat{\chi}(\tau,\mathbf{k})\delta\hat{\chi}(\tau',\mathbf{k}')|0\rangle = \delta^{(3)}(\mathbf{k}+\mathbf{k}')G(k,\tau,\tau'),\\ &\langle \delta\hat{\chi}_{\mathbf{k}_1}\delta\hat{\chi}_{\mathbf{k}_2}\delta\hat{\chi}_{\mathbf{k}_3}\rangle = i\,\lambda\,\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\\ &\qquad\qquad \times\int_{-\infty}^{\tau}\frac{\mathrm{d}\tau'}{H\tau'}\left[\prod_{i=1}^{3}G(k_i,\tau,\tau') - \prod_{i=1}^{3}G^*(k_i,\tau,\tau')\right],\end{aligned}$$

where the Green's function reads

$$(6)\qquad G(k_i,\tau,\tau') = \frac{1}{2k_i}\left(1-\frac{i}{k_i\tau}\right)\left(1+\frac{i}{k_i\tau'}\right)e^{ik_i(\tau'-\tau)}.$$

The bispectrum is (ζ being a function of order 1)

$$(7)\qquad \langle \delta\chi_{\mathbf{k}_1}\delta\chi_{\mathbf{k}_2}\delta\chi_{\mathbf{k}_3}\rangle = \sum_i \nu_3(k_i)\prod_{j\neq i}\frac{H^2}{2k_j^3}\,,$$

where

$$(8)\qquad \nu_3(k_i) = \frac{\lambda}{3H^2}\left[\gamma+\zeta(k_i)+\log(-k_T\tau)\right].$$

Historically, in [33] it was found $f_{\mathrm{NL}}\sim\epsilon^2$ for the standard single-field slow-roll scenario (from non-linearity in the inflaton potential in a fixed de Sitter space-time).

Later, calculations from second-order gravitational corrections during stochastic inflation indicated $f_{\mathrm{NL}}\sim\epsilon,\eta$ [6]. This result has been confirmed in [10, 11], up to numerical factors and momentum-dependent terms, with a full second-order approach.

Finally, Weinberg extended the calculation of the bispectrum to 1-loop [35]. Remarkably, one of the terms gives rise to the so-called "consistency relation", according to which $f_{\mathrm{NL}} = -5\,(n_s-1)/12$.

However, it has been shown that the "consistency relation" term can be gauged away by a non-linear rescaling of coordinates, up to sub-leading terms. Hence the only residual term is proportional to ϵ *i.e.* to the amplitude of tensor modes; see comments on this point, later on.

2˙3. *Shapes of non-Gaussianity from inflation.* – In order to extract the relevant information regarding the amplitude and shape of PNG, it is convenient to write the bispectrum of primordial curvature perturbations as

$$\langle\zeta(\mathbf{k}_1)\zeta(\mathbf{k}_2)\zeta(\mathbf{k}_3)\rangle = (2\pi)^3\delta^{(3)}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)f_{\rm NL}F(k_1,k_2,k_3), \tag{9}$$

where $f_{\rm NL}$ represents the amplitude, while $F(k_1,k_2,k_3)$ encodes the shape of PNG. Note that, usually, we study the function $F(1,x_2,x_3)x_2^2x_3^2$ in terms of the rescaled coordinates $x_2=k_2/k_1$ and $x_3=k_3/k_1$, where momenta satisfy the triangle inequality $x_2+x_3>1$.

Remarkably, there are several possible shapes of non-Gaussianity from inflation, say more than ... stars in the sky. The most famous are:

– local NG, characteristic for multi-field, curvaton, ekpyrotic and cyclic models;

– equilateral NG, associated to non-canonical kinetic and higher-derivative terms, DBI and K-inflation, ghost inflation and EFT approaches;

– orthogonal NG, which distinguishes between variants of non-canonical kinetic term and higher-derivative interactions;

– flattened or folded NG.

More specifically, the bispectrum for the local shape [4, 6-8] peaks for squeezed triangles $k_3 \ll k_1 \sim k_2$, see fig. 1 in [36]. In this case, non-linearities develop outside the horizon during or immediately after inflation (*e.g.* multifield models of inflation). On the other hand, the bispectrum for the equilateral shape [38], see fig. 2, peaks for equilateral triangles $k_1=k_2=k_3$. Generally, in the equilateral family we can find single field models of inflation with non-canonical kinetic term $\mathcal{L}=P(\phi,X)$ with $X=-\frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi$ (*e.g.* DBI or K-inflation) where NG comes from higher-derivative interactions of the inflaton field, such as

$$\mathcal{L}\supset\delta\dot{\phi}(\nabla\delta\phi)^2. \tag{10}$$

Finally, the bispectrum for the flattened shape peaks for flattened (or folded) triangles $k_1=k_2+k_3$, see fig. 3, and can be written in terms of the equilateral and orthogonal shapes [39]. It is characteristic for excited initial states (see [40-42]), higher derivative interactions [37] or models where a Galilean symmetry is imposed [43].

However, there are many other shapes: *e.g.* directionally dependent bispectra, tensor bispectra, etc.

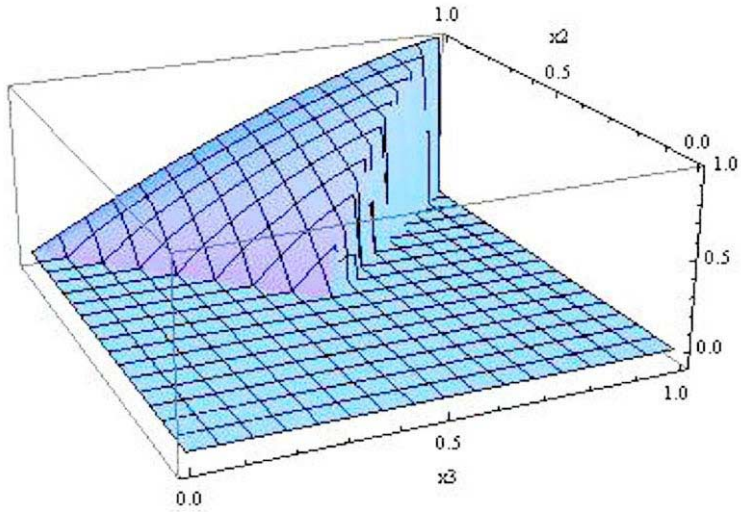


Fig. 2. – Plot of the function $F(1, x_2, x_3)x_2^2x_3^2$ for non-Gaussianities generated by higher derivative interactions. The figure is normalized to have value 1 for equilateral configurations $x_2 = x_3 = 1$ and set to zero outside the region $1 - x_2 \leq x_3 \leq x_2$. From [37].

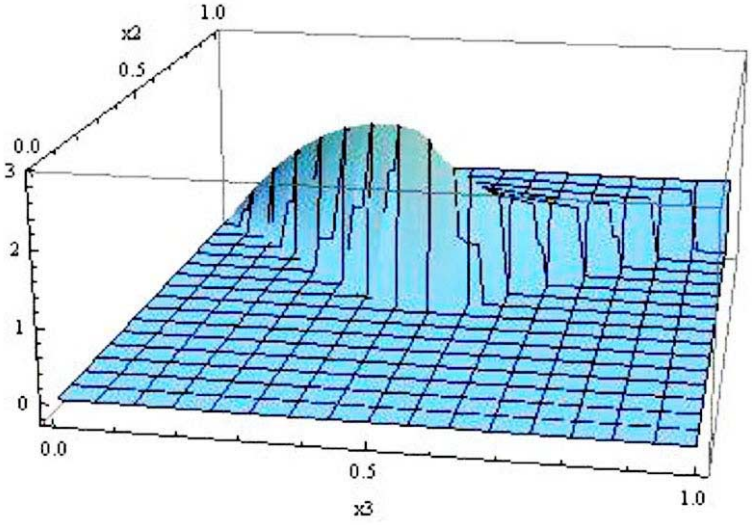


Fig. 3. – The folded template shape $F(x_2, x_3)x_2^2x_3^2$, the maximum is in the flat configuration ($k_1 = 1$, $k_2 = k_3 = \frac{1}{2}$). From [37].

2·4. *The role of $f_{\rm NL}$ and the detection of primordial non-Gaussianity.* – Clearly, detecting a non-zero primordial bispectrum (*e.g.* $f_{\rm NL} \neq 0$) proves that the initial seeds were non-Gaussian. Similarly, for the trispectrum and n-point correlation functions.

However, the opposite is not true, namely detecting $f_{\rm NL} \approx 0$ does not prove Gaussianity. Actually, there are infinitely many ways PNG can evade observational bounds optimized to search for $f_{\rm NL}$ and similar higher-order parameters.

As an example, consider the situation where the linear density contrast δ is non-Gaussian [44]. In this case, by the central-limit theorem, the gravitational potential Φ (which yields large-scale CMB anisotropies) tends to be much more Gaussian. Indeed, consider the non-Gaussian distribution of densities [44]

$$\delta(\mathbf{r}) = \int f(|\mathbf{r} - \mathbf{r}'|)\Delta(\mathbf{r}')\mathrm{d}^3\mathbf{r}', \tag{11}$$

where $\Delta(\mathbf{r})$ is an uncorrelated field with a gamma distribution, and f is chosen to give a Zel'dovich power spectrum ($P(k) \propto k$) for the density field. By solving Poisson's equation, we get for the gravitational potential

$$\Phi(\mathbf{r}) = -Ga^2\bar{\rho} \int \frac{\delta(\mathbf{r}')\mathrm{d}^3\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \tag{12}$$

with the resulting distribution very close to be Gaussian, see fig. 1(a) and 1(b) in [44].

3. – Non-Gaussianity and Cosmic Microwave Background

As mentioned before, the *Planck* satellite has provided accurate measurements of higher-order CMB correlations, resulting in very stringent constraints on PNG.

Planck is a project of the European Space Agency, with instruments provided by two scientific Consortia funded by ESA member states (in particular the lead countries: France and Italy) with contributions from NASA (USA), and telescope reflectors provided in a collaboration between ESA and a scientific Consortium led and funded by Denmark. The *Planck* satellite has measured the CMB temperature and polarization anisotropies with great accuracy, and, since PNG affects both, we can compare the data with the NG CMB simulated maps, shown in fig. 4.

The latest release regarding non-Gaussianity [27] tested the local, equilateral, orthogonal (and many more) shapes for the bispectrum and provided new constraints on the primordial trispectrum parameter $g_{\rm NL}$ (while $\tau_{\rm NL}$ was constrained in the previous release [26]). A new *Planck* legacy release, which will improve the 2015 results in terms of more refined treatment of E-mode polarization, is in preparation.

The standard representation used in the *Planck* analysis for the CMB bispectrum is

$$B^{m_1m_2m_3}_{\ell_1\ell_2\ell_3} \equiv \langle a_{\ell_1m_1} a_{\ell_2m_2} a_{\ell_3m_3} \rangle = \mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3} b_{\ell_1\ell_2\ell_3}, \tag{13}$$

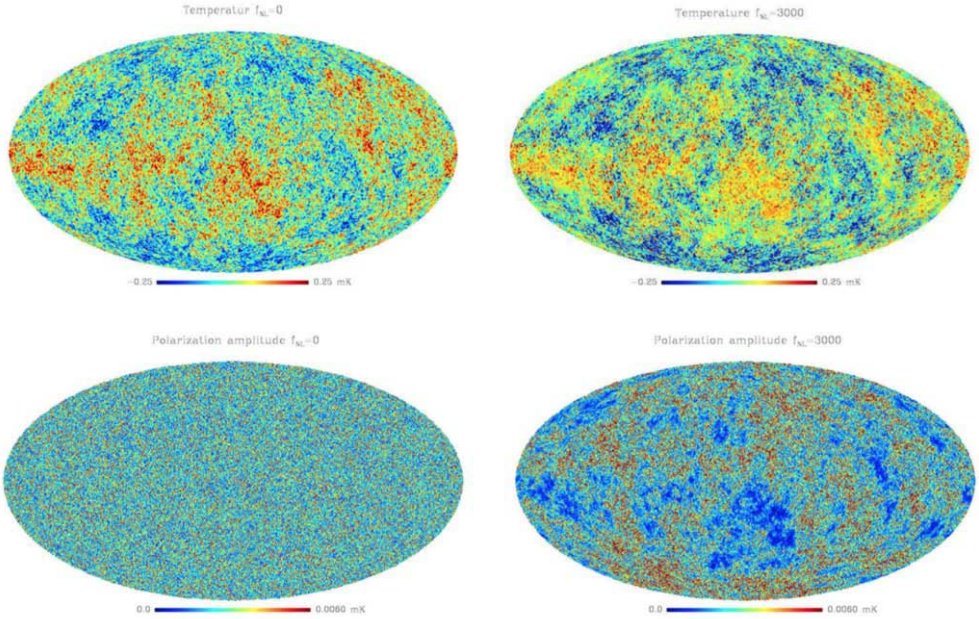


Fig. 4. – Left column: temperature and polarization intensity Gaussian CMB simulations. Right column: temperature and polarization non-Gaussian maps with the same Gaussian seed as in the left column and $f_{\rm NL} = 3000$. Temperatures are in mK. From [45].

where $\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3}$ are the Gaunt integrals

$$\mathcal{G}^{\ell_1\ell_2\ell_3}_{m_1m_2m_3} \equiv \int Y_{\ell_1m_1}(\hat{\mathbf{n}})Y_{\ell_2m_2}(\hat{\mathbf{n}})Y_{\ell_3m_3}(\hat{\mathbf{n}})\mathrm{d}^2\hat{\mathbf{n}} = h_{\ell_1\ell_2\ell_3}\begin{pmatrix}\ell_1 & \ell_2 & \ell_3\\ m_1 & m_2 & m_3\end{pmatrix} \tag{14}$$

and the ℓ_j satisfy the following conditions, see fig. 5:

- triangle condition: $\ell_1 \leq \ell_2 + \ell_3$ for $\ell_1 \geq \ell_2, \ell_3$ and permutations,
- parity condition: $\ell_1 + \ell_2 + \ell_3 = 2\,n$ with $n \in \mathbb{N}$,
- resolution: $\ell_1, \ell_2, \ell_3 \leq \ell_{max}$ with $\ell_1, \ell_2, \ell_3 \in \mathbb{N}$.

As noticed before, the search for PNG is optimized in terms of $f_{\rm NL}$. Leaving aside complications coming from breaking of statistical isotropy (sky-cut, noise, etc.), the general procedure is to fit the theoretical bispectrum template

$$\hat{f}_{\rm NL} = \frac{1}{N}\sum B^{m_1m_2m_3}_{\ell_1\ell_2\ell_3}\left[(C^{-1}a)^{m_1}_{\ell_1}(C^{-1}a)^{m_2}_{\ell_2}(C^{-1}a)^{m_3}_{\ell_3} - 3C^{-1}_{\ell_1m_1\ell_2m_2}(C^{-1}a)^{m_3}_{\ell_3}\right] \tag{15}$$

to the 3-point function obtained analyzing the data. Unfortunately, a brute force implementation scales like ℓ^5_{max}, unfeasible at *Planck* (or WMAP) resolution. On the other

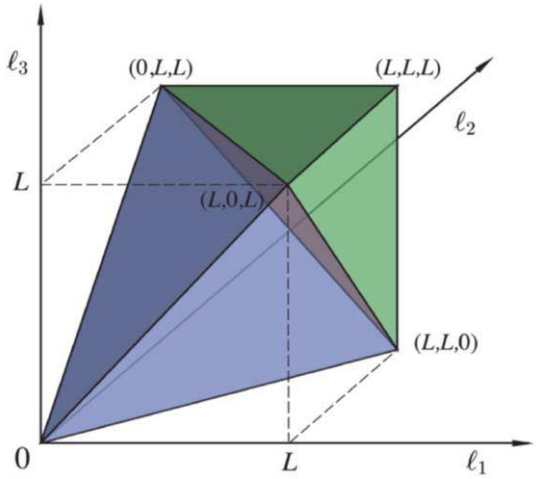


Fig. 5. – Permitted observational domain for the CMB bispectrum $b_{\ell_1\ell_2\ell_3}$. Allowed multipole values (ℓ_1, ℓ_2, ℓ_3) lie inside the shaded "tetrapyd" region (tetrahedron+pyramid), satisfying both the triangle condition and the experimental resolution. From [26].

hand, we can achieve a massive speed improvement (ℓ^3_{max} scaling) if the reduced bispectrum is separable. Generally, there are different ways to write the theoretical template in separable form:

– the KSW [46] separable template fitting and the Skew-Cl extension [47];

– the binned bispectrum presented in [48];

– the modal expansion described in [49].

The alternative implementations differ basically in terms of the separation technique adopted and of the projection domain.

More recent improvements can be found in [50] where the interested reader can find an exact expression for the multi-variate joint probability distribution function (PDF) of non-Gaussian fields, primordially arising from local transformations of a Gaussian field. This expression has been applied to the non-Gaussianity estimation from CMB maps and the halo mass function, obtaining both analytical expressions as well as approximations with specified range of validity.

The results for the CMB gave a fast way to compute the PDF, valid up to more than $7\,\sigma$ for $f_{\rm NL}$ values not ruled out by current observations, expressed as a combination of bispectrum and trispectrum of the temperature maps. Note that such expression is valid for any kind of non-Gaussianity and is not limited to the local type, providing a useful basis for a fully Bayesian analysis of the NG parameter.

Finally, note that, in principle, we could go to higher order. In fact, this may become important if we want to detect NG in observables characterized by a large $f_{\rm NL}$ (*e.g.* in

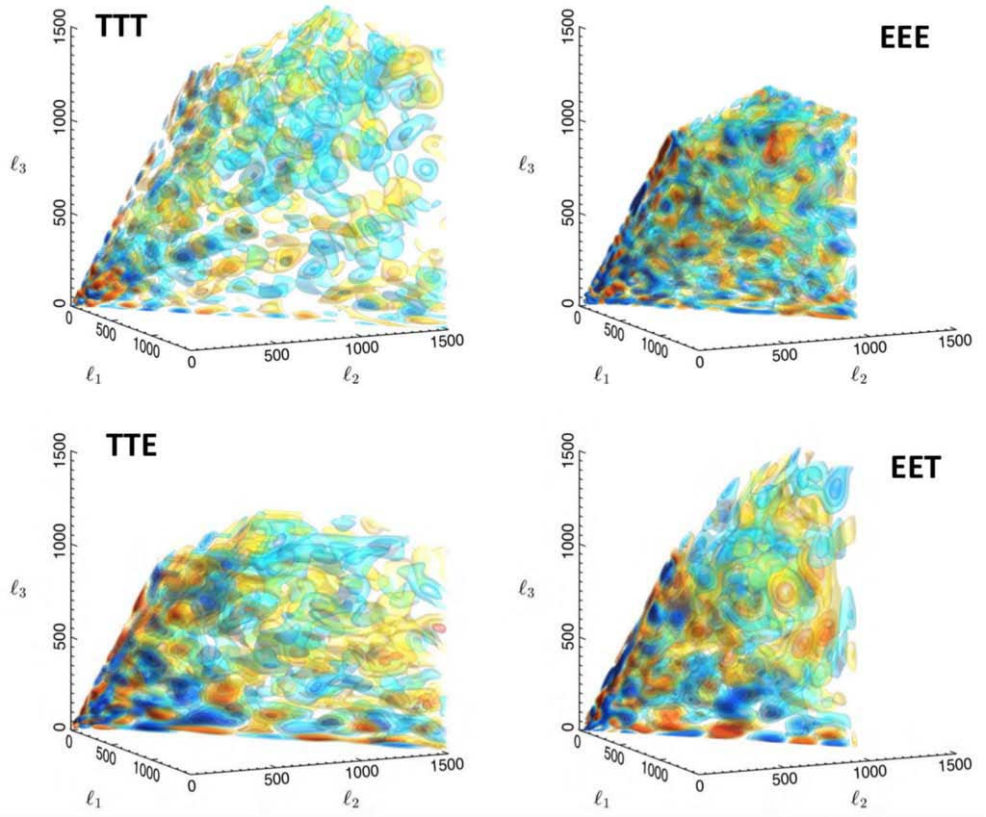


Fig. 6. – CMB temperature and polarization bispectrum reconstructions for *Planck* `SMICA` maps using the full set of polynomial modes with $n_{max} = 2001$ and with signal-to-noise weighting. From [27].

high-redshift probes) and/or if $f_{\rm NL}$ (leading-order bispectrum) and $g_{\rm NL}$ (leading-order trispectrum) are both depending on the same underlying physical coupling constant that we aim at determining. However, the expressions are rather involved and we refer the reader to [50] for details.

3˙1. *Planck results on primordial non-Gaussianity.* – In this section, we briefly present some of the *Planck* results on PNG, focusing on the improvements compared to the 2013 release and the differences between the methods used in the analysis, particularly concerning the Integrated Sachs-Wolfe (ISW) effect.

Let us start with the 2015 *Planck* analysis for the bispectrum in the modal representation, described in fig. 6, for the various combinations TTT, EEE, TTE, EET.

Note that compared to *Planck* 2013 the new constraints on local, equilateral, orthogonal bispectra have improved by up to 15%.

In the 2015 analysis particular attention was devoted to investigate the ISW-lensing effect which strongly affects the constraints on $f_{\rm NL}$ from *Planck* bispectrum. The results

TABLE I. – *Results for the amplitude of the lensing-ISW bispectrum from the* SMICA*,* SEVEM*,* NILC*, and* Commander *foreground-cleaned maps, for different bispectrum estimators. Error bars are* 68% *CL. From* [27].

Method	Lensing-ISW amplitude			
	SMICA	SEVEM	NILC	Commander
T				
KSW	0.79 ± 0.28	0.78 ± 0.28	0.78 ± 0.28	0.84 ± 0.28
Binned . .	0.59 ± 0.33	0.60 ± 0.33	0.68 ± 0.33	0.65 ± 0.36
Modal2 . .	0.72 ± 0.26	0.73 ± 0.26	0.73 ± 0.26	0.78 ± 0.27
$T + E$				
Binned . .	0.82 ± 0.27	0.75 ± 0.28	0.85 ± 0.26	0.84 ± 0.27

are confirmed using several $f_{\rm NL}$ estimators and CMB maps. Actually, different maps have been produced by the SMICA, NILC, SEVEM and Commander-Ruler (or C-R) pipelines. Notice that the SMICA product is considered the preferred one overall. Having said this, the amplitude of the ISW-lensing bispectrum from the SMICA, NILC, SEVEM, and C-R foreground-cleaned maps, for the KSW, binned, and modal (polynomial) estimators are summarized in table I. Remarkably, the coupling between weak lensing and ISW effect is the leading contamination to local NG, as the ISW lensing bispectrum has been detected with a significance of $2.8\,\sigma$ (see fig. 7), and improves to $3.0\,\sigma$ when including polarization. In conclusion, the bias in the three primordial $f_{\rm NL}$ parameters due to the ISW-lensing signal is described in table II.

Actually, the new analysis was the first adopting also polarization data for PNG, and the *Planck* 2015 constraints on $T+E$ have confirmed T results with significantly reduced error bars, see table III for the latest constraints on $f_{\rm NL}$ for the various shapes.

Specifically, the *Planck* 2013 hints of NG in oscillatory feature models remain in T, but decrease significantly when polarization is included. Also, new estimators for high-frequency oscillations cover 10 times more parameter space, compared to the previous analysis.

Remarkably, the improvements in the *Planck* 2015 results have allowed to put new constraints on:

– isocurvature NG, where polarizartion data were crucial in this respect;

– tensor NG, where parity-odd T limits are consistent with WMAP (null result);

– trispectrum due to cubic NG (in particular $g_{\rm NL}$ for a variety of shapes).

Finally, with the 2015 release we could also constraint the three fundamental shapes of

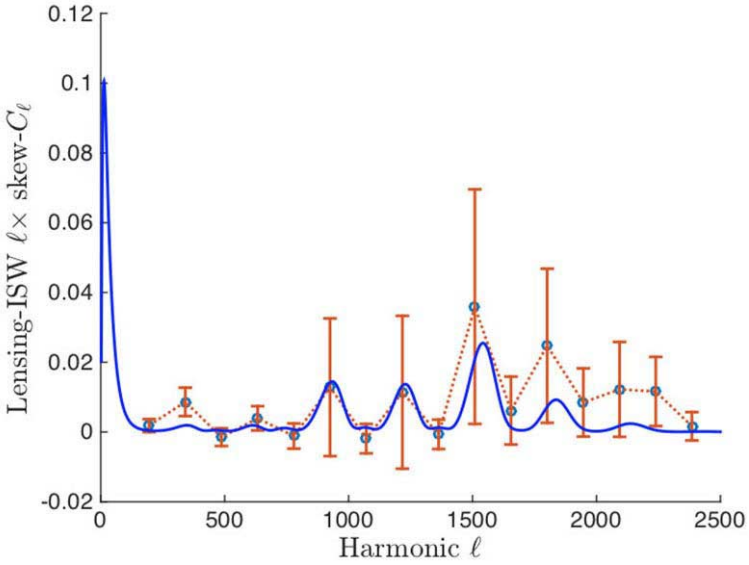


Fig. 7. – The skew-Cl spectrum for the lensing-ISW effect (red line with data points), from the temperature map. The blue curve is the theoretically-expected spectrum. From [27].

Table II. – *Bias in the three primordial $f_{\rm NL}$ parameters due to the lensing-ISW signal for the four component separation methods. From* [27].

Shape	Lensing-ISW $f_{\rm NL}$ bias			
	SMICA	SEVEM	NILC	Commander
T Local	7.5	7.5	7.3	7.0
T Equilateral	1.1	1.2	1.3	1.8
T Orthogonal	−27	−27	−26	−26
E Local	1.0	1.1	1.0	1.1
E Equilateral	2.6	2.7	2.5	2.9
E Orthogonal	−1.3	−1.3	−1.2	−1.5
$T+E$ Local	5.2	5.5	5.1	4.9
$T+E$ Equilateral	3.4	3.4	3.4	3.6
$T+E$ Orthogonal	−10	−11	−10	−10

TABLE III. – *Results for the $f_{\rm NL}$ parameters of the primordial local, equilateral, and orthogonal shapes, determined by the KSW estimator from the* `SMICA` *foreground-cleaned map. Error bars are* 68% *CL. From* [27].

Shape and method	$f_{\rm NL}$(KSW)	
	Independent	ISW-lensing subtracted
`SMICA` (T)		
Local	10.2 ± 5.7	$\mathbf{2.5 \pm 5.7}$
Equilateral	-13 ± 70	$\mathbf{-16 \pm 70}$
Orthogonal	-56 ± 33	$\mathbf{-34 \pm 33}$
`SMICA` ($T + E$)		
Local	6.5 ± 5.0	$\mathbf{0.8 \pm 5.0}$
Equilateral	3 ± 43	$\mathbf{-4 \pm 43}$
Orthogonal	-36 ± 21	$\mathbf{-26 \pm 21}$

the trispectrum

$$g_{\rm NL}^{\rm local} = (-9.0 \pm 7.7) \times 10^4, \tag{16}$$

$$g_{\rm NL}^{\dot\sigma^4} = (-0.2 \pm 1.7) \times 10^6,$$

$$g_{\rm NL}^{(\partial\sigma)^4} = (-0.1 \pm 3.8) \times 10^5.$$

In conclusion, the *Planck* 2015 release contains a largely extended analysis of NG templates, and we expect that the upcoming release (the "*Planck* legacy" paper) will further improve the constraints on standard shapes (owing to refined treatment of E-mode polarization maps) and add some extra shapes (such as scale-dependent $f_{\rm NL}$, conformal symmetry), looking for features both in the power spectrum and the bispectrum.

3˙2. *Implications for inflation.* – One of the most important consequence from *Planck* data is that the simplest inflationary models (standard inflation) are still alive . . . and in very good shape!

Specifically, for standard inflation we refer to a single scalar field ϕ (representing a single clock), characterized by a Bunch-Davies initial vacuum state and a canonical kinetic term $X = -\frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi$, performing a slow-roll dynamics by means of a potential $\mathcal{V}(\phi)$, minimally coupled to gravity, described by general relativity (GR)

$$S = \int {\rm d}^4x\sqrt{-g}\left[X - \mathcal{V}(\phi) + \mathcal{L}_{\rm int}(\phi, A_\mu, \Psi) + \frac{M_{\rm Pl}^2}{2}\,R\right], \tag{17}$$

where $\mathcal{L}_{\rm int}$ is the interaction term between the inflaton and other fields such as gauge bosons A_μ or fermions Ψ.

Actually, standard inflation predicts tiny ($\mathcal{O}(10^{-2})$, thus no presently detectable) PNG.

However, alternatives to standard inflation have also been considered. In particular, results from *Planck* 2015 (which increased the number of modes from 600 to 2000 with respect to *Planck* 2013) constrained $f_{\rm NL}$ for a large number of inflationary models including:

– the equilateral family (DBI, EFT, ghost and K-inflation);

– the flattened shapes (non-Bunch Davies);

– feature models (oscillatory or scale-dependent bispectra);

– direction dependence;

– quasi-single-field;

– parity-odd models.

Since no evidence for NG has been found, we could only put tighter constraints on the parameters from the models above, for example on:

– the curvaton decay fraction $r_D > 19\%$ (from local $f_{\rm NL}$, $T+E$),

– the speed of sound in the Effective Field Theory of Inflation [51] $c_S > 0.024$ (from equilateral and orthogonal $f_{\rm NL}$),

– the speed of sound in DBI inflation $c_S > 0.087$ (from $T+E$).

3˙3. *Primordial non-Gaussianity with CMB spectral distorsions.* – Possible measurements to improve the constraints on PNG include CMB spectral distorsions from acoustic wave dissipation that can probe a large range of scales, much more than CMB/LSS [52].

Actually, if μ-anisotropies were measured we would access to

– $T\mu$ correlations, useful to investigate the primordial local $f_{\rm NL}$ (see [53]) or other squeezed shapes, *e.g.*, excited initial states [54],

– $\mu\mu$ correlations, associated to the primordial local trispectrum, $\tau_{\rm NL}$ [55],

– $TT\mu$ bispectrum, related to the primordial local trispectrum $g_{\rm NL}$ [55].

While for Gaussian initial conditions the dissipated power in small patches (from $\approx 50\,{\rm Mpc}^{-1}$ to $\approx 10^4\,{\rm Mpc}^{-1}$) is isotropically distributed, squeezed bispectra associated to local NG generate couplings between large and small scales.

As a consequence of these coupling between long and short modes, the CMB temperature fluctuations on large scales can be coupled to spectral distortions arising from acoustic wave dissipation at very small scales, resulting in $T\mu$ correlations [56, 53, 57].

More specifically, following [57], consider the curvature perturbation at position $\vec{x}$ in terms of a Gaussian random variable $z(\vec{x})$

$$\zeta(\vec{x}) = z(\vec{x}) + \frac{3}{5} f_{\rm NL}^{\rm loc} z^2(\vec{x}). \tag{18}$$

Splitting $\zeta(\vec{x})$ in long-wavelength and short-wavelength modes as $\zeta(\vec{x}) = \zeta_L(\vec{x}) + \zeta_S(\vec{x})$, and similarly writing $z(\vec{x}) = z_L(\vec{x}) + z_S(\vec{x})$ we get

$$\zeta_L + \zeta_S = z_L + z_S + \frac{3}{5} f_{\rm NL}^{\rm loc} [z_L^2 + 2 z_L z_S + z_S^2], \tag{19}$$

and we conclude that, in the presence of $f_{\rm NL}^{\rm loc} \neq 0$, long and short modes can be coupled. In particular, to linear order in $f_{\rm NL}^{\rm loc}$, the small-scale curvature fluctuation in the presence of some fixed long-wavelength curvature fluctuation is

$$\zeta_S = z_S \left(1 + \frac{6}{5} f_{\rm NL}^{\rm loc} \zeta_L \right) \tag{20}$$

thus modulating the (fractional) chemical-potential fluctuation, given by [56-58]

$$\frac{\delta\mu}{\mu} \approx \frac{\delta\langle\zeta^2\rangle}{\langle\zeta^2\rangle} \approx \frac{12}{5} f_{\rm NL}^{\rm loc} \zeta_L. \tag{21}$$

Similarly, for the large-angle ($\ell \lesssim 100$, probing causally disconnected regions at the last scattering surface) the temperature fluctuation is determined primarily by the curvature fluctuation at the surface of last scatter, given by $\delta T/T \approx \zeta/5$ [56, 57].

As a consequence, the fractional chemical potential fluctuation $\frac{\delta\mu}{\mu}$ and the temperature fluctuation $\frac{\delta T}{T}$ are cross correlated with an angular power spectrum $C_\ell^{\mu T}$ equal to [56, 57]

$$C_\ell^{\mu T} = 12 f_{\rm NL}^{\rm loc} C_\ell^{TT}. \tag{22}$$

4. – Primordial Non-Gaussianity and the Large-Scale Structure

Non-Gaussianity in Large-Scale Structure can be either of primordial origin or associated to gravitational instability.

In particular, to make contact with the CMB definition, PNG in LSS can be defined starting from the DM density fluctuation δ through the Poisson's equation

$$\delta = -\left(\frac{3}{2} \Omega_m H^2 \right)^{-1} \nabla^2 \Phi \tag{23}$$

where we have used the comoving gauge for density fluctuation [59].

As before, we can write

$$\Phi = \phi_L + f_{\rm NL}(\phi_L^2 - \langle\phi_L^2\rangle) + g_{\rm NL}(\phi^3 - \langle\phi_L^2\rangle\phi_L) + \dots, \tag{24}$$

where ϕ_L is the linear Gaussian contribution and $f_{\rm NL}$ and $g_{\rm NL}$ are dimensionless non-linearity parameters([2]).

In order to investigate PNG in LSS, generally N-body simulations have been playing a crucial role [60-67]. The standard equations are

$$\begin{aligned} \Phi &= \phi_L + f_{\rm NL}(\phi_L^2 - \langle\phi_L^2\rangle), \\ \nabla^2(\Phi * T)g(z) &= -4\pi G a^2 \delta\rho_{DM}, \end{aligned} \tag{25}$$

where T is the matter transfer function and g is the growth suppression factor. Typical results are shown in fig. 8.

However, in the mild non-linear regime, analytical approaches have been developed to study the PNG effects on the matter power spectrum. Results for the local shape calculated using the Time Renormalization Group theory [68, 69] compared to N-body are shown in fig. 9, while for the equilateral and folded shapes see fig. 10.

Similar techniques include the renormalized perturbation theory [70-72], renormalization group approach [73], closure theory [74], Lagrangian perturbation theory [75-77], the time-sliced perturbation theory [78] and the Effective Field Theory of LSS (EFTofLSS) [79, 80]. Specifically, the EFTofLSS for non-Gaussian initial conditions have been developed in [81], see also [82-88].

4˙1. *Non-Gaussianity and halo mass function.* – Besides using the standard statistical estimators, like the (mass) bispectrum, trispectrum, etc., one can look at the tails of the distribution, *i.e.* at rare events.

Rare events have the advantage that they often maximize deviations from what is predicted by a Gaussian distribution, but have the obvious disadvantage of being rare! But remember that, according to the Press-Schechter-like schemes, all collapsed DM halos correspond to (rare) high peaks of the underlying density field.

In [91, 92] it was shown that clusters at high redshift ($z > 1$) can probe NG down to $f_{\rm NL} \sim 10^2$, see also [93] for an alternative approach. Actually, many methods have been developed for the determination of mass function, such as

– the stochastic approach (first-crossing of a diffusive barrier) [94-97],

– the ellipsoidal collapse method [98, 99],

– a combination of saddle-point and diffusive barrier [100],

– the Log-Edgeworth expantion [101],

– the excursion sets studied with correlated steps [102, 103].

([2]) CMB and LSS conventions may differ by a factor 1.3 for $f_{\rm NL}$, $(1.3)^2$ for $g_{\rm NL}$.

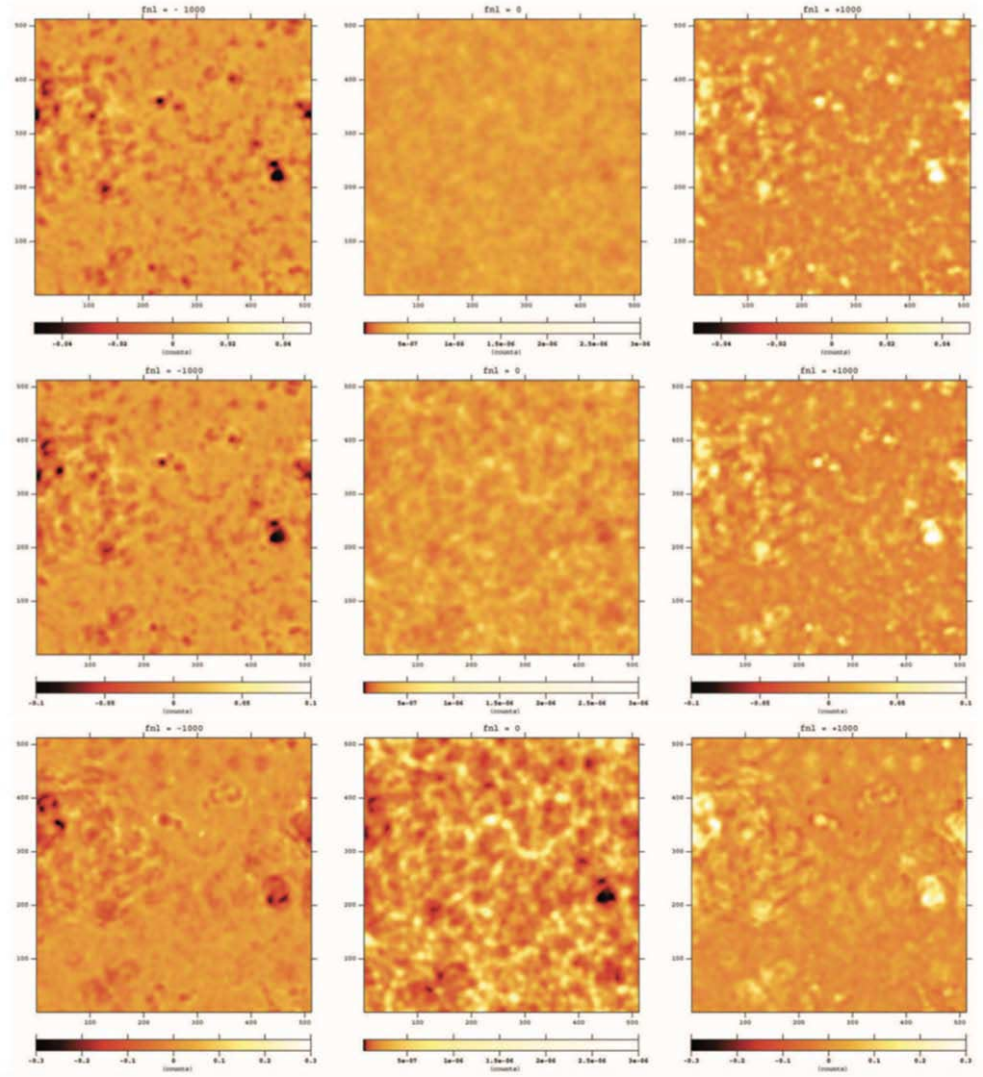


Fig. 8. – Slice maps of simulated mass density fields at $z = 5.15$ (top), $z = 2.13$ (middle) and $z = 0$ (bottom). The number of pixels at a side length is 512 ($500h^{-1}$ Mpc) and that of the thickness is 32 ($31.25h^{-1}$ Mpc). The panels in the middle row show the log of the projected density smoothed with a Gaussian filter of 10 pixels width, corresponding to $9.8h^{-1}$ Mpc. The left and right panels are the relative residuals for the $f_{\rm NL} = \pm 1000$ runs. Each panel has the corresponding color bar and the range considered are different from panel to panel. From [61].

Remarkably, excellent agreement of analytical formulae with N-body simulations (*e.g.*, see fig. 11 for DM halos in NG simulations and fig. 12 for the ration of the non-Gaussian $f_{\rm NL}$ to Gaussian mass function) have been found in [89,64,62], and in many other papers afterwards, *e.g.*, [104-106].

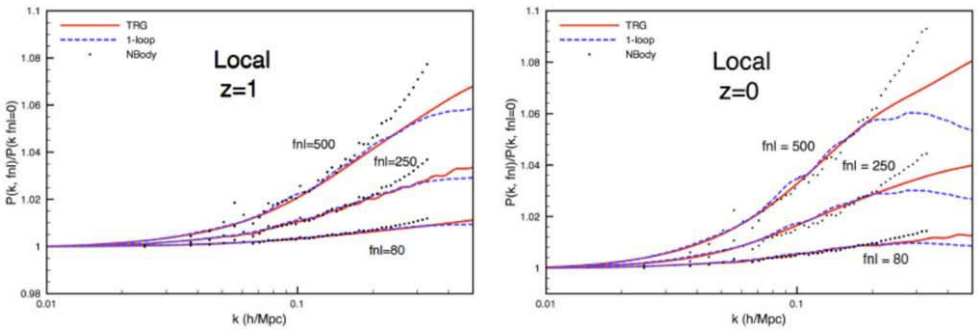


Fig. 9. – Ratio of the non-Gaussian to Gaussian power spectrum for several values of $f_{\rm NL}$ in the local model. The dots correspond to the data from the N-body simulations of [89]. The red (continuous) line is the TRG result of this paper and the blue (dashed) line is the one-loop result. From [90].

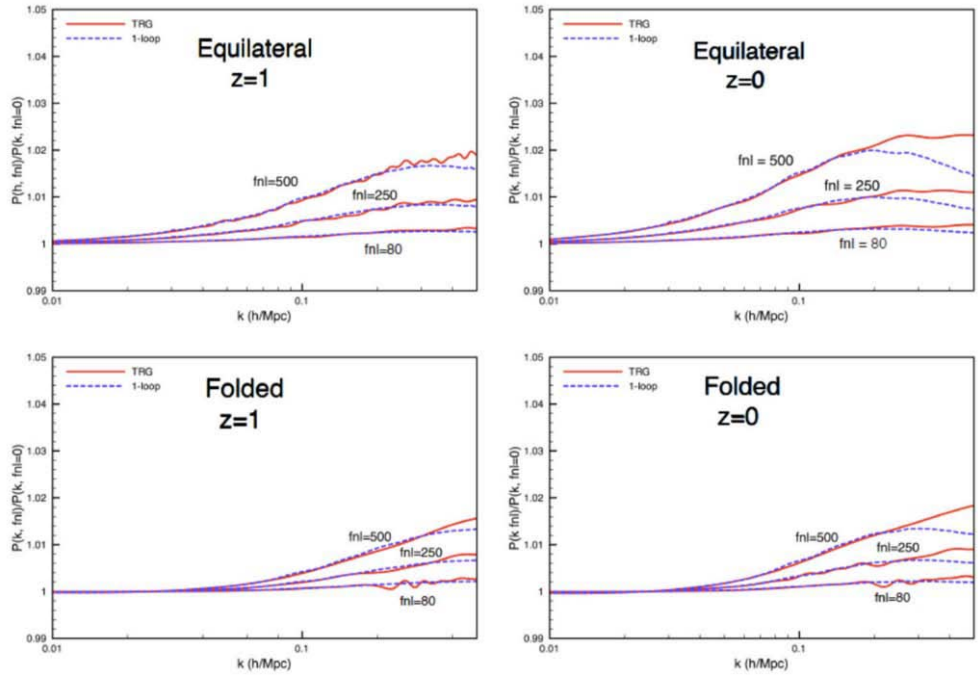


Fig. 10. – Ratio of the non-Gaussian to Gaussian power spectrum for several values of $f_{\rm NL}$ in the equilateral (top panels) and folded (bottom panels) models. The red (continuous) lines are the TRG result of this paper and the blue (dashed) lines are the one-loop result. From [90].

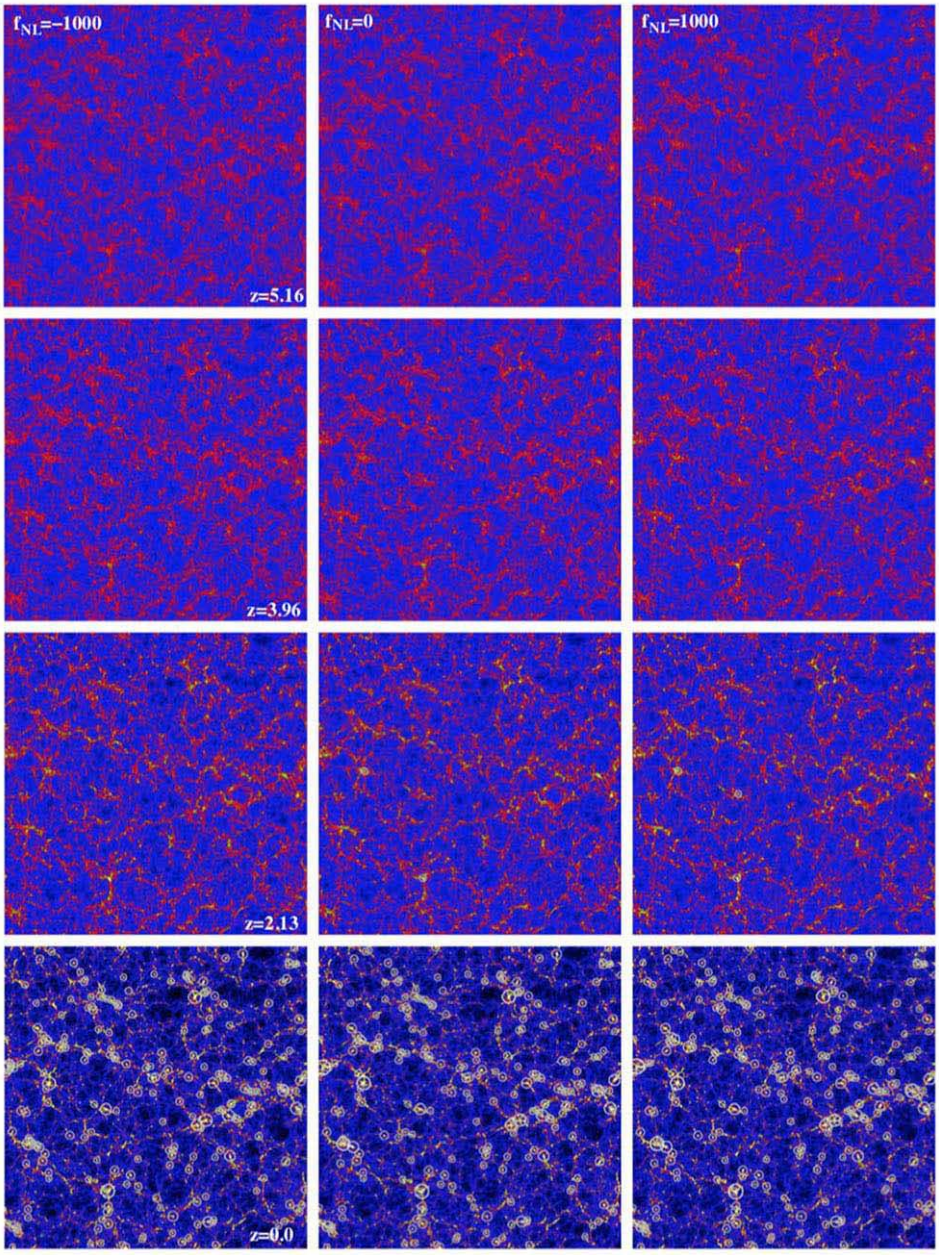


Fig. 11. – Mass density distribution and halo positions in a slice cut across the simulation box. The color-coded contours indicate different density levels ranging from dark (deep blue) underdense regions to bright (yellow) high density peaks. The halo positions are indicated by open circles with size proportional to their masses. Left panels: NG model with $f_{\rm NL} = -1000$. Central panels: Gaussian model. Right panels NG model with $f_{\rm NL} = +1000$. The mass and halo distributions are shown at various epochs, characterized by increasing redshifts (from bottom to top), as indicated in the panels. From [60].

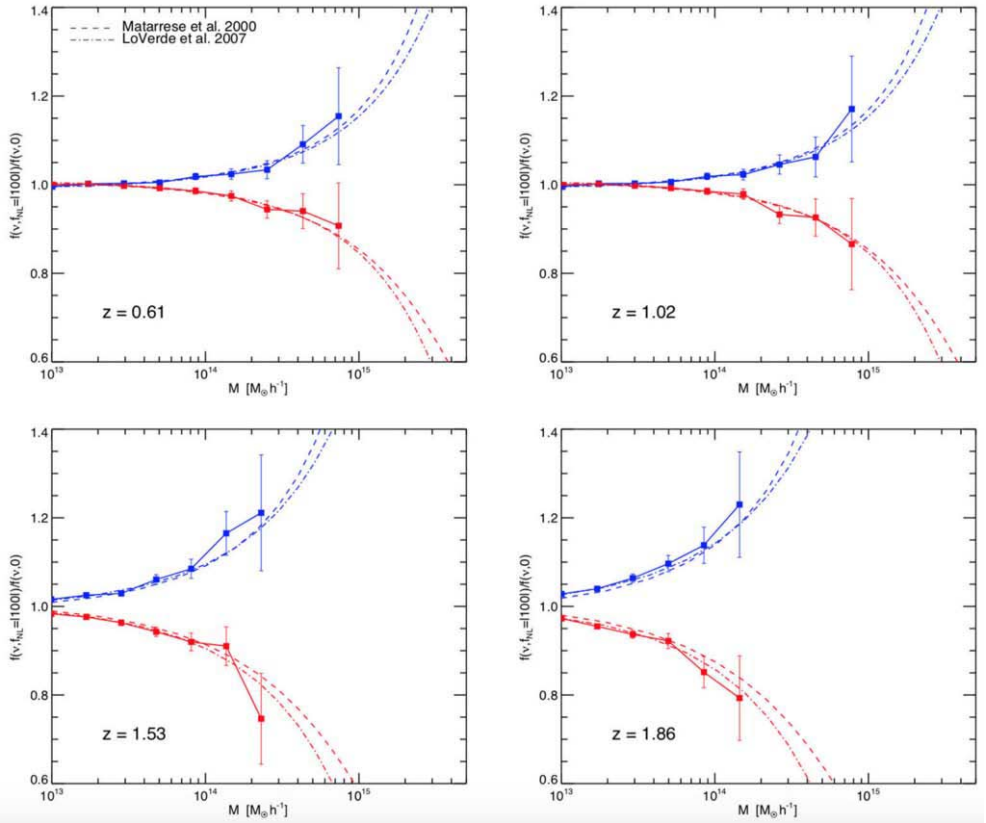


Fig. 12. – Ratio of the non-Gaussian ($f_{\rm NL} = \pm 100$) to Gaussian mass function for different redshift snapshots: top left $z = 0.61$; top right $z = 1.02$; bottom left $z = 1.53$; bottom right $z = 1.86$. The dashed line is the mass function of [92] and the dot-dashed lines are that of [93], both including the q-correction. From [64].

Moreover, halo (galaxy) clustering and halo (galaxy) higher-order correlation functions represent further and more powerful implementations of this general idea.

Actually, the halo mass function (à-la-Press-Schechter) can be a useful tool to probe PNG as it essentially depends exponentially on the PNG parameters [92], by modulating the critical overdensity for collapse. Its calculation can be done along the lines of the original Press-Schechter approach, using a steepest-descent approximation to deal with (small) PNG. However, several effects have to be carefully considered such as non-Markovianity, already there in the Gaussian case, but unavoidable in NG case, or the details of the non-spherical collapse. While analytical treatments are welcome, the validation with N-body simulations is crucial. Still we need to better understand the connection between analytical and numerical quantities and real observables, and to what level is this affecting NG measurements.

– Should we necessarily go on with (extended) Press-Schechter-like approaches?

– Are alternative approaches viable such as Smoluchowski equation for the non-Poissonian random process (or the earliest attempt proposed in [107])?

In conclusion, rare events (*e.g.*, high-z and massive clusters) offer interesting and promising opportunities for the detection of PNG, as both the mass-function of massive haloes and the number-counts of massive haloes are affected.

4˙2. *Halo bias in NG models.* – As it is well known, halos (galaxies) do not trace the underlying (dark) matter distribution. For this reason, following the original proposal [108], we introduce the "bias" parameters or Eulerian bias for galaxy clusters and later for galaxies (for a review see [109]):

$$\delta_{\rm halo}(x) = b_1 \delta_{\rm matter}(x) + b_2 \delta^2_{\rm matter}(x) + \dots \tag{26}$$

that allow to parametrize our ignorance about the way in which dark matter halos cluster in space with respect to the underlying dark matter.

Note that a complete set of the local bias terms (representing all possible local gravitational observables along the fluid trajectory) was presented in [110-112]. For the most general expansion up to second order in the Eulerian framework with Gaussian initial condition see also [113-121]. The various bias parameters can be generally regarded either as purely phenomenological ones (*i.e.* to be fitted to observations) or predicted by a theory (*e.g.* Press-Schecter together with Lagrangian perturbation theory).

Specifically, considering $\delta_{\rm halo}(x) = b\ \delta_{\rm matter}(x)$, it is possible to show that the halo bias is sensitive to PNG through a scale-dependent correction term (in Fourier space), see *e.g.* [122, 19, 123-129]. In particular, we have

$$\frac{\delta b(k)}{b} \sim \frac{2 f_{\rm NL} \delta_c}{k^2}\,. \tag{27}$$

This opens interesting prospects for constraining or measuring NG in LSS but demands for an accurate evaluation of the effects of (general) NG on halo biasing.

The idea is to start from the results obtained in the 80s in [130,131] giving the general expression for the peak 2-point function as a function of N-point connected correlation functions of the background linear (*i.e.* Lagrangian) mass-density field

$$\xi_{h,M}(|\mathbf{x}_1 - \mathbf{x}_2|) = -1 + \exp\left\{ \sum_{N=2}^{\infty} \sum_{n=1}^{N-1} \frac{\nu^N \sigma_R^{-N}}{j!(N-1)!} \xi^{(N)} \left[\underset{j \text{ times}}{\mathbf{x}_1, \dots, \mathbf{x}_1}, \underset{(N-j) \text{ times}}{\mathbf{x}_2, \dots, \mathbf{x}_2} \right] \right\}, \tag{28}$$

which requires many techniques such as the path-integral, the cluster expansion, the multinomial theorem and asymptotic expansion. The analysis of NG models was motivated in [132] on bulk flows.

In [123] this relation was applied to the case of NG of the gravitational potential, obtaining the power-spectrum of dark matter halos modeled as high "peaks" (up-crossing

regions) of height $v = \delta_c/\sigma_R$ of the underlying mass density field (Kaiser's model). Here $\delta_c(z)$ is the critical overdensity for collapse (at redshift z) and σ_R is the root mean square mass fluctuation on scale R ($M \sim R^3$).

The motion of peaks (going from Lagrangian to Eulerian space), which implies [116]

$$1 + \delta_h(\mathbf{x}_{\rm Eulerian}) = (1 + \delta_h(\mathbf{x}_{\rm Lagrangian}))(1 + \delta_R(\mathbf{x}_{\rm Eulerian})) \tag{29}$$

and (to linear order) $b = 1 + b_L$ [133], allows to derive the scale-dependent halo bias in the presence of NG initial conditions. Corrections may arise from second-order bias and GR terms.

Alternative approaches (*e.g.* based on 1-loop calculations) have been developed in [74,134,76,135]. Improvements in the fit with N-body simulations by assuming dependence on gravitational potential have been carried out in [63], while for the extension to bispectrum see [136]. Finally, the inclusion of $g_{\rm NL}$ and $f_{\rm NL}$ in analysis of QSO clustering was performed in [25].

Note that the extension to general (scale and configuration dependent) NG is straightforward [123]. Actually, we can write, in full generality, the f bispectrum as $B_f(k_1, k_2, k_3)$. Then, the relative NG correction to the halo bias is

$$\frac{\Delta b_h}{b_h} = \frac{\Delta_c(z)}{D(z)} \frac{1}{8\pi^2\sigma_R^2} \int \mathrm{d}k_1\, k_1^2 \mathcal{M}_R(k_1) \times \int_{-1}^{1} \mathrm{d}\mu \frac{\mathcal{M}_R(\sqrt{\alpha})}{\mathcal{M}_R(k)} \frac{B_\phi(k_1, \sqrt{\alpha}, k)}{P_\phi(k)}\,, \tag{30}$$

where $\alpha = k_1^2 + k_2^2 + 2k_1k\mu$, P_ϕ is the power-spectrum of a Gaussian gravitational potential, while $\mathcal{M}_R$ is the factor connecting the smoothed linear overdensity with the primordial potential by means of the factor

$$\mathcal{M}_R(k) = \frac{2}{3} \frac{T(k)k^2}{H_0^2\Omega_{m,0}} W_R(k), \tag{31}$$

where $T(k)$ is the transfer function and $W_R(k)$ is the window function defining the radius R of a proto-halo of mass $M(R)$. It also applies to non-local (*e.g.* "equilateral") PNG (corresponding to DBI or ghost inflation) and universal PNG term (see also [126, 137-139]). The halo bias in NG models has been calculated in [123], the result is

$$b_h^{f_{\rm NL}} = 1 + \frac{\Delta_c(z)}{\sigma_R^2 D^2(z)} \left[1 + 2 f_{\rm NL} \frac{\Delta_c(z)}{D(z)} \frac{\mathcal{F}_R(z)}{\mathcal{M}_R(z)}\right], \tag{32}$$

where the form factor is given by

$$\mathcal{F}_R(k) = \frac{1}{8\pi^2\sigma_R^2} \int \mathrm{d}k_1 k_1^2 \mathcal{M}_R(k_1) P_\phi(k_1) \int_{-1}^{1} \mathrm{d}\mu \mathcal{M}_R(\sqrt{\alpha}) \left[\frac{P_\phi(\sqrt{\alpha})}{P_\phi(k)} + 2\right] \tag{33}$$

and plotted, for three different masses, in fig. 13.

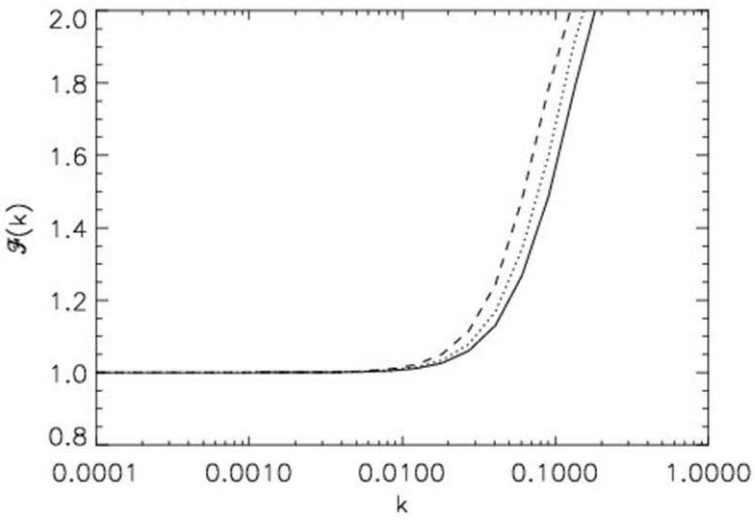


Fig. 13. – The function $\mathcal{F}_R(k)$ for three different masses: $1 \times 10^{14} M_\odot$ (solid), $2 \times 10^{14} M_\odot$ (dotted), $1 \times 10^{14} M_\odot$ (dashed). From [123].

4˙3. *PNG with LSS: the galaxy bispectrum.* – The bispectrum of galaxies can be used to forecast the constraining power of LSS surveys on measuring the amplitude of PNG, see *e.g.* [140, 141]. In this section, we derive, following [142], the galaxy bispectrum.

The starting point is the relation between the linearly evolving density field δ_{lin} and the primordial gravitational potential

$$\delta_{lin}(\mathbf{k}, z) = \alpha(k, z)\Phi_{in}(\mathbf{k}) \tag{34}$$

with $\alpha(k, z)$ defined as

$$\alpha(k, z) \equiv \frac{2k^2 c^2 T(k) D(z)}{3\Omega_m H_0^2}\,, \tag{35}$$

where $T(k)$ is the transfer function. Since the linearly evolving density includes non-Gaussian terms in presence of PNG, we can define the Gaussian part as

$$\delta_G(\mathbf{k}, z) = \alpha(k, z)\varphi_G(\mathbf{k}). \tag{36}$$

With this definition, using second-order (Eulerian) perturbation theory, we get [143]

$$\begin{aligned}\delta^{(2)}(\mathbf{k}, z) = &\int \frac{\mathrm{d}\mathbf{k}_1}{(2\pi)^3} \int \frac{\mathrm{d}\mathbf{k}_2}{(2\pi)^3} \delta^D(\mathbf{k} - \mathbf{k}_1 - \mathbf{k}_2) \\ &\times \left[\mathcal{F}_2(\mathbf{k}_1, \mathbf{k}_2) + f_{\mathrm{NL}} \frac{\alpha(k)}{\alpha(k_1)\alpha(k_2)}\right] \delta_G(\mathbf{k}_1, z)\delta_G(\mathbf{k}_2, z),\end{aligned} \tag{37}$$

where the 2-nd order gravity-kernel $\mathcal{F}_2$ is given by

$$\mathcal{F}_2(\mathbf{k}_1, \mathbf{k}_2) = \frac{5}{7} + \frac{1}{2}\frac{\mathbf{k}_1 \cdot \mathbf{k}_2}{k_1 k_2}\left(\frac{k_1}{k_2} + \frac{k_2}{k_1}\right) + \frac{2}{7}\frac{(\mathbf{k}_1 \cdot \mathbf{k}_2)^2}{k_1^2 k_2^2}\,. \tag{38}$$

The quantity $\delta(x)$ is expressed in the Eulerian frame, with the initial spatial coordinate $\mathbf{q}$ in the Lagrangian frame being related to the evolved Eulerian coordiate $\mathbf{x}$ through the formula

$$\mathbf{x}(\mathbf{q}, \tau) = \mathbf{q} + \mathbf{\Psi}(\mathbf{q}, \tau), \tag{39}$$

where $\mathbf{\Psi}$ is the displacement field. Using this relation, we can rewrite the second-order solution appearing in (37) as [144, 145]

$$\delta^{(2)}(\mathbf{x}, \tau) = \frac{17}{21}(\delta_{lin}(\mathbf{x}, z))^2 + \frac{2}{7}s^2(\mathbf{x}, z) - \mathbf{\Psi}(\mathbf{x}, z) \cdot \nabla\delta(\mathbf{x}, z), \tag{40}$$

where $s^2 = s_{ij}s^{ij}$ and s_{ij} is the trace-free tidal tensor, defined as

$$s_{ij} \equiv \left(\nabla_i \nabla_j - \frac{1}{3}\delta^K_{ij}\nabla^2\right)\nabla^{-2}\delta \tag{41}$$

and δ^K_{ij} is the Kronecker delta. In the following, we will omit the redshift dependence in the density and velocity fields.

We can introduce a long-short splitting of the gravitational potential and DM density field, such that, for local NG one easily finds (for local NG)

$$\delta_{lin,\ell}(\mathbf{k}) = \delta_{G,\ell} + f_{\rm NL}\alpha(\varphi^2_{G,\ell} - \langle\varphi^2_{G,\ell}\rangle) \tag{42}$$

with

$$\varphi_G(\mathbf{q}) = \varphi_{G,\ell}(\mathbf{q}) + \varphi_{G,s}(\mathbf{q}). \tag{43}$$

In the Lagrangian space one can then introduce the expansion

$$\begin{aligned}\delta^L_g(\mathbf{q}) &= \frac{n_g(\mathbf{q}) - \langle n_g\rangle}{\langle n_h\rangle} \\ &= b^L_{10}\delta_{lin} + b^L_{01}\varphi_G + b^L_{20}(\delta_{lin})^2 + b^L_{11}\delta_{lin}\varphi_G + b^L_{02}\varphi^2_G + \ldots,\end{aligned} \tag{44}$$

where the 5 b_{ij} represent our (generally unknown) bias parameters.

The final Eulerian position of the galaxy can be obtained by using the conservation law [116]

$$1 + \delta^E_g(\mathbf{x}, z) = [1 + \delta(\mathbf{x}, z)][1 + \delta^L_g(\mathbf{q}, z)] \tag{45}$$

to find

$$\delta_g^E(\mathbf{k}) = \delta_{10}^E\delta + b_{01}^E\varphi_G + b_{20}^E\delta*\delta + b_{11}^E\delta*\varphi_G + b_{02}^E\varphi_G*\varphi_G - \frac{2}{7}b_{10}^L s^2 - b_{01}n^2, \tag{46}$$

where s^2 and n^2 are suitable expansion terms and the Eulerian bias parameters read:

$$\begin{aligned} b_{10}^E &= 1 + b_{10}^L, \\ b_{01}^E &= b_{01}^L, \\ b_{20}^E &= \frac{8}{21}b_{10}^L + b_{20}^L, \\ b_{11}^E &= b_{01}^L + b_{11}^L, \\ b_{20}^E &= b_{02}^L. \end{aligned} \tag{47}$$

Using the standard definitions for the galaxy power spectrum P_{gg} and bispectrum B_{ggg}

$$\begin{aligned} \left\langle \delta_g^E(\mathbf{k}_1)\delta_g^E(\mathbf{k}_2)\right\rangle &= (2\pi)^3\delta^D(\mathbf{k}_1+\mathbf{k}_2)P_{gg}(\mathbf{k}_1), \\ \left\langle \delta_g^E(\mathbf{k}_1)\delta_g^E(\mathbf{k}_2)\delta_g^E(\mathbf{k}_3)\right\rangle &= (2\pi)^3\delta^D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)B_{ggg}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3). \end{aligned} \tag{48}$$

we can simply write at tree level

$$\begin{aligned} P_{gg}(\mathbf{k}_1) &= E_1^2(\mathbf{k}_1)P(k_1), \\ B_{ggg}(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3) &= 2E_1(\mathbf{k}_1)E_1(\mathbf{k}_2)E_2(\mathbf{k}_1,\mathbf{k}_2)P(k_1)P(k_2) + 2 \text{ cyc.}, \end{aligned} \tag{49}$$

where $P(k)$ is the matter power spectrum for the Gaussian source field φ_G, while the kernels E_i are defined as

$$E_1(\mathbf{k}_1) = b_{10} + \frac{b_{01}}{\alpha(k_1)} \tag{50}$$

with the scale dependent bias term $b_{01}/\alpha(k_1) \propto f_{\rm NL}/k_1^2$, and

$$\begin{aligned} E_2(\mathbf{k}_1,\mathbf{k}_2) = b_{10}&\left[F_2(\mathbf{k}_1,\mathbf{k}_2) + f_{\rm NL}\frac{\alpha(|\mathbf{k}_1+\mathbf{k}_2|)}{\alpha(k_1)\alpha(k_2)}\right] \\ &+ \left[b_{20} - \frac{2}{7}b_{10}^L S_2(\mathbf{k}_1,\mathbf{k}_2)\right] + \frac{b_{11}}{2}\left[\frac{1}{\alpha(k_1)} + \frac{1}{\alpha(k_2)}\right] \\ &+ \frac{b_{02}}{\alpha(k_1)\alpha(k_2)} - b_{01}\left[\frac{N_2(\mathbf{k}_1,\mathbf{k}_2)}{\alpha(k_2)} + \frac{N_2(\mathbf{k}_2,\mathbf{k}_a)}{\alpha(k_1)}\right]. \end{aligned} \tag{51}$$

Note that, general relativistic effects (including also redshift-space distortions, lensing, etc.) have to be taken into account both in the galaxy power-spectrum and bispectrum, as well as in the dark matter evolution.

Table IV. – *Forecasts for $\sigma_{f_{\rm NL}}$, the accuracy of the determination of local $f_{\rm NL}$, from the bispectrum of BOSS, eBOSS, DESI and Euclid. From* [142].

Sample	Power Spectrum		Bispectrum	
	$\sigma_{f_{\rm NL}}$ bias float	$\sigma_{f_{\rm NL}}$ bias fixed	$\sigma_{f_{\rm NL}}$ bias float	$\sigma_{f_{\rm NL}}$ bias fixed
BOSS	21.30	13.28	$1.04^{(0.65)}_{(2.47)}$	$0.57^{(0.35)}_{(1.48)}$
eBOSS	14.21	11.12	$1.18^{(0.82)}_{(2.02)}$	$0.70^{(0.48)}_{(1.29)}$
Euclid	6.00	4.71	$0.45^{(0.18)}_{(0.71)}$	$0.32^{(0.12)}_{(0.35)}$
DESI	5.43	4.37	$0.31^{(0.17)}_{(0.48)}$	$0.21^{(0.12)}_{(0.37)}$
BOSS + Euclid	5.64	4.44	$0.39^{(0.17)}_{(0.59)}$	$0.28^{(0.11)}_{(0.34)}$

The complete expression (very involved) for the galaxy bispectrum was written down for the first time recently, and can be found in [146] to be soon compared with observations, see also [147-149].

We conclude this section describing, following [142], the Fisher matrix forecasts on $\sigma_{f_{\rm NL}}$ (the accuracy of the determination of local non-linear parameter $f_{\rm NL}$) from measurements of the galaxy bispectrum, as well as the constraints on PNG from the galaxy power spectrum and bispectrum in future radio continuum and optical surveys [150]. See also [151] for the constraining power on primordial non-Gaussianity of the Dark Energy Survey (DES) as well as Euclid and WFIRST.

Particularly, the tree-level bispectra with local non-Gaussian initial conditions are shown in table IV in redshift space, where the covariance between different triangles has been neglected.

While many issues are still present (such as the need for full covariance, a better understanding of accurate bias model, the inclusion of general relativistic effects, or the proper implementation of the estimators), still moving from the 2D *Planck* data to the 3D maps of the forthcoming surveys represent a great potential as

– the bispectrum could do better than the power-spectrum,

– we might increase the accuracy to achieve $f_{\rm NL} \sim 1$.

In particular, the LSS bispectrum allows in principle tight constraints also on non-local shapes (*e.g.*, equilateral). However, even if naive mode counting suggest that $\sigma_{f_{\rm NL}} \sim 1$ for the equilateral shape might be achievable by pushing k_{max} high enough, modeling the gravitational bispectrum in the non-linear regime with high accuracy is very challenging, as the equilateral shape is more correlated than local to the non-linear gravitational bispectrum.

TABLE V. – *Summary of* $1\,\sigma$ *limits for the three PNG types considered, from radio continuum and optical surveys derived from combining the power spectrum and bispectrum and accounting for RSD, the trispectrum term and theoretical errors. From* [150].

	Planck	$1\,\mu$Jy	$10\,\mu$Jy	Spectroscopic	Photometric
Local	5.0	0.2	0.6	1.3	0.3
Equilateral	43	244	274	57	184
Orthogonal	21	18	29	18	38

Finally, possible future constraints on the amplitude $f_{\rm NL}$ for the local, equilateral and orthogonal shapes, through galaxy power spectrum and bispectrum measurements on large scales based on radio continuum (with $10\,\mu$Jy and $1\,\mu$Jy flux limits) and optical (spectroscopic and photometric) surveys are presented in table V, see [150] for details. Remarkably, for the local shape, LSS measurements can provide significant improvements over current *Planck* constraints on PNG.

5. – Controversial issues on non-Gaussianity

In this section, we present some aspects regarding non-Gaussianity which we think are still controversial or require further clarifications.

5˙1. *Single-field consistency relation.* – The first issue concerns the single-field consistency relation. Actually, as mentioned before, the common lore is that the "consistency relation", implying $f_{\rm NL} = -5\,(n_s - 1)/12$, can be gauged away by a non-linear rescaling of coordinates, up to sub-leading terms. As a consequence, the only residual term is proportional to ϵ thus of the same order of the amplitude of tensor modes.

More precisely, the bispectrum for single-field inflation can be represented as [6,10,11]

$$B_\zeta(k_1,k_2,k_3) \propto \frac{(\Delta^2_\zeta)^2}{(k_1k_2k_3)^2}\left[(1-n_s)\mathcal{S}_{loc.}(k_1,k_2,k_3) + \frac{5}{3}\epsilon\mathcal{S}_{equil.}(k_1,k_2,k_3)\right], \tag{52}$$

$$n_s = 1 - \eta - 2\epsilon, \quad \text{with} \quad \epsilon \equiv \frac{\dot H}{H^2}\,, \quad \eta \equiv \frac{\dot\epsilon}{H\epsilon}\,.$$

The observability of the so-called "Maldacena consistency relation", related to the above bispectrum for single field inflation, in CMB and LSS data, has led to a long-standing controversy.

Recently, various groups have argued that the $(1 - n_s)$ term is totally unobservable (for single-clock inflation), as, in the strictly squeezed limit (one of the wave-numbers, say k_i, going to 0, $k_i \to 0$), this term can be gauged away by a suitable coordinate tranformation.

However, in [152] it has been argued that the term survives up to a "renormalization" which further reduces it by a factor of ~ 0.1 if one applies Conformal Fermi Coordinates (CFC) to get rid of such a "gauge mode".

– Is this (CFC approach) the only way to deal with this term?

– Can we aim at an exact description, which is not affected by "spurious PNG"?

5˙2. *Non-Gaussian $f_{\rm NL}$-like terms generated by non-linear general relativistic evolution.* – An other important issue regards the role of the non-linear evolution of the matter perturbations in general relativity. Actually, second order DM dynamics in GR leads to (post-Newtonian) $\delta\zeta$-like terms which mimic local primordial non-Gaussianity [153]. For instance, these terms have been included in the halo bias in [125].

For a recent estimate of the effective non-Gaussianity due to general-relativistic light-cone effects mimicking a PNG signal, see [154]([3]).

Remarkably, these GR terms can be recovered by a short-long mode splitting (λ_S and λ_L, respectively) leading to a resummed non-linear contribution $\delta e^{-2\zeta}$ [157]. This comes from the modulation of sub-horizon scales due to modes entering the horizon at any given time. In the comoving gauge (suitable for calculation of halo bias) this would correspond to an $f_{\rm NL} = -5/3$ in the pure squeezed limit.

Then, we may ask the following question

– is such relativistic NG signature detectable via some cosmological observables?

Consider a patch of the Universe, where the comoving spatial element is given by

$$\mathrm{d}s^2_{(3)} = e^{2\zeta}\delta_{ij}\mathrm{d}x^i\mathrm{d}x^j. \tag{53}$$

There is a global background which must be defined with respect to some scale λ_0, at least as large as all the other scales of interest, *i.e.*, at least as large as our presently observable Universe.

Then, there is an other important scale, the separate Universe patches, λ_P, distinguished from λ_0 and we assume $\lambda_0 \gg \lambda_P \gg \lambda_S$. This is large enough for each patch to be treated as locally homogeneous and isotropic, but patches must be stitched together to describe the long-wavelength perturbations on a scale $\lambda_L \gg \lambda_P$, see fig. 14. Thus,

$$\lambda_0 > \lambda_L \gg \lambda_P \gg \lambda_S. \tag{54}$$

The local observer in a separate Universe patch cannot observe the effect of ζ_L, which is locally homogeneous on the patch scale λ_P. However, local coordinates can be defined only locally and the long mode curvature perturbation is observable through a mapping from local to global coordinates.

In the halo bias case the effect is unobservable. Indeed, as pointed out in [159-161], a local physical redefinition of the mass gauges away such a NG effect (in the pure squeezed limit), similarly to Maldacena's single-field NG contribution. This is true provided the halo bias definition is strictly local. We may ask the following questions:

([3]) Note that also dark energy could in principle introduce degeneracies with PNG, see, *e.g.*, [155, 156].

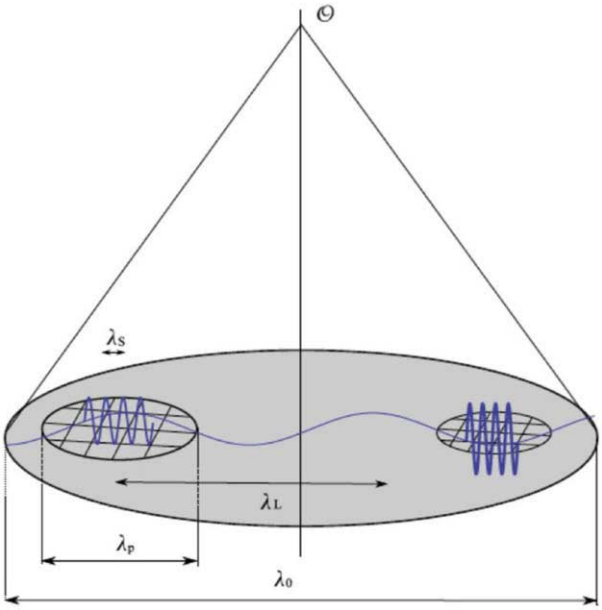


Fig. 14. – Schematic of the various scales in eq. (54). From [158].

– are there significant exceptions?

– are all non-linear GR effects fully accounted for by "projection effects"?

In general, this dynamically generated GR non-linearity is physical and cannot be gauged away by any local mass-rescaling, provided it involves scales larger than the patch required to define halo bias, but smaller than the separation between halos (and the distance of the halo to the observer).

Hence one would expect it to be in principle detectable in the matter bispectrum. Similarly, the observed galaxy bispectrum obtained via a full GR calculation must include all second-order GR non-linearities on such scales (only as projection effects?).

In conclusion, the separate Universe approach is very useful for many applications, but the effect of the external world cannot be always described by linear theory, thus the usual identification of large scales with the linear theory is only qualitative and can become misleading in some cases.

For example, on the one hand, perturbations of order $N \gg 1$ give the leading contribution to N-th order moments, such as $\langle \delta^N \rangle_c$. On the other hand, we know from non-linear Newtonian dynamics that $\langle \delta^N \rangle_c \sim \langle \delta^2 \rangle^{N-1}$ on all scales (for scale-free spectra). Well inside a given separate Universe, the assumption that the only non-linearity is described by Newtonian physics might be too restrictive, as the relevance of non-linear GR effects in sub-patch dynamics depends upon the specific problem.

It would be interesting to see the effects of using, for example, the silent Universe

description ([162]) to account for deviations of the patch from purely spherical behavior. Recall that over-dense patches evolve towards oblate ellipsoids and even under-dense ones can collapse to oblate ellipsoids, owing to tidal effects of surrounding matter. Recent approaches using the local tide approximation [163] go in this direction.

6. – Concluding remarks

The concordance ΛCDM cosmological model describes the evolution of the Universe from 380000 years after the Big Bang to the present time with astonishingly accuracy. However, the mechanism that generated the primordial fluctuations representing the seeds of the structures we observe in the CMB and in the LSS is unknown. Remarkably, inflation provides a causal mechanism for the generation of cosmological perturbations, whose detailed predictions are fully supported by CMB and LSS data.

Clearly, the direct detection of PGW and PNG with the specific features predicted by inflation would provide strong independent support to this framework.

In the previous sections, we have summarized the theoretical motivations for PNG and the present observational status. As stated before, *Planck* has provided stringent limits on $f_{\rm NL}$ which will be improved with polarization maps and using the full data in the upcoming "*Planck* legacy" paper.

In conclusion, among the short term goals, we need to look for more non-Gaussian shapes, such as scale-dependent $f_{\rm NL}$, make use of the bispectrum in 3D data and improve the constraints on $g_{\rm NL}$.

The final goal is to reconstruct the inflationary action by improving the sensitivity on NG parameters, searching for $f_{\rm NL} \sim 1$ for all shapes, taking into account non-linear general relativistic effects and second-order radiation transfer function contributions.

* * *

SM would like to thank the directors of the school, Eugenio Coccia, Joe Silk and Nicola Vittorio, and the Italian Physical Society (SIF) for having invited him to lecture at the International School of Physics Enrico Fermi (from 26th June to 19th July 2017) in the beautiful Villa Monastero, Varenna, Lake Como (Italy). SM acknowledges partial financial support by ASI Grant No. 2016-24-H.0.

REFERENCES

[1] Aghanim N. *et al.*, [arXiv:1807.06209] (2018).
[2] Bardeen J. M., Bond J. R., Kaiser N. and Szalay A. S., *Astrophys. J.*, **304** (1986) 15.
[3] Groth E. J. and Peebles P. J. E., *Astrophys. J.*, **217** (1977) 385.
[4] Salopek D. S. and Bond J. R., *Phys. Rev. D*, **42** (1990) 3936.
[5] Salopek D. S. and Bond J. R., *Phys. Rev. D*, **43** (1991) 1005.
[6] Gangui A., Lucchin F., Matarrese S. and Mollerach S., *Astrophys. J.*, **430** (1994) 447.

[7] Verde L., Wang L.-M., Heavens A. and Kamionkowski M., *Mon. Not. R. Astron. Soc.*, **313** (2000) L141.
[8] Komatsu E. and Spergel D. N., *Phys. Rev. D*, **63** (2001) 063002.
[9] Moscardini L., Lucchin F., Messina A. and Matarrese S., *Ann. N.Y. Acad. Sci.*, **647** (1991) 775.
[10] Acquaviva V., Bartolo N., Matarrese S. and Riotto A., *Nucl. Phys. B*, **667** (2003) 119.
[11] Maldacena J. M., *JHEP*, **05** (2003) 013.
[12] Feldman H. A., Frieman J. A., Fry J. N. and Scoccimarro R., *Phys. Rev. Lett.*, **86** (2001) 1434.
[13] Scoccimarro R., Feldman H. A., Fry J. N. and Frieman J. A., *Astrophys. J.*, **546** (2001) 652.
[14] Verde L. *et al.*, *Mon. Not. R. Astron. Soc.*, **335** (2002) 432.
[15] Marin F. A. *et al.*, *Mon. Not. R. Astron. Soc.*, **432** (2013) 2654.
[16] Gil-Marn H., Norea J., Verde L., Percival W. J., Wagner C., Manera M. and Schneider D. P., *Mon. Not. R. Astron. Soc.*, **451** (2015) 539.
[17] Gil-Marin H., Percival W. J., Verde L., Brownstein J. R., Chuang C.-H., Kitaura F.-S., Rodrguez-Torres S. A. and Olmstead M. D., *Mon. Not. R. Astron. Soc.*, **465** (2017) 1757.
[18] Padmanabhan N. *et al.*, *Mon. Not. R. Astron. Soc.*, **378** (2007) 852.
[19] Slosar A., Hirata C., Seljak U., Ho S. and Padmanabhan N., *JCAP*, **0808** (2008) 031.
[20] Xia J.-Q., Viel M., Baccigalupi C., De Zotti G., Matarrese S. and Verde L., *Astrophys. J.*, **717** (2010) L17.
[21] Xia J.-Q., Baccigalupi C., Matarrese S., Verde L. and Viel M., *JCAP*, **1108** (2011) 033.
[22] Nikoloudakis N., Shanks T. and Sawangwit U., *Mon. Not. R. Astron. Soc.*, **429** (2013) 2032.
[23] Agarwal N., Ho S. and Shandera S., *JCAP*, **1402** (2014) 038.
[24] Karagiannis D., Shanks T. and Ross N. P., *Mon. Not. R. Astron. Soc.*, **441** (2014) 486.
[25] Leistedt B., Peiris H. V. and Roth N., *Phys. Rev. Lett.*, **113** (2014) 221301.
[26] Ade P. A. R. *et al.*, *Astron. Astrophys.*, **571** (2014) A24.
[27] Ade P. A. R. *et al.*, *Astron. Astrophys.*, **594** (2016) A17.
[28] Akrami Y. *et al.*, [arXiv:1807.06211] (2018).
[29] Jimenez R., Raccanelli A., Verde L. and Matarrese S., *JCAP*, **1804** (2018) 002.
[30] Ade P. A. R. *et al.*, *Phys. Rev. Lett.*, **121** (2018) 221301.
[31] Arkani-Hamed N. and Maldacena J., [arXiv:1503.08043] (2015).
[32] Silverstein E., [arXiv:1706.02790] (2017).
[33] Falk T., Rangarajan R. and Srednicki M., *Astrophys. J.*, **403** (1993) L1.
[34] Bernardeau F., Kofman L. and Uzan J.-P., *Phys. Rev. D*, **70** (2004) 083004.
[35] Weinberg S., *Phys. Rev. D*, **72** (2005) 043514.
[36] Babich D., Creminelli P. and Zaldarriaga M., *JCAP*, **0408** (2004) 009.
[37] Bartolo N., Fasiello M., Matarrese S. and Riotto A., *JCAP*, **1008** (2010) 008.
[38] Creminelli P., Nicolis A., Senatore L., Tegmark M. and Zaldarriaga M., *JCAP*, **0605** (2006) 004.
[39] Senatore L., Smith K. M. and Zaldarriaga M., *JCAP*, **1001** (2010) 028.
[40] Meerburg P. D., van der Schaar J. P. and Corasaniti P. S., *JCAP*, **0905** (2009) 018.
[41] Chen X., Huang M.-x., Kachru S. and Shiu G., *JCAP*, **0701** (2007) 002.

[42] Holman R. and Tolley A. J., *JCAP*, **0805** (2008) 001.
[43] Creminelli P., D'Amico G., Musso M., Norena J. and Trincherini E., *JCAP*, **1102** (2011) 006.
[44] Scherrer R. J. and Schaefer R. K., *Astrophys. J.*, **446** (1995) 44.
[45] Liguori M., Yadav A., Hansen F. K., Komatsu E., Matarrese S. and Wandelt B., *Phys. Rev. D*, **76** (2007) 105016, (Erratum: *Phys. Rev. D* **77** (2008) 029902).
[46] Komatsu E., Spergel D. N. and Wandelt B. D., *Astrophys. J.*, **634** (2005) 14.
[47] Munshi D. and Heavens A., *Mon. Not. R. Astron. Soc.*, **401** (2010) 2406.
[48] Bucher M., Van Tent B. and Carvalho C. S., *Mon. Not. R. Astron. Soc.*, **407** (2010) 2193.
[49] Fergusson J. R., Liguori M. and Shellard E. P. S., *Phys. Rev. D*, **82** (2010) 023502.
[50] Verde L., Jimenez R., Alvarez-Gaume L., Heavens A. F. and Matarrese S., *JCAP*, **1306** (2013) 023.
[51] Cheung C., Creminelli P., Fitzpatrick A. L., Kaplan J. and Senatore L., *JHEP*, **03** (2008) 014.
[52] Khatri R. and Sunyaev R. A., *JCAP*, **1306** (2013) 026.
[53] Pajer E. and Zaldarriaga M., *JCAP*, **1302** (2013) 036.
[54] Ganc J. and Komatsu E., *Phys. Rev. D*, **86** (2012) 023518.
[55] Bartolo N., Liguori M. and Shiraishi M., *JCAP*, **1603** (2016) 029.
[56] Pajer E. and Zaldarriaga M., *Phys. Rev. Lett.*, **109** (2012) 021302.
[57] Emami R., Dimastrogiovanni E., Chluba J. and Kamionkowski M., *Phys. Rev. D*, **91** (2015) 123531.
[58] Cabass G., *CMB Anisotropies and Spectral Distortions: Constraining Inflation at Small Scales*, PhD Thesis (2017).
[59] Bardeen J. M., *Phys. Rev. D*, **22** (1980) 1882.
[60] Grossi M., Dolag K., Branchini E., Matarrese S. and Moscardini L., *Mon. Not. R. Astron. Soc.*, **382** (2007) 1261.
[61] Hikage C., Coles P., Grossi M., Moscardini L., Dolag K., Branchini E. and Matarrese S., *Mon. Not. R. Astron. Soc.*, **385** (2008) 1613.
[62] Desjacques V., Seljak U. and Iliev I., *Mon. Not. R. Astron. Soc.*, **396** (2009) 85.
[63] Giannantonio T. and Porciani C., *Phys. Rev. D*, **81** (2010) 063530.
[64] Grossi M., Verde L., Carbone C., Dolag K., Branchini E., Iannuzzi F., Matarrese S. and Moscardini L., *Mon. Not. R. Astron. Soc.*, **398** (2009) 321.
[65] Sefusatti E., Crocce M. and Desjacques V., *Mon. Not. R. Astron. Soc.*, **406** (2010) 1014.
[66] Wagner C., Verde L. and Boubekeur L., *JCAP*, **1010** (2010) 022.
[67] Wagner C. and Verde L., *JCAP*, **1203** (2012) 002.
[68] Matarrese S. and Pietroni M., *JCAP*, **0706** (2007) 026.
[69] Pietroni M., *JCAP*, **0810** (2008) 036.
[70] Crocce M. and Scoccimarro R., *Phys. Rev. D*, **73** (2006) 063519.
[71] Bernardeau F., Crocce M. and Scoccimarro R., *Phys. Rev. D*, **78** (2008) 103521.
[72] Montesano F., Sanchez A. G. and Phleps S., *Mon. Not. R. Astron. Soc.*, **408** (2010) 2397.
[73] McDonald P., *Phys. Rev. D*, **75** (2007) 043514.
[74] Taruya A. and Hiramatsu T., *Astrophys. J.*, **674** (2008) 617.
[75] Buchert T., *Mon. Not. R. Astron. Soc.*, **254** (1992) 729.
[76] Matsubara T., *Phys. Rev. D*, **78** (2008) 083519, (Erratum: *Phys. Rev. D* **78** (2008) 109901).
[77] Zheligovsky V. and Frisch U., *J. Fluid Mech.*, **749** (2014) 404.
[78] Blas D., Garny M., Ivanov M. M. and Sibiryakov S., *JCAP*, **1607** (2016) 052.

[79] Baumann D., Nicolis A., Senatore L. and Zaldarriaga M., *JCAP*, **1207** (2012) 051.
[80] Carrasco J. J. M., Hertzberg M. P. and Senatore L., *JHEP*, **09** (2012) 082.
[81] Assassi V., Baumann D., Pajer E., Welling Y. and van der Woude D., *JCAP*, **1511** (2015) 024.
[82] Hertzberg M. P., *Phys. Rev. D*, **89** (2014) 043521.
[83] Carrasco J. J. M., Foreman S., Green D. and Senatore L., *JCAP*, **1407** (2014) 057.
[84] Porto R. A., Senatore L. and Zaldarriaga M., *JCAP*, **1405** (2014) 022.
[85] Senatore L. and Zaldarriaga M., [arXiv:1409.1225] (2014).
[86] Baldauf T., Mercolli L., Mirbabayi M. and Pajer E., *JCAP*, **1505** (2015) 007.
[87] Angulo R., Fasiello M., Senatore L. and Vlah Z., *JCAP*, **1509** (2015) 029.
[88] Abolhasani A. A., Mirbabayi M. and Pajer E., *JCAP*, **1605** (2016) 063.
[89] Pillepich A., Porciani C. and Hahn O., *Mon. Not. R. Astron. Soc.*, **402** (2010) 191.
[90] Bartolo N., Almeida J. P. B., Matarrese S., Pietroni M. and Riotto A., *JCAP*, **1003** (2010) 011.
[91] Verde L., Jimenez R., Kamionkowski M. and Matarrese S., *Mon. Not. R. Astron. Soc.*, **325** (2001) 412.
[92] Matarrese S., Verde L. and Jimenez R., *Astrophys. J.*, **541** (2000) 10.
[93] LoVerde M., Miller A., Shandera S. and Verde L., *JCAP*, **0804** (2008) 014.
[94] Maggiore M. and Riotto A., *Astrophys. J.*, **711** (2010) 907.
[95] Maggiore M. and Riotto A., *Astrophys. J.*, **717** (2010) 515.
[96] Maggiore M. and Riotto A., *Astrophys. J.*, **717** (2010) 526.
[97] Maggiore M. and Riotto A., *Mon. Not. R. Astron. Soc.*, **405** (2010) 1244.
[98] Lam T. Y. and Sheth R. K., *Mon. Not. R. Astron. Soc.*, **395** (2009) 1743.
[99] Lam T. Y. and Sheth R. K., *Mon. Not. R. Astron. Soc.*, **398** (2009) 2143.
[100] D'Amico G., Musso M., Norena J. and Paranjape A., *JCAP*, **1102** (2011) 001.
[101] LoVerde M. and Smith K. M., *JCAP*, **1108** (2011) 003.
[102] Paranjape A., Lam T. Y. and Sheth R. K., *Mon. Not. R. Astron. Soc.*, **420** (2012) 1429.
[103] Paranjape A. and Sheth R. K., *Mon. Not. R. Astron. Soc.*, **419** (2012) 132.
[104] Sefusatti E., Crocce M. and Desjacques V., *Mon. Not. R. Astron. Soc.*, **425** (2012) 2903.
[105] Lazanu A., Giannantonio T., Schmittfull M. and Shellard E. P. S., *Phys. Rev. D*, **95** (2017) 083511.
[106] Lazanu A., Giannantonio T., Schmittfull M. and Shellard E. P. S., *Phys. Rev. D*, **93** (2016) 083517.
[107] Silk J. and White S. D., *Astrophys. J.*, **223** (1978) L59.
[108] Kaiser N., *Astrophys. J.*, **284** (1984) L9.
[109] Desjacques V., Jeong D. and Schmidt F., *Phys. Rept.*, **733** (2018) 1.
[110] Assassi V., Baumann D., Green D. and Zaldarriaga M., *JCAP*, **1408** (2014) 056.
[111] Senatore L., *JCAP*, **1511** (2015) 007.
[112] Mirbabayi M., Schmidt F. and Zaldarriaga M., *JCAP*, **1507** (2015) 030.
[113] Coles P., *Mon. Not. R. Astron. Soc.*, **262** (1993) 1065.
[114] Fry J. N. and Gaztanaga E., *Astrophys. J.*, **413** (1993) 447.
[115] Fry J. N., *Astrophys. J.*, **461** (1996) L65.
[116] Catelan P., Lucchin F., Matarrese S. and Porciani C., *Mon. Not. R. Astron. Soc.*, **297** (1998) 692.
[117] Catelan P., Porciani C. and Kamionkowski M., *Mon. Not. R. Astron. Soc.*, **318** (2000) 39.

[118] McDonald P. and Roy A., *JCAP*, **0908** (2009) 020.
[119] Elia A., Kulkarni S., Porciani C., Pietroni M. and Matarrese S., *Mon. Not. R. Astron. Soc.*, **416** (2011) 1703.
[120] Chan K. C., Scoccimarro R. and Sheth R. K., *Phys. Rev. D*, **85** (2012) 083509.
[121] Baldauf T., Seljak U., Desjacques V. and McDonald P., *Phys. Rev. D*, **86** (2012) 083540.
[122] Dalal N., Dore O., Huterer D. and Shirokov A., *Phys. Rev. D*, **77** (2008) 123514.
[123] Matarrese S. and Verde L., *Astrophys. J.*, **677** (2008) L77.
[124] Afshordi N. and Tolley A. J., *Phys. Rev. D*, **78** (2008) 123507.
[125] Verde L. and Matarrese S., *Astrophys. J.*, **706** (2009) L91.
[126] Schmidt F. and Kamionkowski M., *Phys. Rev. D*, **82** (2010) 103002.
[127] Desjacques V. and Seljak U., *Adv. Astron.*, **2010** (2010) 908640.
[128] Desjacques V., Jeong D. and Schmidt F., *Phys. Rev. D*, **84** (2011) 063512.
[129] Schmidt F., Jeong D. and Desjacques V., *Phys. Rev. D*, **88** (2013) 023515.
[130] Grinstein B. and Wise M. B., *Astrophys. J.*, **310** (1986) 19.
[131] Matarrese S., Lucchin F. and Bonometto S. A., *Astrophys. J.*, **310** (1986) L21.
[132] Vittorio N., Juszkiewicz R. and Davis M., *Nature*, **323** (1986) 132.
[133] Mo H. J. and White S. D. M., *Mon. Not. R. Astron. Soc.*, **282** (1996) 347.
[134] Taruya A., Koyama K. and Matsubara T., *Phys. Rev. D*, **78** (2008) 123534.
[135] Jeong D. and Komatsu E., *Astrophys. J.*, **703** (2009) 1230.
[136] Baldauf T., Seljak U. and Senatore L., *JCAP*, **1104** (2011) 006.
[137] Scoccimarro R., Hui L., Manera M. and Chan K. C., *Phys. Rev. D*, **85** (2012) 083002.
[138] Schmidt F. and Hui L., *Phys. Rev. Lett.*, **110** (2013) 011301, (Erratum: *Phys. Rev. Lett.* **110** (2013) 059902).
[139] Assassi V., Baumann D. and Schmidt F., *JCAP*, **1512** (2015) 043.
[140] Scoccimarro R., Sefusatti E. and Zaldarriaga M., *Phys. Rev. D*, **69** (2004) 103513.
[141] Sefusatti E. and Komatsu E., *Phys. Rev. D*, **76** (2007) 083004.
[142] Tellarini M., Ross A. J., Tasinato G. and Wands D., *JCAP*, **1606** (2016) 014.
[143] Bernardeau F., Colombi S., Gaztanaga E. and Scoccimarro R., *Phys. Rept.*, **367** (2002) 1.
[144] Peebles P., *The Large-scale Structure of the Universe*, in *Princeton Ser. Phys.* (Princeton University Press) 1980.
[145] Bouchet F. R., Juszkiewicz R., Colombi S. and Pellat R., *Astrophys. J.*, **394** (1992) L5.
[146] Bertacca D., Raccanelli A., Bartolo N., Liguori M., Matarrese S. and Verde L., *Phys. Rev. D*, **97** (2018) 023531.
[147] Di Dio E., Durrer R., Marozzi G. and Montanari F., *JCAP*, **1601** (2016) 016.
[148] Raccanelli A., Montanari F., Bertacca D., Dore O. and Durrer R., *JCAP*, **1605** (2016) 009.
[149] Umeh O., Jolicoeur S., Maartens R. and Clarkson C., *JCAP*, **1703** (2017) 034.
[150] Karagiannis D., Lazanu A., Liguori M., Raccanelli A., Bartolo N. and Verde L., *Mon. Not. R. Astron. Soc.*, **478** (2018) 1341.
[151] Giannantonio T., Porciani C., Carron J., Amara A. and Pillepich A., *Mon. Not. R. Astron. Soc.*, **422** (2012) 2854.
[152] Cabass G., Pajer E. and Schmidt F., *JCAP*, **1701** (2017) 003.
[153] Bartolo N., Matarrese S. and Riotto A., *JCAP*, **0510** (2005) 010.
[154] Koyama K., Umeh O., Maartens R. and Bertacca D., *JCAP*, **1807** (2018) 050.
[155] Sefusatti E., Vale C., Kadota K. and Frieman J., *Astrophys. J.*, **658** (2007) 669.

[156] Hashim M., Giocoli C., Baldi M., Bertacca D. and Maartens R., *Mon. Not. R. Astron. Soc.*, **481** (2018) 2933.
[157] Bruni M., Hidalgo J. C. and Wands D., *Astrophys. J.*, **794** (2014) L11.
[158] Bartolo N., Bertacca D., Bruni M., Koyama K., Maartens R., Matarrese S., Sasaki M., Verde L. and Wands D., *Phys. Dark Univ.*, **13** (2016) 30.
[159] Dai L., Pajer E. and Schmidt F., *JCAP*, **1511** (2015) 043.
[160] Dai L., Pajer E. and Schmidt F., *JCAP*, **1510** (2015) 059.
[161] de Putter R., Dore' O. and Green D., *JCAP*, **1510** (2015) 024.
[162] Bruni M., Matarrese S. and Pantano O., *Astrophys. J.*, **445** (1995) 958.
[163] Ip H. Y. and Schmidt F., *JCAP*, **1702** (2017) 025.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200010

CMB polarization theory

Wayne Hu

Kavli Institute for Cosmological Physics, Department of Astronomy and Astrophysics
Enrico Fermi Institute, University of Chicago - Chicago, IL, USA

Summary. — I review the sources of CMB polarization in the acoustic regime, reionization and gravitational waves from inflation and their distortion by gravitational waves using the ΛCDM model as an illustration.

1. – Introduction

In these Proceedings of the Varenna School, I review the theory of CMB polarization anisotropy as it applies to the standard inflationary ΛCDM model of cosmology (ΛCDM for short). Since the discussion of the current status of CMB measurements and the observational successes and tensions of the ΛCDM paradigm have already become out of date between the time the lectures were given in Summer 2017 and when these Proceedings were written in Summer 2018 (and will only get further out of date by the time they are published), I have retained only the basic theory of polarization anisotropy under the ΛCDM paradigm here from the more widely ranging topics in the lectures themselves, which are available online.

This treatment is based in part on the more extensive reviews in [1,2] where references to the original literature can be found; citations here will only be to other works from which I have adapted figures.

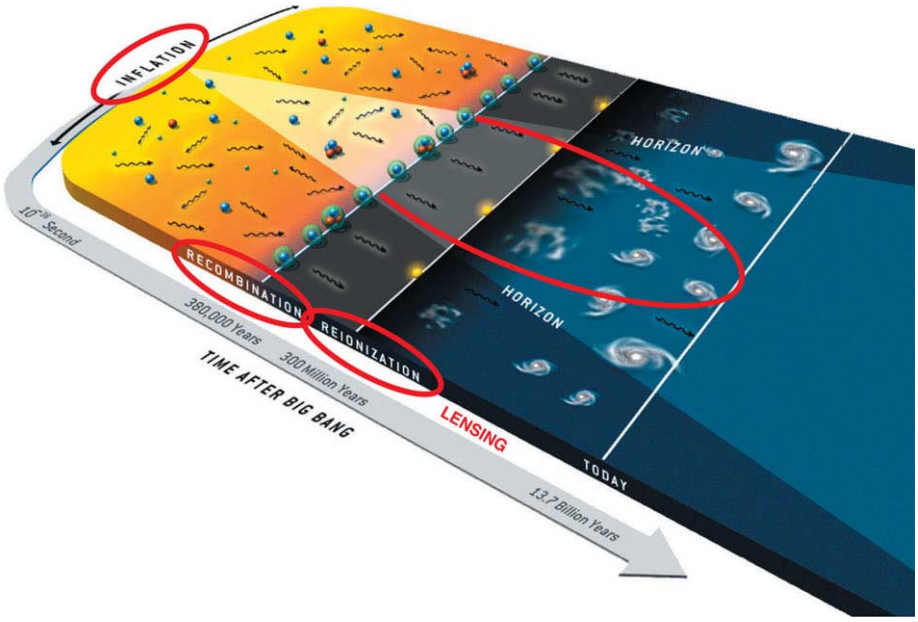


Fig. 1. – The trinity of polarization sources and their latter-day distortion: acoustic polarization (pressure wave E-modes at recombination), reionization (rescattered Sachs-Wolfe quadrupoles), inflation (gravitational wave B-modes), and lensing (distortion of acoustic E-modes to B-modes by deflection). The Varenna school was held in Villa Monastero, a former monastery; to balance the implicit Western bias, let me recount the Four Noble Truths of Polarization: the measure of polarization is suffering, the origin of suffering is desire (for the Nobel prize for inflation), the cessation of suffering is the nirvana of theory, the path to cessation of suffering is deflection. Adapted from [3].

2. – Sources of CMB polarization

Fundamentally the CMB is polarized since we receive its photons after having last scattered off free electrons via Thomson scattering. This last scattering event mostly occurs at recombination, the epoch of which free electrons recombine with protons to form neutral hydrogen, and for a few percent of photons at the epoch of reionization (see fig. 1). The CMB anisotropy is polarized at only a small level since polarization by Thomson scattering requires, but also destroys, anisotropic radiation. Scattering isotropizes radiation and eliminates quadrupole moments.

More precisely, the local quadrupole of the radiation intensity plays a special role in polarization formation as can be seen from fig. 2 and the Thomson differential cross section

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{3}{8\pi}|\hat{\mathbf{E}}' \cdot \hat{\mathbf{E}}|^2 \sigma_T, \tag{1}$$

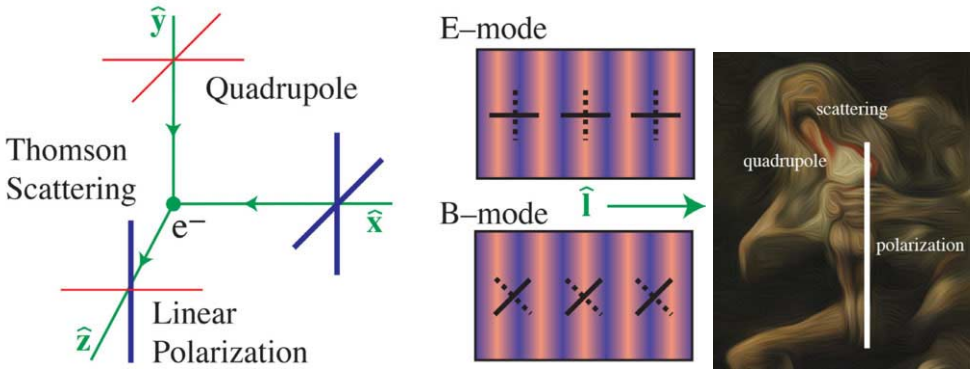


Fig. 2. – Thomson scattering: the originator and the devourer of polarization's source. Scattering of incoming radiation with a quadrupole anisotropy in intensity leads to linear polarization in the outgoing radiation that is aligned with the less intense (red) or "cold" axis of the quadrupole. Polarization can be decomposed into modes that are aligned with their wave vectors $\hat{\mathbf{l}}$ (direction of variation) or crossed, known as E- and B-modes respectively. The level of polarization is at best a small fraction of the temperature anisotropy since quadrupoles are necessary for scattering to create polarization but scattering also randomizes photon directions and destroys the quadrupoles themselves, like Saturn devouring his Son.

where $\sigma_T = 8\pi\alpha^2/3m_e$ is the Thomson cross section, $\hat{\mathbf{E}}'$ and $\hat{\mathbf{E}}$ denote the incoming and outgoing directions of the electric field or polarization vector. In the semi-classical approximation, incoming radiation shakes an electron in the direction of its electric field vector $\hat{\mathbf{E}}'$ causing it to radiate with an outgoing polarization parallel to that direction. However since the outgoing polarization $\hat{\mathbf{E}}$ must be orthogonal to the outgoing direction, incoming radiation that is polarized parallel to the outgoing direction cannot scatter, leaving only one polarization state. If the intensity were completely isotropic the missing polarization state is supplied by radiation incoming from the direction orthogonal to the original one. Only a quadrupole anisotropy in the radiation intensity generates a net linear polarization from Thomson scattering. In particular, the polarization is oriented along the cold axis of the quadrupole moment in the plane transverse to the outgoing direction.

We can formally describe the statistics of the CMB intensity and polarization anisotropy with Stokes parameters based on the intensity matrix. The CMB radiation is measured to be a blackbody and so it is convenient to characterize the intensity in terms of the brightness temperature and the intensity matrix using the Pauli matrices $\boldsymbol{\sigma}_i$

$$\begin{aligned}\mathbf{P} &= C\left\langle \mathbf{E}(\hat{\mathbf{n}})\,\mathbf{E}^{\dagger}(\hat{\mathbf{n}})\right\rangle \\ &= \Theta(\hat{\mathbf{n}})\mathbf{I} + Q(\hat{\mathbf{n}})\,\boldsymbol{\sigma}_3 + U(\hat{\mathbf{n}})\,\boldsymbol{\sigma}_1 + V(\hat{\mathbf{n}})\,\boldsymbol{\sigma}_2,\end{aligned} \tag{2}$$

where $\hat{\mathbf{n}}$ denotes direction on the sky (θ, ϕ) in spherical polar coordinates, and we have chosen the constant of proportionality so that the Stokes parameters (Θ, Q, U, V) are

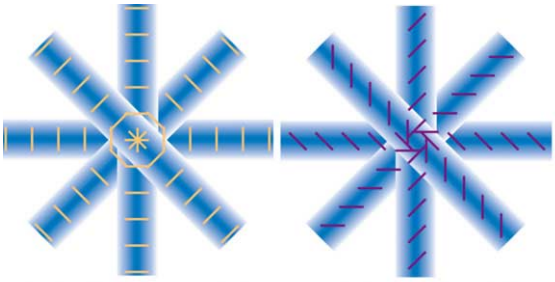


Fig. 3. – E- and B-modes differ geometrically in their symmetry. For E-modes, polarization only knows how to point in the direction to the variation (blue bands representing a crest of a plane wave decomposition) or orthogonal to it. When harmonic modes are superimposed, the result is that the polarization direction is radial or tangential to spots where the polarization is the strongest. For B-modes, the direction is crossed to the variation and lead to a handed swirl pattern around spots.

normalized to the brightness temperature, *e.g.* $\Theta(\hat{\mathbf{n}}) \equiv \Delta T(\hat{\mathbf{n}})/T$ averaged over polarization states. Note that the circular polarization V is absent cosmologically and under a counterclockwise rotation of the coordinate axes by ψ, $Q \pm iU \to e^{\mp 2i\psi}(Q \pm iU)$. This reflects the rotation of linear polarization that Q and U mutually represent.

Since the Stokes parameters depend on the coordinate choice one should choose a coordinate system that carries a geometric meaning. In general, linear polarization is a tensor field with two characteristic directions: the direction along which the polarized fraction varies in space and the direction that the polarization vector is pointing (see fig. 2). In a Fourier decomposition of the field, this former corresponds to the direction of the wave vector (see fig. 2). An E-mode is defined as Q polarization in the coordinates defined by the wave vector and a B-mode is defined as U polarization in that system. Importantly, E-modes possess even parity and when the Fourier modes are resummed display a pattern where the polarization vector is either radial or tangential around regions of high polarization, whereas B-modes display a swirl pattern that carries a handedness or odd parity (see fig. 3).

The final point is that Fourier harmonics must be replaced with spherical harmonics on the sky due to sky curvature. These harmonics are the eigenfunctions of the Laplace operator on the sphere for scalar and tensor fields. Specifically we decompose the temperature and polarization as

$$\Theta_{lm} = \int d\hat{\mathbf{n}}\, Y^*_{lm}(\hat{\mathbf{n}})\Theta(\hat{\mathbf{n}}), \tag{3}$$

$$E_{lm} \pm iB_{lm} = -\int d\hat{\mathbf{n}}_{\pm 2}\, Y^*_{lm}(\hat{\mathbf{n}})[Q(\hat{\mathbf{n}}) \pm iU(\hat{\mathbf{n}})],$$

in terms of the complete and orthogonal set of spin harmonic functions, ${}_sY_{lm}$, which are eigenfunctions of the Laplace operator on a rank s tensor. $Y_{lm} = {}_0Y_{lm}$ is the ordinary

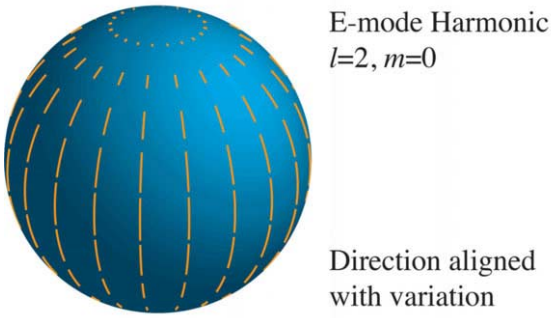


Fig. 4. – Polarization harmonics on the curved sky follow spin spherical harmonics rather than Fourier plane waves. Both are the eigenfunctions of the Laplace operator on a tensor in their respective curved and flat spaces. The fundamental pattern for E-modes in the CMB is based on its ($l = 2$, $m = 0$) harmonic. Higher harmonics result when this pattern is modulated in amplitude by the variation across the sky induced by the wavemode of the perturbation (see fig. 10).

spherical harmonic. An example of an E-mode with $l = 2$, $m = 0$ is shown in fig. 4. Note that the E-mode symmetry is manifest in that the polarization both varies and points in the same north-south direction. For small sections of sky, we can make the correspondence to the Fourier discussion above more explicit. The spin-harmonic expansion becomes a Fourier expansion with $Y_{lm} \to e^{i\mathbf{l}\cdot\hat{\mathbf{n}}}$ and ${}_{\pm2}Y_{lm} \to -e^{\pm 2i\phi_l} e^{i\mathbf{l}\cdot\hat{\mathbf{n}}}$, where ϕ_l is the azimuthal angle of the Fourier wave vector $\mathbf{l}$.

Conversely we can also make the all-sky connection to the quadrupole moment sources in fig. 2. In fig. 5 we show a cross section of the sky along a great circle passing from the equator through the north pole. As with the local scattering discussion, linear polarization is generated in the direction of the cold lobe of the quadrupole in projection on the plane of the sky, orthogonal to the outgoing direction of the scattered photon. Notice that as we sweep through the polar angle θ, the linear polarization from a projection of an $l = 2$, $m = 0$ quadrupole takes on the pattern of the $l = 2$, $m = 0$ E-mode shown in fig. 4. We can therefore understand polarization patterns through the quadrupole moments that generate then.

3. – Acoustic source

Before recombination, the abundance of free electrons keeps photons trapped in the baryons due to their relatively short comoving mean free path

$$\lambda_C \equiv \dot{\tau}^{-1} = \frac{1}{n_e \sigma_T a} \approx 2.5 \left(x_e \frac{\Omega_b h^2}{0.02} \right)^{-1} \left(\frac{a}{10^{-3}} \right)^2 \text{Mpc}. \tag{4}$$

Overdots here and below represent derivatives with respect to conformal time η, x_e is the ionization fraction and the scale factor a is scaled to the epoch of recombination. Electrons are coupled to baryons by Coulomb interactions and so on scales larger than

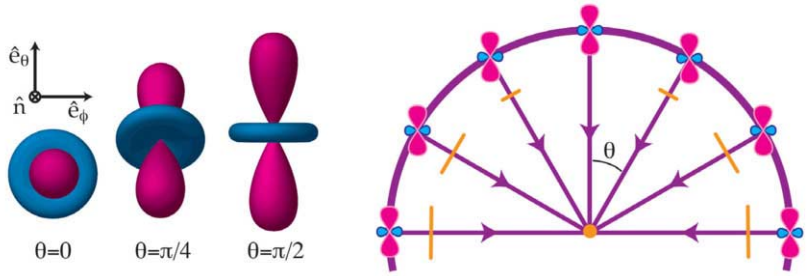


Fig. 5. – The fundamental $l = 2$, $m = 0$ polarization pattern is the projection on the sky of the cold lobe of an $l = 2$, $m = 0$ quadrupole moment, just like one would expect from fig. 2 for generation by scattering. Near the equator at polar angle $\theta = \pi/2$, the lobes are maximal in projection as is the linear polarization. At the poles $\theta = 0$ and the azimuthal symmetry of the quadrupole demands that the projection vanishes as does the linear polarization.

λ_C, the photon-baryon plasma can be considered a nearly perfect fluid. In particular, rapid scattering keeps the photons isotropic in the baryon rest frame and so sets its dipole moment equal to the baryon velocity. In this regime, the quadrupole moment of the temperature fluctuations is small and so polarization can only be generated at the end of this epoch, when the photons last scatter.

During this epoch, evolution of the anisotropy sources is governed by simple fluid dynamics for the density or monopole temperature fluctuation $\Theta(\mathbf{k}) \equiv \sqrt{4\pi}\,\Theta_{00}(\mathbf{k})$, and the dipole or bulk velocity $v_\gamma(\mathbf{k}) = -i\sqrt{4\pi/3}\,\Theta_{10}(\mathbf{k})$. The quadrupole or anisotropic stress $\pi_\gamma(\mathbf{k}) = -(12/5)\sqrt{4\pi/5}\,\Theta_{20}(\mathbf{k})$ is small but important for understanding the polarization. We will hereafter drop the argument $\mathbf{k}$ with the understanding that the Fourier representation is always assumed. For convenience we have chosen the coordinate system so that $\mathbf{z} \parallel \mathbf{k}$ so that the plane waves are azimuthally symmetric and stimulate only the $m = 0$ mode. Likewise we have suppressed the vector dependence of the bulk velocity by the same assumption $\mathbf{v}_\gamma = -iv_\gamma\hat{\mathbf{k}}$. The analogous quantities for the baryons are the density perturbation δ_b and bulk velocity v_b. For gravity, we choose a conformal Newtonian representation where the line element in a spatially flat universe is

$$\mathrm{d}s^2 = a^2[-(1+2\Psi)\mathrm{d}\eta^2 + (1+2\Phi)\mathrm{d}x^2], \tag{5}$$

the gravitational potential perturbations are defined by the Newtonian potential Ψ (time-time metric fluctuation) and the curvature fluctuation Φ (space-space metric fluctuation $\approx -\Psi$).

Covariant conservation of energy and momentum requires that the photons and baryons satisfy separate continuity equations

$$\dot{\Theta} = -\frac{k}{3}v_\gamma - \dot{\Phi}, \qquad \dot{\delta}_b = -kv_b - 3\dot{\Phi}, \tag{6}$$

and coupled Euler equations

$$\dot{v}_\gamma = k(\Theta + \Psi) - \frac{k}{6}\pi_\gamma - \dot{\tau}(v_\gamma - v_b), \qquad \dot{v}_b = -\frac{\dot{a}}{a}v_b + k\Psi + \dot{\tau}(v_\gamma - v_b)/R, \tag{7}$$

where $R = (p_b + \rho_b)/(p_\gamma + \rho_\gamma) \approx 3\rho_b/4\rho_\gamma$ is the photon-baryon momentum density ratio.

The continuity equations represent particle number conservation. For the baryons, $\rho_b \propto n_b$. For the photons, $T \propto n_\gamma^{1/3}$, which explains the 1/3 in the velocity divergence term. The $\dot{\Phi}$ terms come from the fact that Φ is a perturbation to the scale factor and so they are the perturbative analogues of the cosmological redshift and density dilution from the expansion. The Euler equations have similar interpretations. The expansion makes particle momenta decay as a^{-1}. The cosmological redshift of T accounts for this effect in the photons. For the baryons, it becomes the expansion drag on v_b ($\dot{a}/a$ term). Potential gradients $k\Psi$ generate potential flow. For the photons, stress gradients in the fluid, both isotropic ($k\delta p_\gamma/(p_\gamma + \rho_\gamma) = k\Theta$) and anisotropic ($k\pi_\gamma$) counter infall. Thomson scattering exchanges momentum between the two fluids ($\dot{\tau}$ terms).

For scales much larger than the mean free path $\dot{\tau}^{-1}$, the Euler equation may be expanded to leading order in $k/\dot{\tau}$, such that the photons are isotropic in the baryon rest frame $v_\gamma = v_b$ and so the joint Euler equation becomes

$$\frac{\mathrm{d}}{\mathrm{d}\eta}[(1+R)v_\gamma] = k[\Theta + (1+R)\Psi]. \tag{8}$$

Combining this with the continuity equation leads to the oscillator equation

$$\frac{\mathrm{d}}{\mathrm{d}\eta}[(1+R)\dot{\Theta}] + \frac{k^2}{3}\Theta = -\frac{k^2}{3}(1+R)\Psi - \frac{\mathrm{d}}{\mathrm{d}\eta}[(1+R)\dot{\Phi}]. \tag{9}$$

Note that the effective temperature $\Theta + \Psi$ is the local temperature Θ compensated for the loss of energy the photon experiences in climbing out of gravitational potential Ψ at recombination. This combination gives the Sachs-Wolfe effect for modes that are above the horizon at recombination.

Equation (9) is the fundamental relation for acoustic oscillations (see fig. 6). The change in the momentum of the photon-baryon fluid is determined by a competition between the pressure restoring and gravitational driving forces which causes the system to oscillate around its equilibrium. Note that the frequency of the oscillation

$$\omega^2 = \frac{1}{3(1+R)}k^2 = c_s^2 k^2, \tag{10}$$

where c_s is the sound speed of the fluid.

Now let us consider the effects of the finite mean free path of the photons from (4). On scales $\lambda \gg \lambda_C$, the photon and baryon bulk velocities are equal and the photons carry no anisotropy in the rest frame of the baryons due to rapid scattering. For the fluid model this implies that the photon-baryon fluid is nearly ideal: there is no heat

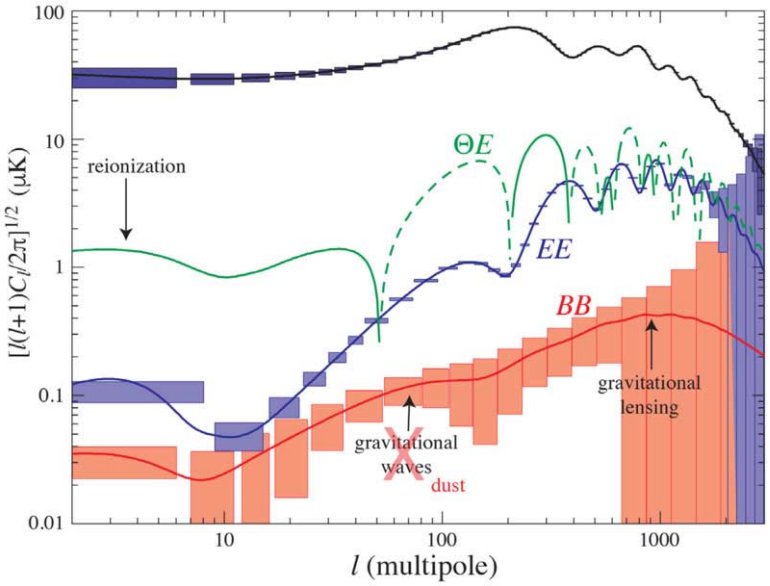


Fig. 6. – The landscape of temperature and polarization anisotropies plotted as the rms fluctuation or $[l(l+1)C_l/2\pi]^{1/2}$ with forecasted measurements based on ΛCDM for the post-Planck era from [2] in 2002. These forecasts have largely come to pass through the remarkable experimental and instrumental progress and the ΛCDM model has surviving stringent tests with only mildly significant hints of internal and external tension. The main exception is the interpretation of the first detection of BB power at $l < 100$ on the sky by the BICEP experiment as gravitational waves in the BB spectrum, which turned to dust.

conduction or viscosity in the fluid. Imperfections in the fluid related to the mean free path first appear at the diffusion length, where photons from hot and cold regions meet and dissipate their fluctuations. In the random walk approximation this is given by the square root of the number of scattering events N in the time period η,

$$\lambda_D = \sqrt{N}\lambda_C = \sqrt{\eta/\lambda_C}\,\lambda_C = \sqrt{\eta\lambda_C}, \tag{11}$$

so that the diffusion length $\lambda_D = 2\pi/k_D$ is the geometric mean between the horizon and mean free path. Since $\lambda_D/\eta_* \sim$ few % at recombination, we would expect the acoustic peaks beyond the third to be affected by dissipation.

Microphysically, the dissipation is due to viscosity and heat conduction. The former is characterized by π_γ the quadrupole moment of the photons and therefore leads to polarization. The dissipation of the acoustic oscillations therefore leaves a signature in the polarization of the CMB in its wake. We expect from fig. 7 that π_γ is generated by streaming of photons when gradients in the dipole or bulk velocity meet. But this generation is suppressed by the isotropization effects of scattering as $1/\dot{\tau}$. Thus the net effect should scale as $\pi_\gamma \sim v_\gamma \frac{k}{\dot{\tau}}$. The full Boltzmann or radiative transfer equations

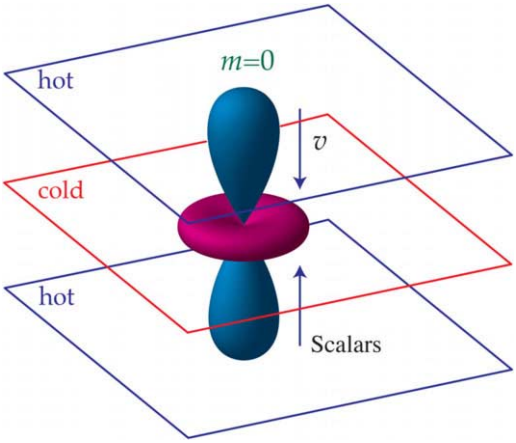


Fig. 7. – The fundamental $l = 2$, $m = 0$ temperature quadrupole is formed from scalar (density) fluctuations when the originally isotropic photons in crests (hot) and troughs (cold) of the plane wave lead to a dipole or bulk motion of the photons and then to a quadrupole as the dipoles flows reverse at the center. Since bulk flows must follow the original direction of temperature variation, the quadrupole is azimuthally symmetric ($m = 0$) around the axis of the plane wave and hence generates an E-mode polarization. If scattering is rapid then this anisotropy is rapidly destroyed by its tendency to randomize photon directions.

expanded to first order in $k/\dot\tau$ in fact give $A_v = 16/15$ of this expectation

$$\pi_\gamma = A_v \frac{k}{\dot\tau} v_\gamma. \tag{12}$$

Expanding the photon and baryon velocity difference in $k/\dot\tau$ gives a similar contribution weighted by $A_h = R^2/(1+R)$ correcting the oscillator equation (9) at order $k/\dot\tau$ to

$$c_s^2 \frac{\mathrm{d}}{\mathrm{d}\eta}(c_s^{-2}\dot\Theta) + \frac{k^2 c_s^2}{\dot\tau}(A_v + A_h)\dot\Theta + k^2 c_s^2 \Theta = -\frac{k^2}{3}\Psi - c_s^2 \frac{\mathrm{d}}{\mathrm{d}\eta}(c_s^{-2}\dot\Phi),$$

and the dispersion relation (10) to

$$-\omega^2 + \frac{k^2 c_s^2}{\dot\tau}(A_v + A_h) i\omega + k^2 c_s^2 = 0. \tag{13}$$

The imaginary component implies an exponential dissipation of the acoustic amplitude.

In fig. 8 we show the numerical solution for acoustic oscillations in temperature and the quadrupole moment. Because the quadrupole moment is azimuthally symmetric around the wave vector, the acoustic polarization is a pure E-mode. The fact that the polarization source is the quadrupole also explains the shape and height of the polarization spectra in fig. 6. Since the quadrupole is of order $kv_\gamma/\dot\tau \sim (k/k_\mathrm{d})(k_D\eta_*)^{-1}v_\gamma$, the polarization spectrum rises as l/l_D to peak at the damping scale with an amplitude of

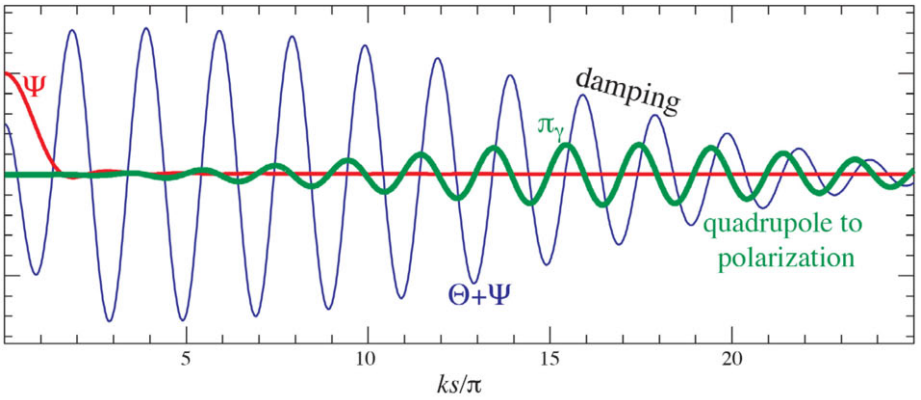


Fig. 8. – Polarization arises in the wake of the dissipation of acoustic oscillations. If you were expecting a diagram with balls and springs, you are too old to be reading this; but if you read this far at all, send me a message. Gravitational potentials (Ψ) drive the acoustic oscillation in the local CMB temperature ($\Theta = \Delta T/T$) in the radiation-dominated regime leading to compression and rarefaction in potential wells. As the photon mean free path becomes large enough due to the reduction in the free electron density, diffusion dissipates the oscillations and allows quadrupole moments (π_γ) to grow. These quadrupole moments then scatter into linear polarization with a structure that reflects the oscillations in π_γ and the gradients in bulk velocity that generate it.

about 10% of the temperature fluctuations before falling due to the elimination of the acoustic source itself due to damping. Since v_γ is out of phase with the temperature, the polarization peaks are also out of phase with the temperature peaks (see fig. 6). Furthermore, the phase relation also tells us that the polarization is correlated with the temperature perturbations. The correlation power $C_l^{\Theta\mathrm{E}}$ being the product of the two, exhibits oscillations at twice the acoustic frequency.

These acoustic peaks in the polarization have several advantages over those in the temperature. The first involves the fact that polarization only arises due to scattering and so it comes from recombination (and reionization, see below) and is unaffected by gravitational effects like the early integrated Sachs-Wolfe effect that can blur out features.

Moreover E-mode polarization from recombination has an intrinsically sharper transfer of features in k space to l space (see fig. 9). We can understand the much cleaner transfer geometrically. For the temperature fluctuations, the spherical decomposition of a plane wave tells us that the multipole components of the plane wave follow the decomposition $e^{i\mathbf{k}\cdot\mathbf{x}} \to j_l(kD)Y_{lm}$, where j_l is the spherical Bessel function. Mathematically then $j_l(kD)$ gives the transfer of k onto l and it has contributions for $l < kD$. Geometrically in fig. 5, if we imagine a plane wave going north-south from the equator to the pole, the projection of the plane wave on to the spherical shell implies that a single k-mode will generate a range of multipoles —the largest multipoles $l \sim kD$ come from

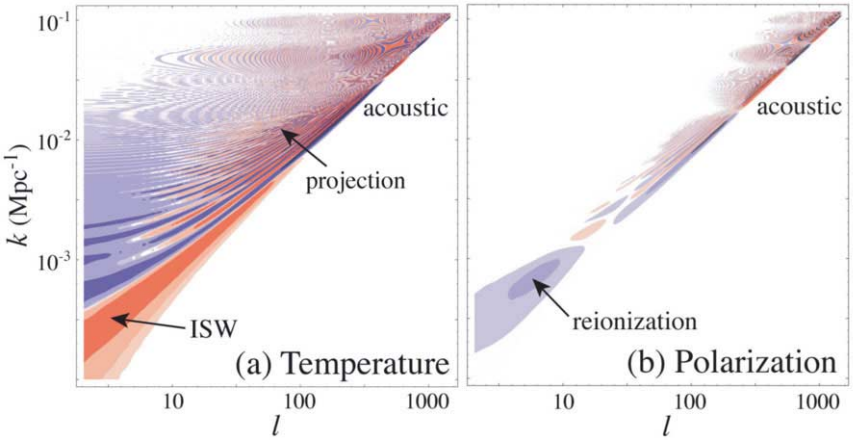


Fig. 9. – CMB transfer function which takes initial power in wave number k-space into multipole l-space. Due to projection effects, a single k-mode in the temperature field contributes to a broad range of multipoles $l < kD$, where D is the distance to the source, even in the acoustic regime and even more so at lower multipoles where the early and late ISW effects contribute. Polarization exhibits a cleaner, more one-to-one, transfer especially in the acoustic regime due to the way in which quadrupole moments project onto polarization (see fig. 5). Adapted from [4].

the equator with a tail of lower multipole contributions as we sweep toward the poles. This broad transfer is even broader for the Doppler effect $j_l \to j_l'$ —so much so that "Doppler" peaks from recombination cannot exist just due to geometry. The sum of all of these temperature contributions is shown in fig. 9 leading to the off-diagonal spread of power to the upper left.

Now consider the E-mode polarization from the projection of quadrupole moments (see fig. 10). At the equator, we have the same effect where the k-mode projects onto the maximum l mode. However the projection of the quadrupole vanishes as we go toward the poles, leaving only highly suppressed contributions at lower multipole moments. Mathematically this is described by the radial eigenfunction

$$\sqrt{\frac{3}{8}\frac{(l+2)!}{(l-2)!}}\frac{j_l(kD)}{(kD)^2}\,. \tag{14}$$

Therefore even though the temperature and polarization acoustic peaks depend on the same physics and the same parameters of ΛCDM, there are advantages for parameter estimation from the polarization. In fig. 11 we highlight the ultimate, cosmic variance limited, sensitivity to $\Omega_c h^2$ and hence to H_0 within ΛCDM. Beyond ΛCDM the consistency between acoustic temperature and polarization is a trigger for new physics.

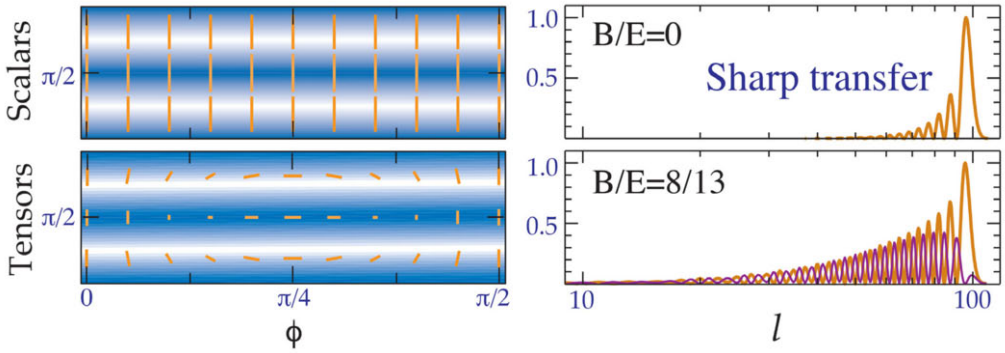


Fig. 10. – The observed polarization from a given scalar k-mode is the product of the fundamental $l = 2$, $m = 0$ pattern of fig. 4 with the modulation of the amplitude with the crests and troughs of the wave (shaded blue region). The modulation determines the angular scale of the polarization anisotropy and for the scalar pattern is dominated by equatorial lines of sight where the spatial variation is nearly in the plane of the sky $l \sim kD$. For lines of sight near the pole, where $l < kD$, the polarization is suppressed by the symmetry of the projected quadrupole moment. This leads to the clean transfer of power for the acoustic polarization shown in fig. 9 and characterized by a pure E-mode. The same considerations for tensor $l = 2$, $m = 2$ fluctuations lead to a much broader transfer onto both E- and B-modes with a geometrically determined ratio between the two (see sect. **4**).

4. – Inflation source

Inflation provides not only a source of cosmological density perturbations but also primordial gravitational wave perturbations. These gravitational wave perturbations carry their own polarization and their effect on the temperature quadrupole anisotropy at recombination (and reionization) leaves a B-mode imprint in the CMB polarization.

The gravitational wave amplitudes, *e.g.* in the + and × polarization states $h_{+,\times}$ obey the same Klein-Gordon type equation as the inflaton ϕ, a scalar field. Just like the inflaton, quantum fluctuations freeze out at horizon crossing during inflation with the power per $\ln k$ given by the Hubble scale H during inflation

$$\Delta^2_{\delta\phi} = \frac{H^2}{(2\pi)^2}\,, \qquad \Delta^2_{+,\times} = \frac{2}{M^2_{\rm pl}}\frac{H^2}{(2\pi)^2}\,, \tag{15}$$

where the different factors of the reduced Planck mass factor $M_{\rm pl} = 1/\sqrt{8\pi G}$ come from the canonical normalization of the fields. From the Friedmann equation, we can relate H to the potential energy scale of inflation $E_i = V^{1/4}$

$$H^2 = \frac{\rho}{3M^2_{\rm pl}} \approx \frac{V}{3M^2_{\rm pl}}\,, \tag{16}$$

and so a measurement of the gravitational wave amplitude will tell us the energy scale of inflation.

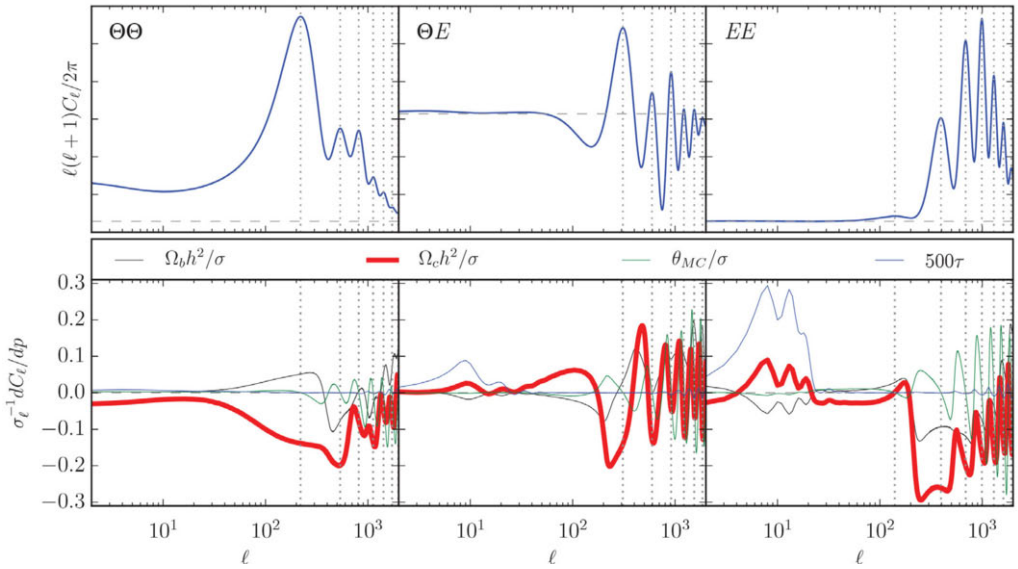


Fig. 11. – Acoustic spectra for ΛCDM and their sensitivity to the parameters of the model (in units of the cosmic variance errors per multipole σ_l). Highlighted in red is the sensitivity to $\Omega_c h^2$ which controls the inference for H_0 from the CMB. Note that the sharper transfer of power creates a sharper response especially around $l \sim 200$. This allows polarization to both improve on ΛCDM parameter estimation and provides a sharp consistency test for the model itself. Adapted from [5].

Gravitational waves are frozen in amplitude outside the horizon and begin oscillating and decaying once they cross back into the horizon after inflation, behaving thereafter like radiation. The change in the gravitational wave amplitude leaves a quadrupolar distortion in the temperature anisotropy of the CMB, much like the deformation of a test ring of particles (see fig. 12). The crucial difference between this type of quadrupole and that of scalar fluctuations (cf. fig. 7) is that the quadrupole is no longer azimuthally symmetric around the plane wave direction **k**. Instead, once aligned to this axis, the quadrupole moment appears as an $m = \pm 2$ state. During the breakdown of tight coupling that occurs at last scattering, any gravitational waves present will also imprint a local quadrupole anisotropy to the photons and hence a linear polarization to the CMB according to the projection of the cold lobe of the quadrupole moment.

In fig. 13, we show the pattern of an $l = 2$, $m = 2$ E-polarization field on the sky. Notice that there are now both Q and U components in the coordinate system of the plane wave. While this corresponds to the infinite wavelength limit with the U contribution being part of an E-mode, if we now consider that for a finite wavelength, the plane wave itself modulates the amplitude of the polarization between crests and troughs of the wave, this becomes a B-mode on the sky.

The gravitational wave contributions to the temperature and polarization power spectra are shown in fig. 14. The peak in the B-mode power spectrum occurs around $l \sim 10^2$

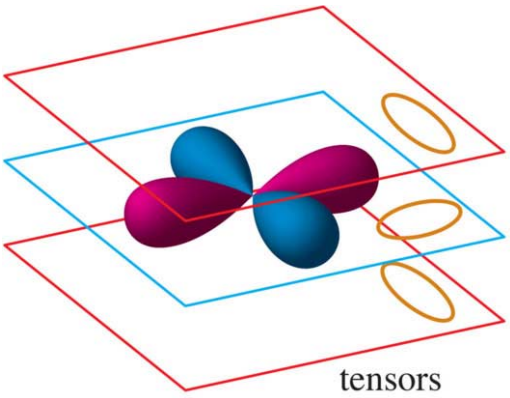


Fig. 12. – A change in the gravitational waves amplitude produces a quadrupolar distortion in the temperature of the CMB, just like its effect on a ring of test particles (inset ellipses). In the coordinate system where $\hat{\mathbf{z}} \parallel \hat{\mathbf{k}}$ (the wave vector of the gravitational wave), the quadrupole is $l = 2$, $m = \pm 2$. Since polarization follows the projection of the quadrupole, there are now two directions to consider: $\hat{\mathbf{k}}$ (along which the amplitude varies) and the polarization direction of the gravitational wave which is responsible for the direction of the CMB polarization after scattering.

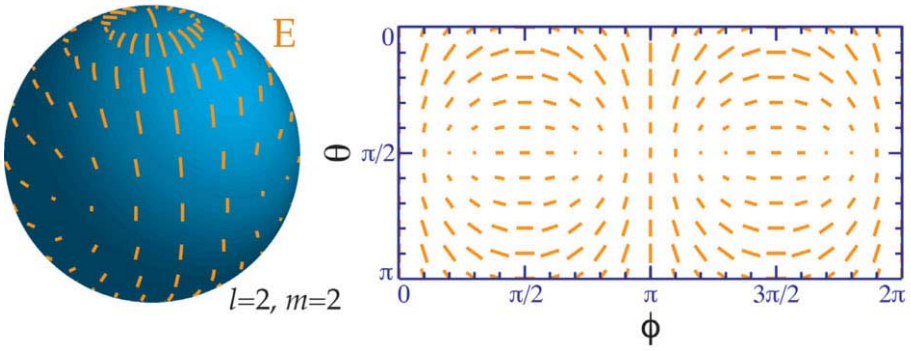


Fig. 13. – Gravitational waves imprint an $l = 2$, $m = \pm 2$, E-mode pattern to the polarization direction once projected on the sky in the coordinate system where the poles are aligned with its wave vector $\hat{\mathbf{k}}$. The polarization has both Q and U components in this system. Once modulated by the gravitational wave plane wave amplitude, as in fig. 10, the U-component becomes a B-mode in the polarization anisotropy.

since this corresponds to gravitational waves that cross the horizon around recombination. The peak amplitude can be translated into the energy scale of inflation roughly as

$$B_{\rm peak} \approx 0.024 \left(\frac{E_i}{10^{16}\,{\rm GeV}}\right)^2 \mu{\rm K}. \tag{17}$$

For the temperature and E-modes of course the observed spectrum will be the sum

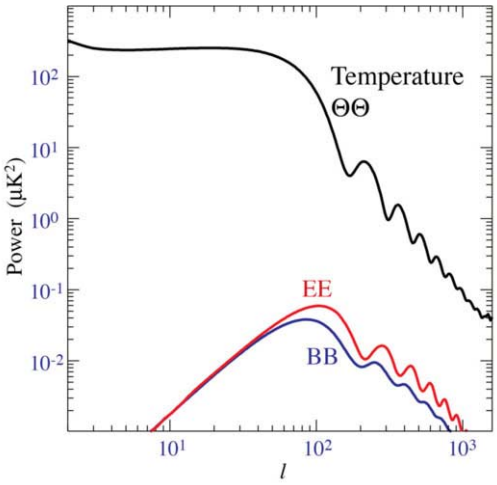


Fig. 14. – Temperature and polarization power spectrum from scale-invariant gravitational waves from inflation. The temperature anisotropy comes from the quadrupolar distortion which their decay inside the horizon produces, modulated by their plane wave amplitude in projection on the sky; for $l > 100$ the waves enter the horizon before recombination and have oscillated and decayed during the fluid epoch when quadrupoles are suppressed by scattering. On the horizon scale at recombination, the quadrupole moments most efficiently transfer onto polarization leading to both EE and BB power leaving peaks in both.

of the scalar and tensor contributions to the power spectrum. This is conventionally characterized by the tensor-scalar ratio

$$r = 4\frac{\Delta_+^2}{\Delta_{\mathcal{R}}^2} = 16\epsilon, \tag{18}$$

where $\Delta_{\mathcal{R}}^2$ is the power spectrum of comoving curvature fluctuations

$$\Delta_{\mathcal{R}}^2 = \frac{1}{2M_{\rm pl}^2\epsilon}\Delta_{\delta\phi}^2 = \frac{H^2}{8\pi^2 M_{\rm pl}^2\epsilon} \tag{19}$$

and ϵ is the slow roll parameter

$$\epsilon = \frac{1}{2M_{\rm pl}^2}\left(\frac{{\rm d}\phi}{{\rm d}\ln a}\right)^2. \tag{20}$$

Since linear scales span ${\rm d}\ln a = {\rm d}\ln k \sim 5$, the inflaton must roll by at least

$$\Delta\phi > 5\frac{{\rm d}\phi}{{\rm d}\ln a} \approx 5\left(\frac{r}{8}\right)^{1/2} M_{\rm pl} \approx 0.6\left(\frac{r}{0.1}\right)^{1/2} M_{\rm pl}^2. \tag{21}$$

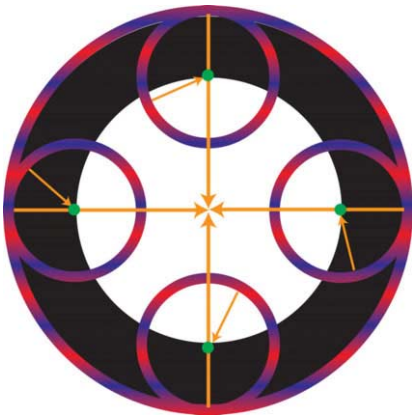


Fig. 15. – When the universe reionizes (transition from black to white), the free electrons in the ionized medium (green) can again scatter CMB photons (orange), but only a small fraction of them due to the density dilution of the expansion since recombination. The photons that are scattered out of the line of sight are replaced by photons from random directions, which were last scattered during recombination in regions with different temperature anisotropy (inset circles). These fluctuations average out, leaving the anisotropy from recombination suppressed by $e^{-\tau}$ in amplitude.

Therefore for tensors to be observable near the current bounds of $r < 0.1$, the inflaton must roll by a Planck distance. A detection at the level would have dramatic implications for the UV completion of inflation as an effective theory. An example of a scalar-tensor ratio that is slightly larger than this bound is shown in fig. 6.

5. – Reionization source

From the lack of a Gunn-Peterson trough in quasar spectra, we have long known that the universe is reionized at low redshift, with current bounds requiring nearly full ionization by $z \sim 6$. Even at this minimal level, a few percent of CMB photons will have last scattered during reionization rather than recombination.

The impact of reionization on the acoustic temperature and polarization is to suppress power spectra by $e^{-2\tau}$, where τ is the optical depth through reionization. In fig. 15 we illustrate the suppression effect in a way that will also illuminate the regeneration of polarization on large scales. Here the universe undergoes sudden reionization, depicted by the transition from black to white. Given the low mean density of the universe at low redshift, most photons do not rescatter but rather stream directly toward the observer in the center. A few percent of photons rescatter of free electrons (green) during reionization. The photons that are scattered out of the line of sight are replaced by photons that are scattered into the line of sight. Since the original directions of these photons are randomized, these photons come from different places at recombination with different temperature anisotropies, shown as the smaller inscribed circles (which are

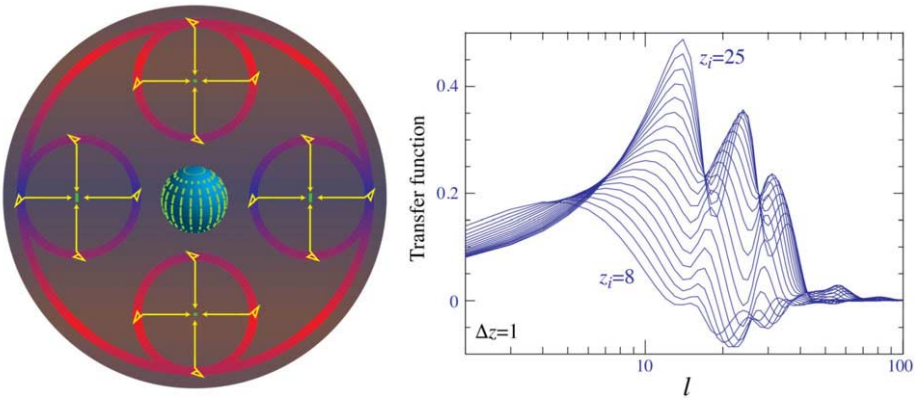


Fig. 16. – On scales near the horizon at the reionization scattering epoch z_i, the photons deflected into the line of sight have a quadrupole anisotropy from their last scattering surface at recombination (inset circles). Rescattering of the azimuthally symmetric quadrupole anisotropy leads to E polarization as usual leaving an imprint on its power spectrum corresponding to the projected scale of the horizon (right panel). The shape of the EE power spectrum thus carries information on both the total optical depth τ during reionization and the ionization history $x_e(z)$. Adapted from [6].

the next-to-last scattering surfaces), and their contributions average to zero. Therefore the original fine scale acoustic anisotropy from the last scattering surface is reduced by rescattering as $e^{-\tau}$ in amplitude or $e^{-2\tau}$ in power. Note that since a long wavelength perturbation generates a nearly uniform temperature inhomogeneity on the scale of the inset last scattering surface, this suppression applies to fluctuations that were subhorizon at reionization.

Now consider fluctuations that are horizon scale or larger at recombination. From our general discussion of polarization generation, we know that if there is a quadrupole anisotropy on these next-to-last scattering surfaces, then the result of rescattering will be the generation of linear polarization. This is depicted in fig. 16. In this case the Sachs-Wolfe temperature anisotropy appears as a quadrupole moment to the free electrons whose projection gives the E-mode polarization pattern shown in the center. Wavelengths around the horizon at rescattering produce the largest quadrupoles and hence reionization polarization provides a handle on both the amount of rescattering and its redshift. Furthermore, since the origin of the quadrupole moment is the Sachs-Wolfe effect, the polarization is correlated with the temperature fluctuations that come from photons that were not rescattered during reionization.

In fig. 16 we show the transfer function of scattering at a given redshift z_i onto the E-mode spectrum

$$T = \frac{\partial \ln C_l^{EE}}{\partial x_e(z_i)} \,. \tag{22}$$

The amplitude and shape of the E-mode power spectrum therefore give information not only on τ but also on the ionization history which is especially useful in constraining any contribution from high redshifts which would be otherwise difficult to detect with other techniques. To recover the full CMB information on reionization from this effect requires cosmic variance limited E-mode measurements out to multipoles $l < 30$.

6. – Lensing distortion

The intervening mass along the line of sight gravitationally lenses the CMB photon trajectories and hence distorts both the temperature and polarization anisotropy fields from recombination. The photons are deflected according to the angular gradient of the potential projected along the line of sight

$$\phi(\hat{\mathbf{n}}) = \int_{\eta_*}^{\eta_{a=1}} \mathrm{d}\eta \, \frac{(D_* - D)}{D\, D_*} [\Phi - \Psi](D\hat{\mathbf{n}}, \eta), \tag{23}$$

where D_* denotes the distance to recombination and $D = \eta_0 - \eta$. Because surface brightness is conserved in lensing, the deflection simply remaps the observed fields as

$$X(\hat{\mathbf{n}}) \to X(\hat{\mathbf{n}} + \nabla\phi), \tag{24}$$

where $X \in \{\Theta, Q, U\}$. The typical deflection angle is of order a few arcminutes but the lines of sight are coherently deflected across scales of a few degrees. Since the coherence scale of the acoustic features is larger than the deflection angle, the lensing effect can be approximated to leading order by Taylor expanding eq. (24). The result is a product of fields so that in harmonic space the modes are coupled to each other across a range $\Delta l \approx 60$ set by the coherence of the deflection. Heuristically, lensing distorts the hot and cold spots formed by the acoustic oscillations and hence the mapping of k to l.

In the temperature power spectrum, this mode coupling smooths the acoustic peaks slightly. For the polarization, the remapping not only smooths the E-polarization peaks but actually generates B-mode polarization (see fig. 17). Remapping by the lenses preserves the orientation of the polarization but warps its spatial distribution and hence does not preserve the symmetry of the original E-mode. Gravitational lensing B-modes represent a fundamental obstacle to the detection of a small tensor-scalar ratio r from inflation (see fig. 6).

Because the lensed CMB distribution is not linear in the fluctuations, the lensed CMB fields are non-Gaussian. In particular, the coupling of multipoles separated by Δl can be used to construct a minimum variance estimator of the deflection potential $\phi(\hat{\mathbf{n}})$ out of pairs of moments which can be iteratively improved.

These lensing effects are easiest to see with the temperature field in the flat-sky limit. Here the Taylor expansion gives

$$\begin{aligned}\Theta(\hat{\mathbf{n}}) &= \tilde{\Theta}(\hat{\mathbf{n}} + \nabla\phi) \\ &= \tilde{\Theta}(\hat{\mathbf{n}}) + \nabla_i\phi(\hat{\mathbf{n}})\nabla^i\tilde{\Theta}(\hat{\mathbf{n}}) + \frac{1}{2}\nabla_i\phi(\hat{\mathbf{n}})\nabla_j\phi(\hat{\mathbf{n}})\nabla^i\nabla^j\tilde{\Theta}(\hat{\mathbf{n}}) + \ldots,\end{aligned}$$

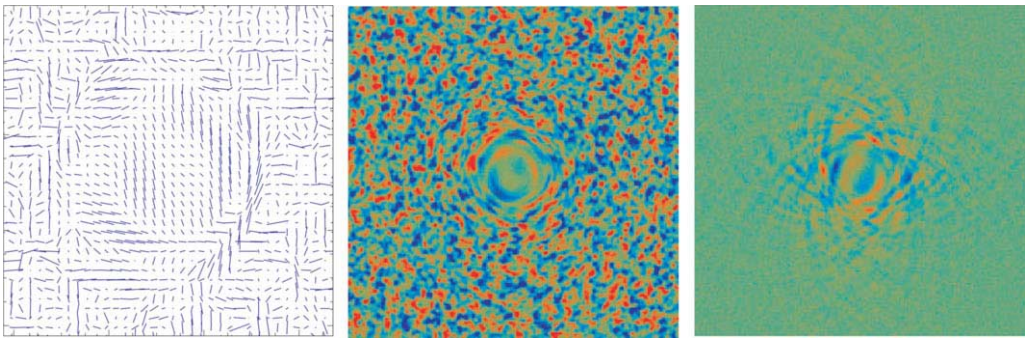


Fig. 17. – Gravitational lensing deflects the paths of CMB photons leading to a distortion of the acoustic polarization from recombination. In this exaggerated example of a spherically symmetric lens, the image in the center is magnified and the edges sheared (left). The radial direction of the lens deflection breaks the symmetry of the intrinsic E-mode (center) and creates a B-mode (right). Adapted from [7].

where tildes denote unlensed fields. In Fourier space

$$\begin{aligned}\Theta(\mathbf{l}) &= \int \mathrm{d}\hat{\mathbf{n}}\,\Theta(\hat{\mathbf{n}})e^{-il\cdot\hat{\mathbf{n}}} \\ &= \tilde{\Theta}(\mathbf{l}) - \int \frac{\mathrm{d}^2\mathbf{l}_1}{(2\pi)^2}\tilde{\Theta}(\mathbf{l}_1)L(\mathbf{l},\mathbf{l}_1),\end{aligned}$$

where

$$\begin{aligned}L(\mathbf{l},\mathbf{l}_1) = {}& \phi(\mathbf{l}-\mathbf{l}_1)\,(\mathbf{l}-\mathbf{l}_1)\cdot\mathbf{l}_1 \\ &+\frac{1}{2}\int \frac{\mathrm{d}^2\mathbf{l}_2}{(2\pi)^2}\phi(\mathbf{l}_2)\phi^*(\mathbf{l}_2+\mathbf{l}_1-\mathbf{l})\,(\mathbf{l}_2\cdot\mathbf{l}_1)(\mathbf{l}_2+\mathbf{l}_1-\mathbf{l})\cdot\mathbf{l}_1.\end{aligned}$$

The temperature power spectrum then becomes

$$C_l^{\Theta\Theta} = \left(1-l^2R\right)\tilde{C}_l^{\Theta\Theta} + \int \frac{\mathrm{d}^2\mathbf{l}_1}{(2\pi)^2}\tilde{C}^{\Theta\Theta}_{|\mathbf{l}-\mathbf{l}_1|}C_{l_1}^{\phi\phi}[(\mathbf{l}-\mathbf{l}_1)\cdot\mathbf{l}_1]^2,$$

where

$$R = \frac{1}{4\pi}\int \frac{\mathrm{d}l}{l}\,l^4C_l^{\phi\phi}.$$

If $\tilde{C}_l^{\Theta\Theta}$ slowly varies then two term cancel

$$\tilde{C}_l^{\Theta\Theta}\int \frac{\mathrm{d}^2\mathbf{l}_1}{(2\pi)^2}C_l^{\phi\phi}(\mathbf{l}\cdot\mathbf{l}_1)^2 \approx l^2R\tilde{C}_l^{\Theta\Theta}.$$

So lensing acts to smooth features in the power spectrum with a smoothing width of $L \sim 60$, the peak of deflection power spectrum. Lensing also correlates the modes so

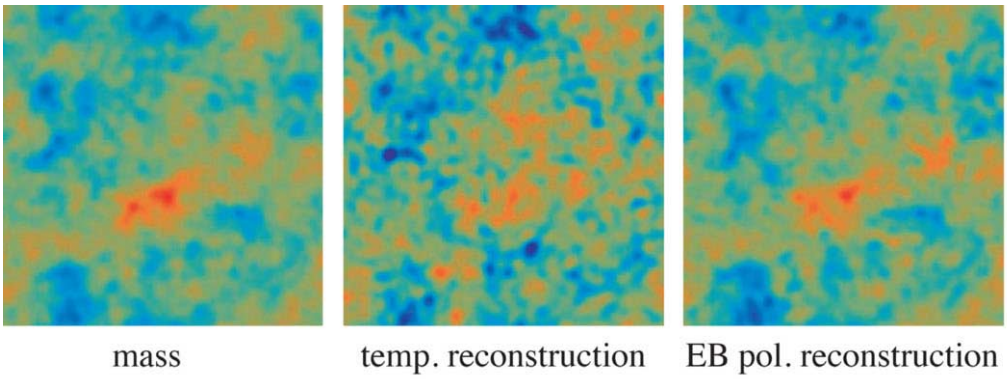


Fig. 18. – Lensing-induced correlations between multipole moments of the temperature and polarization fields can be used to reconstruct the projected mass distribution or convergence field. Because the intrinsic B-modes from recombination are small even with the maximal allowed tensor-scalar ratio r, the highest signal-to-noise pairs involve the observed E- and B-fields (right) leading to a better reconstruction than is available from the temperature multipole pairs. Adapted from [7].

that

$$\langle \Theta(\mathbf{l})\Theta'(\mathbf{l}')\rangle_{\rm CMB} = f(\mathbf{l},\mathbf{l}')\phi(\mathbf{L}), \tag{25}$$

where $\mathbf{L} = \mathbf{l} + \mathbf{l}'$ and

$$f(\mathbf{l},\mathbf{l}') = \tilde{C}^{\Theta\Theta}_{l_1}(\mathbf{L}\cdot\mathbf{l}_1) + \tilde{C}^{\Theta\Theta}_{l_2}(\mathbf{L}\cdot\mathbf{l}_2). \tag{26}$$

Here the average $\langle\ldots\rangle_{\rm CMB}$ denotes an ensemble average over realizations of the unlensed CMB. Each pair of multipoles therefore provides a noisy estimate of $\phi(\mathbf{L})$. We can generalize this to the polarization fields

$$\langle x(\mathbf{l})x'(\mathbf{l}')\rangle_{\rm CMB} = f_\alpha(\mathbf{l},\mathbf{l}')\phi(\mathbf{L}), \tag{27}$$

where $x \in \Theta, E, B$. Combining all of the estimators with minimum variance weights gives the optimal quadratic estimator.

Since small-scale polarization anisotropy is otherwise free of cosmological B-modes, most of the signal-to-noise in the reconstruction lies in the pairing of E-modes to neighboring lensing-generated B-modes. In fig. 18 we show a comparison of lensing reconstruction from the $\Theta\Theta$ modes *vs.* EB modes.

Once the lenses are reconstructed, the polarization of the CMB can be delensed, allowing for a much better measurement or limits on the inflationary tensor-scalar ratio r than can be achieved from just modeling their contribution assuming Gaussian statistics.

7. – Discussion

In these Proceedings, I have reviewed the sources of CMB polarization in the acoustic regime, reionization and gravitational waves from inflation and their distortion by gravitational waves. Throughout I have used the ΛCDM model to illustrate these effects. Whereas scalar fluctuations at recombination provide a source of E-mode polarization only, gravitational waves produce both E- and B-modes and gravitational lensing converts acoustic E-modes into B-modes. The non-Gaussianity of the lensed CMB fields allows the gravitational lenses to be reconstructed from quadratic and higher-order temperature and polarization estimators. Using this reconstruction, the observed polarization can be delensed and the gravitational wave B-modes largely removed. If their removal uncovers the B-modes from inflation, then the inflationary energy scale and field excursion will be determined which would provide invaluable information on the UV completion of inflation as an effective theory.

* * *

WH was supported by the DOE through DE-FG02-13ER41958, NASA ATP through NNX15AK22G, and the Simons Foundation.

REFERENCES

[1] Hu W. and White M. J., *New Astron.*, **2** (1997) 323.
[2] Hu W., *Annals Phys.*, **303** (2003) 203.
[3] Hu W. and White M. J., *Sci. Am.*, **290** (2004) 44, (*Spektrum Wiss.* **2004**, N.5 (2004) 46).
[4] Mortonson M. J., Dvorkin C., Peiris H. V. and Hu W., *Phys. Rev. D*, **79** (2009) 103519.
[5] Obied G., Dvorkin C., Heinrich C., Hu W. and Miranda V., *Phys. Rev. D*, **96** (2017) 083526.
[6] Hu W. and Holder G. P., *Phys. Rev. D*, **68** (2003) 023001.
[7] Hu W. and Okamoto T., *Astrophys. J.*, **574** (2002) 566.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200011

The legacy of *Planck*

C. Burigana

INAF, Istituto di Radioastronomia - Via Piero Gobetti 101, I-40129 Bologna, Italy

Dipartimento di Fisica e Scienze della Terra, Università di Ferrara
Via Giuseppe Saragat 1, I-44122 Ferrara, Italy

INFN, Sezione di Bologna - Via Irnerio 46, I-40126, Bologna, Italy

T. Trombetti

INAF, Istituto di Radioastronomia - Via Piero Gobetti 101, I-40129 Bologna, Italy

Dipartimento di Fisica e Scienze della Terra, Università di Ferrara
Via Giuseppe Saragat 1, I-44122 Ferrara, Italy

INFN, Sezione di Ferrara - Via Giuseppe Saragat 1, I-44122 Ferrara, Italy

on behalf of the *Planck* Collaboration

Summary. — While the final release of products and papers from the *Planck* mission is forthcoming, new much more ambitious projects dedicated to the Cosmic Microwave Background (CMB) are in preparation for the next decade and beyond. We review the *Planck* products, in form of source catalogs and maps, briefly discussing the methods to produce them and extracting the various types of astrophysical and cosmological signals, presenting their main implications. The limits on primordial B-modes associated to stochastic field of gravitational waves expected in inflationary scenarios and the perspectives open by the investigations of expected tiny spectral distortions are discussed. Finally, the studies carried out in the last decade towards future CMB space missions are described.

1. – Introduction

Since the discovery of the Cosmic Microwave Background (CMB) in 1964, and in particular in the past two decades, the improved measurements in temperature fluctuations achieved an unprecedented accuracy in constraining the majority of the cosmological parameters values [1]. Together with the cosmic expansion and the primordial nucleosynthesis, the CMB is one of the fundamental observational evidence of the hot big bang theory [2], representing a powerful tool of investigation in a cosmological context. The presence of small CMB temperature anisotropies would have been, as realized, the seeds for galaxy formation, hence, a direct probe of the formation and evolution of the structures we observe today.

In early 1990's the NASA COsmic Background Explorer (COBE) satellite inaugurated the "precision cosmology" era, determining the CMB spectrum temperature of $T_0 = 2.72548 \pm 0.00057\,\mathrm{K}$ [3-5] and then discovering the CMB temperature fluctuations of few tens of μK on angular scales larger than 7° [6], the so-called anisotropies. Remarkably, already in 1990 [7] argued that a positive cosmological constant is necessary to jointly explain large scale structures observations and CMB anisotropy upper limits at small angular scales in a spatially flat cosmology, anticipating the paradigm transition from a Cold Dark Matter (CDM) to a cosmological constant plus CDM (ΛCDM) model of the Universe accounting for the evidence of the recent cosmic acceleration derived through Type-Ia Supernovae (SNIa) observations [8] and the flatness of the Universe determined by the Balloon Observations Of Millimetric Extragalactic Radiation and Geophysics (BOOMERanG) CMB anisotropy experiment [9].

The great amount of more recent available data coming from ground-based and balloon-borne experiments to space-borne experiments, as the Wilkinson Microwave Anisotropy Probe (WMAP) and *Planck* missions, allows to explore key aspects of cosmology and fundamental physics and to contribute to trace the roadmap for future microwave projects in the coming decades. Indeed, the wealth of information encoded in the CMB has significantly helped cosmology to turn into a precision science, and allowed us to describe cosmic evolution starting from its very primordial stages. The linear polarization pattern of CMB anisotropies [10] has been generated in two distinct cosmological epochs: the recombination era ($z \simeq 1100$) during which, as the Universe expanded and cooled, the diffusion length of photons in an increasingly neutral medium became large enough to reveal the quadrupole moment of the local anisotropy pattern at the last scattering surface and the Universe then became almost neutral and transparent to radiation, since Thomson scattering, the main physical process involved, was no longer efficient in coupling matter and radiation; much later, during the early phases of galaxy and star formation, when photoionizing radiation able to escape from the bound structures was injected in the intergalactic medium that became ionized again (reionization epoch). During this period, free electrons interacted, through Thomson scattering, with CMB photons once more. Owing to the size of the Hubble horizon at reionization, this new contribution is seen at large angular scales, as a characteristic bump at low multipoles ($\ell \lesssim 10$) in the pure polarization CMB spectra, as well as in the temperature to polarization correlation.

2. – The *Planck* mission

Planck([1]) is a project of the European Space Agency (ESA) and a third-generation mission in terms of precision cosmology, designed for the CMB anisotropy analysis. The satellite, launched in 2009, was equipped with a Gregorian dual reflector telescope of 1.5 m diameter, whose focal plane were able to collect the microwave photons owing to two wide band cooled receiver arrays, the Low-Frequency Instruments (LFI) made up of radiometers operative at 20 K, and the High-Frequency Instruments (HFI) composed of bolometers working at 100 mK. Together, the two instruments scanned 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 [11, 12]. *Planck* observed the sky from a large Lissajous orbit around the Sun-Earth L2 Lagrange point designed to offer stable observing conditions and avoids contamination from side lobe pick-up of stray radiation originating from the Sun, Earth, and Moon, scanning the sky while spinning at 1 rpm in almost great circles with a full width half maximum (FWHM) resolution ranging from $33.3'$ to $4.3'$ from 30 GHz to 857 GHz, and a final sensitivity per FWHM^2 resolution element spanning ~ 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, then extracting Stokes T, Q, U parameter maps of CMB anisotropies on the whole sky with an 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_ℓ^{XY}, where X and Y stands for T, E, B, mainly related to the anisotropy amplitude at the multipole $\ell \sim 180°/\theta$, θ being the angular scale of interest. A linear combination of Q, U allows us to define, in Fourier space, the E-mode (gradient component) and B-mode (curl component) polarization patterns and to identify, together with temperature anisotropies, four angular power spectra (APS), C^{TT}, C^{EE}, C^{BB}, C^{TE}, the latter being the temperature polarization cross correlation([2]). Scalar perturbations generate E-mode polarization, tensor perturbations generate both E- and B-mode polarization([3]).

([1]) *Planck* hosted instruments provided by two scientific Consortia funded by ESA member states (in particular the lead countries: France and Italy) with contributions from NASA (USA), and telescope reflectors provided in a collaboration between ESA and a scientific Consortium led and funded by Denmark.

([2]) Other cross-correlators, C^{TB} and C^{EB}, expected to vanish due to the different handedness of the B and (T, E) harmonics because of the parity conserving of standard physics of CMB fluctuations, may be induced by parity-asymmetric gravity dynamics during inflation in the case of a discrepancy among left and right-handed gravitational waves or in particle physics models with non-standard parity-violating interactions, while the presence of a primordial homogeneous or helical magnetic field which would induce Faraday rotation and non-zero TB correlations (see, *e.g.*, [13] and references therein).

([3]) Vector perturbations generate both modes, but, except for particular scenarios, they are expected to be typically subdominant.

Besides the so-called cosmic variance limit, an intrinsic error which has an impact on, primarily, the low multipoles, the main original goal of *Planck* was the accurate, almost definitive measure of the CMB temperature fluctuations, but the increasing relevance of polarization motivated a stringent analysis of polarization performance in the various steps of the mission, from design and ground tests to in flight data analysis and scientific interpretation. The development and refinement of new and already existing methodologies were crucial for the success of the cosmological investigations. The large majority of scientific analyses of sky maps requires suitable pixelization schemes of the sphere improving over previous ones [14-16] to include efficient methods to pass from the real space to the harmonic space representation. It is worthy to mention in this context the crucial development of fast analysis methods based on the relationship between fast Fourier transforms and spherical harmonic expansion [17] for anisotropy maps discretized through pixels located on the sphere parallels and of suitable pixelization schemes also satisfying other extremely useful features, such as pixel area equality and pixel hierarchical structure in passing from different resolutions, and complemented with a wide set of tools and facilities as in the HEALPix tesselation [18, 19], widely adopted in CMB projects and astrophysical studies, or variants implementing an IGLOO scheme [20] or with pixel centers located to optimize quadrature with Gaussian-Legendre method for converting T, Q, U maps into spherical harmonics or spin ± 2 spherical harmonics expansions coefficients as in the GLESP-pol pixelization [21].

3. – Control of systematic effects

To look at this extremely faint signal, it is crucial the almost optimal control of systematics effects, one of the main task in data reduction and interpretation [24-27]. Table I lists the main systematic effects in *Planck* according to their source and summarizes their control and residual impact on science. Systematic effects can be additive or multiplicative spurious contributions independent of the sky signal, or coupled with the sky signal and related with the observational strategy. Furthermore, in the microwaves, different kinds of astrophysical contamination emissions can, in principle, veil the cosmological information. This foreground signal can be diffuse, coming from the Solar System and from the Galaxy, or discrete, due to Galactic sources and extragalactic sources at various cosmic distances.

Table II report the main characteristics of LFI and HFI. Figure 1 shows potential residuals for the EE auto-power spectra of systematic effects for HFI. The level of systematics effects has been in line with expectations, therefore confirmed the belief of a full science exploitation of the CMB signal.

4. – Astrophysical foregrounds

To investigate a great variety of astrophysical and cosmological processes it is crucial the analysis of the microwave to sub-millimetre sky and of the evidences carried by the CMB. The data acquired from different sky signals and at various epochs or maps

TABLE I. – *Main systematic effects in Planck data.*

Category	Effect	Notes – Control/Removal
Effects coupled with sky		
	Main beam distortions	Relevant in both temperature and polarization maps. Accounted for in window function for APS estimation.
	Nearly-lobes pickup	Optical response at angles $< 5^\circ$ from the main beam. Main effect from Galaxy and point sources. Accounted through simulations.
	Side-lobes pickup	Galaxy and CMB dipole pickup by main and sub-reflector spillovers. Small effect on temperature maps, relevant effect for polarization study, particularly at large angular scales. Accounted through simulations.
	Polarization angle uncertainty	Uncertainty in the polarization angle in-flight measurement.
	Bandpass mismatch	Affects primarily radiometric detectors. Negligible impact on temperature. Corrected in polarization at map level exploiting source measurements (*e.g.* Crab Nebula).
	Imperfect photometric calibration	Optics, receiver noise properties changes, and other non-idealities. Adaptive smoothing algorithm using beam characterization, 4-K reference load voltage output, temperature sensor data.
	Pointing	Uncertainties in pointing reconstruction, thermal changes affecting focal plane geometry. Small impact on anisotropy measurements.
	Receiver time response	In HFI, bolometer memory is coupled with beam shape. Filters, window function.
Detectors		
	Cosmic ray hits	Affect bolometric detectors. Timelines cleaning via template fitting.
	$1/f$-type noise	Affects radiometric and bolometric detectors. In LFI, $1/f$ contribution limited to max 3% by differential measurement strategy. Destriping/map-making algorithms.
	Orthomode transducer cross-polarization	In LFI, imperfect polarization separation. Negligible impact.
Electronics		
	1 Hz spikes	Affects LFI data. Removed from timelines by template fitting.
	Analogue-to-digital converter (ADC) nonlinearity	Reduced/removed through template fitting.
Thermal[(a)]		
	300 K fluctuations	In principle affect both instruments. Inherent hardware stability is compliant with scientific requirements.
	20 K fluctuations	Affect mainly LFI. Inherent hardware stability is compliant with scientific requirements.
	4 K fluctuations	Affect both instruments. Inherent hardware stability is compliant with scientific requirements.

[(a)] A combination of differential measurement strategy (specifically for LFI) with calibration and de-striping further reduce the effect.

Table II. – *Main characteristics of LFI and HFI full mission maps.*

Characteristic	Frequency band		
	30 GHz	44 GHz	70 GHz
Centre frequency [GHz]	28.4	44.1	70.4
Effective beam FWHM[a] [arcmin]	32.29	27.00	13.21
Effective beam ellipticity[a]	1.32	1.04	1.22
Temperature noise (1 deg)[b] [μK_{CMB}]	2.5	2.7	3.5
Polarization noise (1 deg)[b] [μK_{CMB}]	3.5	4.0	5.0
Overall calibration uncertainty[c] [%]	0.35	0.26	0.20
Systematic effects uncertainty in Stokes I^d [μK_{CMB}]	0.19	0.39	0.40
Systematic effects uncertainty in Stokes Q^d [μK_{CMB}]	0.20	0.23	0.45
Systematic effects uncertainty in Stokes U^d [μK_{CMB}]	0.40	0.45	0.44

Characteristic	Reference frequency						
	100 GHz	143 GHz	217 GHz	353 GHz	545 GHz	857 GHz	Notes
Number of bolometers	8	11	12	12	3	4	*a*1
Effective beam FWHM$_1$ [arcmin]	9.68	7.30	5.02	4.94	4.83	4.64	*b*1
Effective beam FWHM$_2$ [arcmin]	9.66	7.22	4.90	4.92	4.67	4.22	*b*2
Effective beam ellipticity ε	1.186	1.040	1.169	1.166	1.137	1.336	*b*3
Noise per beam solid angle [μK_{CMB}]	7.5	4.3	8.7	29.7	–	–	*c*1
[kJy sr^{-1}]	–	–	–	–	9.1	8.8	*c*1
Temperature noise [μK_{CMB} deg]	1.29	0.55	0.78	2.56	–	–	*c*2
[kJy sr^{-1} deg]	–	–	–	–	0.78	0.72	*c*2
Polarization noise [μK_{CMB} deg]	1.96	1.17	1.75	7.31	–	–	*c*3
Calibration accuracy [%]	0.09	0.07	0.16	0.78	1.1(+5)	1.4(+5)	*d*1
CIB monopole prediction [MJy/sr]	0.0030	0.0079	0.033	0.13	0.35	0.64	*e*1
Zodiacal light level correction [K_{CMB}]	4.3×10^{-7}	9.4×10^{-7}	3.8×10^{-6}	3.4×10^{-5}	–	–	*e*2
[MJy/sr]	–	–	–	–	0.04	0.12	*e*2

[a] Calculated from the main beam solid angle of the effective beam, Ω_{eff} = mean(Ω). These values are used in the source extraction pipeline.
[b] Noise rms computed after smoothing to 1°.
[c] Sum of the error determined from the absolute and relative calibration.
[d] Estimated rms values over the full sky and after full mission integration. Not included here are gain reconstruction uncertainties, estimated to be of order 0.1%.
[a1] Number of bolometers whose data were used in producing the channel map.
[b1] FWHM of the Gaussian whose solid angle is equivalent to that of the effective beams.
[b2] FWHM of the elliptical Gaussian fit.
[b3] Ratio of the major to minor axis of the bestfit Gaussian averaged over the full sky.
[c1] Estimate of the noise per beam solid angle, as given in *b1*.
[c2] Estimate of the noise in intensity scaled to 1° assuming that the noise is white.
[c3] Estimate of the noise in polarization scaled to 1° assuming that the noise is white.
[d1] Calibration accuracy (at 545 and 857 GHz, the 5% accounts for the model uncertainty).
[e1] Cosmic Infrared Background (CIB) according to the model by [22], whose uncertainty is estimated to be at the 20% level (also for constant νI_ν).
[e2] Zero level correction to be applied on Zodiacal Light corrected maps. From [23]. Credit: Planck Collaboration, A&A 594, A1, 2016, reproduced with permission © ESO.

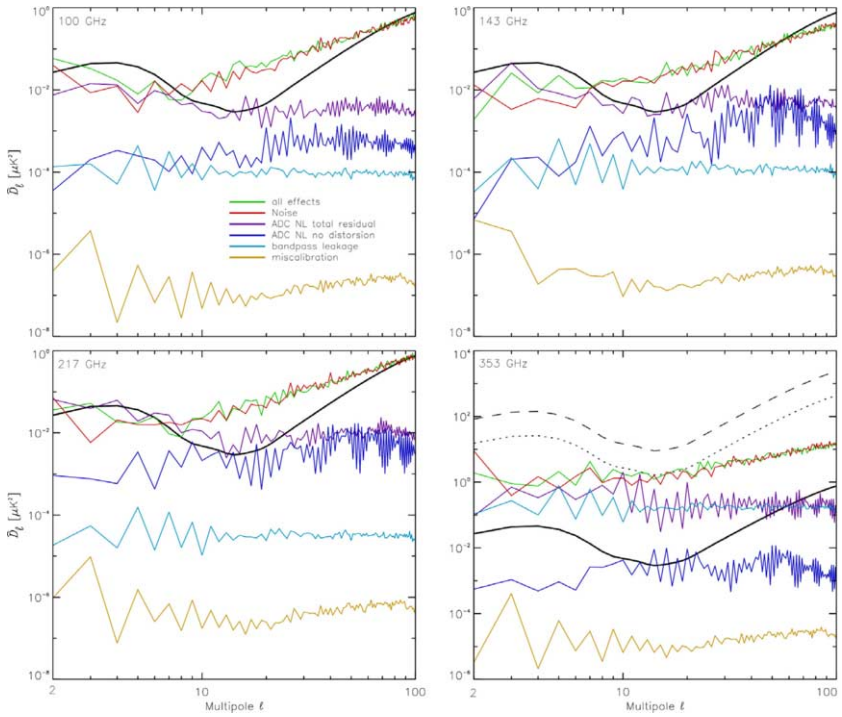


Fig. 1. – Residual *EE* auto-power spectra of systematic effects from the HFI pre-2016 end-to-end simulations computed on 50% of the sky (colours specified in the top left panel apply to all panels). The purple line (ADC NL total residual) shows the sum of all effects associated with ADC nonlinearity. The dark blue line (ADC NL — Analogue-to-digital converter nonlinearity — no distortion) shows the level without the dominant dipole distortion. The plots show also the F-*EE* fiducial model from best-fit *Planck* 2015 cosmological parameter [28] (black curves). The 100-GHz and 143-GHz model scaled to 353 GHz with a dust SED is shown as dashed and dotted lines, respectively. From [29]. Credit: Planck Collaboration, *Astron. Astrophys.*, **596** (2016) A107, reproduced with permission © ESO.

are compared, because of the time dependence and relative positions of observer-source or -interplanetary dust cloud, in order to extract relatively bright variable sources, the diffuse Zodiacal Light Emission and solar system bodies. Once the maps are averaged over valuable number of surveys, several point-like sources and specific diffuse emissions filters are applied to disentangle between all other astrophysical foregrounds and the CMB anisotropies, possibly supplemented by external catalogs. These algorithms exploit the signal frequency dependence and angular correlation, in combination (if feasible) with other frequency bands templates.

4˙1. *Catalogs of sources and clusters of galaxies*. – The Second *Planck* Catalogue of Compact Sources (PCCS2), based on the full mission data, lists Galactic and extragalactic sources detected in single-frequency maps over the entire sky (see fig. 2). At the HFI frequencies this is divided into two sub-catalogues depending on sources sky positions

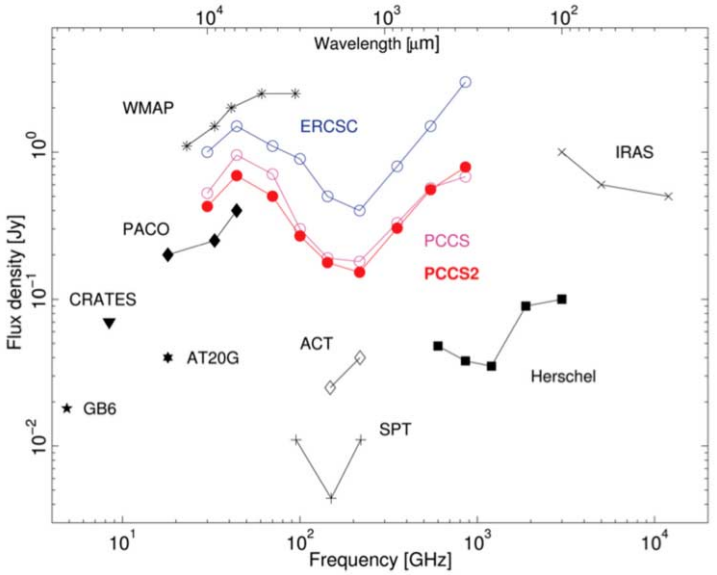


Fig. 2. – Sensitivity (the flux density at 90% completeness) of the PCCS2, PCCS, ERCSC, WMAP and others. For the LFI channels, the sensitivities refer to the full sky. For the HFI channels, the 90% completeness limits plotted for the PCCS were evaluated in a selected extragalactic zone; the regions of sky to which the 90% completeness limits apply are similar but not identical to those of the PCCS2. From [30]. Credit: Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A26, reproduced with permission © ESO.

and reliability, the PCCS2, identified by an integral reliability of 80% or greater, and the PCCS2E, classified by unknown reliability due to diffuse emission contamination. To extract the objects, the sky maps are convolved with a cleaning and denoising algorithm, the Mexican Hat Wavelet 2 (MHW2) [31, 32], specific for each sky map because of their individual characteristics. The filters are applied to square patches, the full maps projections [33], overlapped such to cover the entire sky and avoid multiple detection of the same sources. In LFI, a first comparison of the detection with existing radio sources catalogues is completed and then, for those unclassified outside the Galactic cut, a manual individual source inquiry into archival repositories is carried out. In HFI, because of the absence of external full sky catalogue, two different methods are applied, the simulation and injection reliability. The first represents the number of revealed sources that, injected into simulated maps, match the injected target positions. In the second, the injection occurs on real maps [30]. The first technique, preferable, is dependent on the quality of the simulations and, in the worst cases, the second approach is required. The maps have been masked on the Galactic plane extended to external filamentary structures and all sources located in this masked region categorized in the PCCS2E. The catalogues, based on total intensity measurements, include polarized flux densities estimated at the sources position and at frequencies $30\,\mathrm{GHz} \leq \nu \leq 353\,\mathrm{GHz}$. The analysis of the recov-

ered spectral indexes shows that synchrotron emitters dominate at low frequencies and in the bright tail of extragalactic sources at high frequencies ($\nu \leq 217\,\mathrm{GHz}$) while dusty galaxies prevail at high frequencies. At $\nu \sim 50\,\mathrm{GHz}$, blazar spectra considerably steepens. Characterizing the contamination of unresolved polarized extragalactic sources is a crucial step for accurate CMB polarization analyses and to detect the primordial B-modes. The PCCS2 includes more than 120 polarized objects at $> 99.99\%$ confidence level (c.l.), mainly located at low Galactic latitudes, while the detection of extragalactic polarized sources is still very poor, ranging the unity at 353 GHz. However, it is possible to overcome this lack of information by statistically estimating fainter sources polarization fraction, Π, with different approaches. The first one is the stacking technique, a method to evaluate the polarization behaviour of compact sources by adding up numerous sky regions surrounding previous selected positions, such to increase the signal and reduce the background, and dealing with the noise bias [34,35]. This led to a radio source (dusty galaxies) median Π of $\simeq 1.9\%$ ($\simeq 1.3\%$ and $\simeq 2.0\%$ at 217 and 353 GHz). The second approach, the Intensity Distribution Analysis, validated on other radio source catalogs with polarization data, evaluates the signals at provided sources positions and compares them with random sky locations far from the sources (control fields), assessing the two distributions difference significance with the one-sided Kolmogorov-Smirnov test. Averaging over the frequencies, the radio sources median polarization fraction resulted to be $\simeq 2.83\%$, consistent with previous analyses, while for dusty galaxies a 90% c.l. upper limit of median Π of $\simeq 2.2\%$ at 353 GHz and of $\simeq 3.9\%$ at 217 GHz (where dusty galaxies are fainter) was derived [36].

The Second *Planck* catalogue of Sunyaev-Zeldovich (SZ) galaxy clusters (PSZ2), the largest and deepest all-sky selected galaxy clusters archive, encloses 1653 sources. Typically, a cluster emits in optical and infrared (IR), being formed by different components such as dark matter, cold gas and dust in the galaxy, while it emits in the X-ray band owing to ionized intra cluster medium thermal bremsstrahlung and gives rise to energy boosts of CMB photons via inverse Compton, the so-called thermal SZ effect [37, 38]. This effect induces an intensity spectrum increase above 220 GHz and a decrease at lower frequencies, a suitable range at HFI frequencies to provide high-precision data. The extraction algorithm, divided into 3 steps, consists of a Matched-Multi Filters (MMF) technique [39-41] and a fast, fully Bayesian, multifrequency detection approach, the PowellSnakes (PwS) method [42], designed to identify and characterize compact objects in a diffuse background.

Among the measured sources, 1203 have revealed a counterpart in other datasets external to *Planck* once analyzed in other wavelengths. In particular, the SZ selection recognized a low-redshift X-ray under-luminous clusters population with abundant mass content characteristics, as revealed in optical surveys but absent in X-ray samples. Figure 3 compares PSZ2 with deeper SZ catalogs obtained with South Pole and Atacama Cosmology Telescopes (SPT and ACT) on selected areas. Microwave and X-ray cluster data have been jointly used to independently constrain cosmological parameters. A fundamental property of galaxy cluster abundance is that from their mass and redshift evolution one can infer constraints on matter density fluctuation normalization, mean

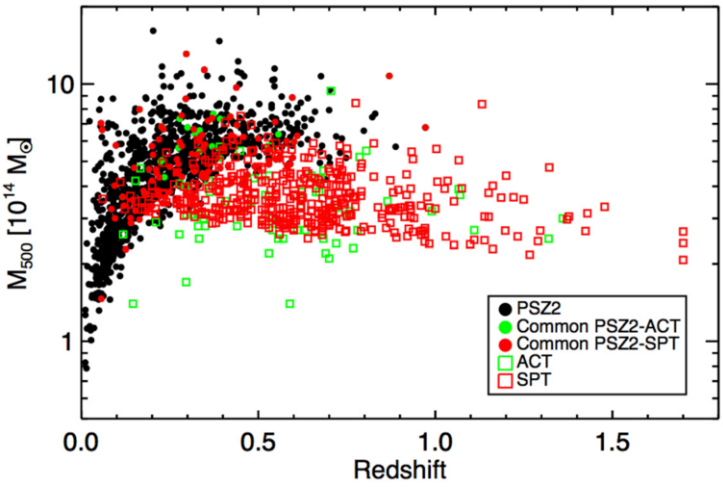


Fig. 3. – Distribution of PSZ2 clusters with associated redshift in the mass-redshift (M500 – z) plane compared to SPT and ACT catalogues. Black circles: PSZ2 clusters; red (green) filled circles: common SPT/PSZ2 (ACT/PSZ2) clusters. Red (green) empty squares: the remaining SPT (ACT) clusters not detected by *Planck*. See also the text. From [45]. Credit: Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A27, reproduced with permission © ESO.

matter density, dark energy density and state equation and on extensions of the minimal cosmological model, including *e.g.* massive neutrinos and modified gravity. The Comptonization parameter, y, and its full-sky *Planck* maps fluctuations [12] represent important tools to investigate integrated information on clusters physics and cosmological evolution as well as to analyze the SZ effect through selected clusters. High-z galaxy evolution and star formation history can be studied in the sub-millimetre/far-IR background fluctuations, specially when performed in privileged *Planck* areas [43]. To investigate these topics, it is also crucial the CMB polarization study, which encloses valuable information on cosmic reionization history [23, 44], structure and star formation and evolution and the ionizing photons they originate in the intergalactic medium.

4˙2. *Galactic diffuse components*. – In order to separately reveal CMB anisotropies and astrophysical diffuse components, the *Planck* single-frequency full-sky maps, have been reduced with four different algorithms of component separation [46]. They have been developed in real space (SEVEM and COMMANDER), real and harmonic (needlet) space (NILC), in harmonic or needlet domain (SMICA) [47]. Even though the cleaned maps were in good agreement, the COMMANDER maps turned out to be the most suitable to trace the sky astrophysical processes owing to the real-space Bayesian technique that allows *a priori* foreground parameters representation [48]. Aiming at synchrotron, spinning and thermal dust, free-free and CO emissions, temperature *Planck* and WMAP [49] measurements were merged together with 408 MHz [50] maps. Among the various polarized diffuse emissions only synchrotron and thermal dust maps were resolved (see fig. 4).

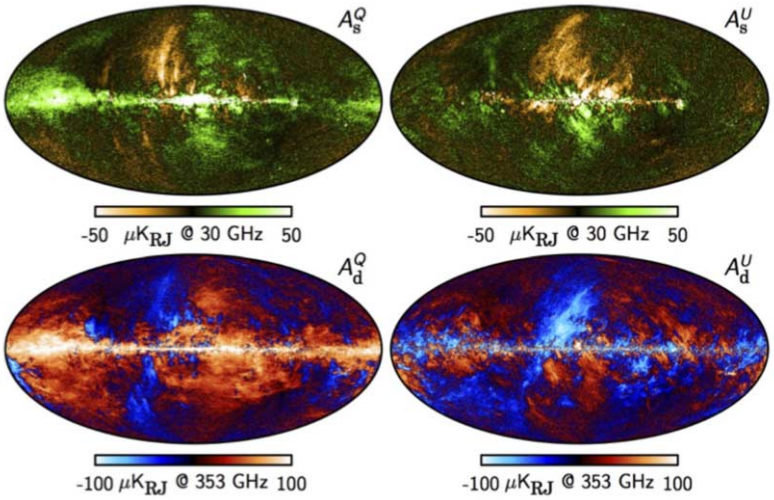


Fig. 4. – Stokes Q and U polarization parameters for the synchrotron at 30 GHz (upper line) and thermal dust at 353 GHz (bottom line). From [23]. Credit: Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A1, reproduced with permission © ESO.

In *Planck* LFI data, the applied methods of component separation were able to reveal and describe a distinct and non-negligible emitting constituent close to the Galactic centre, the so-called "haze" [51]. Recent *Fermi* observations evidenced a strong morphological correlation between the haze and gamma-ray haze/bubbles, supporting the idea of two multi-wavelength views of the same phenomenon.

Figure 5 displays the foreground contribution to the sky fluctuations. The *Planck* mission had the capability to estimate the microwave sky complexity, requiring, for future CMB experiments, an increased number of frequency channels in order to accurately characterize the foregrounds and the tiny CMB polarization features, such as primordial B-modes. *Planck* data also changed our comprehension of the relative weights of the various components. Indeed, in temperature synchrotron fluctuations are smaller than free-free and spinning dust around 30 GHz. The synchrotron and dust polarized emissions APS has been precisely outlined for a variety of wide sky areas [52] and the synchrotron frequency dependence has been investigated as a function of the Galactic latitude [53]. At frequencies above 70 GHz, polarized emission from Galactic dust prevails extensively in the sky and, even though previously occasionally underestimated, *Planck* maps accurately supplied its contamination level in CMB APS [54] for polarization projects, turned to be crucial for BICEP2 and *Keck* array observations [55]. Furthermore, synchrotron and dust polarized emissions APS have been accurately described analyzing different patches in the sky [52] and the synchrotron Galactic latitude dependence was investigated [56].

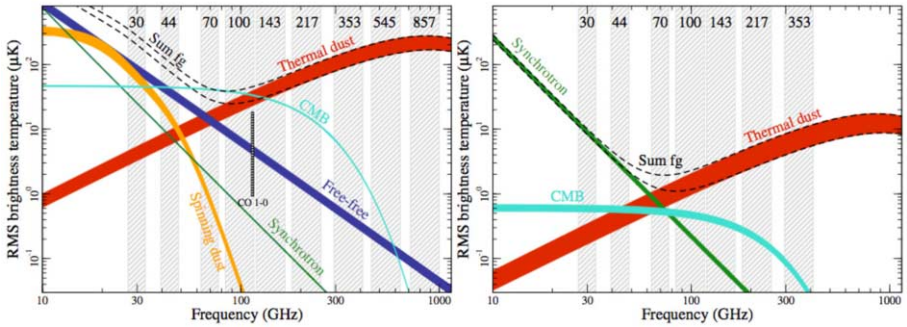


Fig. 5. – Brightness temperature rms as a function of frequency and component for temperature (left), where each component is smoothed to an angular resolution of 1° FWHM and the lower and upper edges of each line are defined by masks covering respectively 81% and 93% of the sky, and polarization (right), where the corresponding smoothing scale is 40′ and the sky fractions are 73% and 93%. From [23]. Credit: Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A1, reproduced with permission © ESO.

5. – Main implications for cosmology and fundamental physics

Latest measurements of CMB in temperature and polarization from *Planck* satellite [28], complemented at smaller scales by recent ground-based experiments [55, 57-59] and combined with other cosmological information coming from, *e.g.*, SNIa, galaxy and galaxy cluster surveys, have reached high precision in estimating all the parameters that describe the current so-called standard cosmological model. Far from representing a fully, physically exhaustive interpretation of the Universe properties, the ΛCDM model phenomenologically describes reasonably well existing data with a simple set of six parameters.

The first step for the extraction of the cosmological parameters from *Planck* maps is represented by the estimation of the CMB APS, C_ℓ^{XY}. Under the assumption of random Gaussian fluctuations, the C_ℓ^{XY} contain all the relevant statistical information. Particularly important for multifrequency experiments is also the analysis of the correlation between different channels. Under the assumption of a blackbody spectrum, CMB anisotropies are expected to be frequency independent when expressed in terms equivalent thermodynamic temperature (denoted also as CMB temperature), thus correlating different frequency channels is a power diagnostic for potential residual systematic effects or imperfect foreground subtraction. Specific methods have been developed to deal with large and small scales (*i.e.* low and high multipoles) anisotropies (hybrid approach). Low multipoles are better analyzed through formally exact maximum-likelihood approaches implemented in real space that allow to exploit the data full covariance matrix as well as to accurately deal with the effect of sky masking necessary to exclude regions potentially affected by residual foregrounds, while at high multipoles, where the size of the full covariance matrix prevents the use of these methods and the very large number of modes

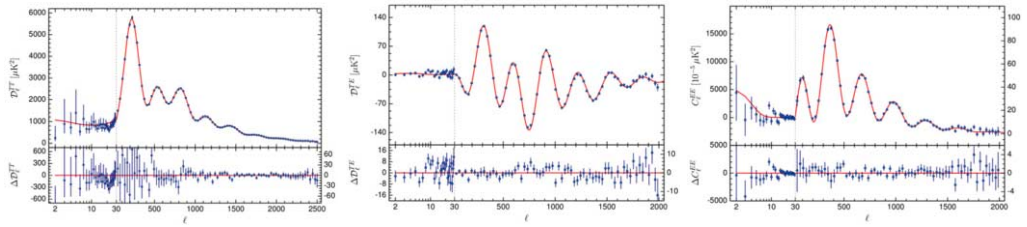


Fig. 6. – *Planck* 2015 CMB spectra, compared with the base ΛCDM fit to *Planck TT* data (red line). The upper panels show the spectra and the lower panels the residuals. In all the panels, the horizontal scale changes from logarithmic to linear at the "hybridization" scale, $\ell = 29$ (the division between the low-ℓ and high-ℓ likelihoods). For the residuals, the vertical axis scale changes as well, as shown by different left and right axes. We show $\mathcal{D}_\ell = \ell(\ell+1)C_\ell/(2\pi)$ for TT and TE, but C_ℓ for EE, which also has different vertical scales at low- and high-ℓ. From [60]. Credit: Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A11, reproduced with permission © ESO.

together with the increasingly Gaussian behaviour of the likelihood function alleviates the need for full covariance matrix inclusion, pseudo-C_ℓ^{XY} estimators, debiased and deconvolved to account for the mask and noise, provide accurate results. Figure 6 shows the *Planck* CMB APS derived in [60].

5˙1. *Cosmological results*. – Markov chain Monte Carlo analyses ingesting the theoretical predictions of Boltzmann codes for CMB APS depending on cosmological model and parameters are then used to estimate cosmological parameters, marginalizing also over instrument and foreground residual parametrizations [28].

Cosmological results based on full-mission *Planck* observations of CMB temperature and polarization anisotropies were presented in the 2015 release. In particular, in the 2015 release the first results of polarization measurements with the LFI at large angular scales were presented. They are in very good agreement with the 2013 analysis of the *Planck* nominal-mission temperature data, but with increased precision.

The temperature and polarization CMB APS are consistent with the standard spatially-flat 6-parameter ΛCDM cosmology with a power-law spectrum of adiabatic scalar perturbations (denoted "base ΛCDM" in this work). For this cosmological model, from the temperature data combined with *Planck* lensing, *Planck* Collaboration found a Hubble constant, $H_0 = (67.8 \pm 0.9)\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$, a matter density parameter $\Omega_\mathrm{m} = 0.308 \pm 0.012$, and a tilted scalar spectral index with $n_\mathrm{s} = 0.968 \pm 0.006$, consistent with the 2013 analysis([4]).

Combined with the *Planck* temperature and lensing data, these measurements give a reionization optical depth of $\tau = 0.066 \pm 0.016$, corresponding to a reionization redshift of $z_\mathrm{re} = 8.8^{+1.7}_{-1.4}$. These results are consistent with those from WMAP polarization

([4]) We report 68% confidence limits on measured parameters and 95% upper limits on other parameters.

Table III. – *Parameters of the base ΛCDM cosmology computed from the* 2015 *baseline Planck likelihoods, illustrating the consistency of parameters determined from the temperature and polarization spectra at high multipoles. Column* [1] *uses the TT spectra at low and high multipoles. Columns* [2] *and* [3] *use only the TE and EE spectra at high multipoles, and only polarization at low multipoles. Column* [4] *uses the full likelihood. The last column lists the relative deviations of the cosmological parameters determined from the Planck TT and Planck full likelihoods.* H_0 *expressed in* $\mathrm{km\,s^{-1}\,Mpc^{-1}}$. *From* [28]. *Credit: Planck Collaboration, Astron. Astrophys.,* **594** (2016) A13, *reproduced with permission* © *ESO.*

Parameter	[1] TT + lowP	[2] TE + lowP	[3] EE + lowP	[4] TT, TE, EE + lowP	$([1]-[4])/\sigma_{[1]}$
$\Omega_{\rm b}h^2$	0.02222 ± 0.00023	0.02228 ± 0.00025	0.0240 ± 0.0013	0.02225 ± 0.00016	-0.1
$\Omega_{\rm c}h^2$	0.1197 ± 0.0022	0.1187 ± 0.0021	$0.1150^{+0.0048}_{-0.0055}$	0.1198 ± 0.0015	0.0
$100\theta_{\rm MC}$ ...	1.04085 ± 0.00047	1.04094 ± 0.00051	1.03988 ± 0.00094	1.04077 ± 0.00032	0.2
τ	0.078 ± 0.019	0.053 ± 0.019	$0.059^{+0.022}_{-0.019}$	0.079 ± 0.017	-0.1
$\ln(10^{10}A_{\rm s})$	3.089 ± 0.036	3.031 ± 0.041	$3.066^{+0.046}_{-0.041}$	3.094 ± 0.034	-0.1
$n_{\rm s}$	0.9655 ± 0.0062	0.965 ± 0.012	0.973 ± 0.016	0.9645 ± 0.0049	0.2
H_0	67.31 ± 0.96	67.73 ± 0.92	70.2 ± 3.0	67.27 ± 0.66	0.0
$\Omega_{\rm m}$	0.315 ± 0.013	0.300 ± 0.012	$0.286^{+0.027}_{-0.038}$	0.3156 ± 0.0091	0.0
σ_8	0.829 ± 0.014	0.802 ± 0.018	0.796 ± 0.024	0.831 ± 0.013	0.0
$10^9 A_{\rm s}e^{-2\tau}$	1.880 ± 0.014	1.865 ± 0.019	1.907 ± 0.027	1.882 ± 0.012	-0.1

measurements cleaned for dust emission using 353 GHz polarization maps from the HFI. The spatial curvature of our Universe is found to be very close to zero, with $|\Omega_K| < 0.005$. The addition of *Planck* polarization data leads to strong constraints on deviations from a purely adiabatic spectrum of fluctuations. No significant evidence for any contribution from isocurvature perturbations or from cosmic defects was found. Combining *Planck* data with other astrophysical data, including SNIa, the equation of state of dark energy is constrained to $w = -1.006 \pm 0.045$, consistent with the expected value for a cosmological constant.

The *Planck* results for base ΛCDM are in good agreement with baryon acoustic oscillation data and with the Joint-Light-curves-Analysis sample of SNIa. However, as in the 2013 analysis, the amplitude of the fluctuation spectrum is found to be higher than that inferred from some analyses of rich cluster counts and weak gravitational lensing. These tensions cannot easily be resolved with simple modifications of the base ΛCDM cosmology. Apart from these tensions, the base ΛCDM cosmology provides an excellent description of the *Planck* CMB observations and many other astrophysical data sets.

Table III summarizes the cosmological parameters found for the base ΛCDM cosmology computed from the 2015 baseline *Planck* likelihoods [28].

Cosmic reionization has been subsequently widely studied combining the *Planck* data in temperature with the low-multipole polarization data, including also HFI channels, to fit ΛCDM models with various parameterizations of the reionization history [61]. A Thomson optical depth $\tau = 0.058 \pm 0.012$ was derived for the commonly adopted in-

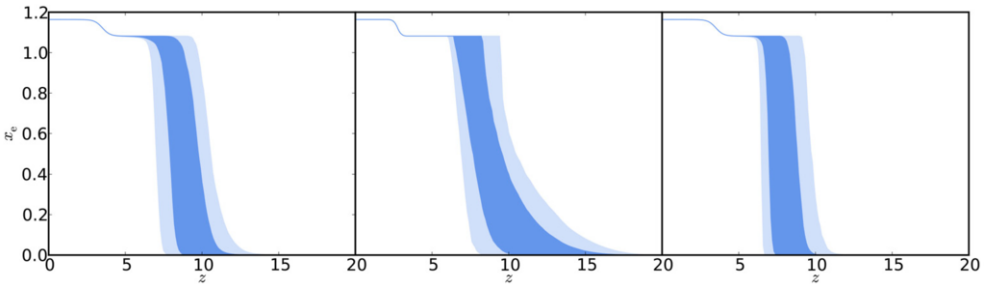


Fig. 7. – Constraints on ionization fraction during reionization. The allowed models, in terms of $z_{\rm re}$ and Δz, translate into an allowed region in $x_{\rm e}(z)$ (68% and 95% in dark blue and light blue, respectively), including the $z_{\rm end} > 6$ prior here. Left: Constraints from CMB data using a redshift-symmetric function ($x_{\rm e}(z)$ as a hyperbolic tangent with $\delta z = 0.5$). Centre: Constraints from CMB data using a redshift-asymmetric parameterization ($x_{\rm e}(z)$ as a power law). Right: Constraints from CMB data using a redshift-symmetric parameterization with additional constraints from the kSZ effect. From [61]. Credit: Planck Collaboration, *Astron. Astrophys.*, **596** (2016) A108, reproduced with permission © ESO.

stantaneous reionization model, in agreement with *Planck* 2015 results combined with other data sets, but reducing the uncertainties. The history of the ionization fraction was reconstructed using either a symmetric or an asymmetric model for the transition between the neutral and ionized phases (see fig. 7). To determine better constraints on the duration of the reionization process, the amplitude of the kinetic SZ (kSZ) effect was included using additional information from the high-resolution ACT and SPT experiments. The derived average redshift at which reionization occurs lies between $z = 7.8$ and 8.8, depending on the model. Using kSZ constraints and a redshift-symmetric reionization model, the upper limit to the width of the reionization period turned to be $\Delta z < 2.8$. The Universe is found to be ionized at less than the 10% level at $z \gtrsim 10$, strongly disfavouring an early onset of reionization. This result also reduces the tension between CMB-based analyses and constraints from other astrophysical sources.

A remarkable example of the success of current cosmology has been obtained from *Planck* data analysing the Integrated Sachs-Wolfe (ISW) effect [62]. Such intrinsically weak predicted signal has been in fact clearly recognized in two classical cosmological probes, CMB anisotropies and galaxy surveys, and in their cross-correlations. The correlation between the gravitational lensing of the CMB anisotropies and the ISW effect gives rise to a secondary CMB bispectrum, clearly detected in *Planck* maps [63]. Subtracting this cosmological secondary non-Gaussian signal and with a careful foreground mitigation performed with different types of estimators, the currently strongest limits on primordial non-Gaussianities have been set from *Planck* maps. The constraints on the amplitudes of local, equilateral, and orthogonal bispectrum as well as of primordial trispectrum in the local configuration and the analysis of trispectrum beyond local configuration [63] are in agreement with ΛCDM model, with cosmological structures sourced by adiabatic, passive, Gaussian, and primordial seed perturbations.

5˙2. *Fundamental physics results.* – Looking at fine signatures in the CMB it is possible to carry out a sort of laboratory tests to constrain particle and fundamental physics.

Planck observations combined with other astrophysical data found no evidence for any departure from base ΛCDM in the neutrino sector of the theory [28]; for example, $N_{\rm eff} = 3.15 \pm 0.23$ for the effective number of relativistic degrees of freedom, consistent with the value $N_{\rm eff} = 3.046$ of the Standard Model of particle physics. The sum of neutrino masses is set to $\sum m_\nu < 0.23\,{\rm eV}$.

The standard big bang nucleosynthesis predictions for the helium and deuterium abundances for the best-fit *Planck* base ΛCDM cosmology are in excellent agreement with observations [28].

Planck data set also stringent constraints on variations in fundamental constants [64], annihilating dark matter and on possible deviations from the standard recombination history [28], and on parity-violating extensions of the standard electromagnetic theory [65], with no evidence for new physics.

5˙3. *Constraints on primordial B-modes.* – The theory of inflation, originally proposed as a solution of the cosmological horizon, smoothness and monopole problems, seems to be a good explanation for the Gaussian nature of primordial perturbations and the density perturbations nearly scale-invariant spectrum. Moreover, it implies that gravitational waves should originate from scalar field perturbations, whose amplitude is expressed in terms of the parameter $r_k = T/S$, *i.e.* the ratio between the amplitudes of primordial tensor and scalar perturbations typically specified at a given scale or wavenumber k.

Adding a tensor component as a single-parameter extension to base ΛCDM, the upper limit on the tensor-to-scalar ratio found from *Planck* 2015 release is $r_{0.002} < 0.11$ [66], consistent with the *Planck* 2013 results and with the B-mode polarization constraints from a joint analysis of BICEP2, *Keck Array*, and *Planck* data discussed below.

BICEP2 and *Keck Array* have observed the same approximately $400\,{\rm deg}^2$ patch of sky centered on RA 0h, Dec. $-57.5°$, together achieving a sensitivity of $\simeq 57\,{\rm nK\,deg}$ in Stokes Q and U parameters in a band centered at 150 GHz. A cross-correlation in B-modes at high significance level was detected considering the 150 GHz and the *Planck* 353 GHz map in the same region. The single- and cross-frequency APS at frequencies $\geq 150\,{\rm GHz}$ were fitted to a lensed-ΛCDM model that includes dust and a possible contribution from inflationary gravitational waves, using *a priori* on the frequency spectral behavior of polarized dust emission from previous *Planck* analysis of other regions of the sky. As previously mentioned, a strong evidence for dust contamination was found with no statistically significant evidence for tensor modes. Exploiting various models including also a synchrotron component, in combination with lower frequency data, and performing an alternative analysis similar to a map-based cleaning of the dust contribution did not change significantly the constraints on r. The final result (see fig. 8), expressed as a likelihood curve for r, yields an upper limit $r_{0.05} < 0.12$ (or $r_{0.002} < 0.09$) at 95% c.l. [55], disfavouring inflationary models with a $V(\phi) \propto \phi^2$ potential.

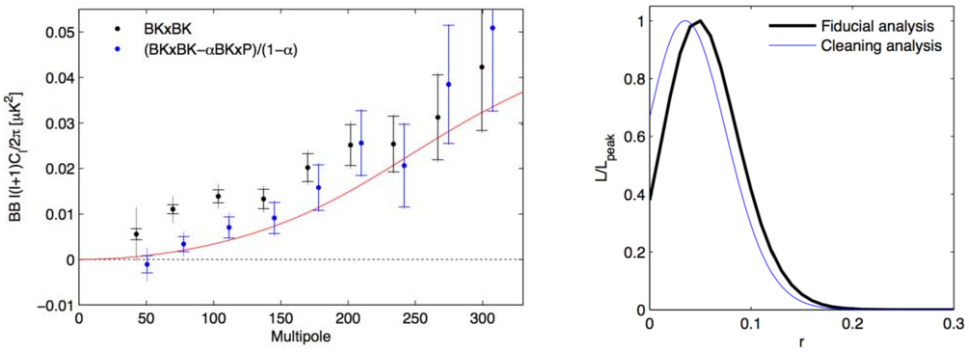


Fig. 8. – Left: *BB* APS of the BICEP2/*Keck Array* maps before and after subtraction of the dust contribution, estimated from the cross-spectrum with *Planck* 353 GHz. The error bars are the standard deviations of simulations, which, in the latter case, have been scaled and combined in the same way. The inner error bars are from lensed-ΛCDM+noise simulations as in the previous plots, while the outer error bars are from the lensed-ΛCDM+noise+dust simulations. The red curve shows the lensed-ΛCDM expectation. Right: constraint on r derived from the cleaned spectrum compared to the fiducial analysis. From [55].

6. – Towards future CMB missions

The energy scale probed by ultra-precise observations of the B-mode polarization of the CMB radiation is $\sim 10^{16}$ GeV, more than 12 orders of magnitude beyond the energy scales accessible to the Large Hadron Collider at CERN. This is also the energy scale where quantum gravity effects start to become relevant. Long-wavelength gravity waves have not yet been detected, despite a vigorous sub-orbital observation programme that has been actively pursued for more than a decade. Cosmic inflation does not provide a unique prediction for the amplitude of the primordial tensor mode, parameterized by r. A precise measurement of, or upper bound on r is essential for constraining inflationary physics, substantially restricting the field of inflationary models allowed by current observations. Because of the uncertainty in the expected value of r, it is not possible to predict which experiment will deliver a first statistically significant detection. However, the comprehensive study of the CMB polarization is a crucial scientific theme that requires observations by a space mission with exquisite sensitivity, control of systematic errors, and broad frequency coverage. A possible detection by a suborbital experiment will require a confirmation from space because of ambiguities in the removal of foreground contamination and would imply to move in the direction of a precise characterization.

To reveal the primordial inflationary B-modes, it is crucial to observe the sky looking for the reionization bump and the recombination epoch, respectively, at $\ell = 2$–10 and $\ell = 20$–100 (angular scales larger than $\simeq 1^\circ$). The main requirements for a B-modes detection mission, to reveal a tensor to scalar ratio $r \sim 10^{-3}$, are an improvement in sensitivity of a factor ~ 100 with respect to *Planck* a control of systematics effects of the order of few tens of nK and an accurate understanding of polarized foregrounds at the level of 1%.

The excellent agreement between the microwave sky emission and the perfect blackbody observed by the COBE/FIRAS instrument is rightfully highlighted as a crucial confirmation of big bang cosmology. However theory predicts that at higher sensitivity this agreement breaks down. Energy injection at any redshift $z \lesssim (\text{few}) \times 10^6$ superimposes spectral distortions relative to a perfect CMB blackbody spectrum. Detecting and characterizing them will represent a powerful way to study dissipation processes at various cosmic times. Remarkably, Thomson scattering of CMB photons from free electrons at the epoch of reionization, producing E-mode polarization on large angular scales, necessarily distorts the unpolarized CMB away from a blackbody.

A secondary, but also important, objective is the collection of high quality full-sky data for a large variety of applications by the astronomical community. They will allow the exploitation of various Galactic and extragalactic science cases. While improved sensitivity and frequency channel number and coverage will allow a better characterization of diffuse emissions, sub-degree (possibly up to $\sim$ few arcmin or better) angular resolution is crucial for producing richer compact source catalogs, both Galactic and extragalactic, and to map the very small scales of diffuse emissions, including galactic magnetic fields, interstellar medium cold dust grain emission and synchrotron emission fine structure. Furthermore, experiments with a fine spectral resolution can assess the study of foreground emission lines and, in principle, going at extreme sensitivity, the chemistry at the recombination epoch, so opening a new window of cosmological investigation.

The above objectives are potentially achievable by different classes of missions proposed to space agencies, described below. The studies carried out to design them largely contributed to define the perspectives of future CMB challenges. Except the Primordial Inflation Explorer (PIXIE), in all the other proposals the spacecraft will observe the sky from an orbit around L2, according to WMAP and *Planck* lessons.

6˙1. *CMB mission proposals at degree resolution.* – In 2009, in response to the ESA Cosmic Vision 2015-2025 call for Proposals, a medium-class space mission has been presented, B-Pol [67], aimed at detecting the primordial gravitational waves occurred during the inflationary epoch on cosmological scales through the detection of the primordial B-mode polarization, thus confirming the quantum origin of the CMB fluctuations. It was designed to host six broad frequency bands detectors, from 45 GHz to 353 GHz, coupled to a set of lens.

PIXIE is an Explorer-class mission submitted to NASA aimed at the CMB mapping in absolute intensity and linear polarization and the diffuse astrophysical foregrounds over the full sky, from 30 GHz to 6 THz (1 cm to 50 μm wavelength), to derive high signal-to-noise ratio polarization maps of the CMB and every Galactic foreground [68]. The four dominant CMB foreground contamination, synchrotron, polarized dust, free-free and electric dipole due to rapidly spinning dust grains emission, are characterized by smooth frequency-dependent functions, and then removed in multi-frequency combined observations. As COBE, the spacecraft will spin from a low Earth orbit (of $\sim$ 660 km altitude, $\sim 3/4$ of that of COBE) but at 4 rpm around the spin axis.

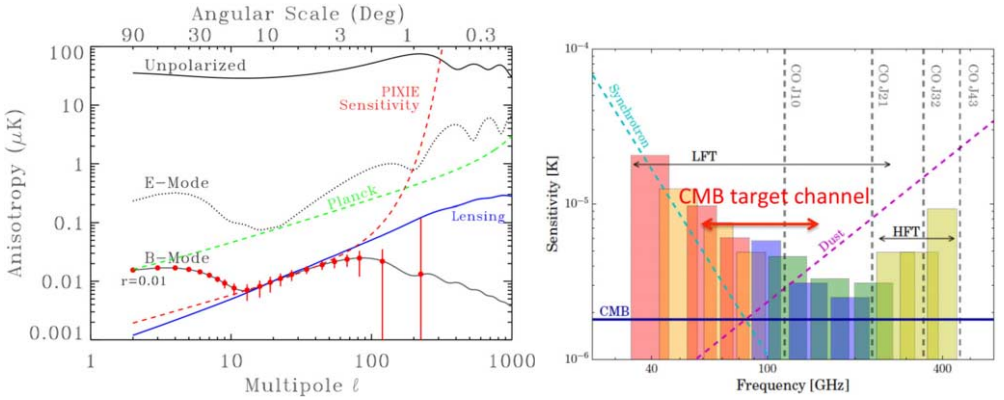


Fig. 9. – Left: CMB APS. The dashed red line shows the PIXIE sensitivity to B-mode polarization. The sensitivity estimate assumes a 4-year mission and includes the effects of foreground subtraction within the cleanest 75% of the sky combining PIXIE data at frequencies $\nu < 600\,\mathrm{GHz}$. Red points and error bars show the response within broader ℓ bins to a B-mode APS with amplitude $r = 0.01$. PIXIE will reach the confusion noise (blue curve) from the gravitational lensing of the E-mode signal by cosmic shear along each line of sight, and has the sensitivity and angular response to measure even the minimum predicted B-mode APS at high statistical confidence. From [68]. © SISSA Medialab Srl. Reproduced by permission of IOP Publishing. All rights reserved. Right: the LiteBIRD band locations centre and the sensitivities for each band. It is also shown the contamination from synchrotron and dust emissions, the CMB signal and the most suitable cosmological frequency channels. We warmly acknowledge Prof. Satoru Uozumi for the kind permission to reproduce this figure. From `http://litebird.jp/wp-content/uploads/2012/03/LNPC_LiteBIRD_uozumi_small.pdf`.

PIXIE uses a polarizing Michelson interferometer to measure the difference spectrum between two orthogonal linear polarizations from two co-aligned beams [69] to detect inflationary epoch polarization with $r < 10^{-3}$. The PIXIE mission concept is an improved version of the FIRAS spectrometer by adding large area detectors, which aims at performing absolute spectroscopy and measuring large-scale CMB B-modes simultaneously. It is designed to be perfectly symmetric in order to reduce its sensitivity to potential systematics since the rotation of a non-circular beam around a non polarized sky induces a spurious polarization signal. Hence, PIXIE distinctive scheme can gather breakthrough sensitivity both in polarization and spectral distortions [70]. Figure 9 (left panel) shows the PIXIE sensitivity compared to the CMB APS.

PIXIE will provide crucial constraints on Universe ionization history and on CIB spectrum and anisotropies, tracing its monopole, dipole and higher order power spectrum to test the matter distribution at $z \sim 3$. Moreover, more stringent constraints on the optical depth parameter, the gas temperature and the Universe reionization at $z \sim 10$ will be performed owing to the cross correlation analysis of the temperature polarization anisotropies.

In February 2015, LiteBIRD, a Lite (Light) Satellite for the studies of B-mode polarization and Inflation from cosmic background Radiation Detection at the extremely early universe, has been formally proposed as a strategic mission to the ISAS, JAXA, and one year later it has started the conceptual design phase-A1 [71, 72]. LiteBird is a highly-targeted, low-cost Japanese B-mode mission concept, in many scientific respects similar to the B-Pol mission. It is designed to detect B-modes at the level of $r \sim 10^{-3}$, a sensitivity that it should be able to achieve assuming that the foregrounds are not too complicated. Its main scientific goal will be the probe of the B-mode range down to one order of magnitude lower than what can be done from the ground. The instrument on board the satellite is equipped with a superconducting detector array cooled to 4 K in order to achieve a target sensitivity of $2\,\mu\mathrm{K} \cdot \mathrm{arcmin}$ over 15 frequency bands from 40 GHz to 400 GHz [71]. Within the assumption $\Delta\nu/\nu \sim 30\%$, each centre band is chosen according to fig. 9 (right panel).

In December 2014, the US team presented the project to the NASA's mission of opportunity to supply the focal plane detectors and the refrigerator system of LiteBIRD, that were partially covered. LiteBIRD is currently in conceptual study phase-A. Other contributions come, or will be likely available, from various European countries.

While the specific B-modes pattern from the primordial gravitational waves is imprinted at angular scales larger than about one degree, gravitational lensing of the CMB by intervening massive structures also generates B-modes, peaking on 10 arcmin angular scales. The major limitation of CMB space missions at degree resolution is represented by their difficulty in accurately subtracting the lensing contribution to B-mode. This problem can be partially solved in combination with ground based super telescopes, as foreseen for CMB-S4 [73]. On the other hand, the atmosphere (with the possible exception of some selected narrow frequency ranges) prevents accurate observations at frequencies above ~ 250 GHz, thus significantly limiting the understanding of Galactic and extragalactic foreground emissions. At small angular scales, this could be a serious problem affecting also an extremely accurate lensing subtraction, crucial at very low values of r.

6˙2. *CMB mission proposals at sub-degree resolution.* – A precise measurement of the lensing effect, required to distinguish between these two sources of B-modes, also provides the opportunity to derive a high-fidelity map of dark matter distribution at times and scales inaccessible otherwise [74, 75], and to address fundamental questions such as the absolute scale of the neutrino masses, possibly better than in laboratory experiments. CMB lensing will also provide crucial inputs concerning the early stages of galaxy formation and their arrangement into large-scale structures. Also, with precise CMB polarization measurements it will be possible the search for primordial non-Gaussianities, with the potential to uncover novel physics. These scientific goals can be achieved with extremely sensitive requirements, full sky measurements, an extremely high level of control of instrumental systematic effects and a broad frequency coverage that only a space mission such with a resolution of (at least) some arcminutes can satisfy.

In December 2010, a new proposal has been submitted to ESA to trace the polarization

in the microwaves [76], the Cosmic Origin Explorer (COrE). The spacecraft, designed to host 15 frequency bands from 45 GHz to 795 GHz with angular resolution comparable to *Planck* and a sensitivity from 10 to 30 times better than *Planck* is addressed to look for primordial gravitational waves generated during inflation up to $r \sim 10^{-3}$ at more than 3σ level.

Few years later, in May 2013, a large class mission has been proposed to ESA, the Polarized Radiation Imaging and Spectroscopy Mission (PRISM) [77], a full sky, high-sensitivity and -resolution observer, in total intensity and polarization, from the microwaves to the IR. PRISM was conceived as the almost definitive CMB/far-IR space mission, able to cover almost all the topics affordable with a single satellite. PRISM consists of two instruments: i) a polarimetric imager with a 3.5 m usable diameter telescope, cooled to below 10 K to maximally reduce the photon noise due to the thermal emission of the mirrors, to map the intensity and polarization anisotropies on the whole sky in 32 broad frequency bands between 30 GHz (1 cm) and 6 THz (50 microns) with unprecedented sensitivity and with an angular resolution ranging from about 17 arcminutes to about 6 arcseconds; ii) a lower-angular-resolution (1.4°) spectrometer that will compare the sky frequency spectrum to a nearly perfect reference blackbody and measure the absolute sky emission over the same frequency range more than three orders of magnitude better than FIRAS.

Because of its broad frequency coverage and extreme stability, PRISM will be able to detect B-modes at 5σ for $r = 5 \times 10^{-4}$, even under pessimistic assumptions concerning the complexity of the astrophysical foreground emissions that must be reliably removed.

While y-distortions from the re-ionized gas as well as from hot clusters constitute a certain detection, μ-distortions and more general spectral distortions (see fig. 10) have the potential to uncover decaying dark matter and to probe the primordial power spectrum on very small scales that cannot be measured by other means, being contaminated by the nonlinearity of gravitational clustering at late times. Finally, the PRISM spectral resolution will range from about 0.5 GHz to 15 GHz at 1.4° angular resolution, and from $\delta\nu/\nu \simeq 0.025$ to 0.25 at the diffraction limit of a 3.5 m telescope (from $\sim 6''$ to $17'$), making in principle possible to detect the most prominent hydrogen and helium lines at recombination (see [77] and references therein).

In 2015, a medium-class space mission, Cosmic Origins Explorer + (COrE+), was proposed to ESA to perform full sky observations of the polarized microwave and submillimetre sky between 60 and 600 GHz in order to explore the origin of stars and cosmic structures on large scales. The primary science goal of COrE+ is to investigate the physics of the very early Universe, which is the source of the entire cosmic web we observe today.

Starting from the COrE and COrE+ studies, in October 2016 the Cosmic ORigins Explorer (CORE) was submitted [78] in response to a call for future medium-sized space mission proposals for the M5 launch opportunity of ESA's Cosmic Vision programme, with the aim of providing the definitive full-sky maps of the CMB polarization anisotropies at

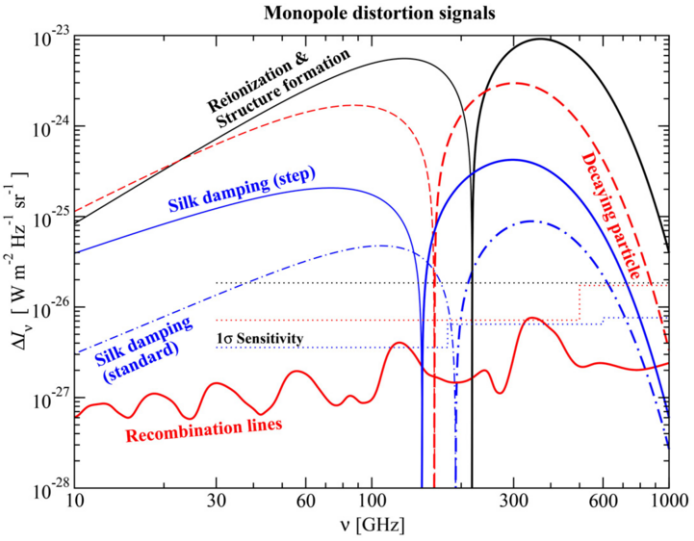


Fig. 10. – Spectral distortions for different scenarios expressed in terms of spectral radiance. Thick lines denote positive, and thinner lines negative signal. The 1σ sensitivities of PRISM for different designs are also indicated. The largest distortion is from the heating of the medium during reionization and structure formation, here for $y \approx 5\times10^{-7}$. The decaying particle scenario is for lifetime $t_X \approx 3.6 \times 10^{-9}$ s and total energy release $\Delta\rho_\gamma/\rho_\gamma \approx 6.3 \times 10^{-7}$. Two cases for the distortion caused by dissipation of small-scale curvature perturbations are shown, one for the standard power spectrum, extrapolated from large-scale CMB measurements all the way to $k \approx$ few $\times\, 10^4\,\mathrm{Mpc}^{-1}$, the other assuming an additional step $\Delta A_\zeta = 10^{-8}$ at $k = 30\,\mathrm{Mpc}^{-1}$ in the power spectrum. The signal caused by recombinations of hydrogen and helium is also, in principle, directly detectable (combining adjacent frequencies). From [77].

large and medium angular scales([5]). CORE [79] is designed to achieve a mean noise level of approximately $2\,\mu\mathrm{K}\cdot$ arcmin (25 times better than the *Planck* mission) with an angular resolution of order 5 arcmin at around 200 GHz, able to allow a significant improvement in cosmological parameter estimation including a variety of model extensions [80] and, in case of a detection of primordial gravitational waves, a determination of the shape of their spectrum over the largest possible range of cosmological scales [81]. The instrument [82] will host about 19 frequency channels, distributed over a broad frequency range span-

([5]) Almost in the same period, the Probe of Inflation and Cosmic Origin (PICO) was proposed to NASA (see `https://zzz.physics.umn.edu/ipsig/_media/missionstudyproposal_science.pdf`). It is currently one of the 8 Probe-Scale – \$ 400 M – \$ 1000 M – space missions whose study is being funded by NASA. The concept and performance of PICO is similar to that of CORE regarding CMB polarization anisotropies, but also the possibility of CMB absolute measurements to carry out spectrum studies with a sensitivity similar or slightly better than that proposed for PIXIE is under investigation.

ning the 60–600 GHz interval, to control astrophysical foreground emission [83]. Indeed, more than 99% of the Galactic emission must be removed from the observed maps, *i.e.* the foreground contribution to the APS of the observations must be modelled with 10^{-4} accuracy. Another key requirement is the measurement accuracy: all systematic effects will be controlled so that no more than approximately 10^{-4} of the intensity leaks into polarization maps, and that no more than about 1% of E-type polarization leaks into B-type modes. The entire sky will be observed repeatedly during four years of continuous scanning, with a combination of three rotations of the spacecraft over different time scales so that each sky pixel is crossed frequently and from many different directions, providing polarization measurements at many different angles. With about 50% of the sky covered every few days, this scan strategy also provides a powerful mitigation of systematic effects to be further corrected [84] in combination with data processing. The feasibility of this approach is already supported by a fairly extensive set of simulations carried out by modifying the *Planck* simulation chain.

Precise inter-frequency calibration foreseen for a mission like CORE will also offer the opportunity to constrain or even detect CMB spectral distortions, particularly from the cosmological reionization epoch, and to significantly refine the determination of CIB spectrum exploiting the frequency dependence of the dipole spectrum [85]. The expected improvement with respect to FIRAS in the recovery of distortion parameters (up to a factor of several hundred for an ideal experiment with the CORE configuration) ranges from a factor of several up to about 50, depending on the quality of foreground removal and relative calibration. Even in the worst case of $\simeq 1\%$ accuracy in both foreground removal and relative calibration at an angular scale of 1°, dipole analyses for a mission like CORE will be able to improve the recovery of the CIB spectrum amplitude by a factor $\simeq 17$ in comparison with current results based on FIRAS. This, in combination with the (only) diffraction-limited resolution up to the highest frequencies implying a considerable decrease of the source confusion and then substantially fainter detection limits, will allow to significantly improve our understanding of galaxy evolution [86], in particular of the high redshift, dust obscured phase.

* * *

We gratefully acknowledge financial support from ASI/INAF agreement n. 2014-024-R.1 for the *Planck* LFI Activity of Phase E2, 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. Some of the results in this paper have been derived using the HEALPix [19] package. We acknowledge the use of the ESA *Planck* Legacy Archive (PLA) and of the Legacy Archive for Microwave Background Data Analysis (LAMBDA, supported by the NASA Office of Space Science). It is a pleasure to thank the many colleagues of the *Planck* and CORE Collaborations for numberless and constructive conversations.

REFERENCES

[1] Planck Collaboration, *Astron. Astrophys.*, **571** (2014) A16.
[2] Danese L. and de Zotti G., *Riv. Nuovo Cimento*, **7** (1977) 277.
[3] Mather J. C., Cheng E. S., Eplee, Jr. R. E. *et al.*, *Astrophys. J.*, **354** (1990) L37.
[4] Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A. and Wright E. L., *Astrophys. J.*, **473** (1996) 576.
[5] Fixsen D. J., *Astrophys. J.*, **707** (2009) 916.
[6] Smoot G. F., Bennett C. L., Kogut A. *et al.*, *Astrophys. J.*, **371** (1991) L1.
[7] Efstathiou G., Sutherland W. J. and Maddox S. J., *Nature*, **348** (1990) 705.
[8] Perlmutter S., Aldering G., Goldhaber G. *et al.*, *Astrophys. J.*, **517** (1999) 565.
[9] de Bernardis P., Ade P. A. R., Bock J. J. *et al.*, *Astrophys. J.*, **564** (2002) 559.
[10] Hu W. and White M., *New Astron.*, **2** (1997) 323.
[11] Tauber J. A., Norgaard-Nielsen H. U., Ade P. A. R. *et al.*, *Astron. Astrophys.*, **520** (2010) A2.
[12] Planck Collaboration, *Astron. Astrophys.*, **536** (2011) A1.
[13] Mandolesi N., Burigana C., Gruppuso A. and Natoli P., *Testing discrete symmetries with the cosmic microwave background: current constraints and Planck forecasts*, *J. Phys. Conf. Ser.*, **335** (2011) 012009.
[14] Chan F. K. and O'Neill E. M., *Computer Sciences Corp Silver Spring Md System Sciences Div, EPRF Tech. Rep.*, **2** (1975) 75.
[15] O'Neill E. M. and Laubscher R. E., *Computer Sciences Corp Silver Spring Md System Sciences Div, NEPRF Tech. Rep.*, **3** (1976) 76.
[16] Tegmark M., *Astrophys. J.*, **470** (1996) L81.
[17] Muciaccia P. F., Natoli P. and Vittorio N., *Astrophys. J.*, **488** (1997) L63.
[18] Górski K. M. *et al.*, *Analysis issues for large CMB data sets*, in *Proceedings of Evolution of Large Scale Structure: From Recombination to Garching*, edited by Banday A. J., Sheth R. K. and da Costa L. N. 1999, p. 37.
[19] Górski K. M., Hivon E., Banday A. J., Wandelt B. D., Hansen F. K., Reinecke M. and Bartelmann M., *Astrophys. J.*, **622** (2005) 759.
[20] Crittenden R. G. and Turok N. G., ArXiv astro-ph e-prints (1998) 9806374.
[21] Doroshkevich A. G., Verkhodanov O. V., Naselsky P. D., Kim J., Novikov D. I., Turchaninov V. I., Novikov I. D., Chiang L.-Y. and Hansen M., *Int. J. Mod. Phys. D*, **20** (2011) 1053.
[22] Béthermin M., Daddi E., Magdis G. *et al.*, *Astrophys. J.*, **757** (2012) L23.
[23] Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A1.
[24] Mennella A., Bersanelli M., Butler R. C. *et al.*, *Astron. Astrophys.*, **536** (2011) A3.
[25] Planck HFI Core Team, Ade P. A. R., Aghanim N., Ansari R. *et al.*, *Astron. Astrophys.*, **536** (2011) A4.
[26] Zacchei A., Maino D., Baccigalupi C. *et al.*, *Astron. Astrophys.*, **536** (2011) A5.
[27] Planck Collaboration, *Astron. Astrophys.*, **536** (2011) A7.
[28] Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A13.
[29] Planck Collaboration, *Astron. Astrophys.*, **596** (2016) A107.
[30] Planck Collaboration, *Astron. Astrophys.*, **594** (2016) A26.
[31] González-Nuevo J., Argüeso F., López-Caniego M., Toffolatti L., Sanz J. L., Vielva P. and Herranz D., *Mon. Not. R. Astron. Soc.*, **369** (2006) 1603.
[32] López-Caniego M., Herranz D., González-Nuevo J., Sanz J. L., Barreiro R. B., Vielva P., Argüeso F. and Toffolatti L., *Mon. Not. R. Astron. Soc.*, **370** (2006) 2047.
[33] Planck Collaboration, *Astron. Astrophys.*, **571** (2014) A28.

[34] BONAVERA L., GONZÁLEZ-NUEVO J., ARGÜESO F. and TOFFOLATTI L., *Mon. Not. R. Astron. Soc.*, **469** (2017) 2401.
[35] BONAVERA L., GONZÁLEZ-NUEVO J., DE MARCO B., ARGÜESO F. and TOFFOLATTI L., *Mon. Not. R. Astron. Soc.*, **472** (2017) 628.
[36] TROMBETTI T., BURIGANA C., DE ZOTTI G., GALLUZZI V. and MASSARDI M., *Astron. Astrophys.*, in press, DOI: 10.1051/0004-6361/201732342, ArXiv e-prints (2017) 1712.08412.
[37] SUNYAEV R. A. and ZELDOVICH Y. B., *Comm. Astrophys. Space Phys.*, **2** (1970) 66.
[38] SUNYAEV R. A. and ZELDOVICH I. B., *Annu. Rev. Astron. Astrophys.*, **18** (1980) 537.
[39] HAEHNELT M. G. and TEGMARK M., *Mon. Not. R. Astron. Soc.*, **279** (1996) 545.
[40] HERRANZ D., SANZ J. L., HOBSON M. P., BARREIRO R. B., DIEGO J. M., MARTÍNEZ-GONZÁLEZ E. and LASENBY A. N., *Mon. Not. R. Astron. Soc.*, **336** (2002) 1057.
[41] MELIN J.-B., BARTLETT J. G. and DELABROUILLE J., *Astron. Astrophys.*, **459** (2006) 341.
[42] CARVALHO P., ROCHA G. and HOBSON M. P., *Mon. Not. R. Astron. Soc.*, **393** (2009) 681.
[43] PLANCK COLLABORATION, *Astron. Astrophys.*, **571** (2014) A30.
[44] TROMBETTI T. and BURIGANA C., *J. Mod. Phys.*, **3** (2012) 1918.
[45] PLANCK COLLABORATION, ADE P. A. R., AGHANIM N., ARNAUD M., ASHDOWN M., AUMONT J., BACCIGALUPI C., BANDAY A. J., BARREIRO R. B., BARRENA R. *et al.*, *Astron. Astrophys.*, **594** (2016) A27.
[46] BURIGANA C., TROMBETTI T., PAOLETTI D., MANDOLESI N. and NATOLI P., *Astrophysical components from Planck maps*, in *Particle Physics at the Year of Light*, edited by STUDENIKIN A. I., *Proceedings of the Seventeenth Lomonosov Conference on Elementary Particle Physics* (2015), p. 320.
[47] PLANCK COLLABORATION, *Astron. Astrophys.*, **571** (2014) A12.
[48] LEACH S. M., CARDOSO J.-F., BACCIGALUPI C. *et al.*, *Astron. Astrophys.*, **491** (2008) 597.
[49] BENNETT C. L., LARSON D., WEILAND J. L. *et al.*, *Astrophys. J. Suppl.*, **208** (2013) 20.
[50] WIELEBINSKI R., *Astron. Astrophys.*, **500** (2009) 245.
[51] PLANCK COLLABORATION, *Astron. Astrophys.*, **554** (2013) A139.
[52] PLANCK COLLABORATION, *Astron. Astrophys.*, **594** (2016) A10.
[53] PLANCK COLLABORATION, *Astron. Astrophys.*, **582** (2015) A28.
[54] PLANCK COLLABORATION, *Astron. Astrophys.*, **586** (2016) A133.
[55] BICEP2/KECK COLLABORATION and PLANCK COLLABORATION, *Phys. Rev. Lett.*, **114** (2015) 101301.
[56] PLANCK COLLABORATION, *Astron. Astrophys.*, **594** (2016) A25.
[57] DAS S., LOUIS T., NOLTA M. R. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2014) 014.
[58] THE POLARBEAR COLLABORATION, *Astrophys. J.*, **794** (2014) 171.
[59] GEORGE E. M., REICHARDT C. L., AIRD K. A. *et al.*, *Astrophys. J.*, **799** (2015) 177.
[60] PLANCK COLLABORATION, *Astron. Astrophys.*, **594** (2016) A11.
[61] PLANCK COLLABORATION, *Astron. Astrophys.*, **596** (2016) A108.
[62] PLANCK COLLABORATION, *Astron. Astrophys.*, **594** (2016) A21.
[63] PLANCK COLLABORATION, *Astron. Astrophys.*, **594** (2016) A17.
[64] PLANCK COLLABORATION, *Astron. Astrophys.*, **580** (2015) A22.
[65] PLANCK COLLABORATION, *Astron. Astrophys.*, **596** (2016) A110.
[66] PLANCK COLLABORATION, *Astron. Astrophys.*, **594** (2016) A20.
[67] DE BERNARDIS P., BUCHER M., BURIGANA C. and PICCIRILLO L., *Exper. Astron.*, **23** (2009) 5.
[68] KOGUT A., FIXSEN D. J., CHUSS D. T. *et al.*, *J. Cosmol. Astropart. Phys.*, **7** (2011) 025.

[69] Kogut A. J., Chuss D. T., Dotson J. L. *et al.*, *The Primordial Inflation Explorer (PIXIE) Mission*, in *Space Telescopes and Instrumentation 2010: Optical, Infrared, and Millimeter Wave*, *Proc. SPIE*, **7731** (2010) 77311S.
[70] Kogut A. J., Chuss D. T., Dotson J. L. *et al.*, *The Primordial Inflation Explorer (PIXIE)*, in *American Astronomical Society Meeting Abstracts*, **223** (2014) 439.01.
[71] Matsumura T., Akiba Y., Borrill J. *et al.*, *J. Low Temp. Phys.*, **176** (2014) 733.
[72] Ishino H., Akiba Y., Arnold K. *et al.*, *LiteBIRD: lite satellite for the study of B-mode polarization and inflation from cosmic microwave background radiation detection*, *Proc. SPIE*, **9904** (2016) 99040X.
[73] Abazajian K. N., Adshead P., Ahmed Z. *et al.*, ArXiv e-prints (2016) 1610.02743.
[74] Planck Collaboration, *Astron. Astrophys.*, **596** (2016) A102.
[75] Challinor A., Allison R., Carron J. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 018.
[76] The COrE Collaboration, ArXiv e-prints (2011) 1102.2181.
[77] André P., Baccigalupi C., Banday A. *et al.*, *J. Cosmol. Astropart. Phys.*, **2** (2014) 006.
[78] Delabrouille J., de Bernardis P., Bouchet F. R. and the CORE Collaboration, A proposal in response to the ESA call for a Medium Size space mission for launch in 2029-2030 (2016).
[79] Delabrouille J., de Bernardis P., Bouchet F. R. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 014.
[80] Di Valentino E., Brinckmann T., Gerbino M. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 017.
[81] Finelli F., Bucher M., Achúcarro A. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 016.
[82] de Bernardis P., Ade P. A. R., Baselmans J. J. A. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 015.
[83] Remazeilles M., Banday A. J., Baccigalupi C. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 023.
[84] Natoli P., Ashdown M., Banerji R. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 022.
[85] Burigana C., Carvalho C. S., Trombetti T. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 021.
[86] De Zotti G., González-Nuevo J., Lopez-Caniego M. *et al.*, *J. Cosmol. Astropart. Phys.*, **4** (2018) 020.

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200012

Future steps in cosmology using spectral distortions of the cosmic microwave background

Jens Chluba(*)

Jodrell Bank Centre for Astrophysics, School of Physics and Astronomy, University of Manchester - Oxford Road, Manchester M13 9PL, UK

Summary. — Since the measurements of COBE/FIRAS in the mid-90's we know that the energy spectrum of the cosmic microwave background (CMB) is extremely close to that of a perfect blackbody at an average temperature $T_0 \simeq 2.726\,\mathrm{K}$. However, a number of early-universe processes are expected to create *CMB spectral distortions* —departures of the average CMB energy spectrum from a blackbody— at a level that is within reach of present-day technology. This provides strong motivation to study the physics of CMB spectral distortions and ask what these small signals might be able to tell us about the Universe we live in. In this lecture, I will give a broad-brush overview of recent theoretical and experimental developments, explaining why future spectroscopic measurements of the CMB will open an unexplored new window to early-universe and particle physics. I will give an introduction about the different types of distortions, how they evolve and thermalize and highlight some of the physical processes that can cause them. I hope to be able to convince you that CMB spectral distortions could open an exciting new path forward in CMB cosmology, which is complementary to planned and ongoing searches for primordial B-mode polarization signals. Spectral distortions should thus be considered very seriously as part of the activities in the next decades.

(*) E-mail: Jens.Chluba@Manchester.ac.uk

1. – Overview and motivation

Cosmology is now a precise scientific discipline, with detailed theoretical models that fit a wealth of very accurate measurements. Of the many cosmological data sets, the cosmic microwave background (CMB) temperature and polarization anisotropies provide the most stringent and robust constraints to theoretical models, allowing us to determine the key parameters of our Universe with unprecedented precision and address fundamental questions about inflation and early-universe physics. Clearly, by looking at the statistics of the CMB anisotropies with different experiments over the past decades we have learned a lot about the Universe we live in, entering the *era of precision cosmology* and establishing the ΛCDM concordance model [1-3].

But the quest continues. Today we are in the position to ask exciting questions about extensions of the standard cosmological model [4-7]. For instance, what do the CMB anisotropies tell us about Big Bang Nucleosynthesis (BBN) and in particular the primordial helium abundance, $Y_{\rm p}$? How many neutrino species are there in our Universe? This question is often addressed through the effective number of relativistic degrees of freedom, $N_{\rm eff}$. What are the neutrino masses and their hierarchy? Are there some decaying or annihilating particles? What about dark radiation? And regarding the initial conditions of our Universe: what is the running of the power spectrum of curvature perturbations? How about the gravitational wave background, parametrized through the tensor-to-scalar ratio, r, which determines the energy scale of inflation, at least when assuming the standard inflation scenario. And to top it up, what about dark energy and the accelerated expansion of our Universe?

All these questions are extremely exciting and define todays cutting-edge research in cosmology, driving present-day theoretical and experimental efforts. The CMB anisotropies in combination with large-scale structure, weak lensing and supernova observations deliver ever more precise answers to these questions [8, 9]. But the CMB holds another, complementary and independent piece of invaluable information: its *frequency spectrum*. Departures of the CMB frequency spectrum from a pure blackbody — commonly referred to as *spectral distortion*— encode information about the thermal history of the early Universe (from when it was a few month old until today). Since the measurements with COBE/FIRAS in the early 90's, the average CMB spectrum is known to be extremely close to a perfect blackbody at a temperature $T_0 = (2.726 \pm 0.001)\,\mathrm{K}$ [10,11] at redshift $z = 0$, with possible distortions limited to one part in 10^5. This impressive measurement was awarded the Nobel Prize in Physics 2006 and already rules out cosmologies with extended periods of significant energy release, disturbing the thermal equilibrium between matter and radiation in the Universe.

1˙1. *Why are spectral distortions so interesting today.* – So far no spectral distortion of the average CMB spectrum was found. Thus, why is it at all interesting to think about spectral distortions now? First of all, there is a long list of processes that could lead to spectral distortions. These include: *reionization* and *structure formation*; *decaying* or *annihilating particles*; *dissipation of primordial density fluctuations*; *cosmic strings*;

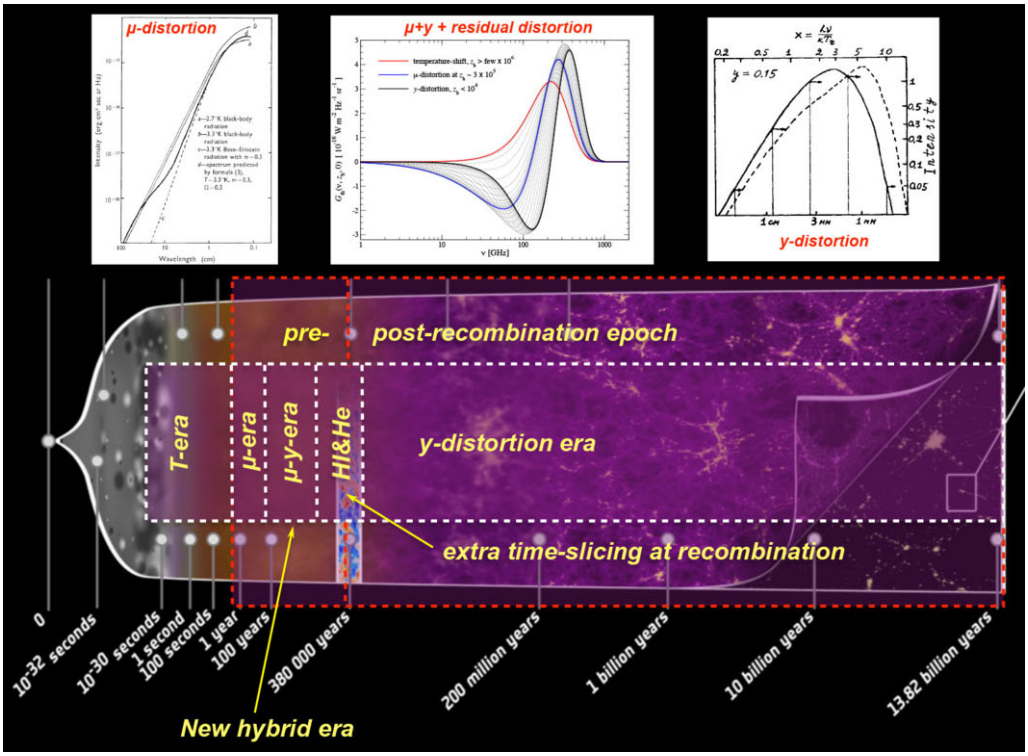


Fig. 1. – CMB spectral distortions probe the thermal history of the Universe at various stages during the pre- and post-recombination era. Energy release at $z \gtrsim$ few $\times 10^6$, when the Universe was only a few month old, merely causes a change of the CMB temperature. A μ-type distortion arises from energy release at $3 \times 10^5 \lesssim z \lesssim$ few $\times 10^6$, while y-type distortions are created at $z \lesssim 10^4$. The signal caused during the μ/y-transition era ($10^4 \lesssim z \lesssim 3 \times 10^5$) is described by a superposition of μ- and y-distortion with some small *residual* distortion that allows probing the time-dependence of the energy-release mechanism. In the recombination era ($10^3 \lesssim z \lesssim 10^4$), additional spectral features appear due to atomic transitions of hydrogen and helium. These could allow us to distinguish pre- from post-recombination y-distortions.

primordial black holes; *small-scale magnetic fields*; *adiabatic cooling of matter*; *cosmological recombination*; and several new physics examples [12-17]. This certainly makes theorists very happy, but most importantly, many of these processes (*e.g.*, reionization and cosmological recombination) are part of our standard cosmological model and therefore should lead to guaranteed signals to search for. This shows that studies of spectral distortions offer both the possibility to *constrain well-known physics* but also to open up a *discovery space* for non-standard physics, potentially adding new *time-dependent information* to the picture (fig. 1).

The second reason for spectral distortion being interesting is due to impressive technological advances since COBE. Although measurements of the CMB temperature and

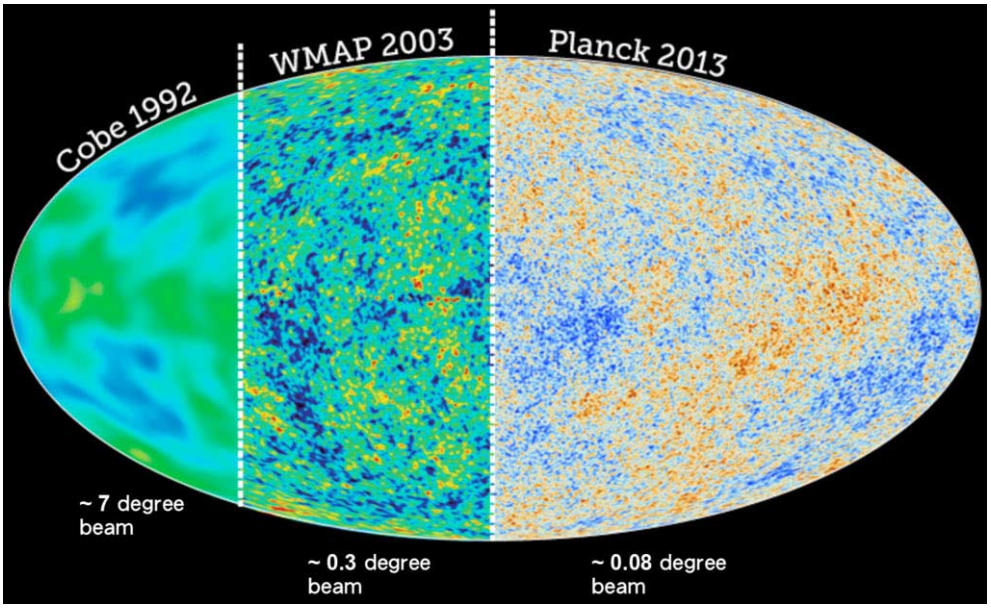


Fig. 2. – Over the past decades, CMB experiments have seen a dramatic improvement in sensitivity and angular resolution, illustrated here with a comparison of COBE, WMAP and PLANCK. In contrast, CMB spectral distortion measurements are still in the state of some 25 years ago, with COBE/FIRAS defining the unchallenged standard.

polarization anisotropies have improved significantly in terms of angular resolution and sensitivity since COBE/DMR, our knowledge of the CMB energy spectrum is still in a similar state as more than 25 years ago (fig. 2). Already in 2002, improvements by a factor of $\simeq 100$ over COBE/FIRAS were deemed feasible [18], and today even more ambitious experimental concepts like PIXIE [19,20] and PRISM [4], possibly reaching $\gtrsim 10^3$ in spectral sensitivity, are being seriously considered. These types of experiments provide a unique way to learn about processes that are otherwise hidden from us. At this stage, CMB spectral distortion measurements at high frequencies are furthermore only possible from space, so that, in contrast to B-mode polarization science, competition from the ground is largely excluded, making CMB spectral distortions a unique target for future CMB space missions [21]. These efforts could be complemented from the ground at low frequencies ($\nu \lesssim 10\,\mathrm{GHz}$), targeting the cosmological recombination ripples, as suggested for APSERa [22], or μ and y-distortions using COSMO.

The immense potential of spectral distortions was realized in the NASA 30-year Roadmap study, where improved characterization of the CMB spectrum was declared as one of the future targets [23]. The strong synergy between spectral distortion and B-mode polarization measurements in terms of challenges related to foregrounds and systematic effects further motivate serious consideration of both science cases as part of the future experimental activities.

1˙2. *Overview and goal of the lecture*. – The main goal of the lectures is to convince you that CMB spectral distortion studies provide us with a new and immensely rich probe of early-universe physics, making it an exciting direction of cosmology for the future. These notes are based on extensive lectures on thermalization physics given as part of the CUSO lecture series in 2014, with extended lecture notes available at `www.chluba.de/science`. I will briefly review the physics of CMB spectral distortions, explaining the different types of distortions and how to compute them for different scenarios. I will then highlight different sources of distortions and what we might learn by measuring distortion signals in the future. Particular attention will be payed to the dissipation of small-scale perturbations and decaying particle scenarios, which illustrate the potential of distortion science. I will also briefly talk about the recombination era and the associated distortion signals and then mention a few of the challenges related to CMB foregrounds. This will also emphasize some of the synergies of distortion and B-mode searches.

2. – The physics of CMB spectral distortions

In this section, I briefly review the main ingredients to describe CMB spectral distortions. The pioneering works on this topic are mainly due to Yakov Zeldovich and Rashid Sunyaev in the 60's and 70's [24-27]. These early works were later extended by [28, 29], to include the effect of double Compton emission, and [30, 31], with refined numerical and analytical treatments. Latest considerations of spectral distortion and their science can be found in [12-17] and [32-34] for the recombination radiation.

2˙1. *Simple blackbody relations*. – Before talking about CMB spectral distortions, let us briefly remind ourselves of a few important blackbody relations. We shall denote the blackbody intensity or *Planckian* as, $B_\nu(T)$, where ν is the frequency and T the blackbody temperature. The *Planck* law reads:

$$B_\nu(T) = \frac{2h}{c^2}\frac{\nu^3}{\mathrm{e}^{h\nu/kT}-1} = \frac{2h\,\nu^3}{c^2}\,n_\nu^{\rm bb}(T) = I_0\,\frac{x^3}{\mathrm{e}^x-1}\,, \tag{1}$$

having units $[B_\nu(T)] = \mathrm{ergs\,s^{-1}\,cm^{-2}\,Hz^{-1}\,sr^{-1}} = 10^{17}\,\mathrm{MJy\,sr^{-1}}$. The spectrum of the Sun is approximately represented by this expression (let us be theorists and forget about all the Fraunhofer lines and existence of the atmosphere with all its absorption bands) with a temperature $T_{\rm ph} \simeq 6000\,\mathrm{K}$ (photosphere). Also, we already heard about the CMB blackbody spectrum, which is unbelievably close to a blackbody at a temperature $T_0 = 2.726\,\mathrm{K}$ [10, 11].

In eq. (1), we also indicate the connection of B_ν to the blackbody occupation number, $n_\nu^{\rm bb}(T) = 1/(\mathrm{e}^{h\nu/kT}-1) = 1/(\mathrm{e}^x-1)$, and transformed to the dimensionless frequency, $x = h\nu/kT$ (redshift-independent), introducing $I_0(T) = (2h/c^2)(kT/h)^3 \approx 270\,\mathrm{MJy\,sr^{-1}}(T/2.726\,\mathrm{K})^3$. It is useful to remember that $x = 1$ corresponds to $\nu \approx 56.8\,\mathrm{GHz}$ for the CMB. Also, the maximum of the blackbody spectrum (Wien's displacement law) is located at $\nu_{\rm max} \approx 160\,\mathrm{GHz}[\frac{T}{2.726\,\mathrm{K}}]$ or $x_{\rm max} \approx 2.821$. We furthermore have

the important limiting cases

$$B_\nu(T) \approx \begin{cases} \dfrac{2\nu^2}{c^2}kT, & \text{for } h\nu \ll kT \quad \text{(Rayleigh-Jeans limit)}, \\ \dfrac{2h\nu^3}{c^2}\mathrm{e}^{-h\nu/kT}, & \text{for } h\nu \gg kT \quad \text{(Wien law)}, \end{cases} \tag{2}$$

for the blackbody spectrum. In the Wien part of the spectrum, very few photons are found but their energy is large. The opposite is true in the Rayleigh-Jeans part.

2˙2. *Photon energy and number density*. – For our discussions, the total photon number and energy densities, ρ_γ and N_γ, will be important. These are defined by the integrals, $\rho_\gamma = \int \frac{I_\nu}{c}\mathrm{d}\nu\,\mathrm{d}\Omega$ and $N_\gamma = \int \frac{I_\nu}{c\,h\nu}\mathrm{d}\nu\,\mathrm{d}\Omega$, over all photon energies and directions. Here, I_ν is the photon intensity. For blackbody radiation, this simply gives

$$\begin{aligned} \rho_\gamma^{\rm Pl} &= \frac{2h}{c^3}\int \frac{\nu^3}{\mathrm{e}^x-1}\mathrm{d}\nu\,\mathrm{d}\Omega = \frac{8\pi h}{c^3}\left(\frac{kT}{h}\right)^4\int\frac{x^3\mathrm{d}x}{\mathrm{e}^x-1} = \frac{8\pi^5(kT)^4}{15\,c^3h^3} \\ &= a_R T^4 \approx 5.10\times10^{-7}\,m_{\rm e}c^2\,{\rm cm}^{-3}\left(\frac{T}{2.726\,{\rm K}}\right)^4 \approx 0.26\,{\rm eV\,cm}^{-3}\left(\frac{T}{2.726\,{\rm K}}\right)^4 \end{aligned} \tag{3a}$$

$$\begin{aligned} N_\gamma^{\rm Pl} &= \frac{2}{c^3}\int \frac{\nu^2}{\mathrm{e}^x-1}\mathrm{d}\nu\,\mathrm{d}\Omega = \frac{8\pi}{c^3}\left(\frac{kT}{h}\right)^3\int\frac{x^2\mathrm{d}x}{\mathrm{e}^x-1} = \frac{16\pi\zeta_3(kT)^3}{c^3h^3} \\ &= b_R T^3 \approx 410\,{\rm cm}^{-3}\left(\frac{T}{2.726\,{\rm K}}\right)^3, \end{aligned} \tag{3b}$$

where ζ_i denotes the Riemann ζ-function. Here, $a_R = 4\sigma/c \approx 7.566\times10^{-15}\,{\rm ergs\,cm^{-3}\,K^{-4}}$ is the *radiation constant*, where σ is the *Stefan-Boltzmann constant*. We also have the useful relation $\rho_\gamma^{\rm Pl} \approx 2.701 kT N_\gamma^{\rm Pl}$. In particular, we have $\rho_\gamma^{\rm Pl} \propto T^4$ and $N_\gamma^{\rm Pl} \propto T^3$, the crucial blackbody relations.

2˙3. *What we need to do to change the blackbody temperature*. – The blackbody spectrum is fully characterized by one number, its temperature T. Thus, one simple question is, *what do we have to do to shift the temperature to* $T' \neq T$? Let us suppose we increase the temperature by adding some energy to the photon field (let us say we just move all photons upwards in frequency in some way; no change of the volume or photon number), $\epsilon = \Delta\rho_\gamma/\rho_\gamma^{\rm Pl}(T) \equiv (T'/T)^4 - 1$, then the expected change in the photon temperature is

$$\frac{\Delta T}{T} = (1+\epsilon)^{1/4} - 1 \approx \frac{1}{4}\frac{\Delta\rho_\gamma}{\rho_\gamma^{\rm Pl}}, \tag{4}$$

for small $\Delta\rho_\gamma/\rho_\gamma^{\rm Pl}$. Clearly, if we stopped here, the new spectrum cannot be a blackbody anymore, since we did not change the photon number density. Thus, pure energy release/extraction inevitably leads to a spectral distortion, no matter how the photons are distributed in energy.

To keep the blackbody relation, $N_\gamma^{\rm Pl} \propto T^3$, unchanged we *simultaneously* need to add

$$\frac{\Delta N_\gamma}{N_\gamma^{\rm Pl}} = (T'/T)^3 - 1 = (1+\epsilon)^{3/4} - 1 \approx 3\frac{\Delta T}{T} \quad \Longrightarrow \quad \frac{\Delta N_\gamma}{N_\gamma^{\rm Pl}} \approx \frac{3}{4}\frac{\Delta \rho_\gamma}{\rho_\gamma^{\rm Pl}} \tag{5}$$

of photons to avoid creating a non-blackbody spectrum. This condition is *necessary but not sufficient*, since it does not specify *how* the missing photons are distributed in energy! For example, let us assume we add photons to the blackbody spectrum at one frequency only. Then $\Delta\rho_\gamma = h\nu \Delta N_\gamma$ and $\epsilon \approx (h\nu/2.701kT)\,\Delta N_\gamma/N_\gamma^{\rm Pl}$. To satisfy the condition eq. (5), we just need to tune the frequency to $h\nu/kT \approx (4/3)\,2.70 \approx 3.60$. Clearly, a blackbody spectrum with a single narrow line at $h\nu \simeq 3.6\,kT$ is no longer a blackbody even if eq. (5) is satisfied. We thus also need to add photons to the CMB spectrum in just the right way and the question is how?

To go from one blackbody with temperature T to another at temperature T', we need to have a change of the photon occupation number by

$$\begin{aligned}\Delta n_\nu &= n_{\rm bb}(T') - n_{\rm bb}(T) = \frac{1}{{\rm e}^{x'}-1} - \frac{1}{{\rm e}^{x}-1} = -x\partial_x n_{\rm bb}\frac{\Delta T}{T} + \mathcal{O}\left(\frac{\Delta T}{T}\right)^2 \\ &= \frac{x{\rm e}^x}{({\rm e}^x-1)^2}\frac{\Delta T}{T} + \mathcal{O}\left(\frac{\Delta T}{T}\right)^2\end{aligned}$$

with $x' = x\,T/T'$. In what follows, we will frequently use the definition

$$G(x) = -x\partial_x n_{\rm bb} = \frac{x\,{\rm e}^x}{({\rm e}^x-1)^2} \approx \begin{cases}\dfrac{1}{x}\,, & \text{for } x \ll 1,\\ x\,{\rm e}^{-x}, & \text{for } x \gg 1,\end{cases} \tag{6}$$

which determines the *spectrum of a temperature shift*: $T\partial_T B_\nu \propto x^3 G(x)$, for small $\Delta T/T$. Its spectral shape is shown in fig. 5. It is easy to prove that a change with this spectral distribution does not lead to any distortion as long as $\Delta T/T$ is sufficiently small. We will thus refer to $G(x)$ as the spectrum of a temperature shift. In the thermalization problem, it is created through the combined action of Compton scattering and photon creation processes (*i.e.*, double Compton and Bremsstrahlung emission).

2˙4. *What is the thermalization problem all about.* – When considering the cosmological thermalization problem we are asking: *how was the present CMB spectrum really created?* Assuming that everything starts off with a pure blackbody spectrum, the uniform adiabatic expansion of the Universe alone (absolutely no collisions and spatial perturbations here!) leaves this spectrum unchanged —a blackbody thus remains a blackbody at all times. However, as the simple discussion in the proceeding section already showed, processes leading to photon production/destruction or energy release/extraction should inevitably introduce momentary distortions to the CMB spectrum. Then the big question is: *was there enough time from the creation of the distortion until today to fully restore the blackbody shape, pushing distortions below any observable level?* For this, we

need to redistribute photons in energy through Compton scattering by free electrons. However, this is not enough to restore the blackbody spectrum. We also need to adjust the number of photons through double Compton and Bremsstrahlung. By understanding the thermalization problem and studying the CMB spectrum in fine detail we can thus learn about different early-universe processes and the thermal history of our Universe. This can open a new window to the early Universe, allowing us to peek behind the last scattering surface which is so important for the formation of the CMB temperature and polarization anisotropies.

2˙5. *General conditions relevant to the thermalization problem.* – In the early Universe, photons undergo many interactions with the other particles. We shall mainly concern ourselves with the *average* CMB spectrum and neglect distortion anisotropies when describing their evolution([1]). Distortion anisotropies can be created through anisotropic energy release processes; however, these are usually very small, such that we only briefly touch on them below. We also assume that the distortions are always minor in amplitude, so that the problem can be linearized. This allows us to resort to a Green's function approach when solving the thermalization problem [38,39], which greatly simplifies explicit thermalization calculations for different energy release scenarios as can be carried out using the full thermalization code `CosmoTherm` [12].

We furthermore assume the standard ΛCDM background cosmology [9] with standard ionization history computed using `CosmoRec` [40]. Also, the electron and baryon distribution functions are given by Maxwellians at a common temperature, $T_{\rm e}$, down to very low redshifts ($z \lesssim 10$), when thermalization process is already extremely inefficient. We furthermore need not worry about the evolution of distortions before the electron-positron annihilation era ($z \gtrsim 10^7$–10^8), since in this regime rapid thermalization processes always ensure that the CMB spectrum is very close to that of a blackbody. We are thus just dealing with non-relativistic electrons, protons and helium nuclei immersed in a bath of CMB photons. We can also neglect the traces of other light elements for the thermalization problem and usually assume that neutrinos and dark matter are only important for determining the expansion rate of the Universe.

2˙6. *Photon Boltzmann equation for average spectrum.* – The study of the formation and evolution of CMB fluctuations in both real and frequency space begins with the radiative transport, or *Boltzmann equation* for the photon phase space distribution, $n(x^\mu, p^\mu)$. Here, we are only interested in the evolution of the average spectrum. In this case, perturbations can be neglected, such that $n(x^\mu, p^\mu) \rightarrow n(t, p)$ and we may express the photon Boltzmann equation as

$$\frac{\partial n}{\partial t} - H\, p \frac{\partial n}{\partial p} = \mathrm{C}[n], \tag{7}$$

([1]) Some distortion anisotropies are created by SZ clusters [35]. Primordial distortion anisotropies can also be created by anisotropic acoustic heating [36,37].

omitting any spatial dependence. Here, $H(t)$ is the standard Hubble expansion rate and $\mathrm{C}[n]$ denotes the collision term, which accounts for interactions of photons with the other species in the Universe. The collision term incorporates several important effects. Most importantly, Compton scattering couples photons and electrons, keeping the two in close thermal contact until low redshifts, $z \lesssim 100$–200. Bremsstrahlung and double Compton emission allow adjusting the photon number and are especially fast at low frequencies, as we explain below.

Neglecting collisions ($\mathrm{C}[n] = 0$), we directly recover $n(t, p) = n[t_0, p\,a(t)/a(t_0)]$, which means that the shape of the photon distribution is conserved by the universal expansion and only the photon momenta are redshifted. Introducing the variable $x = p/kT_\gamma(t) = h\nu/kT_\gamma(t)$, with $T_\gamma(t) = T_\gamma(t_0)\,a(t_0)/a(t) \propto (1+z)$, the photon Boltzmann equation takes the more compact form $\partial n(t,x)/\partial t = \mathrm{C}[n(t,x)]$ (see [12] for more details), which highlights the conservation law.

2˙7. *Collision term for Compton scattering.* – We already mentioned that Compton scattering is responsible for redistributing photons in energy. This problem has been studied a lot in connection with X-rays from compact objects [41,42] and the cosmological context [24,25]. In reality, electron-photon scattering also helps isotropizing the photon field (Thomson scattering limit), although for this energy exchange is not as crucial [43, 44].

To account for the Comptonization of photons by free thermal electrons, we can use the so-call *Kompaneets equation* [45]:

$$\left.\frac{\partial n}{\partial \tau}\right|_{\mathrm{CS}} \approx \frac{\theta_\mathrm{e}}{x_\mathrm{e}^2}\frac{\partial}{\partial x_\mathrm{e}} x_\mathrm{e}^4 \left[\frac{\partial}{\partial x_\mathrm{e}} n + n(1+n)\right] \equiv \frac{\theta_\mathrm{e}}{x^2}\frac{\partial}{\partial x} x^4 \left[\frac{\partial}{\partial x} n + \frac{T_\gamma}{T_\mathrm{e}}\, n(1+n)\right], \tag{8}$$

where $\mathrm{d}\tau = N_\mathrm{e}\sigma_\mathrm{T} c\,\mathrm{d}t$ is the Thomson optical depth, $\theta_\mathrm{e} = kT_\mathrm{e}/m_\mathrm{e}c^2$ is the dimensionless electron temperature and $x_\mathrm{e} = h\nu/kT_\mathrm{e}$ is the chosen frequency variable. This expression can be obtained by computing the Compton collision term in the limit $h\nu \ll kT_\mathrm{e}$ and $kT_\mathrm{e} \ll m_\mathrm{e}c^2$, keeping only terms up to first order in θ_e and $h\nu/m_\mathrm{e}c^2$ [46]. This is equivalent to considering the first two moments of the photons frequency shift, $\Delta\nu/\nu$ over the scattering kernel [47]. The Kompaneets equation can be used to describe the repeated scattering of photons by thermal electrons in the isotropic medium. The first term in the brackets describes *Doppler broadening* and *Doppler boosting* and the last term accounts for the *recoil effect* and *stimulated recoil*. These latter terms are especially important for reaching full equilibrium in the limit of many scatterings.

Below we discuss some analytic solutions of the Kompaneets equation in limiting cases. Here, a couple of words about limitations of this equation. First of all, we assumed that the change in the energy of the photon by the scattering is small. For hot electrons this is no longer correct and one has to go beyond the lowest orders in $\Delta\nu/\nu$. This is for example important for the *Sunyaev-Zeldovich effect* of very hot clusters [48-50], but this procedure only converges *asymptotically* [51, 52]. The second limitation is that if the photon distribution has sharp features (narrower than the width of the scattering kernel) then the shape of the scattered photon distribution is not well represented with

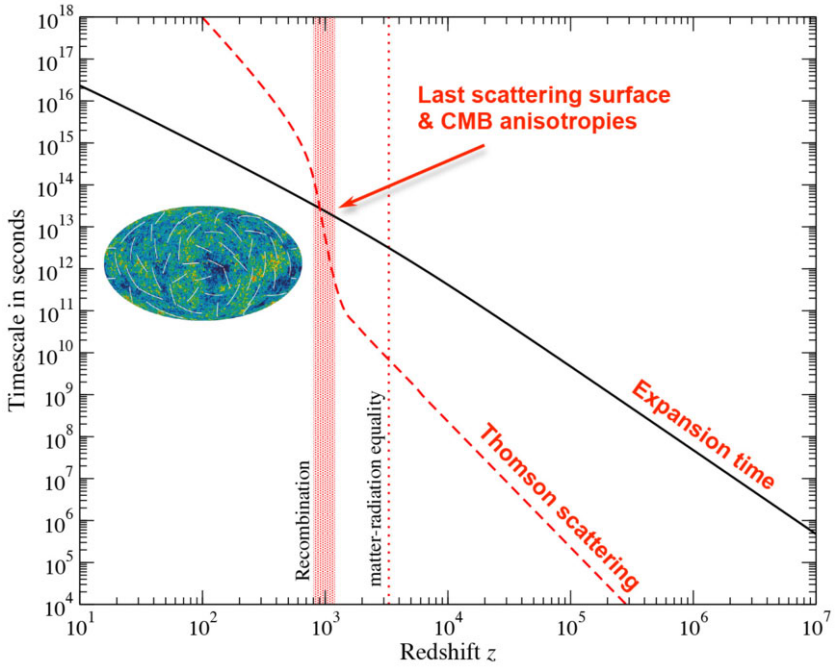


Fig. 3. – Comparison of the Thomson scattering time-scale with the Hubble expansion time-scale.

the diffusion approximation. In this case, a scattering kernel approach can be used to describe the scattering problem [47], although efficient numerical scheme for many scatterings are cumbersome.

2˙7.1. Comptonization efficiency. With the Kompaneets equation, we can already understand some of the important aspects of Comptonization, by simply looking at characteristic time-scales. One important quantity is the Thomson scattering time-scale, $t_{\rm T} = (\sigma_{\rm T} N_{\rm e} c)^{-1}$. It describes how rapidly photons scatter with electrons. For the standard cosmology with 24% of helium (by mass), we have

$$(9)\quad t_{\rm T} = (\sigma_{\rm T} N_{\rm e} c)^{-1} \simeq 2.7\times10^{20}\, X_{\rm e}^{-1}(1+z)^{-3}\,{\rm s} \simeq 4.0\times10^{4}\left[\frac{X_{\rm e}}{0.16}\right]^{-1}\left[\frac{1+z}{1100}\right]^{-3}{\rm years},$$

where $X_{\rm e} = N_{\rm e}/N_{\rm H}$ is the free electron fraction relative to the number of hydrogen nuclei. At $z = 1100$, this corresponds to $\simeq 40\,000$ years between scatterings! To put this into perspective we have to compare with the typical expansion time-scale given by the inverse Hubble rate:

$$(10)\qquad t_{\rm exp} = H^{-1} \simeq \begin{cases} 4.8\times10^{19}\,(1+z)^{-2}\,{\rm s} & \text{(radiation domination)},\\ 8.4\times10^{17}\,(1+z)^{-3/2}\,{\rm s} & \text{(matter domination)}, \end{cases}$$

where the transition between matter and radiation (photons + neutrinos) domination occurs around $z_{\rm eq} \simeq 3400$. From fig. 3 we see that the Thomson scattering rate (shorter time-scales) is much higher than the Hubble expansion rate until after decoupling around $z \simeq 10^3$. But even then, the time-scale for scattering only exceeds the expansion time by a factor of $\simeq 10^2$–10^4. However, this is when the isotropization process of CMB temperature and polarization anisotropies becomes inefficient and we start seeing the primordial CMB fluctuations.

The most important aspect of Comptonization is energy exchange between electrons and photons. The time-scale on which electrons transfer energy to the photons is [24, 53]

$$t_{\rm e\gamma} \approx \frac{t_{\rm T}}{4\theta_{\rm e}} \simeq 4.9\times10^5\, t_{\rm T}\left[\frac{1+z}{1100}\right]^{-1} \simeq 1.2\times10^{29}(1+z)^{-4}\,{\rm s}. \tag{11}$$

In simple words, the time-scale for scattering is $t_{\rm T} \simeq [N_{\rm e}\sigma_{\rm T}c]^{-1}$ and per scattering the fractional energy-exchange between photons and electrons is $\Delta\nu/\nu \simeq 4\theta_{\rm e}$. Comparing $t_{\rm e\gamma}$ with the Hubble rate one finds that at $z_{\mu y} \simeq 5\times10^4$, *Comptonization* becomes inefficient (see fig. 4). At this redshift, the characteristic of spectral distortions changes, transitioning from a so-called μ-distortion to a y-type distortion (see below). Evidently, the transition is not abrupt and the characteristic shape of the distortion changes over a range of redshift between $z \simeq 10^4$–3×10^5 (*e.g.*, see [38]).

The Comptonization time-scale is quite long compared to the time-scale over which electrons are heated by photons. The big difference is that every electron has $\simeq 1.9\times10^9$ photons to scatter with, making the number of interactions much larger. This fact influences many phases in the history of the Universe. For example, the cosmological recombination process is delayed until the temperature of the CMB has dropped below $kT_\gamma \simeq 0.26\,{\rm eV}$, which is two orders of magnitude smaller than the ionization potential, $E_{\rm ion} \simeq 13.4\,{\rm eV}$. Similarly, BBN occurs significantly later than what is expected from naively assuming $kT \simeq mc^2$.

From $\rho_{\rm th} = (3/2)\sum_i N_i kT_{\rm e} = (3/2)N_{\rm H}(1+f_{\rm He}+X_{\rm e})\,kT_{\rm e}$ for the thermal energy of the plasma, by comparison with the energy density of photons, we have

$$t_{\gamma{\rm e}} = \frac{\rho_{\rm th}}{\rho_\gamma}\,t_{\rm e\gamma} \simeq \frac{3N_{\rm H}(1+f_{\rm He}+X_{\rm e})}{8\rho_\gamma/(m_{\rm e}c^2)}\,t_{\rm T} \simeq 0.31\,t_{\rm T}\,(1+z)^{-1} \simeq 7.3\times10^{19}(1+z)^{-4}\,{\rm s}, \tag{12}$$

where for the estimate we used $X_{\rm e} = 1+2f_{\rm He}$ (fully ionized) and $f_{\rm He} \approx Y_{\rm p}/[4(1-Y_{\rm p})] \approx 0.079$. Before recombination, the Compton cooling time is about $\simeq 1.6\times10^9$ times shorter than the Comptonization time. This means that electrons and baryons (through Coulomb scatterings) remain in full thermal contact with the photon field until very late. From fig. 4 one can see that thermal decoupling is expected to happen somewhere around $z \simeq 100$–200 [54]. This is when the earliest signals from the 21 cm era are produced [55].

2˙8. *Bremsstrahlung and double Compton emission.* – So far we have only considered the redistribution of photons in energy. As discussed above, this alone is insufficient for thermalizing the radiation field. In addition, we need to adjust the photon number, which

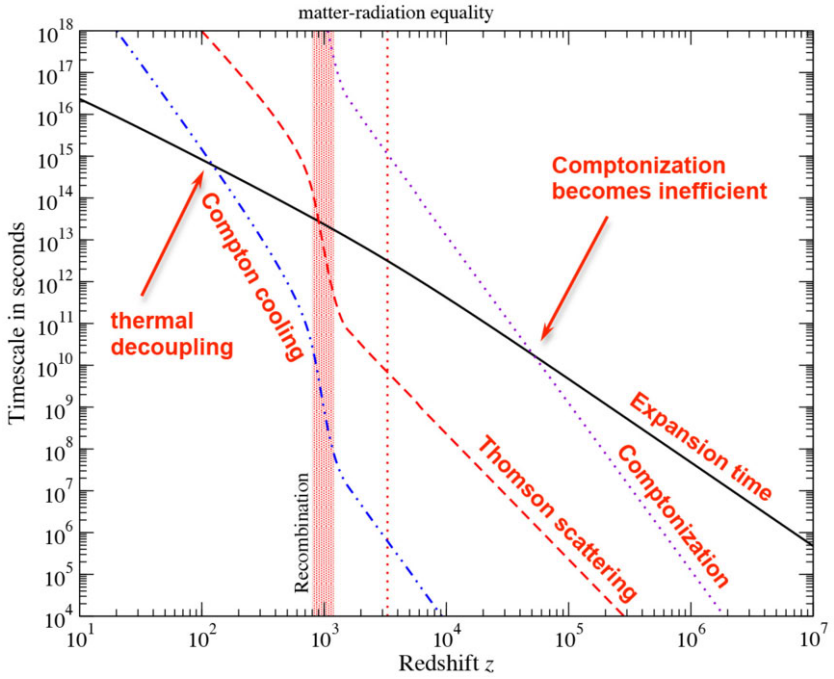


Fig. 4. – Comparison of the Comptonization, Compton cooling and Hubble expansion time-scale.

in the expanding early Universe is achieved by thermal Bremsstrahlung (BR) and double Compton (DC) emission. BR is the first and most obvious suspect for photon production and absorption in the early Universe. However, it turns out that in our Universe DC emission is much more important [29]. Nevertheless, at late times BR has to be included for accurate computations [30, 31, 12].

The collision term for BR and DC emission can be expressed as

$$\text{(13a)} \qquad \left.\frac{\partial n(\tau,x)}{\partial \tau}\right|_{\rm em/abs} = \frac{K_{\rm BR}\,{\rm e}^{-x_{\rm e}} + K_{\rm DC}\,{\rm e}^{-2x}}{x^3}\left[1-n(\tau,x)\,({\rm e}^{x_{\rm e}}-1)\right],$$

$$\text{(13b)} \quad K_{\rm BR}=\frac{\alpha}{2\pi}\frac{\lambda_{\rm e}^3}{\sqrt{6\pi}\theta_{\rm e}^{7/2}}\left(\frac{T_{\rm e}}{T_\gamma}\right)^3\sum_i Z_i^2 N_i\,\bar{g}_{\rm ff}(Z_i,T_{\rm e},T_\gamma,x_{\rm e}),\quad K_{\rm DC}=\frac{4\alpha}{3\pi}\theta_\gamma^2 I_{\rm dc} g_{\rm dc}(T_{\rm e},T_\gamma,x),$$

$$\text{(13c)} \quad \bar{g}_{\rm ff}(x_{\rm e}) \approx \begin{cases} \dfrac{\sqrt{3}}{\pi}\ln\left(\dfrac{2.25}{x_{\rm e}}\right) & \text{for } x_{\rm e}\le 0.37,\\ 1 \text{ otherwise,}\end{cases} \qquad g_{\rm dc} \approx \frac{1+\frac{3}{2}x+\frac{29}{24}x^2+\frac{11}{16}x^3+\frac{5}{12}x^4}{1+19.739\theta_\gamma-5.5797\theta_{\rm e}}.$$

where $\alpha \approx 1/137$, $I_{\rm dc} = \int x^4 n(1+n){\rm d}x \approx 4\pi^4/15$ and $\lambda_{\rm e} = h/m_{\rm e}c \simeq 2.43\times10^{-10}\,{\rm cm}$. The approximation for the DC Gaunt factor, $g_{\rm dc}$, was taken from [12] and is based on [56]. It should provide a very good approximation for our purpose. For the BR Gaunt factors, $\bar{g}_{\rm ff}$, we normally use fits from [57] in numerical calculations or the above approximation for estimates.

One can already see that both BR and DC push the radiation field into equilibrium with a blackbody at the temperature of the electrons, $n_{\rm e}=1/({\rm e}^{x_{\rm e}}-1)$. Also, due to the $1/x^3$ scaling of the emissivity it is clear that BR and DC emission both are most important at low frequencies. Inserting typical numbers for $z\gtrsim 10^3$ and assuming $T_{\rm e}\approx T_\gamma$, we have

$$K_{\rm BR}\simeq 1.4\times10^{-6}\,\left[\frac{\bar{g}_{\rm ff}}{3.0}\right]\left[\frac{\Omega_{\rm b}h^2}{0.022}\right](1+z)^{-1/2}, \tag{14a}$$

$$K_{\rm DC}\simeq 1.7\times10^{-20}\,(1+z)^2. \tag{14b}$$

This implies that at $z_{\rm dc,br}\simeq 3.7\times10^5([\frac{\bar{g}_{\rm ff}}{3.0}][\frac{\Omega_{\rm b}h^2}{0.022}])^{2/5}$ BR and DC emission are similarly important [30, 31]. At $z>z_{\rm dc,br}$, DC emission is more crucial, while at lower redshifts BR dominates.

3. – Types of spectral distortions from energy release

We are now in the position to discuss the main types of spectral distortions created by energy release. In sect. **2**˙7, we learned that around $z_{\mu y}\simeq 5\times10^4$ the Comptonization time-scale (transfer of energy from electrons to photons) becomes longer than the Hubble time. It is clear that this marks an important transition in the efficiency of Compton scattering and redistribution of photons. Let us try to quantify this a little better by looking at the photon evolution equation, for now neglecting photon emission

$$\frac{\partial n}{\partial\tau}\equiv\frac{\theta_{\rm e}}{x^2}\frac{\partial}{\partial x}x^4\left[\frac{\partial}{\partial x}n+\frac{T_\gamma}{T_{\rm e}}\,n(1+n)\right], \tag{15}$$

setting $n=n(\tau,x)$. This equation has no general analytic approximation, but we can solve it for limiting cases. As we explain next, the Compton-y distortion is created by scatterings with inefficient energy exchange between electrons and photons, while the chemical potential μ-distortion is formed in the regime of extremely efficient energy exchange.

3˙1. *Scattering of CMB photons in the limit of small y-parameter.* – Assuming that at $\tau=0$ we start with $n=n_{\rm bb}=1/({\rm e}^x-1)$, then after a very short time $\Delta\tau\ll1$ we find

$$\begin{aligned}\Delta n&\approx\frac{\Delta\tau\theta_{\rm e}}{x^2}\frac{\partial}{\partial x}x^4\left[\frac{\partial}{\partial x}n_{\rm bb}+\frac{T_\gamma}{T_{\rm e}}n_{\rm bb}(1+n_{\rm bb})\right]\approx\frac{\Delta\tau(\theta_\gamma-\theta_{\rm e})}{x^2}\frac{\partial}{\partial x}x^4n_{\rm bb}(1+n_{\rm bb})\\&\approx\Delta\tau\,(\theta_\gamma-\theta_{\rm e})\left[4xn_{\rm bb}(1+n_{\rm bb})-x^2n_{\rm bb}(1+n_{\rm bb})(1+2n_{\rm bb})\right]\\&\approx\Delta\tau\,(\theta_{\rm e}-\theta_\gamma)\,G(x)\,\left[x\,\frac{{\rm e}^x+1}{{\rm e}^x-1}-4\right]\equiv\Delta\tau\,(\theta_{\rm e}-\theta_\gamma)\,Y_{\rm SZ}(x),\end{aligned} \tag{16}$$

where we used $\partial_x n_{\rm bb}=-n_{\rm bb}(1+n_{\rm bb})=-{\rm e}^x/({\rm e}^x-1)^2=-G(x)/x$ and $(1+2n_{\rm bb})=({\rm e}^x+1)/({\rm e}^x-1)=\coth(x/2)$. This is the definition of the so-called *Compton-y* distortion, $Y_{\rm SZ}(x)$, which arises in the limit of scatterings with inefficient energy exchange. This

distortion of the CMB was first studied by [24] and then applied to hot electrons residing inside the potential wells of clusters of galaxies, giving rise to the *thermal Sunyaev-Zeldovich (SZ) effect*. The important variable is the *Compton-y parameter*

$$y = \int_0^\tau \frac{k(T_\mathrm{e}-T_\gamma)}{m_\mathrm{e}c^2}\mathrm{d}\tau' = \int_0^t \frac{k(T_\mathrm{e}-T_\gamma)}{m_\mathrm{e}c^2}\sigma_\mathrm{T}N_\mathrm{e}c\mathrm{d}t', \tag{17}$$

which depends on the number of scattering (related to τ) and the net energy exchange([2]), $\Delta\nu/\nu \simeq 4(\theta_\mathrm{e}-\theta_\gamma) \ll 1$, per scattering. Clearly, for $T_\mathrm{e} \equiv T_\gamma$ one has $y = 0$ and $\Delta n = 0$, no matter how many scattering actually take place! The solution eq. (16) for the distortion is thus valid as long as $|y| \ll 1$. This also ensures that the electron temperature does not change much by the scattering. One possible way to violate this condition even if the number of scattering is tiny ($\tau \ll 1$) is by having a very large difference in the electron and photon temperature. Note, however, that $\theta_\mathrm{e} \ll 1$ is needed since otherwise relativistic corrections to the Compton process appear, which are not accounted for by the Kompaneets equation [58]. For the cosmological thermalization problem, we are always in the situation that the y-parameter is increased beyond unity by increasing the number of scatterings. In this case, Compton scattering pushes electrons and photons into kinetic equilibrium until a μ-distortion is formed (sect. **3**·2).

Assuming that we are in the regime $|y| \ll 1$, there are two cases of interest:

– $y > 0$: energy transferred from the electrons to the photons → *Comptonization*,

– $y < 0$: energy flows from the photons to the electrons → *Compton cooling*.

For most conditions in our Universe, $y > 0$ is relevant, since most processes tend to heat the matter in the Universe. Therefore *negative y-distortions* are usually not being considered, however, the adiabatic cooling of matter in the expanding Universe (in the absence of heating) allows $T_\mathrm{e} < T_\gamma$, so that $y < 0$ does occur [59, 12, 60].

In fig. 5, we illustrate the frequency dependence of the y-distortion for $T_0 = 2.725\,\mathrm{K}$. It has a very characteristic shape, with a deficit of photons in the Rayleigh-Jeans part and an increment of photon in the Wien tail of the CMB spectrum. The limiting behaviors are

$$Y_\mathrm{SZ}(x) = G(x)\left[x\,\frac{\mathrm{e}^x+1}{\mathrm{e}^x-1} - 4\right] \approx \begin{cases} -\dfrac{2}{x}\,, & \text{for } x \ll 1, \\ x(x-4)\mathrm{e}^{-x}, & \text{for } x \gg 1. \end{cases} \tag{18}$$

This corresponds to $\Delta I/I \simeq \Delta T/T \simeq -2y$ for $x \ll 1$ and $\Delta T/T \simeq (x-4)y$ for $x \gg 1$. The y-distortion vanishes close to $\nu \simeq 217\,\mathrm{GHz}$ ($\equiv x \simeq 3.830$), which in principle makes

([2]) To some extent it would be better to immediately write $y^* = \int_0^\tau 4\frac{k(T_\mathrm{e}-T_\gamma)}{m_\mathrm{e}c^2}\mathrm{d}\tau'$, so that $y^* = 4y = \Delta\rho_\gamma/\rho_\gamma$ evidently gives the total amount of energy transfer.

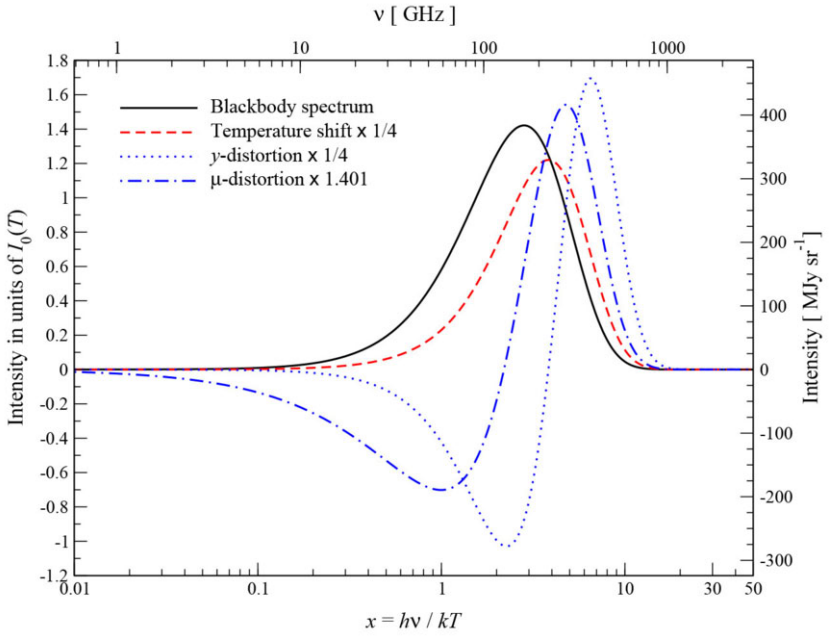


Fig. 5. – Comparison of a Compton y-distortion, $Y_{\rm SZ}(x)$, and μ-distortion, $M(x)$, with the blackbody spectrum and temperature shift, $G(x)$. For convenience, we plot the spectrum as a function of $x = h\nu/kT$ and normalize the left y-axis by $I_0(T) = (2h/c^2)(kT/h)^3 \approx 270\,{\rm MJy\,sr^{-1}}(T/2.726\,{\rm K})^3$. The y-distortion has its crossover frequency around $x \simeq 3.830$ ($\equiv$ 217 GHz), while the μ-distortion has its zero around $x \simeq 2.192$ ($\equiv$ 124 GHz). The upper x-axis and right y-axis also give the corresponding frequency and spectral intensity for $T = 2.726\,{\rm K}$.

it distinguishable from the μ-distortion (sect. **3**˙2). One can easily verify that for a y-distortion $\Delta N_\gamma = 0 \propto \int x^2 Y_{\rm SZ}(x)$ and $\Delta\rho_\gamma = 4y\,\rho_\gamma^{\rm Pl} \propto \int x^3 Y_{\rm SZ}(x){\rm d}x$. Clearly, Compton scattering should not change the number of photons, as reflected by $\Delta N_\gamma = 0$. The second relation means that $4y \equiv \Delta\rho_\gamma/\rho_\gamma^{\rm Pl}$ defines the fractional energy exchange of the electrons with the initial blackbody spectrum. Thus, starting from a pure blackbody, by computing $y = (1/4)\Delta\rho_\gamma/\rho_\gamma^{\rm Pl} \ll 1$ one can directly give a simple approximation for the distortion [24]. In detail, it may be a little more involved to compute $\Delta\rho_\gamma/\rho_\gamma^{\rm Pl}$ for some process, but all one really needs to know is how much energy was pumped into the CMB by energy exchange with the thermal electrons.

3˙1.1. Thermal Sunyaev-Zeldovich effect. Clusters of galaxies are the largest virialized objects in our Universe, with typical masses $M \simeq (10^{13}$–$10^{14})M_\odot$ ($M_\odot \approx 2\times10^{33}\,{\rm g}$) and up to $\simeq 10^3$ galaxies. Cluster also host a hot plasma with free electrons at temperature $T_{\rm e} \simeq$ few $\times\,10^7\,{\rm K}$ ($\equiv$ few $\times$ keV) at typical densities $N_{\rm e} \simeq 10^{-3}\,{\rm cm^{-3}}$. We know this already for a while since clusters show a X-ray glow produced by thermal Bremsstrahlung. The hot electrons can scatter CMB photons and create a Compton-y distortion. The

typical y-parameter of massive clusters is $y \simeq$ few $\times 10^{-5} - 10^{-4}$ with $\theta_{\rm e} \simeq$ few $\times 10^{-2}$ and $\tau \simeq$ few $\times 10^{-3}$. Because for clusters $T_{\rm e} \gg T_\gamma$, the y-parameter reduces to

$$y = \int_0^\tau \frac{kT_{\rm e}}{m_{\rm e}c^2} {\rm d}\tau' \approx \theta_{\rm e}\,\tau \tag{19}$$

and thus directly probes the *integrated electron pressure*, $\bar{P}_{\rm e} \simeq \int N_{\rm e} T_{\rm e}\,{\rm d}l$, through the cluster medium. More than 10^3 clusters have been now seen using the SZ effect [61].

One of the great properties of the thermal SZ effect that is it independent of redshift (ignoring evolutionary effects) [62,63,35]. The reason is that CMB temperature increases $\propto (1+z)$ with redshift, so that the "*light bulb*" illuminating the hot electrons residing inside the cluster becomes brighter the higher the redshift. The cosmological redshift dimming of the signals, which for example reduces the X-ray fluxes for high redshift clusters, is therefore compensated since the CMB itself brightens, and no matter what the redshift of the cluster is it will have the same signal relative to the CMB. The redshift-independence of the SZ signal makes SZ clusters a powerful cosmological probe, since one can in principle track the growth of structures out to high redshifts ($z \simeq$ 1–2) and thus constrain cosmological parameters and the evolution of dark energy [64, 35, 65].

But the thermal SZ effect is even more rich. For a cluster with $kT_{\rm e} = 5\,{\rm keV}$, the thermal velocities of the electrons are $v_{\rm th} \simeq \sqrt{2\theta_{\rm e}}c \simeq 0.14c$. That is quite fast and relativistic corrections become important. In this regime the Kompaneets equation is no longer valid and one has to include higher order corrections [50, 49, 48, 66]. In addition, if the cluster is moving with respect to the CMB, the Doppler kick adds a change in the CMB temperature towards the cluster by $\Delta I \simeq \beta_{\rm c}\tau\, T\partial_T B_\nu(T)$, also knows as *kinematic SZ effect* [62]. This can in principle be used to study large-scale bulk flows in the Universe.

3˙2. *Chemical potential or μ-distortion.* – We now understand that for inefficient energy exchange between electrons and photons (*i.e.*, $y \ll 1$) the shape of the distortion is determined by the y-parameter and has a spectral dependence, $Y_{\rm SZ}(x) = G(x)[x\coth(x/2) - 4]$, shown in fig. 5. Let us now consider the other extreme, when many scatterings are taking place and the redistribution of photons in frequency is very efficient (*i.e.*, $y \gg 1$). In the early Universe, this regime is found at $z \gtrsim 5\times 10^4$ and the distortion is given by a μ-distortion.

3˙2.1. Compton equilibrium solution. When many scatterings occur, the spectrum is driven towards an equilibrium with respect to Compton scattering. Neglecting emission and absorption processes, the kinetic equation thus becomes quasi-stationary

$$0 \approx \frac{\theta_{\rm e}}{x^2}\frac{\partial}{\partial x} x^4 \left[\frac{\partial}{\partial x} n + \frac{T_\gamma}{T_{\rm e}}\, n(1+n)\right]. \tag{20}$$

One solution of this equation is $n_{\rm bb} = 1/({\rm e}^x - 1)$ if $T_{\rm e} \equiv T_\gamma$, since $\partial_x n_{\rm bb} = -n_{\rm bb}(1+n_{\rm bb})$, as it should be for full equilibrium. However, this is not the general solution of the problem. To find a more general solution we have to solve the equation $\partial_x n = -\frac{T_\gamma}{T_{\rm e}}\, n(1+$

n). The factor T_γ/T_e can be absorbed by redefining the frequency scaling $x \to x_\mathrm{e}$ so that this becomes $\partial_{x_\mathrm{e}} n = -n(1+n)$. This can be integrated to $\ln(1+n) - \ln(n) \equiv x_\mathrm{e} + \mathrm{const}$, or

$$n_\mathrm{BE} = \frac{1}{\mathrm{e}^{x_\mathrm{e}+\mu_0} - 1}\,, \tag{21}$$

where we introduced the integration constant μ_0. This is a *Bose-Einstein spectrum* with constant chemical potential([3]) μ_0. Let us pause for a moment. Photons have no rest mass, so the chemical potential should vanish, shouldn't it? This statement is only true if we are in full equilibrium, *i.e.*, we have a blackbody at the temperature of the medium. More generally, for fixed photon number and energy densities the chemical potential can be non-zero.

The chemical potential can in principle be both positive or negative:

– $\mu_0 > 0$: *fewer* photons than a in blackbody at $T_\mathrm{e} \to$ *energy release/photon destruction*,

– $\mu_0 \equiv 0$: blackbody at temperature $T_\mathrm{e} \to$ *full equilibrium*,

– $\mu_0 < 0$: *more* photons than a in blackbody at $T_\mathrm{e} \to$ *energy extraction /photon injection*.

In practice, the solution $\mu_0 < 0$ is unphysical unless μ_0 is actually a function of frequency. The reason is that $x_\mathrm{e} + \mu_0$ can vanish at $x_\mathrm{e} = -\mu_0 > 0$, but this state is never reached or even passed through during the evolution, since instead excess photons would form a Bose-condensate at $x = 0$ with $\mu_0 = 0$ elsewhere [67, 68]. In a real plasma, BR and DC emission will prevent this from happening though [27, 69].

3˙2.2. Definition of the μ-distortion. In the previous section, we found that $n = 1/(\mathrm{e}^{x_\mathrm{e}+\mu_0} - 1)$ is approached for many scatterings in the plasma. But how do we fix the constant μ_0 and what is the definition of the distortion really? Let us assume we start with a blackbody and electrons at temperature $T_\gamma = T_\mathrm{e} = T_i$. Let us change the number and energy densities of the photon field by some $\epsilon_N = \Delta N_\gamma / N_\gamma^\mathrm{Pl}(T_i)$ and $\epsilon_\rho = \Delta\rho_\gamma/\rho_\gamma^\mathrm{Pl}(T_i)$, respectively, and then wait until everything has equilibrated by Compton scattering. This means

$$N_\gamma^\mathrm{BE} = N_\gamma^\mathrm{Pl}(T_i)(1+\epsilon_N) \equiv \frac{N_\gamma^\mathrm{Pl}(T_f)}{G_2^\mathrm{Pl}} \int \frac{x_\mathrm{f}^2 \mathrm{d}x_\mathrm{f}}{\mathrm{e}^{x_\mathrm{f}+\mu_0} - 1}\,, \tag{22a}$$

$$\rho_\gamma^\mathrm{BE} = \rho_\gamma^\mathrm{Pl}(T_i)(1+\epsilon_\rho) \equiv \frac{\rho_\gamma^\mathrm{Pl}(T_f)}{G_3^\mathrm{Pl}} \int \frac{x_\mathrm{f}^3 \mathrm{d}x_\mathrm{f}}{\mathrm{e}^{x_\mathrm{f}+\mu_0} - 1}\,, \tag{22b}$$

where T_f is the final electron temperature in the distorted (Bose-Einstein spectrum) radiation field, $x_\mathrm{f} = h\nu/kT_f$ and $G_2^\mathrm{Pl} \approx 2.404$ and $G_3^\mathrm{Pl} \approx 6.494$. These two equations

([3]) Notice that the sign is different from the normal convention used in thermodynamics.

allow us to fix T_f and μ_0 as a function of the parameters ϵ_N and ϵ_ρ. Assuming that all changes are small we have

$$N_\gamma^{\rm BE} \approx N_\gamma^{\rm Pl}(T_f)\,[1-\mu_0\mathcal{M}_2^{\rm c}] \approx N_\gamma^{\rm Pl}(T_i)\left[1+3\frac{\Delta T}{T_i}-\mu_0\mathcal{M}_2^{\rm c}\right], \tag{23a}$$

$$\rho_\gamma^{\rm BE} \approx \rho_\gamma^{\rm Pl}(T_f)\,[1-\mu_0\mathcal{M}_3^{\rm c}] \approx \rho_\gamma^{\rm Pl}(T_i)\left[1+4\frac{\Delta T}{T_i}-\mu_0\mathcal{M}_3^{\rm c}\right], \tag{23b}$$

where $\mathcal{M}_2^{\rm c}\approx 1.3684$ and $\mathcal{M}_3^{\rm c}\approx 1.1106$. With the conditions eq. (22), we then find [25,53]

$$\mu_0 \approx \frac{3}{\kappa^{\rm c}}\left[\frac{\Delta\rho_\gamma}{\rho_\gamma}-\frac{4}{3}\frac{\Delta N_\gamma}{N_\gamma}\right] \approx 1.401\left[\frac{\Delta\rho_\gamma}{\rho_\gamma}-\frac{4}{3}\frac{\Delta N_\gamma}{N_\gamma}\right], \tag{24a}$$

$$\frac{\Delta T}{T_i} \approx \frac{\mathcal{M}_2^{\rm c}}{\kappa^{\rm c}}\frac{\Delta\rho_\gamma}{\rho_\gamma}-\frac{\mathcal{M}_3^{\rm c}}{\kappa^{\rm c}}\frac{\Delta N_\gamma}{N_\gamma} \approx 0.6389\frac{\Delta\rho_\gamma}{\rho_\gamma}-0.5185\frac{\Delta N_\gamma}{N_\gamma} \approx 0.4561\mu_0+\frac{1}{3}\frac{\Delta N_\gamma}{N_\gamma} \tag{24b}$$

with $\kappa^{\rm c}=4\mathcal{M}_2^{\rm c}-3\mathcal{M}_3^{\rm c}\approx 2.1419$. From eq. (24a) we see that for $\Delta\rho_\gamma/\rho_\gamma \equiv (4/3)\Delta N_\gamma/N_\gamma$ we have no distortion ($\mu_0=0$), as we already understood from the adiabatic condition, eq. (5). In this case, only the temperature of the blackbody is increased([4]) after Compton scattering redistributed all photons, $\Delta T/T_i \approx \frac{1}{3}\Delta N_\gamma/N_\gamma$.

In fig. 6 we illustrate a Bose-Einstein spectrum with $\mu_0=0.5$ and $T_i=T_0=2.726\,{\rm K}$. Only energy was added to the photons but the number of photons was not changed with respect to the initial CMB spectrum. One can see that in the Rayleigh-Jeans tail of the CMB the Bose-Einstein spectrum shows a deficit of photons, while in the Wien tail more photons than in the CMB blackbody spectrum are present. We have $n_{\rm BE}\approx n_{\rm bb}$ at $\nu_\mu\approx 124\,{\rm GHz}$ although for large chemical potential $\nu_\mu\approx 124\,{\rm GHz}\,(1-0.304\,\mu\ln\mu)$ is more accurate [58].

3˙2.3. But how do we define the distortion? To derive the expressions from above, we used

$$n_{\rm BE}=\frac{1}{{\rm e}^{x_{\rm e}+\mu_0}-1}\approx\frac{1}{{\rm e}^{x_{\rm e}}-1}-\frac{G(x_{\rm e})}{x_{\rm e}}\mu_0+\mathcal{O}(\mu_0^2) \tag{25}$$

for $\mu_0\ll 1$. This suggest that $\Delta n=-G(x_{\rm e})\,\mu_0/x_{\rm e}$ could be called the distortion with respect to the blackbody part at temperature $T_{\rm e}$ and in fact this definition has been used frequently. However, since also the final electron temperature, $T_{\rm e}=T_f$, depends on μ_0, this definition does not separate the distortion cleanly. Motivated by the fact that Compton scattering conserves photon number, one natural definition is to fix the μ-distortion such that $\int x^2 M(x){\rm d}x=0$. Integrating Δn gives $\int x^2\Delta n\,{\rm d}x=-2\mu_0\int x\,{\rm d}x/({\rm e}^x-1)=-\mu_0\,\pi^2/3\approx -3.2899\,\mu_0$, so that $M(x)=G(x)[\alpha_\mu-1/x]$ with $\alpha_\mu=\pi^2/18\zeta(3)\approx 0.4561$

([4]) We neglect the small heat capacity of the electrons and baryons.

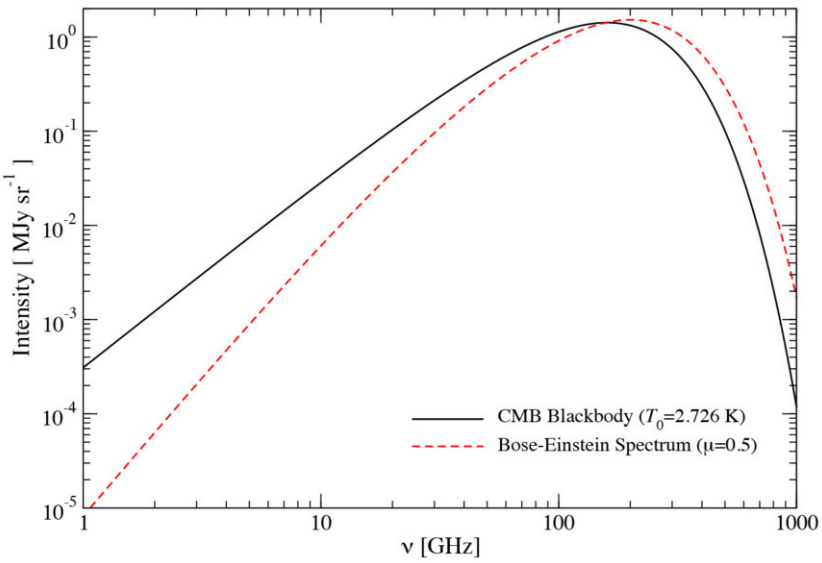


Fig. 6. – Bose-Einstein spectrum for large chemical potential $\mu = 0.5$ and $T_i = T_0 = 2.726\,\mathrm{K}$. Only energy was added to the photon field, but the number of photons was not changed with respect to the initial CMB spectrum. For large chemical potential, the cross over frequency shifts towards higher frequencies according to $\nu_\mu \approx 124\,\mathrm{GHz}\ (1 - 0.304\,\mu \ln \mu) \approx 158\,\mathrm{GHz}$. The figure was taken from [58].

fulfills $\int x^2 M(x)\mathrm{d}x = 0$. If in addition we now normalize the relative change of the photon energy density to unity ($\Delta\rho_M/\rho^{\mathrm{Pl}} = 1$), we obtain the spectral shape of the μ-distortion

$$(26)\quad M^*(x) = \frac{3}{\kappa^{\mathrm{c}}} M(x) \approx 1.401 G(x)\left[0.4561 - \frac{1}{x}\right] \approx \begin{cases} -\dfrac{1.401}{x^2}\,, & \text{for } x \ll 1, \\ 0.6390\, x\, \mathrm{e}^{-x}, & \text{for } x \gg 1, \end{cases}$$

where $3/\kappa^{\mathrm{c}} \approx 1.401$. This implies $\Delta I/I \simeq \Delta T/T \simeq -\mu_0/x$ for $x \ll 1$ and $\Delta T/T \simeq 0.4561\,\mu_0$ at $x \gg 1$. The frequency dependence of $M(x)$ is illustrated in fig. 5 in comparison with the y-distortion and spectrum of a temperature shift. The important feature of a μ-distortion is that it is shifted towards lower frequencies with respect to the y-distortion. This makes it in principle distinguishable and observing a μ-distortion is a clear indication for a signal created in the pre-recombination era, deep into the thermal history of our Universe.

3˙3. *Simple description of primordial distortions.* – We now have all the pieces for a simplest, zeroth-order description of primordial distortions. At late times, ($z \lesssim z_{\mu y} \simeq 5 \times 10^4$), the redistribution of photons by Compton scattering becomes inefficient and a

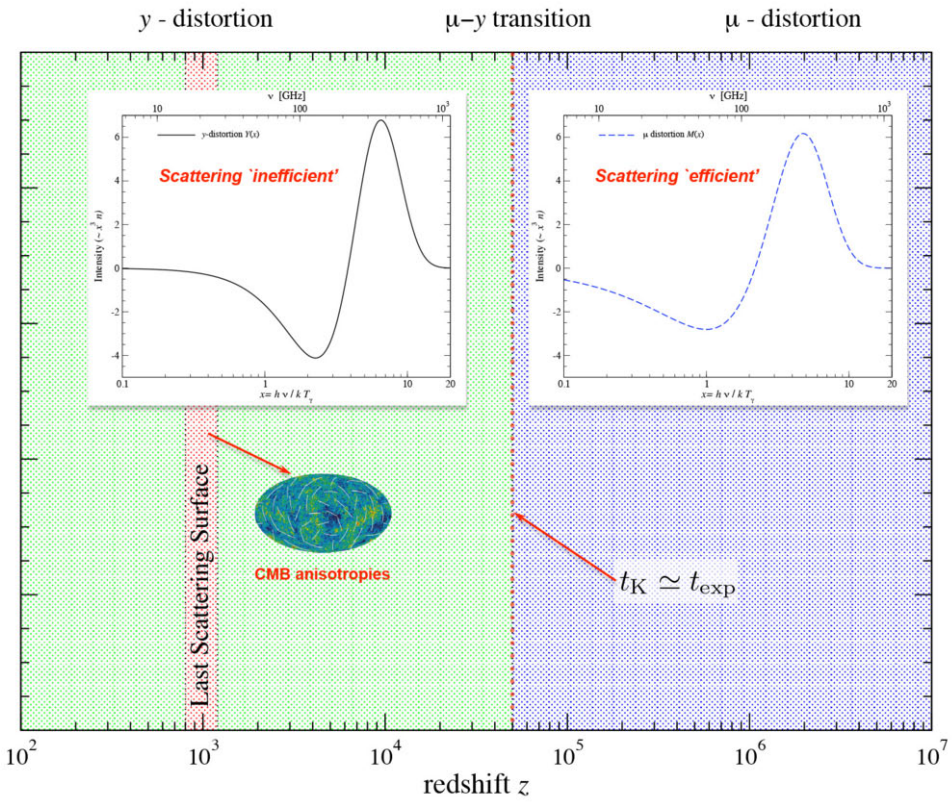


Fig. 7. – Simplest zeroth-order picture for the formation of primordial distortions. At low redshifts ($z \lesssim 5\times10^4$), a y-distortion is formed, while at high redshifts we expect a μ-distortion. At this point we have not included any photon production and we will see that this strongly attenuates the amplitude of the μ-distortion at $z \gtrsim 2\times10^6$.

y-type distortion is formed, in the other extreme we have a μ-distortion (see fig. 7) with the approximations [24, 25]

$$y \approx \frac{1}{4}\frac{\Delta\rho_\gamma}{\rho_\gamma}\bigg|_y, \tag{27a}$$

$$\mu_0 \approx 1.401\left[\frac{\Delta\rho_\gamma}{\rho_\gamma}\bigg|_\mu - \frac{4}{3}\frac{\Delta N_\gamma}{N_\gamma}\bigg|_\mu\right], \tag{27b}$$

such that the total distortion is given by $\Delta n \approx Y_{\rm SZ} y + M(x)\mu_0$. Here, we indicate that to estimate the distortion one needs to consider the *partial energy release* and *photon production* relative to the CMB blackbody in the respective y- and μ-era. If extra photons are injected in the y-era (*e.g.*, by particle decay), the distortion generally is not just a y-distortion, since these extra photons are not redistributed very efficiently, but in the

μ-era they can be ingested and modify the effective chemical potential. Photon injection during the y-era was considered in detail in [39].

Two important aspects are still missing. Firstly, we have not included any thermal-photon production by BR or DC but assumed that only Compton scattering changes the photon field. Photon production will be mostly relevant for the evolution of μ-distortions, implying that not all energy release or photon production is eventually visible as a distortion. That is, at very early phases the *distortion visibility* (see explanation below) is smaller than unity because thermalization reduces the effective amount of energy release that survives as a distortion. This is implicitly hidden in the definition of $\Delta\rho_\gamma/\rho_\gamma|_\mu$ and $\Delta N_\gamma/N_\gamma|_\mu$. We will consider this problem in sect. **3**˙3.1. The second point is that the transition between μ and y distortions is not abrupt at $z \simeq 5\times10^4$ but occurs over a range of redshifts, where in the intermediate regime the distortion is not only given by the superposition of μ and y-distortion. This makes the distortion signal much richer, as pointed out only recently [12, 70, 38]. We will consider this problem in sect. **3**˙4.

3˙3.1. Inclusion of photon production in the μ-era. It is straightforward to approximately include the effect of BR and DC in the μ-distortion era. For a detailed discussion of the approximations and its limitations we refer the interested reader to [58]. Since scattering is efficient, we can assume that the spectrum evolves along a sequence of quasi-stationary stages. However, now we also have to account for emission and absorption, such that

$$0 \approx \frac{\theta_\mathrm{e}}{x^2}\frac{\partial}{\partial x}x^4\left[\frac{\partial}{\partial x}n + \frac{T_\gamma}{T_\mathrm{e}}\,n(1+n)\right] + \frac{K}{x^3}\left[1-n(\mathrm{e}^{x_\mathrm{e}}-1)\right] \tag{28}$$

determines the CMB spectrum. Inserting $n \approx 1/(\mathrm{e}^{x_\mathrm{e}}-1) - \mu(z,x_\mathrm{e})\,G(x_\mathrm{e})/x_\mathrm{e}$ and assuming $x_\mathrm{e} \ll 1$ yields a simple differential equation for $\mu(z,x_\mathrm{e}) \ll 1$, which has the approximate solution

$$\mu(z,x_\mathrm{e}) \approx \mu_0(z)\,\mathrm{e}^{-x_\mathrm{c}(z)/x_\mathrm{e}}. \tag{29}$$

This solution was first derived by [25]. Including both DC and BR, the critical frequency, x_c, which is determined by the competition between photon emission and absorption and Compton up-scattering of photons, is usually $x_\mathrm{c}(z) \simeq 10^{-3}$–$10^{-2}$ during the thermalization period [30, 31].

Equation (29) shows that at $x \gg x_\mathrm{c}$, the chemical potential becomes constant, $\mu(z,x_\mathrm{e}) \approx \mu_0(z)$, while at low frequencies it vanishes exponentially, returning to a blackbody at the temperature of the electrons, with a smooth transition between these regimes around $x \simeq x_\mathrm{c}$. The solution has the expected limiting behavior, even if strictly speaking it is only valid at low frequencies. Indeed, the correct high-frequency behavior is $\mu(z,x) \simeq \mu_0^*(z) + \mathrm{C}(z)\ln x$, where the coefficient, $\mathrm{C}(z)$, is related to the time derivative of the electron temperature [58].

With eq. (29), one can now compute the total photon production rate at any redshift. From that one can estimate how the high-frequency photon chemical potential is affected

by photon production. This essentially boils down to a differential equation for $\mu_0(z)$, which for single energy release $\Delta\rho_\gamma/\rho_\gamma|_i$ has the solution

$$\mu_0(z) \approx 1.401 \frac{\Delta\rho_\gamma}{\rho_\gamma}\bigg|_i \mathrm{e}^{-(z_i/z_\mu)^{5/2}+(z/z_\mu)^{5/2}} = \mu_i\, \mathcal{J}(z_i, z), \tag{30}$$

with $z_{\rm dc} \approx 1.98 \times 10^6$ [30,31]. Here, we defined $\mu_i = 1.401 \Delta\rho_\gamma/\rho_\gamma|_i$. The factor $\mathcal{J}(z_i, z)$ defines the *spectral distortion visibility* between the injection redshift z_i and z (with $z < z_i$). It determines the fraction of energy injected at z_i that is still visible as a distortion at z. For $\mathcal{J}(z_i, z) \simeq 1$, most of the energy is still stored in the distortion, while for $\mathcal{J}(z_i, z) \ll 1$, most of the energy was thermalized and converted into a temperature shift. This implies that after a single energy release event, today's remaining chemical potential is heavily suppressed if the energy injection happens at $z \gtrsim z_{\rm dc} \approx 1.98 \times 10^6$ or some $\simeq 3$ months after the big bang. For continuous energy release in the μ-era, we can then estimate the final distortion measured today using [25, 30, 31]

$$\mu_0 \approx 1.401 \frac{\Delta\rho_\gamma}{\rho_\gamma}\bigg|_\mu \approx 1.401 \int_{z_{\mu y}}^{\infty} \frac{\mathrm{d}(Q/\rho_\gamma)}{\mathrm{d}z} \mathcal{J}(z', 0) \mathrm{d}z'. \tag{31}$$

Here, $\mathrm{d}(Q/\rho_\gamma)/\mathrm{d}z$ describes the energy release relative to the CMB blackbody and depends on the specific energy release mechanism (see sect. **4**). We neglected any extra photon production, but refer to [39] for additional discussion.

3˙3.2. The importance of double Compton emission. In the above, the DC process dominated in the definition of the thermalization redshift [29-31]

$$z_{\rm dc} \approx 1.98 \times 10^6 \left[\frac{\Omega_{\rm b} h^2}{0.022}\right]^{-2/5} \left[\frac{T_0}{2.725\,\mathrm{K}}\right]^{1/5} \left[\frac{(1 - Y_{\rm p}/2)}{0.88}\right]^{-2/5}, \tag{32}$$

assuming $N_{\rm eff} = 3.046$. At $z \gg z_{\rm dc}$, thermalization is very efficient and the distortion visibility drops exponentially. If alternatively we only include BR emission, we find [25, 29, 31]

$$\mathcal{J}_{\rm BR}(z_{\rm h}) = \exp(-[z_{\rm h}/z_{\rm br}]^{1.328}), \tag{33}$$

with $z_{\rm br} \approx 5.27 \times 10^6$. In the classical result, given first by [25], the power-law coefficient is $5/4 = 1.25$ because a different approximation for the BR Gaunt factor was utilized. This shows that the thermalization redshift is significantly higher when only BR is included. In addition, the distortion visibility function drops less steeply at $z \gtrsim 5.27 \times 10^6$.

In fig. 8 we compare the distortion visibility functions for DC and BR only with the full numerical result for the distortion visibility obtained from CosmoTherm [12, 58]. Clearly, DC emission increases the thermalization efficiency significantly. If only BR were taken into account, we would still expect to see some small distortion even from the tail of the electron-positron annihilation era around $z \simeq 2 \times 10^7$! In full detail, this would be

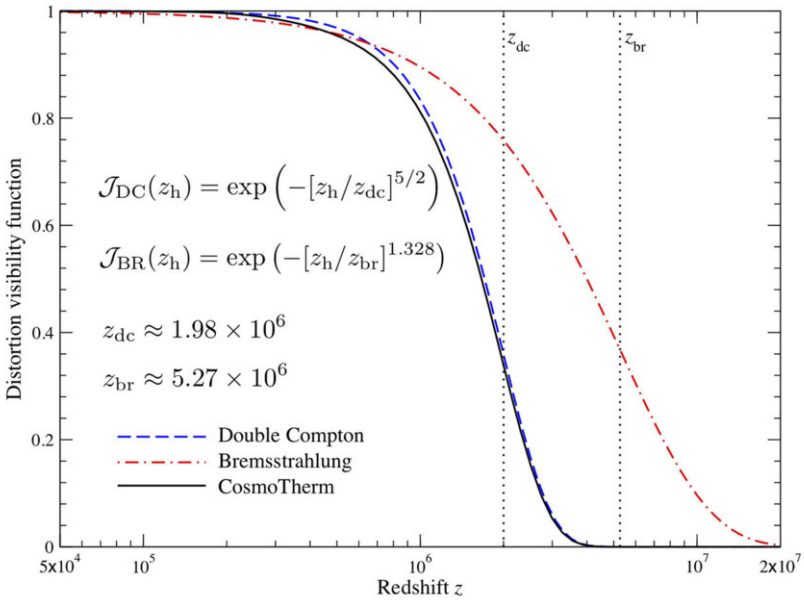


Fig. 8. – Distortion visibility function (adapted from [58]). We compare $\mathcal{J}_{\rm DC}(z_{\rm h})$, $\mathcal{J}_{\rm BR}(z_{\rm h})$ and the numerical result obtained with CosmoTherm. DC emission significantly changes the thermalization efficiency.

quite complicated to compute, but luckily the distortion visibility is exceedingly small, even if only DC is included, providing a rough but tiny upper limit. Comparing with the full numerical result, $\mathcal{J}_{\rm DC}(z_{\rm h}) = \exp(-[z_{\rm h}/z_{\rm dc}]^{5/2})$, provides a very good approximation, which, for simple estimates, is more than sufficient. Improvements to $\mathcal{J}$ can be added analytically [58, 71], but for refined computations it is easier to simply use the Green's function method (*e.g.*, [38, 39]), described in the next section.

3˙4. *Modeling the transition between μ and y.* – In what was presented so far we modeled the transition between μ- and y-era as a simple step-function around $z \simeq z_{\mu y}$. Even in the early studies of the evolution of distortions it was realized that this is not quite correct and that the transition is much more gradual [27, 30, 53]. However, only more recently it was explicitly highlighted that the distortion in the intermediate-regime ($z \simeq 10^4$–3×10^5), contains valuable additional information, allowing us in principle to distinguish different types of distortions [12, 14, 70, 72].

In fig. 9, we illustrate the shape of the distortion caused by a decaying particle scenario at different stages of the evolution. The final distortion is not simply given by the sum of μ and y distortions and thus could allow determining the lifetime of the particle [12]. The description of the distortion in the intermediate regime was later refined by [70] and [38]. For single energy release, the distortion response ($\leftrightarrow$ Green's function) is illustrated in fig. 10. Eliminating the leading-order μ- and y-distortion contributions, one is left with a smaller signal, the so-called *residual distortion* or r-type distortion (see fig. 10),

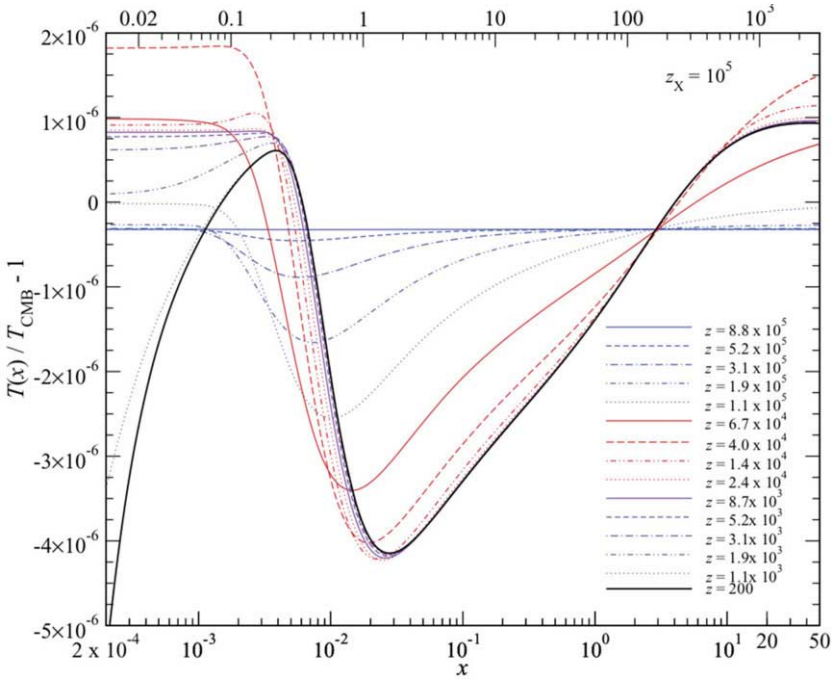


Fig. 9. – Spectral distortion signal (in terms of brightness temperature) caused by a decaying particle scenario at different stages of the evolution (figure taken from [12]). The total energy release was $\Delta\rho_\gamma/\rho_\gamma \simeq 1.3\times10^{-6}$ assuming a particle lifetime $t_X \simeq 2.4\times10^9$ s or $z_X = 10^5$. The final distortion is not described by a simple superposition of μ and y and thus contains valuable time-dependent information teaching us about the lifetime of the partcle.

which can be conveniently parametrized using distortion eigenmodes [72]. The r-type distortion is what contains the extra time-dependent information and detection limits for different energy release scenarios are presented in [72]. However, the thermalization problem is even richer when including the effect of pre-recombination ($z \gtrsim 10^3$) atomic transitions [73, 74]. This might allow us to reach even deeper into the μ- and y-eras by using spectral features of the cosmological recombination radiation [33].

For computational purposes, efficient modeling of the r-type distortion is best handled using the Green's function method [38]. For even more accurate results, the flexible thermalization code `CosmoTherm` [12] can be used, which now runs in $\simeq 30$ s for a given model on a standard laptop. However, we can still improve the analytical description of the μ- and y-distortion contributions using simple representations of the total distortion. From the full Green's function response, one can determine the best-fitting μ- and y-distortion representation. The obtained approximation can then be used to improve the distortion visibility functions in the different regimes. This approach was used in [38]

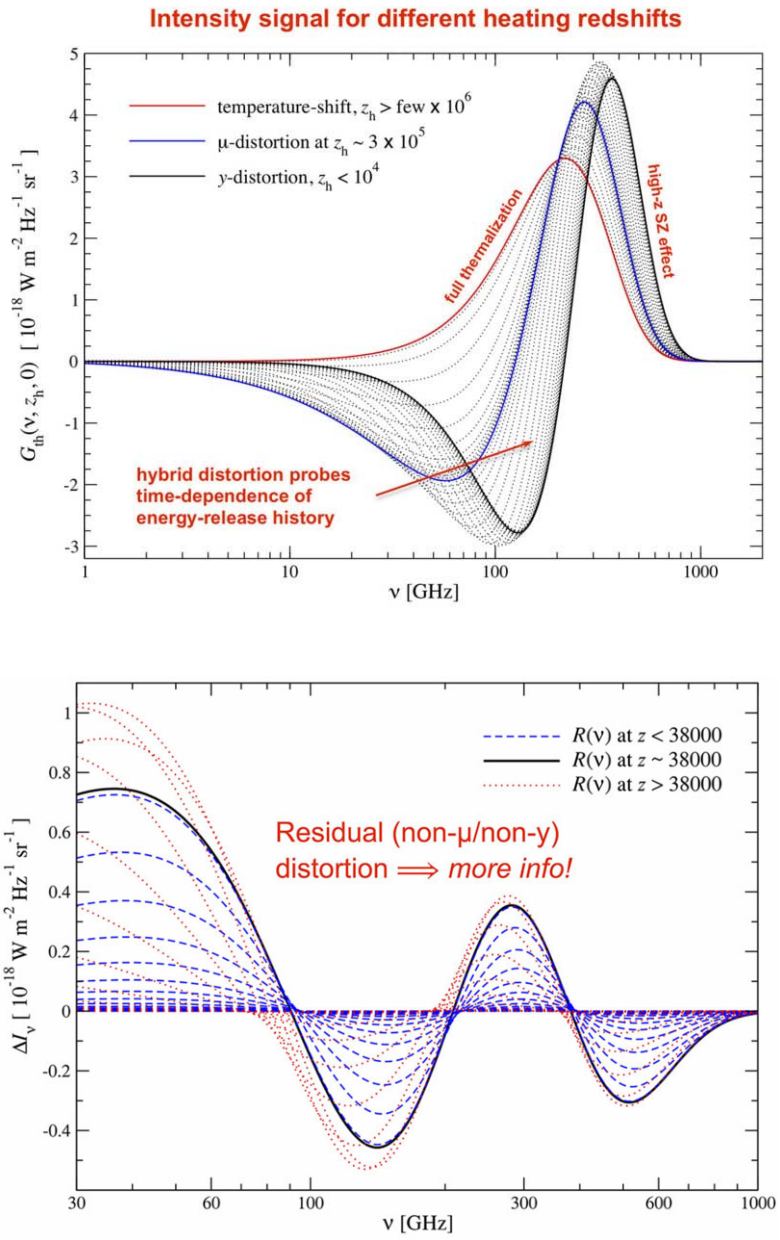


Fig. 10. – Change in the CMB spectrum after a single energy release at different heating redshifts, $z_{\rm h}$. Top panel: thermalization Green's function, $G_{\rm th}(\nu, z_{\rm h})$; lower panel: residual distortion. At $z \gtrsim$ few $\times 10^6$, a temperature shift is created. Around $z \simeq 3 \times 10^5$ a pure μ-distortion appears, while at $z \lesssim 10^4$ a pure y-distortion is formed. At all intermediate stages, the signal is given by a superposition of these extreme cases with a small residual (non-μ/non-y) distortion that contains information about the time-dependence of the energy-release process (figures adapted from [38] and [72]).

and can be summarized using

$$y = \frac{1}{4}\frac{\Delta\rho_\gamma}{\rho_\gamma}\bigg|_y = \frac{1}{4}\int_0^\infty \mathcal{J}_y(z')\,\frac{\mathrm{d}(Q/\rho_\gamma)}{\mathrm{d}z'}\mathrm{d}z', \tag{34a}$$

$$\mu = 1.401\frac{\Delta\rho_\gamma}{\rho_\gamma}\bigg|_\mu = 1.401\int_0^\infty \mathcal{J}_\mu(z')\frac{\mathrm{d}(Q/\rho_\gamma)}{\mathrm{d}z'}\mathrm{d}z', \tag{34b}$$

with the distortion visibilities

$$\mathcal{J}_y(z) \approx \begin{cases} \left(1+\left[\dfrac{1+z}{6\times10^4}\right]^{2.58}\right)^{-1}, & \text{for } z_{\rm rec}\simeq 10^3 \leq z, \\ 0, & \text{otherwise}, \end{cases} \tag{35a}$$

$$\mathcal{J}_\mu(z) \approx \mathcal{J}_{\rm bb}(z)\left[1-\exp\left(-\left[\frac{1+z}{5.8\times10^4}\right]^{1.88}\right)\right]. \tag{35b}$$

These expression should represent the exact fractions of μ and y to $\simeq 10\%$–20% precision. To ensure full energy conservation (no leakage of energy to the r-distortion), instead one can use $\mathcal{J}_\mu(z) \approx [1-\mathcal{J}_y(z)]\,\mathcal{J}_{\rm bb}(z)$. These approximation were presented in [75] and [17]. The distortion visibilities are illustrated in fig. 11.

3˙5. *Distortions from photon injection.* – To finish our discussion of spectral distortion physics, we briefly mention distortions created by photon injection. As shown by [39], these can have a much richer phenomenology than just the simple and broad μ- and y-distortion created by energy release. This is illustrated in fig. 12 for several cases, showing that the final distortion depends on both the injection time and frequency.

In terms of physics, distortions created by photon injection do not directly heat the electrons or baryons. Only once Comptonization becomes relevant do the electrons start heating or cooling. The net effect depends on the injection frequency of the photons. For frequencies $x_i \gtrsim 3.6$–3.8, photons on average loose energy heating the matter. This causes a broad μ- and y-type contribution to the total distortion signal, which for extremely high-frequency injection, $x_i \gtrsim 10$, can dominate. At lower frequencies, cooling of the medium occurs since photons are on average up-scattered. This can create negative μ- and y-type contributions [39].

Photon injection distortions are by no means exotic. For example, the cosmological recombination radiation [33], one of the standard ΛCDM distortions, is created by photon injection. Injection of photons can also occur in decaying or annihilating particle scenarios or evaporation of primordial black holes. In light of recent measurements of EDGES [76] and the ARCADE low-frequency excess [77-79], photon injection distortions of the CMB could become very interesting. This is because these observations potentially point towards a connection with photon injection (or absorption) from decaying or annihilating particles and their low energy by-products in form of non-thermal Bremsstrahlung or synchrotron emission [39, 80-82].

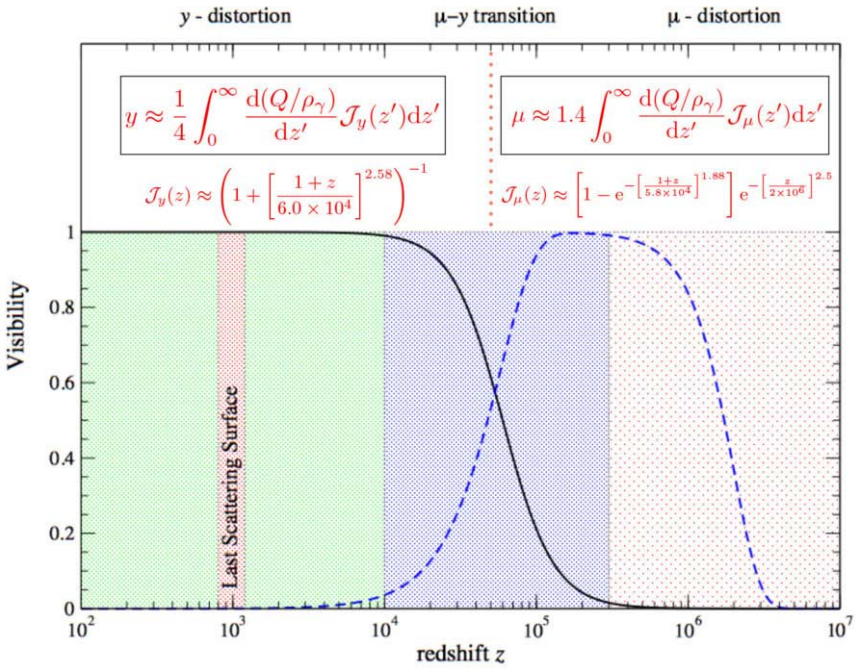


Fig. 11. – Improved picture for the formation of primordial distortions. At low redshifts ($z \lesssim z_{\mu y} \simeq 5\times10^4$), a y-distortion is formed with distortion visibility close to unity, while at high redshifts a μ-distortion appears. The energy release has to be weighted with distortion visibility function which drops exponentially at $z_{\rm dc} \gtrsim 2\times10^6$, leading to a pure temperature shift in that regime from inside the *cosmic photosphere*.

4. – CMB spectral distortion signals from various scenarios

Several exhaustive reviews on various spectral distortion scenarios exist [12-17], covering both standard and non-standard processes. Here we highlight some of the main distortion signals expected within ΛCDM and only briefly mention more exotic sources of distortions. A summary of the relevant ΛCDM distortions is shown in fig. 13. The distortion templates are available at `www.Chluba.de/CosmoTherm`.

4˙1. *Reionization and structure formation.* – The first sources of radiation during reionization [84,85], supernova feedback [86] and structure formation shocks [87-90] heat the intergalactic medium at low redshifts ($z \lesssim 10$), producing hot electrons (in a wide range of temperatures $T_{\rm e} \simeq 10^4\,{\rm K}$–$10^6\,{\rm K}$) that partially up-scatter CMB photons, causing a Compton y-distortion [24]. Although this is the *largest* expected average distortion of the CMB caused within ΛCDM, its amplitude is quite uncertain and depends on the detailed structure and temperature of the medium, as well as scaling relations (*e.g.*, between halo mass and temperature). Several estimates for this contribution were obtained, yielding values for the total y-parameter at the level $y \simeq$ few $\times10^{-6}$ [89,91,83,92,16].

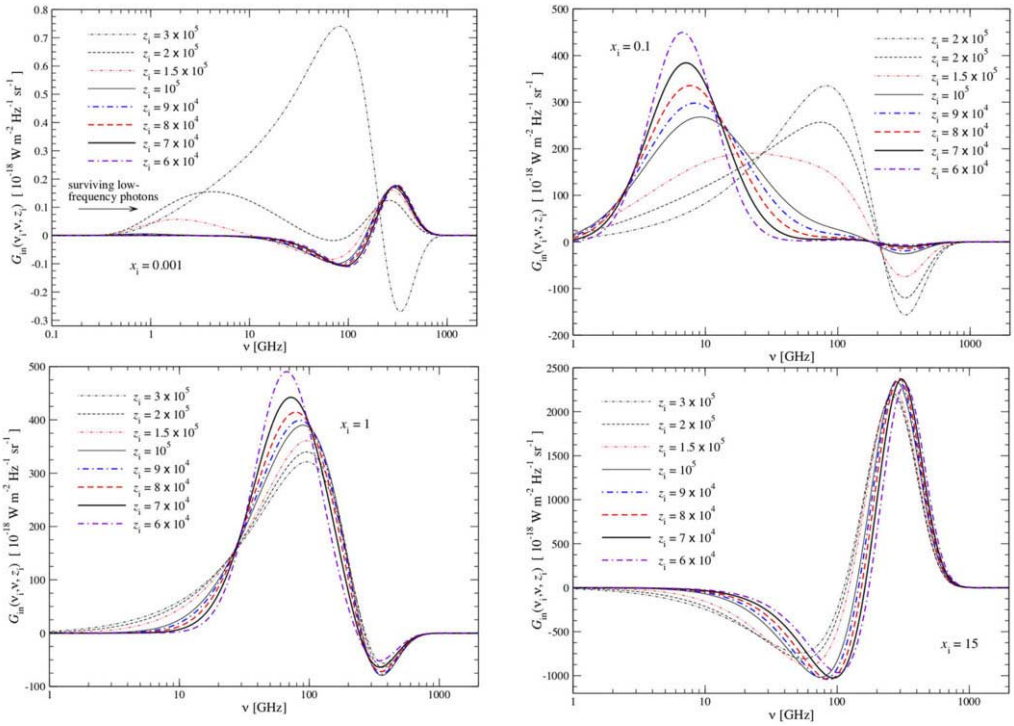


Fig. 12. – Spectral distortions created by photon injection at different frequencies and initial redshifts. The figure is taken from [39].

Following [83], we use a fiducial value of $y = 2{\times}10^{-6}$ (see fig. 13). This is dominated by the low-mass end of the halo function ($M \simeq 10^{13}\, M_\odot$) and the signal should be detectable with a PIXIE-type experiment at more than $10^3\,\sigma$. The detection significance reduces to a few hundred σ when including estimates for the CMB foregrounds [93], but still this provide a sensitive probe of reionization and structure formation physics. Future CMB imagers (*e.g.*, CORE and PICO) furthermore have the potential to separate the spatially varying signature caused by the warm hot intergalactic medium (often referred to as WHIM) and proto-clusters [89, 91], if the challenge of accurate channel intercalibration can be overcome.

Because the signal is so easily detectable, small corrections due to the high gas temperature ($kT_\mathrm{e} \simeq 1\,\mathrm{keV}$) become noticeable [83]. The relativistic correction can be computed using the temperature moment method of `SZpack` [66,52] and differs from the distortions produced in the early Universe (see fig. 14). This correction should be detectable with PIXIE at $\simeq$ 10–20 σ [83, 93] and could teach us about the average temperature of the intergalactic medium, promising a way to solve the missing baryon problem [88]. Both distortion signals are illustrated in fig. 13.

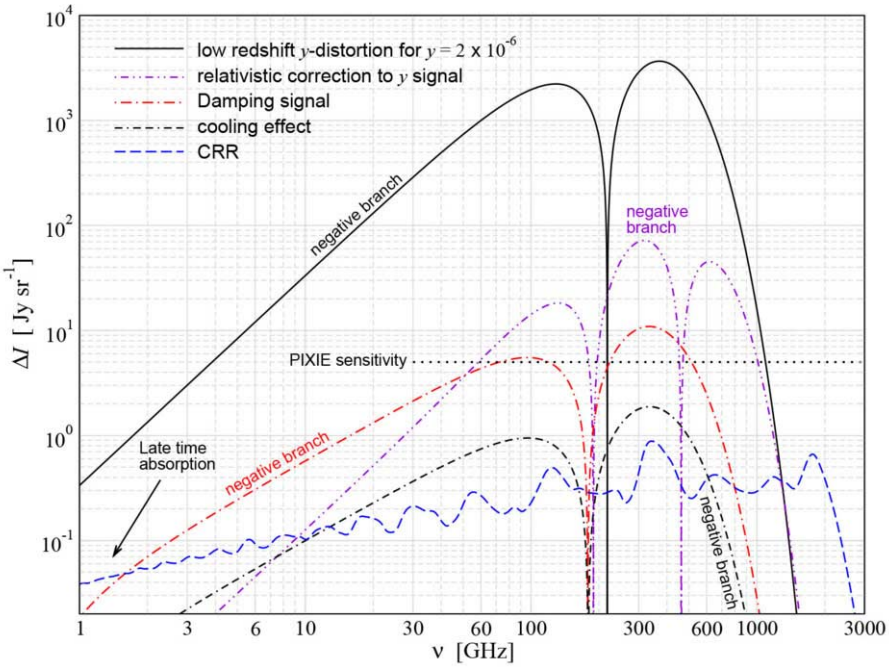


Fig. 13. – Comparison of several CMB monopole distortion signals produced in the standard ΛCDM cosmology. The low-redshift distortion created by reionization and structure formation is close to a pure Compton-y distortion with $y \simeq 2\times10^{-6}$. Contributions from the hot gas in low-mass haloes give rise to a noticeable relativistic temperature correction, which is taken from [83]. The damping and adiabatic cooling signals were explicitly computed using `CosmoTherm` [12]. The cosmological recombination radiation (CRR) was obtained with `CosmoSpec` [34]. The estimated sensitivity ($\Delta I_\nu \approx 5\,\mathrm{Jy/sr}$) of PIXIE is shown for comparison (dotted line). The figure is taken from [17].

4˙2. *Damping of primordial small-scale perturbations.* – The damping of small-scale fluctuations of the CMB temperature set up by inflation at wavelength $\lambda < 1\,\mathrm{Mpc}$ causes another inevitable distortion of the CMB spectrum [95-99]. The idea behind this mechanism is extremely simple and just based on the mixing of blackbodies with varying temperatures through Thomson scattering (see fig. 15). However, the process was only recently described rigorously [43,100], allowing us to perform detailed computations of the associated distortion signal for different early-universe models [43, 101-103, 14, 104, 105]. The distortion is sensitive to the amplitude and shape of the power spectrum at very small scales (wavenumbers $1\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 2\times10^4\,\mathrm{Mpc}^{-1}$ corresponding to multipoles $10^5 \lesssim \ell \lesssim 10^8$) and thus provides a promising new way for constraining inflation while modes are still evolving in the linear regime.

In the early days of CMB cosmology, this effect was already used to derive first upper limits on the spectral index of scalar perturbations, yielding $n_\mathrm{S} \lesssim 1.6$ from

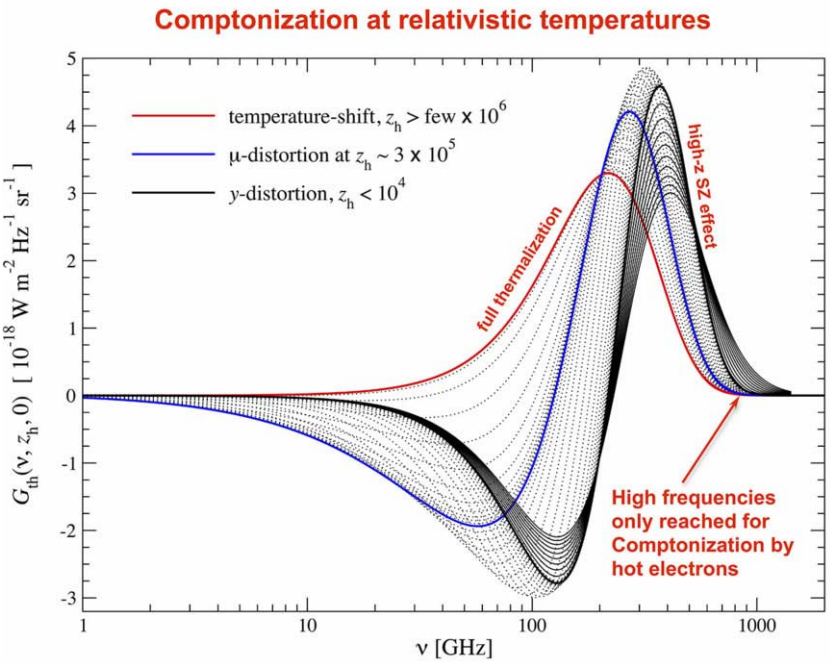


Fig. 14. – Illustration for the effect of relativistic temperature corrections on the distortion signal. In the primordial Universe, electrons hardly reach temperatures $\simeq 1\,\mathrm{keV}$ during the thermalization era ($z \lesssim 10^6$). Therefore even repeated Compton scattering cannot push the distortion signals beyond the standard non-relativistic y-distortion signal. Inside clusters of galaxies, electrons can have temperatures $kT_\mathrm{e} \gtrsim 1\,\mathrm{keV}$. In this case, the distortion signals can extend to much higher frequencies.

COBE/FIRAS [98]. Perturbation modes with $1\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 50\,\mathrm{Mpc}^{-1}$ create y-distortions, while modes with $50\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 2\times10^4\,\mathrm{Mpc}^{-1}$ yield μ-distortions. These scales are hard to access by any other means but spectral distortions provide a new sensitive probe in this regime (fig. 16).

For a given initial power spectrum of perturbations, the effective heating rate in general has to be computed numerically [43]. However, at high redshifts the tight coupling approximation can be used to simplify the calculation. An excellent approximation for the effective heating rate can be obtained using(5) [43, 106]

$$\frac{\mathrm{d}(Q/\rho_\gamma)}{\mathrm{d}z} \approx 4A^2\partial_z k_\mathrm{D}^{-2}\int_{k_\mathrm{min}}^{\infty}\frac{k^4\mathrm{d}k}{2\pi^2}P_\zeta(k)\,\mathrm{e}^{-2k^2/k_\mathrm{D}^2}, \tag{36}$$

where $P_\zeta(k) = 2\pi^2\,A_\mathrm{s}\,k^{-3}\,(k/k_0)^{n_\mathrm{S}-1+\frac{1}{2}\,n_\mathrm{run}\ln(k/k_0)}$ defines the usual curvature power

(5) Here, we define the heating rate such that $\int_z^\infty \frac{\mathrm{d}(Q/\rho_\gamma)}{\mathrm{d}z}\mathrm{d}z > 0$.

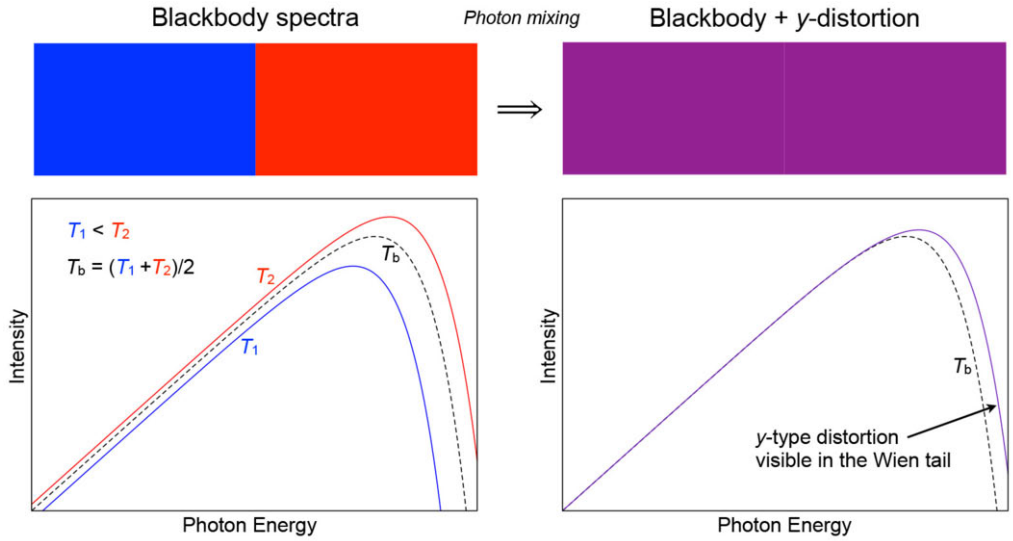


Fig. 15. – llustration for the superposition of blackbodies. We envision blackbody photons inside a box at two temperatures T_1 and T_2, and mean $T_\mathrm{b} = \frac{1}{2}(T_1 + T_2)$ initially (left panel). Thomson scattering mixes the two photon distributions without changing the photon number or energy. The averaged distribution is not a pure blackbody but at second order in the temperature difference exhibits a y-type distortion in the Wien tail (right panel). This then starts the thermalization process and repeated Compton scattering slowly converts the distortion to a μ-distortion.

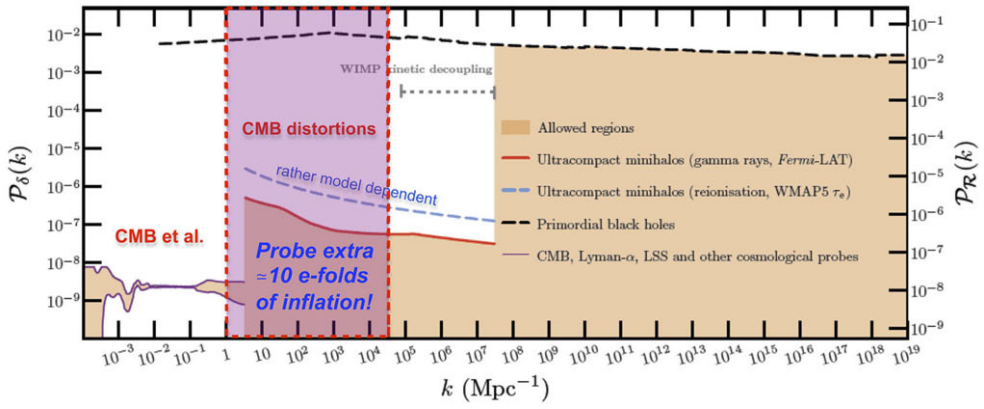


Fig. 16. – Current constraints on the small-scale power spectrum. At large scales ($k \lesssim 3\,\mathrm{Mpc}^{-1}$), CMB anisotropies and large scale structure measurements provide very stringent limits on the amplitude and shape of the primordial power spectrum. At smaller scales, the situation is much more uncertain and at $3\,\mathrm{Mpc}^{-1} \lesssim k \lesssim 10^4\,\mathrm{Mpc}^{-1}$ which can be targeted with CMB spectral distortion measurements wiggle room of at least two orders of magnitude is present. CMB distortion measurements could improve these limits to a level similar to the large-scale constraints. The figure is adapted from [94].

spectrum of scalar perturbations and $k_{\rm D}$ is the photon damping scale [107, 108], which early on scales as $k_{\rm D} \approx 4.048 \times 10^{-6}\,(1+z)^{3/2}\,{\rm Mpc}^{-1}$. For adiabatic modes, we obtain a heating efficiency $A^2 \approx (1+4R_\nu/15)^{-2} \approx 0.813$, where $R_\nu \approx 0.409$ for $N_{\rm eff} = 3.046$. The k-space integral is truncated at $k_{\rm min} \approx 0.12\,{\rm Mpc}^{-1}$, which reproduces the full heating rate across the recombination era quite well [14]. With this we can directly compute the associated distortion using `CosmoTherm` [12]. The various isocurvature perturbations can be treated in a similar manner [106]; however, in the standard inflation model these should be small. Tensor perturbations also contribute to the dissipation process, but the associated heating rate is orders of magnitudes lower than for adiabatic modes even for very blue tensor power spectra and thus can be neglected [109, 110].

For $A_{\rm s} = 2.207 \times 10^{-9}$, $n_{\rm S} = 0.9645$ and $n_{\rm run} = 0$ [9], we present the result in fig. 13. The adiabatic cooling distortion (see sect. **4**˙3) was simultaneously included. The signal is uncertain to within $\simeq 10\%$ in ΛCDM, simply because of the remaining uncertainties in the measurement of $A_{\rm s}$ and $n_{\rm S}$. It is described by a sum of μ- and y-distortion with $\mu \approx 2.0 \times 10^{-8}$ and $y \approx 3.6 \times 10^{-9}$ and a non-vanishing overall residual at the level of $\simeq 20\%$–30% [17]. In terms of raw sensitivity, this signal is close to the detection limit of a PIXIE-like experiment; however, foregrounds in particular at low frequencies make a detection more challenging [93]. Still, a PIXIE-like experiment could place interesting upper limits on the amplitude of scalar fluctuations around $k \simeq 10^3\,{\rm Mpc}^{-1}$ [101, 72], potentially helping to shed light on the small-scale crisis [111] and rule out models of inflation with increased small-scale power [104, 112].

The damping signal is also sensitive to primordial non-Gaussianity in the squeezed-limit, leading to a spatially varying spectral signal that correlates with CMB temperature anisotropies as large angular scales [36, 37]. This effect therefore provides a unique way for studing the scale-dependence of $f_{\rm NL}$ [113-118]. CMB spectral distortions hence deliver a complementary and independent probe of early-Universe physics, which allows capitalizing on the synergies with large-scale B-mode polarization measurements.

4˙3. *Adiabatic cooling for baryons.* – The adiabatic cooling of ordinary matter continuously extracts energy from the CMB photon bath by Compton scattering, leading to another small but guaranteed distortion that directly depends on the baryon density and helium abundance. The distortion is characterized by *negative* μ- and y-parameters at the level of $\simeq$ few $\times 10^{-9}$ [59, 12, 69]. The effective energy extraction history is given by

$$\begin{aligned}(37)\quad \frac{{\rm d}(Q/\rho_\gamma)}{{\rm d}z} &= -\frac{3}{2}\frac{N_{\rm tot} k T_\gamma}{\rho_\gamma (1+z)} \\ &\approx -\frac{5.7\times 10^{-10}}{(1+z)}\left[\frac{(1-Y_{\rm p})}{0.75}\right]\left[\frac{\Omega_{\rm b}h^2}{0.022}\right]\left[\frac{(1+f_{\rm He}+X_{\rm e})}{2.25}\right]\left[\frac{T_0}{2.726\,{\rm K}}\right]^{-3},\end{aligned}$$

where $N_{\rm tot} = N_{\rm H}(1+f_{\rm He}+X_{\rm e})$ is the number density of all thermally coupled baryons and electrons; $N_{\rm H} \approx 1.881 \times 10^{-6}\,(1+z)^3\,{\rm cm}^{-3}$ is the number density of hydrogen nuclei; $f_{\rm He} \approx Y_{\rm p}/4(1-Y_{\rm p}) \approx 0.0819$ and $X_{\rm e} = N_{\rm e}/N_{\rm H}$ is the free electron fraction,

which can be computed accurately with `CosmoRec` [40]. For *Planck* 2015 parameters, the signal is shown in fig. 13. It is uncertain at the $\simeq 1\%$ level in ΛCDM and cancels part of the damping signal; however, it is roughly one order of magnitude weaker and cannot be separated at the currently expected level of sensitivity of next generation CMB spectrometers.

Additional interactions of dark matter with photons, electrons or protons could further increase the cooling distortion [60]. This allows placing interesting constraints on the nature of dark matter and its interactions with the standard sectors. The recent EDGES measurements [76] have spurred increased interest in this possiblity [119-121].

4˙4. *The cosmological recombination radiation.* – The cosmological recombination process is associated with the emission of photons in free-bound and bound-bound transitions of hydrogen and helium [54,122,123]. This causes a small distortion of the CMB and the redshifted recombination photons should still be visible as the cosmological recombination radiation (CRR), a tiny spectral distortion ($\simeq$ nK-μK level) present at mm to dm wavelength (for overview see [33]). The amplitude of the CRR depends directly on the number density of baryons in the Universe. The helium abundance furthermore affects the detailed shape of the recombination lines, while the number of neutrinos has a minor effect [34]. Finally, the line positions and widths depend on when and how fast the Universe recombined. The CRR thus provides an independent way to constrain cosmological parameters and map the recombination history [124].

Several computations of this CRR have been carried out in the past [125-128, 32, 129-133]. These calculations were very time-consuming, taking a few days of supercomputer time for one cosmology [130, 133]. This big computational challenge was recently overcome [134, 34], today allowing us to compute the CRR in about 15 seconds on a standard laptop using `CosmoSpec`([6]) [34]. The *fingerprint* from the recombination era shows several distinct spectral features that encode valuable information about the recombination process (fig. 13). Many subtle radiative transfer and atomic physics processes [130, 132, 40, 135] can now be included by `CosmoSpec`, yielding the most detailed and accurate predictions of the CRR in the standard ΛCDM model to date (see fig. 17). In ΛCDM, the CRR is uncertain at the level of a few percent, with the error being dominated by atomic physics rather than cosmological parameter values [34].

The CRR is currently roughly $\simeq 6$ times below the estimated detection limit of PIXIE (cf. fig. 13) and a detection from space will require several times higher sensitivity [136]. In the future, this could be achieved by experimental concepts similar to PRISM [4] or Millimetron [137]. At low frequencies ($1\,\mathrm{GHz} \lesssim \nu \lesssim 10\,\mathrm{GHz}$), the significant spectral variability of the CRR may also allow us to detect it from the ground with APSERa [22]. This could open a new way for directly studying the conditions of the Universe at $z \simeq 10^3$ (HI-recombination), $z \simeq 2000$ (HeI-recombination) and $z \simeq 6000$ (HeII-recombination). Furthermore, if something unexpected happened during different stages of the recom-

([6]) `www.Chluba.de/CosmoSpec`.

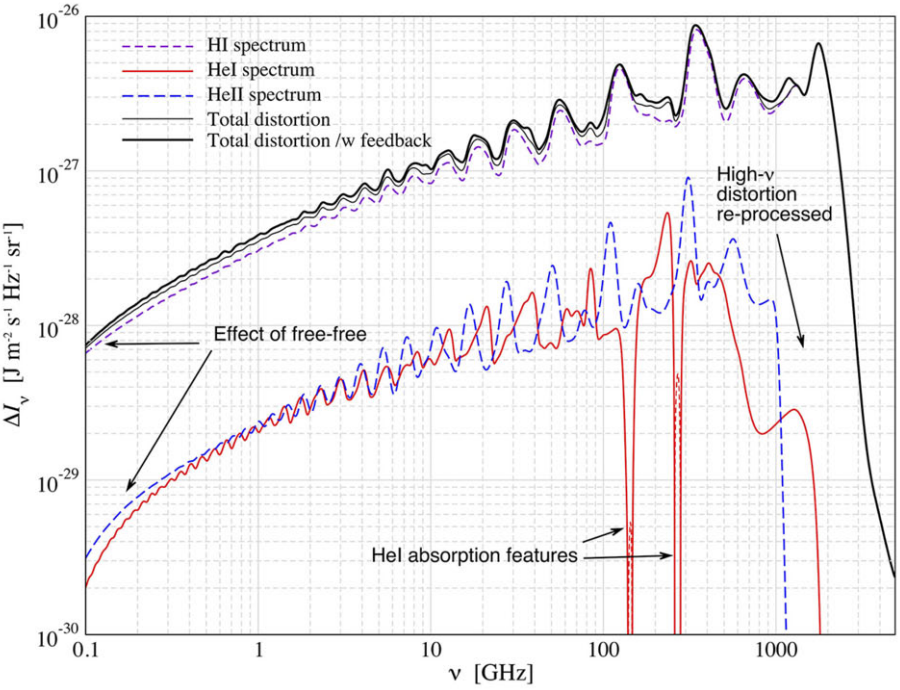


Fig. 17. – CRR from hydrogen and helium for 500-shell calculations. The different curves show individual contributions (without feedback) as well as the total distortion with and without feedback processes. At low frequencies, free-free absorption becomes noticeable. The effect is stronger for the contributions from helium due to the larger free-free optical depth before recombination ends at $z \simeq 10^3$. In total, some 6.1γ are emitted per hydrogen atom when all emission and feedback are included. Hydrogen alone contributes about $5.4\gamma/N_{\rm H}$ and helium $\simeq 0.7\gamma/N_{\rm H}$ ($\simeq 8.9\gamma/N_{\rm He}$). The figure is taken from [34].

bination epoch, atomic species will react to this [74] and produce additional distortion features that can exceed those of the normal recombination process. This will provide a unique way to distinguish pre- from post-recombination energy release [74, 138].

To appreciate the importance of the cosmological recombination process at $z \simeq 10^3$ a little more, consider that today measurements of the CMB anisotropies are sensitive to uncertainties of the ionization history at a level of $\simeq 0.1\%$–1% [139, 140]. For a precise interpretation of CMB data, uncertainties present in the original recombination calculations had to be reduced by including several previously omitted atomic physics and radiative transfer effects [141,139]. This led to the development of the new recombination modules CosmoRec [40] and HyRec [135] which are used in the analysis of Planck data [3]. Without these improve treatments of the recombination calculation the value for $n_{\rm S}$ would be biased by $\Delta n_{\rm S} \simeq -0.01$ to $n_{\rm S} \simeq 0.95$ instead of $\simeq 0.96$ [140]. We would be discussing different inflation models [142] without these corrections taken into

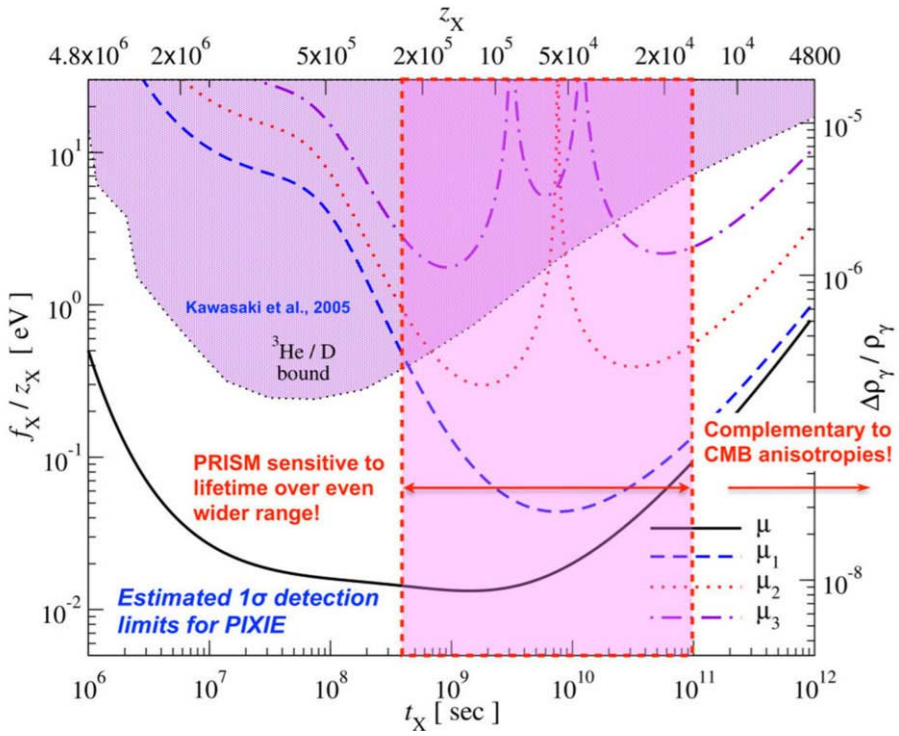


Fig. 18. – Decaying particle detection limits (1σ) for a PIXIE-like experiment. The eigenamplitudes μ_i characterize the non-μ/non-y distortion signal [72], which provides time-dependent information of the energy release history. CMB distortion limits could be $\simeq 50$ times tighter than those derived from light element abundances [169, 143]. A separate determination of lifetime and particle abundance could be possible for lifetimes $t_X \simeq 10^8\,\mathrm{s}$–$10^{11}\,\mathrm{s}$, being complementary to constraints derived using the CMB anisotropies [144-146]. The figure is adapted from [72].

account! Conversely, this emphasizes how important it is to experimentally confirm the recombination process and CMB spectral distortions provide a way to do so.

4˙5. *Dark matter annihilation.* – Today, cold dark matter is a well-established constituent of our Universe [2, 3, 9]. However, the nature of dark matter is still unclear and many groups are trying to gather any new clue to help unravel this big puzzle [147-151, 94, 152]. Similarly, it is unclear how dark matter was produced, however, within ΛCDM, the WIMP scenario provides one viable solution [153, 154]. In this case, dark matter should annihilate at a low level throughout the history of the Universe and even today.

For specific dark matter models, the level of annihilation around the recombination epoch is tightly constrained with the CMB anisotropies [148, 155-157, 151, 158, 159, 9]. The annihilation of dark matter can cause changes in the ionization history around last scattering ($z \simeq 10^3$), which in turn can lead to changes of the CMB temperature

and polarization anisotropies [144, 145, 160, 161]. Albeit significant dependence on the interaction of the annihilation products with the primordial plasma [162, 157, 163-165], the same process should lead to distortions of the CMB [166, 138, 12]. Sadly, it turns out that for the standard WIMP scenario with s-wave annihilation cross section, the expected signal is even smaller than the adiabatic cooling distortion [14]. We will thus not go into more details here.

4˙6. *Decaying particle scenarios*. – The CMB spectrum also allows us to place stringent limits on decaying particles in the pre-recombination epoch [167-170, 166, 138, 12]. This is especially interesting for decaying particles with lifetimes $t_{\rm X} \simeq 10^8\,{\rm s}$–$10^{11}\,{\rm s}$ [14, 72], as the exact shape of the distortion encodes when the decay occurred. Decays associated with significant low-energy photon production could furthermore create a unique spectral signature that can be distinguished from simple energy release [39]. This would provide an unprecedented probe of early-universe particle physics (*e.g.*, dark matter in excited states [171, 172]), with many natural particle candidates found in supersymmetric models [173, 174]. This could also shed light on gravitino physics [168, 175], axions [176] and primordial black holes [177, 178].

The expected 1σ detection limits for a PIXIE-like experiment are illustrated in fig. 18. The bounds obtained from measurements of light-elements [169, 143] could be superseded by more than one order of magnitude. Similar improvements from light-elements are not expected any time soon, and most recent updated only improved the limits by $\simeq$ 10% [179]. Spectral distortions thus provide a powerful new probe of particle physics.

4˙7. *Anisotropic CMB distortions*. – To close the discussion of different distortion signals, we briefly mention anisotropic ($\leftrightarrow$ *spectral-spatial*) CMB distortions. Even in the standard ΛCDM cosmology, anisotropies in the spectrum of the CMB are expected. The largest source of anisotropies is due to the Sunyaev-Zeldovich effect caused by the hot plasma inside clusters of galaxies [24, 62, 64, 35], as mentioned above. The y-distortion power spectrum has already been measured directly by Planck [180, 181] and encodes valuable information about the atmospheres of clusters [89, 182-186, 92]. Similarly, the warm hot intergalactic medium contributes and should become visible [91, 92].

In the primordial Universe, anisotropies in the μ- and y-distortions are expected to be tiny (relative perturbations $\lesssim 10^{-4}$, *e.g.*, see [187]) unless strong spatial variations in the primordial heating mechanism are expected [43]. As mentioned above, this could in principle be caused by non-Gaussianity of perturbations in the squeezed limit [36, 37, 113-116]; however, a present detectable levels of non-Gaussianity are beyond ΛCDM cosmology (see [117] for discussion of some of the foreground issues) and will not be considered further.

Another guaranteed anisotropic signal is due to Rayleigh scattering of CMB photons in the Lyman-series resonances of hydrogen around the recombination era [188, 189]. The signal is strongly frequency dependent, can be modeled precisely and may be detectable with future CMB imagers (*e.g.*, COrE+) or possibly PIXIE at large angular scales [189]. In a very similar manner, the resonant scattering of CMB photons by metals appearing

in the dark ages [190-193] or scattering in the excited levels of hydrogen during recombination [194, 193] can lead to anisotropic distortions. To measure these signals, precise channel intercalibration and foreground rejection is required.

Due to our motion relative to the CMB rest frame, the spectrum of the CMB dipole should also be distorted simply because the CMB monopole has a distortion [195, 196]. The signal associated with the large late-time y-distortion could be detectable with PIXIE at the level of a few σ [196]. Since for these measurements no absolute calibration is required, this effect will allow us to check for systematics. In addition, the dipole spectrum can be used to constrain monopole foregrounds [196, 16, 197].

Finally, due to the superposition of blackbodies of different temperatures (caused by the spherical harmonic expansion of the intensity map), the CMB quadrupole spectrum is also distorted, exhibiting a y-distortion related to our motion [198, 199]. The associated effective y-parameter is $y_{\rm Q} = \beta^2/6 \approx (2.525 \pm 0.012) \times 10^{-7}$ and should be noticeable with PIXIE and future CMB imagers [197].

5. – Conclusions

CMB spectral distortion measurements provide a unique way for studying physical processes leading to energy release or photon injection in the pre- and post-recombination eras. In the future, this could open a new unexplored window to early-universe and particle physics, delivering independent and complementary pieces of information about the Universe we live in. We highlighted several processes that should lead to distortions at a level within reach of present-day technology. Different distortion signals can be computed precisely and efficiently for various scenarios using both analytical and numerical schemes. Time-dependent information, beyond the standard μ- and y-type parametrization, may allow us to independently constrain lifetime and abundance of decaying relic particles, learn about the shape and amplitude of the small-scale power spectrum of primordial perturbations and shed light on dark matter. The cosmological recombination radiation will allow us to check our understanding of the recombination processes at redshifts of $z \simeq 10^3$. It furthermore should allow us to distinguish pre- from post-recombination y-distortions. All this emphasizes the immense potential of CMB spectroscopy, both in terms of *discovery* and *characterization* science, and we should make use of this invaluable source of information with the next CMB space mission and worldwide ground-based efforts.

* * *

JC cordially thanks the organizers of the Varenna International School of Physics "Enrico Fermi", Nicola Vittorio and Joseph Silk, for their great hospitality and continued encouragement to finish these notes. JC would also like to thank Boris Bolliet and Andrea Ravenni for their comments on the manuscript. JC is supported by the Royal Society as a Royal Society University Research Fellow and an European Research Council Consolidator Grant (*CMBSPEC*, No. 725456) at the University of Manchester, UK.

REFERENCES

[1] SMOOT G. F., BENNETT C. L., KOGUT A., WRIGHT E. L., AYMON J., BOGGESS N. W., CHENG E. S., DE AMICI G., GULKIS S., HAUSER M. G., HINSHAW G., JACKSON P. D., JANSSEN M., KAITA E., KELSALL T. and KEEGSTRA P., *Astrophys. J. Lett.*, **396** (1992) L1.

[2] BENNETT C. L., HALPERN M., HINSHAW G., JAROSIK N., KOGUT A., LIMON M., MEYER S. S., PAGE L., SPERGEL D. N., TUCKER G. S., WOLLACK E., WRIGHT E. L., BARNES C., GREASON M. R., HILL R. S., KOMATSU E., NOLTA M. R., ODEGARD N., PEIRIS H. V. and VERDE L., *Astrophys. J. Suppl.*, **148** (2003) 1.

[3] PLANCK COLLABORATION, ADE P. A. R., AGHANIM N., ARMITAGE-CAPLAN C., ARNAUD M., ASHDOWN M., ATRIO-BARANDELA F., AUMONT J., BACCIGALUPI C., BANDAY A. J. *et al.*, *Astron. Astrophys.*, **571** (2014) A16.

[4] ANDRÉ P., BACCIGALUPI C., BANDAY A., BARBOSA D., BARREIRO B., BARTLETT J., BARTOLO N., BATTISTELLI E., BATTYE R., BENDO G., BENO#523T A., BERNARD J.-P., BERSANELLI M., BÉTHERMIN M., BIELEWICZ P., BONALDI A., BOUCHET F., BOULANGER F., BRAND J., BUCHER M., BURIGANA C., CAI Z.-Y., CAMUS P., CASAS F., CASASOLA V., CASTEX G., CHALLINOR A., CHLUBA J., CHON G., COLAFRANCESCO S., COMIS B., CUTTAIA F., D'ALESSANDRO G., DA SILVA A., DAVIS R., DE AVILLEZ M., DE BERNARDIS P., DE PETRIS M., DE ROSA A., DE ZOTTI G., DELABROUILLE J., DÉSERT F.-X., DICKINSON C., DIEGO J. M., DUNKLEY J., ENSSLIN T., ERRARD J., FALGARONE E., FERREIRA P., FERRIÈRE K., FINELLI F., FLETCHER A., FOSALBA P., FULLER G., GALLI S., GANGA K., GARCÍA-BELLIDO J., GHRIBI A., GIARD M., GIRAUD-HÉRAUD Y., GONZALEZ-NUEVO J., GRAINGE K., GRUPPUSO A., HALL A., HAMILTON J.-C., HAVERKORN M., HERNANDEZ-MONTEAGUDO C., HERRANZ D., JACKSON M., JAFFE A., KHATRI R., KUNZ M., LAMAGNA L., LATTANZI M., LEAHY P., LESGOURGUES J., LIGUORI M., LIUZZO E., LOPEZ-CANIEGO M., MACIAS-PEREZ J., MAFFEI B., MAINO D., MANGILLI A., MARTINEZ-GONZALEZ E., MARTINS C., MASI S., MASSARDI M., MATARRESE S., MELCHIORRI A., MELIN J.-B., MENNELLA A., MIGNANO A., MIVILLE-DESCHÊNES M.-A., MONFARDINI A., MURPHY A., NASELSKY P., NATI F., NATOLI P., NEGRELLO M., NOVIELLO F., O'SULLIVAN C., PACI F., PAGANO L., PALADINO R., PALANQUE-DELABROUILLE N., PAOLETTI D., PEIRIS H., PERROTTA F., PIACENTINI F., PIAT M., PICCIRILLO L., PISANO G., POLENTA G., POLLO A., PONTHIEU N., REMAZEILLES M., RICCIARDI S., ROMAN M., ROSSET C., RUBINO-MARTIN J.-A., SALATINO M., SCHILLACI A., SHELLARD P., SILK J., STAROBINSKY A., STOMPOR R., SUNYAEV R., TARTARI A., TERENZI L., TOFFOLATTI L., TOMASI M., TRAPPE N., TRISTRAM M., TROMBETTI T., TUCCI M., VAN DE WEIJGAERT R., VAN TENT B., VERDE L., VIELVA P., WANDELT B., WATSON R. and WITHINGTON S., *JCAP*, **2** (2014) 6.

[5] ABAZAJIAN K. N., ARNOLD K., AUSTERMANN J., BENSON B. A., BISCHOFF C., BOCK J., BOND J. R., BORRILL J., BUDER I., BURKE D. L., CALABRESE E., CARLSTROM J. E., CARVALHO C. S., CHANG C. L., CHIANG H. C., CHURCH S., COORAY A., CRAWFORD T. M., CRILL B. P., DAWSON K. S., DAS S., DEVLIN M. J., DOBBS M., DODELSON S., DORÉ O., DUNKLEY J., FENG J. L., FRAISSE A., GALLICCHIO J., GIDDINGS S. B., GREEN D., HALVERSON N. W., HANANY S., HANSON D., HILDEBRANDT S. R., HINCKS A., HLOZEK R., HOLDER G., HOLZAPFEL W. L., HONSCHEID K., HOROWITZ G., HU W., HUBMAYR J., IRWIN K., JACKSON M., JONES W. C., KALLOSH R., KAMIONKOWSKI M., KEATING B., KEISLER R., KINNEY W., KNOX L., KOMATSU E., KOVAC J., KUO C.-L., KUSAKA A., LAWRENCE C., LEE A. T., LEITCH E., LINDE A., LINDER E., LUBIN P., MALDACENA J., MARTINEC E., MCMAHON J., MILLER A., MUKHANOV V., NEWBURGH L., NIEMACK M. D., NGUYEN H., NGUYEN H. T., PAGE L., PRYKE C.,

Reichardt C. L., Ruhl J. E., Sehgal N., Seljak U., Senatore L., Sievers J., Silverstein E., Slosar A., Smith K. M., Spergel D., Staggs S. T., Stark A., Stompor R., Vieregg A. G., Wang G., Watson S., Wollack E. J., Wu W. L. K., Yoon K. W., Zahn O. and Zaldarriaga M., *Astropart. Phys.*, **63** (2015) 55.

[6] Abazajian K. N., Arnold K., Austermann J., Benson B. A., Bischoff C., Bock J., Bond J. R., Borrill J., Calabrese E., Carlstrom J. E., Carvalho C. S., Chang C. L., Chiang H. C., Church S., Cooray A., Crawford T. M., Dawson K. S., Das S., Devlin M. J., Dobbs M., Dodelson S., Doré O., Dunkley J., Errard J., Fraisse A., Gallicchio J., Halverson N. W., Hanany S., Hildebrandt S. R., Hincks A., Hlozek R., Holder G., Holzapfel W. L., Honscheid K., Hu W., Hubmayr J., Irwin K., Jones W. C., Kamionkowski M., Keating B., Keisler R., Knox L., Komatsu E., Kovac J., Kuo C.-L., Lawrence C., Lee A. T., Leitch E., Linder E., Lubin P., McMahon J., Miller A., Newburgh L., Niemack M. D., Nguyen H., Nguyen H. T., Page L., Pryke C., Reichardt C. L., Ruhl J. E., Sehgal N., Seljak U., Sievers J., Silverstein E., Slosar A., Smith K. M., Spergel D., Staggs S. T., Stark A., Stompor R., Vieregg A. G., Wang G., Watson S., Wollack E. J., Wu W. L. K., Yoon K. W. and Zahn O., *Astropart. Phys.*, **63** (2015) 66.

[7] Abazajian K. N., Adshead P., Ahmed Z., Allen S. W., Alonso D., Arnold K. S., Baccigalupi C., Bartlett J. G., Battaglia N., Benson B. A., Bischoff C. A., Borrill J., Buza V., Calabrese E., Caldwell R., Carlstrom J. E., Chang C. L., Crawford T. M., Cyr-Racine F.-Y., De Bernardis F., de Haan T., di Serego Alighieri S., Dunkley J., Dvorkin C., Errard J., Fabbian G., Feeney S., Ferraro S., Filippini J. P., Flauger R., Fuller G. M., Gluscevic V., Green D., Grin D., Grohs E., Henning J. W., Hill J. C., Hlozek R., Holder G., Holzapfel W., Hu W., Huffenberger K. M., Keskitalo R., Knox L., Kosowsky A., Kovac J., Kovetz E. D., Kuo C.-L., Kusaka A., Le Jeune M., Lee A. T., Lilley M., Loverde M., Madhavacheril M. S., Mantz A., Marsh D. J. E., McMahon J., Meerburg P. D., Meyers J., Miller A. D., Munoz J. B., Nguyen H. N., Niemack M. D., Peloso M., Peloton J., Pogosian L., Pryke C., Raveri M., Reichardt C. L., Rocha G., Rotti A., Schaan E., Schmittfull M. M., Scott D., Sehgal N., Shandera S., Sherwin B. D., Smith T. L., Sorbo L., Starkman G. D., Story K. T., van Engelen A., Vieira J. D., Watson S., Whitehorn N. and Kimmy Wu W. L., arXiv:1610.02743 (2016).

[8] Calabrese E., Hlozek R. A., Battaglia N., Battistelli E. S., Bond J. R., Chluba J., Crichton D., Das S., Devlin M. J., Dunkley J., Dünner R., Farhang M., Gralla M. B., Hajian A., Halpern M., Hasselfield M., Hincks A. D., Irwin K. D., Kosowsky A., Louis T., Marriage T. A., Moodley K., Newburgh L., Niemack M. D., Nolta M. R., Page L. A., Sehgal N., Sherwin B. D., Sievers J. L., Sifón C., Spergel D. N., Staggs S. T., Switzer E. R. and Wollack E. J., *Phys. Rev. D*, **87** (2013) 103012.

[9] Planck Collaboration, Ade P. A. R., Aghanim N., Arnaud M., Ashdown M., Aumont J., Baccigalupi C., Banday A. J., Barreiro R. B., Bartlett J. G. *et al.*, *Astron. Astrophys.*, **594** (2016) A13.

[10] Fixsen D. J., Cheng E. S., Gales J. M., Mather J. C., Shafer R. A. and Wright E. L., *Astrophys. J.*, **473** (1996) 576.

[11] Fixsen D. J., *Astrophys. J.*, **707** (2009) 916.

[12] Chluba J. and Sunyaev R. A., *Mon. Not. R. Astron. Soc.*, **419** (2012) 1294.

[13] Sunyaev R. A. and Khatri R., *Int. J. Mod. Phys. D*, **22** (2013) 30014.

[14] Chluba J., *Mon. Not. R. Astron. Soc.*, **436** (2013) 2232.

[15] TASHIRO H., *Prog. Theor. Exp. Phys.*, **2014** (2014) 060000.
[16] DE ZOTTI G., NEGRELLO M., CASTEX G., LAPI A. and BONATO M., *JCAP*, **3** (2016) 047.
[17] CHLUBA J., *Mon. Not. R. Astron. Soc.*, **460** (2016) 227.
[18] FIXSEN D. J. and MATHER J. C., *Astrophys. J.*, **581** (2002) 817.
[19] KOGUT A., FIXSEN D. J., CHUSS D. T., DOTSON J., DWEK E., HALPERN M., HINSHAW G. F., MEYER S. M., MOSELEY S. H., SEIFFERT M. D., SPERGEL D. N. and WOLLACK E. J., *JCAP*, **7** (2011) 25.
[20] KOGUT A., CHLUBA J., FIXSEN D. J., MEYER S. and SPERGEL D., *The Primordial Inflation Explorer (PIXIE) Proc. SPIE*, **9904** (2016) 99040W.
[21] SILK J. and CHLUBA J., *Science*, **344** (2014) 586.
[22] SATHYANARAYANA RAO M., SUBRAHMANYAN R., UDAYA SHANKAR N. and CHLUBA J., *Astrophys. J.*, **810** (2015) 3.
[23] KOUVELIOTOU C., AGOL E., BATALHA N., BEAN J., BENTZ M., CORNISH N., DRESSLER A., FIGUEROA-FELICIANO E., GAUDI S., GUYON O., HARTMANN D., KALIRAI J., NIEMACK M., OZEL F., REYNOLDS C., ROBERGE A., STRAUGHN K. S. A., WEINBERG D. and ZMUIDZINAS J., arXiv:1401.3741 (2014).
[24] ZELDOVICH Y. B. and SUNYAEV R. A., *Astrophys. Space Sci.*, **4** (1969) 301.
[25] SUNYAEV R. A. and ZELDOVICH Y. B., *Astrophys. Space Sci.*, **7** (1970) 20.
[26] ILLARIONOV A. F. and SUNYAEV R. A., *Astronom. Zh.*, **51** (1974) 1162.
[27] ILLARIONOV A. F. and SUNYAEV R. A., *Sov. Astron.*, **18** (1975) 413.
[28] DANESE L. and DE ZOTTI G., *Riv. Nuovo Cimento*, **7** (1977) 277.
[29] DANESE L. and DE ZOTTI G., *Astron. Astrophys.*, **107** (1982) 39.
[30] BURIGANA C., DANESE L. and DE ZOTTI G., *Astron. Astrophys.*, **246** (1991) 49.
[31] HU W. and SILK J., *Phys. Rev. D*, **48** (1993) 485.
[32] RUBIÑO-MARTÍN J. A., CHLUBA J. and SUNYAEV R. A., *Mon. Nat. R. Astron. Soc.*, **371** (2006) 1939.
[33] SUNYAEV R. A. and CHLUBA J., *Astronom. Nachrich.*, **330** (2009) 657.
[34] CHLUBA J. and ALI-HAÏMOUD Y., *Mon. Nat. R. Astron. Soc.*, **456** (2016) 3494.
[35] CARLSTROM J. E., HOLDER G. P. and REESE E. D., *Annu. Rev. Astron. Astrophys.*, **40** (2002) 643.
[36] PAJER E. and ZALDARRIAGA M., *Phys. Rev. Lett.*, **109** (2012) 021302.
[37] GANC J. and KOMATSU E., *Phys. Rev. D*, **86** (2012) 023518.
[38] CHLUBA J., *Mon. Nat. R. Astron. Soc.*, **434** (2013) 352.
[39] CHLUBA J., *Mon. Nat. R. Astron. Soc.*, **454** (2015) 4182.
[40] CHLUBA J. and THOMAS R. M., *Mon. Nat. R. Astron. Soc.*, **412** (2011) 748.
[41] POZDNIAKOV L. A., SOBOL I. M. and SUNYAEV R. A., *Astron. Astrophys.*, **75** (1979) 214.
[42] SUNYAEV R. A. and TITARCHUK L. G., *Astron. Astrophys.*, **86** (1980) 121.
[43] CHLUBA J., KHATRI R. and SUNYAEV R. A., *Mon. Nat. R. Astron. Soc.*, **425** (2012) 1129.
[44] CHLUBA J. and DAI L., *Mon. Nat. R. Astron. Soc.*, **438** (2014) 1324.
[45] KOMPANEETS A., *Sov. Phys. JETP*, **31** (1956) 876.
[46] RYBICKI G. B. and LIGHTMAN A. P., *Radiative processes in astrophysics* (New York, Wiley-Interscience) 1979, p. 393.
[47] SAZONOV S. Y. and SUNYAEV R. A., *Astrophys. J.*, **543** (2000) 28.
[48] ITOH N., KOHYAMA Y. and NOZAWA S., *Astrophys. J.*, **502** (1998) 7.
[49] SAZONOV S. Y. and SUNYAEV R. A., *Astrophys. J.*, **508** (1998) 1.
[50] CHALLINOR A. and LASENBY A., *Astrophys. J.*, **499** (1998) 1.
[51] CHLUBA J., NAGAI D., SAZONOV S. and NELSON K., *Mon. Nat. R. Astron. Soc.*, **426** (2012) 510.
[52] CHLUBA J., SWITZER E., NELSON K. and NAGAI D., *Mon. Nat. R. Astron. Soc.*, **430** (2013) 3054.

[53] Hu W., arXiv:astro-ph/9508126 (1995).
[54] Zeldovich Y. B., Kurt V. G. and Syunyaev R. A., *Zh. Eksp. Teor. Fiz.*, **55** (1968) 278.
[55] Pritchard J. R. and Loeb A., *Phys. Rev. D*, **78** (2008) 103511.
[56] Chluba J., Sazonov S. Y. and Sunyaev R. A., *Astron. Astrophys.*, **468** (2007) 785.
[57] Itoh N., Sakamoto T., Kusano S., Nozawa S. and Kohyama Y., *Astrophys. J. Suppl.*, **128** (2000) 125.
[58] Chluba J., *Mon. Nat. R. Astron. Soc.*, **440** (2014) 2544.
[59] Chluba J., *Spectral Distortions of the Cosmic Microwave Background*, PhD Thesis, LMU München (Mar. 2005).
[60] Ali-Haïmoud Y., Chluba J. and Kamionkowski M., *Phys. Rev. Lett.*, **115** (2015) 071304.
[61] Planck Collaboration, Ade P. A. R., Aghanim N., Armitage-Caplan C., Arnaud M., Ashdown M., Atrio-Barandela F., Aumont J., Aussel H., Baccigalupi C. *et al.*, *Astron. Astrophys.*, **571** (2014) A29.
[62] Sunyaev R. A. and Zeldovich I. B., *Mon. Nat. R. Astron. Soc.*, **190** (1980) 413.
[63] Rephaeli Y., *Annu. Rev. Astron. Astrophys.*, **33** (1995) 541.
[64] Birkinshaw M., *Phys. Rep.*, **310** (1999) 97.
[65] Bolliet B., Comis B., Komatsu E. and Macías-Pérez J. F., *Mon. Nat. R. Astron. Soc.*, **477** (2018) 4957.
[66] Chluba J., Nagai D., Sazonov S. and Nelson K., *Mon. Nat. R. Astron. Soc.*, **426** (2012) 510.
[67] Zeldovich Y. B. and Levich E. V., *Sov. Phys. JETP*, **28** (1969) 1287.
[68] Zeldovich Y. B. and Sunyaev R. A., *Zh. Eksp. Teor. Fiz.*, **62** (1972) 153.
[69] Khatri R., Sunyaev R. A. and Chluba J., *Astron. Astrophys.*, **540** (2012) A124.
[70] Khatri R. and Sunyaev R. A., *JCAP*, **9** (2012) 16.
[71] Khatri R. and Sunyaev R. A., *JCAP*, **6** (2012) 38.
[72] Chluba J. and Jeong D., *Mon. Nat. R. Astron. Soc.*, **438** (2014) 2065.
[73] Lyubarsky Y. E. and Sunyaev R. A., *Astron. Astrophys.*, **123** (1983) 171.
[74] Chluba J. and Sunyaev R. A., *Astron. Astrophys.*, **501** (2009) 29.
[75] Chluba J., Hamann J. and Patil S. P., *Int. J. Mod. Phys. D*, **24** (2015) 1530023.
[76] Bowman J. D., Rogers A. E. E., Monsalve R. A., Mozdzen T. J. and Mahesh N., *Nature*, **555** (2018) 67.
[77] Fixsen D. J., Kogut A., Levin S., Limon M., Lubin P., Mirel P., Seiffert M. and Wollack E., *Astrophys. J.*, **612** (2004) 86.
[78] Seiffert M., Fixsen D. J., Kogut A., Levin S. M., Limon M., Lubin P. M., Mirel P., Singal J., Villela T., Wollack E. and Wuensche C. A., *Astrophys. J.*, **734** (2011) 6.
[79] Fixsen D. J., Kogut A., Levin S., Limon M., Lubin P., Mirel P., Seiffert M., Singal J., Wollack E., Villela T. and Wuensche C. A., *Astrophys. J.*, **734** (2011) 5.
[80] Feng C. and Holder G., *Astrophys. J. Lett.*, **858** (2018) L17.
[81] Hektor A., Hütsi G., Marzola L., Raidal M., Vaskonen V. and Veermäe H., arXiv:1803.09697 (2018).
[82] Moroi T., Nakayama K. and Tang Y., arXiv:1804.10378 (2018).
[83] Hill J. C., Battaglia N., Chluba J., Ferraro S., Schaan E. and Spergel D. N., *Phys. Rev. Lett.*, **115** (2015) 261301.
[84] Hu W., Scott D. and Silk J., *Phys. Rev. D*, **49** (1994) 648.
[85] Barkana R. and Loeb A., *Phys. Rep.*, **349** (2001) 125.
[86] Oh S. P., Cooray A. and Kamionkowski M., *Mon. Nat. R. Astron. Soc.*, **342** (2003) L20.
[87] Sunyaev R. A. and Zeldovich Y. B., *Astron. Astrophys.*, **20** (1972) 189.

[88] CEN R. and OSTRIKER J. P., *Astrophys. J.*, **514** (1999) 1.
[89] REFREGIER A. *et al.*, *Phys. Rev. D*, **61** (2000) 123001.
[90] MINIATI F., RYU D., KANG H., JONES T. W., CEN R. and OSTRIKER J. P., *Astrophys. J.*, **542** (2000) 608.
[91] ZHANG P., PEN U.-L. and TRAC H., *Mon. Nat. R. Astron. Soc.*, **355** (2004) 451.
[92] DOLAG K., KOMATSU E. and SUNYAEV R., *Mon. Nat. R. Astron. Soc.*, **463** (2016) 1797.
[93] ABITBOL M. H., CHLUBA J., HILL J. C. and JOHNSON B. R., *Mon. Nat. R. Astron. Soc.*, **471** (2017) 1126.
[94] BRINGMANN T., SCOTT P. and AKRAMI Y., *Phys. Rev. D*, **85** (2012) 125027.
[95] SUNYAEV R. A. and ZELDOVICH Y. B., *Astrophys. Space Sci.*, **9** (1970) 368.
[96] DALY R. A., *Astrophys. J.*, **371** (1991) 14.
[97] BARROW J. D. and COLES P., *Mon. Nat. R. Astron. Soc.*, **248** (1991) 52.
[98] HU W., SCOTT D. and SILK J., *Astrophys. J. Lett.*, **430** (1994) L5.
[99] HU W. and SUGIYAMA N., *Astrophys. J.*, **436** (1994) 456.
[100] PAJER E. and ZALDARRIAGA M., *JCAP*, **2** (2013) 36.
[101] CHLUBA J., ERICKCEK A. L. and BEN-DAYAN I., *Astrophys. J.*, **758** (2012) 76.
[102] DENT J. B., EASSON D. A. and TASHIRO H., *Phys. Rev. D*, **86** (2012) 023514.
[103] KHATRI R. and SUNYAEV R. A., *JCAP*, **6** (2013) 26.
[104] CLESSE S., GARBRECHT B. and ZHU Y., *JCAP*, **10** (2014) 046.
[105] CABASS G., MELCHIORRI A. and PAJER E., *Phys. Rev. D*, **93** (2016) 083515.
[106] CHLUBA J. and GRIN D., *Mon. Nat. R. Astron. Soc.*, **434** (2013) 1619.
[107] WEINBERG S., *Astrophys. J.*, **168** (1971) 175.
[108] KAISER N., *Mon. Nat. R. Astron. Soc.*, **202** (1983) 1169.
[109] OTA A., TAKAHASHI T., TASHIRO H. and YAMAGUCHI M., *JCAP*, **10** (2014) 029.
[110] CHLUBA J., DAI L., GRIN D., AMIN M. A. and KAMIONKOWSKI M., *Mon. Nat. R. Astron. Soc.*, **446** (2015) 2871.
[111] NAKAMA T., CHLUBA J. and KAMIONKOWSKI M., *Phys. Rev. D*, **95** (2017) 121302.
[112] CLESSE S. and GARCÍA-BELLIDO J., *Phys. Rev. D*, **92** (2015) 023524.
[113] BIAGETTI M., PERRIER H., RIOTTO A. and DESJACQUES V., *Phys. Rev. D*, **87** (2013) 063521.
[114] EMAMI R., DIMASTROGIOVANNI E., CHLUBA J. and KAMIONKOWSKI M., *Phys. Rev. D*, **91** (2015) 123531.
[115] CHLUBA J., DIMASTROGIOVANNI E., AMIN M. A. and KAMIONKOWSKI M., *Mon. Nat. R. Astron. Soc.*, **466** (2017) 2390.
[116] RAVENNI A., LIGUORI M., BARTOLO N. and SHIRAISHI M., *JCAP*, **9** (2017) 042.
[117] REMAZEILLES M. and CHLUBA J., *Mon. Nat. R. Astron. Soc.* (2018).
[118] CABASS G., PAJER E. and VAN DER WOUDE D., ArXiv e-prints 1805.08775 (2018).
[119] BARKANA R., *Nature*, **555** (2018) 71.
[120] MUÑOZ J. B. and LOEB A., arXiv:1802.10094 (2018).
[121] BERLIN A., HOOPER D., KRNJAIC G. and MCDERMOTT S. D., arXiv:1803.02804 (2018).
[122] PEEBLES P. J. E., *Astrophys. J.*, **153** (1968) 1.
[123] DUBROVICH V. K., *Sov. Astron. Lett.*, **1** (1975) 196.
[124] CHLUBA J. and SUNYAEV R. A., *Astron. Astrophys.*, **478** (2008) L27.
[125] RYBICKI G. B. and DELL'ANTONIO I. P., *Spectral Distortions in the CMB from Recombination.*, in *Observational Cosmology*, edited by CHINCARINI G. L., IOVINO A., MACCACARO T. and MACCAGNI D., *Proc. ASP Conf. Ser.*, Vol. **51** (Astronomical Society of the Pacific, San Francisco, CA) 1993, p. 548.
[126] DUBROVICH V. K. and STOLYAROV V. A., *Astron. Astrophys.*, **302** (1995) 635.
[127] DUBROVICH V. K. and STOLYAROV V. A., *Astron. Lett.*, **23** (1997) 565.

[128] Kholupenko E. E., Ivanchik A. V. and Varshalovich D. A., *Gravit. Cosmol.*, **11** (2005) 161.
[129] Chluba J. and Sunyaev R. A., *Astron. Astrophys.*, **458** (2006) L29.
[130] Chluba J., Rubiño-Martín J. A. and Sunyaev R. A., *Mon. Nat. R. Astron. Soc.*, **374** (2007) 1310.
[131] Rubiño-Martín J. A., Chluba J. and Sunyaev R. A., *Astron. Astrophys.*, **485** (2008) 377.
[132] Chluba J. and Sunyaev R. A., *Mon. Nat. R. Astron. Soc.*, **402** (2010) 1221.
[133] Chluba J., Vasil G. M. and Dursi L. J., *Mon. Nat. R. Astron. Soc.*, **407** (2010) 599.
[134] Ali-Haïmoud Y., *Phys. Rev. D*, **87** (2013) 023526.
[135] Ali-Haïmoud Y. and Hirata C. M., *Phys. Rev. D*, **83** (2011) 043513.
[136] Desjacques V., Chluba J., Silk J., de Bernardis F. and Doré O., *Mon. Nat. R. Astron. Soc.*, **451** (2015) 4460.
[137] Smirnov A. V., Baryshev A. M., Pilipenko S. V., Myshonkova N. V., Bulanov V. B., Arkhipov M. Y., Vinogradov I. S., Likhachev S. F. and Kardashev N. S., *Space mission Millimetron for terahertz astronomy*, *SPIE Conf. Ser.*, **8442** (2012) .
[138] Chluba J., *Mon. Nat. R. Astron. Soc.*, **402** (2010) 1195.
[139] Rubiño-Martín J. A., Chluba J., Fendt W. A. and Wandelt B. D., *Mon. Nat. R. Astron. Soc.*, **403** (2010) 439.
[140] Shaw J. R. and Chluba J., *Mon. Nat. R. Astron. Soc.*, **415** (2011) 1343.
[141] Fendt W. A., Chluba J., Rubiño-Martín J. A. and Wandelt B. D., *Astrophys. J. Suppl.*, **181** (2009) 627.
[142] Planck Collaboration, Ade P. A. R., Aghanim N., Armitage-Caplan C., Arnaud M., Ashdown M., Atrio-Barandela F., Aumont J., Baccigalupi C., Banday A. J. *et al.*, *Astron. Astrophys.*, **571** (2014) A22.
[143] Kawasaki M., Kohri K. and Moroi T., *Phys. Rev. D*, **71** (2005) 083502.
[144] Adams J. A., Sarkar S. and Sciama D. W., *Mon. Not. R. Astron. Soc.*, **301** (1998) 210.
[145] Chen X. and Kamionkowski M., *Phys. Rev. D*, **70** (2004) 043502.
[146] Zhang L., Chen X., Kamionkowski M., Si Z. and Zheng Z., *Phys. Rev. D*, **76** (2007) 061301.
[147] Adriani O., Barbarino G. C., Bazilevskaya G. A., Bellotti R., Boezio M., Bogomolov E. A., Bonechi L., Bongi M., Bonvicini V., Bottai S., Bruno A., Cafagna F., Campana D., Carlson P., Casolino M., Castellini G., de Pascale M. P., de Rosa G., de Simone N., di Felice V., Galper A. M., Grishantseva L., Hofverberg P., Koldashov S. V., Krutkov S. Y., Kvashnin A. N., Leonov A., Malvezzi V., Marcelli L., Menn W., Mikhailov V. V., Mocchiutti E., Orsi S., Osteria G., Papini P., Pearce M., Picozza P., Ricci M., Ricciarini S. B., Simon M., Sparvoli R., Spillantini P., Stozhkov Y. I., Vacchi A., Vannuccini E., Vasilyev G., Voronov S. A., Yurkin Y. T., Zampa G., Zampa N. and Zverev V. G., *Nature*, **458** (2009) 607.
[148] Galli S., Iocco F., Bertone G. and Melchiorri A., *Phys. Rev. D*, **80** (2009) 023505.
[149] CDMS II Collaboration, Ahmed Z., Akerib D. S., Arrenberg S., Bailey C. N., Balakishiyeva D., Baudis L., Bauer D. A., Brink P. L., Bruch T., Bunker R., Cabrera B., Caldwell D. O., Cooley J., Cushman P., Daal M., DeJongh F., Dragowsky M. R., Duong L., Fallows S., Figueroa-Feliciano E., Filippini J., Fritts M., Golwala S. R., Grant D. R., Hall J., Hennings-Yeomans R., Hertel S. A., Holmgren D., Hsu L., Huber M. E., Kamaev O., Kiveni M., Kos M., Leman S. W., Mahapatra R., Mandic V., McCarthy K. A., Mirabolfathi N., Moore D., Nelson H., Ogburn R. W., Phipps A., Pyle M., Qiu X., Ramberg E., Rau W., Reisetter A., Saab T., Sadoulet B., Sander J., Schnee R. W., Seitz

D. N., SERFASS B., SUNDQVIST K. M., TARKA M., WIKUS P., YELLIN S., YOO J., YOUNG B. A. and ZHANG J., *Science*, **327** (2010) 1619.
[150] ZAVALA J., VOGELSBERGER M., SLATYER T. R., LOEB A. and SPRINGEL V., *Phys. Rev. D*, **83** (2011) 123513.
[151] HÜTSI G., CHLUBA J., HEKTOR A. and RAIDAL M., *Astron. Astrophys.*, **535** (2011) A26.
[152] ASLANYAN G., PRICE L. C., ADAMS J., BRINGMANN T., CLARK H. A., EASTHER R., LEWIS G. F. and SCOTT P., arXiv:1512.04597 (2015).
[153] JUNGMAN G., KAMIONKOWSKI M. and GRIEST K., *Phys. Rep.*, **267** (1996) 195.
[154] BERTONE G., HOOPER D. and SILK J., *Phys. Rep.*, **405** (2005) 279.
[155] CIRELLI M., IOCCO F. and PANCI P., *JCAP*, **10** (2009) 9.
[156] HÜTSI G., HEKTOR A. and RAIDAL M., *Astron. Astrophys.*, **505** (2009) 999.
[157] SLATYER T. R., PADMANABHAN N. and FINKBEINER D. P., *Phys. Rev. D*, **80** (2009) 043526 http://link.aps.org/abstract/PRD/v80/e043526.
[158] GIESEN G., LESGOURGUES J., AUDREN B. and ALI-HAÏMOUD Y., *JCAP*, **12** (2012) 8.
[159] DIAMANTI R., LOPEZ-HONOREZ L., MENA O., PALOMARES-RUIZ S. and VINCENT A. C., *JCAP*, **2** (2014) 017.
[160] PADMANABHAN N. and FINKBEINER D. P., *Phys. Rev. D*, **72** (2005) 023508.
[161] ZHANG L., CHEN X., LEI Y. and SI Z., *Phys. Rev. D*, **74** (2006) 103519.
[162] SHULL J. M. and VAN STEENBERG M. E., *Astrophys. J.*, **298** (1985) 268.
[163] VALDÉS M., EVOLI C. and FERRARA A., *Mon. Nat. R. Astron. Soc.*, **404** (2010) 1569.
[164] GALLI S., SLATYER T. R., VALDES M. and IOCCO F., *Phys. Rev. D*, **88** (2013) 063502.
[165] SLATYER T. R., *Phys. Rev. D*, **93** (2016) 023521.
[166] MCDONALD P., SCHERRER R. J. and WALKER T. P., *Phys. Rev. D*, **63** (2001) 023001.
[167] SARKAR S. and COOPER A. M., *Phys. Lett. B*, **148** (1984) 347.
[168] ELLIS J., NANOPOULOS D. V. and SARKAR S., *Nucl. Phys. B*, **259** (1985) 175.
[169] ELLIS J., GELMINI G. B., LOPEZ J. L., NANOPOULOS D. V. and SARKAR S., *Nucl. Phys. B*, **373** (1992) 399.
[170] HU W. and SILK J., *Phys. Rev. Lett.*, **70** (1993) 2661.
[171] POSPELOV M. and RITZ A., *Phys. Lett. B*, **651** (2007) 208.
[172] FINKBEINER D. P. and WEINER N., *Phys. Rev. D*, **76** (2007) 083519.
[173] FENG J. L., RAJARAMAN A. and TAKAYAMA F., *Phys. Rev. D*, **68** (2003) 063504.
[174] FENG J. L., *Annu. Rev. Astron. Astrophys.*, **48** (2010) 495.
[175] DIMASTROGIOVANNI E., KRAUSS L. M. and CHLUBA J., *Phys. Rev. D*, **94** (2016) 023518.
[176] TASHIRO H., SILK J. and MARSH D. J. E., *Phys. Rev. D*, **88** (2013) 125024.
[177] CARR B. J., KOHRI K., SENDOUDA Y. and YOKOYAMA J., *Phys. Rev. D*, **81** (2010) 104019.
[178] POULIN V., SERPICO P. D., CALORE F., CLESSE S. and KOHRI K., *Phys. Rev. D*, **96** (2017) 083524.
[179] KAWASAKI M., KOHRI K., MOROI T. and TAKAESU Y., *Phys. Rev. D*, **97** (2018) 023502.
[180] PLANCK COLLABORATION, ADE P. A. R., AGHANIM N., ARMITAGE-CAPLAN C., ARNAUD M., ASHDOWN M., ATRIO-BARANDELA F., AUMONT J., BACCIGALUPI C., BANDAY A. J. *et al.*, arXiv e-prints (2013).
[181] PLANCK COLLABORATION, AGHANIM N., ARNAUD M., ASHDOWN M., AUMONT J., BACCIGALUPI C., BANDAY A. J., BARREIRO R. B., BARTLETT J. G., BARTOLO N. *et al.*, arXiv:1502.01596 (2015).
[182] KOMATSU E. and SELJAK U., *Mon. Nat. R. Astron. Soc.*, **336** (2002) 1256.
[183] DIEGO J. M. and MAJUMDAR S., *Mon. Nat. R. Astron. Soc.*, **352** (2004) 993.
[184] BATTAGLIA N., BOND J. R., PFROMMER C., SIEVERS J. L. and SIJACKI D., *Astrophys. J.*, **725** (2010) 91.

[185] SHAW L. D., NAGAI D., BHATTACHARYA S. and LAU E. T., *Astrophys. J.*, **725** (2010) 1452.
[186] MUNSHI D., JOUDAKI S., SMIDT J., COLES P. and KAY S. T., *Mon. Nat. R. Astron. Soc.*, **429** (2013) 1564.
[187] PITROU C., BERNARDEAU F. and UZAN J.-P., *JCAP*, **7** (2010) 19.
[188] YU Q., SPERGEL D. N. and OSTRIKER J. P., *Astrophys. J.*, **558** (2001) 23.
[189] LEWIS A., *JCAP*, **8** (2013) 053.
[190] LOEB A., *Astrophys. J. Lett.*, **555** (2001) L1.
[191] ZALDARRIAGA M. and LOEB A., *Astrophys. J.*, **564** (2002) 52.
[192] BASU K., HERNÁNDEZ-MONTEAGUDO C. and SUNYAEV R. A., *Astron. Astrophys.*, **416** (2004) 447.
[193] HERNÁNDEZ-MONTEAGUDO C., RUBIÑO-MARTÍN J. A. and SUNYAEV R. A., *Mon. Nat. R. Astron. Soc.*, **380** (2007) 1656.
[194] RUBIÑO-MARTÍN J. A., HERNÁNDEZ-MONTEAGUDO C. and SUNYAEV R. A., *Astron. Astrophys.*, **438** (2005) 461.
[195] DANESE L. and DE ZOTTI G., *Astron. Astrophys.*, **94** (1981) L33.
[196] BALASHEV S. A., KHOLUPENKO E. E., CHLUBA J., IVANCHIK A. V. and VARSHALOVICH D. A., *Astrophys. J.*, **810** (2015) 131.
[197] BURIGANA C., CARVALHO C. S., TROMBETTI T., NOTARI A., QUARTIN M., GASPERIS G. D., BUZZELLI A., VITTORIO N., DE ZOTTI G., DE BERNARDIS P., CHLUBA J., BILICKI M., DANESE L., DELABROUILLE J., TOFFOLATTI L., LAPI A., NEGRELLO M., MAZZOTTA P., SCOTT D., CONTRERAS D., ACHÚCARRO A., ADE P., ALLISON R., ASHDOWN M., BALLARDINI M., BANDAY A. J., BANERJI R., BARTLETT J., BARTOLO N., BASAK S., BERSANELLI M., BONALDI A., BONATO M., BORRILL J., BOUCHET F., BOULANGER F., BRINCKMANN T., BUCHER M., CABELLA P., CAI Z.-Y., CALVO M., CASTELLANO M. G., CHALLINOR A., CLESSE S., COLANTONI I., COPPOLECCHIA A., CROOK M., D'ALESSANDRO G., DIEGO J.-M., DI MARCO A., DI VALENTINO E., ERRARD J., FEENEY S., FERNÁNDEZ-COBOS R., FERRARO S., FINELLI F., FORASTIERI F., GALLI S., GÉNOVA-SANTOS R., GERBINO M., GONZÁLEZ-NUEVO J., GRANDIS S., GREENSLADE J., HAGSTOTZ S., HANANY S., HANDLEY W., HERNÁNDEZ-MONTEAGUDO C., HERVIAS-CAIMAPO C., HILLS M., HIVON E., KIIVERI K., KISNER T., KITCHING T., KUNZ M., KURKI-SUONIO H., LAMAGNA L., LASENBY A., LATTANZI M., LESGOURGUES J., LIGUORI M., LINDHOLM V., LOPEZ-CANIEGO M., LUZZI G., MAFFEI B., MANDOLESI N., MARTINEZ-GONZALEZ E., MARTINS C. J. A. P., MASI S., MATARRESE S., MCCARTHY D., MELCHIORRI A., MELIN J.-B., MOLINARI D., MONFARDINI A., NATOLI P., PAIELLA A., PAOLETTI D., PATANCHON G., PIAT M., PISANO G., POLASTRI L., POLENTA G., POLLO A., POULIN V., REMAZEILLES M., ROMAN M., RUBIÑO-MARTÍN J.-A., SALVATI L., TARTARI A., TOMASI M., TRAMONTE D., TRAPPE N., TUCKER C., VÄLIVIITA J., VAN DE WEIJGAERT R., VAN TENT B., VENNIN V., VIELVA P., YOUNG K. and ZANNONI M., *JCAP*, **4** (2018) 021.
[198] KAMIONKOWSKI M. and KNOX L., *Phys. Rev. D*, **67** (2003) 063001.
[199] CHLUBA J. and SUNYAEV R. A., *Astron. Astrophys.*, **424** (2004) 389.

This page intentionally left blank

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200013

Recent developments in the analysis of galaxy surveys

Will J. Percival(*)

Department of Physics and Astronomy, University of Waterloo
200 University Ave W, Waterloo, ON N2L 3G1, Canada

Perimeter Institute for Theoretical Physics - 31 Caroline St. North
Waterloo, ON N2L 2Y5, Canada

Summary. — These are advanced lecture notes covering recent developments in the methodology used to analyse galaxy surveys. The focus is particularly on direct measurements of the galaxy power spectrum although we also discuss its Fourier transform, the correlation function for comparison. These 2-point statistics, under the assumption that the overdensity field has Gaussian statistics on large scales, contain the majority of the cosmological signal available from the galaxy distribution. Recent developments in multipole measurements, dealing with systematics, convolving theoretical models with the survey window function, the approximation of covariance matrices, and weighting schemes for measuring evolution with redshift are considered. The focus is on analytic explanation of the issues involved rather than on recent analyses or simulation results. These notes only loosely follow the lectures I gave in Varenna, which were more wide ranging and contained more introductory material. However, I have in the not-too-distant past created lecture notes for an introductory course on galaxy survey analysis, and I did not want to duplicate those notes, but instead write something new for these proceedings.

(*) E-mail: will.percival@uwaterloo.ca

1. – Introduction

This review of recent developments in the analysis of galaxy surveys is meant to be useful for someone who already has a basic understanding of the field. For a general introduction see, for example, [1].

2. – The overdensity field

The dimensionless overdensity field is defined as

$$\delta(\mathbf{s}) = \frac{\rho(\mathbf{s}) - \langle\rho(\mathbf{s})\rangle}{\langle\rho(\mathbf{s})\rangle}\,, \tag{1}$$

where $\rho(\mathbf{s})$ is the observed galaxy density and $\langle\rho(\mathbf{s})\rangle$ is the expected density.

At early times, and on large scales at present day, $\delta(\mathbf{s})$ has a distribution that is close to that of Gaussian, adiabatic fluctuations [2], and in these limits the statistical distribution is completely described by the two-point functions of this field.

In ref. [3] (commonly referred to as the FKP paper) the first full analysis pipeline for a galaxy survey in Fourier space is presented. The start of this pipeline is to define the function

$$F(\mathbf{s}) = \frac{w(\mathbf{s})}{I^{1/2}}\left[n(\mathbf{s}) - \langle n(\mathbf{s})\rangle\right], \tag{2}$$

where $n(\mathbf{s})$ is the observed number density of galaxies around location $\mathbf{s}$. The expected value $\langle n(\mathbf{s})\rangle$ is commonly determined by means of a synthetic catalog of *random* points, Poisson sampled with the same mask and selection function as the survey. In this case, $\langle n(\mathbf{s})\rangle \equiv \alpha n_r(\mathbf{s})$, where α normalizes the weighted random catalogue with density $n_r(\mathbf{s})$ to match the weighted galaxy catalogue. The random catalogue is usually set to have $\gtrsim 50$ times as many points as galaxies in order that the shot noise contribution is subdominant compared to that of the galaxy catalogue.

In contrast to $\delta(\mathbf{s})$ in eq. (1), the denominator in the expression for $F(\mathbf{s})$ in eq. (2) is not a function of $\mathbf{s}$. Thus, fluctuations in the expected density of galaxies across the survey are not normalised out, and we have to accept that the power spectrum of the field $F(\mathbf{s})$ is convolved with a window function. In general, it is not easy to divide by the density when calculating the power spectrum as regions outside of the survey, or where no galaxies are expected because of discreteness effects when using a random catalogue to specific the survey selection function, would cause divide-by-zero problems in the calculation of $F(\mathbf{s})$.

The factor I normalizes this expression such that the observed monopole moment of the power has the correct amplitude in a universe with no window, $I \equiv \int \mathrm{d}\mathbf{s}\, w^2\bar{n}^2(\mathbf{s})$.

The match between galaxy and random catalogues creates what is known as the Integral Constraint (IC), forcing the average density within the survey region to be zero, ignoring larger-than-survey fluctuations in density. In sect. **11**, we discuss how to

calculate α in such a way that the effect of this IC can be included in models to be fitted to the data.

The FKP paper showed how the galaxies in the survey should be optimally weighted to allow for variations in density across a survey, balancing sample variance and shot noise. Each galaxy is weighted by

$$w_{\rm FKP}(\mathbf{r}) = \frac{1}{1+\bar{n}(\mathbf{r})\bar{P}(k)}\,, \tag{3}$$

where $\bar{n}(\mathbf{r})$ is the expected density of galaxies, and $\bar{P}(k)$ the expected power spectrum, usually fixed at a fiducial value. This optimal weight depends on a number of assumptions —particularly that the galaxies Poisson sample the density field, and that the galaxies all have the same clustering strength. Revised weighting schemes have been proposed for more realistic models of these effects (*e.g.* [4, 5]).

3. – Line-of-sight assumptions

Most theoretical models (*e.g.* [6]) make predictions for an idealised survey in which the line of sight (LOS) to every galaxy is assumed to be parallel. However, for a real galaxy survey the LOS to different galaxies are not parallel, and analysing a survey under the global plane-parallel assumption only gives results close to the theory for distant surveys with small angular coverage (*e.g.* WiggleZ, [7]). For surveys covering a wide angular region (*e.g.* the Baryon Oscillation Spectroscopic Survey BOSS, [8]), such a global approximation gives a poor match to the standard theory.

In order to make a measurement that can be matched to theory (in which it is assumed that LOS to galaxies are parallel), it is far better to make a local plane-parallel approximation when analysing a survey: that is, when measuring 2-point clustering we split the survey into pairs of galaxies, and make a local plane-parallel approximation for each pair. This approximation will still break down for pairs of galaxies separated by a wide angle, but gives results closer to the global plane-parallel clustering amplitude for pairs whose angular separation is small. There is a subtlety in that when analysing data, we can choose whether to define the single line of sight as matching that for one galaxy in a pair, or as the direction to the pair centre, and then furthermore define how the pair centre is calculated. However, this only produces a minor perturbation on the overall effect.

The difference between clustering measurements made using the local plane-parallel approximation (see next section), and the model provided in the global plane-parallel framework, is commonly called the wide-angle effect. This gets worse for wide-angle pairs, and the severity of the problem depends on the distribution of pair separation angles in a survey. Note that the plane-parallel approximation is also commonly called the distant observer approximation.

In order to allow for wide-angle effects, one could imagine trying to model the power spectrum calculated under the local plane-parallel assumption using the full wide-angle

theory [9, 10]. However, it is not clear that there would be a gain compared with using the correlation function, for which it would be simple to split pair counts into bins in separation, pair-centre angle to the LOS, and angular separation, and then model these directly.

4. – Multipole moments

Models of the Alcock-Packynski (AP, [11]) effect and of linear Redshift-Space Distortions (RSD, [6]) show that, to first order, under the global plane-parallel assumption, the cosmological information of interest is contained within the first three even power-law moments of the correlation function or power spectrum with respect to μ, the cosine of the angle that the pair or that the Fourier mode makes with respect to the line of sight (LOS). Decomposing into a Legendre polynomial basis instead of power-law moments has the advantage of giving independent moments of the power spectrum in the absence of a window function in the global plane-parallel limit. Remembering that the first 3 even Legendre polynomials are

$$L_0(\mu) = 1, \tag{4}$$

$$L_2(\mu) = \frac{1}{2}(3\mu^2 - 1)\,, \tag{5}$$

$$L_4(\mu) = \frac{1}{8}(35\mu^4 - 30\mu^2 + 3); \tag{6}$$

we see that the Legendre polynomial moments can be determined by the trivial linear combination of power-law moments. Thus we use both interchangeably in the following depending on which is simplest to adopt.

The Legendre polynomial moments of the unconvolved correlation function and power spectrum in the global plane-parallel approximation are related by

$$\xi_\ell(s) = (2\ell+1)\int_0^1 \mathrm{d}\mu\, \xi(\mathbf{s})\mathcal{L}_\ell(\mu_s), \tag{7}$$

$$P_\ell(k) = (2\ell+1)\int_0^1 \mathrm{d}\mu\, P(\mathbf{k})\mathcal{L}_\ell(\mu_k), \tag{8}$$

with reciprocal formulae

$$\xi(\mathbf{s}) = \sum_\ell \xi_\ell(s)\mathcal{L}_\ell(\mu_s), \tag{9}$$

$$P(\mathbf{k}) = \sum_\ell P_\ell(k)\mathcal{L}_\ell(\mu_k). \tag{10}$$

The power spectrum and correlation function moments are related by the Hankel transform

$$P_\ell(k) = 4\pi i^\ell \int \mathrm{d}s\, s^2 \xi_\ell(s) j_\ell(sk). \tag{11}$$

In the following, quantities that include the survey window function are denoted by a prime, while unconvolved quantities are not. We do not make a distinction between measured and model quantities as this should be clear from the context.

5. – Correlation function estimators in the local plane-parallel formalism

The correlation function is most commonly measured using the Landy-Szalay estimator [12]

$$\bar{\xi}(s,\mu) = \frac{DD(s,\mu) - 2DR(s,\mu) + RR(s,\mu)}{RR(s,\mu)}\,, \tag{12}$$

where we can define μ with respect to the LOS to the pair centre. $DD(s)$ is the number of galaxy-galaxy pairs within a bin with centre s normalised to the maximum possible number of galaxy-galaxy pairs, and $RR(s)$ and $DR(s)$ are the normalised number of random-random pairs, and galaxy-random pairs, respectively.

This estimate is biased by the integral constraint, a consequence of the fact that the total number of galaxy-galaxy pairs is estimated from the sample itself (equivalent to determining α in eq. (2) from the galaxies).

$$\langle 1 + \bar{\xi}(s,\mu)\rangle = \frac{1+\xi(s,\mu)}{1+\xi_\Omega(s)}\,, \tag{13}$$

where $\xi_\Omega(s)$ is the mean of the two-point correlation function over the mask [12]. For modern large galaxy surveys this correction is negligibly small unless we are interested in the clustering on very large scales. See sect. **11** to see how this affects the power spectrum measurement.

From this, we can calculate the multipoles by integrating over μ as in eq. (7). Note that this integral should be carried out separately from the calculation of $\xi(s,\mu)$ of eq. (12) because the distribution of pairs in a survey is usually not uniformly distributed in μ as required by the integral in eq. (7). An alternative is to define a pseudo-multipole estimator where the kernel is not exactly the Legendre polynomial, which generally complicates the analysis.

Note that eq. (12) estimates the unconvolved correlation function, and so can be directly compared with models. As we will see later, this is not necessarily true for direct estimators of the power spectrum.

6. – Power spectrum estimators in the global plane-parallel formalism

The Quadratic Maximum Likelihood (QML) estimator [13] correctly accounts for correlations between modes when optimally measuring the power spectrum from data. In the limit of uncorrelated modes with equal noise per mode in each bin this simplifies to the FKP estimator [3].

The QML estimator is given by

$$P(k_i) = \sum_j \mathbf{N}_{ij}^{-1} \mathbf{p}_j, \tag{14}$$

where the power is a convolution of the inverse of a normalisation matrix $\mathbf{N}_{ij}$ and a weighted two-point function

$$\mathbf{p}_j \equiv \sum_{\alpha,\beta} F^*(\mathbf{k}_\alpha) \mathbf{E}_{\alpha\beta}(k_j) F(\mathbf{k}_\beta). \tag{15}$$

The weight is given by the estimator matrix

$$\mathbf{E}(k_j) = -\frac{\partial \mathbf{C}^{-1}}{\partial P(k_j)}, \tag{16}$$

which describes how the inverse of the density field covariance matrix $\mathbf{C}$ changes with respect to the prior of the power spectrum of the respective bin. If the QML normalisation is proportional to the Fisher information,

$$\mathbf{N}_{ij} = \mathrm{tr}\left\{ \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial P(k_i)} \mathbf{C}^{-1} \frac{\partial \mathbf{C}}{\partial P(k_j)} \right\}, \tag{17}$$

the QML estimator is the optimal maximum likelihood estimator of the variance of a field that obeys a multivariate Gaussian distribution [13]. *I.e.* assuming a Gaussian density field, the QML estimator therefore provides an estimate of the power spectrum with minimal errors.

Under the assumption that all modes are independent, the QML estimator reduces to the FKP estimator

$$P(k_i) = \frac{1}{N_i} \sum_{\mathbf{k}_\alpha \in \mathrm{bin}\, i} |F(\mathbf{k}_\alpha)|^2, \tag{18}$$

which simply averages the power in the Fourier modes. The FKP-style estimator is commonly applied even when the assumptions required for optimality are not valid. This concept of averaging rather than performing the optimised combination of on- and off-diagonal modes is also used in the estimator in the local plane-parallel formalism given in the next section.

7. – Power spectrum estimators in the local plane-parallel formalism

Following the ethos behind the FKP power spectrum estimator, in the local plane-parallel approximation, we can define as the statistic that we want to reduce the data to as the order-n power-law moments of the window-convolved power spectrum [14]

$$P'_n(k) = \frac{1}{4\pi} \int \mathrm{d}\Omega_k \int \mathrm{d}^3 s_1 \int \mathrm{d}^3 s_2 \, (\hat{\mathbf{k}} \cdot \hat{\mathbf{s_1}})^n F(\mathbf{s_1}) F(\mathbf{s_2}) e^{i\mathbf{k}\cdot(\mathbf{s}_1 - \mathbf{s}_2)} - P_s, \tag{19}$$

where we have assumed that the LOS of the pair of galaxies lies along direction $\mathbf{s}_1$, $\mathrm{d}\Omega_k$ is the solid angle element in k-space, and we denote the window convolved power spectrum P'. P_s is the shot noise term. The local plane-parallel approximation is used to both define a single LOS to each pair of galaxies, and to define that LOS as the direction to one of the galaxies.

By writing the LOS in terms of only $\mathbf{s}_1$, we can split the integrals in eq. (19), such that

$$P'_n(k) = \frac{1}{4\pi} \int \mathrm{d}\Omega_k A'_n(\mathbf{k})[A'_0(\mathbf{k})]^* - P_s, \tag{20}$$

where

$$A'_n(\mathbf{k}) = \int \mathrm{d}^3 s\, F(\mathbf{s})(\hat{\mathbf{k}} \cdot \hat{\mathbf{s}})^n e^{i\mathbf{k}\cdot\mathbf{s}}. \tag{21}$$

In refs. [15] and [16] it was shown that these can be solved using Fast Fourier Transforms (FFTs) on a Cartesian grid after substituting the trivial decomposition

$$\hat{\mathbf{k}} \cdot \hat{\mathbf{s}} = \frac{k_x s_x + k_y s_y + k_z s_z}{ks} \tag{22}$$

into eq. (21). Recently, in ref. [17] a Legendre polynomial decomposition of $\hat{\mathbf{k}} \cdot \hat{\mathbf{s}}$ was proposed which allowed fewer FFTs to be used than in the simple approach above. Many FFT libraries are available, with FFTW being a commonly used example (`http://www.fftw.org/`). Thus this statistic can be quickly measured for a given catalogue.

8. – Grid assignment, aliasing and interlacing

When using FFTs to Fourier transform $F(\mathbf{s})$ (eq. (2)) we must sample $F(\mathbf{s})$ on a grid. Small-scale modes, unresolved due to the finite grid can alias large-scale modes, leading to the wrong power spectrum measurement. One way to avoid this is to use direct summation rather then FFT for the Fourier transform. However this is slow for large galaxy surveys, particularly when applied to a large random catalogue.

The effect of aliasing can be reduced through the choice of grid assignment scheme to interpolate $F(\mathbf{s})$ onto the regular grid, and by use of interlacing. In ref. [18] an excellent review of these issues is provided, and is the source for the ideas presented in this section. Considering calculating $F(\mathbf{s})$ at the grid points $\mathbf{s}_j$, we see that we can assign galaxies to $n(\mathbf{s})$ as given by

$$n(\mathbf{s}_j) = \frac{1}{H^3} \sum_{i=1}^{N_{\mathrm{gal}}} W^{(p)}(\Delta s_x/H)\, W^{(p)}(\Delta s_y/H)\, W^{(p)}(\Delta s_z/H), \tag{23}$$

where $\Delta\mathbf{s} = \mathbf{s}_j - \mathbf{s}_i = (\Delta s_x, \Delta s_y, \Delta s_z)$, and H is the grid size. If a random catalogue is used to define the survey mask, then the assignment of these points to the grid follows the same procedure.

It is common to consider piecewise polynomial functions for the 1-dimensional functions $W^{(p)}$, which simply correspond to convolving a top-hat function with itself $(p-1)$ times:

Nearest Grid Point (NGP)

$$W^{(1)}(t) = \begin{cases} 1, & \text{for } |t| < \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{24}$$

Cloud-In-Cell (CIC)

$$W^{(2)}(t) = \begin{cases} 1 - |t|, & \text{for } |t| < \frac{1}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{25}$$

Triangular Shaped Cloud (TSC)

$$W^{(3)}(t) = \begin{cases} \frac{3}{4} - t^2, & \text{for } |t| < \frac{1}{2}, \\ \frac{1}{2}\left(\frac{3}{2} - |t|\right)^2, & \text{for } \frac{1}{2} \le |t| < \frac{3}{2}, \\ 0, & \text{otherwise.} \end{cases} \tag{26}$$

Piecewise Cubic Spline (PCS)

$$W^{(4)}(t) = \begin{cases} \frac{1}{6}(4 - 6\,t^2 + 3\,|t|^3), & \text{for } 0 \le |t| < 1, \\ \frac{1}{6}(2 - |t|)^3, & \text{for } 1 \le |t| < 2, \\ 0, & \text{otherwise.} \end{cases} \tag{27}$$

The interpolation function acts as a convolution in configuration-space, and hence is a multiplicative factor in Fourier space and can be removed by dividing $F(\mathbf{k})$ by the Fourier transform of the window, $W^{(p)}(k_x)\,W^{(p)}(k_y)\,W^{(p)}(k_z)$, where

$$W^{(p)}(k) = \left[\frac{\sin(kH/2)}{(kH/2)}\right]^p. \tag{28}$$

While this corrects for the magnitude of the convolution its effects live on in the amplitude of the aliasing effect.

A method to partially correct for aliasing based on the interlacing of two grids is discussed in the classic text of [19]. The key idea is to perform an additional, configuration-space interpolation onto a grid shifted by $H/2$ in all spatial directions, and then take the average of the two

$$F(\mathbf{k}) = \frac{1}{2}\left[F_1(\mathbf{k}) + F_2(\mathbf{k})\right], \tag{29}$$

where $F_1(\mathbf{k})$ and $F_2(\mathbf{k})$ represent the individual transforms. This removes the leading-order aliasing terms.

9. – Linking Fourier and Fourier-Bessel bases

The link between a Fourier-space decomposition and a decomposition into a basis consisting of spherical harmonics and spherical Bessel functions (hereafter known as a Fourier-Bessel basis) is given by the Rayleigh expansion of a plane wave. In terms of Legendre polynomials this is written

$$e^{i\mathbf{k}\cdot\mathbf{s}} = \sum_\ell i^\ell (2\ell+1) j_\ell(ks)\mathcal{L}_\ell(\hat{\mathbf{k}}\cdot\hat{\mathbf{s}}), \tag{30}$$

and in terms of Spherical Harmonics

$$e^{i\mathbf{k}\cdot\mathbf{s}} = 4\pi \sum_{\ell m} i^\ell j_\ell(ks) Y_{\ell m}(\hat{\mathbf{k}}) Y^*_{\ell m}(\hat{\mathbf{s}}). \tag{31}$$

Using this, we can see how eq. (21) can be solved using a Fourier-Bessel basis. As explained by [20], consider differentiating eq. (31) with respect to ks n times to give

$$i^n(\hat{\mathbf{k}}\cdot\hat{\mathbf{s}})^n e^{iks(\hat{\mathbf{k}}\cdot\hat{\mathbf{s}})} = 4\pi \sum_{\ell m} i^\ell j_\ell^{(n)}(ks) Y_{\ell m}(\hat{\mathbf{k}}) Y^*_{\ell m}(\hat{\mathbf{s}}), \tag{32}$$

where $j_\ell^{(n)}(ks)$ is the n-th derivative of the spherical Bessel function with respect to ks.

In ref. [20] the trick is used of taking the derivative of the plane wave expansion to directly link power spectrum multipoles to a Fourier-Bessel decomposition. This was applied to the linear model RSD to write $\delta_{\ell m}(k)$ in terms of expansions in spherical Bessel functions and their derivatives. The work of [20] was applied by [21, 22] to investigate the impact of wide-angle effects from Fourier-based multipole measurements.

To see how this trick can allow us to measure the statistic in eq. (21), using a Fourier-Bessel rather than Fourier basis, we substitute the plane-wave expansion into eq. (21) to give

$$A'_n(\mathbf{k}) = 4\pi \int \mathrm{d}^3 s F(\mathbf{s}) \sum_{\ell m} i^{\ell-n} j_\ell^{(n)}(ks) Y_{\ell m}(\hat{\mathbf{k}}) Y^*_{\ell m}(\hat{\mathbf{s}}). \tag{33}$$

If we also expand $F(\mathbf{s})$ in a basis of Spherical Harmonics and the n-th derivative of spherical Bessel functions

$$F_{\ell m}^{(n)}(k) = \int \mathrm{d}^3 s\, F(\mathbf{s}) Y_{\ell m}^{\star}(\hat{\mathbf{s}}) j_{\ell}^{(n)}(ks), \tag{34}$$

then we can write $A_n(\mathbf{k})$ in terms of $F_{\ell m}^{(n)}(k)$ as

$$A_n'(\mathbf{k}) = 4\pi \sum_{\ell m} i^{\ell-n} F_{\ell m}^{(n)}(k) Y_{\ell m}(\hat{\mathbf{k}}). \tag{35}$$

To intuitively see how the Fourier and Fourier-Bessel solutions for $A_n(\mathbf{k})$ are related, note that derivatives of spherical Bessel functions can be rewritten as the difference between standard spherical Bessel functions. For example

$$j_{\ell}^{(1)}(ks) = -j_{\ell+1}(ks) + \frac{\ell}{ks} j_{\ell}(ks). \tag{36}$$

Subsequent derivatives can be related to standard spherical Bessel functions by recursive application of this formula. As $\ell/(ks)$ is equal to $k_{\perp}/k$ defined locally at position $\mathbf{s}$ for modes of wavenumber ℓ in a Fourier-Bessel function decomposition, we can see that taking the derivatives in eq. (32) is directly related to the local multipole expansion as considered in eq. (22). *I.e.* instead of using the derivatives of the spherical Bessel functions to get eq. (35), it would have instead been possible to use the local definition of $\hat{\mathbf{k}} \cdot \hat{\mathbf{s}}$ in the integral to get the same result.

Substituting eq. (35) for n and $n = 0$ into eq. (20), and using the orthogonality relations for spherical harmonics removes the angular integral and gives the simple result that

$$P_n'(k) = 4\pi \sum_{\ell m} F_{\ell m}^{(n)}(k) [F_{\ell m}^{(0)}(k)]^* - P_s. \tag{37}$$

Thus we see that we can use either Fourier or Fourier-Bessel bases to measure $P_n'(k)$ from a galaxy redshift survey under the local plane-parallel approximation. The use of a Fourier-Bessel basis does not alleviate wide-angle effects: these are built in to the power spectrum multipole definition (eq. (19)). Indeed, wide-angle effects are fundamentally model dependent, so we cannot remove them completely using a model-independent or fiducial-model based estimator.

Given the same end point for the measurements, the Fourier approach is preferred as it allows the use of FFTs to perform the transforms required, saving computational resources. Using a Fourier-Bessel basis may help when calculating the cross-power spectrum of a galaxy redshift survey with an angular survey.

10. – Window convolution of models

Given that we measure the power spectrum convolved with the window function, we need a fast mechanism to convolve models to be compared to $P'_n(k)$. In ref. [23] it was shown that we can perform the required 3D convolution quickly using the Hankel transform relation between the window-convolved multipole moments in configuration and Fourier space,

$$P'_\ell(k) = 4\pi i^\ell \int \mathrm{d}s\, s^2 \xi'_\ell(s) j_\ell(sk). \tag{38}$$

Crucially, this equation holds for both the unconvolved and convolved power spectrum and correlation function pairs in both the global plane-parallel [23] and local plane-parallel [24] limits. The Hankel tranform can be quickly solved using a 1D FFT, although care has to be taken given the oscillatory nature of the integrand, as described in [25].

By using this transform we can calculate the convolved model power spectra using multiplications in real space. The Legendre moments of the convolved correlation function are defined

$$\xi'_\ell(s) = \frac{2\ell+1}{4\pi} \int \mathrm{d}\Omega_s\, \xi(\mathbf{s}) W^2(\mathbf{s}) \mathcal{L}_\ell(\hat{\mathbf{s}} \cdot \hat{\mathbf{x}}_1), \tag{39}$$

where $\xi(\mathbf{s})$ and $W^2(\mathbf{s})$ are the anisotropic correlation function and window function. $\mathcal{L}_\ell$ is the Legendre polynomial of order ℓ, here written as a function of the LOS to one galaxy in each pair $\mathbf{x}_1$, matching the assumptions of eq. (19).

We define the moments of the window function as

$$W^2_p(s) = \frac{2p+1}{4\pi} \int \mathrm{d}\Omega_s \int \mathrm{d}\mathbf{x}_1 W(\mathbf{x}_1) W(\mathbf{x}_1 + \mathbf{s}) \mathcal{L}_p(\mu_s), \tag{40}$$

which can be calculated by using the random catalogue to perform the integrations with a Monte Carlo based technique. Substituting this into eq. (39), and also expanding the unconvolved correlation function in Legendre moments means that we can rewrite this equation as

$$\xi'_\ell(s) = (2\ell+1) \sum_L \xi_L(s) \sum_p \frac{1}{2p+1} W^2_p(s) a^p_{L\ell}, \tag{41}$$

where $a^p_{L\ell}$ are the solutions to the equation $\mathcal{L}_\ell(\mu)\mathcal{L}_L(\mu) = \sum_p a^p_{\ell L} \mathcal{L}_p(\mu)$, and can be obtained by substituting in the polynomials and equating powers of μ. The window convolution spreads the linear information to $\xi'_\ell(s)$ with $\ell > 4$ and the $\xi'_\ell(s)$ modes with $\ell \leq 4$ depend on higher-order moments. However, it is common to only fit to the first three even multipoles, ignoring the potential information at higher orders, and furthermore only apply the window convolution to the linear model, which is reasonable

as the window effect diminishes to smaller scales. Expanding eq. (41) gives that the relevant expansion components are

$$\xi'_0(s) = \xi_0 W_0^2 + \frac{1}{5}\xi_2 W_2^2 + \frac{1}{9}\xi_4 W_4^2, \tag{42}$$

$$\xi'_2(s) = \xi_0 W_2^2 + \xi_2\left[W_0^2 + \frac{2}{7}W_2^2 + \frac{2}{7}W_4^2\right] + \xi_4\left[\frac{2}{7}W_2^2 + \frac{100}{693}W_4^2 + \frac{25}{143}W_6^2\right], \tag{43}$$

$$\begin{aligned}\xi'_4(s) = {} & \xi_0 W_4^2 + \xi_2\left[\frac{18}{35}W_2^2 + \frac{20}{77}W_4^2 + \frac{45}{143}W_6^2\right] \\ & + \xi_4\left[W_0^2 + \frac{20}{77}W_2^2 + \frac{162}{1001}W_4^2 + \frac{20}{143}W_6^2 + \frac{490}{2431}W_8^2\right],\end{aligned} \tag{44}$$

keeping ξ_ℓ terms with $\ell \leq 4$, and including all of the relevant window multipole moments. This matches the premise that linear theory is complete to $\ell = 4$, but the window function has no such constraint. Given $\xi'_\ell(s)$, the model power spectrum multipoles can be calculated using eq. (38).

11. – Power spectrum integral constraint

This formalism for the window also makes it easy to see how the integral constraint can be included in models [24]. For the power spectrum, the integral constraint is relevant because, to formulate $F(\mathbf{s})$ as in eq. (2), we have matched $\langle n(\mathbf{s})\rangle$ to the actual observed density of galaxies. We can assume that the variations in the expected distribution of $\langle n(\mathbf{s})\rangle$ as a function of $\mathbf{s}$ are known, but that the normalisation is incorrect, so

$$\langle n(\mathbf{s})\rangle_{\rm assumed} = (1+C)\langle n(\mathbf{s})\rangle_{\rm true}, \tag{45}$$

where C is a constant. The multiplicative nature of C with $\langle n(\mathbf{s})\rangle$ means that the constant is inside the window function convolution and is equivalent to an additive contribution to ξ_0. We ignore the possibility that a mistake has also been made in the calculation of I.

The size of the correction depends on how the randoms have been matched to the galaxies. If the total number of weighted pairs has been matched, then we have forced that $P'_0(0) = 0$. In this case, the model to be compared to the data is

$$P'_{\ell,\text{ic-corrected}}(k) = P'_\ell(k) - \frac{P'_0(k=0)}{W_0^2(k=0)}W_\ell^2(k), \tag{46}$$

where

$$W_\ell^2(k) = 4\pi\int {\rm d}s\, s^2 W_\ell^2(s) j_\ell(sk). \tag{47}$$

Considering this from a different stand-point, by defining α by matching the total number of weighted pairs in the estimator, we have a simple expression for the integral constraint

to be included in the model to match the measurement. Thus this is the preferred method for calculating α.

12. – Covariance matrix under Gaussian assumption

In order to make statistical inferences from the measured power, we need to model the distribution from which it is drawn. It is common to assume that the power spectrum multipoles are drawn from a multi-variate Gaussian population, in which case the Likelihood for the power spectrum is

$$\mathcal{L}(\mathbf{x}|\mathbf{p}, \Psi_{\rm true}) = \frac{|\Psi_{\rm true}|}{\sqrt{2\pi}} \exp\left[-\frac{1}{2}\chi^2(\mathbf{x}, \mathbf{p}, \Psi_{\rm true})\right] , \tag{48}$$

where

$$\chi^2(\mathbf{x}, \mathbf{p}, \Psi_{\rm true}) \equiv \left[\mathbf{x}^d - \mathbf{x}(\mathbf{p})\right] \Psi_{\rm true} \left[\mathbf{x}^d - \mathbf{x}(\mathbf{p})\right] . \tag{49}$$

In the example considered in these notes, the data $\mathbf{x}^d$, and model for the data $\mathbf{x}(\mathbf{p})$, would be the power spectra (or correlation function), with the parameter $\mathbf{p}$ being the cosmological parameters of interest. $\Psi_{\rm true}$ is the true inverse covariance matrix.

For power spectrum measurements the covariance is

$$\mathrm{Cov}\left[P'_n(k_i), P'_{n'}(k_j)\right] = \langle P'_n(k_i)\, P'_{n'}(k_j)\rangle - \langle P'_n(k_i)\rangle\langle P'_{n'}(k_j)\rangle, \tag{50}$$

where $P'_n(k_i)$ is the window-convolved power-law moment of the power spectrum, calculated in the local plane-parallel approximation, and binned into k-bin i. Each of the window convolution (including weighting), local plane-parallel geometry, power-law moment and binning effects will complicate the covariance from the simple form without these, which is given by

$$\mathrm{Cov}\left[P(\mathbf{k}),\, P(\mathbf{k}')\right] = \frac{2}{V}\left[P(\mathbf{k}) + \frac{1}{\bar{n}}\right]^2 \delta^D(\mathbf{k} - \mathbf{k}'). \tag{51}$$

Here, δ^D is the Dirac delta function and the shot noise term assumes that the galaxies Poisson sample the underlying matter field. V is the volume of the survey. This expression follows from Wick's theorem for a Gaussian random field with zero mean

$$\langle\delta_1\,\delta_2\,\delta_3\,\delta_4\rangle = \langle\delta_1\,\delta_2\rangle\langle\delta_3\,\delta_4\rangle + \langle\delta_1\,\delta_3\rangle\langle\delta_2\,\delta_4\rangle + \langle\delta_1\,\delta_4\rangle\langle\delta_2\,\delta_3\rangle, \tag{52}$$

and the standard expression for the variance of a Poisson sampling.

For the local plane-parallel power spectrum estimator of eq. (19), applying Wick's theorem to $A'_n(\mathbf{k})$ leads to

$$\begin{aligned}\mathrm{Cov}\left[P'_n(k_i), P'_{n'}(k_j)\right] = \frac{1}{(4\pi)^2}\int_i \mathrm{d}\Omega_k \int_j \mathrm{d}\Omega_{k'} \langle A'_n(\mathbf{k}) A'_{n'}(\mathbf{k}')^*\rangle\langle A'_0(\mathbf{k}) A'_0(\mathbf{k}')^*\rangle \\ + \langle A'_n(\mathbf{k}) A'_0(\mathbf{k}')^*\rangle\langle A'_0(\mathbf{k}) A'_{n'}(\mathbf{k}')^*\rangle.\end{aligned} \tag{53}$$

No way of writing this equation in a form that allows its calculation using FFTs and Hankel transforms has yet been found without also applying simplifying assumptions. A common assumption to make is that the power spectrum is constant over the extent of the window function, so that the convolution breaks down [3, 26, 27]. In ref. [27] it was shown that this results in a form for the covariance that can be solved using only FFTs.

Because of these complications it is common to estimate the covariance matrix in a brute-force way, from a large set of mock catalogues that match the survey geometry and analysis method

$$\mathrm{Cov}\left[P'_n(k_i), P'_{n'}(k_j)\right] = \frac{1}{N_m - 1}\sum_{m=1}^{N_m}\left[P'_{n,m}(k_i) - \overline{P'}_n(k_i)\right]\left[P'_{n',m}(k_j) - \overline{P'}_{n'}(k_j)\right], \tag{54}$$

where the sum is over N_m mock catalogues, and $\overline{P'}$ is the mean power spectrum over those mocks. This method automatically includes all of the linear effects discussed above, and can also include non-linear effects to the extent that they are included in the method used to create the mocks, and the Gaussian assumption holds. Given fast methods for creating the mocks and complications due to non-linear shot-noise, galaxy bias and Redshift Space Distortions (RSD) there are many advantages to such an approach.

The disadvantages include that this covariance matrix lacks fluctuations caused by k-modes larger than the simulation box —so-called supersample covariance [28]. This can be included by adjusting the mocks to force each to have slightly different cosmological parameters [26].

An additional problem is that the estimate of the covariance matrix has errors that systematically distort the likelihood. Formally, eq. (54) gives an estimate drawn from a Wishart distribution. We can allow for the bias induced using a perturbative analysis [29, 30], which has to be adjusted if parameter errors are estimated from the likelihood surface [31]. Alternatively, as demonstrated by [32], one should perform a joint likelihood analysis of both the power spectrum and covariance matrix. With the assumption of a Jeffreys prior allowing us to use Bayes theorem to determine the distribution of the true covariance matrix given our estimate from mocks, we need to adjust the likelihood of eq. (48) to

$$\mathcal{L}(\mathbf{x}^d|\mathbf{x}(\mathbf{p}), \Psi) \propto \left[1 + \frac{[\mathbf{x}^d - \mathbf{x}(\mathbf{p})]\Psi[\mathbf{x}^d - \mathbf{x}(\mathbf{p})]}{N_m - 1}\right]^{-\frac{N_m}{2}}. \tag{55}$$

This offers a neater and statistically more rigorous method for correcting for the approximate covariance matrix of eq. (54) compared with the perturbative solution.

13. – 1-point systematics

The measurement of the power spectrum from a galaxy survey is likely to be contaminated by systematic effects that alter the observed galaxy density such that the fluctuations are not driven only by astrophysical processes. Due to the way most galaxy

redshift surveys to date have been constructed by spectroscopic follow-up observation of targets selected from imaging data, it is natural to think that there is a split in contaminants between angular and radial directions. However, this is not necessarily the case: for example, removing faint targets in a patch of the sky would tend to remove high-redshift galaxies in an apparent magnitude-limited sample. Consequently, we do not make any such separation here.

Using the observed target distribution, coupled with maps of the distribution of causes of potential problems —for example, imaging depth maps, or maps of bright star locations— one can look for fluctuations in target density. We expect the cosmological fluctuations to be independent of the systematics, and so any statistically significant correlation is indicative of a problem in the sample. In refs. [33, 34] a careful analysis of potential systematics in BOSS is under taken and a set of multiplicative weights is developed that correct the galaxy density for these fluctuations, in effect creating a different window for the galaxies compared to the random catalogue, so that eq. (2) is changed to

$$F(\mathbf{s}) = \frac{w(\mathbf{s})}{I^{1/2}}[w_{\rm sys} n(\mathbf{s}) - \langle n(\mathbf{s})\rangle], \tag{56}$$

where $w_{\rm sys}$ are the systematic weights. These weights increase the noise in our estimator. This would be reduced by weighting instead the randoms to match the galaxies [35]. To see this, consider the toy example of a Poisson distribution with mean and variance N, weighted by a set of weights with variance σ_w^2. The variance of the weighted sample is $N(1+\sigma_w^2)$, and so is always greater than the unweighted field.

Weighting the randoms rather than the galaxies also shows that this multiplicative weighting is equivalent to an additive contaminant [36]

$$F(\mathbf{s}) = \frac{w(\mathbf{s})}{I^{1/2}}[n(\mathbf{s}) - \langle n(\mathbf{s})\rangle] + f_{\rm sys}(\mathbf{s}). \tag{57}$$

The benefit of writing the effect like this is that we can equate the weighting applied to correct for systematics to the mode deprojection technique. Mode deprojection works by setting the covariance of modes to be removed to be infinite in the covariance matrix of the unbinned power

$$\mathrm{Cov}\,[P'(\mathbf{k}), P'(\mathbf{k}')] \to \mathrm{Cov}\,[P'(\mathbf{k}), P'(\mathbf{k}')] + \lim_{\sigma\to\infty} \sigma f_{\rm sys}(\mathbf{k}) f_{\rm sys}(\mathbf{k}')^*. \tag{58}$$

In ref. [37] it was shown that this is mathematically equivalent to weighting the randoms, modulo making the correct normalisation when calculating the power spectrum. The standard procedure can easily be modified to include the required renormalisation when measuring the power —adjusting the effective number of modes as required.

It is possible to extend these ideas to remove multiple contaminants, and to remove sets of contaminants that span a space where systematic errors are suspected. However, as discussed in [36], the problem in general lies not in removing the contaminants, but knowing which modes are affected: the removal of contaminants only works for known unknowns, and fails for unknown unknowns.

14. – 2-point systematics

We now consider a common problem induced by the mechanics of fiber-fed multi-object spectrographs. Most have a physical limit on how close the ends of fibers can be placed in the focal plane of the telescope such that they cannot observe close targets in a single pass of the instrument on the sky. Thus there is a geometrical difference between the samples selected for observation and the parent sample from which it was selected. Furthermore, in a sample of galaxies, close pairs tend to have a higher bias as they are located in higher mass haloes compared with isolated galaxies. Thus the lack of close pairs of galaxies due to fibre collisions changes the clustering between observed and parent samples even at large separations due to the change in mean bias. This would not be a problem if this difference occurred uniformly across the sky as we would then be simply selecting a lower bias galaxy population compared with the full target sample. However, this lack of pairs is often avoided in regions of overlapping observations, leading to an anisotropic mean bias in the observed sample. In addition, obviously, the small-scale clustering is strongly affected as we lose small-separation pairs, such that there are no angular pairs with separation smaller than the instrumental cut-off in 1-pass regions.

The problem described above is inherently of higher order than the issues raised in sect. **13**. In that section, we considered issues that were equivalent to changing the window through which the survey was observed. In contrast, close-pair effects are 2-point in origin as they depend on the overdensity at two positions: we cannot observe one galaxies if there is another nearby. To correct for these, we cannot easily use the techniques described in sect. **13**. Surveys such as DESI [38, 39] have more complicated, but related problems due to experimental limitations on how the fibres can be placed in the focal plane of the telescope.

In ref. [40] the Pairwise Inverse Probability (PIP) method is proposed to correct for the kind of 2-point systematic arising from hardware limitations. For this we need a complete parent sample from which the observed subsample is selected. The PIP method estimates the probability that a pair of objects can be observed by counting how many times it is observed in a set of possible surveys selected for observation from the parent sample. This set can be created by translating or rotating the survey, or by rerunning the algorithm used to select the observed subsample with different randomly chosen priorities for different objects. The key thing is that all surveys in the set are equally likely, and the number of pairs of objects in the parent sample that are never observed in any realisation of the survey is negligible.

By weighting each pair by the inverse of this probability when counting pairs in order to estimate the correlation function, we recover pair counts with the same expected value as those of the full parent sample. As an example, consider the situation where close pairs are only observed in parts of the survey. In order to ensure that there are no zero-probability pairs in the parent catalogue, we need to move the survey when determining the set of samples, so that any close-pair in the parent has a chance of falling into an overlap region in some surveys in the set. Because they are only observed in selected regions, close-pairs will be given a lower probability than wide-separation pairs, leading

the counts to be upweighted in the sums for any realisation, correcting for this effect.

One issue with the technique is the time it takes to perform the calculation. For a galaxy survey with 10^7 galaxies, there are 10^{14} pairs, and if we create a set of 10^3 possible survey realisations, the PIP calculation is of order 10^{17}. The computational burden can be minimised by calculating the weights on-the-fly while pair counting to estimate the correlation function, based on storing the selection of galaxies in each survey in a bit-wise way (it is a yes/no decision on whether each object in the parent makes it into a particular survey). The weight can then be quickly calculated using a bit-wise sum [40].

In ref. [41] it is shown how the angular clustering measurement in the parent can be incorporated into the method to improve signal, while in [42] and [43] it was shown that the method works for the Dark Energy Spectroscopic Survey mocks and the VIMOS Public Extragalactic Redshift Survey (VIPERS), respectively.

A similar method to debias power spectrum measurements using only FFTs and Hankel transforms has yet to be developed, and it would also be useful to have a method to correct the overdensity field as used in reconstruction [44] for such effects.

15. – Binning in redshift and redshift-dependent weighting

Future surveys such as DESI [38,39] and Euclid [45] will cover a wide range in redshift, such that there will be significant evolution in the populations of galaxies observed. Thus we either need to allow for this evolution when analysing the data, or divide the survey in redshift prior to analysis. Dividing galaxies based on their redshifts into shells will tend to miss pairs of galaxies where galaxies are in different bins. It would also be possible to split instead by radial pair-centre rather than galaxy position, which also mitigates for the effect of the window on RSD measurements [46].

An alternative is to perform multiple analyses of the full sample using sets of weights optimised to measure the evolving quantities of interest. *I.e.* binning can be seen as using a set of top-hat weights in redshift, and this is not necessarily the optimal choice. Using Fisher matrix based techniques, one can find sets of weights optimised for BAO [47], RSD [48], and primordial non-Gaussianity [49] measurements.

These ideas were recently applied to the quasar sample in the extended-Baryon Oscillation Spectroscopic Survey (eBOSS), recovering BAO and RSD based measurements on evolving parameters using multiple analyses with different sets of weights [50-53].

16. – Reconstruction

While the bulk motion of material in the Universe drives structure growth, it also acts to smooth the primordial overdensity field, leading to the degradation of the BAO feature on small scales. The basic idea behind reconstruction is to move the late-time over-densities back to their initial positions, sharpening the BAO peak [44]. In terms of information, the bulk motion moves the small-scale 2-point information into higher-order terms, and reconstruction recovers this information [54].

Reconstruction requires us to know the displacement field linking Eulerian and Lagrangian positions. This displacement field can be approximated using the standard Zeldovich displacements

$$\nabla \cdot \boldsymbol{\Psi} + \frac{f}{b} \nabla \left(\boldsymbol{\Psi} \cdot \hat{\mathbf{r}} \right) \hat{\mathbf{r}} = -\frac{\delta_g}{b} \,, \tag{59}$$

where $\boldsymbol{\Psi}$ is the displacement field, $f = \mathrm{d} \ln D(a) / \mathrm{d} \ln a$, $D(a)$ is the linear growth rate, and σ_8 normalises the amplitude of the linear power spectrum.

Solving eq. (59) is complicated by the LOS-dependent RSD term, and that the RSD field has a non-zero curl component. Two approaches have been proposed: solving this equation on a grid spanning the survey using finite-difference techniques [55], and a FFT based technique that iteratively solves for the LOS dependence [56]. Both of these techniques have been successfully applied to the analysis of data (*e.g.* [57]). Equation (59) is solved after smoothing the observed field in order to focus on the large-scale bulk motions. This changes the shape of the recovered power spectrum, requiring changes in BAO fitting routines [58].

It is easy to imaging that we can do better than solving eq. (59) as a way to recover the bulk motions. For example there is extra information available —such as that the initial distribution of over-densities was homogeneous. The development of algorithms to reconstruct the initial density field from an evolved field has a long history (*e.g.* [59]), stretching back even before the improvement of BAO observations was considered. Many methods have been proposed. Recent highlights include: Iterative reconstruction [60,61], which removes the need to specify a smoothing scale. More complicated schemes have been proposed based on limiting the information used, allowing perturbation-theory-based solutions [62-64]. The extension of such methods to biased tracers [65] and including RSD [66], have also been recently considered. Clearly, for the next generation of experiment, reconstruction will be improved compared to the algorithms used for BOSS, and we will have many methods to choose from.

17. – Conclusions

This update on the lecture notes I provided 5 years ago at a previous graduate school in Varenna [1] clearly shows that the best analysis method applied to measure clustering in galaxy surveys continues to change, with better techniques being developed alongside improvements in the experiments themselves. The techniques being used in analyses now are very different and more robust than those used 5 years ago, and are resulting in more accurate measurements. Given the excellent data becoming available in the next few years from DESI and Euclid, it is clear that there is a strong driver for techniques to continue to be developed. I look forward to writing the next set of notes in 5 years time, if I am invited back to lecture again at a Varenna school.

* * *

The author acknowledges useful conversations with Davide Bianchi, Cullan Howlett and Faizan Mohammad.

REFERENCES

[1] PERCIVAL W. J., ArXiv/1312.5490 (2013).
[2] PLANCK COLLABORATION *et al.*, ArXiv/1807.06211 (2018).
[3] FELDMAN H. A., KAISER N. and PEACOCK J. A., *Astrophys. J.*, **426** (1994) 23.
[4] PERCIVAL W. J., VERDE L. and PEACOCK J. A., *Mon. Not. R. Astron. Soc.*, **347** (2004) 645.
[5] SELJAK U., HAMAUS N. and DESJACQUES V., *Phys. Rev. Lett.*, **103** (2009) 091303.
[6] KAISER N., *Mon. Not. R. Astron. Soc.*, **227** (1987) 1.
[7] BLAKE C. *et al.*, *Mon. Not. R. Astron. Soc.*, **406** (2010) 803.
[8] DAWSON K. S. *et al.*, *Astron. J.*, **145** (2013) 10.
[9] SZAPUDI I., *Astrophys. J.*, **614** (2004) 51.
[10] PÁPAI P. and SZAPUDI I., *Mon. Not. R. Astron. Soc.*, **389** (2008) 292.
[11] ALCOCK C. and PACZYNSKI B., *Nature*, **281** (1979) 358.
[12] LANDY S. D. and SZALAY A. S., *Astrophys. J.*, **412** (1993) 64.
[13] TEGMARK M., HAMILTON A. J. S., STRAUSS M. A., VOGELEY M. S. and SZALAY A. S., *Astrophys. J.*, **499** (1998) 555.
[14] YAMAMOTO K., NAKAMICHI M., KAMINO A., BASSETT B. A. and NISHIOKA H., *Publ. Astron. Soc. Jpn.*, **58** (2006) 93.
[15] BIANCHI D., GIL-MARÍN H., RUGGERI R. and PERCIVAL W. J., *Mon. Not. R. Astron. Soc.*, **453** (2015) L11.
[16] SCOCCIMARRO R., *Phys. Rev. D*, **92** (2015) 083532.
[17] HAND N., LI Y., SLEPIAN Z. and SELJAK U., *J. Cosmol. Astropart. Phys.*, **7** (2017) 002.
[18] SEFUSATTI E., CROCCE M., SCOCCIMARRO R. and COUCHMAN H. M. P., *Mon. Not. R. Astron. Soc.*, **460** (2016) 3624.
[19] HOCKNEY R. W. and EASTWOOD J. W., *Computer Simulation Using Particles* (Taylor and Francis) 1981.
[20] DATTA K. K., CHOUDHURY T. R. and BHARADWAJ S., *Mon. Not. R. Astron. Soc.*, **378** (2007) 119.
[21] CASTORINA E. and WHITE M., *Mon. Not. R. Astron. Soc.*, **476** (2018) 4403.
[22] CASTORINA E. and WHITE M., *Mon. Not. R. Astron. Soc.*, **479** (2018) 741.
[23] WILSON M. J., PEACOCK J. A., TAYLOR A. N. and DE LA TORRE S., *Mon. Not. R. Astron. Soc.*, **464** (2017) 3121.
[24] BEUTLER F. *et al.*, *Mon. Not. R. Astron. Soc.*, **466** (2017) 2242.
[25] HAMILTON A. J. S., *Mon. Not. R. Astron. Soc.*, **312** (2000) 257.
[26] HOWLETT C. and PERCIVAL W. J., *Mon. Not. R. Astron. Soc.*, **472** (2017) 4935.
[27] BLAKE C., CARTER P. and KODA J., *Mon. Not. R. Astron. Soc.*, **479** (2018) 5168.
[28] TAKADA M. and HU W., *Phys. Rev. D*, **87** (2013) 123504.
[29] DODELSON S. and SCHNEIDER M. D., *Phys. Rev. D*, **88** (2013) 063537.
[30] TAYLOR A., JOACHIMI B. and KITCHING T., *Mon. Not. R. Astron. Soc.*, **432** (2013) 1928.
[31] PERCIVAL W. J. *et al.*, *Mon. Not. R. Astron. Soc.*, **439** (2014) 2531.
[32] SELLENTIN E. and HEAVENS A. F., *Mon. Not. R. Astron. Soc.*, **456** (2016) L132.
[33] ROSS A. J. *et al.*, *Mon. Not. R. Astron. Soc.*, **417** (2011) 1350.
[34] ROSS A. J. *et al.*, *Mon. Not. R. Astron. Soc.*, **424** (2012) 564.

[35] BAUTISTA J. E. *et al.*, *Astrophys. J.*, **863** (2018) 110.
[36] KALUS B., PERCIVAL W. J., BACON D. J., MUELLER E.-M., SAMUSHIA L., VERDE L., ROSS A. J. and BERNAL J. L., *Mon. Not. R. Astron. Soc.*, **482** (2019) 453.
[37] KALUS B., PERCIVAL W. J., BACON D. J. and SAMUSHIA L., *Mon. Not. R. Astron. Soc.*, **463** (2016) 467.
[38] DESI COLLABORATION *et al.*, ArXiv/1611.00036 (2016).
[39] DESI COLLABORATION *et al.*, ArXiv/1611.00037 (2016).
[40] BIANCHI D. and PERCIVAL W. J., *Mon. Not. R. Astron. Soc.*, **472** (2017) 1106.
[41] PERCIVAL W. J. and BIANCHI D., *Mon. Not. R. Astron. Soc.*, **472** (2017) L40.
[42] BIANCHI D., BURDEN A., PERCIVAL W. J., BROOKS D., CAHN R. N., FORERO-ROMERO J. E., LEVI M., ROSS A. J. and TARLE G., *Mon. Not. R. Astron. Soc.*, **481** (2018) 2338.
[43] MOHAMMAD F. G. *et al.*, *Astron. Astrophys.*, **619** (2018) 17.
[44] EISENSTEIN D. J., SEO H.-J., SIRKO E. and SPERGEL D. N., *Astrophys. J.*, **664** (2007) 675.
[45] LAUREIJS R. *et al.*, ArXiv/1110.3193 (2011).
[46] NOCK K., PERCIVAL W. J. and ROSS A. J., *Mon. Not. R. Astron. Soc.*, **407** (2010) 520.
[47] ZHU F., PADMANABHAN N. and WHITE M., *Mon. Not. R. Astron. Soc.*, **451** (2015) 236.
[48] RUGGERI R., PERCIVAL W. J., GIL-MARÍN H., ZHU F., ZHAO G.-B. and WANG Y., *Mon. Not. R. Astron. Soc.*, **464** (2017) 2698.
[49] MUELLER E.-M., PERCIVAL W. J. and RUGGERI R., ArXiv/1702.050088 (2017).
[50] RUGGERI R., PERCIVAL W. J., MUELLER E.-M., GIL-MARIN H., ZHU F., PADMANABHAN N. and ZHAO G.-B., ArXiv/1712.03997 (2017).
[51] ZHU F. *et al.*, *Mon. Not. R. Astron. Soc.*, **480** (2018) 1096.
[52] WANG D. *et al.*, *Mon. Not. R. Astron. Soc.*, **477** (2018) 1528.
[53] RUGGERI R. *et al.*, *Mon. Not. R. Astron. Soc.*, **483** (2019) 3878.
[54] SCHMITTFULL M., FENG Y., BEUTLER F., SHERWIN B. and CHU M. Y., *Phys. Rev. D*, **92** (2015) 123522.
[55] PADMANABHAN N., XU X., EISENSTEIN D. J., SCALZO R., CUESTA A. J., MEHTA K. T. and KAZIN E., *Mon. Not. R. Astron. Soc.*, **427** (2012) 2132.
[56] BURDEN A., PERCIVAL W. J. and HOWLETT C., *Mon. Not. R. Astron. Soc.*, **453** (2015) 456.
[57] ALAM S. *et al.*, *Mon. Not. R. Astron. Soc.*, **470** (2017) 2617.
[58] SEO H.-J., BEUTLER F., ROSS A. J. and SAITO S., *Mon. Not. R. Astron. Soc.*, **460** (2016) 2453.
[59] NARAYANAN V. K. and CROFT R. A. C., *Astrophys. J.*, **515** (1999) 471.
[60] SCHMITTFULL M., BALDAUF T. and ZALDARRIAGA M., *Phys. Rev. D*, **96** (2017) 023505.
[61] HADA R. and EISENSTEIN D. J., *Mon. Not. R. Astron. Soc.*, **478** (2018) 1866.
[62] TASSEV S. and ZALDARRIAGA M., *J. Cosmol. Astropart. Phys.*, **10** (2012) 006.
[63] WANG X., YU H.-R., ZHU H.-M., YU Y., PAN Q. and PEN U.-L., *Astrophys. J. Lett.*, **841** (2017) L29.
[64] SHI Y., CAUTUN M. and LI B., *Phys. Rev. D*, **97** (2018) 023505.
[65] WANG X. and PEN U.-L., *Astrophys. J.*, **870** (2019) 116.
[66] ZHU H.-M., YU Y. and PEN U.-L., *Phys. Rev. D*, **97** (2018) 043502.

Proceedings of the International School of Physics "Enrico Fermi"
Course 200 "Gravitational Waves and Cosmology", edited by E. Coccia, J. Silk and N. Vittorio
(IOS, Amsterdam; SIF, Bologna) 2020
DOI 10.3254/ENFI200014

Nonlinear astrophysical probes of screened modified gravity

David F. Mota

Institute of Theoretical Astrophysics, University of Oslo
PO Box 1029 Blindern, 0315 Oslo, Norway

Summary. — We review the effects of modified gravity theories, in particular the symmetron and $f(R)$ gravity, on the nonlinear regime of structure formation. In particular, we investigate the velocity dispersion of galaxy clusters as a function of the halo masses, how the matter power spectra depends on the coupling, range, and screening scale of the fifth force, and on possible ways of detecting violations of the equivalence principle using the mass inferred via lensing methods *versus* the mass inferred via dynamical methods. Furthermore, we show how one could use different voids statistics as one of the most promising probes of modified gravity.

1. – Introduction

Extended theories of gravity have been considered as a new paradigm to cure shortcomings of General Relativity at infrared and ultraviolet scales. They are an approach that, by preserving the undoubtedly positive results of Einstein's theory, are aimed to address problems recently emerged in astrophysics, cosmology and high-energy Physics. In particular, problems like dark energy and dark matter [1-7].

Gravity theories beyond General Relativity may possibly explain several cosmological puzzles, specifically the present accelerated expansion of the Universe [1]. Modified

gravity must, however, comply with strong requirements: One is that the model must have similar cosmological predictions to those of ΛCDM for the background evolution and the linear large-scale structures [8]. Another condition is that the modifications to General Relativity are suppressed at small scales [9]. This requirement is assured through the so-called screening mechanisms [10].

Since different modified gravity theories can be degenerate with regard to both the background cosmology and the growth rate of linear perturbations, it is crucial to identify new probes that can be used to break these degeneracies. In this paper we study the effects of a class of screened modified gravity models in the nonlinear regime of structure formation. The aim is to predict possible smoking guns of modified gravity and of screening mechanisms at cluster of galaxy scales.

2. – Theoretical models

Scalar-tensor theories are an extension of General Relativity that add a scalar field φ to the standard Einstein-Hilbert Lagrangian. A general Lagrangian for the scalar field is

$$\mathcal{L} = -\frac{1}{2}\partial_\mu\varphi\partial_\nu\varphi - V(\varphi) + \beta(\varphi)T^\mu_\mu, \tag{1}$$

where $V(\varphi)$ is the self-interacting potential, $\beta(\varphi)$ is a coupling function and T^μ_μ is the trace of the matter energy-momentum tensor. The scalar field gives rise to an additional *fifth force*, which can be quantified by $\gamma \equiv |\mathbf{F}_{\rm Fifth}|/|\mathbf{F}_{\rm N}|$, where $\mathbf{F}_{\rm N}$ is the Newtonian force. Experiments constrain $\gamma \ll 1$ in the Solar System. A screening mechanism with the aim to hide this field from local gravity experiments can be realised in two different ways:

Density dependent mass: If the mass of the field $m^2(\varphi)$ is large in dense environments, then the fifth force mediated by the scalar field is suppressed on scales above its Compton wavelength. On the other hand, in low-density environments, the mass can be light and the scalar field mediates a long-range fifth force. This is the so-called *Chameleon* screening [11].

Density dependent coupling: If the coupling to matter $\beta(\varphi)$ is small in the region of high density, the strength of the fifth force $\mathbf{F}_{\rm Fifth}$ is weak and the modifications to gravity are suppressed. On the other hand, in low-density environments, the size of the fifth force can be of the same order as standard gravity. This idea is the so-called Symmetron mechanism [12].

2˙1. *Chameleon-f(R) gravity.* – The Lagrangian for this theory is

$$S = \int \mathrm{d}^4x\sqrt{-g}\frac{1}{16\pi G}(R + f(R)) + S_m(g_{\mu\nu}, \psi_i). \tag{2}$$

In the quasi-static and weak-field limits the equations of motion become

$$\nabla^2\Phi = \frac{16\pi G}{3}a^2\delta\rho_m + \frac{1}{6}a^2\delta R, \qquad \nabla^2 f_R = -\frac{a^2}{3}[\delta R + 8\pi G\delta\rho_m], \tag{3}$$

where $\delta\rho_m = \rho_m - \bar{\rho}_m$ and $\delta R = R - \bar{R}$ are the density and Ricci scalar perturbations (overbars denote background quantities), and $f_R = \mathrm{d}f(R)/\mathrm{d}R$. In this formulation, f_R plays the role of the scalar degree of freedom φ that determines the fifth force.

We choose the Hu-Sawicki model [13] as a working example, which is given by

$$f(R) = -m^2 \frac{c_1 (R/m^2)^n}{c_2 (R/m^2)^n + 1}\,, \tag{4}$$

where $m^2 = H_0^2 \Omega_m$ is a mass scale and c_1, c_2 and n are model parameters. One recovers a ΛCDM expansion history by setting $c_1/c_2 = 6\Omega_\Lambda/\Omega_m$. In this paper we consider $n = 1$, and we consider models with $|\bar{f}_{R0}| = 10^{-4}$, $|\bar{f}_{R0}| = 10^{-5}$ and $|\bar{f}_{R0}| = 10^{-6}$.

Notice that the modified Poisson equation, eq. (3), can also be written as

$$\nabla^2 \Phi = \nabla^2 \Phi_N - \frac{1}{2}\nabla^2 f_R. \tag{5}$$

This makes explicit that in $f(R)$ models the total gravitational force is governed by a modified gravitational potential $\Phi = \Phi_N - \frac{1}{2} f_R$. It is the nonlinearity of $f(R)$ in eq. (4) that gives rise to the Chameleon screening, and the screening of the fifth force is determined by the depth of the gravitational potential Φ_N.

2˙2. *Symmetron.* – The Symmetron model [12] action is given by

$$S = \int \mathrm{d}x^4 \sqrt{-g} \left[\frac{R}{16\pi G} - \frac{1}{2}(\partial\varphi)^2 - V(\varphi) \right] + S_m(\tilde{g}_{\mu\nu}, \psi). \tag{6}$$

The matter fields ψ couple to the Jordan frame metric $\tilde{g}_{\mu\nu}$, which relates to the Einstein frame metric $g_{\mu\nu}$ as $\tilde{g}_{\mu\nu} = A^2(\varphi) g_{\mu\nu}$. The coupling function $A(\varphi)$ is

$$A(\varphi) = 1 + \frac{1}{2}\left(\frac{\varphi}{M}\right)^2, \tag{7}$$

where M is a mass scale. The total force felt by matter is given by

$$\mathbf{F} = \nabla\left(\Phi_N + \frac{1}{2}\frac{\varphi^2}{M^2}\right) = \nabla\Phi_N + \frac{\varphi\nabla\varphi}{M^2}\,. \tag{8}$$

The potential is taken to be

$$V(\varphi) = V_0 - \frac{1}{2}\mu^2\varphi^2 + \frac{1}{4}\lambda\varphi^4, \tag{9}$$

where the value of V_0 is determined by the condition that the model gives rise to the observed accelerated expansion of the Universe [14]. The field equation for φ reads

$$\Box\varphi = V_{\mathrm{eff},\varphi}, \tag{10}$$

where for nonrelativistic matter the effective potential is given by

$$V_{\rm eff}(\varphi) = V_0 + \frac{1}{2}\left(\frac{\rho_m}{M^2} - \mu^2\right)\varphi^2 + \frac{1}{4}\lambda\varphi^4. \tag{11}$$

In the quasi-static limit, eq. (10) becomes

$$\nabla^2\chi = \frac{a^2}{2\lambda_0^2}\left(\frac{\rho_m}{\rho_{\rm SSB}} - 1 + \chi^2\right)\chi, \tag{12}$$

where $\chi = \varphi/\varphi_0$.

Screening in the Symmetron model is very similar to the Chameleon-$f(R)$ case in the sense that the condition for screening is determined by the local gravitational potential. The important difference is that the coupling $\beta(\varphi) = \frac{\beta_0\varphi}{\varphi_0}$, which is constant in $f(R)$ gravity, now depends on the local field value. In high-density regions, $\rho_m > \rho_{\rm SSB}$, the field falls into the minima $\varphi = 0$, and since the coupling is proportional to φ, the fifth force vanishes.

We define three physical parameters L, β and $z_{\rm SSB}$ which are the range of the field, the coupling strength to matter and the redshift of symmetry breaking:

$$L = \frac{\lambda_0}{\rm Mpc/\mathit{h}} = \frac{3000H_0}{\sqrt{2}\mu}\,, \qquad \beta = \frac{\phi_0 M_{\rm pl}}{M^2} = \frac{\mu M_{\rm pl}}{\sqrt{\lambda}M^2}\,, \qquad (1+z_{\rm SSB})^3 = \frac{\mu^2 M^2}{\rho_{m0}}\,. \tag{13}$$

We simulate four symmetron models: symm_A ($z_{ssb} = 1$, $\beta = 1$, $L = 1$), symm_B ($z_{ssb} = 2$, $\beta = 1$, $L = 1$), symm_C ($z_{ssb} = 1$, $\beta = 2$, $L = 1$), and symm_D ($z_{ssb} = 3$, $\beta = 1$, $L = 1$).

3. – Efficiency of screening mechanisms

The common feature to all the screening mechanisms proposed in the literature is that they are built, and their efficiency tested, assuming the so-called quasistatic approximation for the field equations. For instance, in scalar-tensor theories, a scalar degree of freedom is introduced into the standard Einstein-Hilbert action. This field follows the Klein-Gordon equation of motion, which determines both its time and spatial variations. When constructing screening mechanisms to hide the scalar field within the accurately tested regimes, the quasistatic approximation is invariably applied to the equations of motion for the field. This simplifies the calculations by implying that the scalar field is at rest in the minimum of the local effective potential at all points in space and time. This reduces the equation of motion to a Poisson-like equation, which is readily solved to find the approximated scalar field value at any point.

Notice, however, that the full equation of motion for the scalar field is, in fact, a second-order differential equation in time, more similar to a wave equation. Therefore, ignoring the time evolution of the field, via the quasistatic approximation, is to shortfall effects that are only possible to realize when considering the full equation of motion [15, 16].

3˙1. *Solar System constraints.* – In order to test how screening mechanisms work in the Solar System, the community generally chooses a static, spherically symmetric matter distribution to mimic the Galaxy. We follow this approach and choose the Navarro-Frenk-White (NFW) density profile with the characteristics to represent the Milky Way Galaxy, specifically with a virial radius of $r_{\rm vir} = 137\,{\rm kpc}/h$ and concentration $c = 28$, resulting in a halo mass of $1.0 \times 10^{12} M_\odot$ and a circular velocity of 220 km/s at 8 kpc. The reason for the high value of the concentration is simply that we are modeling not only dark matter, but the total matter of the Milky Way, which is more concentrated than the pure dark matter halo. We also did the calculations with an Einasto profile with identical virial mass, and found that the results presented are not very sensitive to the choice of distribution. Because of limitations of spherical symmetry, we did not model a galactic disc.

One of the most precisely measured gravity parameters to probe deviations from General Relativity is the parametrized post-Newtonian (PPN) parameter γ. It can be expressed as the ratio of the metric perturbations in the Jordan frame, Ψ_J and Φ_J. We find the expression for $\gamma - 1$ to be

$$\gamma - 1 = -\frac{\phi^2}{M^2} \frac{2}{\frac{\phi^2}{M^2} - 2\Psi_E - 2\Psi_E \frac{\phi^2}{M^2}} \,. \tag{14}$$

In General Relativity, $\gamma = 1$ exactly. The strongest constraint to date, measured by the Cassini spacecraft [17], is $\gamma - 1 = (2.1 \pm 2.3) \times 10^{-5}$.

The screening mechanism of the symmetron model works by modifying the effective potential such that the field value is pushed towards zero in high-density regions — like the inner regions of the Galaxy. This results in $\gamma - 1 \to 0$, such that the deviations from General Relativity in the proximity of the Solar System are small. The same occurs for the fifth force F_ϕ associated to the scalar field.

We calculate the γ parameter arising from the smoothed matter distribution of the Milky Way. Note that, by using this method, we find an upper bound on the actual value of $|\gamma - 1|$ in the inner Solar System. This is because we do not include the presence of massive bodies like the Sun, which will increase the screening to some degree. Nevertheless, most of the screening is believed to come from the matter distribution of the Galaxy because, in the symmetron model, the Solar System cannot screen itself in vacuum, and therefore, the theory depends on a working screening from the Galaxy.

3˙2. *Simulations.* – Since the equation of motion is a hyperbolic partial differential equation, it can be solved as an initial value and boundary condition problem. The initial condition at $t = 0$ is chosen to be the static solution of the nonlinear Klein-Gordon equation of motion. With a constant boundary condition, this would imply that the field will stay at rest forever. The boundary condition at the edge of the simulation at $r_{\rm max}$ is chosen to emulate incoming sinusoidal waves in the scalar field, specifically $\chi(r_{\rm max}, t) = \chi_0(r_{\rm max}) + A \sin(\omega t)$. Possible sources of such waves will be discussed later.

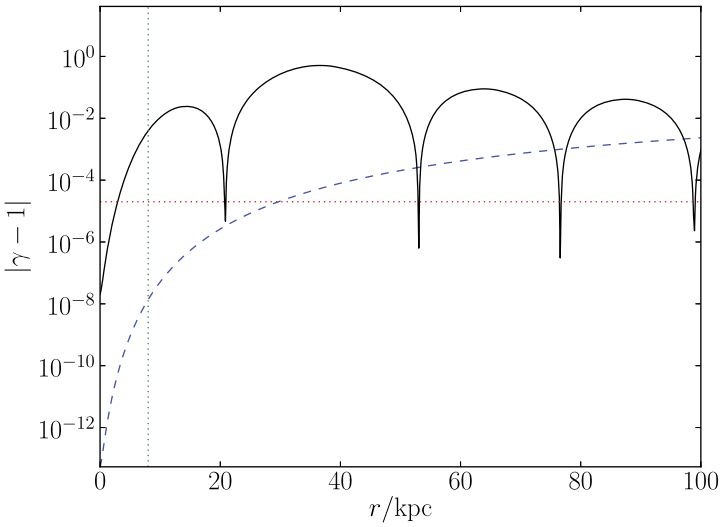


Fig. 1. – The PPN parameter $|\gamma - 1|$, plotted against distance from the center of the Galaxy. The curves show $|\gamma - 1|$ in the quasistatic case (blue dashed line), as well as after a scalar wave has entered the halo (black solid line). The vertical (green dotted) line indicates the position of the Solar System, and the horizontal (red dotted) line indicates the highest allowed value of $\gamma - 1$ in the Solar System from the Cassini experiment. When the wave enters the Milky Way, it increases the value of $|\gamma - 1|$ by several orders of magnitude [15].

We set up a radial grid, divided into linearly spaced steps Δr up to $r_{\max} = 4\,\mathrm{Mpc}/h$. On each of the grid points we specify the matter density according to the NFW halo. Starting from the initial value and with the inclusion of incoming waves, we evolve eq. (10) forward in time steps of size Δt, using the leapfrog algorithm for time integration in each grid point. Tests of this technique applied to the symmetron are presented in [18,19]. We are only interested in events that happen during the last few megayears of cosmic time, meaning that we take the approximations $z \approx 0$ and $a \approx 1$ in all computations. Spatial derivatives are found using a finite-difference method in spherical coordinates, assuming all derivatives in the tangential directions vanish. The code outputs the evolution of the scalar field and, more importantly, the value of $|\gamma - 1|$ at 8 kpc from the center — corresponding to the position of the Solar System in the Milky Way.

We confirm that the values used for technical parameters of our solver give a stable solution by running convergence tests. These are performed by increasing the resolution in factors of two (both temporal and spatial resolution separately) until the resulting scalar field at some later time $t_{\max}$ does not change significantly with resolution.

3˙3. *Results*. – Figure 1 shows an example of how the PPN parameter γ changes when a wave enters the inner 100 kpc of the Milky Way. The vertical line shows the position of the Solar System, which we assume to be 8 kpc from the Galactic Center. The

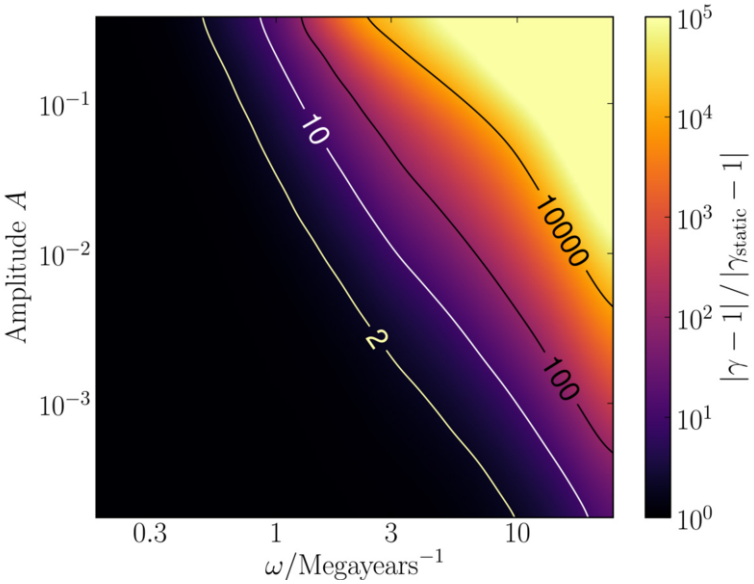


Fig. 2. – Maximum increase in the PPN parameter $|\gamma - 1|$ due to incoming scalar field waves at the position of the Sun in the Galaxy (8 kpc from the center) as a function of amplitude and frequency of the incoming waves. The color indicates by which factor $|\gamma - 1|$ is increased when compared to the quasistatic case with no waves [15].

modifications to gravity are initially screened very well in the regions around this position, with $|\gamma - 1| < 10^{-8}$ (blue dashed line). However, after the wave has arrived (black solid line), the scalar field is perturbed enough to breach the Solar System constraints, $|\gamma - 1| > 2 \times 10^{-5}$. In other words, the screening mechanism breaks down under these circumstances. The wave in this particular simulation has an amplitude $A = 0.01$ and a frequency $\omega = 40\,\mathrm{Myr}^{-1}$. The cusps are regions where the scalar field is zero, which exist since the wave oscillates both above and below $\chi = 0$.

When measuring γ arising from a single sinusoidal wave with low frequency, there is a possibility that the local wave is between two extrema at the time of measurement. This could render this kind of detection difficult for several thousand years. Nevertheless, given that various astrophysical events — such as supernovae — can generate waves, the probability that one of the wavefronts would bring us away from the minima at the present time is not negligible.

In order to investigate how our result depends on the frequency ω and amplitude A of the waves, we simulate incoming waves with several values of these two parameters. Figure 2 shows the maximum growth of $|\gamma - 1|$ that we found at 8 kpc from the Galactic center. Brighter colors mean a larger increase of $|\gamma - 1|$ compared to the quasistatic approximation. The values of the frequency and amplitude that lie in the black region of the plot, give waves that do not significantly impact γ compared to the quasistatic

solution. Therefore, in this region of parameter space, the screening mechanism is efficient and hides the extra degree of freedom from gravity experiments.

From fig. 2, it is possible to conclude that higher frequencies and amplitudes for the incoming scalar waves give larger deviations from the General Relativity result (*i.e.*, $\gamma = 1$). The limit where amplitude and frequency go to zero is equivalent to the quasistatic limit, where no waves are produced and their energy is zero. As one goes into the high-frequency and -amplitude regime, the waves carry more energy, and therefore, the PPN parameter γ starts deviating significantly from the quasistatic limit. Note that, since in the symmetron model, the fifth force is $F_\phi \propto \nabla\phi^2/M^2$, these values can be immediately extrapolated to the impact of the waves on this quantity.

The dependence of the γ PPN parameter on the wave amplitude is straightforward to understand: When a wave propagates through the screened regions of the halo, a larger-amplitude wave will lead to larger displacements of the field from the screening value $\phi \approx 0$. Therefore, $|\gamma - 1| \propto \phi^2$ will increase accordingly.

The frequency dependence of the γ parameter is a consequence of the following: The effective potential of the symmetron grows steeper and narrower in high-density areas. In other words, the mass of the field increases towards the center of the halo. Therefore, it becomes more difficult to perturb the field away from the minimum, and a higher wave energy is needed to displace it. Specifically, if the energy of the external waves is small compared to the mass of the field, the field will not be perturbed and the γ parameter will not be affected.

The results obtained imply that if waves with sufficient amplitude or frequency can somehow be generated in a given model for modified gravity, they will have to be taken into account when constraining the model parameters. Cosmic tsunamis, resulting from extreme events, could even completely ruin the screening mechanisms in modified gravity by increasing the deviations from General Relativity by several orders of magnitude compared to the quasistatic case. A subject that must be discussed now is the generation of such waves. Extreme events on small scales, such as collision of neutron stars, stellar, or super-massive black holes are obvious examples. Generation of waves by pulsating stars are another possibility.

In the specific case of the symmetron model, it is possible to obtain waves from events that occur on cosmological scales. First, the symmetron model undergoes a phase transition when the density falls below a specific threshold. This transition first occurs in voids when the expansion factor is close to $a_{\rm SSB}$ [18, 19]. When this happens, the scalar field receives a kick, which produces waves traveling from the center of the voids towards the dark matter halos. By doing postprocessing of simulations presented in [20], we find that, in a symmetron model with slightly different parameters, the amplitude of cosmological waves is typically smaller that 0.1 and the associated frequencies are of the order of 1/Myr. Note that these values depend on the model parameters and, hence, must be taken only as indicative. Scalar waves can also be created through the collapse of topological defects, which are known to exist in any model in which such phase transition occurs.

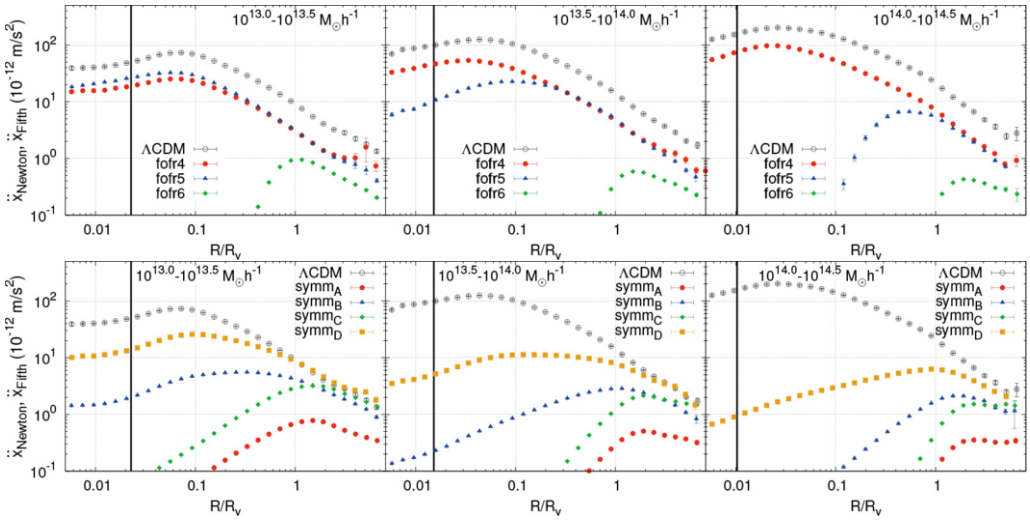


Fig. 3. – Acceleration $|\ddot{\mathbf{x}}|_{\rm Fifth}$ in the $f(R)$ and symmetron models as a function of radius for three different mass bins. Also, the pure Newtonian gravitational force $|\ddot{\mathbf{x}}|_{\rm Newton}$ is included from the ΛCDM data set (black, unfilled circles) [21-24].

4. – Distribution of fifth force in dark matter haloes

To understand the main features of nonlinear structure formation within screened modified gravity, it is important to investigate the magnitude of the fifth force inside the dark matter haloes [2, 21-24]. For that, we run N-body simulations for each model. The simulations were run with the code `Isis` [25]. The background cosmology for all the models is $(\Omega_{m0}, \Omega_{\Lambda 0}, H_0) = (0.267, 0.733, 71.9)$. The data sets are taken at $z = 0$, and all simulations were run with the same initial conditions.

The halos were identified using the spherical overdensity halo finder `AHF` [26] Amiga Halo Finder. For the analysis we used only halos consisting of at least 100 particles which limits the smallest halo we can probe to $M \sim 3 \times 10^{12} M_\odot h^{-1}$. In the high-mass end the simulation-box limits the maximum halo-masses we can study and the largest halos in our simulations has mass $M \sim 2$–$3 \times 10^{15} M_\odot h^{-1}$. We have checked that the mass-function of our simulation agrees to $\sim 10\%$–20% to simulations with larger box-size and also to the fit to the mass-function in the range $M \in [10^{13}, 8 \times 10^{14}] M_\odot h^{-1}$. The total number of halos in this mass range in our simulations, which is what we used for the upcoming analysis, is ~ 8000.

Figure 3 shows the value of the fifth force acceleration as a function of the cluster radius at redshift $z = 0$. All the curves show a characteristic maximum, whose position is dependent on the model parameters and the mass of the haloes. In the $f(R)$ case, the maximum of the fifth force profile moves towards larger radii when increasing mass, and decreasing $|f_{R0}|$. This is a consequence of the screening that is activated in the centre

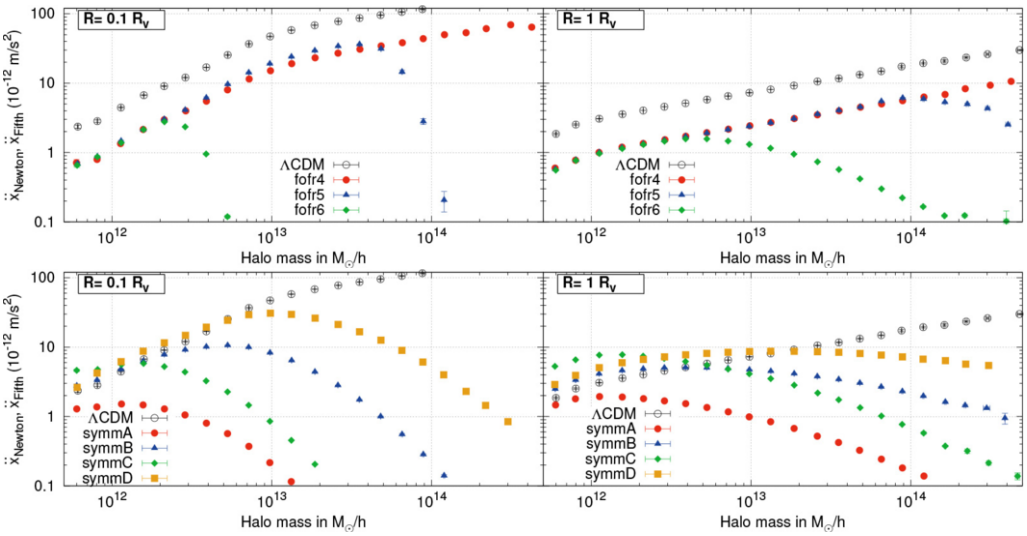


Fig. 4. – Acceleration $|\ddot{\mathbf{x}}|_{\rm Fifth}$ in the $f(R)$ and symmetron models as a function of halo mass for two different radial bins. The black unfilled circles represent $|\ddot{\mathbf{x}}|_{\rm Newton}$ [21-24].

of the haloes, and when decreasing $|f_{R0}|$. For the symmetron model, an increase of β or $z_{\rm SSB}$ leads to a stronger fifth force. A greater β value increases the $|\ddot{\mathbf{x}}_{\rm Fifth}|$ values by a constant factor (symm_A *vs.* symm_C) while altering z_{ssb} changes the shape of the fifth force profile in general.

We show in fig. 4 the acceleration $|\ddot{\mathbf{x}}_{\rm Fifth}|$ as a function of halo mass as measured at two different radii. It is clear that the fifth force is screened in large massive haloes (since they have a higher density) while the force is active in the low-mass range (where the density is lower).

The dependence of the symmetron fifth force on the redshift of symmetry breaking $z_{\rm SSB}$ is clear from the figure: the higher the symmetry breaking, the higher the masses that are unscreened and affected by the fifth force. The low-mass end of the distribution is insensitive to changes in $z_{\rm SSB}$, since in there the density is low and the force is unscreened. Differences in the models are also dependent of the strength of the coupling β.

5. – The matter and the velocity power spectra

Due to the presence of the fifth force, it is expected in modified theories of gravity, that the acceleration felt by particles is in general higher than in General Relativity [27]. Therefore, a promising observable and probe of modified gravity is the measure of velocity distributions in galaxy clusters.

The global statistical properties of the velocity field are described by the velocity divergence power spectrum. Normalising the divergence of the velocity field $\vec{\nabla}\cdot\mathbf{v}$ with the Hubble parameter gives the dimensionless expansion scalar $\theta = \frac{1}{H}\vec{\nabla}\cdot\mathbf{v}$.

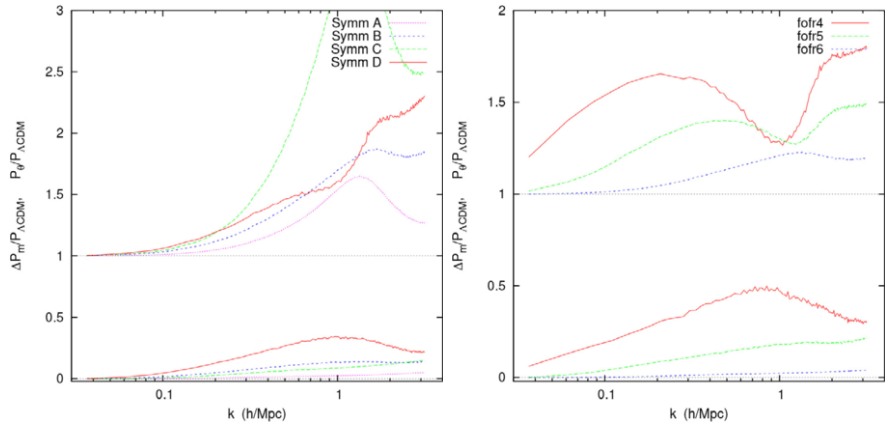


Fig. 5. – The fractional difference of the velocity divergence (above) and matter (below) power-spectra with respect to ΛCDM for the Symmetron and $f(R)$ simulations [27].

We compute the power-spectrum of θ from our simulations. To characterize this quantity, we study the relative particle velocity, which is simply defined as $v_{\rm rel} = \sqrt{(\mathbf{v} - \mathbf{v}_{\mathcal{H}})^2}$, where $\mathbf{v}$ and $\mathbf{v}_{\mathcal{H}}$ are the particle velocity and its halo velocity, respectively. For the latter, we use the core velocity of the halo.

In fig. 5 we show the fractional difference in the velocity divergence with respect to ΛCDM for both modified gravity models. For comparison we also show the matter power-spectrum. For our $f(R)$ simulations we find that the difference with respect to ΛCDM in the velocity divergence spectrum can be roughly two times as large as the difference in the matter power spectrum. For the Symmetron the difference can be much larger. For the symm_C model (which is the model with the largest value of the coupling strength β), we see that $(\Delta P/P)_{\rm m} \approx 10\%$ at $k = 1h/Mpc$ while $(\Delta P/P)_\theta \approx 200\%$. The symm_$C$ model has a fifth force in unscreened regions four times that of the other Symmetron models and this is likely why we get this extreme signal.

The reason why generally we have $(\Delta P/P)_\theta \gtrsim (\Delta P/P)_{\rm m}$ is that the velocity divergence field is not mass-weighted in any way. Hence, low-density regions (voids) will contribute a large part of the signal in the velocity divergence power-spectrum (since voids contribute a large part of the volume in the Universe), which is not the case for the matter power-spectrum. Now the fifth force is generally not screened in low-density regions, so consequently velocities are boosted to significantly higher relative values (when compared with ΛCDM) in voids opposed to in clusters. This indicates that low-density regions like cosmic voids, as we would expect, is the place where the strongest signals of modified gravity can be found.

6. – The dynamical and lensing masses

A general prediction of Modified Gravity theories with screening mechanisms is violations of the equivalence principle. The equivalence principle can be tested by observing

differences between the gravitational mass and the inertial mass of objects, in cosmology through measurements of galaxy clusters masses. Astronomers infer the masses of galaxy clusters using lensing observations and dynamical methods. The former is a probe of the gravitational mass, while the later is a measure of the inertial mass. Differences between these two masses would be a smoking gun of modified gravity and screening mechanisms. In order to make predictions within the chamleon-$f(R)$ and the symmetron gravity, we first need to compute the Bardeen potentials.

In the Jordan-frame we have

$$\mathrm{d}s^2 = a^2(\eta)\left[-\mathrm{d}\eta^2(1+2\Phi) + (1-2\Psi)\mathrm{d}\vec{x}^2\right], \tag{15}$$

where $\Phi \simeq \Phi_N + \delta A(\phi)$, $\Psi \simeq \Phi_N - \delta A(\phi)$, with

$$\delta A(\phi) \equiv A(\phi) - 1 = \frac{1}{2}\left(\frac{\phi}{M}\right)^2. \tag{16}$$

The fifth-force potential is given by the difference in the above two potentials

$$\Phi_- = \frac{\Phi - \Psi}{2} = \delta A(\phi). \tag{17}$$

Lensing on the other hand is affected by the lensing potential

$$\Phi_+ = \frac{\Phi + \Psi}{2} = \Phi_N, \tag{18}$$

which satisfies the Poisson equation

$$\nabla^2\Phi_+ = 4\pi G a^2 \delta\rho_m. \tag{19}$$

This is the same equation as in GR since it is conformally invariant and therefore photons are not affected by the fifth force. The lensing mass is defined as

$$M_L = \frac{1}{4\pi G a^2}\int \nabla^2\Phi_+ \mathrm{d}V, \tag{20}$$

which is the gravitational mass of the halo. It is determined from the simulations by counting the number of particles within a given radius. For spherical symmetry, using the Stokes theorem, we have

$$M_L(r) \propto r^2 \frac{\mathrm{d}\Phi_+}{\mathrm{d}r}. \tag{21}$$

The dynamical mass $M_D(r)$ of a halo is defined as the mass contained within a radius r as inferred from the gravitational potential Φ, *i.e.*

$$M_D(r) = \frac{1}{4\pi G a^2}\int \nabla^2\Phi \,\mathrm{d}V, \tag{22}$$

where the integration is over the volume of the body out to radius r. For spherical symmetry, and using Stokes theorem, we find

$$M_D(r) \propto \int r^2 \frac{\mathrm{d}\Phi(r)}{\mathrm{d}r} = r^2 \left(\frac{\mathrm{d}\Phi_N}{\mathrm{d}r} + \frac{\phi}{M^2} \frac{\mathrm{d}\Phi}{\mathrm{d}r} \right). \tag{23}$$

The terms in the brackets are recognised as the sum of the gravitational force and the fifth-force. Observationally, M_D can be determined from measurements of velocity dispersion of galaxies in clusters.

In GR the lensing mass is the same as the dynamical mass, but they can be significantly different in modified gravity. We follow [28] and define the relative difference

$$\Delta_M(r) = \frac{M_D}{M_L} - 1 = \frac{\mathrm{d}\Phi_-/\mathrm{d}r}{\mathrm{d}\Phi_+/\mathrm{d}r}. \tag{24}$$

This allows us to quantify the difference between the two masses in the simulations. In GR we have $\Delta_M \equiv 0$ while in the symmetron model Δ_M will vary depending on the mass of the halo and its environment. The theoretical maximum is achieved for small objects in a low-density environment where the screening is negligible and reads

$$\Delta_M^{\mathrm{Max}}(r) = 2\beta^2. \tag{25}$$

In fig. 6 we show $\Delta_M(r_{340})$ as a function of the halo mass in both high-density (purple circles) and low-density (blue circles) environments. This figure shows that GR is recovered for larger halos, independent of the environment. For low-mass halos we see a significant dispersion of Δ_M from 0 to the maximum value obtained in low-density environments for the same mass ranges. This is because low-mass halos cannot efficiently screen themselves and must rely on the environment to get the screening. As expected, we see a clear trend that the small halos which are efficiently screened generally reside in high-density environments, while those which are less screened lie in low-density environments. Massive halos on the other hand can screen themselves efficiently and the environment only plays a small role in their total screening.

In fig. 7 we show $\Delta_M(r)$ as a function of the distance r from the halo centre, for small and large halos in high- and low-density environments, respectively. Again we see a large difference between large halos in dense environments and small halos in low-density environments. The r-dependence of $\Delta_M(r)$ is shown to be rather weak in high-density environments since the value of the scalar field inside the halo is mainly determined by the environment, while in low-density environments the value of the scalar field mainly depends on the mass of the halo, which leads to a stronger r-dependence.

Note also that in all the figures above the deviation from GR is stronger for higher symmetry-breaking redshift z_{SSB}, and for larger values of the coupling β, which implies a stronger fifth force and therefore a stronger effect.

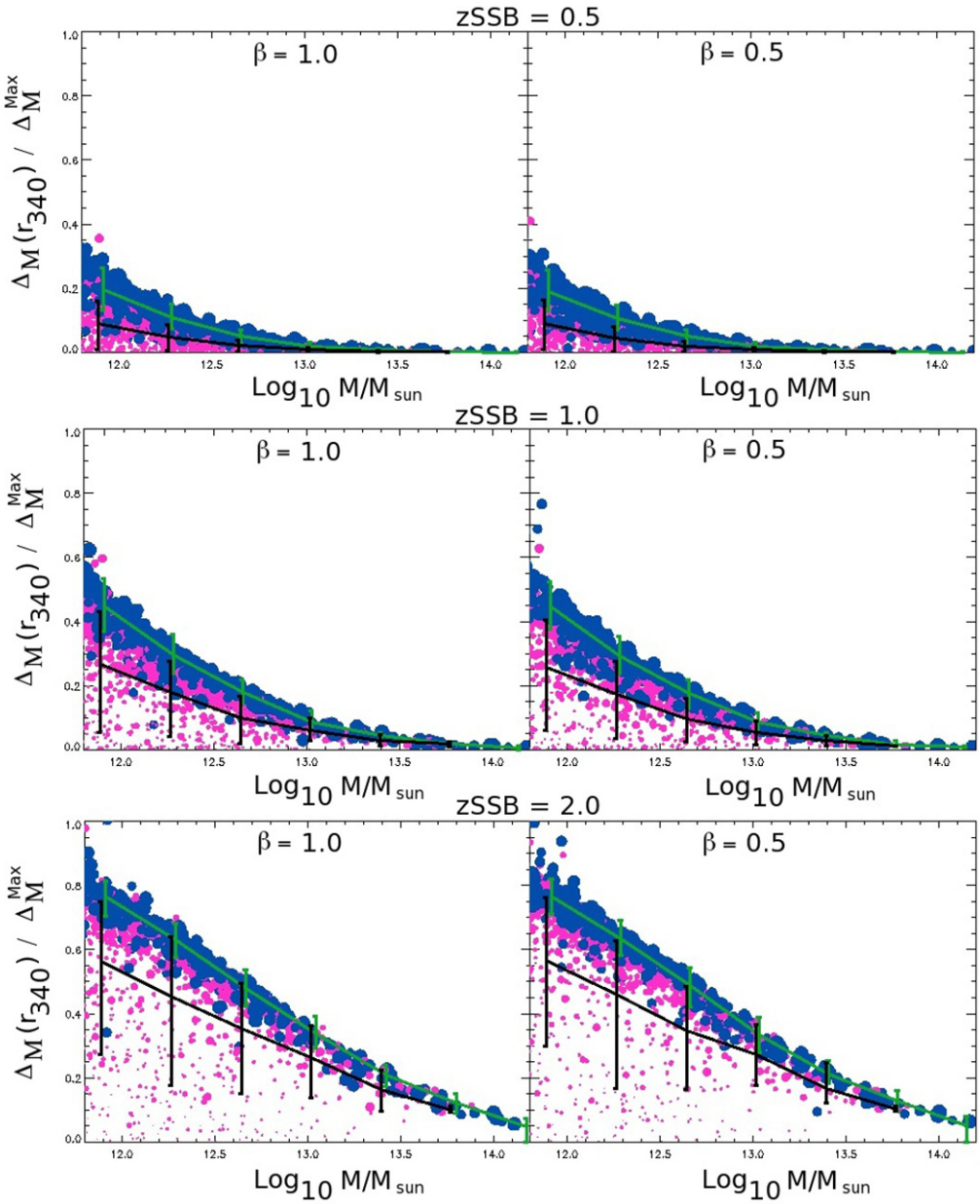


Fig. 6. – $\Delta_M(r_{340})/\Delta_M^{\rm Max}$ as a function of the halo lensing mass M_L for high-density environments (purple) and low-density environments (blue) where $\Delta_M^{\rm Max} = 2\beta^2$. The error bars are 1σ. For the high-mass halos we recover GR independent of the environment as the effectiveness of the screening increases with mass [28].

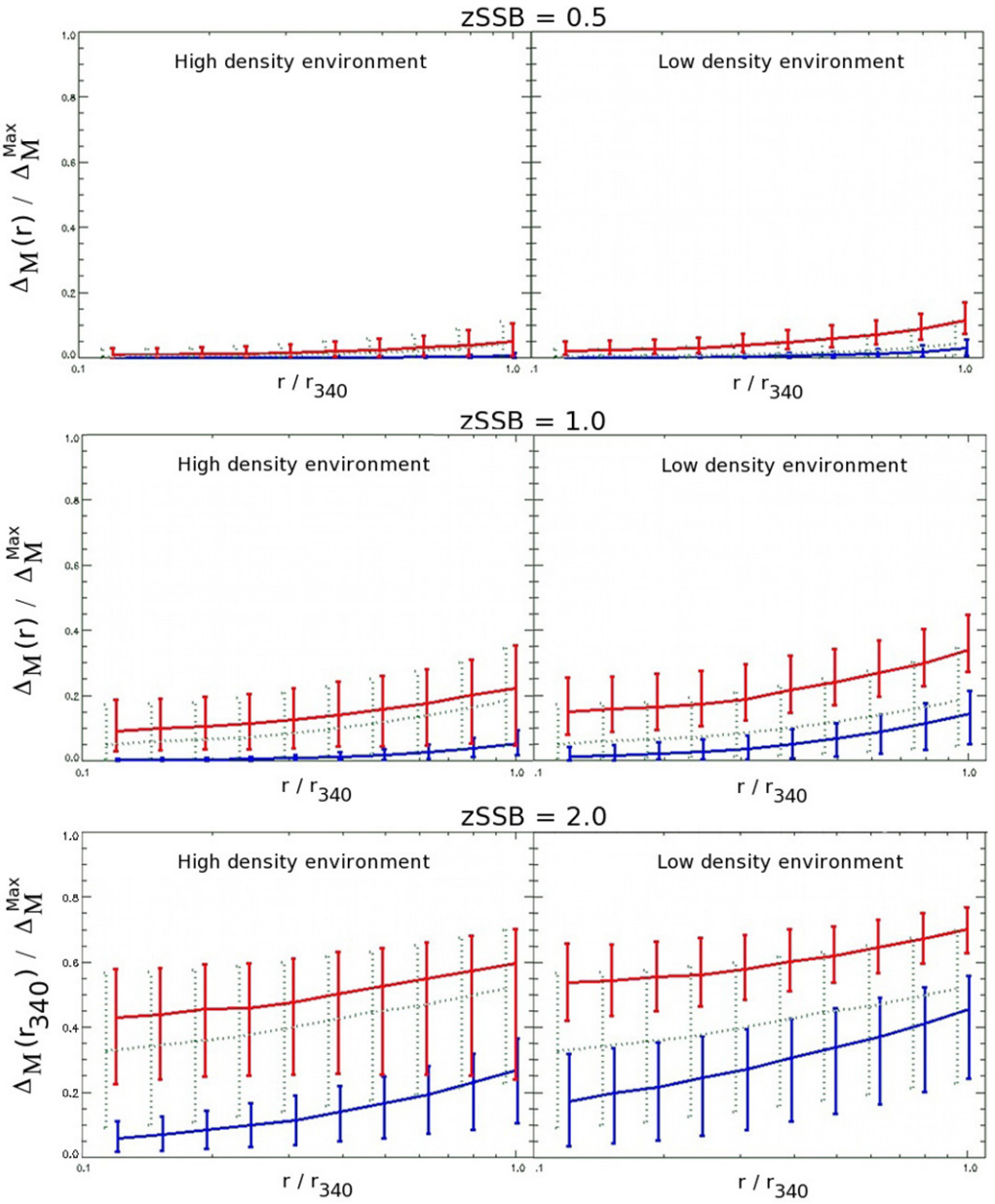


Fig. 7. – $\Delta_M(r)/\Delta_M^{\rm Max}$ as a function of the rescaled halo radius r/r_{340}, where $\Delta_M^{\rm Max} = 2\beta^2$, for high- and low-density environments and small (red) and large (blue) halos. The error bars are 1σ [28].

7. – Thermal *versus* lensing mass measurements

In order to compare the lensing and thermal mass of the clusters we took measurements from [29] and [30]. These two data-sets provide both thermal and lensing mass measurements and uncertainties for a total of 58 clusters in the mass range $M/M_{\odot}h^{-1} \in [5 \times 10^{13}, 3 \times 10^{15}]$ so there was no need to combine the mass estimates in a similar fashion as in the previous section. We divided the data for the thermal mass measurements by the data for the lensing mass measurements while properly propagating the error. As we're interested in a systematic deviation, we binned the data in 6 (lensing) mass bins which we stratified so that roughly the same number of halos are in each bin.

An important point to bear in mind when working with thermal mass estimates is the fact that the measured quantity in this case is the temperature of the intracluster gas. The conversion to a mass assumes hydrostatic-equilibrium, as done, *e.g.*, in [29,30]. However, it has been shown that in reality the pressure of the intracluster medium will have a significant non-thermal component generated by random gas motions and turbulence. This means the inferred thermal mass given a temperature T will be slightly lower than the true mass of the cluster.

While empirical models exist in order to quantify the magnitude of this deviation (where the non-thermal component yield variations to the mass from 10% to 30%) we want to stress that these were calibrated against pure ΛCDM simulations, and thus their results cannot be taken into account when dealing with modified gravity. One has to consider instead that if gravity is truly enhanced, the temperature of the intracluster medium will be hotter and, thus, the inferred thermal mass will be greater (as shown in sect. **5**). This means the effect of any non-thermal physics (such as cosmic rays) is degenerate with modified gravity and, consequently, at the present time thermal measurements cannot be used to constrain modified gravity [31-33].

Figure 8 illustrates this degeneracy. Here, we compare the $M_{\rm therm}/M_{\rm lens}$ results of the Symmetron D model and our ΛCDM simulation to hypothetical measurements where we modelled the contribution of non-thermal pressure as

$$P_{\text{non-thermal}} = P_{\text{total}}\tilde{g}(M_{200}), \tag{26}$$

which resembles the functional forms fitted to ΛCDM simulations (see also subsect. **7**˙1). Thus, our proposed non-thermal contribution is not unreasonable. Modelling a non-thermal contribution as given by eq. (26) while keeping the total pressure $P_{\rm total} = P_{\rm therm} + P_{\text{non-thermal}}$ (and, thus, the halo structure) constant is equivalent to rescaling the temperature as $T \leftarrow T(1 - \tilde{g})$ since naturally $P_{\rm therm} \propto T$.

As fig. 8 shows in the case of a non-thermal contribution the $M_{\rm therm}$ measurement (which is carried out the same way as done by observations, *i.e.*, assuming no non-thermal contribution) matches the lensing mass reasonably well *in the case of modified gravity*. We achieved this by choosing the functional form of $\tilde{g}$ in eq. (26) as

$$\tilde{g}(M) = \frac{1}{1 + \tilde{a}M_{13}^{\tilde{\alpha}}}\,, \tag{27}$$

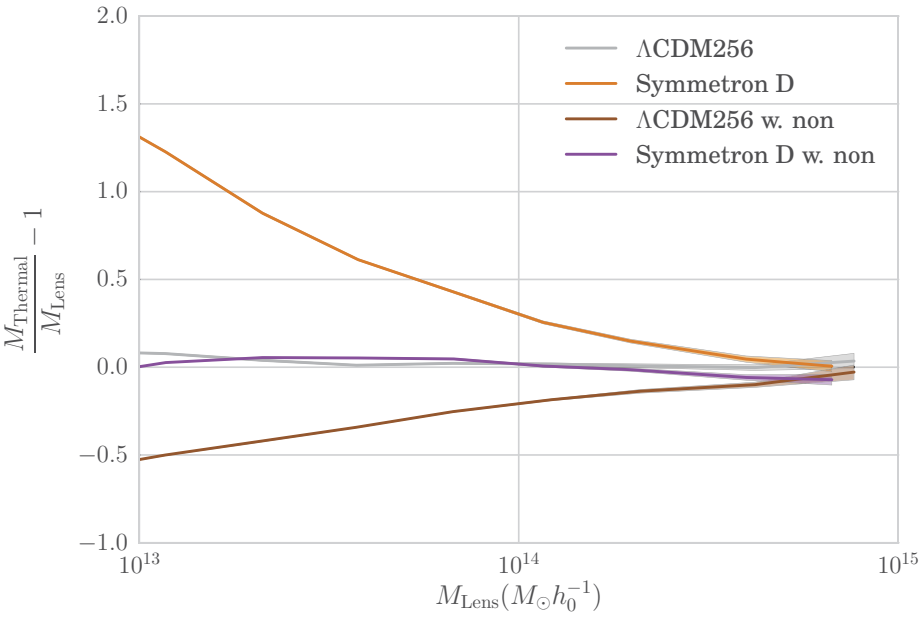


Fig. 8. – Ratio of thermal and lensing mass for ΛCDM and Symmetron D, the symmetron model with the largest deviations from ΛCDM. Also included is the hypothetical measurements obtained by including the non-thermal contribution described in sect. **7** [31-33].

with $M_{13} \equiv M/(10^{13} M_\odot h^{-1})$ and $(\tilde{a}, \tilde{\alpha}) = (3/4, 2/3)$. This serves as an example of how unknown non-thermal physics can cancel out any signal originating from modified gravity — which is a severe problem when trying to place constraints on the modifications of gravity using thermal measurements.

This problem will be alleviated once the contribution of non-thermal effects can be directly quantified using observational data (*e.g.*, by measuring directly the intra-cluster turbulence). In the sequel of this subsection, we assume this has been done and is has been shown the contribution of the non-thermal components is negligible. We do this in order to show which constraints on modified gravity can be placed hypothetically using thermal mass estimates.

7˙1. *Including the non-thermal pressure component.* – It is known that the pressure of the intracluster medium will have a significant non-thermal component generated by random gas motions and turbulence, so that the total pressure $P_{\rm Tot}$ of a cluster is

$$P_{\rm Tot}(< r) = P_{\rm thermal}(r) + P_{\rm non\text{-}thermal}(r). \tag{28}$$

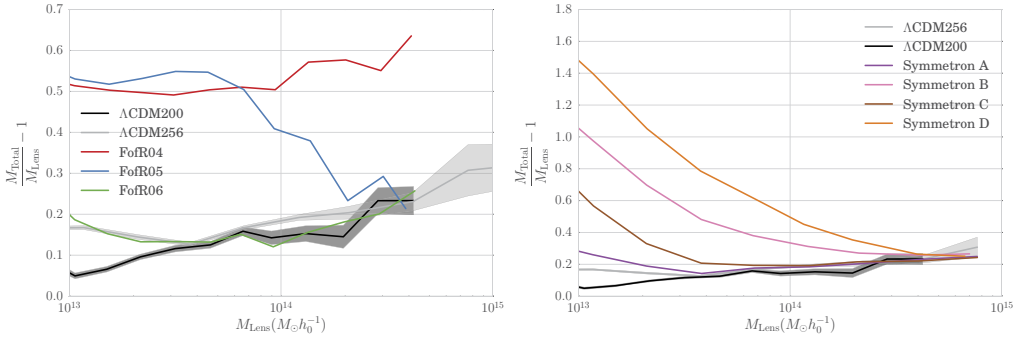


Fig. 9. – Ratio of the combined thermal and non-thermal mass and lensing mass for the analyzed $f(R)$ models (left panel) and Symmetron models (right panel) [31].

This results in the mass estimates will consist of a thermal and non-thermal component as well

$$M(< r) = M_{\rm thermal}(r) + M_{\rm non\text{-}thermal}(r), \tag{29}$$

where

$$M_{\rm thermal}(r) = -\frac{r^2}{G\rho_{\rm gas}(r)}\frac{{\rm d}P_{\rm thermal}(r)}{{\rm d}r}\,, \tag{30}$$

$$M_{\rm non\text{-}thermal}(r) = -\frac{r^2}{G\rho_{\rm gas}(r)}\frac{{\rm d}P_{\rm non\text{-}thermal}(r)}{{\rm d}r}\,. \tag{31}$$

By using $P_{\rm thermal} = kn_{\rm gas}T_{\rm gas}$, where $\rho_{\rm gas} = \mu m_p n_{\rm gas}$, we find that

$$\frac{{\rm d}P_{\rm thermal}}{{\rm d}r} = \frac{kT_{\rm gas}(r)}{\mu m_p}\left(\frac{{\rm d}\rho_{\rm gas}(r)}{{\rm d}r} + \frac{\rho_{\rm gas}(r)}{T_{\rm gas}(r)}\frac{{\rm d}T_{\rm gas}(r)}{{\rm d}r}\right) \tag{32}$$

so that

$$M_{\rm therm} = -\frac{k_B r^2 T_{\rm thermal}(r)}{\mu m_p G}\left(\frac{{\rm d}\ln\rho_{\rm thermal}}{{\rm d}r} + \frac{{\rm d}\ln T_{\rm thermal}}{{\rm d}r}\right), \tag{33}$$

as show earlier in the paper.

Often, the non-thermal pressure is expressed as a fraction of the total pressure

$$P_{\rm non\text{-}thermal}(r) = g(r)P_{\rm total}(r) = \frac{g(r)}{1-g(r)}P_{\rm thermal}\,, \tag{34}$$

with the derivative

$$\frac{\mathrm{d}P_{\text{non-thermal}}}{\mathrm{d}r} = \frac{1}{1-g(r)}\left(g(r)\frac{\mathrm{d}P_{\text{thermal}}}{\mathrm{d}r} + \frac{\mathrm{d}g(r)}{\mathrm{d}r}P_{\text{thermal}}\right. \tag{35}$$

$$\left. + \frac{g(r)}{1-g(r)}\frac{\mathrm{d}g(r)}{\mathrm{d}r}P_{\text{thermal}}\right). \tag{36}$$

A fit to the g-function has been found from ΛCDM simulations to be

$$g(r) = \alpha_{\text{nt}}(1+z)^{\beta_{\text{nt}}}\left(\frac{r}{r_{500}}\right)^{n_{\text{nt}}}\left(\frac{M_{200}}{3\times 10^{14}M_{\odot}}\right)^{n_m}, \tag{37}$$

where the free variables have the ΛCDM best-fit values $\alpha_{\text{nt}} = 0.18$, $\beta_{\text{nt}} = 0.5$, $n_{\text{nt}} = 0.8$, and $n_M = 0.2$. The derivative of the g-factor is

$$\frac{\mathrm{d}g(r)}{\mathrm{d}r} = \frac{n_{\text{nt}}}{r}g(r). \tag{38}$$

Using the best fit we redo the analysis from before, now including the non-thermal pressure contribution.

As we can see the results now differs: with the non-thermal pressure component having introduced a strong mass dependence. However, we want to stress that this is just one particular example as the current expression of the non-thermal pressure contribution is derived from standard gravity simulations is strongly model dependent. Thus, we cannot simply use the expression as is for the modified gravity models.

In spite of this complication, we want to note that *in principle* it is possible to use the ratio between the thermal and lensing mass to constrain screened modified gravity theories, and also — when including the kinetic mass — to rule out certain combinations of non-universal coupling. All this, however, requires the contribution of the non-thermal pressure to be "under control", *i.e.*, the magnitude of the intra-cluster turbulence are at least limited by observations.

8. – Modelling void abundance in modified gravity

Due to the fact that screening mechanisms are much less efficient in low-density regions, then voids are clearly one of the most promising astrophysical objects to observe signatures of a strong and active fifth force, and, therefore, to probe possible deviations from General Relativity and evidence for Modified Gravity. In this section we review in detail how one can use voids statistics to perform such work.

8˙1. *Linear power spectrum.* – The spherical evolution model is usually the first step to investigate the abundance of virialized objects tracing the Universe structure, such as halos, and likewise it is a promising tool for voids [34]. It also offers a starting point to

study the collapse of non-spherical structures and the parameters required to quantify the abundance of these objects within extended models.

The large-scale structure of the Universe is well characterized by the evolution of dark matter, which interacts only gravitationally and can be approximated by a pressureless perfect fluid. The line element for a perturbed Friedmann-Lemaître-Robertson-Walker (FLRW) metric in the Newtonian gauge is given by

$$\mathrm{d}s^2 = -a^2(1+2\Psi)\mathrm{d}\tau^2 + a^2(1-2\Phi)\mathrm{d}l^2, \tag{39}$$

where a is the scale factor, τ is the conformal time related to the physical time t by $a\mathrm{d}\tau = \mathrm{d}t$, $\mathrm{d}l^2$ is the line element for the spatial metric in a homogeneous and isotropic Universe and Ψ and Φ are the gravitational potentials.

For a large class of modified gravity models, the perturbed fluid equations in Fourier space are given by

$$\dot{\delta} = -(1+\delta)\theta, \tag{40}$$

$$\dot{\theta} + 2H\theta + \frac{1}{3}\theta^2 = k^2\Phi, \tag{41}$$

$$-k^2\Phi = 4\pi G\mu(k,a)\bar{\rho}_{\mathrm{m}}\delta, \tag{42}$$

where $\delta = (\rho_{\mathrm{m}} - \bar{\rho}_{\mathrm{m}})/\bar{\rho}_{\mathrm{m}}$ is the matter density contrast, θ is the velocity divergence, $H = \dot{a}/a$ is the Hubble parameter and dots denote derivatives with respect to physical time t.

The first is the continuity equation, the second the Euler equation and the last is the modified Poisson equation, where modified gravity effects are incorporated within the function $\mu(a,k)$. In general this function depends on the scale factor a as well as physical scale or wave number k in Fourier space.

Combining these equations we obtain an evolution equation for spherical perturbations in modified gravity given by

$$\delta'' + \left(\frac{3}{a} + \frac{E'}{E}\right)\delta' - \frac{4}{3}\frac{(\delta')^2}{1+\delta} = \frac{3}{2}\frac{\Omega_m}{a^5E^2}\mu(k,a)\delta(1+\delta), \tag{43}$$

where primes denote derivatives with respect to the scale factor a, $E(a) = H(a)/H_0$, $H(a)$ is the Hubble parameter at a, H_0 is the Hubble constant and Ω_m is the present matter density relative to critical. Clearly the growth of perturbations is scale-dependent — a general feature of modified theories of gravity.

The linearized version of eq. (43) is given by

$$\delta'' + \left(\frac{3}{a} + \frac{E'}{E}\right)\delta' = \frac{3}{2}\frac{\Omega_m}{a^5E^2}\mu(k,a)\delta, \tag{44}$$

and can be used to determine linear quantities, such as the linear power spectrum. Notice that this matter linear equation is valid more generally and does not not require spherical perturbations.

The function $\mu(k, a)$ above is given by

$$\mu(k,a) = \frac{(1+2\beta^2)k^2 + m^2a^2}{k^2 + m^2a^2}\,, \tag{45}$$

where β is the coupling between matter and the fifth force and m is the mass of the scalar field propagating the extra force.

It is important to stress that the parameterization in eq. (45) does not fully account for modified gravity perturbative effects, containing only effects of the background and linear perturbations for extra fields related to modified gravity. This is enough for the linearized eq. (44), but is only an approximation in eq. (43). For instance the parameterization in eq. (45) does not contain effects from the screening mechanisms, which would turn μ into a function not only of scale k, but of *e.g.* the local density or gravitational potential.

We start by defining the linear density contrast field $\delta(R)$ smoothed on a scale R around $\mathbf{x} = 0$([1])

$$\delta(R) = \int \frac{\mathrm{d}^3k}{(2\pi)^3}\tilde{\delta}(\mathbf{k})\tilde{W}(k,R), \tag{46}$$

where tildes denote quantities in Fourier space and $W(\mathbf{x}, R)$ is the window function that smooths the original field $\delta(\mathbf{x})$ on scale R.

The variance $S(R) = \sigma^2(R)$ of the linear density field can be written as

$$S(R) = \langle|\delta(R)|^2\rangle = \int \frac{\mathrm{d}k}{2\pi^2}k^2P(k)|\tilde{W}(k,R)|^2, \tag{47}$$

where $P(k)$ is the linear power spectrum defined via

$$\langle\tilde{\delta}(\mathbf{k})\tilde{\delta}(\mathbf{k}')\rangle = (2\pi)^3\delta_D(\mathbf{k}-\mathbf{k}')P(k), \tag{48}$$

and $\delta_D(\mathbf{k} - \mathbf{k}')$ is the Dirac delta function. Clearly the linear power spectrum will play a key role in describing the effects of modified gravity on void properties. For GR computations, we use `CAMB` to compute the linear power spectrum. For modified gravity, we may use `MGCAMB`, a modified version of `CAMB` which generates the linear spectrum for a number of alternative models, such as the Hu and Sawicki $f(R)$ model [13] and others. However it does not compute the linear spectrum for instance for the symmetron model. Therefore we also construct the linear power spectrum independently for an arbitrary gravity theory parametrized by eqs. (44) and (45).

Our independent estimation of the spectrum is accomplished by evolving eqs. (44) and (45) with parameters from specific gravity theories (*e.g.* for $f(R)$ and for symmetron

([1]) The choice $\mathbf{x} = 0$ is irrelevant because of translational invariance in a homogeneous Universe, and is used for simplificity here, as we are interested in the behaviour of δ as a function of scale R.

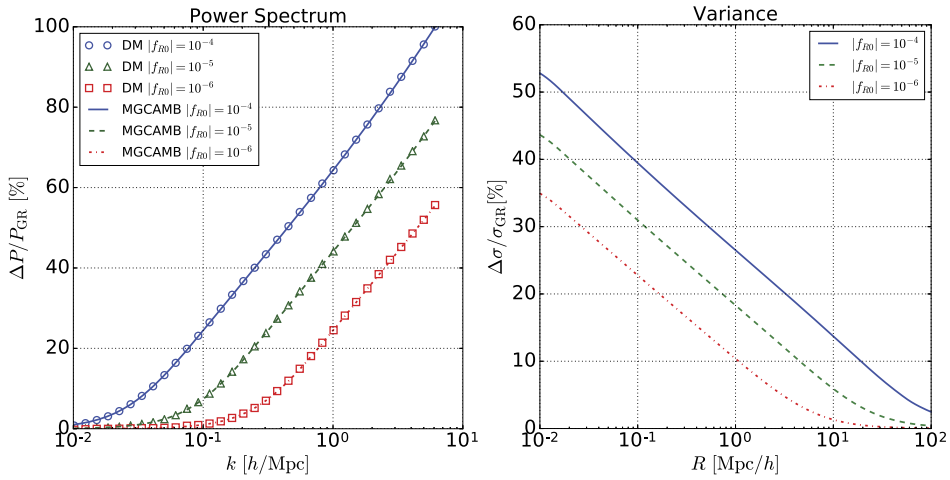


Fig. 10. – Left: Relative percent deviation in the linear matter power spectrum $P(k)$ at $z = 0$ of $f(R)$ modified gravity with respect to the GR spectrum $P_{GR}(k)$ in ΛCDM. Results are shown for spectra obtained from `MGCAMB` (lines) as well as from evolving eq. (44) for dark matter perturbations (open dots), for $|f_{R0}| = 10^{-4}$ (blue solid line and circles), 10^{-5} (green dashed line and triangles) and 10^{-6} (red dot-dashed line and squares). Right: Percent deviation with respect to GR of the mean square density $\sigma(R) = S(R)^{1/2}$ smoothed at scale R, computed from eq. (47) at $z = 0$ for the $f(R)$ model. In this case, the power spectrum was evaluated from eq. (44) [34].

models) for a set of initial conditions at matter domination. Since at sufficiently high redshifts viable gravity models reduce to GR, we take initial conditions given by `CAMB` at high redshifts ($z \approx 100$), when gravity is not yet modified and the Universe is deep into matter domination. We also compute initial conditions for $\dot{\delta}$ numerically by using the ΛCDM power spectrum at two closeby redshifts, *e.g.* at $z = 99$ and $z = 100$.

The results of using this procedure are shown (open dots) on the left panel of fig. 10 and compared with the results from `MGCAMB` (lines) for the Hu and Sawicki model with $n = 1$ and three values of the parameter $|f_{R0}| = 10^{-4}, 10^{-5}, 10^{-6}$. We can see that solving eq. (44) for the power spectrum produces results nearly identical to the full solution from `MGCAMB` on all scales of interest. The percent level differences may be traced to the fact that the simplified equation solved does not contain information about photons and baryons, but only dark matter. For our purposes, this procedure can be used to compute the linear power spectrum for other modified gravity models that reduce to GR at high redshifts, such as the symmetron model.

On the right panel of fig. 10 we see that the relative difference of $\sigma(R) = S(R)^{1/2}$ for the $f(R)$ model with respect to GR can be significant on the scales of interest ($1\,\mathrm{Mpc}/h < R < 20\,\mathrm{Mpc}/h$). Therefore we expect a similar impact on void properties derived from σ and the linear power spectrum.

8˙2. *Spherical collapse*. – Because of the void-in-cloud effect([2]), the linearly extrapolated density contrast δ_c for the formation of halos is important in describing the properties of voids as both are clearly connected. Within theoretical calculations of the void abundance using the excursion set formalism, δ_c corresponds to another absorbing barrier, whose equivalent is not present for halo abundance. Therefore calculating δ_c in the gravity theory of interest gives us important hints into the properties of both halos and voids.

The computation of δ_c is done similarly to that of the GR case, but using eqs. (43) and (44) with the appropriate modified gravity parameterization $\mu(k, a)$ (GR is recovered with $\mu(k, a) = 1$).

We start with appropriate initial conditions([3]) for δ and $\dot{\delta}$ and evolve the linear eq. (44) until a_c. The value of δ obtained is δ_c, the density contrast linearly extrapolated for halo formation at $a = a_c$. In this work, since we only study simulation outputs at $z = 0$, we take $a_c = 1$ in all calculations. The only modification introduced by a nontrivial parameterization $\mu(k, a)$ is that the collapse parameters will depend on the scale k of the halo. As mentioned previously, the parameterization of eq. (45) only takes into account the evolution of the scalar field in the background, and does not account for the dependence of the collapse parameters on screening effects. Even though our calculation is approximated, it does approach the correct limits at sufficiently large and small scales.

For a Universe with only cold dark matter (CDM) under GR, the collapse equations can be solved analytically yielding $\delta_c = 1.686$. For a ΛCDM Universe, still within GR, δ_c changes to a slightly lower value, whereas for stronger gravity it becomes slightly larger. In fig. 11 we show δ_c as a function of scale for the $f(R)$ model. The value of δ_c starts at its ΛCDM value $\delta_c = 1.675$ on scales larger than the Compton scale ($k/a \ll m$; weak-field limit where $\mu \approx 1$) and approaches the totally modified value $\delta_c = 1.693$ on smaller scales ($k/a \gg m$; strong-field limit where $\mu \approx 1 + 2\beta^2 = 4/3$) where the modification to the strength of gravitational force is maximal. These values were computed at the background cosmology described in subsect. **8**˙4. Note that δ_c reaches its strong field limit faster for larger values of $|f_{R0}|$ (value of the extra scalar field today), as expected. In the approximation of eq. (43), δ_c varies with k less than in the full collapse, indicating that the no-screening approximation may not be sufficient. As a full exact calculation is beyond the scope of this work and given that δ_c does not change appreciably, in our abundance models we will fix δ_c to its ΛCDM value and encapsulate modified gravity effects on the linear power spectrum and on other model parameters.

8˙2.1. Spherical expansion. We now compute δ_v, the analog of δ_c for voids, *i.e.* the density contrast linearly extrapolated to today for the formation of a void. We follow a procedure similar to spherical collapse, but in this case the initial values for δ_i are

([2]) The fact that voids inside halos are eventually swallowed and disappear.
([3]) This initial condition is actually determined by a shooting method, evolving the nonlinear eq. (43) for multiple initial values and checking when collapse happens ($\delta \to \infty$) at $a = a_c$.

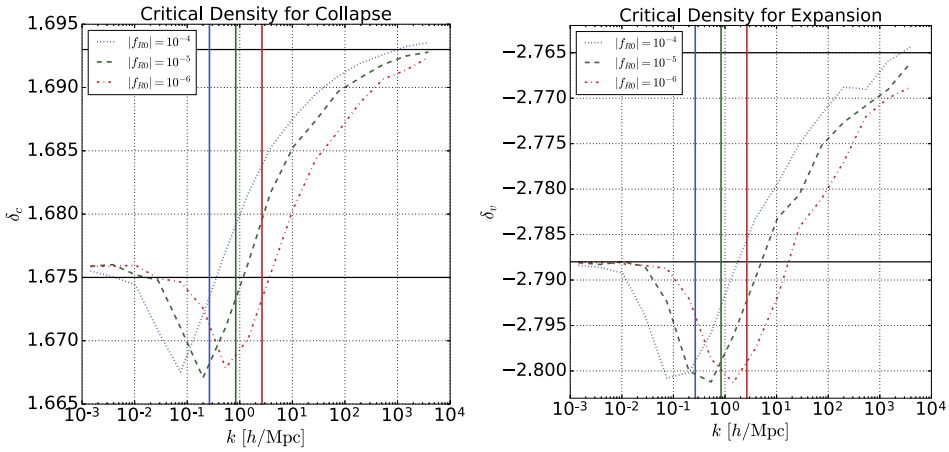


Fig. 11. – Left: The critical density δ_c for collapse of a halo at $z = 0$ as a function of halo scale k in $f(R)$ modified gravity with $|f_{R0}| = 10^{-4}$ (blue dotted line), 10^{-5} (green dashed line) and 10^{-6} (red dot-dashed line). The upper horizontal black line is the value expected for the strong-field limit ($\mu = 4/3$) and the lower line for the weak-field limit, *i.e.* GR ($\mu = 1$). The vertical lines indicate the Compton scales for each gravity with the same corresponding line colors. Right: Same for the critical density δ_v for void formation at $z = 0$ [34].

negative. We also set a criterium in the nonlinear field δ for the formation of a void to be $\delta_{sc} = -0.8$ or equivalently $\Delta_{sc} = 1 + \delta_{sc} = 0.2$. This quantity is somewhat the analogue for voids of the virial overdensity $\Delta_{vir} \approx 180$ for halo formation in an Einstein-de-Sitter (EdS) Universe. Despite the value of Δ_{vir} being only strictly appropriate for an EdS Universe, halos are often defined with this overdensity or other arbitrary values that may be more appropriate for specific observations. Similarly, $\delta_{sc} = -0.8$ is only strictly appropriate for shell-crossing in an EdS Universe. Here we will employ $\delta_{sc} = -0.8$, but we should keep in mind that this is an arbitrary definition of our spherical voids. When we fix this criterium for void formation we also fix the factor by which the void radius R expands with respect to its linear theory radius R_L. This factor is given by $R/R_L = (1 + \delta_{sc})^{-1/3} = 1.717$, and comes about from mass conservation throughout the expansion. Differently from halos, voids are not virialized structures and continue to expand faster than the background. Again environmental dependences are not incorporated in our computations as these values will depend only on scale factor a and the scale k or size of the void.

The right panel of fig. 11 displays the behaviour of δ_v as a function of k, which is very similar to that of δ_c. This is important when modelling the absorbing barriers used for evaluating the void abundance distribution function. Again the values of δ_v vary with k less than in the full calculation.

The spherical collapse and expansion calculations can be performed similarly for the symmetron model, with the appropriate change in the expression for the mass and coupling of the scalar field. For $f(R)$ gravity the change in parameters does not seem to

be relevant and we fix these parameters to their ΛCDM values. In order to treat both gravity models in the same way, we do the same for the symmetron model. Therefore we do not show explicit calculations of δ_c and δ_v for symmetron.

8·3. *Void abundance function.* – We now compute the void abundance distribution function as a function of void size using an extended Excursion Set formalism, which consists in solving the Fokker-Planck equation with appropriate boundary conditions([4]).

Differently from the halo description, for voids it is necessary to use two boundary conditions, because of the void-in-cloud effect. In this case we use two Markovian stochastic barriers with linear dependence in the density variance S, which is a simple generalization from the conventional problem with a constant barrier. The barriers can be described statistically as

$$\begin{aligned} \langle B_c(S)\rangle &= \delta_c + \beta_c S, \\ \langle B_c(S)B_c(S')\rangle &= D_c \min(S,S'), \\ \langle B_v(S)\rangle &= \delta_v + \beta_v S, \\ \langle B_v(S)B_v(S')\rangle &= D_v \min(S,S'), \end{aligned} \tag{49}$$

where $B_c(S)$ is the barrier associated with halos and $B_v(S)$ the barrier associated with voids. Notice that the two barriers are uncorrelated, *i.e.* $\langle B_c(S)B_v(S')\rangle = 0$. Here β_c describes the linear relation between the mean barrier and the variance S, $\delta_{c,v}$ is the mean barrier as $S \to 0$ ($R \to \infty$), and $D_{c,v}$ describes the barrier diffusion coefficient.

As we consider different scales R, the smoothed density field $\delta(R)$ performs a random walk with respect to a *time coordinate* S, and we have([5])

$$\begin{aligned} \langle \delta(S)\rangle &= 0, \\ \langle \delta(S)\delta(S')\rangle &= \min(S,S'). \end{aligned} \tag{50}$$

The field δ satisfies a Langevin equation with white noise and therefore the probability density $\Pi(\delta, S)$ to find the value δ at variance S is a solution of the Fokker-Planck equation

$$\frac{\partial \Pi}{\partial S} = \frac{1}{2}\frac{\partial^2 \Pi}{\partial \delta^2}\,, \tag{51}$$

([4]) This procedure is valid when the barrier (boundary conditions) is linear in S and the random walk motion is Markovian.

([5]) This occurs when the window function in eq. (46) S is sharp in k-space. For a window that is sharp in real space the motion of δ is not Markovian and the second equation in (50) is not true. In that case a more sophisticated method is necessary, and the solution presented here represents the zeroth-order approximation for the full solution.

with boundary conditions

$$\Pi(\delta = B_c(S), S) = 0 \quad \text{and} \quad \Pi(\delta = B_v(S), S) = 0, \tag{52}$$

and initial condition

$$\Pi(\delta, S = 0) = \delta_D(\delta), \tag{53}$$

where δ_D is the Dirac delta function and notice that $S \to 0$ corresponds to void radius $R \to \infty$. In order to solve this problem, it is convenient to introduce the variable

$$Y(S) = B_v(S) - \delta(S). \tag{54}$$

Making the *simplifying* assumption that $\beta \equiv \beta_c = \beta_v$([6]) and using the fact that all variances can be added in quadrature, the Fokker-Planck equation, eq. (51), becomes

$$\frac{\partial \Pi}{\partial S} = -\beta \frac{\partial \Pi}{\partial Y} + \frac{1+D}{2} \frac{\partial^2 \Pi}{\partial Y^2}\,, \tag{55}$$

where $D = D_v + D_c$.

We define $\delta_T = |\delta_v| + \delta_c$ and notice that $\delta(S) = B_v(S)$ implies $Y(S) = 0$, $\delta(S) = B_c(S)$ implies $Y(S) = -\delta_T$ (only occurs because we set $\beta_c = \beta_v$) and $\delta(0) = 0$ implies $Y(0) = \delta_v$. Therefore the boundary conditions become

$$\Pi(Y = 0, S) = 0 \quad \text{and} \quad \Pi(Y = -\delta_T, S) = 0, \tag{56}$$

and the initial conditions

$$\Pi(Y, 0) = \delta_D(Y - \delta_v). \tag{57}$$

Rescaling the variable $Y \to \tilde{Y} = Y/\sqrt{1+D}$ and factoring the solution in the form $\Pi(\tilde{Y}, S) = U(\tilde{Y}, S) \exp[c(\tilde{Y} - cS/2 - \tilde{Y}_0)]$ where $c = \beta/\sqrt{1+D}$ and $\tilde{Y}_0 = \delta_v/\sqrt{1+D}$. The function $U(\tilde{Y}, S)$ obeys a Fokker-Planck equation like eq. (51), for which the solution is known. Putting it all together the probability distribution function becomes

$$\begin{aligned} \Pi(Y, S) = \exp\left[\frac{\beta}{1+D}\left(Y - \frac{\beta S}{2} - \delta_v\right)\right] \\ \times \sum_{n=1}^{\infty} \frac{2}{\delta_T} \sin\left(\frac{n\pi\delta_v}{\delta_T}\right) \sin\left(\frac{n\pi}{\delta_T} Y\right) \exp\left[-\frac{n^2\pi^2(1+D)}{2\delta_T^2} S\right]. \end{aligned} \tag{58}$$

([6]) Notice that β here should not be confused with the coupling between matter and the extra scalar in eq. (45).

The ratio of walkers that cross the barrier $B_v(S)$ is then given by

$$\mathcal{F}(S) = \frac{\partial}{\partial S}\int_{\infty}^{0} \mathrm{d}Y \Pi(Y,S) = \frac{1+D}{2}\frac{\partial \Pi}{\partial Y}\bigg|_{Y=0}, \tag{59}$$

where we used the modified Fokker-Planck equation, eq. (55), and the first boundary condition from eq. (56). The void abundance function, defined as $f(S) = 2S\mathcal{F}(S)$, for this model is then given by

$$\begin{aligned} f(S) = 2(1+D)\exp\left[-\frac{\beta^2 S}{2(1+D)} + \frac{\beta\delta_v}{(1+D)}\right] \\ \times \sum_{n=1}^{\infty}\frac{n\pi}{\delta_T^2} S \sin\left(\frac{n\pi\delta_v}{\delta_T}\right)\exp\left[-\frac{n^2\pi^2(1+D)}{2\delta_T^2}S\right]. \end{aligned} \tag{60}$$

There are four important limiting cases to consider:

– $D = \beta = 0$: This is the simplest case of two static barries. It is given by

$$\begin{aligned} f_{D=\beta=0}(S) = 2\sum_{n=1}^{\infty}\frac{n\pi}{\delta_T^2} S \sin\left(\frac{n\pi\delta_v}{\delta_T}\right) \\ \times \exp\left(-\frac{n^2\pi^2}{2\delta_T^2}S\right), \end{aligned} \tag{61}$$

This is one of the functional forms tested in this work and the only case with no free parameters. We refer to this case as that of 2 static barriers (2SB).

– $D = 0$ and $\beta \neq 0$: This case considers that the barriers depend linearly on S but are not difusive. In this case the expression is given by

$$\begin{aligned} f_{D=0}(S) = 2e^{-\frac{\beta^2 S}{2}} e^{\beta\delta_v}\sum_{n=1}^{\infty}\frac{n\pi}{\delta_T^2} S \sin\left(\frac{n\pi\delta_v}{\delta_T}\right) \\ \times \exp\left(-\frac{n^2\pi^2}{2\delta_T^2}S\right) \end{aligned} \tag{62}$$

Note that these authors define the barrier with a negative slope, therefore our β is equal to their $-\beta$, but $\delta_v < 0$ in our case;

– $\beta = 0$ and $D \neq 0$: Here we have a barrier that does not depend on S but which is diffusive. In this case we have

$$\begin{aligned} f_{\beta=0}(S) = 2(1+D)\sum_{n=1}^{\infty}\frac{n\pi}{\delta_T^2} S \sin\left(\frac{n\pi\delta_v}{\delta_T}\right) \\ \times \exp\left(-\frac{n^2\pi^2(1+D)}{2\delta_T^2}S\right) \end{aligned} \tag{63}$$

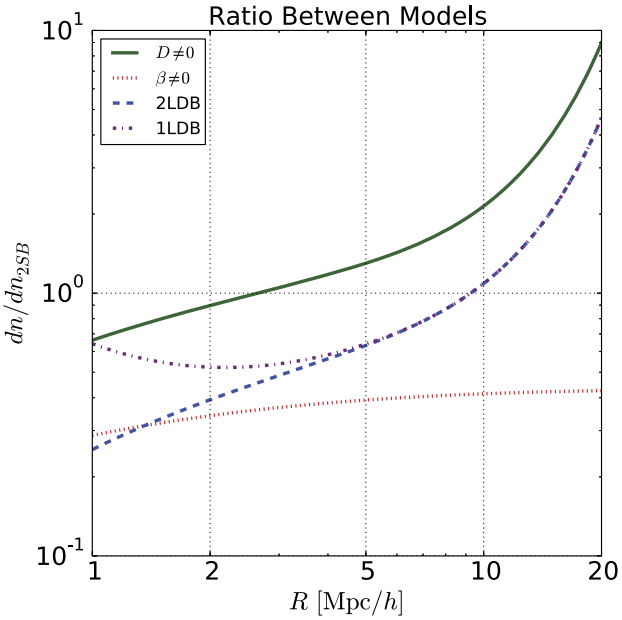


Fig. 12. – Ratio of multiple models for void abundance relative to the model with two static barriers (2SB) eq. (61) ($\beta = D = 0$). We show models with only $D \neq 0$ (green solid line), with only $\beta \neq 0$ (red dotted line), the 1LDB model (purple dotted-dashed line) and the 2LDB model (blue dashed line). The latter two cases are the main models considered in this work and differ only at small radii ($R \lesssim 4\,\mathrm{Mpc}/h$), as a manifestation of the void-in-cloud effect [34].

– *Large void radius*: As discussed in [35] and [36], for large radii R the void-in-cloud effect is not important as we did not expect to find big voids inside halos. In others words, when $S \to 0$ ($R \to \infty$) the abundance becomes equal to that of a one-barrier problem. Even though we do not attempt to properly consider the limit of eq. (60) when $S \to 0$, this expression can be directly compared to the function of the problem with one linear diffusive barrier (1LDB), given by

$$f_{1\mathrm{LDB}}(S) = \frac{|\delta_v|}{\sqrt{S(1+D_v)}}\sqrt{\frac{2}{\pi}}\exp\left[-\frac{(|\delta_v| + \beta_v S)^2}{2S(1+D_v)}\right] \tag{64}$$

In fig. 12, we compare the void abundance from multiple cases by taking their ratio with respect to the abundance of the 2SB model. The abundance of the model with $D \neq 0$ is substantially higher than 2SB, whereas that of the model with $\beta \neq 0$ is significantly lower. The cases with two linear diffusive barriers (2LDB) eq. (60) and one linear diffusive barrier (1LDB) eq. (64) are the main models considered in this work. The void abundance of the 1LDB and 2LDB models are nearly identical for $R > 4\,\mathrm{Mpc}/h$, when the same values of β and D are used.

Given the ratio of walkers that cross the barrier $B_v(S)$ with a radius given by $S(R)$,

the number density of voids with radius between R_L and $R_L + dR_L$ in linear theory is given by

$$\frac{\mathrm{d}n_L}{\mathrm{d}\ln R_L} = \frac{f(\sigma)}{V(R_L)} \frac{\mathrm{d}\ln \sigma^{-1}}{\mathrm{d}\ln R_L}\bigg|_{R_L(R)}, \tag{65}$$

where the subscript L denotes linear theory quantities, $V(R_L)$ is the volume of the spherical void of linear radius R_L and recall $S = \sigma^2$.

Whereas for halos the number density in linear theory is equal to the final nonlinear number density, for voids this is not the case. In fact, Jennings *et al.* [36] shows that such criterium produces nonphysical void abundances, in which the volume fraction of the Universe occupied by voids becomes larger than unity. Instead, to ensure that the void volume fraction is physical (less than unity) the authors of [36] impose that the volume density is the conserved quantity when going from the linear-theory calculation to the nonlinear abundance. Therefore, when a void expands from $R_L \to R$ it combines with its neighbours to conserve volume and not number. This assumption is quantified by the equation

$$V(R)\mathrm{d}n = V(R_L)\mathrm{d}n_L\big|_{R_L(R)}, \tag{66}$$

which implies

$$\frac{\mathrm{d}n}{\mathrm{d}\ln R} = \frac{f(\sigma)}{V(R)} \frac{\mathrm{d}\ln \sigma^{-1}}{\mathrm{d}\ln R_L} \frac{\mathrm{d}\ln R_L}{\mathrm{d}\ln R}\bigg|_{R_L(R)}, \tag{67}$$

where recall in our case $R = (1+\delta_{sc})^{-1/3} R_L = 1.717 R_L$ is the expansion factor for voids. Therefore we have trivially $\mathrm{d}\ln R_L/\mathrm{d}\ln R = 1$ above.

The expression in eq. (67) — referred as the Vdn model — along with the function in eq. (60) provide the theoretical prediction for the void abundance distribution in terms of void radius, which will be compared to the abundance of spherical voids found in N-body simulations of GR and modified gravity.

8˙4. *Voids from simulations.* – We used the N-body simulations that were run with the Isis code [25] for ΛCDM, $f(R)$ Hu-Sawicki and symmetron cosmological models. For the $f(R)$ case we fixed $n = 1$ and considered $|f_{R0}| = 10^{-4}$, 10^{-5} and 10^{-6}. For symmetron, we fix $\beta_0 = 1$ and $L = 1$ and used simulations symm_A, symm_B, symm_D, which have $z_{\mathrm{SSB}} = 1, 2, 3$, respectively. Each simulation has 512^3 particles in a box of size $256\,\mathrm{Mpc}/h$, and cosmological parameters $(\Omega_b, \Omega_{dm}, \Omega_\Lambda, \Omega_\nu, h, T_{CMB}, n_s, \sigma_8) = (0.045, 0.222, 0.733, 0.0, 0.72, 2.726\mathrm{K}, 1.0, 0.8)$. These represent the baryon density relative to critical, dark matter density, effective cosmological constant density, neutrino density, Hubble constant, CMB temperature, scalar spectrum index and spectrum normalization. The normalization is actually fixed at high redshifts, so that $\sigma_8 = 0.8$ is derived for the ΛCDM simulation, but is larger for the modified gravity simulations. In terms of spatial resolution, seven levels of refinement were employed on top of a uniform

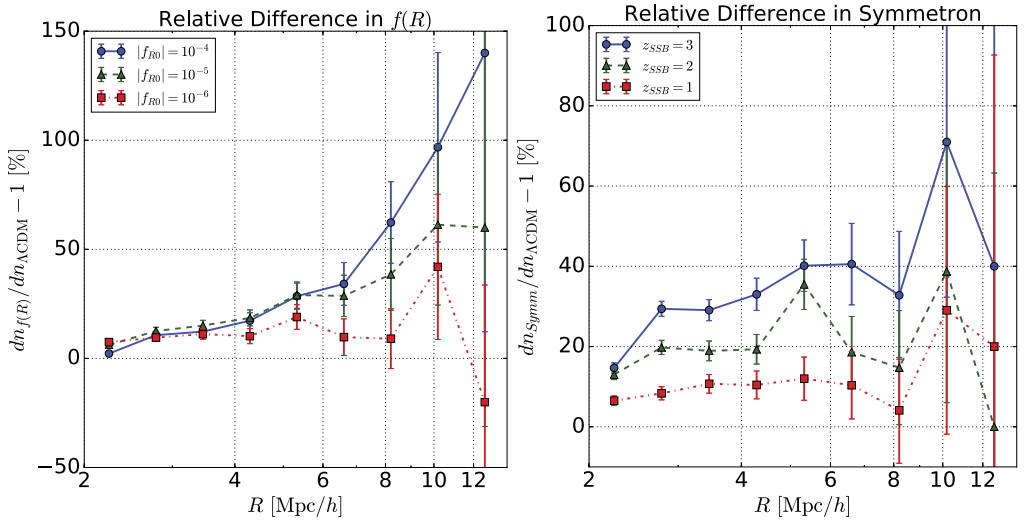


Fig. 13. – Relative difference between void abundance in modified gravity models and in standard GR (ΛCDM model). Left: Relative difference of $f(R)$ theories, for parameters $|f_{R0}| = 10^{-6}$ (red squares with dotted-dashed line), 10^{-5} (green triangles with dashed line) and 10^{-4} (blue circles with solid line). Right: Relative difference of symmetron theories, for parameters $z_{\rm SSB} = 1$ (red squares with dotted-dashed line), 2 (green triangles with dashed line) and 3 (blue circles with solid line) [34].

grid with 512 nodes per dimension. This gives an effective resolution of 32678 nodes per dimension, which corresponds to $7.8\,{\rm kpc}/h$. The particle mass is $9.26 \times 10^9 M_\odot/h$.

We ran the ZOBOV void-finder algorithm — based on Voronoi tessellation — on the simulation outputs at $z = 0$ in order to find underdense regions and define voids, and compared our findings to the Vdn model of eq. (67) [36] with the various multiplicity functions $f(\sigma)$ proposed above (2SB, 1LDB and 2LDB models).

First, we used ZOBOV to determine the position of the density minima locations within the simulations and rank them by signal-to-noise S/N significance. Next, we started from the minimum density point of highest significance and grew a sphere around this point, adding one particle at a time in each step, until the overdensity $\Delta = 1 + \delta$ enclosed within the sphere was 0.2 times the mean background density of the simulation at $z = 0$. Therefore we defined *spherical* voids, which are more closely related to our theoretical predictions based on spherical expansion.

We also considered growing voids around the *center-of-volume* from the central Voronoi zones. The center-of-volume is defined similarly to the center-of-mass, but each particle position is weighted by the volume of the Voronoi cell enclosing the particle, instead of the particle mass. Using the center-of-volume produces results very similar to the previous prescription, so we only present results for the centers fixed at the density minima.

In fig. 13 we compare the void abundance inferred from simulations for the three $f(R)$ and the three symmetron theories relative to the ΛCDM model. Since the differential

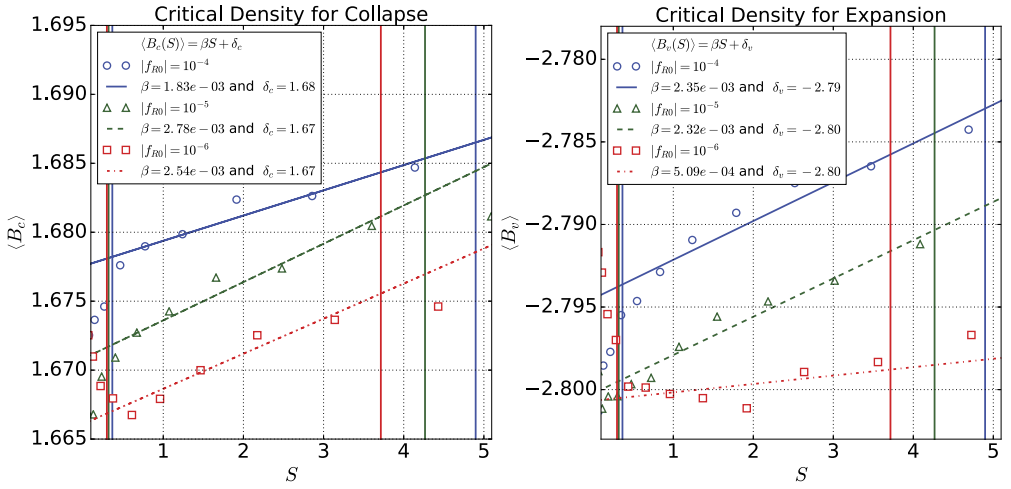


Fig. 14. – Left: Average barrier $\langle B_c \rangle$ for halos as a function of variance S, for the $f(R)$ parameters: $|f_{R0}| = 10^{-6}$ (red squares), 10^{-5} (green triangles) and 10^{-4} (blue circles), and corresponding fits for each case in same colors and with dotted-dashed, dashed and solid lines, respectively. Vertical lines indicate the limits used for the fits, which also correspond to the range of interest for the study of voids in our case (2.0–14.0 Mpc/h). Right: Same for the void barrier $\langle B_v \rangle$ [34].

abundance as a function of void radius is denoted by $\mathrm{d}n/\mathrm{d}\ln R$, we denote the relative difference between the $f(R)$ and ΛCDM abundances by $\mathrm{d}n_{f(R)}/\mathrm{d}n_{\Lambda\mathrm{CDM}} - 1$ and show the results in terms of percent differences. The error bars shown here reflect shot-noise from voids counts in the simulation runs. In the $f(R)$ simulation this relative difference is around 100% at radii $R > 10\,\mathrm{Mpc}/h$ (for the $|f_{R0}| = 10^{-4}$ case). In the symmetron simulation, the difference is around 40% (for the $z_{\mathrm{SSB}} = 3$ case), for radii $R \sim 8\,\mathrm{Mpc}/h$. This indicates that void abundance is a potentially powerful tool for constraining modified gravity parameters.

8˙5. *Results*

8˙5.1. Fitting β and D from simulations. In order to use the theoretical expression in eq. (60) to predict the void abundance we need values for the parameters β and D. The usual interpretation of β is that it encodes, at the linear level, the fact that the true barrier in real cases is not constant. In other words, the contrast density for the void (or halo) formation depends on its size/scale. This can occur because halos/voids are not perfectly spherical and/or because the expansion (or collapse) intrinsically depends on scale (Birkhoff's theorem is generally not valid in modified gravity). The scale dependency induced by modified gravity can be calculated using our model for spherical collapse (expansion) by fitting a linear relationship between δ_c (δ_v) or average barrier $\langle B_c \rangle$ ($\langle B_v \rangle$) as a function of the variance $S(R)$. Here we use $k = 2\pi/R$ to convert wave number to scale R.

In fig. 14 we show the average barriers $\langle B_c \rangle$, $\langle B_v \rangle$ as functions of variance S for multiple gravity theories, and empirical fits for the parameters δ_c, δ_v, β_c, β_v from eqs. (49).

These fits indicate that the barriers depend weakly on scale in the range of interest. The values of δ_c, δ_v are nearly constant and those of β_c, β_v are of order 10^{-3} while the corresponding values for halos in ΛCDM are of order 10^{-1}. Even though voids are quite spherical, the small values of β indicate that the main contribution to β may come from more general aspects of nonspherical evolution. The small fitted values of β can also be due to errors induced by the approximations in the nonlinear equation eq. (43), which does not capture screening effects of modified gravity.

Given these issues, and as it is beyond the scope of this work to consider more general collapse models or study the exact modified gravity equations, we will instead keep the values of δ_c and δ_v fixed to their ΛCDM values and treat β as a free parameter to be fitted from the abundance of voids detected in the simulations.

Likewise, the usual interpretation of D is that it encodes stochastic effects of possible problems in our void (halo) finder [37], such as an intrinsic incompleteness or impurity of the void sample, or other peculiarities of the finder, which may even differ from one algorithm to another. Therefore D is also taken as a free parameter in our abundance models.

In fig. 15 we show the abundance of voids $\mathrm{d}n/\mathrm{d}\ln R$ as measured from simulations (open dots), as well as three theoretical models, namely the 2SB [36], 1LDB eq. (64) and 2LDB eq. (60) models. Multiple panels show results for ΛCDM and $f(R)$ models. In fig. 16 we show the same for ΛCDM and symmetron models.

We can see that linear-diffusive-barrier models (1LDB and 2LDB) work best in all gravities, relative to the static barriers model (2SB). In fact, these two models describe the void abundance distribution within 10% precision for $R \lesssim 10\,\mathrm{Mpc}/h$. As expected, the model with two linear diffusive barriers (2LDB) better describes the abundance of small voids ($R \lesssim 3\,\mathrm{Mpc}/h$), due to the void-in-cloud effect, more relevant for small voids [35].

As both parameters β and D have an explicit dependence on the modified gravity strength, next we fit a relationship between the abundance parameters β and D and the gravity parameters $\log_{10}|f_{R0}|$ and z_{SSB}. In these fits we set the value $\log_{10}|f_{R0}| = -8$ to represent the case of ΛCDM cosmology, as this is indeed nearly identical to ΛCDM for purposes of large-scale structure observables, *i.e.* $\log_{10}|f_{R0}| = -8 \simeq -\infty$.

As we expect β and D to depend monotonically on the modified gravity parameters, we fit for them using simple two-parameter functions. For β case we use a straight line, and for D a second-order polynomial with maximum fixed by the ΛCDM value. These fits are shown in the multiple panels of fig. 17.

Our values of β and D as a function of gravity parameters fluctuate considerably around the best fit. This occurs at least partially because we have used only one simulation for each gravity model, and we expect this oscillation to be reduced with a larger number of simulations. At present, the use of the fits is likely more robust than the use of exact values obtained for each parameter/case.

8˙5.2. Constraining modified gravity. Given the fits for β and D obtained in the last subsection, we now check for the power of constraining modified gravity from the void

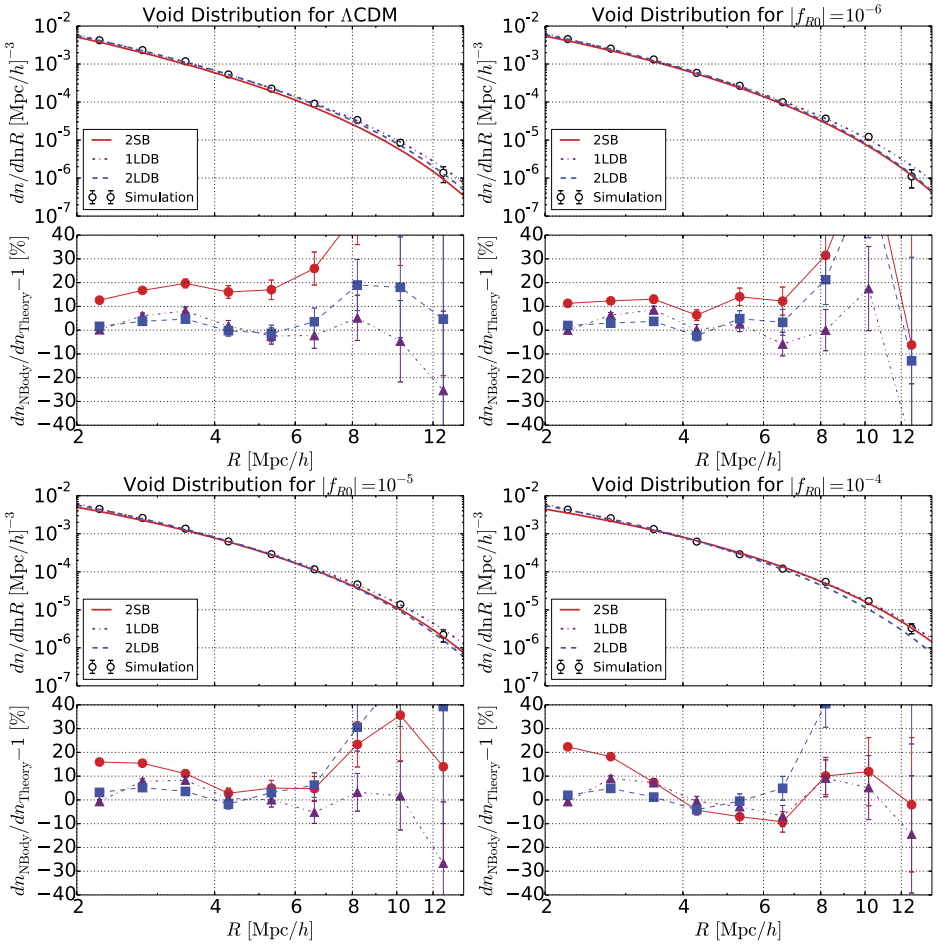


Fig. 15. – Top left: The upper sub-panel shows the void differential abundance distribution $\mathrm{d}n/\mathrm{d}\ln R$ as a function of void radius R for GR (ΛCDM) from simulations (*open dots*), along with theory predictions from the 2SB model [36] (red solid curve), from the 1LDB eq. (64) (purple dotted-dashed curve) and the 2LDB model eq. (60) (blue dashed line). The lower sub-panel shows the relative difference between simulation data and each theory model. Top right: Same for $f(R)$ modified gravity with $|f_{R0}| = 10^{-6}$. Bottom left: Same for $|f_{R0}| = 10^{-5}$ Bottom right: Same for $|f_{R0}| = 10^{-4}$ [34].

distribution function in each of the three void abundance models considered, namely 2SB, 1LDB and 2LDB. We take the abundance of voids actually found in simulations (described in sect. **4**) to represent a hypothetical real measurement of voids and compare it to the model predictions, evaluating the posterior for $\log_{10}|f_{R0}|$ and $z_{\rm SSB}$, thus assessing the constraining power of each abundance model in each gravity theory. Obviously the constraints obtained in this comparison are optimistic — since we are taking

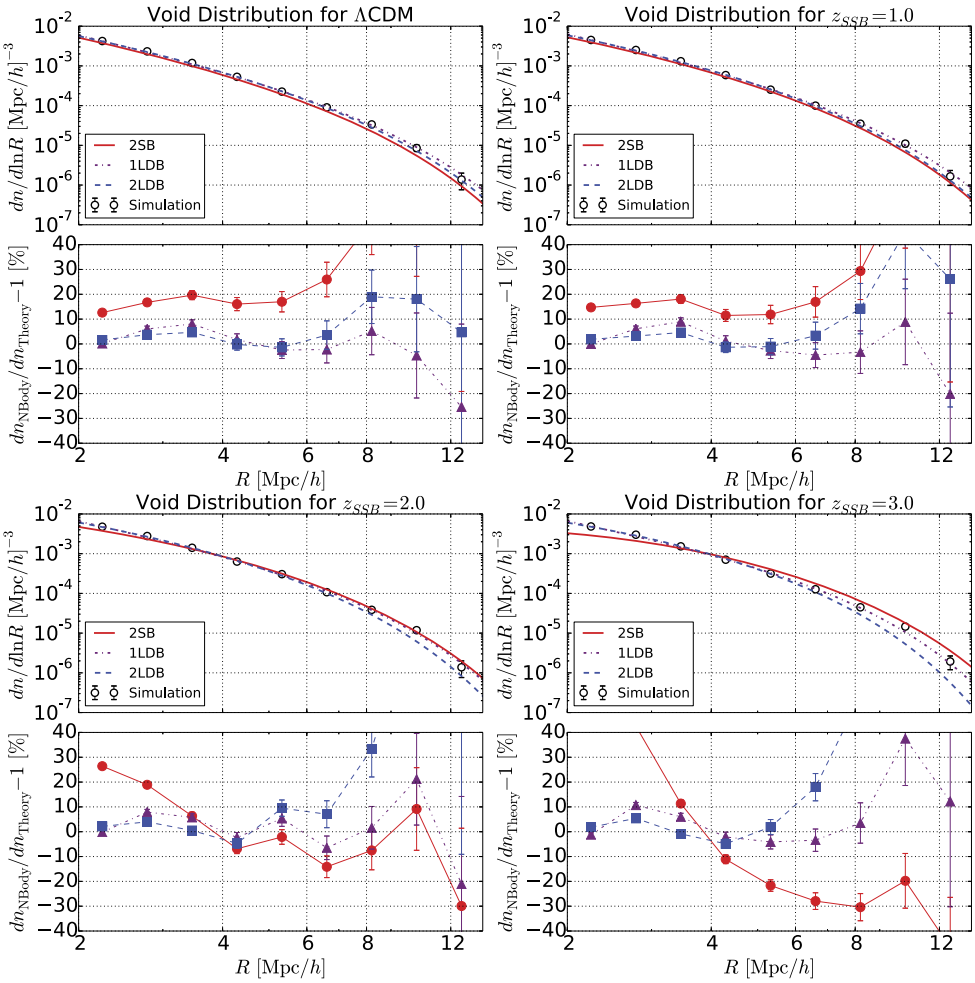


Fig. 16. – Same as fig. 15, but for the symmetron model with $z_{\rm SSB} = 1$ top right, 2 bottom left and 3 bottom right [34].

as real data the same simulations used to fit for the abundance model parameters — but they provide us with idealized constraints similar in spirit to a Fisher analysis around a fiducial model.

The posteriors for the gravity parameters are shown in figs. 18 and 19, as well as the mean values and 1σ errors in each case. For the results shown here all cosmological parameters from subsect. **8**˙4 have been fixed to their true values. We also considered the case where we apply Planck priors on Ω_{dm} and h and let them vary freely in the MCMC, keeping other parameters fixed. In the latter case, the mean values and errors found for $\log_{10}|f_{R0}|$ are slightly worse, but the errors remain less than twice those found for the case of all fixed parameters. Moreover, the errors derived for Ω_{dm} and h reduce to half of their original Planck priors.

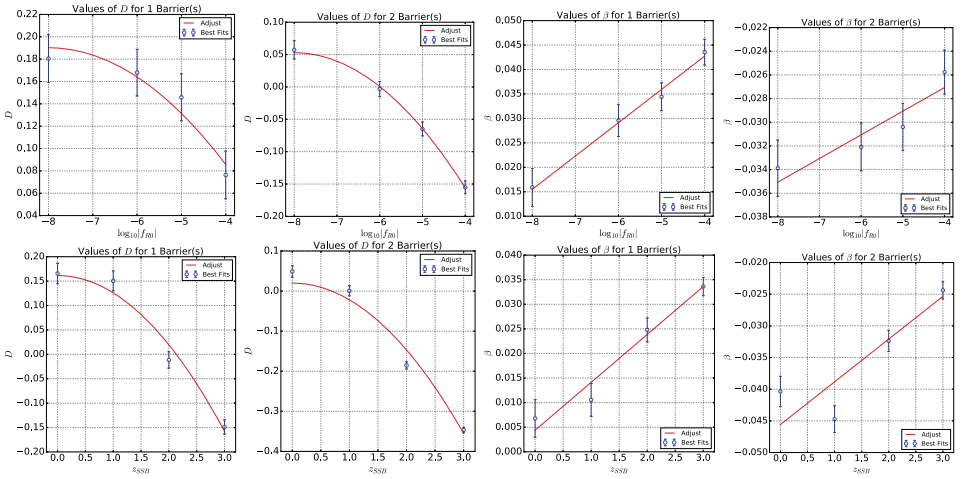


Fig. 17. – Top row: Fits of D and β as a function of $\log_{10}|f_{R0}|$ in $f(R)$ gravity. These fits are shown for D in the 1LDB and 2LDB models, and for β in the 1LDB and 2LDB models, respectively from left to right. Bottom row: Same for fits as a function of $z_{\rm SSB}$ in symmetron gravity [34].

In fig. 18 we can see that the 2SB model predicts values for the $f(R)$ parameter ($\log_{10}|f_{R0}|$) which are incorrect by more than 3σ for all cases. In fact, this model predicts incorrect values even for general relativity. Therefore we find this model to be highly inappropriate to describe the abundance of dark matter voids, and focus on models with linear diffusive barriers.

Both the 1LDB and 2LDB models predict correct values for the gravity parameters within 1σ in most cases. We find that the 1LDB model presents results similar to 2LDB, despite being a simpler model and providing a worse fit to the data (larger reduced χ^2). For ΛCDM both posteriors go to $\log_{10}|f_{R0}| = 10^{-8}$, which represents the GR case by assumption. This shows that within the $f(R)$ framework, we can also constrain GR with reasonable precision from void abundance, using one of these two abundance models with diffusive barriers (1LDB, 2LDB).

For the symmetron Model, we can see in fig. 19 that the parameter $z_{\rm SSB}$ is also well constrained, similarly to f_{R0} in $f(R)$. Again the 2SB model has the worst result in all cases, and the 1LDB and 2LDB models produce similar results.

8˙5.3. Voids in galaxy samples. In real observations it is much harder to have direct access to the dark matter density field. Instead we observe the galaxy field, a biased tracer of the dark matter. Therefore it is important to investigate the abundance of voids defined by galaxies and the possibility of constraining cosmology and modified gravity in this case.

We introduce galaxies in the original dark matter simulations using the Halo Occupation Distribution (HOD) model from [38]. In [39] the authors investigated similar void

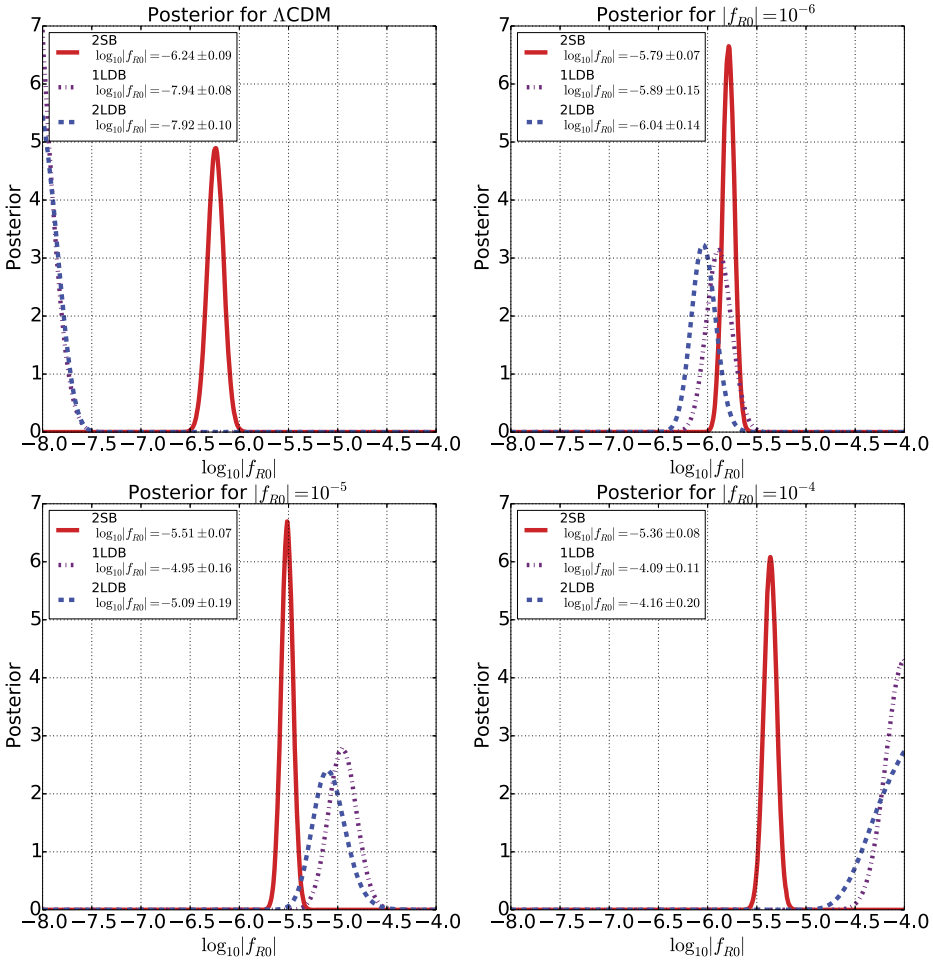


Fig. 18. – Posterior distribution for $\log_{10}|f_{R0}|$ and for the three abundance models considered in the text, 2SB model [36] (red continuous line), 2LDB model eq. (60) (blue dashed line) and 1LDB model eq. (64) (purple dotted dashed line). The mean and 1σ values of $\log_{10}|f_{R0}|$ in each case are indicated in the legend. Top left: Posterior for the ΛCDM simulation. Top right: Posterior for the $|f_{R0}| = 10^{-6}$. Bottom left: Posterior for the $|f_{R0}| = 10^{-5}$. Bottom right: Posterior for the $|f_{R0}| = 10^{-4}$ [34].

properties but did not considered spherical voids, using instead the direct outputs of the `VIDE` [40] void finder.

In our implementation, first we find the dark matter halos in the simulations using the overdensities outputted by `ZOBOV`. We grow a sphere around each of the densest particles until its enclosed density is 200 times the mean density of the simulation. This process is the reverse analog of the spherical void finder described in sect. **4**, the only difference being the criterium used to sort the list of potential halo centers. Here we sort them

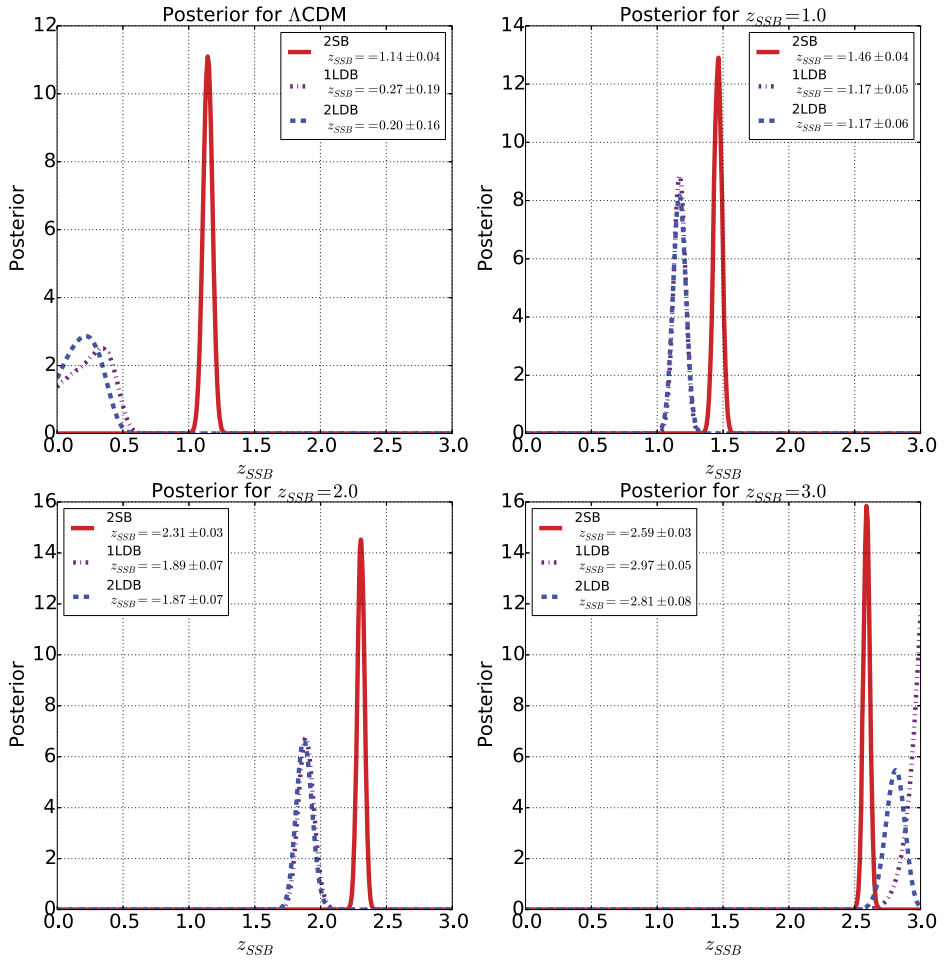


Fig. 19. – Same as fig. 18, but for the symmetron model with $z_{\rm SSB} = 1$ top right, 2 bottom left and 3 bottom right [34].

using the value of the point density, not a S/N significance, as the latter is not provided by ZOBOV in the case of halos.

We populate these halos with galaxies using the HOD model of [38]. This model consists of a mean occupation function of central galaxies given by

$$\langle N_{cen}(M)\rangle = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\log M - \log M_{min}}{\sigma_{\log M}}\right)\right], \tag{68}$$

with a nearest-integer distribution. The satellite galaxies follow a Poisson distribution

with mean given by

$$\langle N_{sat}(M)\rangle = \langle N_{cen}(M)\rangle \left(\frac{M - M_0}{M_1'}\right)^{\alpha}. \tag{69}$$

Central galaxies are put in the center of halo, and the satellite galaxies are distributed following a Navarro Frenk and White profile.

We use parameter values representing the sample *Main 1* of [39], namely: $(\log M_{min}, \sigma_{\log M}, \log M_0, \log M_1', \alpha) = (12.14, 0.17, 11.62, 13.43, 1.15)$. These parameters give a mock galaxy catalogue with galaxy bias $b_g = 1.3$ and mean galaxy density $\bar{n}_g = 5.55 \times 10^{-3}\ (h/\mathrm{Mpc})^3$ in ΛCDM.

We then find voids in this galaxy catalogue using the same algorithm applied to the dark matter catalogue (described in sect. **4**). We use the same criterium that a void is a spherical, non-overlapping structure with overdensity equal to 0.2 times the background galaxy density. However, as the galaxies are a biased tracer of the dark matter field, if we find galaxy voids with 0.2 times the mean density, we are really finding regions which are denser in the dark matter field. In fact, if $\delta_g = b_g\delta$ is the galaxy overdensity, with galaxy bias b_g and δ is the dark matter overdensity we have

$$\Delta = 1 + \delta = 1 + \frac{\delta_g}{b_g}. \tag{70}$$

Therefore, if we find voids with $\delta_g = -0.8$ and $b_g = 1.3$ we have $\Delta = 0.38$, *i.e.* the galaxy voids enclose a region of density 0.38 times the mean density of the dark matter field. Therefore it is this value that must be used in the previous theoretical predictions.

Using this value, the relation between linear and nonlinear radii is $R = 1.37R_L$, and the density parameter for the spherical void formation —calculated using the spherical expansion equations— is $\delta_v = -1.33$. We insert these new values into the theoretical predictions and compare to the measured galaxy void abundance. The result is shown in fig. 20 for the ΛCDM case. We see that both original models, 2SB and 2LDB (blue curves), with $R = 1.71R_L$ and $\delta_v = -2.788$, provide incorrect predictions for the abundance of galaxy voids. However when corrected for the galaxy bias (red curves), these models are in good agreement with the data. We also see that the 2LDB provides a slightly better fit, which is not significant given the error bars.

The main problem of our galaxy catalogues is the low number density of objects. Larger box sizes (or a galaxy population intrinsically denser) might help decrease the error bars sufficiently in order to constrain modified gravity parameters. In fig. 21 we show the relative difference between the abundance for the three modified gravity models and GR as inferred from our simulations. We see that it is not possible to constrain the gravity model using the abundance of galaxy voids, as extracted from mock galaxy catalogues of the size considered here, due to limited statistics. Further investigations using larger or multiple boxes, or else considering a galaxy population with larger intrinsic number density should decrease Poisson errors significantly, allowing for a better investigation of

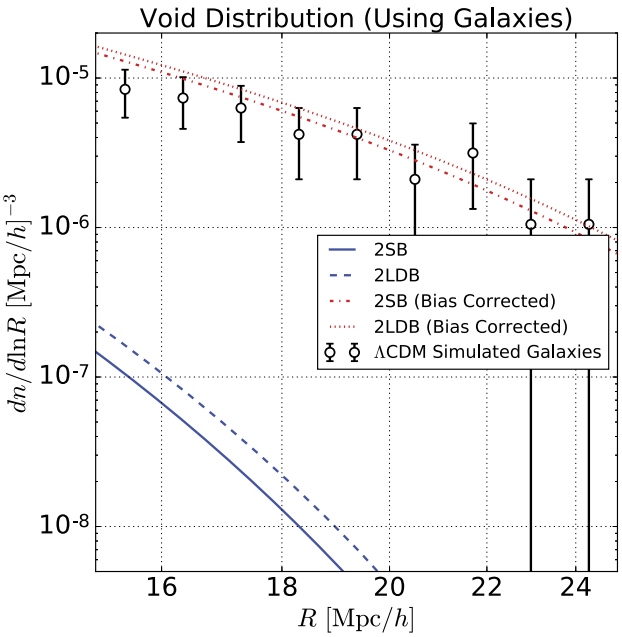


Fig. 20. – Void abundance distribution as a function of void radius for voids detected in the *galaxy* mock catalogue for ΛCDM (open circles). Also shown are the abundance predictions from the 2SB and 2LDB models with no corrections due to galaxy bias (blue solid and dashed lines, respectively), as well as the same model predictions with the bias corrections (red dotted dashes and dotted lines, respectively) [34].

void abundance in the large data sets expected for current and upcoming surveys, such as the SDSS-IV, DES, DESI, Euclid and LSST.

9. – Conclusions and perspectives

The rise of a fifth force in modified gravity theories leads to a stronger clustering of matter. Therefore, the matter power spectra has in general a higher amplitude than in GR. In screened modified gravity theories, the range of the fifth force at cosmological scales is around Mpc (in order to avoid the strong local gravity constraints); therefore, the linear power spectra is similar to ΛCDM, and the differences occur in the small scales of the nonlinear regime.

Overall, we find that halo velocity profiles are an excellent direct tracer of the fifth force: large deviations in the relative velocity of particles are found. Moreover, we find that the velocity field in modified gravity simulations is more affected by the presence of the fifth force than the density field for $f(R)$-gravity. For the Symmetron model we found this to be even more apparent. A particular striking example of this is the Symm_C model, with boosts of up to $(\Delta P/P)_\theta \gtrsim 3$, whereas $(\Delta P/P)_\mathrm{m} \sim 0.1$.

In order to find smoking guns of screening mechanisms, we studied the environment dependence of the masses of dark matter halos in the symmetron modified gravity

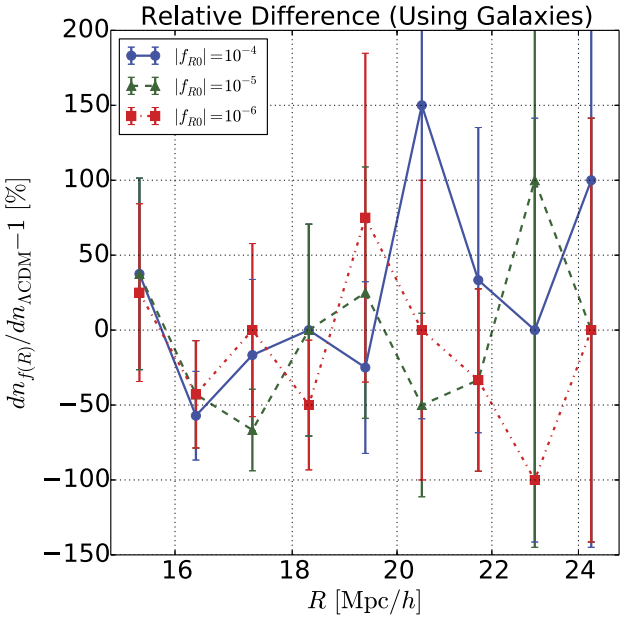


Fig. 21. – Relative difference in galaxy void abundance as measured in $f(R)$ gravity simulations and GR simulations. The difference is shown for $|f_{R0}| = 10^{-6}$ (red squares), 10^{-5} (green triangles) and 10^{-6} (blue circles) [34].

scenario. The potential governing the dynamics of the matter fields ($\Phi_- + \Phi_+$) can differ significantly from the lensing potential Φ_+ in this model, which leads to a clear difference between the mass of the halo as obtained from dynamical measurements and that obtained from gravitational lensing. Such an effect found in the symmetron model can be significantly stronger than in $f(R)$ gravity. This signature, which is unique to modified gravity, can in practice be measured by combining dynamical (*e.g.*, velocity dispersion) and lensing mass measurements of clusters of galaxies or even single galaxies. We find that the environmental dependence is strongest for small halos as very large halos are sufficiently massive to be able to screen themselves.

This discovered feature of environmental dependence also allows us, in principle, to distinguish between different modified gravity scenarios such as $f(R)$, more general chameleons, and the symmetron. In $f(R)$ the maximum fraction of the fifth force to the Newtonian force in halos are around 30% while in chameleon/symmetron scenarios this fraction can be either smaller or larger, depending on the value of the coupling strength β.

Although the theoretical nature of screening mechanisms is different, we find common features in both the matter and the velocity properties. In particular, our findings suggest that one can classify screening mechanisms into three general categories: 1) the fully screened regime where GR is recovered, 2) an unscreened regime where the strength of the fifth force is large, and, 3) a partially screened regime where screening occurs in the

inner part of a halo, but the fifth force is active at larger radii. Any observable sensitive to this regimes and environments can be a future probe of screening mechanisms and modified gravity theories beyond General Relativity.

* * *

DFM thanks M. Gronke, R. Hagala, A. Hammami, C. Llineares, M. Lima, R. Voivodic and H. Winther whose collaborations resulted in the articles which this review is based on. This work was partially supported by the Research Council of Norway and the simulations where performed in NOTUR.

REFERENCES

[1] CLIFTON T., FERREIRA P. G., PADILLA A. and SKORDIS C., *Phys. Rep.*, **513** (2012) 1 doi:10.1016/j.physrep.2012.01.001 [arXiv:1106.2476 [astro-ph.CO]].
[2] SKORDIS C., MOTA D. F., FERREIRA P. G. and BOEHM C., *Phys. Rev. Lett.*, **96** (2006) 011301 doi:10.1103/PhysRevLett.96.011301 [astro-ph/0505519].
[3] KOIVISTO T. S., MOTA D. F. and ZUMALACARREGUI M., *Phys. Rev. Lett.*, **109** (2012) 241102 doi:10.1103/PhysRevLett.109.241102 [arXiv:1205.3167 [astro-ph.CO]].
[4] AKRAMI Y., KOIVISTO T. S., MOTA D. F. and SANDSTAD M., *JCAP*, **1310** (2013) 046 doi:10.1088/1475-7516/2013/10/046 [arXiv:1306.0004 [hep-th]].
[5] BURGESS C. P., EASTHER R., MAZUMDAR A., MOTA D. F. and MULTAMAKI T., *JHEP*, **0505** (2005) 067 doi:10.1088/1126-6708/2005/05/067 [hep-th/0501125].
[6] THORSRUD M., MOTA D. F. and HERVIK S., *JHEP*, **1210** (2012) 066 doi:10.1007/JHEP10(2012)066 [arXiv:1205.6261 [hep-th]].
[7] BARROW J. D. and MOTA D. F., *Class. Quantum Grav.*, **20** (2003) 2045 doi:10.1088/0264-9381/20/11/307 [gr-qc/0212032].
[8] ADE P. A. R. *et al.* (PLANCK COLLABORATION), arXiv:1502.01590 [astro-ph.CO].
[9] WILL C. M., *Living Rev. Rel.*, **9** (2006) 3 doi:10.12942/lrr-2006-3 [gr-qc/0510072].
[10] BRAX P., *Acta Phys. Polon. B*, **43** (2012) 2307. doi:10.5506/APhysPolB.43.2307 [arXiv:1211.5237 [hep-th]].
[11] KHOURY J. and WELTMAN A., *Phys. Rev. Lett.*, **93** (2004) 171104 doi:10.1103/PhysRevLett.93.171104 [astro-ph/0309300].
[12] HINTERBICHLER K. and KHOURY J., *Phys. Rev. Lett.*, **104** (2010) 231301.
[13] HU W. and SAWICKI I., *Phys. Rev. D*, **76** (2007) 064004. [arXiv:0705.1158 [astro-ph]].
[14] DAVIS A. C., LI B., MOTA D. F. and WINTHER H. A., *Astrophys. J.*, **748** (2012) 61.
[15] HAGALA R., LLINARES C. and MOTA D. F., *Phys. Rev. Lett.*, **118** (2017) 101301 doi:10.1103/PhysRevLett.118.101301.
[16] HAGALA R., LLINARES C. and MOTA D. F., *Astron. Astrophys.*, **585** (2016) A37 doi:10.1051/0004-6361/201526439 [arXiv:1504.07142 [astro-ph.CO]].
[17] BERTOTTI B. *et al.*, *Nature*, **425** (2003) 374.
[18] LLINARES C. and MOTA D., *Phys. Rev. Lett.*, **110** (2013) 161101 doi:10.1103/PhysRevLett.110.161101 [arXiv:1302.1774 [astro-ph.CO]].
[19] LLINARES C. and MOTA D. F., *Phys. Rev. D*, **89** (2014) 084023 doi:10.1103/PhysRevD.89.084023 [arXiv:1312.6016 [astro-ph.CO]].
[20] HAGALA R., LLINARES C. and MOTA D. F., *Astron Astrophys.*, **585** (2016) A37, arXiv:1504.07142.
[21] GRONKE M. B., LLINARES C. and MOTA D. F., *Astron. Astrophys.*, **562** (2014) A9 doi:10.1051/0004-6361/201322403 [arXiv:1307.6994 [astro-ph.CO]].

[22] GRONKE M., HAMMAMI A., MOTA D. F. and WINTHER H. A., *Astron. Astrophys.*, **595** (2016) A78 doi:10.1051/0004-6361/201628644.
[23] GRONKE M., MOTA D. F. and WINTHER H. A., *Astron. Astrophys.*, **583** (2015) A123 doi:10.1051/0004-6361/201526611 [arXiv:1505.07129 [astro-ph.CO]].
[24] GRONKE M., LLINARES C., MOTA D. F. and WINTHER H. A., *Mon. Not. R. Astron. Soc.*, **449** (2015) 2837 doi:10.1093/mnras/stv496 [arXiv:1412.0066 [astro-ph.CO]].
[25] LLINARES C., MOTA D. F. and WINTHER H. A., *Astron. Astrophys.*, **562** (2014) A78.
[26] KNOLLMANN S. R. and KNEBE A., *Astrophys. J. Suppl.*, **182** (2009) 608 doi:10.1088/0067-0049/182/2/608 [arXiv:0904.3662 [astro-ph.CO]].
[27] GRONKE M., LLINARES C., MOTA D. F. and WINTHER H. A., *Mon. Not. R. Astron. Soc.*, **449** (2015) 2837 doi:10.1093/mnras/stv496 [arXiv:1412.0066 [astro-ph.CO]].
[28] WINTHER H. A., MOTA D. F. and LI B., *Astrophys. J.*, **756** (2012) 166.
[29] ZHANG Y. Y. *et al.*, *Astrophys. J.*, **711** (2010) 1033 doi:10.1088/0004-637X/711/2/1033 [arXiv:1001.0780 [astro-ph.CO]].
[30] MAHDAVI A., HOEKSTRA H., BABUL A., BILDFELL C., JELTEMA T. and HENRY J. P., *Astrophys. J.*, **767** (2013) 116 doi:10.1088/0004-637X/767/2/116 [arXiv:1210.3689 [astro-ph.CO]].
[31] HAMMAMI A. and MOTA D. F., *Astron. Astrophys.*, **584** (2015) A57 doi:10.1051/0004-6361/201526606 [arXiv:1505.06803 [astro-ph.CO]].
[32] HAMMAMI A. and MOTA D. F., *Astron. Astrophys.*, **598** (2017) A132 doi:10.1051/0004-6361/201629003 [arXiv:1603.08662 [astro-ph.CO]].
[33] HAMMAMI A., LLINARES C., MOTA D. F. and WINTHER H. A., *Mon. Not. R. Astron. Soc.*, **449** (2015) 3635 doi:10.1093/mnras/stv529.
[34] VOIVODIC R., LIMA M., LLINARES C. and MOTA D. F., *Phys. Rev. D*, **95** (2017) 024018. doi:10.1103/PhysRevD.95.024018.
[35] SHETH R. K., MO H. J. and TORMEN G., *Mon. Not. R. Astron. Soc.*, **323** (2001) 1 doi:10.1046/j.1365-8711.2001.04006.x [astro-ph/9907024]; SHETH R. K. and VAN DE WEYGAERT R., *Mon. Not. R. Astron. Soc.*, **350** (2004) 517 doi:10.1111/j.1365-2966.2004.07661.x [astro-ph/0311260].
[36] JENNINGS E., LI Y. and HU W., *Mon. Not. R. Astron. Soc.*, **434** (2013) 2167 doi:10.1093/mnras/stt1169 [arXiv:1304.6087 [astro-ph.CO]].
[37] MAGGIORE M. and RIOTTO A., *Astrophys. J.*, **717** (2010) 515 doi:10.1088/0004-637X/717/1/515 [arXiv:0903.1250 [astro-ph.CO]].
[38] ZHENG Z., COIL A. L. and ZEHAVI I., *Astrophys. J.*, **667** (2007) 760 doi:10.1086/521074 [astro-ph/0703457 [ASTRO-PH]].
[39] NADATHUR S. and HOTCHKISS S., *Mon. Not. R. Astron. Soc.*, **454** (2015) 2228 doi:10.1093/mnras/stv2131 [arXiv:1504.06510 [astro-ph.CO]]; NADATHUR S. and HOTCHKISS S., *Mon. Not. R. Astron. Soc.*, **454** (2015) 889 doi:10.1093/mnras/stv1994 [arXiv:1507.00197 [astro-ph.CO]]; NADATHUR S., HOTCHKISS S., DIEGO J. M., ILIEV I. T., GOTTLÖBER S., WATSON W. A. and YEPES G. IAU Symp. **308** [*IAU Symp.*, **308** (2016) 542] doi:10.1017/S1743921316010541 [arXiv:1412.8372 [astro-ph.CO]].
[40] SUTTER P. M. *et al.*, arXiv:1406.1191 [astro-ph.CO].

International School of Physics "Enrico Fermi"

Villa Monastero, Varenna

Course 200

3 – 12 July 2017

Gravitational Waves and Cosmology

Directors

Eugenio Coccia
Gran Sasso Science Institute
Viale Francesco Crispi 7
67100 L'Aquila
Italy
tel. ++39 08624280249
eugenio.coccia@gssi.infn.it

Joseph Silk
Paris Institute of Astrophysics
98 bis Boulevard Arago
75014 Paris
France
tel. ++33 0144328054
silk@iap.fr

Nicola Vittorio
Dipartimento di Fisica
Università di Roma "Tor Vergata"
Via della Ricerca Scientifica 1
00133 Roma
Italy
tel. ++39 0672594581
nicola.vittorio@uniroma2.it

Scientific Secretary

Irene Sartini
Gran Sasso Science Institute
Viale Francesco Crispi 7
67100 L'Aquila
Italy
irene.sartini@gssi.infn.it

Lecturers

Giuseppe Bono
Dipartimento di Fisica
Università di Roma "Tor Vergata"
Via della Ricerca Scientifica 1
00133 Roma
Italy
tel. ++39 3493942895
bono@roma2.infn.it

Marica Branchesi
Dipartimento di Scienze di Base
e Fondamenti
Università di Urbino "Carlo Bo"
Via S. Chiara 27
61029 Urbino
Italy
tel. ++39 0722303393
marica.branchesi@gmail.com

Carlo Burigana
Istituto di Astrofisica Spaziale
e Fisica Cosmica di Bologna
Area della Ricerca
Via Piero Gobetti 101
40129 Bologna
Italy
tel. ++39 0516398705
burigana@iasfbo.inaf.it

Jens Chluba
Royal Society URF
Jodrell Bank Centre for Astrophysics
Alan Turing Building
Room 3.213
University of Manchester
Oxford Road
Manchester M13 9PL
UK
tel. ++44 1613062765
jens.chluba@manchester.ac.uk

Nelson Christensen
ARTEMIS UMR 7250
Observatoire de la Côte d'Azur
Boulevard de l'Observatoire
CS 34229
06304 Nice Cedex 4
France
tel. ++33 4 92003166
nelson.christensen@oca.eu

Jacques Delabrouille
Laboratoire Astroparticule & Cosmologie
Rue Alice Domon et Lèonie Duquet 10
75013 Paris Cedex 13
France
tel. ++33 157276040
delabrouille@apc.univ-paris7.fr

Viviana Fafone
Dipartimento di Fisica
Università di Roma "Tor Vergata"
Via della Ricerca Scientifica 1
00133 Roma
Italy
tel. ++39 0672594563
viviana.fafone@roma2.infn.it

Federico Ferrini
European Gravitational Observatory
Via E. Amaldi
56021 Cascina (PI)
Italy
tel. ++39 050 752300
federico.ferrini@ego-gw.it

Francesco Fidecaro
Dipartimento di Fisica "Enrico Fermi"
Università di Pisa and INFN
Largo Bruno Pontecorvo 3
56127 Pisa
Italy
tel. ++39 0502214909
Francesco.fidecaro@unipi.it

David Fonseca Mota
Institute of Theoretical Astrophysics
University of Oslo
Svein Rossendalds hus Sem Sælands vei 13
0371 Oslo
Norway
tel. ++49 22857581
d.f.mota@astro.uio.no

Nicolao Fornengo
Dipartimento di Fisica
Università di Torino e INFN
Via P. Giuria 1
10125 Torino
Italy
tel. ++39 0116707225
fornengo@to.infn.it

Gianluca Maria Guidi
Dipartimento di Scienze Pure
e Applicate
Università di Urbino "Carlo Bo"
Via Aurelio Saffi 2
61029 Urbino
Italy
tel. ++0722 303386
gianluca.guidi@uniurb.it

Wayne Hu
Kavli Institute for Cosmological Physics
University of Chicago
933 East 56th St.
Chicago, IL 60637
USA
tel. ++1 7737020160
whu@background.uchicago.edu

Michael Kramer
Max Planck Institute
for Radio Astronomy
Auf dem Hügel 69
D-53121 Bonn
Germany
tel. ++49 228525278/299
michael@mpifr-bonn.mpg.de

Michela Mapelli
INAF - Osservatorio Astronomico
di Padova
Vicolo dell'Osservatorio 5
35122 Padova PD
Italy
tel. ++39 0498293527
michela.mapelli@oapd.inaf.it

Jerome Martin
Paris Institute of Astrophysics
98 bis Boulevard Arago
75014 Paris
France
tel. ++33 0144328167
jmartin@iap.fr

Sabino Matarrese
Dipartimento di Fisica
e Astronomia "Galileo Galilei"
Università di Padova
Via 8 febbraio 2
35122 Padova
Italy
tel. ++39 0498277120
sabino.matarrese@unipd.it

William Percival
Institute of Cosmology and Gravitation
University of Porthsmouth
Dennis Sciama Building
Burnaby Road
Portsmouth, PO1 3FX
UK
tel. ++44 2392 843107
will.percival@port.ac.uk

Fulvio Ricci
Dipartimento di Fisica
Sapienza Università di Roma
Piazzale Aldo Moro 2
00185 Rom
Italy
tel. ++39 0649914261
fulvio.ricci@roma1.infn.it

Douglas Scott
Department of Physics and Astronomy
Hennings Building
University of British Columbia
6224, Agricultural Road
Vancouver, BC V6T 1Z1
Canada
tel. ++1 604 8222802
dscott@phas.ubc.ca

Alicia Sintes
Physics Department
University of the Balearic Islands
Mateu Orfila i Rotger Building
Cra. de Valldemossa, km 7.5
07122 Palma de Mallorca
Spain
tel. ++34 971259834
Alicia.sintes@uib.es

Patrick Sutton
School of Physics and Astronomy
Cardiff University
Queen's Builings
The Parade
Cardiff CF24 3AA
UK
tel. ++44 29 208 74043
Patrick.Sutton@astro.cf.ac.uk

Stefano Vitale
Dipartimento di Fisica
Università di Trento
Via Mesiano 77
38123 Trento
Italy
tel. ++39 0461281568
stefano.vitale@unitn.it

Simon D. M. White
Max Planck Institute for Astrophysics
Karl-Schwarzschild-Str. 1
Postfach 1523
85740 Garching
Germany
tel. ++49 89300002211
swhite@mpa-garching.mpg.de9

Students

Lorenzo Aiello	Gran Sasso Science Institute, Italy
Sandeep Balakrishna Haridasu	Gran Sasso Science Institute, Italy
Giampaolo Benevento	Università di Padova, Italy
Matteo Bonetti	Università dell'Insubria, Italy
Alessandro Buzzelli	Sapienza Università di Roma, Italy
Claudio Casentini	INFN, Sezione di Roma 2, Italy
Eleonora Castelli	Università di Trento, Italy
Marco Celoria	Gran Sasso Science Institute, Italy
Gong Cheng	Chinese Academy of Science, China
Francesco Cipriano	Observatoire Côte D'Azure, France
Alessio Cirone	INFN, Sezione di Genova, Italy
Rocco D'Agostino	Università di Roma "Tor Vergata", Italy
Matthias Dahlmanns	Universität zu Köln, Germany
Francesco Di Renzo	Università di Pisa, Italy
Jose Maria Ezquiaga	Instituto De Fisica Teorica UAM-CSIC, Spain
Lorenzo Galante	Università di Torino, Italy
Maria Chiara Guzzetti	Università di Padova, Italy
Robert Hagala	University of Oslo, Norway
Odysse Halim	Gran Sasso Science Institute, Italy
Huanchen Hu	Università di Padova, Italy
Kirtika Juhi Hurgobin	University of Cape Town, South Africa
Jay Vijay Kalinani	Università di Roma "Tor Vergata", Italy
Oleg Kalinin	Università di Roma "Tor Vergata", Italy
Arpine Kozmanyan	Università di Roma "Tor Vergata", Italy
Mark Linton	University of Portsmouth, UK
Vladimir Luković	Università di Roma "Tor Vergata", Italy
Diana Lumaca	Università di Roma "Tor Vergata", Italy
Andrew Miller	Sapienza Università di Roma, Italy
Eugenio Noda	Università di Parma, Italy
Rishikesh Pandit	Università "Roma Tre", Italy
Chris Pattison	University of Portsmouth, UK
Andrea Ravenni	Università di Padova, Italy
William Wright	University of Portsmouth, UK

Observers

Marcus C. Werner	Kyoto University, Japan

This page intentionally left blank

Finito di stampare
nel mese di luglio 2020

This page intentionally left blank