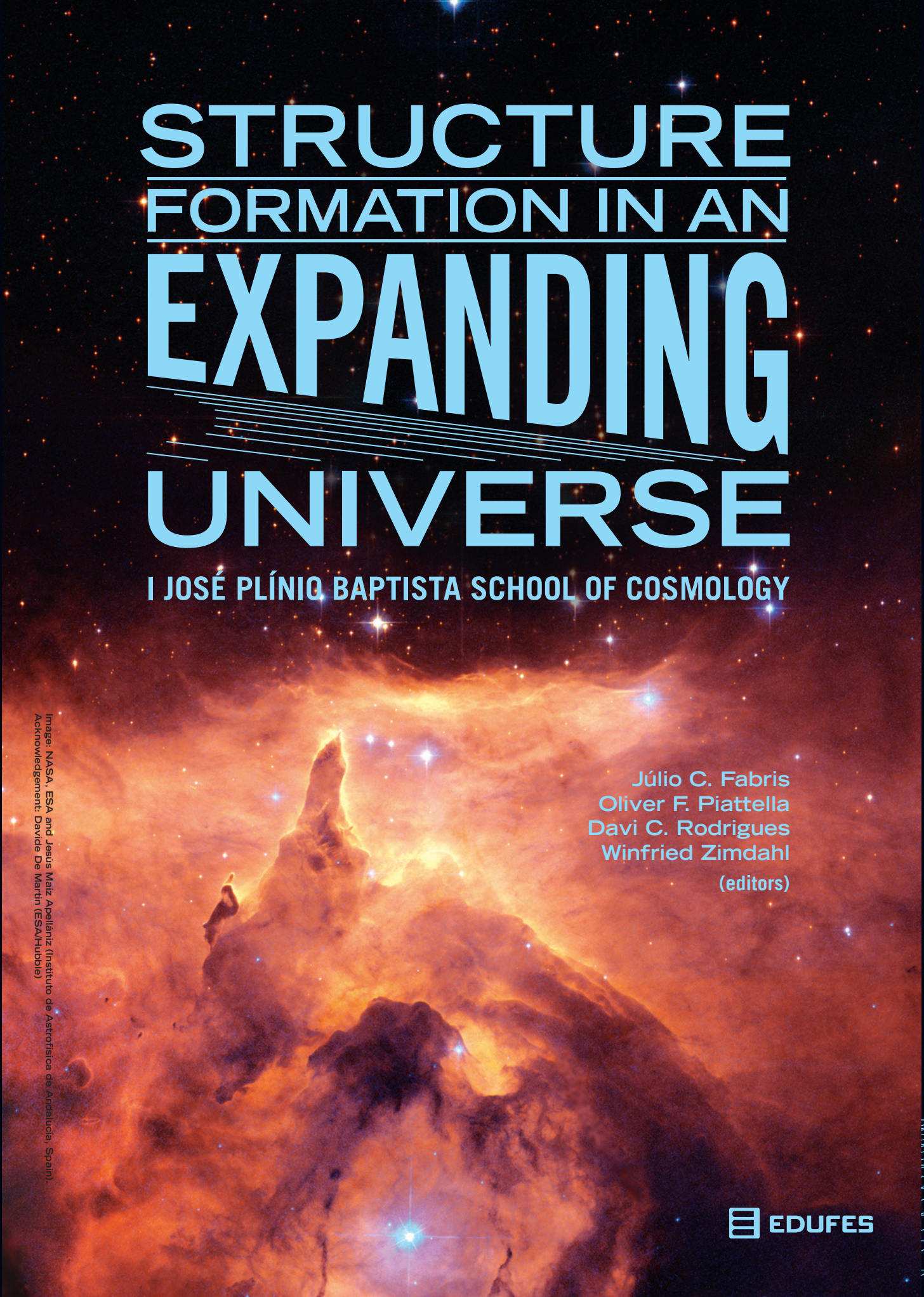
STRUCTURE FORMATION IN AN EXPANDING UNIVERSE

I JOSÉ PLÍNIO BAPTISTA SCHOOL OF COSMOLOGY



Júlio C. Fabris
Oliver F. Piattella
Davi C. Rodrigues
Winfried Zimdahl
(editors)

EDUFES

Image: NASA, ESA and Jesús Maíz Apellániz (Instituto de Astrofísica de Andalucía, Spain).
Acknowledgement: Davide De Martin (ESA/Hubble)

STRUCTURE FORMATION IN AN EXPANDING UNIVERSE

EDUFES
EDITORA DA UNIVERSIDADE
FEDERAL DO ESPÍRITO SANTO

Editora filiada à Associação Brasileira das Editoras Universitárias (Abeu)
Av. Fernando Ferrari, 514, Campus de Goiabeiras
CEP 29 075 910, Vitória - Espírito Santo, Brasil
Tel.: +55 (27) 4009-7852, E-mail: edufes@ufes.br
www.edufes.ufes.br

Reitor | Reinaldo Centoducatte
Vice-Reitora | Ethel Leonor Noia Maciel
Superintendente de Cultura e Comunicação | Ruth de Cássia dos Reis
Secretário de Cultura | Rogério Borges de Oliveira
Coordenador da Edufes | Washington Romão dos Santos

Conselho Editorial
Agda Felipe Silva Gonçalves, Cleonara Maria Schwartz, Eneida Maria Souza Mendonça, Gilvan Ventura da Silva, Glícia Vieira dos Santos, José Armínio Ferreira, Julio César Bentivoglio, Maria Helena Costa Amorim, Rogério Borges de Oliveira, Ruth de Cássia dos Reis, Sandra Soares Della Fonte

Secretário do Conselho Editorial | Douglas Salomão

Projeto Gráfico, Diagramação e Revisão Final | Os organizadores
Capa | Willi Piske Jr.

Dados Internacionais de Catalogação-na-publicação (CIP)
(Biblioteca Central da Universidade Federal do Espírito Santo, ES, Brasil)

J83s José Plínio Baptista School of Cosmology (1. : 2012 : Anchieta, ES)
Structure formation in an expanding universe / Júlio C. Fabris ... [et al.], editors. – Vitória : EDUFES, 2014.
292 p. : il. ; 27 cm

Seminário realizado no período de 14 a 19 de outubro de 2012.
ISBN : 978-85-7772-191-7

1. Cosmologia – Congressos. 2. Gravitação – Congressos. 3. Universo em expansão – Congressos. 4. Astronomia – Congressos. 5. Astrofísica – Congressos. I. Fabris, Júlio César. II. Título.

CDU: 52

STRUCTURE FORMATION IN AN EXPANDING UNIVERSE

I JOSÉ PLÍNIO BAPTISTA SCHOOL OF COSMOLOGY

Ubu - Anchieta, ES, Brazil
October 14-19, 2012

Júlio C. Fabris
Oliver F. Piattella
Davi C. Rodrigues
Winfried Zimdahl

(editors)

CAPES - Coordenação de aperfeiçoamento de pessoal de nível superior
CNPq - Conselho nacional de desenvolvimento científico e tecnológico
FAPES - Fundação de amparo à pesquisa do Espírito Santo
UFES - Universidade Federal do Espírito Santo

(support)

EDUFES
Vitória, 2014

Apresentação

A Cosmologia atravessa uma fase de grande efervescência graças a significativos avanços tanto do ponto de vista teórico quanto do ponto de vista observacional. Até algumas décadas atrás, a Cosmologia era essencialmente especulativa. Hoje, no entanto, a quantidade e a qualidade de dados observacionais requer de qualquer proposta teórica neste domínio um alto nível de sofisticação e de precisão. Tornou-se inclusive comum enfatizar que a Cosmologia abandonou o seu caráter puramente especulativo para se tornar uma ciência no sentido amplo do termo. Antigos problemas estão podendo ser resolvidos, e novos e mais desafiadores problemas estão surgindo requerendo uma compreensão mais profunda das teorias usadas para descrever o universo, entre elas, e principalmente, a Relatividade Geral, complementada em muitos momentos pela Teoria Quântica de Campos.

A Escola José Plínio Baptista de Cosmologia (JPBCosmo) foi concebida visando fornecer uma visão ampla de tópicos de atualidade em Cosmologia e Gravitação. O objetivo é apresentar cursos e seminários avançados concentrados em um tópico específico em Cosmologia, que seja de grande atualidade. Assim, a JPBCosmo é uma escola temática. Ela se destina a estudantes seguindo formação no domínio da Cosmologia, e a pesquisadores atuando nessa área do conhecimento. Devido ao caráter fortemente interdisciplinar da Cosmologia, a escola JPBCosmo pretende atrair estudantes e pesquisadores de áreas correlatas, como Astrofísica, Astronomia e Física de Altas Energias.

A JPBCosmo é uma atividade promovida pelo grupo de pesquisa em Gravitação e Cosmologia da UFES, Brasil, contando com apoio institucional da própria UFES como também das diversas agências de fomento à pesquisa existentes no Espírito Santo e, de uma forma mais geral, no Brasil. Esta escola traz o nome do prof. José Plínio Baptista, que foi um dos fundadores desse grupo de pesquisa em meados da década de 70, o que o torna um dos grupos de pesquisa mais antigos em atividade no Brasil nesta área do conhecimento científico. Prof. Plínio, como ele era chamado, faleceu em 2006. Ao dar o seu nome a esta escola, pretendemos homenageá-lo pelo seu pionerismo e pelo seu entusiasmo por esta ciência, a Cosmologia.

A primeira escola JPBCosmo foi organizada em Ubu, Espírito Santo, em outubro de

2012, tendo sido dedicada ao problema de formação de estruturas no universo. Tendo periodicidade bi-anual, a próxima escola está prevista para março de 2014, sendo agora dedicada aos estudos da radiação cósmica de fundo. A publicação dos textos oriundos dos cursos e palestras ministrados durante cada escola pretende oferecer aos estudantes e pesquisadores atuando na área uma descrição do estado de arte dos mais importantes e atuais problemas em Gravitação e Cosmologia.

Agradecemos em especial aos palestrantes e participantes pelo seu entusiasmo, promovendo um ambiente agradável e produtivo durante a semana que passamos em Ubu. A escola recebeu apoio essencial para sua realização vindo dos órgãos CNPq, CAPES, FAPES e UFES. Agradecemos também à EDUFES pelas orientações, diagramação e pela publicação do livro.

Os organizadores.

Preface

The problem of formation of structure in an expanding universe is one of the most important issues today not only in cosmology but also in physics in general. This problem concerns the origin and evolution of the observed structures in the universe, like galaxies, clusters of galaxies, filaments, voids. Since the universe is expanding and is globally homogenous and isotropic, the understanding process of local agglomeration of matter presents many challenges requiring fundamental additions to the standard cosmological model. These challenges became more strong due to the existence today of a plethora of observational data, that allow to map how matter is distributed at local scales. It is necessary to have a theoretical explanation of these highly detailed map of structures of the universe.

In the last decades a lot of effort has been devoted to this problem. It was quickly realized that it is necessary to provide a mechanism to generate the initial perturbations that would lead to the observed structures, fitting the observed spectrum of matter agglomeration. It was one of the great success of theoretical cosmology to give a convenient and elegant answer to this issue through the concept of quantum fluctuations in the primordial universe. Given this initial spectrum, it is necessary to follow the evolution of the perturbations either at linear and non-linear levels. In doing so, it became evident the necessity of a new matter component in the universe that could lead to the amplification of the initial, tiny quantum fluctuations to the huge structures observed in the universe. This component has been called dark matter, and it can agglomerate even in the deep radiative phase of the universe, when any fluctuations in the baryonic ordinary matter are strongly suppressed. The dark matter paradigm became an important part of the standard structure formation scenario, even if its nature is yet object of debate, and no direct detection of it has been unambiguously reported until now.

All this theoretical aspects of the structure formation became still more intriguing due to observational programs trying to map how matter agglomerates in the universe. The 2dFRGS and SDSS projects are two examples of the efforts made in this sense, giving accurate data of the distribution of matter. More recently, other projects in this sense are running. Together with this observational programs, the simulations

trying to reproduce the matter agglomeration patterns has knew impressive progress, even if important gaps must be still be filled. The observations of the anisotropy of the cosmic microwave background radiation (CMB) through the satellites WMAP and Planck furnish important clues to understand the initial spectrum and some aspects of its subsequent evolution. Now, due to all these observational data any theoretical proposal to explain the structures of the universe must pass by the confrontation with the observational data, many of them highly precise.

The series of texts composing this book are based on the lectures presented during the I José Plínio Baptista School of Cosmology, held in Ubu, Espírito Santo, Brazil, between 14 and 19 october 2012. This I JBPCosmo has been entirely devoted to the problem of the understanding of structure formation in the universe. These lectures try to give a general overview on the many aspects of this problem. It begins with a revision of the quantum mechanism for the origin of the perturbations, followed by the linear perturbation theory, the non-linear process, up to more recent efforts in the sense to understand the patterns of structures at local scales through huge numerical simulations. A series of advanced seminars on specific topics completes this general overview. These texts are addressed to students following their formation in gravitation and cosmology, and to researchers working in these domain.

We want thank specially the speakers and the participants for their enthusiasm, providing a very nice ambiance during the week we have passed together in Ubu, discussing this important problem of modern cosmology. The I JBPCosmo has been supported by CNPq, CAPES, FAPES and UFES. We also thank EDUFES for the advices, layout and for publishing the book.

The editors.

Contents

Mini courses

Theory of Primordial Cosmological Perturbations

Nelson Pinto-Neto[a],[1] Carolina M.S. Barbosa[b],[2] and Júlio.C. Fabris[c2]

[1] *ICRA-CBPF, Rua Xavier Sigaud, 150,*
Urca, Rio de Janeiro, CEP: 22290-180,
Rio de Janeiro, Brazil

[2] *Universidade Federal do Espírito Santo,*
Departamento de Física
Av. Fernando Ferrari,
514, Campus de Goiabeiras,
CEP 29075-910, Vitória,
Espírito Santo, Brazil

One of the main tasks of a cosmological theory is to explain the formation of the observed structures in the universe, like galaxies, voids and clusters of galaxies. To reach this goal is necessary to follow the evolution of perturbations originated in the primordial universe until today. It is believed, in the context of the standard cosmological scenario, that the spectrum associated to the structures observed today can be explained if the initial perturbations originated from quantum fluctuations in the primordial universe. Such mechanism can be implement if there was an initial inflationary phase in the begining of the history of the universe, or if the initial cosmological singularity is replaced by a quantum bounce. The theory of primordial perturbations of quantum origin is revised, using the gauge invariant variables. The formalism is applied to both alternative scenarios for the primordial universe.

I. INTRODUCTION

In cosmology the explanation of the structures observed in the universe at large scales is one of the main challenges to be surmounted. Structures are formed in the matter dominated period of evolution of the universe from fluctuations in a globally homogeneous and isotropic background. One of the main questions to be answered concerns the origin of the initial fluctuations that evolves latter to give birth to galaxies, clusters of galaxies, etc. Pure statistical fluctuations on the matter distribution are

[a] E-mail: nelson.pinto@pq.cnpq.br
[b] E-mail: carolmsb@gmail.com
[c] E-mail: fabris@pq.cnpq.br

insufficient to lead to the observed structures [1]. The standard cosmological lore gives a very elegant way to surmount this drawback: the observed structures are consequences of small primordial fluctuations of quantum origin, that are amplified during the evolution of the universe. In order to construct a viable scenario in this sense, it is necessary to suppose the existence of primordial quantum fields that result in these initial quantum seeds for the observed structures. These primordial quantum fields must, at the same time, drive an inflationary cosmic phase, or a quantum bounce near the initial singularity.

The inflationary mechanism was conceived initially as a way to solve the initial condition problems connected with isotropy, homogeneity and flatness of the universe. The inflationary scenario predicts a primordial phase of exponential expansion - a quasi-de Sitter phase - resulting in an isotropic and flat observed universe in large scales [2]. However, it does not give a satisfactory explanation to the homogeneity problem [3]. This exponential expansion is driven by a scalar field, the inflaton, whose origin must be traced back to the fundamental fields of the grand-unified theories, for example. Quantum fluctuations of the inflaton field may be the seeds for the structure formation.

Another mechanism evoked in the literature more recently is the existence of a contraction phase prior the expansion phase, the so-called bouncing scenarios [4]. The bounce can also be driven by quantum effects [5–8], and primordial quantum vacuum fluctuations can also lead to structures in the universe, and to a spectrum of initial perturbations in good agreement with observations, as in the inflationary case.

One main task of cosmological perturbation theory is to describe the origin of the primordial perturbations, predicting the resulting spectrum, and following the evolution of these perturbations up to the formation of the structures in the universe. This program of research has been inaugurated in the seminal paper of Lifshitz in 1946 [9, 10] using the synchronous coordinate condition. From that moment to now, enormous progress has been achieved, both from the point of view of the formalism employed in the investigation of the perturbations as well as in the construction of the observables quantities. The present text focuses on the construction of perturbed quantities in the context of General Relativity Theory, the definition of the observables,

mainly the primordial spectrum, and in the mechanism of quantum fluctuations as primordial seeds of the structures of the universe. In doing so, we follow the so called *gauge invariant* approach, first proposed by Bardeen [11], and set out in details by Mukhanov et al. [12, 13]. In this approach, the quantization of the perturbations can be performed very elegantly. The mechanism of primordial perturbations of quantum origin has been initially proposed mainly by Lukash and Mukhanov [14–16], and it is now one of the cornerstones of the structure formation scenario. However, it must be stressed that the same quantization of the primordial perturbations can be performed in the original synchronous coordinate condition used originally by Lifshitz [17].

The goal of these lectures is to provide a self-contained introduction to linear perturbation theory, the gauge invariant formalism and the application to the determination of the primordial spectrum, both for inflationary and bounce models.

II. THE UNPERTURBED UNIVERSE

Our standard theory of gravity is represented by General Relativity which is described by the Einstein's equations:

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} R = 8\pi G T_{\mu\nu}, \tag{1.1}$$

$$T^{\mu\nu}{}_{;\mu} = 0, \tag{1.2}$$

where $R_{\mu\nu}$ is the Ricci tensor, R is the Ricci scalar and $T_{\mu\nu}$ is the energy-momentum tensor. In General Relativity, gravity is represented by the curvature of the four-dimensional space-time which is sourced by the distribution of matter and energy. The geometry is described by the infinitesimal spatial-temporal distance, given by

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu, \tag{1.3}$$

where $g_{\mu\nu}$ is the metric of the space-time.

The Ricci tensor and the Ricci scalar are constructed out of the metric:

$$R_{\mu\nu} = \partial_\rho \Gamma^\rho_{\mu\nu} - \partial_\nu \Gamma^\rho_{\mu\rho} + \Gamma^\rho_{\mu\nu} \Gamma^\sigma_{\rho\sigma} - \Gamma^\rho_{\mu\sigma} \Gamma^\sigma_{\nu\rho}, \tag{1.4}$$

$$R = g^{\mu\nu} R_{\mu\nu}, \tag{1.5}$$

$$\Gamma^\rho_{\mu\nu} = \frac{1}{2} g^{\rho\sigma} (\partial_\mu g_{\sigma\nu} + \partial_\nu g_{\sigma\mu} - \partial_\sigma g_{\mu\nu}). \tag{1.6}$$

In the last expression $g^{\mu\nu}$ is the inversed metric, such that $g^{\mu\rho}g_{\rho\nu} = \delta^{\mu}_{\nu}$, where δ^{μ}_{ν} is the Kronnecker's delta.

The matter content, in the right hand side of the Einstein's equation, is described by the energy-momentum tensor. For a perfect fluid, the energy-momentum tensor is given by

$$T_{\mu\nu} = (\rho + p)u_{\mu}u_{\nu} - pg_{\mu\nu}, \tag{1.7}$$

where ρ is the density, p is the pressure and u_{μ} is the four-velocity associated to the fluid. For a self interacting scalar field ϕ, which plays a fundamental rôle in the primordial universe, the energy-momentum tensor reads,

$$T_{\mu\nu} = \phi_{;\mu}\phi_{;\nu} - \frac{1}{2}g_{\mu\nu}\phi_{;\rho}\phi^{;\rho} + g_{\mu\nu}V(\phi), \tag{1.8}$$

where $V(\phi)$ is the potential term, specifying the self-interacting model.

The Einstein's equations described above can be obtained from an action by using the variational principle. The action is given by,

$$\mathcal{A} = \int d^4x(\sqrt{-g}R + \mathcal{L}_m), \tag{1.9}$$

where $\mathcal{L}_m$ is the Lagrangian density leading to the energy-momentum tensor by the relation,

$$T^{\mu\nu} = -\frac{2}{\sqrt{-g}}\frac{\delta\mathcal{L}_m}{\delta g_{\mu\nu}}. \tag{1.10}$$

The Einstein's equations can be generalized in order to include a cosmological term Λ, leading to the most general gravitational equations in four dimensions containing up to second derivatives and satisfying the Bianchi's identities:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R - \Lambda g_{\mu\nu} = 8\pi G T_{\mu\nu}. \tag{1.11}$$

The Bianchi's identities imply that the covariant divergence of such equation is zero, leading the conservation of the energy-momentum tensor. The cosmological constant can be put in the right hand side of the Einstein's equation, being possible to interpret it as a perfect fluid with an equation of state $p = -\rho$. This fact allows to consider

a cosmological term as a manifestation of the quantum vacuum energy, an important interpretation in consctructing the inflationary scenario, as we will see later.

The universe is supposed to be homogenous and isotropic at large scales. This hypothesis is supported by many observations, as a large scale map of galaxies in the Universe, which indicates that there are evidences for approximate homogeneity and isotropy for scales higher than $100\,Mpc$. The high isotropy of the cosmic microwave background radiation (CMBR), by a factor of 10^{-5}, is also evoked. The CMBR is a map of the universe of about 13 billions years ago, when radiation decoupled from the baryonic component.

A homogeneous and isotropic universe can be represented by the Friedmann-Lemaître-Robertson-Walker metric. This is named the *background metric*, representing the properties of the universe out of the small, local, inhomogeneities. The background metric is given by,

$$\begin{aligned} ds_0^2 &= dt^2 - a^2(t)\gamma_{ij}dx^i dx^j, \\ &= a^2(\eta)[-d\eta^2 + \gamma_{ij}dx^i dx^j], \end{aligned} \tag{1.12}$$

where a is the scale factor, t is the cosmic time, η is the conformal time, and γ_{ij} is the metric on the spatial three-dimensional section with constant curvature. The conformal time is related to the cosmic time t by the relation $dt = ad\eta$. The conformal time, as we will see later, is particularly convenient in the perturbative analysis of cosmological model.

This metric (1.12), when applied to the Einstein's equations, leads to the following equations of motion:

$$\left(\frac{a'}{a}\right)^2 + \frac{k}{a^2} = \frac{8\pi G}{3}a^2\rho, \tag{1.13}$$

$$2\frac{a''}{a} - \left(\frac{a'}{a}\right)^2 + \frac{k}{a^2} = -8\pi G a^2 p, \tag{1.14}$$

$$\rho' + 3\frac{a'}{a}(\rho + p) = 0. \tag{1.15}$$

The primes mean derivatives with respect to the conformal time. These equations can be combined leading to the so-called Raychaudhuri's equation,

$$\frac{a''}{a} - \left(\frac{a'}{a}\right)^2 = -\frac{4\pi G}{3}(\rho + 3p). \tag{1.16}$$

In the cosmic time parametrization, the Raychaudhuri's equation reads,

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}(\rho + 3p). \tag{1.17}$$

This equation gives the condition to have a superluminal ($\rho + 3p < 0$) or subluminal ($\rho + 3p > 0$) expansion.

The solution of the equations (1.13,1.14,1.15) depends on the choice of the fluid component. This choice is characterized by the specification of the equation of state of the fluid. Let us suppose a perfect fluid with an equation of state $p = \omega\rho$. The solutions for the matter density and for the scale factor read,

$$\rho = \rho_0 a^{3(1+\omega)}, \quad a = a_0 \eta^{\frac{2}{1+3\omega}}. \tag{1.18}$$

where ρ_0 and a_0 are constants. For pressureless matter, $\omega = 0$, $a \propto \eta^2$. For a radiative fluid, $\omega = \frac{1}{3}$, $a \propto \eta$. In these both important cases, as we will see later, the universe expands decelerating. In order to have an accelerated expansion, $\omega < -\frac{1}{3}$. Parametrizing the scale factor solution as,

$$a = a_0 \eta^{1+\beta}, \quad \beta = \frac{1 - 3\omega}{1 + 3\omega}, \tag{1.19}$$

the accelerated expansion means $-\infty < \beta < -2$ for $-1 < \omega < -1/3$, $-2 < \beta < -1$ for $\omega < -1$ (the phantom case), with $-\infty < \eta < 0_-$ in both cases. For $\beta > -1$, the universe expands decelerating, and $0 < \eta < \infty$. An accelerated expansion is called an *inflationary phase*. An important case is given by $\beta = -2$, which corresponds to the de Sitter phase. In the cosmic time parametrization, the de Sitter phase is described by an exponential expansion, $a \propto e^{Ht}$, H being constant.

The standard cosmological model contains two main phases: a radiation dominated universe with $\omega = \frac{1}{3}$ and $\rho_r \propto a^{-4}$; and a matter dominated phase with $\omega = 0$ and $\rho_m \propto a^{-3}$. The evidences for the radiation dominated phase, which takes place firstly, are the primordial nucleosynthesis and the existence of a cosmic microwave background radiation; the evidences for a matter dominated phase rely on the existence of structures in the universe (galaxies, clusters of galaxies, etc.) dominated by pressureless matter, which must grow by gravitational instability, a mechanism that requires a pressure equal or very near zero.

More recently, it was discovered that the present universe is in an accelerated expansion stage. This implies that the evolution of the universe entered, after the matter dominated phase, in a period dominated by an exotic fluid with negative pressure, called *dark energy*, which implies $\omega < -\frac{1}{3}$. Such present accelerated phase corresponds to a second inflationary regime in the history of the universe, since in the begining of the history of the universe a primordial inflationary phase may have ocurred. The evidences for this present accelerated phase comes mainly from the supernova type Ia observations and the spectrum of the anisotropies in the CMBR. The nature of this *dark energy* component is object of intense debate now. All proposals made to account for this new component of unknown nature face successes and drawbacks. It seems clear, on the other hand, that dark energy must have negative pressure, violating the energy conditions, since under these conditions it can accelerate the universe, without spoil the structure formation mechanism. Another option, is to consider General Relativity as an incomplete theory even at classical level. For a detailed discussion about dark energy, see Ref. [18].

The standard cosmological model explains succesfully the primordial nucleosynthesis, the existence of the cosmic background radiation, and the structure formation. This last process requires, however, the existence of another exotic component, called *dark matter*, which has no baryonic origin, but must exhibit zero pressure or almost zero pressure. Dark matter is the second *exotic* component of the standard cosmological model, together with dark energy. Its nature is also object of intense debate, like in the dark energy case. Hence, the standard cosmological model requires the existence of four components: radiation, baryons, dark matter and dark energy. We must add to this cosmic budget a fifth component, the neutrinos, that play an important rôle in the primordial universe.

However, the standard cosmological model has important drawbacks essentially connected with the initial conditions necessary to fit the observed universe today. They are the following:

1. The horizon problem. The universe is, in the transition from the radiative phase to the matter phase, or even before, in thermal equilibrium even if it is composed

of many causally disconnected regions during that transition.

The origin of the horizon problem is connected with propagation of light in an expanding universe. The physical distance covered by a light ray from the origin to a given time t, called *horizon* is, for a universe that evolves as a power law $a \propto t^n$,

$$d_l = a(t) \int_0^t \frac{dt'}{a(t')} \propto t. \tag{1.20}$$

However, if the universe evolves in a subluminal expansion, such that $n < 1$, the physical distances grow slowly than the horizon. Hence, scales that are today inside a horizon were outside the horizon in the past, without casual contact. The high isotropy of the CMBR implies, on the other hand, that there was thermal equilibrium for these regions even in time they were not connected causally. This requires the imposition of very particular initial conditions.

2. The flatness problem. The universe is, today, essentially spatially flat. However any small initial departure of flatness in the beginning of the history of the universe would imply a very non-flat universe today. Essentially, the universe should have "born" in a very flat state in order to fit the present observations.

3. The origin of the fluctuation that led to the observed structure of the universe today can not be naturally implemented, since it requires non-causal processes and an ad-hoc initial spectrum. This is connected somehow to the horizon problem described before. The horizon distance grows linearly with the cosmic time t. But, the perturbation scales grows with the scale factor $a \propto t^n$. Since the perturbations leading today the observed structures are inside the horizon, in the very past they should be outside the horizon for a subluminal expanding universe, requiring a non-casual mechanism for their generation. In the standard cosmological model described until now, the universe expands in a subluminal regime always, and this problem is inevitable.

4. There is an initial singularity, where the curvature, density and temperature diverge, leading to the breakdown of the usual known physical laws.

These problems can be treated in two alternative ways:

1. Supposing an initial, short, but strongly accelerated phase, called *inflationary phase*, driven by an hypothetical inflaton field, which is essentially a self interacting scalar field. The potential term dominates over the kinetic term during the inflationary phase. In its original formulation, the inflationary scenario was based in a de Sitter for which $p = -\rho$, with $\rho =$ constant, due to the conservation equation, implying an exponential expansion for the universe. But, the models based on phase transitions of fundamental fields in the primordial universe uses the less stringent condition $p \sim -\rho$, a quasi-de Sitter phase. For more details on this proposal, see the reference [19]

2. Supposing a contracting phase, prior to the present expansion phase, the transition from the contraction phase to the expansion phase being assured by quantum effects near the minimum of the scalar factor, called bounce. For a review of this proposal see, for example, the references [20, 21].

The inflationary phase addresses conveniently the first three points listed above. However, it has no word to say about the forth one, as well as for the initial homogeneity conditions. The bounce universe addresses, in principle, successfully the four drawbacks listed above, but requires some special conditions to have sufficient homogeneity and isotropy.

III. THE PERTURBED UNIVERSE AND THE GAUGE INVARIANT VARIABLES

The universe described in the previous section concerns the average background. The universe seems to display homogeneity and isotropy from scales of $100\,Mpc$ on. However, at small scales, the observed universe shows a large variety of inhomogeneities, galaxies, clusters of galaxies, voids, filaments, the great wall, etc. One of the main aims of cosmology is to explain all the complex structures, using a given theory of gravity, initial perturbations on the homogenous and isotropic universe and a mechanism of amplification of these initial perturbations.

Our standard gravity theory is General Relativity, described in the previous section. Its general features for the description of the universe were shown before, leading to the Standard Cosmological Model, given by a radiative phase, a matter dominated phase and the present accelerated phase. This forms the essential steps in the evolution of the universe described by the ΛCDM model. The ΛCDM model must be complemented by the primordial cosmological scenario, what is a source of intense debate, since it is expected that quantum effects must play an important rôle, and the inclusion of quantum effects in cosmology is still a controversial issue.

In general, the inflationary model is the most popular way to complete the Standard Cosmological Model with a primordial phase. Inflation has many important open questions (origin of the inflaton, construction of a realistic model inspired on fundamental theories, etc.), besides which it does not solve the singularity and homogeneity problems. In this sense, bouncing models appear as viable alternative. This issue will be discussed later.

The inhomogeneities observed at small scales are supposed to have originated from the gravitational instability process, driven by initial small fluctuations of the quantum fields in primordial universe. But, let us first review how we perturb a homogenous and isotropic universe.

Given a background which is homogeneous and isotropic, for example the scenarios described in the previous section, we can introduce local fluctuations on it. The background is represented by a metric $g_{\mu\nu}$, a density ρ, a velocity field u^μ, and a pressure p. Hence, the fluctuations lead to,

$$\begin{aligned} \tilde{g}_{\mu\nu} &= g_{\mu\nu} + h_{\mu\nu}, \quad \tilde{\rho} = \rho + \delta\rho, \\ \tilde{u}^\mu &= u^\mu + \delta u^\mu, \quad \tilde{p} = p + \delta p, \end{aligned} \tag{1.21}$$

where $h_{\mu\nu}$, $\delta\rho$, δu^μ and δp are the local fluctuations on the metric, density, velocity and pressure. We will call from now on these fluctuations as *perturbations*.

One fundamental aspect of linear perturbations is that they can separated into three classes: scalar, vector and tensor perturbations. In order to illustrate this splitting, let us focus on the metric perturbations, since the other perturbations are coupled to them.

1. Scalar perturbations:

$$ds^2 = a^2(\eta)[-(1+2\phi)d\eta^2 + 2B_{|i}dx^i d\eta + [(1-2\psi)\gamma_{ij} + 2E_{|i|j}]dx^i dx^j\} \quad (1.22)$$

There are generally four metric scalar modes: ϕ,ψ, B and E. The bar in the sub-indices means covariant derivative with respect to the spatial hyper-surface with constant curvature.

2. Vector perturbations:

$$\delta g^V_{\mu\nu} = a^2(\eta) \begin{pmatrix} 0 & -\delta_i \\ -\delta_j & F_{i|j} + F_{j|i} \end{pmatrix}, \quad (1.23)$$

with $\delta^i_{|i} = 0$ and $F^i_{|i} = 0$

3. Tensor perturbations:

$$\delta g^T_{\mu\nu} = a^2(\eta) \begin{pmatrix} 0 & 0 \\ 0 & W_{ij} \end{pmatrix}, \quad (1.24)$$

with $W^j_{i|j} = 0$, $W_{ij}\gamma^{ij} = 0$.

In all the expressions above, the bar "—" means covariant derivative with respect to the spatial hypersurface. The decomposition into scalar, vectorial and tensorial modes have a clear physical meaning. The scalar modes are connected with density perturbations, while the vectorial mode are connected with rotations and the tensorial modes are related to the propagation of gravitational waves.

The metric containing simultaneously the three types of metric perturbations is given by,

$$\begin{aligned} ds^2 = a^2(\eta)\{&-(1+2\phi)d\eta^2 + (2B_{,i} + 2\delta_i)dx^i d\eta \\ &+ [(1-2\psi)\gamma_{ij} + 2E_{|i|j} + F_{i|j} + F_{j|i} + W_{ij}]dx^i dx^j\}. \end{aligned} \quad (1.25)$$

Among the functions presented in the perturbed metric described above, there are some that have no physical meaning, representing just a coordinate deformation of the

background metric. One important problem in perturbation theory is how to identify and to eliminate these non-physical functions. This can be done in many different ways. One can, following a similar problem appearing in classical electrodynamics, try to fix some coordinate conditions (corresponding to gauge conditions in electrodynamics). Or one can construct dynamical variables that are invariant by coordinate transformations.

Let us first revise some important concepts connected with this issue. A crucial one is the notion of Lie derivative.

Let us consider the infinitesimal coordinate transformation such that,

$$\bar{x}^\alpha = x^\alpha + \xi^\alpha(x). \tag{1.26}$$

Under this transformation, the metric transforms as,

$$\delta g_{\mu\nu} = \bar{g}_{\mu\nu}(x^\alpha) - g_{\mu\nu}(x^\alpha) = -\xi_{\mu;\nu} - \xi_{\nu;\mu}. \tag{1.27}$$

A vector field transforms as,

$$\bar{A}_\mu(x) - A_\mu(x) = \xi^\alpha \partial_\alpha A_\mu(x) - A^\alpha(x)\partial_\alpha \xi_\mu. \tag{1.28}$$

The covariant derivative of the transformation parameter can be expanded as,

$$\xi_{\mu;\nu} = \xi_{\mu,\nu} - \Gamma^\alpha_{\mu\nu}\xi_\alpha \tag{1.29}$$

Moreover, the transformation parameter ξ^α can have a scalar and vectorial nature:

- Scalar

$$\xi^\alpha = (\xi^0, \gamma^{ij}\xi_{,j}). \tag{1.30}$$

- Vector:

$$\xi^\alpha_\nu = (0, \xi^i_{TR}), \tag{1.31}$$

with

$$\xi^i_{TR,i} = 0. \tag{1.32}$$

We must investigate now how the metric components transform under such infinitesimal coordinate transformation. We will from now on concentrate on scalar perturbations, and take all the expressions for a spatially flat universe for which, $\gamma_{ij} = \delta_{ij}$.

For the component g_{00} of the metric, we have,

$$\delta g_{00} = -2a^2(\bar{\phi} - \phi) = -2\xi_{0;0}. \tag{1.33}$$

Remembering that,

$$\xi_{0;0} = -a^2\xi^{0'} - aa'\xi^0, \tag{1.34}$$

we find

$$\bar{\phi} = \phi - \xi^{0'} - \frac{a'}{a}\xi^0. \tag{1.35}$$

From now on, let us concentrate on the scalar perturbations, since they are directly connected with density perturbations. For the component g_{0i} of the metric, we have,

$$\delta g_{0i} = a^2(\bar{B}_{,i} - B_{,i}) = -\xi_{0;i} - \xi_{i;0}, \tag{1.36}$$

leading to,

$$\bar{B} = B + \xi^0 - \xi' \tag{1.37}$$

Finally, for the component g_{ij}, we have

$$\delta g_{ij} = 2a^2[(\bar{E} - E)_{,i,j} - (\bar{\psi} - \psi)\gamma_{ij}] = -\xi_{i;j} - \xi_{j,i}, \tag{1.38}$$

implying,

$$\bar{E} = E - \xi, \quad \bar{\psi} = \psi + \frac{a'}{a}\xi^0. \tag{1.39}$$

We define the new variables,

$$\Phi_B = \phi + \frac{1}{a}[(B - E')a]', \quad \psi_B = \psi - \frac{a'}{a}(B - E'). \tag{1.40}$$

They are called the Bardeen's variable, and they are invariant by infinitesimal transformations, as described above. For a scalar field, following similar considerations, we have,

$$\delta\varphi = -\varphi_{,\alpha}\xi^{\alpha} = -\varphi'\xi^{0}, \tag{1.41}$$

$$\delta\varphi^{(c)} = \delta\varphi + \varphi^{0'}(B - E'). \tag{1.42}$$

The quantity $\delta\varphi^{c}$ is also invariant with respect to these gauge transformations.

IV. THE PERTURBED EINSTEIN'S EQUATIONS

Let us go back to the Einstein's equations. They can be written as

$$G^{\nu}_{\mu} = kT^{\mu}_{\nu}, \tag{1.43}$$

$$k = 8\pi G \tag{1.44}$$

At the linear level, following the same procedure as before, we have the perturbed equation:

$$G^{\mu}_{\nu} + \delta G^{\mu}_{\nu} = k(T^{\mu}_{\nu} + \delta T^{\mu}_{\nu}), \tag{1.45}$$

$$\delta G^{\mu}_{\nu} = k\delta T^{\mu}_{\nu}, \tag{1.46}$$

$$\delta\bar{G}^{\mu}_{\nu} = \delta G^{\mu}_{\nu} - G^{\mu}_{\nu,\alpha}\xi^{\alpha} - G^{\mu}_{\alpha}\xi^{\alpha}_{,\nu} - G^{\alpha}_{\nu}\xi^{\mu}_{,\alpha}, \tag{1.47}$$

$$\delta\bar{G}^{0}_{0} = \delta G^{0}_{0} - G^{0'}_{0}\xi^{0}, \tag{1.48}$$

$$\delta\bar{G}^{0}_{i} = \delta G^{0}_{i} - \left(G^{0}_{0} - \frac{1}{3}G^{k}_{k}\right)\xi^{0}_{,i}, \tag{1.49}$$

$$\delta\bar{G}^{i}_{j} = \delta G^{i}_{j} - (G^{i}_{j})'\xi^{0}. \tag{1.50}$$

The gauge invariant perturbations of the Einstein's tensor can then be written:

$$\delta G^{(c)0}_{0} = \delta G^{0}_{0} + G^{0}_{0}(B - E'), \tag{1.51}$$

$$\delta G^{(c)0}_{i} = \delta G^{0}_{i} + (G^{(c)0}_{0} - \frac{1}{3}G^{k}_{k})(B - E')_{,i}, \tag{1.52}$$

$$\delta G^{(c)j}_{i} = \delta G^{j}_{i} + G^{j}_{i}(B - E'), \tag{1.53}$$

where the subscript c indicates that quantities are gauge invariant.

The perturbed components of the Einstein's tensor G^{μ}_{ν} are the following [12]:

$$\delta G^0_0 = \frac{2}{a^2}\Big\{ -\mathcal{H}(\mathcal{H}\phi + \psi') + \nabla^2\psi - 3\mathcal{H}(-\mathcal{H}' + \mathcal{H}^2)(B - E')\Big\}, \tag{1.54}$$

$$\delta G^0_i = \frac{2}{a^2}\partial_i[\mathcal{H}\phi + \psi' + (\mathcal{H}' - \mathcal{H}^2)(B - E')], \tag{1.55}$$

$$\delta G^i_j = -\frac{2}{a^2}\Big\{[(2\mathcal{H}' + \mathcal{H}^2)\phi + \mathcal{H}\phi' + \psi'' + 2\mathcal{H}\psi' + \frac{1}{2}\nabla^2 D + (\mathcal{H}" - \mathcal{H}\mathcal{H}' - \mathcal{H}^3)(B - E')]\delta^i_j$$
$$- \frac{1}{2}\partial_i\partial_j D\Big\}, \tag{1.56}$$

where $D = \phi - \Psi$ and $\mathcal{H} \equiv a/a$. In the absence of anisotropic pressure stress, the out-diagonal component of the perturbed energy-momentum tensor is zero, implying $\phi = \psi$, a result that we will use from now on.

If we fix $E = B = 0$ we have the perturbed equations in the *newtonian gauge*. In this gauge, the perturbed quantities ϕ and ψ coincides withe the gauge invariant perturbations Φ_B and Ψ_B. From now on, we will use the newtonian gauge. A long but quite direct calculation using the metric leads to the following equations:

$$\nabla^2\Phi - 3\mathcal{H}(\mathcal{H}\Phi + \Phi') + 3k\psi = \frac{k}{2}a^2\delta_c T^0_0, \tag{1.57}$$

$$(\Phi' + \mathcal{H}\Phi)_{,i} = \frac{k}{2}a^2\delta_{(c)}T^0_i, \tag{1.58}$$

$$[\Phi'' + 3\mathcal{H}\Phi' + (2\mathcal{H}' + \mathcal{H}^2)\Phi]\delta^i_j = -\frac{k}{2}a^2\delta T^i_j, \tag{1.59}$$

with $\phi = \psi = \Phi$.

The equations (1.57,1.58,1.59) are the fundamental perturbed equations we must solve. But first we must specify the right-hand-side, representing the perturbations on the energy-momentum tensor. Since we are interested in the primordial spectrum, we will concentrate, in the next section, on the energy-momentum tensor of a self-interacting scalar field.

V. SCALAR FIELDS

In order to be complete, we must specify the right-hand side of the perturbed Einstein's equations. In primordial cosmology, a special rôle is played by scalar fields.

Hence, let us consider the following action,

$$S = \int p(X,\phi)\sqrt{-g}d^4x, \tag{1.60}$$

where $X = \frac{1}{2}g^{\mu\nu}\partial_\mu\varphi\partial_\nu\varphi$ and $p(X,\varphi) = X - V(\varphi)$. The energy-momentum for this scalar Lagrangian is obtained from,

$$T_{\mu\nu} = -2\frac{\delta S}{\delta g^{\mu\nu}}, \tag{1.61}$$

leading to

$$T_{\mu\nu} = \varphi_{;\mu}\varphi_{;\nu} - \frac{1}{2}g_{\mu\nu}\varphi_{;\rho}\varphi^{;\rho} + g_{\mu\nu}V(\varphi). \tag{1.62}$$

The energy-momentum tensor of the scalar field can be identified with a perfect fluid energy-momentum tensor, such that,

$$T_{\mu\nu} = (\rho + p)u_\mu u_\nu - pg_{\mu\nu}, \tag{1.63}$$

$$\rho = \frac{1}{2}\varphi_{;\rho}\varphi^{;\rho} + V(\varphi), \tag{1.64}$$

$$p = \frac{1}{2}\varphi_{;\rho}\varphi^{;\rho} - V(\varphi), \tag{1.65}$$

$$u_\mu = \frac{\partial_\mu\varphi}{\sqrt{2X}}. \tag{1.66}$$

Using the cosmological metric, with the conformal time, the equations for a self-interacting scalar field coupled to gravity, reads:

$$3\left(\frac{a'}{a}\right)^2 = \frac{1}{2}\varphi'^2 + a^2V(\varphi), \tag{1.67}$$

$$\varphi'' + 2\frac{a'}{a}\varphi' = -a^2V_\varphi. \tag{1.68}$$

If the self-interacting term dominates over the kinetic term, the potential acts similarly to a cosmological constant term, and the conditions to have a inflationary behavior can be implemented.

The scalar field discussed above may represent the fundamental fields playing the dominant rôle in the primordial universe. In particular it can be the source of the inflationary phase. As it has already been stated, the main purposes of the inflationary phase is to lead to the initial tiny fluctuations which would later generate the structure

observed at large scale today. Hence, it is mandatory to investigate the perturbation issues of such gravity/scalar system.

To do so, we return to the perturbation theory sketched before. For simplicity, we will work in the longitudinal gauge for which we have $B = E = 0$. Hence, the Bardeen's variable become,

$$\Phi_B = \phi \quad , \quad \bar{\psi}_B = \Psi \quad , \quad \delta\varphi^{(c)} = \delta\varphi. \tag{1.69}$$

The perturbed metric takes the form,

$$ds^2 = a^2(\eta)[(1+2\phi)d^2\eta - (1-2\psi)\delta_{ij}dx^i dx^j]. \tag{1.70}$$

The Friedmann's equation, the Raychaudhuri's equation and the conservation's equation take the form,

$$\mathcal{H}^2 = \frac{8\pi G}{3}a^2\rho, \tag{1.71}$$

$$\mathcal{H}' - \mathcal{H}^2 = 4\pi G(\rho + p)a^2, \tag{1.72}$$

$$\rho' = \rho_{,X}X + \rho_{,\varphi}\varphi = -3\mathcal{H}(\rho + p). \tag{1.73}$$

In these expressions, we have made the identifications described before between the fluid and scalar field representation of the energy-momentum tensor.

At the perturbative level, we have,

$$\delta X = \frac{1}{2}\delta g^{00}\varphi_0'^2 + g^{00}\varphi_0'\delta\varphi' = 2X_0\left(-\phi + \frac{\delta\varphi'}{\varphi_0'}\right), \tag{1.74}$$

where φ_0 is a solution of the background.

The sound speed for a scalar field can be written as [23],

$$c_s^2 = \frac{p_{,X}}{\rho_{,X}} = \frac{\rho + p}{2X\rho_{,X}}. \tag{1.75}$$

The perturbation of the energy component of the scalar field energy-momentum tensor is,

$$\delta T_0^0 = \delta\rho = \rho_{,X}\delta X + \rho_{,\varphi}\delta\varphi. \tag{1.76}$$

The perturbed equations read now:

$$\nabla^2\psi - 3\mathcal{H}(\psi' + \mathcal{H}\phi) = 4\pi G a^2(\rho + p)\left\{\frac{1}{c_s^2}\left[\left(\left(\frac{\delta\varphi}{\varphi_0'}\right)' + \mathcal{H}\frac{\delta\varphi}{\varphi_0'} - \phi\right) - 3\mathcal{H}\frac{\delta\varphi}{\varphi_0}\right]\right\}, \tag{1.77}$$

$$\psi' + \mathcal{H}\phi = 4\pi G a^2(\rho + p)\frac{\delta\varphi}{\varphi_0'}. \tag{1.78}$$

Since

$$\delta T_j^i = 0, \quad i \neq j, \quad \rightarrow \phi = \psi, \tag{1.79}$$

one can write the equations above as

$$\nabla^2\psi = \frac{4\pi G a^2(\rho + p)}{c_s^2}\mathcal{H}\left[\mathcal{H}\frac{\delta\varphi}{\varphi_0'} + \psi\right]',$$
$$\left(a^2\frac{\psi}{H}\right)' = \frac{4\pi G a^4}{\mathcal{H}^2}(\rho + p)\left(\mathcal{H}\frac{\delta\varphi}{\varphi_0'} + \psi\right). \tag{1.80}$$

In order to grasp the meaning of the perturbations of the self-interacting scalar field, and the rôle it plays in setting the spectrum of the primordial perturbations, it is more convenient to go back to the scalar field, and to define new variables. These new variables reproduce more deeply the physical content of the scalar fluctuations. In order to do so, let us define the following new variables:

$$u = \frac{\psi}{4\pi G(\rho + p)^{\frac{1}{2}}}, \tag{1.81}$$

$$v = \sqrt{\rho_{,x}}a(\delta\varphi + \frac{\varphi_0'}{\mathcal{H}}\psi), \tag{1.82}$$

then Eqs. (1.80) read

$$c_s\nabla^2 u = z(\frac{v}{z})', \tag{1.83}$$

$$c_s v = \theta(\frac{u}{\theta})', \tag{1.84}$$

where

$$z \equiv \frac{a^2(\rho + p)^{\frac{1}{2}}}{\mathcal{H}}, \tag{1.85}$$

$$\theta = \frac{1}{c_s z}. \tag{1.86}$$

Remark that all these expressions become senseless when $p = -\rho$, corresponding to a pure de Sitter phase. As we will see later, we will be interested in a quasi-de Sitter phase.

One can write these equations as

$$v'' - \nabla^2 v - \frac{z''}{z} v = 0, \tag{1.87}$$

$$u'' - \nabla^2 u - \frac{\theta''}{\theta} v = 0. \tag{1.88}$$

The perturbed action reads, after discarding a total time derivative,

$$S = \int \mathcal{L} d\eta d^3x = \frac{1}{2} \int \left[v'^2 - c_s^2 (\vec{\nabla} v)^2 + \left(\frac{z'}{z} \right)^2 v^2 - 2 \frac{z'}{z} v v \right] d\eta d^3x. \tag{1.89}$$

This Lagrangian leads to the equation

$$v'' - c_s^2 \nabla^2 v - \frac{z''}{z} v = 0. \tag{1.90}$$

Since we will restrain ourselves to the perturbation linear regime, a Fourier decomposition can be performed:

$$v = \int v_{\vec{k}}^{(n)} e^{i \vec{k} . \vec{x}} d^3k, \tag{1.91}$$

implying,

$$v_{\vec{k}}'' + k^2 c_s^2 v_{\vec{k}} - \frac{z''}{z} v_{\vec{k}} = 0, \tag{1.92}$$

where $k \equiv |\vec{k}|$. This equation represents a parametric harmonic oscillator well-known from classical mechanics [24]. In classical mechanics it can represent oscillations with growing amplitude, for example. In quantum field in curved space-time, it represents the process of particle creation in a variable background field [25]. In our primordial context, it can represent quantum initial perturbations that are amplified generating classical perturbations during the expansion of the universe.

Even if equation (1.92) has analytical solution only for very special cases, two asymptotic regimes can be easily identified:

- If $k^2 \gg \frac{z''}{z}$ the perturbation oscillates and

$$v_{\vec{k}} \simeq \frac{1}{\sqrt{c_s}} e^{\pm i k \int c_s d\eta}. \tag{1.93}$$

- If $k^2 c_s^2 v \ll \frac{z''}{z}$, which usually implies that $k\eta << 1$, there are growing and decreasing modes:

$$v_{\vec{k}} \cong c_1(\vec{k})z + c_2(\vec{k})z \int_{\eta_i} \frac{d\eta}{z^2} d\eta + O((k\eta)^2) + ... \tag{1.94}$$

VI. QUANTIZING THE PERTURBATIONS

The existence of fundamental fields that determine the evolution of the universe during its primordial phase, has led to the concept of quantum perturbations that induces the formation of structure of the universe during its later phase. We will sketch this mechanism for the case of the inflationary universe, but many features are kept in other primordial scenarios, like the bouncing one.

Inflation is driven by a scalar field, that is supposed to be a fundamental field. At the background level, the scalar field, called generally the *inflaton*, is homogeneous. However, it is subjected to fluctuations as any quantum field. Such fluctuations induce perturbations in the metric functions, that are stretched out of the horizon during the inflationary phase. Later, after the end of the inflationary phase, this fluctuations re-enters in the horizon, during the radiation or matter dominated phase. At this moment, they induce fluctuations in the matter, leading to the structures observed today.

In order to implement this idea, let us consider the action representing the perturbations:

$$S = \frac{1}{2} \int \left[(v'^2 - c_s^2 (\vec{\nabla} v)^2 + + \left(\frac{z'}{z} \right)^2 v^2 - 2 \frac{z'}{z} v v \right] d^3x dt. \tag{1.95}$$

The corresponding equation of motion is

$$v" - \left(\nabla^2 v - \frac{z"}{z} \right) v = 0 \tag{1.96}$$

Performing the Fourier decomposition,

$$v(\eta, \vec{x}) = \int \frac{d^3k}{(2\pi)^{\frac{2}{3}}} v_{\vec{k}}(\eta) e^{i\vec{k}.\vec{x}}, \tag{1.97}$$

the equation of motion becomes:

$$v_{\vec{k}}" + \left(k^2 - \frac{z"}{z}\right) v_{\vec{k}} = 0 \tag{1.98}$$

All this framework may be put in a Hamiltonian form, more suitable to perform the quantization:

$$H_k = \int d^3k [p_{\vec{k}} p^*_{\vec{k}} + k^2 v_{\vec{k}} v^*_{\vec{k}} + \frac{z'}{z}(p_{\vec{k}} v^*_{\vec{k}} - v_{\vec{k}} p^*_{\vec{k}})] \tag{1.99}$$

One can now quantize the system. In the Heisenberg picture, the operators evolve in time according to the operator version of Eq. (1.98). We will call $\hat{y}_{\vec{k}}$ the operator associated with $v_{\vec{k}}$.

If at some initial time η_i one defines annihilation operators as

$$\hat{c}_{\vec{k}}(\eta_i) = \frac{1}{\sqrt{2}} \left(\sqrt{k} \hat{y}_{\vec{k}}(\eta_i) + \frac{i}{\sqrt{k}} \hat{p}_{\vec{k}}(\eta_i) \right), \tag{1.100}$$

and the creation operator as $\hat{c}^\dagger_{\vec{k}}(\eta_i)$, one can write the field operator as

$$\hat{y}(\eta, \vec{x}) = \sqrt{\frac{\hbar c}{(2\pi)^3}} \int d^3k [v_{\vec{k}}(\eta) \hat{c}_{\vec{k}}(\eta_i) e^{i\vec{k}.\vec{x}} + v^*_{\vec{k}}(\eta) \hat{c}^\dagger_{\vec{k}}(\eta_i) e^{-i\vec{k}.\vec{x}}], \tag{1.101}$$

where $v_{\vec{k}} \equiv v_{|\vec{k}|}$ satisfies the same mode equation (1.98) as the classical mode $v_{\vec{k}}$. Hence the mode operators at arbitrary η read

$$\hat{y}_{\vec{k}}(\eta) = \sqrt{2k} \mathrm{Re}[v_{\vec{k}}(\eta) \hat{y}_{\vec{k}}(\eta_i)] - \sqrt{\frac{2}{k}} \mathrm{Im}[v_{\vec{k}}(\eta) \hat{p}_{\vec{k}}(\eta_i)], \tag{1.102}$$

and

$$\hat{p}_{\vec{k}}(\eta) = \sqrt{\frac{2}{k}} \mathrm{Re}[g_{\vec{k}}(\eta) \hat{p}_{\vec{k}}(\eta_i)] + \sqrt{2k} \mathrm{Im}[g_{\vec{k}}(\eta) \hat{y}_{\vec{k}}(\eta_i)], \tag{1.103}$$

where,

$$g_{\vec{k}} = i \left(v'_{\vec{k}} - \frac{z'}{z} v_{\vec{k}} \right), \tag{1.104}$$

and $v_{\vec{k}}$ should satisfy Eq. (1.98).

In inflationary models, and also in bouncing models, it is argued that the initial state of the perturbations is the vacuum state:

$$\hat{c}_{\vec{k}}(\eta_i)|0,\eta_i>=0. \tag{1.105}$$

At an arbitrary time, in the Schrödinger picture, one has

$$\hat{c}_{\vec{k}}(\eta)|0,\eta>=0 \quad \rightarrow \quad [\hat{y}_{\vec{k}}(\eta_i)+i\gamma_{\vec{k}}^{-1}(\eta)\hat{p}_{\vec{k}}(\eta_i)]|0,\eta>=0, \tag{1.106}$$

where

$$\gamma_{\vec{k}} \equiv \frac{1}{2|v_{\vec{k}}|^2}\left[1-2i\mathrm{Im}(v_{\vec{k}}^* g_k)\right]. \tag{1.107}$$

In the field representation, the wave functional $\Psi(y)$ satisfies a Schrödinger functional equation. If one writes this wave functional as a product of mode wave functions for each mode $\vec{k}$, $\Psi=\pi_{\vec{k}}\Psi_{\vec{k}}$, this Schrödinger functional equation implies infinite many independent (there are no self-interactions, the Hamiltonian is quadratic in y) ordinary Schrödinger equations for each mode:

$$i\frac{\partial\Psi_{\vec{k}}}{\partial\eta}=\left[-\frac{\partial^2}{\partial y_{\vec{k}}^*\partial y_{\vec{k}}}+k^2 y_{\vec{k}}^* y_{\vec{k}}-i\frac{z'}{z}\left(\frac{\partial}{\partial y_{\vec{k}}^*}y_{\vec{k}}^*+y_{\vec{k}}\frac{\partial}{\partial y_{\vec{k}}}\right)\right]\Psi_{\vec{k}}. \tag{1.108}$$

The imposition of a vacuum initial state (see Eq. (1.106)) implies that the mode wave function must also satisfies

$$\left(y_{\vec{k}}+\gamma_{\vec{k}}^{-1}\frac{\partial}{\partial y_{\vec{k}}^*}\right)\Psi_{\vec{k}}=0. \tag{1.109}$$

The unique normalized solution of Eqs. (1.108) and (1.109) reads

$$\Psi_{\vec{k}}(y_{\vec{k}},y_{\vec{k}}^*,\eta)=\frac{\exp\left(-\frac{|y_{\vec{k}}(\eta_i)|^2}{2|v_{\vec{k}}(\eta)|^2}\right)}{\sqrt{\pi}|v_{\vec{k}}(\eta)|}\exp\left\{i\left[\left(\frac{|v_{\vec{k}}(\eta)|'}{|v_{\vec{k}}(\eta)|}-\frac{z'}{z}\right)|y_{\vec{k}}(\eta_i)|^2-\int\frac{d\eta}{2|v_{\vec{k}}(\eta)|^2}\right]\right\}. \tag{1.110}$$

When $\eta \to \eta_i$ and we are very close to the Minkowski vacuum (wavelengths much smaller than the curvature scale), we know that the mode wave function reduces to the usual Gaussian ground state given by

$$\Psi_{\vec{k}}(y_{\vec{k}}, y^*_{\vec{k}}, \eta) \propto \exp\left(-k|y_{\vec{k}}(\eta_i)|^2\right) \exp\left(-i\eta k\right), \tag{1.111}$$

where the phase corresponds to the ground state energy of the mode k, $E_k = k$.

Comparing Eqs. (1.110) and (1.111) one can see that $|v_{\vec{k}}(\eta)| \approx 1/\sqrt{2k}$. As $v_{\vec{k}}(\eta)$ must satisfy Eq. (1.98) and as in this regime $k^2 >> z''/z$, one can infer that

$$v_{\vec{k}}(\eta) = \frac{\exp[-ik(\eta - \eta_i)]}{\sqrt{2k}} \left(1 + \frac{A_1(k)}{\eta} + ...\right). \tag{1.112}$$

This is written as an expansion in $1/\eta$ because in inflationary and bouncing models, generally, $|\eta_i| >> 1$.

Let us now turn to the statistical predictions. We can consider the two-point correlation function

$$< 0|\hat{y}(\eta, \vec{x})\hat{y}(\eta, \vec{x} + \vec{r})|0 > \tag{1.113}$$

$$= \int \mathcal{D}y |\Psi(y, \eta)|^2 y(\vec{x}) y(\vec{x} + \vec{r}), \tag{1.114}$$

which can be calculated to yield

$$< 0|\hat{y}(\eta, \vec{x})\hat{y}(\eta, \vec{x} + \vec{r})|0 >= \frac{1}{2\pi^2} \int \frac{dk}{k} \frac{\sin kr}{kr} k^3 |v_{\vec{k}}(\eta)|^2, \tag{1.115}$$

in the case of the ground state, which motivates the power spectrum definition of v, $P_v(k) = k^3 v_{\vec{k}}$. This ensemble average should approximately be equal to the spatial average of an actual field configuration for the universe, and then compared with the measured temperature anisotropies of the cosmic microwave background radiation. However, what is directly related to $\Delta T/T$ is the two-point correlation function of the Bardeen operator $\hat{\phi}$, which, when $\phi = \psi$ (the case with no anisotropic pressures), is directly related to the $\hat{y}$ operator through the operator version of equations (1.81,1.83). Hence, from

$$< 0|\hat{\phi}(\eta,\vec{x})\phi(\eta,\vec{y})|0 >= \int |\phi_k|^2 k^3 \frac{sen(kr)}{kr}\frac{dk}{k}, \tag{1.116}$$

with $r = |\vec{x}-\vec{y}|$, we see that the main quantity to be evaluated is the power spectrum

$$P(k) = |\phi_k|^2 k^3, \tag{1.117}$$

where,

$$-\frac{c_s k^2 \phi_k}{4\pi G(\rho+p)^{\frac{1}{2}}} = \left(\frac{v_{\vec{k}}}{z}\right)'. \tag{1.118}$$

Then one defines the spectral index

$$n_s \equiv 1 + \frac{d\ln P(k)}{d\ln k}, \tag{1.119}$$

which observations indicate to be a little less than its scale invariant value ($n_s = 1$), $n_s \approx 0.97$.

We will now see what inflation and bouncing models say about the spectral index.

VII. APPLICATIONS

Let us consider now two examples. The first is the computation of the primordial power spectrum in the inflationary scenario. The second one is the primordial spectrum in a bouncing scenario.

1. The power law inflation example.

   Power law inflation is obtained from scalar field models where the potential of the scalar field is an exponential function of the scalar field, yielding the equation of state $p = \omega\rho$, with $\omega =$ constant. In this case, the scale factor is given by,

   $$a = l_0|\eta|^{1+\beta}, \tag{1.120}$$

   where l_0 is a constant. The condition to have inflation is,

   $$-2 > \beta > -\infty \quad , \quad \beta = \frac{1-3\omega}{1+3\omega}. \tag{1.121}$$

The case $\beta = -2$ represents the de Sitter expansion. In terms of ω, the condition to have inflation is given by $-1 < \omega < -\frac{1}{3}$. In this case,

$$z \propto \frac{a\sqrt{\mathcal{H}^2 - \mathcal{H}}}{\mathcal{H}} \propto a, \tag{1.122}$$

where $\mathcal{H} = \frac{a'}{a}$.

There are two regimes:

(a) When

$$k^2 >> \frac{a"}{a}, \tag{1.123}$$

we have,

$$v_{\vec{k}} = \frac{1}{\sqrt{2k}} e^{-ik\eta} \left(1 + \frac{A_1}{\eta} + \dots\right), \tag{1.124}$$

corresponding to the vacuum state in a flat space-time, leading to,

$$\Phi_k^I = -\frac{e^{-ik\eta}}{k^{\frac{3}{2}}\eta^{2+\beta}} \left[i + \frac{(1+\beta)}{k\eta} + \dots\right]. \tag{1.125}$$

(b) When

$$k^2 \ll \frac{a"}{a}, \tag{1.126}$$

we find,

$$v_{\vec{k}} = c_1 a + c_2 a \int \frac{d\eta}{a^2} = c_1 |\eta|^{1+\beta} + c_2 |\eta|^{-\beta}, \tag{1.127}$$

leading to

$$\Phi_k^{II} = \frac{D_1}{|\eta|^{3+2\beta}} + D_2 \tag{1.128}$$

Matching the asymptotic perturbed solutions at $k^2 \propto \frac{a"}{a}$, which in the power law case means $k \propto 1/\eta$ we find:

$$\Phi^I(\eta = \frac{1}{k}) = \Phi^{II}(\eta = \frac{1}{\eta}), \tag{1.129}$$

$$\Phi'^I = (\eta = \frac{1}{\eta}) = \Phi'^{II}(\eta = \frac{1}{\eta}), \tag{1.130}$$

$$\rightarrow D_2 \propto k^{\frac{(1+2\beta)}{2}}. \tag{1.131}$$

The final Bardeen potential, for $\eta >> 1$ reduces to the constant mode D_2, and the power spectrum is given by:

$$P(k) = k^3 |\Phi_k|^2 \propto k^{4+2\beta}, \tag{1.132}$$

yielding $n_s = 5 + 2\beta$. For inflation, $-2 > \beta > -\infty$. An invariant power spectrum is obtained when $\beta = -2$, that is, when inflation corresponds to a de Sitter primordial phase. Observations yielding $n_s \approx 0.97$ indicates that inflation should have happened close to de Sitter expansion, but not so fast.

2. Bouncing Models.

The bouncing universe can be achieved in many situations. One of them, is when quantum effects are taking into account in the primordial universe. For example, in the context of quantum cosmology where the matter fields play the rôle of the time variable, a general expression for the scale facto can be written as [26],

$$a = a_0 \Big(1 + T^2)^{\frac{1}{3(1-\omega)}}, \tag{1.133}$$

where T is a time variable which depends on the equation of state parameter ω. For a radiation fluid, for example, it is just the conformal time. In the asymptotic regions $T \to \pm\infty$, the classical behaviour characterized by a fluid with an equation $p = \omega\rho$ is obtained, while near the origin quantum effects lead to a bounce, avoiding the singularity.

Then we we suppose again in a initial, $p = \omega\rho$. This can represent the asymptotic expression for the equation of state, for example, in the beginning of the bounce. We have then the equation,

$$v_{\vec{k}}'' + \left(k^2 - \frac{a"}{a} \right) v_{\vec{k}} = 0, \tag{1.134}$$

with $a(\eta) \propto |\eta|^{\frac{2}{1+3\omega}}$, $-\infty < \eta < 0$. The perturbed equation becomes (see Refs. [5–7]),

$$v_{\vec{k}}" + \left[\omega k^2 + \frac{2|3\omega - 1|}{(1+3\omega)^2 \eta^2} \right] v_{\vec{k}} = 0. \tag{1.135}$$

Hence,

$$\phi_k \propto k^{\frac{3(\omega-1)}{2(1+3\omega)}} \left[\text{const.} + \frac{1}{\eta^{\frac{5+3\omega}{1+3\omega}}} \right]. \tag{1.136}$$

The power spectrum is now,

$$P(k) = k^3 |\Phi_k|^2 = k^{\frac{12\omega}{1+3\omega}}. \tag{1.137}$$

An almost invariant power spectrum is obtained when $\omega \approx 0$, that is, when a pressureless fluid dominates during the contracting phase.

VIII. CONCLUSIONS

The goal of the present text was to revise the theory of cosmological perturbations and its application to the determination of the primordial spectrum of fluctuations that leads to large scales structures observed in the universe. The gauge invariant formalism was developed in details. Since it is generally believed today that the primordial fluctuations come from quantum vacuum fluctuations, an special emphasis has been given to the mechanism connected with this phenomena. The Mukhanov-Sasaki variables, that accounts for the quantum fluctuations were set out, and the quantization procedure was described in its fundamental lines. Application to inflationary and bouncing models shows that, in order for such models be compatible to observations, the first one should present an initial accelerated expansion close to de Sitter expansion, while the second should have its contracting phase dominated by dust.

This quantum mechanism gives specific predictions for the primordial spectrum of perturbations, that evolve later to originate the observed structures in the universe. Hence, clear predictions are obtained, that can be compared with observations. This very elegant mechanism create a window to test theories for the primordial universe, like the inflationary scenario and the bouncing models, as it has been discussed.

---

[1] P.J.E. Peebles, **Principles of physical cosmology**, Princeton University Press, Princeton(1993).

[2] E.W. Kolb and M.S. Turner, **The early universe**, Addison-Wesley, New York(1990).

[3] R.S. Perez and N. Pinto-Neto, Grav.&Cosm. **17**, 136(2011).

[4] M. Novello and S.E.Perez Bergliaffa, Phys. Rep. **463** 127 (2008).

[5] P. Peter, E. J. C. Pinho and Nelson Pinto-Neto, Phys. Rev. **D75**, 023516 (2007).

[6] P. Peter and Nelson Pinto-Neto, Phys. Rev. **D78**, 063506 (2008).

[7] F. T. Falciano and N. Pinto-Neto, Phys. Rev. **D79**, 023507 (2009).

[8] S. D. P. Vitenti, F. T. Falciano and N. Pinto-Neto, Phys. Rev. **D87**, 103503 (2013).

[9] E. Lifshitz, Zh. Eksp. Teor. Fiz. **16**, 587(1946).

[10] E. Lifshitz and I. Khalatnikov, Adv. Phys. **12**, 185(1963).

[11] J. Bardeen, Phys. Rev. **D22**, 1882(1980).

[12] V.F. Mukhanov, H.A. Feldman and R.H. Brandenberger, Phys. Rep. **215**, 203(1992).

[13] V.F. Mukhanov, **Physical Foundations of Cosmology**, Cambridge University Press, Cambridge(2005).

[14] V.N. Lukash, Sov. Phys. JETP**52**, 807(1980).

[15] V.N. Lukash, JETP Lett. **31**, 596(1980).

[16] V.F. Mukhanov, Sov. Phys. JETP **67**, 1297(1988).

[17] L.P. Grishchuk, Phys. Rev. **D50**, 7154(1994).

[18] L. Amendola and S. Tsujikawa, **Dark energy**, Cambridge University Press, Cambridge(2010).

[19] A.R. Liddle and D.H. Lyth, **Cosmological inflation and large-scale structure**, Cambridge University Press, Cambridge(2000).

[20] N. Pinto-Neto and J.C. Fabris, Class. Quant. Grav. **30**, 143001(2013).

[21] M. Bojowald, Class. Quant. Grav. **29**, 213001(2012).

[22] S.W. Hawking and G.F.R. Ellis, **The large scale structure of space-time**, Cambridge University Press, Cambridge(1973).

[23] O.F. Piattella, J.C. Fabris and N. Bilić, *Note on the thermodynamics and the speed of sound of a scalar field*, arXiv:1309.4282.

[24] L. Landau and E. Lifchitz, **Mécanique**, Mir, Moscow(1966).

[25] V.F. Mukhanov and S. Winitzki, **Introduction to quantum effects in gravity**, Cambridge University Press, Cambridge(2007).

[26] F.G. Alvarenga, J.C. Fabris, N.A. Lemos and G.A. Monerat, Gen. Rel. Grav. **34**, 651(2002).

Linear cosmological perturbation theory

Christos G. Tsagas[1]

[1]*Section of Astrophysics, Astronomy and Mechanics, Department of Physics, Aristotle University of Thessaloniki, Thessaloniki 54124, Greece.*

Cosmological perturbation theory is believed to provide the theoretical framework for understanding and explaining the large-scale structure of our universe. The physical principle behind the mechanism that started the whole process is known as Jeans instability. The detailed physics, however, which led to the formation of the galaxies, the galaxy clusters and the voids we observe today are still the subject of debate. Nevertheless, it is widely accepted that all begun when small irregularities in an otherwise uniform cosmic fluid started to grow under their own gravity. These irregularities come in three different types, namely scalar, vector and tensor, but as long as they are small they grow linearly and independent of each other. The aim of these lecture notes is to discuss key aspects of the aforementioned linear regime of structure formation and present the main theoretical issues and results.

I. INTRODUCTION

The high isotropy of the temperature distribution in the spectrum of the Cosmic Microwave Background (CMB), combined with the Cosmological Principle, suggest that our universe was extremely uniform at recombination. Today, on the other hand, we see structure everywhere. The universe is filled with galaxies, clusters and superclusters of galaxies and extensive low-density regions known as "voids". The obvious question is how the high uniformity at the time of last scattering, when the universe was only few hundred thousand years old, gave way to the complicated structure we observe today, about ten billion years later. Cosmological perturbation theory attempts to answer this question.

The physical mechanism that set the whole process of structure formation into motion is believed to be "gravitational instability". The principle behind the

aforementioned mechanism, which is also known as "Jeans' instability", is fairly simple. Small irregularities in the largely uniform distribution of the cosmic medium start growing under their own gravity. As long as these irregularities/perturbations remain small, they grow slowly and independent of each other. Afterwards, that is once the perturbations have become sufficiently large, this is no longer the case. The long initial period of slow growth, which is commonly referred to as the "linear regime" of structure formation, is the focal point of the present lecture notes.

We distinguish between three different types of perturbations, namely scalar, vector and tensor. A typical example of scalar-type distortions are overdensities, or underdensities, in the matter distribution of the universe. These usually appear in the literature as density perturbations. Vector distortions, on the other hand, are rotational irregularities. Ordinary kinematic vorticity, or density vortices, are examples of vector perturbations. Finally, (pure) tensor distortions are associated with gravitational waves. During the linear regime, the above mentioned three types of perturbations are separable and evolve independently.

In order to study the evolution of cosmological perturbations, we first need to select the so-called "background" model. The latter defines our (fictitious) unperturbed universe and it is usually described by an exact solution of the Einstein field equations. Perturbing away from the aforementioned background should then provide us with a more realistic model for the universe we live in. In the majority of the studies, the unperturbed universe is the homogeneous and isotropic Friedmann-Robertson-Walker (FRW) spacetime and, in particular, the spatially flat member of the family. This is the simplest cosmological solution of Einstein's equations and also appears to be in very good agreement with the observations (primarily with the CMB), as well as with our theoretical prejudice (namely with the Cosmological Principle).

According to the data of the COBE mission, the temperature anisotropy of the CMB is of the order of 10^{-5}. Taken at face value, this implies that at recombination our universe was an almost-FRW model with an accuracy of one part in hundred thousand. It is relatively straightforward to show that the Friedmann universes are unstable to linear density perturbations. The other two types do not seem to grow, but their presence could have affected the subsequent nonlinear regime of structure formation.

The overall instability of the FRW models means that small irregularities, like those reported by COBE, will grow gravitationally and (given enough time) they could form the large-scale structure that we observe today. However, the simplest scenarios had serious problems in reproducing the observed picture, because the linear growth was too slow. The solution came in the form of non-baryonic dark matter, the introduction of which accelerated the early growth of density perturbations. For a period, structure formation models employed two types of dark matter, namely Hot Dark Matter (HDM) and Cold Dark Matter (CDM). The former was relativistic and helped to reproduce the very large-scale structure of the universe. The latter was of low-energy particles and worked well on relatively small scales. This picture changed around the turn of the millennium, when the supernovae observations suggested a late accelerated expansion for our universe. The result, was the "concordance" cosmological paradigm, which contains approximately 70% Dark Energy (in the form of a cosmological constant or of "dynamical" dark energy), 25% CDM and only 5% ordinary baryonic matter.[1] This is the ΛCDM scenario that seems in good agreement with the observations and has been the main cosmological model for the last fifteen years or so. There is still scepticism in the community, however, based on theoretical and observational arguments. The main theoretical issue appears to be the nature of Dark Energy, which remains an unknown and essentially free parameter that has been fine-tuned to agree with the observations. Observationally as well there are problems, mainly in the low multiples of the CMB. Future data, as well as theoretical work, will hopefully settle these matters, either by firmly establishing the ΛCDM model, or by pointing to new directions.

The outline of these notes is as follows. In section 2 we set up the mathematical framework that will be used to study relativistic cosmological models. The formalism, which is known as the covariant approach to relativity and cosmology, introduces an 1+3 splitting of the spacetime into time and space. Time is the fundamental direction, determined by the observers' worldlines, while space is defined subsequently by projecting orthogonal to the fundamental 4-velocity field. One can then proceed to decompose every variable, every operator and every equation into their irreducible

[1] The latest results from the Planck satellite mission have reported no evidence of dynamical Dark Energy, have slightly increased the percentage of CMD (at the expense of Dark Energy) and lowered the value of the Hubble parameter to $H_0 \simeq 67$ km/secMpc [1].

timelike and spacelike parts. These are used to determine the kinematical and dynamical evolution of the cosmological model in hand. The equations presented in § II are nonlinear and apply to arbitrary spacetimes filled with a general imperfect fluid. Section 3 discusses spatially homogeneous cosmological models, putting particular emphasis to the FRW spacetimes, which have been the cornerstone of modern cosmology for several decades. There is also a brief reference to the Bianchi family of models. These are spatially homogeneous, but anisotropic, spacetimes and some of them contain the FRW solutions as special cases. This fact, as well as the generally richer behaviour of the Bianchi models, explains why these spacetimes have been occasionally employed as viable alternatives to the Friedmann universes. In section 4, we provide the formalism that will be used to study osmological perturbations within the covariant framework. We begin by discussing the so-called "gauge problem" that has long been known to plague cosmological perturbation theory. Here, we bypass all the gauge-related issues by defining and using variables that describe perturbations in a gauge-invariant manner. We then proceed to derive the full set of the nonlinear expressions that determine the dynamical evolution of a universe filled with a general imperfect fluid. The cases of a perfect medium and of a minimally coupled scalar field are also considered. The full equations are linearised around an FRW background in § V. We outline the linearisation rule and then solve the linear expressions assuming a flat Friedmann universe as our background. Among others, our solutions show how density perturbations grow on large scales, while they oscillate (due to pressure support) on small enough lengths. In the process, we also introduce concepts like the Hubble horizon, the Jeans length and the Jeans mass and discuss their dynamical significance. Section 5 also contains a brief discussion on dissipative processes, such as the Silk and the Landau damping, which are known to smooth out perturbations on relatively large scales. We also discuss scalar field perturbations, as well as vector (i.e. rotational) and tensor (i.e. gravitational wave) distortions and provide analytical solutions showing the deacy all these perturbations in an expanding universe. Finally, we employ two-fluid models to demonstrate how the radiation pressure can lead to the "stagnation" of baryonic perturbations before recombination and how the CDM species can accelerate the gravitational contraction of baryonic matter after last scattering.

The last section, § VI, starts with a brief discussion of the supernovae observations and outlines how the data lead the majority of the cosmological community to embrace the idea of an accelerated universe driven by Dark Energy. We explain, in particular, why accelerated expansion in FRW models requires either a positive cosmological constant, or dynamical Dark Energy with negative gravitational mass. In perturbed Friedmann models, however, this is not necessarily the case. There, it is conceivable that local inhomogeneities and anisotropies can "backreact" on the background expansion and may thus lead to global acceleration. Such "backreaction" scenarios have been proposed in an attempt to explain the recent universal acceleration by means of conventional physics. We therefore close § VI, with a brief outline of a simple paradigm, where the overall effect of inhomogeneity and anisotropy is encoded within a single "backreaction term" that has been averaged over a large spatial domain of the observable universe.

II. COVARIANT RELATIVISTIC COSMOLOGY

The covariant formalism utilises the kinematic quantities, the matter energy-momentum tensor and the gravito-electromagnetic parts of the Weyl tensor, instead of the metric. The key equations are the Ricci and Bianchi identities, while Einstein's equations are incorporated via algebraic relations between the Ricci and the energy-momentum tensor.

A. Local spacetime splitting

Consider a general spacetime with metric g_{ab} of signature $(-, +, +, +)$ and introduce a family of fundamental observers with 4-velocity [2][2]

$$u^a = \frac{\mathrm{d}x^a}{\mathrm{d}\tau}\,, \tag{2.1}$$

where τ is the proper time and $u_a u^a = -1$. The u_a-field defines the observer's time direction and introduces a local 1+3 'threading' of the spacetime into time and space.

[2] Latin indices vary between 0 and 3 and refer to arbitrary coordinate or tetrad frames. Greek indices run from 1 to 3. We use geometrised units with $c = 1 = 8\pi G$, which means that all geometrical variables have physical dimensions that are integer powers of length.

Using the above 4-velocity field, we also define the symmetric tensor $h_{ab} = g_{ab} + u_a u_b$, which projects orthogonal to u_a and into the observers' instantaneous rest space at each event.

The u_a-field and its tensor counterpart h_{ab} allow for a unique decomposition of every spacetime quantity into its irreducible timelike and spacelike parts. They are are also used to define the covariant time and spatial derivatives of any tensor field $S_{ab\cdots}{}^{cd\cdots}$ according to

$$\dot{S}_{ab\cdots}{}^{cd\cdots} = u^e \nabla_e S_{ab\cdots}{}^{cd\cdots} \quad \text{and} \quad \mathrm{D}_e S_{ab\cdots}{}^{cd\cdots} = h_e{}^s h_a{}^f h_b{}^p h_q{}^c h_r{}^d \cdots \nabla_s S_{fp\cdots}{}^{qr\cdots} , \tag{2.2}$$

respectively.

B. Gravitational field

In the geometrical interpretation of General Relativity, gravity is spacetime curvature and the gravitational field is monitored by the Riemann curvature tensor R_{abcd}. Also, the interaction between mater and spacetime geometry follows form Einstein's equations,

$$R_{ab} - \frac{1}{2} R g_{ab} = T_{ab} - \Lambda g_{ab} , \tag{2.3}$$

where $R_{ab} = R_{acb}{}^c$ is the Ricci tensor (with trace R), T_{ab} is the total energy-momentum tensor of the matter and Λ is the cosmological constant. Then, the twice contracted Bianchi identities guarantee that $\nabla^b T_{ab} = 0$ and total energy-momentum conservation (see § II E below).

The Ricci tensor describes the local gravity. Information about the non-local (long-range) gravitational field, which includes gravitational waves and tidal forces, is encoded in the Weyl conformal curvature tensor C_{abcd}. This splitting of the total gravitational field into a local and a non-local part is reflected in the following decomposition of the Riemann tensor,

$$R_{abcd} = C_{abcd} + \frac{1}{2}\left(g_{ac}R_{bd} + g_{bd}R_{ac} - g_{bc}R_{ad} - g_{ad}R_{bc}\right) - \frac{1}{6} R \left(g_{ac}g_{bd} - g_{ad}g_{bc}\right) . \tag{2.4}$$

Note that $R_{abcd} = R_{cdab}$, $R_{abcd} = R_{[ab][cd]}$ and $R^{a}{}_{[bcd]} = 0$ and same symmetries are shared by the Weyl tensor, which is also trace-free (i.e. $C^{c}{}_{acb} = 0$). The conformal curvature tensor splits further into a electric and a magnetic component, defined by [3, 4]

$$E_{ab} = C_{acbd}u^{c}u^{d} \qquad \text{and} \qquad H_{ab} = \frac{1}{2}\,\varepsilon_{a}{}^{cd}C_{cdbe}u^{e}\,. \tag{2.5}$$

Then, we may write

$$C_{abcd} = (g_{abqp}g_{cdsr} - \eta_{abqp}\eta_{cdsr})\,u^{q}u^{s}E^{pr} - (\eta_{abqp}g_{cdsr} + g_{abqp}\eta_{cdsr})\,u^{q}u^{s}H^{pr}\,, \tag{2.6}$$

where η_{abcd} is the Levi-Civita tensor of the spacetime and $g_{abcd} = g_{ac}g_{bd} - g_{ad}g_{bc}$. The electric Weyl field generalises the tidal tensor of the Newtonian gravitational potential, but H_{ab} has no Newtonian counterpart. For the propagation of gravitational waves, both of the Weyl components must be present.

C. Matter fields

Relative to the fundamental 4-velocity field, the energy-momentum tensor of a general (imperfect) fluid decomposes as[3]

$$T_{ab} = \rho u_{a}u_{b} + ph_{ab} + 2q_{(a}u_{b)} + \pi_{ab}\,. \tag{2.7}$$

where $\rho = T_{ab}u^{a}u^{b}$ represents the matter energy density, $p = T_{ab}h^{ab}/3$ the effective isotropic pressure, $q_{a} = -h_{a}{}^{b}T_{bc}u^{c}$ the total energy-flux vector, and $\pi_{ab} = h_{\langle a}{}^{c}h_{b\rangle}{}^{d}T_{cd}$ the anisotropic pressure.[4] For a perfect the fluid, we have $q_{a} = 0 = \pi_{ab}$ and expression (2.7) reduces to

$$T_{ab} = \rho u_{a}u_{b} + ph_{ab}\,. \tag{2.9}$$

For a barotropic medium we have $p = p\,(\rho)$, which reduces to $p = 0$ when dealing with pressure-free 'dust' (baryonic or not). In general, however, the equation of state

[3] In a multi-component medium, or when allowing for peculiar velocities, one needs different 4-velocities for the matter components and the observers (e.g. see § 2.4 in [2]).

[4] Angled brackets denote the symmetric and trace-free part of spatially projected second-rank tensors and the projected part of vectors according to

$$S_{\langle ab\rangle} = h_{\langle a}{}^{c}h_{b\rangle}{}^{d}S_{cd} = h_{(a}{}^{c}h_{b)}{}^{d}S_{cd} - \frac{1}{3}\,h^{cd}S_{cd}h_{ab} \qquad \text{and} \qquad V_{\langle a\rangle} = h_{a}{}^{b}V_{b}\,, \tag{2.8}$$

respectively (with $S_{\langle ab\rangle}h^{ab} = 0$).

takes the form $p = p(\rho, s)$, where s is the specific entropy. Finally, we note that expression (2.7) describes any type of matter, including electromagnetic fields, scalar fields, etc.

Using the above given expressions of the energy-momentum tensor, we can recast the Einstein field equations into a set of very useful algebraic relations. To begin with, recall that $R = 4\Lambda - T$, with $T = T_a{}^a$ – see (2.3). Then, Einstein's equations read

$$R_{ab} = T_{ab} - \frac{1}{2}\, T g_{ab} + \Lambda g_{ab}\,. \tag{2.10}$$

Assuming that T_{ab} is that of an imperfect fluid (see Eq. (2.7)), the successive contraction of the above, leads to

$$R_{ab}u^a u^b = \frac{1}{2}\,(\rho + 3p) - \Lambda\,, \qquad\qquad h_a{}^b R_{bc} u^c = -q_a \tag{2.11}$$

and

$$h_a{}^c h_b{}^d R_{cd} = \frac{1}{2}\,(\rho - p) h_{ab} + \Lambda h_{ab} + \pi_{ab}\,. \tag{2.12}$$

D. Kinematics

In principle, there are various alternative options for the fundamental 4-velocity field. In cosmology, however, the u_a-frame is typically defined as the coordinate system in which the CMB dipole vanishes. Once the u_a-field has been specified, its integral curves define the worldlines of the fundamental observers introduced in § II A.

The motion of the fundamental frame is characterised by the irreducible kinematical quantities of the u_a-congruence. These emerge after decomposing of the 4-velocity gradient as

$$\nabla_b u_a = \sigma_{ab} + \omega_{ab} + \frac{1}{3}\,\Theta h_{ab} - A_a u_b\,. \tag{2.13}$$

Here, $\sigma_{ab} = \mathrm{D}_{\langle b} u_{a\rangle}$ monitors shear anisotropies, $\omega_{ab} = \mathrm{D}_{[b} u_{a]}$ describes rotation and $\Theta = \nabla^a u_a = \mathrm{D}^a u_a$ monitors the volume expansion/contrcation of the u_a-field. Also, $A_a = \dot{u}_a = u^b \nabla_b u_a$ is the 4-acceleration. This vector reflects the presence of non-gravitational forces and vanishes when matter moves under gravity

alone. Note that $\sigma_{ab}u^a = 0 = \omega_{ab}u^a = A_a u^a$ by construction. The volume scalar determines the average separation between two neighbouring observers and is used to introduce a representative length scale (a – with $\dot{a}/a = \Theta/3$). Also, employing the 3-dimensional Levi-Civita tensor (ε_{abc} – where $\varepsilon = \eta_{abcd}u^d$), one defines the vorticity vector $\omega_a = \varepsilon_{abc}\omega^{bc}/2$ (so that $\omega_{ab} = \varepsilon_{abc}\omega^c$).[5]

The covariant kinematics are encoded in two sets of three propagation and three constraint equations. These are purely geometrical in origin and emerge after applying the Ricci identities $2\nabla_{[a}\nabla_{b]}u_c = R_{abcd}u^d$ to the fundamental 4-velocity vector. The former of the aforementioned sets contains Raychaudhuri's formula

$$\dot{\Theta} = -\frac{1}{3}\,\Theta^2 - \frac{1}{2}\,(\rho + 3p) - 2(\sigma^2 - \omega^2) + \mathrm{D}^a A_a + A_a A^a + \Lambda\,, \tag{2.14}$$

the shear propagation equation

$$\dot{\sigma}_{\langle ab\rangle} = -\frac{2}{3}\,\Theta\sigma_{ab} - \sigma_{c\langle a}\sigma^c{}_{b\rangle} - \omega_{\langle a}\omega_{b\rangle} + \mathrm{D}_{\langle a}A_{b\rangle} + A_{\langle a}A_{b\rangle} - E_{ab} + \frac{1}{2}\,\pi_{ab} \tag{2.15}$$

and the vorticity evolution formula

$$\dot{\omega}_{\langle a\rangle} = -\frac{2}{3}\,\Theta\omega_a - \frac{1}{2}\,\mathrm{curl}\,A_a + \sigma_{ab}\omega^b\,. \tag{2.16}$$

In the above, $\sigma^2 = \sigma_{ab}\sigma^{ab}/2$, $\omega^2 = \omega_{ab}\omega^{ab}/2 = \omega_a\omega^a$ and $\mathrm{curl}v_a = \varepsilon_{abc}\mathrm{D}^b v^c$ for any orthogonally projected vector v_a. The three constraints, on the other hand, are

$$\mathrm{D}^b\sigma_{ab} = \frac{2}{3}\,\mathrm{D}_a\Theta + \mathrm{curl}\,\omega_a + 2\varepsilon_{abc}A^b\omega^c - q_a\,, \qquad \mathrm{D}^a\omega_a = A_a\omega^a\,, \tag{2.17}$$

and

$$H_{ab} = \mathrm{curl}\,\sigma_{ab} + \mathrm{D}_{\langle a}\omega_{b\rangle} + 2A_{\langle a}\omega_{b\rangle}\,, \tag{2.18}$$

with $\mathrm{curl}v_{ab} = \varepsilon_{cd\langle a}\mathrm{D}^c v_{b\rangle}{}^d$ for any spacelike, symmetric and second-rank tensor v_{ab}.

Raychaudhuri's formula is the key equation of gravitational collapse, as it describes the evolution of the average separation between two neighbouring observers. For this reason, expression (2.14) has been at the core of all the singularity theorems and also plays a fundamental role in cosmology. Note that negative terms in the right-hand

[5] The 3-dimensional Levi-Civita tensor satisfies the relations $\mathrm{D}_d\varepsilon_{abc} = 0$ and $\dot{\varepsilon}_{abc} = 3u_{[a}\varepsilon_{bc]d}A^d$. In addition, recalling that $\eta_{abcd}\eta^{efpq} = -4!\delta_{[a}{}^e\delta_b{}^f\delta_c{}^p\delta_{d]}{}^q$, one can show that $\varepsilon_{abc}\varepsilon^{def} = 3!h_{[a}{}^d h_b{}^e h_{c]}{}^f$.

side of (2.14) lead to contraction and positive resist the collapse. This means that conventional (non-phantom) matter with positive effective gravitational mass/energy (i.e. with $\rho + 3p > 0$) is always attractive. A positive cosmological constant, on the other hand, does the opposite.

E. Conservation laws

The twice contracted Bianchi identities guarantee that $\nabla^b T_{ab} = 0$ and therefore total energy-momentum conservation. The associated formulae are the continuity and the Navier-Stokes equations, which take the covariant form

$$\dot{\rho} = -\Theta(\rho + p) - \mathrm{D}^a q_a - 2A^a q_a - \sigma^{ab}\pi_{ab} \tag{2.19}$$

and

$$[(\rho + p)h_{ab} + \pi_{ab}]\, A^b = -\mathrm{D}_a p - \dot{q}_{\langle a\rangle} - \frac{4}{3}\,\Theta q_a - (\sigma_{ab} + \omega_{ab})q^b - \mathrm{D}^b\pi_{ab}\,, \tag{2.20}$$

respectively. For perfect media, namely when $q_a = 0 = \pi_{ab}$, expressions (2.19) and (2.20) respectively reduce to

$$\dot{\rho} = -\Theta(\rho + p) \qquad \text{and} \qquad (\rho + p)A_a = -\mathrm{D}_a p\,, \tag{2.21}$$

respectively. It follows, from (2.20) and (2.21b), that the fluid pressure also contributes to the effective inertial mass/energy of the system. When dealing with a perfect medium, in particular, the sum $\rho + p$ provides a measure of the effective total inertial mass.

F. Scalar fields

Scalar-fields have come into prominence because of inflation, since they can be associated to fluids with unconventional equations of state (see [5–9] for covariant treatments).

1. Minimally coupled scalar fields

The stress-energy tensor of a scalar field, which is minimally coupled to gravity, in a general spacetime has the form

$$T^{(\varphi)}_{ab} = \nabla_a\varphi\nabla_b\varphi - \left[\frac{1}{2}\,\nabla_c\varphi\nabla^c\varphi + V(\varphi)\right] g_{ab}\,, \tag{2.22}$$

where $V(\varphi)$ is the scalar field potential. Applying the conservation law $\nabla^b T_{ab} = 0$, to the above and assuming that $\nabla_a\varphi \neq 0$, leads to the Klein-Gordon equation

$$\nabla^a\nabla_a\varphi - V'(\varphi) = 0\,, \tag{2.23}$$

with the primes indicating differentiation with respect to φ.[6]

2. Scalar-field kinematics

A fluid-description of scalar fields can be obtained by assigning a 4-velocity to the φ-field. Suppose that $\nabla_a\varphi$ is timelike (i.e. that $\nabla_a\varphi\nabla^a\varphi < 0$). In this case, $\nabla_a\varphi$ is normal to the spacelike hypersurfaces $\varphi(x^a) =$ constant and we may define our 4-velocity field as

$$u_a = -\frac{1}{\dot\varphi}\,\nabla_a\varphi\,, \tag{2.24}$$

with $\dot\varphi = u^a\nabla_a\varphi \neq 0$ [6, 10]. This ensures that $\dot\varphi^2 = -\nabla_a\varphi\nabla^a\varphi > 0$ and $u_a u^a = -1$. The metric of the 3-space orthogonal to u_a is given by the projector

$$h_{ab} = g_{ab} + \frac{1}{\dot\varphi^2}\,\nabla_a\varphi\nabla_b\varphi\,, \tag{2.25}$$

which also defines the covariant derivative operator $\mathrm{D}_a = h_a{}^b\nabla_b$ and guarantees that

$$\mathrm{D}_a\varphi = 0\,, \tag{2.26}$$

always. The latter result is key to our adopted (1+3 covariant) fluid description of the φ-field and essentially dictates its evolution.

---

[6] When $\nabla_a\varphi = 0$ expression (2.22) reduces to $T^{(\varphi)}_{ab} = -V(\varphi)g_{ab}$, ensuring that $\nabla_a V(\varphi) = 0$ (since $\nabla^b T_{ab} = 0$). Then, φ behaves as a cosmological constant rather than a dynamical scalar field.

The kinematics of $u_a = -\nabla_a\varphi/\dot\varphi$, in covariant terms, follow from decomposition (2.13). Starting from definition (2.24), involving the Klein-Gordon equation and taking into account that $\mathrm{D}_a\varphi = 0$, we arrive at [11]

$$\Theta = -\frac{1}{\dot\varphi}\left[\ddot\varphi + V'(\varphi)\right]\,, \qquad \omega_{ab} = 0\,, \qquad \sigma_{ab} = -\frac{1}{\dot\varphi}h_a{}^c h_b{}^d\nabla_c\nabla_d\varphi + \frac{1}{3\dot\varphi}\left[\ddot\varphi + V'(\varphi)\right]h_{ab} \tag{2.27}$$

and

$$A_a = -\frac{1}{\dot\varphi}\,\mathrm{D}_a\dot\varphi\,. \tag{2.28}$$

These are the irreducible kinematical variables of a minimally coupled scalar field, ensuring that the 4-velocity (2.24) is irrotational and that $\dot\varphi$ acts as an acceleration potential.

3. Scalar fields as perfect fluids

Introducing the timelike velocity field (2.24) also facilitates a fluid-description of scalar fields, by recasting the energy-momentum tensor (2.22) in the perfect-fluid form[7]

$$T^{(\varphi)}_{ab} = \rho^{(\varphi)}u_a u_b + p^{(\varphi)}h_{ab}\,, \tag{2.29}$$

where

$$\rho^{(\varphi)} = \frac{1}{2}\,\dot\varphi^2 + V(\varphi) \qquad \text{and} \qquad p^{(\varphi)} = \frac{1}{2}\,\dot\varphi^2 - V(\varphi)\,. \tag{2.30}$$

Hence, demanding a positive definite energy density for the φ-field, implies $\dot\varphi^2 + 2V(\varphi) > 0$.

Two characteristic cases follow from the above fluid-description of a minimally coupled scalar field. The first is that of a free scalar field with purely kinetic energy (i.e. $V(\varphi) = 0$) and $p^{(\varphi)} = \rho^{(\varphi)}$, analogous to the equation of state of stiff-matter. In the second case, the field's energy is purely potential with $p^{(\varphi)} = -\rho^{(\varphi)}$. This corresponds to the slow-rolling regime of standard inflation. In general, however, we have either

$$p^{(\varphi)} = \rho^{(\varphi)} - 2V(\varphi)\,, \qquad \text{or} \qquad p^{(\varphi)} = \dot\varphi^2 - \rho^{(\varphi)}\,. \tag{2.31}$$

[7] A non-minimally coupled scalar field corresponds to an imperfect medium [6].

In other words, minimally coupled scalar fields do not always behave like barotropic media [5].

4. Conservation laws

After the timelike 4-velocity vector (2.24) has been introduced, we may use condition $\mathrm{D}_a\varphi = 0$ to write $\nabla_a\varphi = -\dot{\varphi}u_a$. Then, the Klein-Gordon equation leads to

$$\ddot{\varphi} + \Theta\dot{\varphi} + V'(\varphi) = 0\,, \tag{2.32}$$

which can be seen as the energy-density conservation law of a minimally coupled scalar field. The momentum conservation law is given by (2.28). To verify this insert Eqs. (2.30) into the momentum conservation law of a single perfect fluid (see Eqs. (2.21b) in § II E).

G. Spatial curvature

When there is no rotation, the u_a field is hypersurface orthogonal and the projection tensor becomes the metric of the observer's instantaneous rest space. In the presence of vorticity, however, this is no longer the case (due to the Frobenius' theorem – see [12, 13]) and the observers' rest spaces do not mesh together smoothly.

The information regarding the 3-dimensional (spatial) curvature is encoded in the projected Riemann tensor. The latter is defined as follows

$$\mathcal{R}_{abcd} = h_a{}^q h_b{}^s h_c{}^f h_d{}^p R_{qsfp} - v_{ac}v_{bd} + v_{ad}v_{bc}\,, \tag{2.33}$$

where $v_{ab} = \mathrm{D}_b u_a$ is the relative flow tensor between two neighbouring observers (see § II D). On using Eqs. (2.3)-(2.11a), together with decompositions (2.4) and (2.6), one can express $\mathcal{R}_{abcd}$ in terms of the irreducible dynamic and kinematic variables of the spacetime (see [2] for more details and for the full expressions). It then follows that $\mathcal{R}_{abcd} = \mathcal{R}_{[ab][cd]}$ and that $\mathcal{R}_{abcd} \neq \mathcal{R}_{cdab}$. Note that the latter inequality holds only when $\omega_{ab} \neq 0$. In the opposite case, the spatial Riemann tensor possesses all the symmetries of its 4-dimensional counterpart.

In direct analogy with the 4-dimensions, we can define the spatial Ricci tensor and the associated Ricci scalar by taking successive contractions of the corresponding Riemann tensor. The former, in particular, is defined by $\mathcal{R}_{ab}=h^{cd}\mathcal{R}_{acbd}=\mathcal{R}^{c}{}_{acb}$ and satisfies the so-called Gauss-Codacci formula

$$\mathcal{R}_{ab}=E_{ab}+\frac{2}{3}\left(\rho-\frac{1}{3}\Theta^2+\sigma^2-\omega^2+\Lambda\right)h_{ab}+\frac{1}{2}\,\pi_{ab}-\frac{1}{3}\,\Theta(\sigma_{ab}+\omega_{ab})+\sigma_{c\langle a}\sigma^{c}{}_{b\rangle}$$
$$-\omega_{c\langle a}\omega^{c}{}_{b\rangle}+2\sigma_{c[a}\omega^{c}{}_{b]}\,. \tag{2.34}$$

Further contraction leads to the following expression for the spatial Ricci scalar

$$\mathcal{R}=h^{ab}\mathcal{R}_{ab}=2\left(\rho-\frac{1}{3}\,\Theta^2+\sigma^2-\omega^2+\Lambda\right)\,. \tag{2.35}$$

The above, which may also be seen as the generalised Friedmann equation, combines with Eq. (2.34) to give an alternative version for the Gauss-Codacci formula

$$\mathcal{R}_{ab}=\frac{1}{3}\,\mathcal{R}h_{ab}+E_{ab}+\frac{1}{2}\,\pi_{ab}-\frac{1}{3}\,\Theta(\sigma_{ab}+\omega_{ab})+\sigma_{c\langle a}\sigma^{c}{}_{b\rangle}-\omega_{c\langle a}\omega^{c}{}_{b\rangle}+2\sigma_{c[a}\omega^{c}{}_{b]} \tag{2.36}$$

H. Weyl curvature

The Weyl field satisfies the once contracted Bianci identities, namely the set $\nabla^{d}C_{abcd}=\nabla_{[b}R_{a]c}+\frac{1}{6}\,g_{c[b}\nabla_{a]}R$, which in a sense act as the field equations of the non-local component of the gravitational field. These relations split into a set of two propagation and two constraint formulae that monitor tidal forces and gravitational waves. The propagation equations are

$$\dot{E}_{\langle ab\rangle}=-\Theta E_{ab}-\frac{1}{2}\,(\rho+p)\sigma_{ab}+\,\mathrm{curl}\,H_{ab}-\frac{1}{2}\,\dot{\pi}_{ab}-\frac{1}{6}\,\Theta\pi_{ab}-\frac{1}{2}\,\mathrm{D}_{\langle a}q_{b\rangle}-A_{\langle a}q_{b\rangle}$$
$$+3\sigma_{\langle a}{}^{c}\left(E_{b\rangle c}-\frac{1}{6}\,\pi_{b\rangle c}\right)+\varepsilon_{cd\langle a}\left[2A^{c}H_{b\rangle}{}^{d}-\omega^{c}\left(E_{b\rangle}{}^{d}+\frac{1}{2}\,\pi_{b\rangle}{}^{d}\right)\right] \tag{2.37}$$

and

$$\dot{H}_{\langle ab\rangle}=-\Theta H_{ab}-\,\mathrm{curl}\,E_{ab}+\frac{1}{2}\,\mathrm{curl}\,\pi_{ab}+3\sigma_{\langle a}{}^{c}H_{b\rangle c}-\frac{3}{2}\,\omega_{\langle a}q_{b\rangle}$$
$$-\varepsilon_{cd\langle a}\left(2A^{c}E_{b\rangle}{}^{d}-\frac{1}{2}\,\sigma^{c}{}_{b\rangle}q^{d}+\omega^{c}H_{b\rangle}{}^{d}\right)\,. \tag{2.38}$$

The constraints, on the other hand, are given by

$$\mathrm{D}^{b}E_{ab}=\frac{1}{3}\,\mathrm{D}_{a}\rho-\frac{1}{2}\mathrm{D}^{b}\pi_{ab}-\frac{1}{3}\,\Theta q_{a}+\frac{1}{2}\,\sigma_{ab}q^{b}-3H_{ab}\omega^{b}+\varepsilon_{abc}\left(\sigma^{b}{}_{d}H^{cd}-\frac{3}{2}\,\omega^{b}q^{c}\right) \tag{2.39}$$

and

$$\mathrm{D}^b H_{ab} = (\rho + p)\omega_a - \frac{1}{2}\,\mathrm{curl}\, q_a + 3E_{ab}\omega^b - \frac{1}{2}\,\pi_{ab}\omega^b - \varepsilon_{abc}\sigma^b{}_d\left(E^{cd} + \frac{1}{2}\,\pi^{cd}\right) . \quad (2.40)$$

The time derivatives of (2.37) and (2.38) lead to a pair of wavelike equations for the electric and the magnetic Weyl tensor, showing how curvature distortions propagate like ripples in the spacetime fabric. Finally, it is worth noting that the above four expressions are intriguingly similar to Maxwell's formulae, which explains the names of E_{ab} and H_{ab}.

III. SPATIALLY HOMOGENEOUS COSMOLOGIES

So far we have looked at inhomogeneous and anisotropic spacetimes. The current observational data, however, together with our theoretical prejudice support a homogeneous and isotropic universe. We therefore believe that, at least on large enough scales, our universe is closely described by the FRW models.

A. The FRW universes

The simplest non-static solution of the Einstein field equations corresponds to the Robertson-Walker metric, which in suitable (comoving) coordinates reads

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\left[\mathrm{d}r^2 + f_K^2(r)(\mathrm{d}\theta^2 + \sin^2\theta\mathrm{d}\phi^2)\right] , \quad (2.41)$$

where the form of $f_K(r)$ reflects the geometry of the 3-D hypersurfaces. The scale factor (a) defines a characteristic length scale and leads to the Hubble parameter $H = 3\dot{a}/a$, which determines the rate of the expansion. The isotropy of the 3-space ensures that $\mathcal{R} = 6K/a^2$ (see Eq. (2.45) below), with the curvature index K normalised to $0, \pm 1$. In summary,

$$f_K(r) = \begin{cases} \sin r & \text{for } K = +1\,, \\ r & \text{for } K = 0\,, \\ \sinh r & \text{for } K = -1\,. \end{cases} \quad (2.42)$$

When $K = 0$ the 3-space is flat and the geometry Euclidean. Alternatively, we have closed spatial sections for $K = +1$ and open when $K = -1$, corresponding to spherical and hyperbolic geometry respectively. In the $K = +1$ case the 3-D volume is finite, but in the other two is unbounded (unless nontrivial topologies are employed).

B. The FRW dynamics

The high symmetry of the Friedmann spacetimes allows only for scalar variables that depend on time only. Therefore, in covariant terms an FRW universe has $\Theta = 3H(t) \neq 0$, $\sigma_{ab} = 0 = \omega_a = A_a$ and $E_{ab} = 0 = H_{ab}$. Also, due to the isotropy of the Friedmann models, matter is always in the form of a perfect-fluid (with $\rho = \rho(t)$ and $p = p(t)$). Finally, spatial homogeneity demands that all the orthogonally projected gradients (e.g. $\mathrm{D}_a\rho$, $\mathrm{D}_a p$, etc) vanish. As a result, the only nontrivial formulae are those following from Eqs. (2.14), (2.19) and (2.35), namely

$$\dot{H} = -H^2 - \frac{1}{6}(\rho + 3p) + \frac{1}{3}\Lambda\,, \qquad\qquad \dot{\rho} = -3H(\rho + p) \tag{2.43}$$

and

$$H^2 = \frac{1}{3}\rho - \frac{K}{a^2} + \frac{1}{3}\Lambda\,. \tag{2.44}$$

We also note that the 3-Riemann tensor of the FRW models simplifies to

$$\mathcal{R}_{abcd} = \frac{K}{a^2}\left(h_{ac}h_{bd} - h_{ad}h_{bc}\right). \tag{2.45}$$

Introducing the density parameters $\Omega_\rho = \rho/3H^2$, $\Omega_\Lambda = \Lambda/3H^2$ and $\Omega_K = -K/(aH)^2$, expression (2.44) takes the form

$$1 = \Omega_\rho + \Omega_K + \Omega_\Lambda\,. \tag{2.46}$$

Thus, when $\Lambda = 0 = K$, the matter density takes the critical value $\rho = \rho_c = 3H^2$. Allowing for the cosmological constant, we may recast Eq. (2.43a) into

$$qH^2 = \frac{1}{6}(\rho + 3p) - \frac{1}{3}\Lambda\,, \tag{2.47}$$

where $q = -\ddot{a}a/\dot{a}^2 = -[1 + (\dot{H}/H^2)]$ is the deceleration parameter. When $q < 0$ the universe accelerates, which means that in exact FRW models with zero Λ one needs

to violate the strong energy condition (i.e. set $\rho + 3p < 0$) to achieve accelerated expansion.

The expansion rate of the universe also defines an additional representative length scale, known as the Hubble radius/length,

$$\lambda_H = H^{-1} \,. \tag{2.48}$$

In Friedmann models with conventional matter, the Hubble length effectively coincides with the particle horizon ($d_H \propto t$), which determines the size of the causally connected regions.

The scale factor of an FRW spacetime with non-Euclidean 3-geometry also defines the curvature scale ($\lambda_K = a$) of the model. This is the threshold at which the non-Euclidean geometry of the spatial hypersurfaces starts becoming important. In the absence of a cosmological constant, the Hubble scale and the curvature radius are related by (see Eq. (2.44))

$$\left(\frac{\lambda_K}{\lambda_H}\right)^2 = -\frac{K}{1-\Omega_\rho} \,. \tag{2.49}$$

C. The FRW evolution

To close the system of (2.43a) and (2.43b) we need an equation of state for the matter. Assuming a barotropic medium with $p = w\rho$, where w is the (constant) barotropic index, the continuity equation (see (2.43b)) gives

$$\rho = \rho_0 \left(\frac{a_0}{a}\right)^{3(1+w)} \,. \tag{2.50}$$

Substituting the above result into the Friedmann equation and assuming that $\Lambda = 0 = K$, we arrive at

$$a = a_0 \left(\frac{t}{t_0}\right)^{2/3(1+w)} \,, \tag{2.51}$$

when $w \neq -1$. For non-relativistic matter with $w = 0$ (e.g. baryonic 'dust' or non-baryonic cold dark matter), we find that $a \propto t^{2/3}$. Alternatively, we obtain $a \propto t^{1/2}$ in the case of relativistic species (e.g. isotropic radiation) and $a \propto t^{1/3}$ for a stiff medium

with $w = 1$. A special case is that of $w = -1/3$ (matter with zero gravitational mass/energy), which leads to 'coasting' expansion with $a \propto t$. To address the $w = -1$ (slow-rolling scalar field) we recall that then $\rho = \rho_0 =$ constant (see Eq. (2.50)). Substituted into (2.44), the latter leads to $H = H_0 =$ constant and subsequently to exponential expansion (inflation) with $a \propto \mathrm{e}^{H_0(t-t_0)}$.

The continuity equation is independent of the 3-geometry, which makes expression (2.50) applicable to models with nonzero spatial curvature. In FRW cosmologies with non-Euclidean spatial geometry it helps to express the scale-factor in terms of the conformal time (η – with $\dot{\eta} = 1/a$). Then, for $K = +1$, $\Lambda = 0$ and $w \neq -1/3, -1$ Eqs. (2.44), (2.50) combine to give

$$a = a_0 \left\{ \frac{\sin[(1+3w)\eta/2]}{\sin[(1+3w)\eta_0/2]} \right\}^{2/(1+3w)} , \tag{2.52}$$

where $(1+3w)\eta/2 \in (0, \pi)$. Then, the $\eta = \pi/(1+3w)$ threshold denotes the time maximum expansion when $a = a_{max} = a_0\{\sin[(1+3w)\eta_0/2]\}^{-2/(1+3w)}$. For $w = 0$ and $w = 1/3$ the above solution reduces to $a \propto \sin^2(\eta/2)$ and $a \propto \sin\eta$ respectively. When $w = -1/3$ we have to use Eq. (2.43a), instead of the above, which leads to $a \propto t$. For the $w = -1$ case we employ expressions (2.43b) and (2.44). As in the spatially flat models, the former of these relations ensures that $\rho = \rho_0 =$ constant, while the letter leads to $a(1 + \sqrt{3/\rho_0}H) \propto \mathrm{e}^{\sqrt{(\rho_0/3)}\,t}$.

Applied to FLRW cosmologies with hyperbolic spatial geometry, zero cosmological constant and $w \neq -1/3, -1$, the above described analysis leads to

$$a = a_0 \left\{ \frac{\sinh[(1+3w)\eta/2]}{\sinh[(1+3w)\eta_0/2]} \right\}^{2/(1+3w)} , \tag{2.53}$$

where now $(1+3w)\eta/2 > 0$. Assuming pressure-free 'dust' and radiation, we find $a \propto \sinh^2(\eta/2)$ and $a \propto \sinh\eta$ respectively. Also, as before, the system of (2.43) and (2.44) ensures that $a \propto t$ when $w = -1/3$ and $a(1+\sqrt{3/\rho_0}H) \propto \mathrm{e}^{\sqrt{(\rho_0/3)}\,t}$ for $w = -1$.

D. The Bianchi universes

Despite the success of the Friedmann models, the observable universe is neither homogeneous nor isotropic (at least on certain scales and to a certain extent). The

spatially homogeneous, but anisotropic, Bianchi models have long been used to understand the observed level of isotropy in our universe, as well as to probe the nature of the initial singularity.

1. Classification of the Bianchi models

We generally distinguish between two differents kinds of Bianchi cosmologies. The non-tilted (or orthogonal) models, where the fluid flow-lines are normal to the hypersurfaces of homogeneity, and the tilted models where this is no longer the case.

Three are the basic ways of classifying the orthogonal Bianchi models and are all based on the commutation laws of the associated tetrad basis. Consider the tetrad $\{\mathbf{e}_a,\, a = 0, 1, 2, 3\}$, where $\mathbf{e}_0$ is normal to the hypersurfaces of homogeneity and

$$[\mathbf{e}_a, \mathbf{e}_b] = \gamma^c{}_{ab}\mathbf{e}_c\,, \tag{2.54}$$

with the $\gamma^a{}_{bc} = \gamma^a{}_{bc}(t)$ functions treated as dynamical variables themselves [14]. The spatial commutators $\gamma^\alpha{}_{\beta\gamma}$, with $\alpha, \beta, \gamma = 1, 2, 3$, are then decomposed into the pair $n_{\alpha\beta}$ and a_α that satisfies the condition

$$n_{\alpha\beta}a^\beta = 0\,. \tag{2.55}$$

Choosing our tetrad so that $n_{\alpha\beta}$ is diagonalisable (i.e. $n_{\alpha\beta} = \mathrm{diag}(n_1, n_2, n_3)$) and $a_\alpha = (a, 0, 0)$, the above reduces to $n_1 a = 0$. One may therefore immediately define two major classes of Bianchi models. Those with $a = 0$ are the class A models and spacetimes with $a \neq 0$ are termed class B, while further classification follows from the signs of n_1, n_2, n_3. In the class-B case, one may also introduce the scalar parameter h, so that

$$a^2 = hn_2n_3\,. \tag{2.56}$$

The general Bianchi classification is given in Tab. I. There one can see that some of the Bianchi models incorporate the FRW spacetimes as special cases.

Table I. The non-tilted Bianchi spacetimes classified into two group classes and ten group types (see [15]).

Group class	Group type	n_1	n_2	n_3	FLRW as special case
A ($a=0$)	I	0	0	0	$K=0$
	II	+	0	0	−
	VI_0	0	+	−	−
	VII_0	0	+	+	$K=0$
	$VIII$	−	+	+	−
	IX	+	+	+	$K=+1$
B ($a\neq 0$)	V	0	0	0	$K=-1$
	IV	0	0	+	−
	VI_h	0	+	−	−
	VII_h	0	+	+	$K=-1$

E. The Bianchi I models

The non-tilted Bianchi I models are the simplest anisotropically expanding cosmologies that contain the spatially flat FRW universe as a special case. In comoving coordinates, the associated line element takes the form

$$\mathrm{d}s^2=-\mathrm{d}t^2+X^2(t)\mathrm{d}x^2+Y^2(t)\mathrm{d}y^2+Z^2(t)\mathrm{d}z^2\,. \tag{2.57}$$

This allows for different expansion rates along the three principal spatial directions, with the average scale factor and the mean Hubble parameter given by $a=\sqrt[3]{XYZ}$ and $H=\Theta/3=\dot{a}/a$ respectively. Covariantly, the Bianchi I spacetimes are characterised by

$$\omega_a=0=A_a=H_{ab}=\mathcal{R}_{ab}\,, \tag{2.58}$$

which means that the flow-lines are irrotational geodesics. Also, the only nonzero quantities left, are the volume scalar, the shear tensor and the electric part of the

Weyl field. Finally, we should point out that the Bianchi I family can support imperfect fluids with non-vanishing anisotropic pressure but with zero energy flux (i.e. $\pi_{ab} \neq 0$ though $q_a = 0$). This makes the type-I models natural hosts of large-scale magnetic fields [16].

The formulae monitoring the evolution of the Bianchi I cosmologies follow from the general expressions given in § II. In particular, written in a type-I environment and for $\Lambda = 0$, Eqs. (2.14), (2.19) and (2.35) reduce to

$$\dot{H} = -H^2 - \frac{1}{6}\,(\rho + 3p) - \frac{2}{3}\,\sigma^2\,, \qquad \dot{\rho} = -3H(\rho + p) - \sigma_{ab}\pi^{ab} \tag{2.59}$$

and

$$H^2 = \frac{1}{3}\left(\rho + \sigma^2\right)\,. \tag{2.60}$$

The latter of the above, which is the Bianchi I analogue of the Friedmann equation, also reads

$$1 = \Omega_\rho + \Sigma\,, \tag{2.61}$$

with $\Sigma = \sigma^2/3H^2$ measuring the model's shear anisotropy. Similarly, expressions (2.15), (2.36) and (2.37) simplify to

$$\dot{\sigma}_{ab} = -2H\sigma_{ab} - \sigma_{c\langle a}\sigma^{c}{}_{b\rangle} - E_{ab} + \frac{1}{2}\,\pi_{ab}\,, \qquad E_{ab} = H\sigma_{ab} - \sigma_{c\langle a}\sigma^{c}{}_{b\rangle} - \frac{1}{2}\,\pi_{ab} \tag{2.62}$$

and

$$\dot{E}_{ab} = -3HE_{ab} - \frac{1}{2}\,(\rho + p)\sigma_{ab} - \frac{1}{2}\,(\dot{\pi}_{ab} + H\pi_{ab}) + 3\sigma_{\langle a}{}^{c}\left(E_{b\rangle c} - \frac{1}{6}\,\pi_{b\rangle c}\right)\,, \tag{2.63}$$

respectively. Finally, on using (2.62b), Eq. (2.62a) recasts into

$$\dot{\sigma}_{ab} = -3H\sigma_{ab} + \pi_{ab}\,. \tag{2.64}$$

The latter ensures that, in the absence of anisotropic pressures, the shear depletes as a^{-3}, where a is the average (over the three spatial directions) scale factor.

The set (2.59)-(2.64) fully governs the Bianchi I dynamics, once the matter equation of state has been introduced. For a barotropic perfect fluid with $p = w\rho$, expression

(2.59b) integrates to $\rho \propto a^{-3(1+w)}$. Following (2.64) and (2.60), this means that the shear will dominate the early expansion, unless $w = 1$. These shear-dominated early stages are essentially identical to the Kasner vacuum, in which case $a \propto t^{1/3}$ (combine Eqs. (2.60), (2.64)). The line element of the Kasner solution is

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + t^{2p_1}\mathrm{d}x^2 + t^{2p_2}\mathrm{d}y^2 + t^{2p_3}\mathrm{d}z^2\,, \tag{2.65}$$

where $p_1 + p_2 + p_3 = 1 = p_1^2 + p_2^2 + p_3^2$ [17–19]). These conditions guarantee that either one of the three exponents is negative or two are zero. In the former case the spacetime expands in two directions and contracts in the third, with a cigar-like initial singularity. In the latter case we have expansion in one direction only, which corresponds to a pancake-type singularity.

F. Isotropisation of Bianchi models

Bianchi cosmologies have been traditionally studied by means of dynamical system methods [15, 20]. The existing studies have revealed that many Bianchi models exhibit "intermediate isotropisation". This happens because a number of Bianchi spacetimes have phase planes where the FRW solutions are saddle points. This means that, although these models are generically very different from the Friedmann spacetimes, they can isotropise and look very much like an FRW universe over an extended period of their evolution.

Another issue is whether the aforementioned isotropisation occurs at early or at late times. Following [21], Bianchi models with conventional matter do not isotropise in the future. It has been shown, however, that Bianchi cosmologies with nonzero cosmological constant tend towards late-time isotropy, This seems to suggest that inflation should smooth the anisotropy of these models out. It should be noted, however, that the existing results depend on the amount of the initial anisotropy and apply to the non-tilted Bianchi types.

IV. INHOMOGENEOUS AND ANISOTROPIC COSMOLOGIES

The simplest inhomogeneous cosmologies are the spherically symmetric Lemaitre-Tolman-Bondi (LTB) models. There are also nhomogenoeus solutions of the Einstein equations, such as the Szekeres spacetime or the so-called Swiss-Cheese models, which posses no symmetry. Our universe is believed to be free of global symmetries.

A. The gauge problem

Cosmological perturbation theory has long been known to suffer from what is known as the *gauge problem*. This stems from the fact that when study perturbations we deal with two spacetimes [22–26]. The physical spacetime ($\mathcal{W}$), which corresponds to the real universe, and a fictitious one ($\overline{\mathcal{W}}$) that defines our idealised (unperturbed) background. In most cosmological studies the background spacetime is an FRW model.

To proceed with the analysis, one needs to establish a one-to-one correspondence, a gauge $\phi : \overline{\mathcal{W}} \to \mathcal{W}$, between the two aforementioned spacetimes. Then, Although the mapping is generally arbitrary, depending on the case in hand, some may be more suitable than others. Changes in $\phi : \overline{\mathcal{W}} \to \mathcal{W}$, which are known as *gauge transformations*, change the event of $\mathcal{W}$ that is associated with a given event of $\overline{\mathcal{W}}$. Gauge transformations are therefore different from coordinate transformations which merely relabel events. The gauge problem stems from our inherent freedom to make gauge transformations.

The perturbation of a given quantity is defined as the difference between its value at some event in the real spacetime and its value at the corresponding (through the gauge) event in the background. This means that variables (even scalar ones) with nonzero and position-dependent background values, will have gauge-dependent perturbations [27, 28]. To demonstrate this in a simple way, let us consider the familiar density perturbation $\delta\rho = \rho - \overline{\rho}$, where ρ is the matter density. The difference in the right-hand side is between two corresponding points in the background and the real spacetime. A gauge transformation will generally change this correspondence and therefore the perturbation value. This makes the value of $\delta\rho$ is gauge-dependent and

arbitrary. For example, one can assume a gauge that associates the surfaces of constant density in $\bar{\mathcal{W}}$ with the surfaces of constant density in $\mathcal{W}$, thus setting $\delta\rho$ to zero [26].

B. Gauge-invariant perturbations

We can address the gauge problem by fixing the point-identification map between the background and the real spacetime. Determining the best gauge for a given physical problem, however, is not a trivial task. Alternatively, one can fix the gauge partially, leaving some residual gauge freedom, or employ gauge-invariant variables [25, 26].

1. Criteria for gauge invariance

Gauge-independent quantities are unaffected by gauge transformations between the idealised and the realistic spacetime. The Stewart & Walker lemma shows that, at the linear perturbative level, the simplest gauge-invariant variables are scalars that are constant in the background, or tensors which vanish there [27]. An alternative, though less straightforward option, is to use tensors that are linear combinations of products of the Kronecker deltas with constant coefficients. Note that he same general criteria also apply to second-order perturbations. This time, however, the Stewart & Walker requirements must be satisfied by the first-order variables [29].

In most cosmological studies the background spacetime is that of an FRW model. It is therefore important to know which quantities satisfy this Stewart & Walker criterion on Friedmann backgrounds. As it turns out, the most useful gauge-invariant quantities are those that vanish in FRW environments. Then, the symmetries of the Friedmann models ensure that such variables are those describing spatial inhomogeneity or anisotropy.

2. Gauge-invariant inhomogeneities

Spatial inhomogeneities in the distribution of a physical quantity are described by its orthogonally projected (spatial) gradient. In structure formation studies, the key

variable is the dimensionless gradient of the matter energy density [26]

$$\Delta_a = \frac{a}{\rho}\,\mathrm{D}_a\rho\,. \tag{2.66}$$

The above monitors density variations between a pair of neighbouring fundamental observers, which means that vanishes identically in spatially homogeneous spacetimes [26]. More specifically, we have

$$\mathrm{D}_a\rho = h_a{}^b\nabla_b\rho = h_a{}^0\nabla_0\rho + h_a{}^\alpha\nabla_\alpha\rho = 0\,, \tag{2.67}$$

since $h_a{}^0 = 0$ in a comoving frame and $\nabla_\alpha\rho = 0$ because $\rho = \rho(t)$. Therefore, Δ_a satisfies the Stewart & Walker lemma [27], which means that it describes linear density inhomogeneities in a gauge independent way. This way, one can exploit the high symmetry of the FRW models to construct additional gauge-invariant variables. From the perspective of structure formation, the most useful auxiliary gauge-independent quantity is given by the gradient

$$\mathcal{Z}_a = a\mathrm{D}_a\Theta\,, \tag{2.68}$$

which monitors spatial inhomogeneities in the volume expansion.

C. Inhomogeneous single-fluid cosmologies

The current view is that the structure we observe in our universe today is the result of a physical mechanism known as "gravitational, or Jeans, instability". In this approach small inhomogeneities (perturbations) in the initial density distribution of the matter grow gravitationally to form the galaxies, the galaxy clusters and the voids seen in the universe today.

1. Imperfect fluids

As a first step towards understanding the mechanism of gravitational instability, let us consider a general spacetime filled with a single imperfect fluid. Spatial inhomogeneities in the matter density, as measured between a pair of neighbouring

observers, are monitored by the dimensionless comoving gradient $\Delta_a = (a/\rho)\mathrm{D}_a\rho$. The covariant derivative of the above, together with the energy and the momentum conservation laws (see (2.19) and (2.20)), gives

$$\dot{\Delta}_{\langle a\rangle} = \frac{p}{\rho}\,\Theta\Delta_a - \left(1+\frac{p}{\rho}\right)\mathcal{Z}_a + \frac{a\Theta}{\rho}\left(\dot{q}_{\langle a\rangle} + \frac{4}{3}\,\Theta q_a\right) - \frac{a}{\rho}\,\mathrm{D}_a\mathrm{D}^b q_b + \frac{a\Theta}{\rho}\,\mathrm{D}^b\pi_{ab}$$
$$-\left(\sigma^b{}_a + \omega^b{}_a\right)\Delta_b - \frac{a}{\rho}\,\mathrm{D}_a\left(2A^b q_b + \sigma^{bc}\pi_{bc}\right) + \frac{a\Theta}{\rho}\,(\sigma_{ab}+\omega_{ab})\,q^b + \frac{a\Theta}{\rho}\,\pi_{ab}A^b$$
$$+\frac{1}{\rho}\left(\mathrm{D}^b q_b + 2A^b q_b + \sigma^{bc}\pi_{bc}\right)(\Delta_a - aA_a)\,. \tag{2.69}$$

In the right-hand side we see a number of agents acting as sources of density perturbations. In other words, even if Δ_a was originally zero, it is not expected to remain so. One of the sources is $\mathcal{Z}_a$, the volume expansion gradient, the nonlinear evolution of which is monitored by

$$\dot{\mathcal{Z}}_{\langle a\rangle} = -\frac{2}{3}\,\Theta\mathcal{Z}_a - \frac{1}{2}\,\rho\Delta_a - \frac{3}{2}\,a\mathrm{D}_a p - a\left[\frac{1}{3}\,\Theta^2 + \frac{1}{2}\,(\rho+3p) - \Lambda\right]A_a + a\mathrm{D}_a\mathrm{D}^b A_b$$
$$-\left(\sigma^b{}_a + \omega^b{}_a\right)\mathcal{Z}_b - 2a\mathrm{D}_a\left(\sigma^2-\omega^2\right) + 2aA^b\mathrm{D}_a A_b$$
$$-a\left[2\left(\sigma^2-\omega^2\right) - \mathrm{D}^b A_b - A^b A_b\right]A_a\,. \tag{2.70}$$

The latter is obtained after taking the (proper) time derivative of (2.68) and then using Raychaudhuri's formula.

2. Perfect fluids

Most theoretical studies assume perfect fluids, in which case $q_a = 0 = \pi_{ab}$ by default and Eq. (2.69) reduces to

$$\dot{\Delta}_{\langle a\rangle} = \frac{p}{\rho}\,\Theta\Delta_a - \left(1+\frac{p}{\rho}\right)\mathcal{Z}_a - \left(\sigma^b{}_a + \omega^b{}_a\right)\Delta_b\,. \tag{2.71}$$

Assuming an ideal medium does not change the form of expression (2.70). The only difference is that now Δ_a is monitored by (2.71) and A_a by (2.21b), instead of (2.20). In the barotropic case (i.e. for $p = p(\rho)$), we have the additional relation $\mathrm{D}_a p = c_{\mathrm{s}}^2\mathrm{D}_a\rho$, between the pressure and the density gradients, with $c_{\mathrm{s}}^2 = \dot{p}/\dot{\rho}$ representing the adiabatic sound speed.

3. Scalar fields

Recalling that $\mathrm{D}_a\varphi = 0$ always (see Eq. (2.26)), which guarantees that $\mathrm{D}_a V(\varphi) = 0$ as well, the orthogonally projected gradients of (2.31) gives

$$\mathrm{D}_a p^{(\varphi)} = \mathrm{D}_a \rho^{(\varphi)} = \dot{\varphi}\mathrm{D}_a\dot{\varphi}\,. \tag{2.72}$$

Spatial inhomogeneities in the (effective) energy density of the φ-field are described by the dimensionless gradient $\Delta_a^{(\varphi)} = (a/\rho^{(\varphi)})\mathrm{D}_a\rho^{(\varphi)}$. On using the latter, expression (2.72) leads to

$$a\mathrm{D}_a p^{(\varphi)} = \rho^{(\varphi)}\Delta_a^{(\varphi)}\,, \tag{2.73}$$

or, equivalently, to $\Delta_a^{(\varphi)} = (a\dot{\varphi}/\rho^{(\varphi)})\mathrm{D}_a\dot{\varphi}$. It follows that, despite their non-barotropic equation of state, scalar fields behave as effective stiff fluids in hydrodynamic terms. This 'duality' of the scalar-fields nature, which represents a major departure from the conventional perfect-fluid behaviour, is reflected in the statement that scalar-field perturbations are 'non-adiabatic'.

Setting $\Delta_a^{(\varphi)} = (a/\rho^{(\varphi)})\mathrm{D}_a\rho^{(\varphi)}$, assuming zero cosmological constant and using relation (2.73), the nonlinear formulae of the previous sections are adapted to the case of a minimally coupled scalar field. In particular, Eqs. (2.69) and (2.70) become

$$\dot{\Delta}^{(\varphi)}_{\langle a\rangle} = \frac{p^{(\varphi)}}{\rho^{(\varphi)}}\,\Theta\Delta_a^{(\varphi)} - \left(1+\frac{p^{(\varphi)}}{\rho^{(\varphi)}}\right)\mathcal{Z}_a - \left(\sigma^b{}_a + \omega^b{}_a\right)\Delta_b^{(\varphi)}\,, \tag{2.74}$$

and

$$\begin{aligned}\dot{\mathcal{Z}}_{\langle a\rangle} = &-\frac{2}{3}\,\Theta\mathcal{Z}_a - 2\rho^{(\varphi)}\Delta_a^{(\varphi)} - a\left[\frac{1}{3}\,\Theta^2 + \frac{1}{2}\,(\rho^{(\varphi)}+3p^{(\varphi)})\right]A_a + a\mathrm{D}_a\mathrm{D}^bA_b - \left(\sigma^b{}_a+\omega^b{}_a\right)\mathcal{Z}_b\\ &-2a\mathrm{D}_a\left(\sigma^2-\omega^2\right) + 2aA^b\mathrm{D}_aA_b - a\left[2\left(\sigma^2-\omega^2\right) - \mathrm{D}^bA_b - A^bA_b\right]A_a\,,\end{aligned} \tag{2.75}$$

respectively. Also, combining relations (2.28) and (2.73), leads to the following expression for the 4-acceleration (see Eq. (2.24) in § II F 2)

$$a(\rho^{(\varphi)}+p^{(\varphi)})A_a = -\rho^{(\varphi)}\Delta_a^{(\varphi)}\,. \tag{2.76}$$

Relations (2.74)-(2.76) monitor the nonlinear evolution of density perturbations in scalar-field cosmologies.

V. LINEAR COSMOLOGICAL PERTURBATIONS

The isotropy of the CMB, combined with the Cosmological Principle, lead to the conclusion that the universe was extremely homogeneous at the time of last scattering. In fact, taken at face value, the CMB data ensure that $\Delta \simeq 10^{-5}$ at recombination. Therefore, the initial stages of structure formation can be adequately studied within the linear approximation.

A. Linearisation

When linearising the full equations, one first has to select an exact solution that describes the unperturbed background universe ($\bar{W}$ – see § IV A). The nonlinear formulae will be then linearised around the chosen background. In our linearisation scheme, terms with nonzero unperturbed value are of zero perturbative order, and those that vanish in the background are weak first order perturbations [26, 30]. This guarantees that all the linear variables are gauge-invariance [27]. Finally, products between first-order quantities are neglected.

Most studies use the homogeneous and isotropic Friedmann models for their unperturbed vackground. This means that the only zero-order quantities are the matter energy density (ρ), isotropic pressure (p) and the volume expansion ($\Theta = 3H$). When the 3-geometry is non-Euclidean, these are supplemented by the 3-Ricci scalar ($\mathcal{R}$).

B. Single-fluid perturbations

1. Linear evolution equations

Let us consider a perturbed, almost-FRW universe with a single barotropic perfect fluid. On this background, Eq. (2.71) linearises to

$$\dot{\Delta}_a = 3wH\Delta_a - (1+w)\mathcal{Z}_a\,, \tag{2.77}$$

where $w = p/\rho$ is the barotropic index of the fluid and $c_{\rm s}^2 = {\rm d}p/{\rm d}\rho$ the square of the adiabatic sound speed. The former obeys the evolution law

$$\dot{w} = -3H(1+w)(c_s^2 - w)\,, \tag{2.78}$$

which means that $w =$ constant when $c_s^2 = w$ (provided $H \neq 0$ and $w \neq -1$).

Similarly, keeping up to first-order terms, expression (2.70) becomes

$$\dot{\mathcal{Z}}_a = -2H\mathcal{Z}_a - \frac{1}{2}\,\rho\Delta_a - \frac{3}{2}\,a{\rm D}_a p - a\left[3H^2 + \frac{1}{2}\,\rho(1+3w) - \Lambda\right]A_a + a{\rm D}_a{\rm D}^b A_b\,, \tag{2.79}$$

with $\rho(1+w)A_a = -{\rm D}_a p$ to linear order – see (2.21b). Also, the linearised momentum conservation law reads

$$a(1+w)A_a = -c_{\rm s}^2\Delta_a\,. \tag{2.80}$$

Then, solving for A_a and substituting the result into Eq. (2.79), we obtain

$$\dot{\mathcal{Z}}_a = -2H\mathcal{Z}_a - \frac{1}{2}\,\rho\Delta_a - \frac{c_{\rm s}^2}{1+w}\left({\rm D}^2\Delta_a + \frac{K}{a^2}\,\Delta_a\right) - 6ac_{\rm s}^2 H\,{\rm curl}\,\omega_a\,, \tag{2.81}$$

where $K = 0, \pm 1$ is the background 3-curvature index and ${\rm D}^2 = {\rm D}^a{\rm D}_a$ is the orthogonally projected (spatial) Laplacian operator.

The linear expansion is determined by the first-order Raychaudhuri equation (see expression (2.14) in § II D)). On using (2.80), Raychaudhuri's formula reads

$$qH^2 = \frac{1}{6}\,\rho(1+3w) + \frac{c_s^2}{3a^2(1+w)}\,\Delta - \frac{1}{3}\,\Lambda\,, \tag{2.82}$$

where q is the deceleration parameter of the perturbed spacetime and $\Delta = a{\rm D}^a\Delta_a$ describes scalar density perturbations (see § V B 2 below). When positive, Δ represents overdensities in the matter distribution and adds to the overall gravitational pull. Underdensities, on the other hand, tend to accelerate the expansion.

To first order, the propagation of shear anisotropies follows from the linear version of Eq. (2.15) after using the momentum-conservation law (2.80),

$$\dot{\sigma}_{ab} = -2H\sigma_{ab} - E_{ab} - \frac{c_s^2}{a^2(1+w)}\Delta_{\langle ab\rangle}\,. \tag{2.83}$$

Here, the symmetric and trace-free tensor $\Delta_{\langle ab\rangle}$ represents distortions in the shape of the perturbation (see § V B 2 below). In an analogous way, substituting (2.80) into the linear counterpart of (2.16), provides the linear formula for the vorticity propagation

$$\dot{\omega}_a = -2\left(1-\frac{3}{2}\,c_s^2\right)H\omega_a\,. \tag{2.84}$$

Therefore, to first order, kinematic vortices dilute with the expansion, unless the matter equation of state is 'stiffer' than $w = 2/3$.

These propagation formulae are supplemented by the following set of linear constraints (see (2.17) and (2.18) for their nonlinear expressions),

$$\mathrm{D}^b\sigma_{ab} = \frac{2}{3}\,\mathrm{D}_a\Theta + \mathrm{curl}\,\omega_a\,, \qquad\qquad \mathrm{D}^a\omega_a = 0 \tag{2.85}$$

and

$$H_{ab} = \mathrm{curl}\,\sigma_{ab} + \mathrm{D}_{\langle a}\omega_{b\rangle}\,. \tag{2.86}$$

Note that, when the right-hand side of (2.85a) vanishes, the shear is transverse and describes pure-tensor perturbations (i.e. gravitational waves – see § V D 1 below). Also, in the absence of rotation the magnetic Weyl component is fully determined by the shear.

Additional linear constraints can be derived from the linearised Gauss-Codacci formula (see § II G), which here takes the form

$$\mathcal{R}_{ab} = \frac{1}{3}\,\mathcal{R}h_{ab} - H(\sigma_{ab}+\omega_{ab}) + E_{ab}\,, \tag{2.87}$$

where

$$\mathcal{R} = 2\left(\rho - \frac{1}{3}\,\Theta^2 + \Lambda\right)\,. \tag{2.88}$$

Note that the latter relation is the linear counterpart of Friedmann's equation.

2. Three types of inhomogeneity

The gradient Δ_a contains collective information on all three types of density inhomogeneities, that is density perturbations (scalar modes), vortices (vector modes)

and shape distortions (trace-free tensor modes). We can decode all this information, by splitting the dimensionless projected gradient $\Delta_{ab} = a\mathrm{D}_b\Delta_a$ into its irreducible components as

$$\Delta_{ab} = \Delta_{\langle ab\rangle} + \Delta_{[ab]} + \frac{1}{3}\,\Delta h_{ab}\,, \tag{2.89}$$

with $\Delta_{\langle ab\rangle} = a\mathrm{D}_{\langle b}\Delta_{a\rangle}$, $\Delta_{[ab]} = a\mathrm{D}_{[b}\Delta_{a]}$ and $\Delta = a\mathrm{D}^a\Delta_a$ [31]. The first of these variables describes anisotropies in the density distribution (e.g. pancakes or cigar-like structures). The second is related to vortex-like distortions and the third monitors overdensities (when $\Delta > 0$), or underdensities (when $\Delta < 0$).

3. Density perturbations

Taking the orthogonally projected gradients of Eqs. (2.77) and (2.79), keeping up to first-order terms and then extracting the trace of the resulting expressions we arrive at

$$\dot{\Delta} = 3wH\Delta - (1+w)\mathcal{Z} \tag{2.90}$$

and

$$\dot{\mathcal{Z}} = -2H\mathcal{Z} - \left[\frac{1}{2}\,\rho + \frac{3Kc_{\mathrm{s}}^2}{a^2(1+w)}\right]\Delta - \frac{c_{\mathrm{s}}^2}{1+w}\mathrm{D}^2\Delta\,, \tag{2.91}$$

respectively. It should be noted that the last expression follows after using the linear constraint $\mathrm{D}^a\,\mathrm{curl}\,\omega_a = 0$. The above system monitors the gravitational clumping of matter in a perturbed almost-FLRW universe that contains a single perfect fluid. We may decouple the equations by taking the time derivative of (2.90) and then using expression (2.78), together with the background relations (2.44) and (2.44). When we further assume that $w =$ constant (e.g. recall that $w = 1/3$ in the radiation era and $w = 0$ after equality), the result reads

$$\ddot{\Delta} = -2\left(1 - \frac{3}{2}\,w\right)H\dot{\Delta} + \left[\frac{1}{2}\,\rho(1-w)(1+3w) + 2\Lambda w\right]\Delta + c_{\mathrm{s}}^2\mathrm{D}^2\Delta\,, \tag{2.92}$$

which is independent of the background 3-curvature [30]. The above is a wave-like equation, with extra terms due to gravity and the universal expansion, governing the

linear evolution of matter aggregations in a almost-FLRW universe filled with a single barotropic perfect fluid.

Our last step is to introduce the scalar harmonics $\mathcal{Q}^{(k)}$ and write $\Delta = \Sigma_k \Delta^{(k)} \mathcal{Q}^{(k)}$, with $\mathrm{D}_a \Delta^{(k)} = 0 = \dot{\mathcal{Q}}^{(k)}$ and k representing the comoving wavenumber of the perturbed mode. Then, after dropping the spatial dependence, expression (2.92) takes the form

$$\ddot{\Delta}^{(k)} = -2\left(1 - \frac{3}{2}\,w\right) H\dot{\Delta}^{(k)} + \left\{\frac{1}{2}\,[\rho(1-w)(1+3w) + 4\Lambda w] - \frac{k^2 c_{\mathrm{s}}^2}{a^2}\right\}\Delta^{(k)}\,. \quad (2.93)$$

The first term on the right-hand side carries the diluting effects of the universal expansion and acts as an effective friction (when $w < 2/3$). The last term, on the other hand, reflects the competition between gravitational attraction and pressure support. In particular, gravity wins when the quantity inside the angled brackets in positive. Hence, since $\lambda_k = a/k$ is the physical wavelength of the perturbation, gravitational contraction occurs on scales larger than

$$\lambda_J \simeq \frac{c_{\mathrm{s}}}{\sqrt{\rho(1-w)(1+3w) + 4\Lambda w}}\,, \quad (2.94)$$

that defines the *Jeans length*. The latter also determines the *Jeans mass*, which during the radiation era (i.e. for $c_s^2 = 1/3$) in the absence of a cosmological constant is given by

$$M_J \propto \rho_b \lambda_J^3 \simeq 10^{16}\left(\frac{\Omega_b}{\Omega}\right)\left(\Omega h^2\right)^{-1/2}\,\mathrm{M}_\odot\,, \quad (2.95)$$

corresponding to a super cluster of galaxies. Note that Ω and Ω_b are respectively the total and the baryonic density parameters, while $H = 100h$ km sec^{-1} Mpc^{-1} (e.g. see [32]).

Let us now look for analytical solutions to Eq. (2.93), assuming that $K = 0 = \Lambda$. Consider first the radiation era, during which $w = 1/3 = c_{\mathrm{s}}^2$), $H = 1/(2t)$ and $\rho = 3/(4t^2)$. Then, in a comoving frame we have

$$\frac{\mathrm{d}^2\Delta^{(k)}}{\mathrm{d}t^2} + \frac{1}{2t}\frac{\mathrm{d}\Delta^{(k)}}{\mathrm{d}t} - \frac{1}{2t^2}\left[1 - \frac{1}{6}\left(\frac{k}{aH}\right)^2\right]\Delta^{(k)} = 0\,, \quad (2.96)$$

where $k/aH = \lambda_H/\lambda_k$ and $\lambda_H = 1/H = 2t$. Thus, the solution is scale-dependent. In particular, on scales larger than the Hubble radius, where $k/aH \ll 1$ and the pressure

support is negligible, Eq. (2.96) leads to

$$\Delta = \Delta_1 t + \Delta_2 t^{-1/2}\,, \tag{2.97}$$

with $\Delta_{1,2} =$ constant. Therefore, prior to equipartition, large-scale matter aggregations grow as $\Delta \propto a^2$. Well inside the the Hubble length, on the other hand, $k/aH \gg 1$. There, pressure gradients in the matter distribution support against gravitational collapse and the over/under-density oscillates according to

$$\Delta^{(k)} = \Delta_1^{(k)} \sin\left[\sqrt{3}\frac{k}{a_0 H_0}\left(\frac{t}{t_0}\right)^{1/2}\right] + \Delta_2^{(k)} \cos\left[\sqrt{3}\frac{k}{a_0 H_0}\left(\frac{t}{t_0}\right)^{1/2}\right], \tag{2.98}$$

where $\Delta_{1,2}^{(k)} =$ constant and the zero suffix indicates a given initial time.

After matter-radiation equality $w = 0 = c_{\rm s}^2$. Then, for $K = 0 = \Lambda$, we have $H = 2/(3t)$, $\rho = 4/(3t^2)$ and Eq. (2.93) leads to the scale-independent expression

$$\Delta = \Delta_1 t^{2/3} + \Delta_2 t^{-1}\,. \tag{2.99}$$

Consequently, matter aggregations in the post-recombination universe grow proportionally to the scale factor. Note that the same result can be obtained through a purely Newtonian treatment as well [32–35].

Before closing this section we should point out that, after equipartition, photons can no longer support against gravitational contraction. Nevertheless, on small enough scales, this role can be played by ordinary baryonic gas pressure. At the time of recombination, in particular, latter is $p^{(b)} \simeq n^{(b)} k_B T_{rec}$, while its radiation counterpart is given by $p^{(\gamma)} \simeq n^{(\gamma)} k_B T_{rec}$ (with k_B representing Boltzmann's constant). Recalling that $n^{(b)} \simeq 10^{-8} n^{(\gamma)}$, these two formulae show why pressure support drops drastically at decoupling. As a result, soon after recombination the Jeans mass reduces to

$$M_J \simeq 10^4 \left(\frac{\Omega_b}{\Omega}\right)\left(\Omega h^2\right)^{-1/2}\,{\rm M}_\odot\,, \tag{2.100}$$

which is close to that of a star cluster [32].

4. Density vortices

Rotational, vortex-like, distortions in the density of the cosmic medium are studied by means of the antisymmetric tensor $\Delta_{[ab]}$ – see decomposition (2.89). Within the

geometrical framework of general relativity, the latter is related to the (kinematic) vorticity tensor via the first-order relation $\Delta_{[ab]} = -3a^2(1+w)H\omega_{ab}$. Similarly, $\mathcal{Z}_{[ab]} = 3a^2\dot{H}\omega_{ab}$, which means that

$$\mathcal{Z}_{[ab]} = -\frac{\dot{H}}{(1+w)H}\Delta_{[ab]}\,, \tag{2.101}$$

at the linear level. Using the above, together with the background relations (2.43a) and (2.44), the linearised skew part of (2.77) leads to

$$\dot{\mathcal{W}}_a = -\frac{(1-w)\rho}{2H}\,\mathcal{W}_a\,, \tag{2.102}$$

where $\mathcal{W}_a \equiv \varepsilon_{abc}\Delta^{bc}/2$. Note that in deriving the above, which applies to all scales, we have set the cosmological constant to zero and assumed Euclidean 3-spaces.

Introducing the vector harmonics $\mathcal{Q}_a^{(k)}$ (with $\dot{\mathcal{Q}}_a^{(k)} = 0 = \mathrm{D}^a\mathcal{Q}_a^{(k)}$), we may write $\mathcal{W}_a = \Sigma_k \mathcal{W}^{(k)}\mathcal{Q}_a^{(k)}$ with $\mathrm{D}_a\mathcal{W}^{(k)} = 0$. Then, the k-th mode evolves as

$$\dot{\mathcal{W}}^{(k)} = -\frac{(1-w)\rho}{2H}\,\mathcal{W}^{(k)}\,, \tag{2.103}$$

giving $\mathcal{W} \propto t^{-1/2}$ when radiation dominates and $\mathcal{W} \propto t^{-1}$ for dust. In other words, linear vortices in the density distribution of a perturbed FRW universe (with $K = 0 = \Lambda$) decay with time on all scales.

5. Dissipative effects

So far we have looked at the purely gravitational evolution of matter perturbations and ignored any dissipative effects. When dissipation is accounted for, there are other processes that generally modify the "standard" picture. When dealing with baryons, the most important effect come from their coupling to the pre-recombination photons. The latter diffuse from high-density to low-density regions, dragging the baryons along and erasing inhomogeneities in the baryonic distribution [36, 37]. This is known as *Silk damping* and wipes out fluctuations in the baryon component on scales smaller than the *Silk mass*

$$M_S \propto \rho_b \ell_S^3 \simeq 6.2\times10^{12}\left(\frac{\Omega}{\Omega_b}\right)^{3/2}\left(\Omega h^2\right)^{-5/4}\mathrm{M}_\odot\,, \tag{2.104}$$

where ℓ_S is the associated *Silk length* (e.g. see [32]). As a result, before recombination, perturbations smaller than M_S are obliterated by Silk damping, those between M_S and M_J oscillate and only modes with $M > M_J$ (with M_J given by Eq. (2.95) in § V B 3) can grow.

Similarly, the free motion of the collisionless (dark matter) species erases structures that try to form in their small-scale distribution. This process is known as *free streaming*, or *Landau damping*, and its proper study requires integrating the collisionless Boltzmann equation of the related species. One can still obtain an estimate, however, by calculating the maximum distance traveled by a free-streaming particle. Thus, for hot thermal relics we have

$$\ell_{FS} \simeq 0.5 \left(\frac{m_{DM}}{1\,\mathrm{keV}}\right)^{-4/3} \left(\Omega_{DM} h^2\right)^{1/3} \mathrm{Mpc}\,, \tag{2.105}$$

with m_{DM} representing the mass of the species in units of 1 keV. Therefore, the minimum scale to survive collisionless dissipation depends crucially on the particles' mass. In particular, the lighter the dark matter component, the less power survives on small scales. For neutrinos with $m_\nu \simeq 30\,\mathrm{eV}$, the above gives $\ell_{FS} \simeq 28\,\mathrm{Mpc}$ and a mass-scale of roughly $10^{15}\,\mathrm{M}_\odot$.

Cold thermal relics (CDM) have very small dispersion velocities and their free-streaming masses are very low. So, perturbations grow unaffected by damping on all scales of cosmological interest, although they suffer stagnation because of the Meszaros effect until the time of matter-radiation equality. After recombination, the potential wells of the collisionless species serve to boost the growth of baryonic perturbations (see § V B 7 next).

6. A radiation and dust universe

A flat almost-FRW cosmology filled with radiation and dust is believed to be a good approximation of the universe prior to decoupling. When the radiative component is identified with the photons and the neutrinos, we have $\rho^{(r)} = \rho^{(\gamma)} + \rho^{(\nu)}$. For the non-relativistic species, namely baryons and cold dark matter, we may write $\rho^{(d)} = \rho^{(b)} + \rho^{(c)}$. Then, the total energy is $\rho = \rho^{(r)} + \rho^{(d)}$, the total pressure is $p = \rho^{(r)}/3$

and the effective total sound speed is $c_s^2 = 4\rho^{(r)}/[3(4\rho^{(r)} + 3\rho^{(d)})]$. Also, neglecting any photon-baryon interactions leads to $\rho^{(r)} = \rho_0^{(r)}(a_0/a)^4$ and $\rho^{(d)} = \rho_0^{(d)}(a_0/a)^3$, where the scale factor satisfies the background Friedmann equation $3H^2 = \rho$.

Suppose that the radiation field is homogeneously distributed (i.e. $\Delta^{(r)} = 0$), which holds on small scales damped by diffusion. One may then consider density perturbations in the dust component only. These are monitored by the system

$$\dot{\Delta}^{(d)} = -\mathcal{Z} - a\mathrm{D}^2 v^{(d)}\,, \qquad\qquad \dot{\mathcal{Z}} = -2H\mathcal{Z} - \frac{1}{2}\,\rho^{(d)}\Delta^{(d)} \tag{2.106}$$

and

$$\dot{v}^{(d)} = -Hv^{(d)}, \tag{2.107}$$

where $v^{(d)}$ is the peculiar velocity of the dust relative to the background Hubble flow [38].Taking the time derivative of (2.106a), and using the linear commutation law $(\mathrm{D}^2 v^{(d)})^{\cdot} = \mathrm{D}^2\dot{v}^{(d)} - 2H\mathrm{D}^2 v^{(d)}$, the above system reduces to the scale independent equation [39]

$$\Delta''^{(d)} = -\frac{2+3a}{2a(1+a)}\,\Delta'^{(d)} + \frac{3}{2a(1+a)}\,\Delta^{(d)}\,. \tag{2.108}$$

Note that primes are derivatives with respect to the scale factor and $a = 1$ at matter-radiation equality [38]. The general solution has the form [40]

$$\Delta^{(d)} = \mathcal{C}_1\left(1+\frac{3}{2}\,a\right) - \mathcal{C}_2\left[\left(1+\frac{3}{2}\,a\right)\ln\left(\frac{\sqrt{1+a}+1}{\sqrt{1+a}-1}\right) - 3\sqrt{1+a}\,\right]. \tag{2.109}$$

According to this result, $\Delta^{(d)}$ grows proportionally to the scale factor at late times, in agreement with a single-fluid Einstein-de Sitter model. Deep into the radiation era, however, $a \ll 1$ and the density contrast grows only logarithmically. This "stagnation", or "freezing-in", of matter perturbations prior to equality is sometimes referred to as the Meszaros effect [39].

7. A CDM and baryon universe

The tight coupling between the photons and baryons in the pre-recombination era, means that purely baryonic scenarios cannot explain the observed structure of the

universe. Dark matter, however, is not subjected to the photon drag because it only interacts gravitationally. There is a long catalogue of dark matter candidates. Thermal relics, which are in thermal equilibrium with the rest of the universe until they decouple, are classified as Hot Dark Matter (HDM) and Cold Dark Matter (CDM) species.[8] Of these, CDM has small dispersion velocities and does not suffer free-streaming dissipation.

Following § V B 6, CDM perturbations grow between equipartition and recombination by a factor of $a_{\rm rec}/a_{\rm eq} = T_{\rm eq}/T_{\rm rec} \simeq 21\Omega h^2$, where $h = H/100\,{\rm km\,s^{-1}\,Mpc^{-1}}$ is the rescaled, dimensionless Hubble parameter. After decoupling the universe becomes transparent to radiation and baryonic perturbations start growing, driven by the gravitational potential of the collisionless species. For a mixture of CDM and dust, the system of (2.106)-(2.107) becomes

$$\dot{\Delta}^{(b)} = -\mathcal{Z} - a{\rm D}^2 v^{(b)}\,, \qquad\qquad \dot{\mathcal{Z}} = -2H\mathcal{Z} - \frac{1}{2}\,\rho\Delta \tag{2.110}$$

and

$$\dot{v}^{(b)} = -Hv^{(b)}, \tag{2.111}$$

where $\rho = \rho^{(c)} + \rho^{(b)}$ is the total density and $\rho\Delta = \rho^{(c)}\Delta^{(c)} + \rho^{(b)}\Delta^{(b)}$. Taking the time derivative of (2.110a), and using the rest of the above given relations leads to

$$\ddot{\Delta}^{(b)} + 2H\dot{\Delta}^{(b)} = \frac{1}{2}\,\rho^{(c)}\Delta^{(c)}\,. \tag{2.112}$$

Suppose that the dark sector dominates the baryonic component (i.e. $\rho^{(b)} \ll \rho^{(c)}$) and that $\Delta^{(b)} \ll \Delta^{(c)}$ just after recombination. Recalling that $\Delta^{(c)} \propto a$ after recombination (see § V B 6) and that $\rho^{(c)} \propto a^{-3}$, we find [32, 35]

$$\Delta^{(b)} = \Delta^{(c)}\left(1 - \frac{a_{\rm rec}}{a}\right)\,. \tag{2.113}$$

The above shows that $\Delta^{(b)} \to \Delta^{(c)}$ when $a \gg a_{\rm rec}$. Put another way, baryonic fluctuations quickly catch up with perturbations in the dark-matter component, after

[8] Non-thermal relics, such as axions, magnetic monopoles and cosmic strings, remain out of equilibrium throughout their lifetime.

falling into the potential wells of the collisionless species. This= result demonstrates how the non-baryonic species can accelerate the collapse of ordinary matter.

Cold relics were first proposed in the early 1980's in order to reproduce the current small-scale structure of the universe (see [41–45] and references therein). Purely CDM models do not seem to agree with observations, but CDM plus dark-energy (the latter as an effective cosmological constant) appear in very good agreement with the data [46].

C. Scalar-field perturbations

1. The effective fluid characteristics

Keeping with the (effective) fluid description of scalar fields, we may introduce the dimensionless parameter

$$w_\varphi = \frac{p^{(\varphi)}}{\rho^{(\varphi)}} = \frac{\dot\varphi^2 - 2V(\varphi)}{\dot\varphi^2 + 2V(\varphi)}\,. \tag{2.114}$$

Demanding that $\rho^{(\varphi)} > 0$ leads to $\dot\varphi^2/2 + V(\varphi) > 0$ and the condition $-1 \leq w_\varphi \leq 1$ follows from $V(\varphi) \geq 0$. Taking the time derivative of (2.114) and using the Klein-Gordon equation – see expression (2.32), we arrive at

$$\dot w_\varphi = -3H(1+w_\varphi)\left(\frac{\dot p^{(\varphi)}}{\dot\rho^{(\varphi)}} - w_\varphi\right)\,. \tag{2.115}$$

Formalistically, the above agrees with the evolution law of the w-parameter in conventional perfect-fluid cosmologies (see Eq. (2.78) in § V B 1). In the scalar-field case, however, the ratio

$$\frac{\dot p^{(\varphi)}}{\dot\rho^{(\varphi)}} = 1 + \frac{2V'(\varphi)}{3H\dot\varphi}\,, \tag{2.116}$$

no longer represents a thermodynamic sound speed.

2. Density perturbations

Aggregations in the effective energy density of a minimally coupled scalar field are described by $\Delta^{(\varphi)} = a\mathrm{D}^a\Delta_a^{(\varphi)}$, which is given by $\Delta^{(\varphi)} = (a^2/\rho^{(\varphi)})\mathrm{D}^2\rho^{(\varphi)} =$

$(a^2\dot{\varphi}/\rho^{(\varphi)})\mathrm{D}^2\dot{\varphi}$ to linear order. Also, the linearised divergences of (2.74) and (2.75) read

$$\dot{\Delta}^{(\varphi)} = 3w^{(\varphi)}H\,\Delta^{(\varphi)} - \left(1+w^{(\varphi)}\right)\mathcal{Z}\,, \tag{2.117}$$

and

$$\dot{\mathcal{Z}} = -2H\mathcal{Z} - \left[\frac{1}{2}\,\rho^{(\varphi)} + \frac{3K}{a^2(1+w_\varphi)}\right]\Delta^{(\varphi)} - \frac{1}{1+w_\varphi}\,\mathrm{D}^2\Delta^{(\varphi)}\,, \tag{2.118}$$

respectively. Following (2.117), Δ decouples from $\mathcal{Z}$ when $w^{(\varphi)} = -1$ (see § V C 3 below). When $w^{(\varphi)} \neq -1$ and $K = 0$, the time derivative of (2.117) leads to the wavelike equation

$$\ddot{\Delta}^{(\varphi)}_{(k)} = -2\left(1 - 3w_\varphi + \frac{3\dot{p}^{(\varphi)}}{2\dot{\rho}^{(\varphi)}}\right)H\dot{\Delta}^{(\varphi)}_{(k)} + \frac{3}{2}\left(1 + 8w_\varphi - 3w_\varphi^2 - 6\frac{\dot{p}^{(\varphi)}}{\dot{\rho}^{(\varphi)}}\right)H^2\Delta^{(\varphi)}_{(k)} - \left(\frac{k}{a}\right)^2\Delta^{(\varphi)}_{(k)}\,, \tag{2.119}$$

for the k-th harmonic mode [11]. It is therefore clear that, to a large extent, the evolution of $\Delta^{(\varphi)}_{(k)}$ is determined by the effective equation of state of the φ-field.

3. Standard slow-roll inflation

Standard inflation corresponds to a slowly rolling scalar field, with $\dot{\varphi}^2 \ll V(\varphi)$. This leads to approximately exponential expansion, where H and $\rho^{(\varphi)}$ are nearly constants and $w^{(\varphi)} \simeq -1$. Then, the linear evolution of $\Delta^{(\varphi)}_{(k)}$ – see Eq. (2.119) – is governed by

$$\ddot{\Delta}^{(\varphi)}_{(k)} = -5H\dot{\Delta}^{(\varphi)}_{(k)} - 6H^2\left[1 + \frac{1}{6}\left(\frac{k}{aH}\right)^2\right]\Delta^{(\varphi)}_{(k)}\,, \tag{2.120}$$

where $H \simeq$ constant. For modes that have crossed outside the Hubble radius (i.e. when $k \ll aH$), the solution is (e.g. see [32])

$$\Delta^{(\varphi)} = \mathcal{C}_1\,\mathrm{e}^{-2Ht} + \mathcal{C}_2\,\mathrm{e}^{-3Ht}\,. \tag{2.121}$$

Hence, during de Sitter-type expansion $\Delta^{(\varphi)} \propto a^{-2}$. As a result, any overdensities (or underdensities) that may exist in the spatial distribution of the inflaton field will decay exponentially.

The above apply to a slowly rolling scalar field, with $w^{(\varphi)} \simeq -1$. When $w^{(\varphi)} = -1$, the kinetic energy of the φ-field vanishes and $p^{(\varphi)} = -\rho^{(\varphi)} = -V(\varphi)$. Then, $\mathrm{D}_a p^{(\varphi)} = -\mathrm{D}_a \rho^{(\varphi)} = -V'(\varphi)\mathrm{D}_a\varphi = 0$. This result, which implies that there are no inhomogeneities in the (effective) energy density and pressure of the φ-field, reflects our spacetime slicing (where $\mathrm{D}_a\varphi = 0$).

D. Gravitational wave perturbations

Gravitational waves are propagating ripples in the spacetime fabric that are usually described as weak metric perturbations. Alternatively, one can use the electric and magnetic parts of the Weyl tensor, which describes the long-range (free) gravitational field (see § II B).

1. Isolating tensor modes

Gravitational waves are (covariantly) described by the transverse components of the electric and the magnetic Weyl tensors. The transversality guarantees that we have isolated the pure tensor modes of the free gravitational field. This condition is also imposed on the shear and on any other orthogonally projected and traceless second-rank tensor that might be present. In perturbed FRW model, the above translate into the linear constrains (see § II B and § II D)

$$\mathrm{D}^b E_{ab} = \frac{1}{3}\,\mathrm{D}_a\rho = 0\,, \qquad \mathrm{D}^b H_{ab} = \rho(1+w)\omega_a = 0 \tag{2.122}$$

and

$$\mathrm{D}^b\sigma_{ab} = \frac{2}{3}\,\mathrm{D}_a\Theta + \mathrm{curl}\,\omega_a = 0\,, \tag{2.123}$$

which must hold all times [47]. This is achieved by switching the vorticity off and by demanding that $\mathrm{D}_a\rho = 0 = \mathrm{D}_a p = \mathrm{D}_a\Theta$ initially (for a barotropic medium it is enough to set $\mathrm{D}_a\rho = 0 = \mathrm{D}_a\Theta$). These constraints are self-consistent (i.e. preserved in time) and also guarantee that the 4-acceleration vanishes as well. Then, the only nontrivial linear constraints left are

$$H_{ab} = \mathrm{curl}\,\sigma_{ab} \qquad \text{and} \qquad \mathcal{R}_{\langle ab\rangle} = -H\sigma_{ab} + E_{ab} \tag{2.124}$$

where $\mathrm{curl}\,\sigma_{ab} = \varepsilon_{cd\langle a}\mathrm{D}^c\sigma_{b\rangle}{}^d$ (see Eqs. (2.18) and (2.34) in § II D and § II G respectively).

2. The gravitational-wave energy density

In a perturbed Friedmann universe, the energy density of gravitational radiation is determined by the pure tensor part ($\mathcal{H}^{TT}_{\alpha\beta}$, with $\alpha, \beta = 1, 2, 3$) of the metric perturbation, according to

$$\rho_{GW} = \frac{(\mathcal{H}^{TT}_{\alpha\beta})'(\mathcal{H}^{\alpha\beta}_{TT})'}{2a^2}\,, \tag{2.125}$$

with primes indicating conformal-time derivatives [48]. Then, in a comoving frame with $u^a = \delta^a_0 u^0$, we have [49, 50]

$$\sigma_{\alpha\beta} = a(\mathcal{H}^{TT}_{\alpha\beta})' \qquad \text{and} \qquad \sigma^{\alpha\beta} = a^{-3}(\mathcal{H}^{\alpha\beta}_{TT})'\,, \tag{2.126}$$

so that [51]

$$\rho_{GW} = \sigma^2\,. \tag{2.127}$$

3. Evolution of gravitational waves

In a FRW spacetime the Weyl tensor vanishes identically, which means that E_{ab} and H_{ab} are gauge-invariant variables. The same also holds for the shear. In fact the high symmetry of the Friedmann universes ensures that gravitational waves can be described solely by the transverse component of the shear. On a spatially flat background the latter obeys the wave equation equation

$$\ddot{\sigma}_{ab} = -5H\dot{\sigma}_{ab} - \frac{1}{2}\,\rho(1-3w)\sigma_{ab} + \mathrm{D}^2\sigma_{ab}\,. \tag{2.128}$$

which has been obtained after taking the pure tensor component of (2.83) and using the linear relation $\mathrm{curl}H_{ab} = -\mathrm{D}^2\sigma_{ab}$ (see [2] for more details). Next, we introduce the tensor harmonics $\mathcal{Q}^{(k)}_{ab}$, where $\mathcal{Q}^{(k)}_{ab} = \mathcal{Q}^{(k)}_{\langle ab\rangle}$, $\dot{\mathcal{Q}}^{(k)}_{ab} = 0 = \mathrm{D}^b\mathcal{Q}^{(k)}_{ab}$ and

$\mathrm{D}^2\mathcal{Q}^{(k)}_{ab} = -(k/a)^2\mathcal{Q}^{(k)}_{ab}$. Then, writing $\sigma_{ab} = \sum_k \sigma_{(k)}\mathcal{Q}^{(k)}_{ab}$, with $\mathrm{D}_a\sigma_{(k)} = 0$, Eq. (2.128) reads

$$\ddot{\sigma}_{(k)} = -5H\dot{\sigma}_{(k)} - \frac{1}{2}\,\rho\left[1 - 3w + \frac{2}{3}\left(\frac{\lambda_H}{\lambda_k}\right)^2\right]\sigma_{(k)}\,, \tag{2.129}$$

where $\lambda_H = 1/H$, $\lambda_k = a/k$ and $k \geq 0$.

During the radiation era $w = 1/3$, $a \propto t^{1/2}$, $H = 1/(2t)$ and $\rho = 3/(4t^2)$, in which case $\lambda_H/\lambda_k = k/aH \propto t^{1/2}$ and expression (2.129) simplifies to

$$\ddot{\sigma}_{(k)} = -\frac{5}{2t}\,\dot{\sigma}_{(k)} - \frac{1}{4t^2}\,\left(\frac{\lambda_H}{\lambda_k}\right)^2\sigma_{(k)}\,. \tag{2.130}$$

The above solves to give

$$\sigma_{(k)} = t^{-1}\left[\mathcal{C}_1\,\sin\left(\frac{\lambda_H}{\lambda_k}\right) + \mathcal{C}_2\,\cos\left(\frac{\lambda_H}{\lambda_k}\right)\right] + t^{-3/2}\left[\mathcal{C}_3\,\cos\left(\frac{\lambda_H}{\lambda_k}\right) - \mathcal{C}_4\,\sin\left(\frac{\lambda_H}{\lambda_k}\right)\right]. \tag{2.131}$$

Accordingly, the amplitude of small-scale gravitational waves decays as a^{-2}. On super-Hubble scales, on the other hand, $\lambda_H/\lambda_k \ll 1$ and (2.131) takes the power-law form

$$\sigma_{(k)} = \sigma_0 + \frac{2}{3}\,\dot{\sigma}_0 t_0\left[1 - \left(\frac{t}{t_0}\right)^{3/2}\right], \tag{2.132}$$

suggesting that $\sigma_{(k)}$ = constant beyond the horizon. After equipartition $w = 0$, $a \propto t^{2/3}$, $H = 2/(3t)$, $\rho = 4/(3t^2)$ and Eq. (2.129) accepts the large-scale solution

$$\sigma_{(k)} = \mathcal{C}_1\,t^{-1/3} + \mathcal{C}_2\,t^{-2}\,. \tag{2.133}$$

Hence, after equality, superhorizon-sized gravitational wave perturbations decay as $a^{-1/2}$.

VI. STRUCTURE FORMATION AND COSMIC ACCELERATION

The supernovae observations around the turn of the millennium have changed the course of modern cosmology, leading the majority of the scientific community to embrace the idea of an accelerating universe. Dark energy has so far been the most popular explanation of such universal acceleration. There are other schools of thought, however, one of which suggests that the large-scale structure of our universe today may hold the answer to the supernovae data.

A. The recent universal acceleration

Just before the end of the last century, observations of distant supernovae, of type Ia, provided modern cosmology with one of its greatest puzzles. According to the data, the aforementioned stellar objects, which are treated as "standard candles" and are therefore used for measuring distances on cosmological scales, were dimmer than expected. A number of solutions were proposed to explain the data. These included conventional answers, such as absorption effects, or differences in the chemical composition of these high redshift objects. One by one, these explanations were dismissed however. Then, the only answer remaining on the table was that the observed supernovae dimming was caused by the expansion rate of the universe. Our cosmos was expanding at an accelerated pace. Moreover, the acceleration appeared to be a relatively recent event that started approximately two billion years ago.

There are a number of assumptions along the road leading from the supernovae observations to the idea of universal acceleration. The most crucial perhaps is that we live in an FRW universe. In that case, the luminosity distance (D_L) of a source at a given redshift (z), is related to the Hubble and the deceleration parameters of the universe via the expression

$$D_L = (1+z)H_0^{-1}\int_0^z \mathrm{e}^{-\int_0^x (1+q)\mathrm{d}[\ln(1+y)]}\mathrm{d}x\,. \tag{2.134}$$

When tested against the measured luminosity distance of remote type Ia supernovae, the above relation repeatedly returned negative values for the deceleration parameter, indicating a universe that had recently entered a phase of accelerating expansion [52, 53]. It appears, in particular, that the universal acceleration started at around $z \simeq 0.5$ and that the value of the deceleration parameter is $q \simeq -0.5$ on average.

B. The dark energy paradigm

Adopting the idea of universal acceleration is not the end of the story however. As it usually happens in these occasions, the answer gives rise to a new question, which in our case is what causes the acceleration. To find out, one needs to look

at the Raychaudhuri equation. Assuming that on cosmological scales the universe is described by an FRW model, Raychaudhuri's formula (see Eq. (2.135) in § III B) takes the simple form

$$qH^2 = \frac{1}{6}\,(\rho + 3p) - \frac{1}{3}\,\Lambda\,, \tag{2.135}$$

with q representing the deceleration parameter. The latter can take negative values either when a (positive) cosmological constant dominates the dynamics of the universe, or when the the dominant matter species have negative effective gravitational energy, namely for $\rho+3p<0$. When this theoretical result was combined to additional (largely independent) observational data, it lead to the current concordance cosmological model. According to the later, Dark Energy (in the form of a cosmological constant (Λ), or of a dynamical field with $\rho+3p<0$) makes 70% of the matter in the universe, non-baryonic CDM makes another 25% and only the remaining 5% are ordinary baryons. These percentages have been broadly confirmed by the latest Planck results, which however have slightly decreased the contribution of Dark Energy, in favour of CDM, and did not find evidence of dynamical Dark Energy [1].

The ΛCDM paradigm has been very successful in explaining a host of largely independent cosmological data, like the CMB angular anisotropies, the spatial galactic correlations measured in large-scale galaxy surveys and the above mentioned dimming of the supernovae-Ia luminosities. All these with the help of a relatively small number of free parameters (e.g. see [54, 55]). There is still scepticism, however, both on theoretical and observational grounds. Theoretically, the main issue is the mysterious nature of Dark Energy (e.g. see [56] for a comprehensive discussion). Observationally, as well, there have been a number of puzzling data that do not appear to fit with the ΛCDM scenario (e.g. see [57, 58]). As a result, a number of alternatives have appeared in the literature. Given that the supenovae observations are at the core of the ΛCDM paradigm, most of the aforementioned scenarios attempt to explain this data without involving Dark Energy. Some do so by abandoning General Relativity, others by accounting for the effects of the observed inhomogeneity and anisotropy of the universe. Next, we will discuss one of these alternative scenarios in more detail.

C. The backreaction scenario

According to the observations, our universe started to accelerate around two billion years ago, which is roughly the time galaxy formation moved from the linear into the nonlinear phase. This coincidence has led a number of cosmologists to suggest that cosmic acceleration could be the result of structure formation, reflecting the fact that our universe is not an FRW model, at least on certain scales. These ideas are behind the so-called 'kinematic backreaction' scenarios.

1. Spatial averaging

The problem of averaging general-relativistic spacetimes, which emerges from the generic non-linearity of the theory, is a long-standing one [59, 60]. In fact, the ongoing ambiguity has lead a number of cosmologists to argue that the averaging problem may be crucial in our understanding of the recent expansion of the visible universe. The literature contains more than one averaging methods that have been applied to cosmology [61–65]. The simplest and most straightforward approach is probably that of Buchert, which confines to scalar variables only. Following [63, 64], the spatial average of a scalar field $\phi = \phi(x^a)$ over a simply connected domain D is given by

$$\langle\phi\rangle_D = V_D^{-1}\int_D \phi\mathcal{H}\mathrm{d}^3x^a\,, \tag{2.136}$$

where V_D is the volume of the averaging domain and $\mathcal{H} = \sqrt{\det(h_{ab})}$. The key point, resulting from the time-dependence of the volume, is that spatial averaging and time evolution do not commute. Mathematically, the non-commutativity between these two operations is formulated in the simple rule

$$\langle\phi\rangle^{\cdot}_D - \langle\dot{\phi}\rangle_D = \langle\Theta\phi\rangle_D - \langle\Theta\rangle_D\langle\phi\rangle_D\,, \tag{2.137}$$

with Θ representing the local volume expansion rate.

2. The averaged equations

The aim is to write the averaged equations in the form of their local (non-averaged) counterparts and then compare the two sets. Any differences will then be treated as backreaction effects reflecting the averaged action of inhomogeneity and anisotropy. Here, we will write the averaged formulae in a Friedmann-type form. Assuming irrotational dust and applying rule (2.137), the continuity, the Friedmann and the Raychaudhuri equations average to

$$\langle\rho\rangle^{\cdot} + \langle\Theta\rangle\langle\rho\rangle = 0\,, \tag{2.138}$$

$$\langle\Theta\rangle^{\cdot} + \frac{1}{3}\langle\Theta\rangle^2 + \frac{1}{2}\langle\rho\rangle - \Lambda = \frac{2}{3}\langle(\Theta - \langle\Theta\rangle)^2\rangle - 2\langle\sigma^2\rangle\,, \tag{2.139}$$

$$\frac{1}{2}\langle\mathcal{R}\rangle - \langle\rho\rangle + \frac{1}{3}\langle\Theta\rangle^2 - \Lambda = -\frac{1}{3}\langle(\Theta - \langle\Theta\rangle)^2\rangle + \langle\sigma^2\rangle\,. \tag{2.140}$$

The above provide a (non-closed) set of equations monitoring the spatially averaged scalars in non-rotating, inhomogeneous universes that contain pressure-free matter [66, 67]. It should be noted that Eqs. (2.139), (2.140) are exact and there is no need to assume that the inhomogeneity and the anisotropy are small perturbations. Crucially, the averaging process has added extra terms to the right-hand side of (2.139) and (2.140), the effect of which is encoded in the domain-dependent scalar

$$\mathcal{Q}_D = \frac{2}{3}\langle(\Theta - \langle\Theta\rangle)^2\rangle - 2\langle\sigma^2\rangle = -\frac{2}{3}\left(\langle\Theta\rangle^2 - \langle\Theta^2\rangle\right) - 2\langle\sigma^2\rangle\,. \tag{2.141}$$

This quantity, which in principle can be either positive or negative, is interpreted as the kinematic backreaction of spatial averaging upon pressure-free FRW models.

3. Acceleration from backreaction

Following (2.139) and (2.141), in the absence of a cosmological constant, the backreaction effects will force a domain of the averaged universe to accelerated expansion when

$$\mathcal{Q}_D > \frac{1}{2}\langle\rho\rangle\,. \tag{2.142}$$

This means that $\mathcal{Q}_D$ needs to be positive, which implies that shear fluctuations must be superseded by those in the volume expansion (see (2.141)). In the opposite case, $\mathcal{Q}_D$ will be negative and the domain will decelerate further instead of accelerating. [9] Condition (2.142) also implies that, if backreaction is to work, $\mathcal{Q}_D$ should decay slower than the average density. [10]

We may not have clear information on the sign of $\mathcal{Q}_D$, but we can monitor its dynamical evolution by means of a consistency/integrability condition. In particular, taking the time derivative of (2.140) and then using (2.138) and (2.139) one arrives at

$$\dot{\mathcal{Q}}_D + 2\langle\Theta\rangle\mathcal{Q}_D = -\langle\mathcal{R}\rangle^{\cdot} - \frac{2}{3}\langle\Theta\rangle\langle\mathcal{R}\rangle\,. \tag{2.143}$$

The above means that if $\langle\mathcal{R}\rangle \propto a_D^{-2}$, namely if the averaged spatial curvature behaves like its FRW counterpart, the backreaction term will scale as $\mathcal{Q}_D \propto a_D^{-6}$ mimicking a "stiff" fluid. Clearly, the same evolution law holds when $\langle\mathcal{R}\rangle = 0$ as well. In either case, the density term in (2.139) will quickly dominate the backreaction effects, even when condition (2.142) is initially satisfied. Recall that $\langle\rho\rangle \propto a_D^{-3}$, just like in the standard Einstein-de Sitter model. Therefore, a change in the scale-factor dependance of the average 3-Ricci scalar seems necessary for the backreaction idea to work [70].

VII. DISCUSSION

Structure formation is a long and slow process. The key is the Jeans instability, a mechanism that enhances gravitationally small inhomogeneities in an otherwise smooth matter distribution. Such tiny distortions have been detected in the CMB spectrum, which shows inhomogeneities of the order of 10^{-5} in the last-scattering surface. These are believed to be the seeds of all the structure that we see in the universe today. The details of the process, however, are still missing. Nevertheless, it is almost certain that the aforementioned small perturbations had been growing slowly and independently for long, until they became strong enough to start affecting each other's evolution.

[9] The sign of the backreaction term has yet to be decided. In fact, because of this sign-ambiguity, the same kinematic backreaction term, which is now proposed as a conventional solution to the dark-energy problem, was earlier suggested as an effective dark-matter source [68].

[10] It is conceivable that one can arrive to the same qualitative result, namely that backreaction effects can lead to accelerated expansion, through a perturbative approach as well (e.g. see [69]).

This period of slow and independent growth, which is known as the linear regime of structure formation, has been the focus of the present lecture notes.

When studying cosmological perturbations, one needs first to select a smooth background model. This is generally described by an analytical solution of the Einstein field equations and represents an idealistic (fictitious) universe. In the majority of the studies the unperturbed model coincides with the spatially flat FRW spacetime. By perturbing away from the chosen background, one obtains a more realistic cosmological model, which will hopefully describe the observable universe with sufficient accuracy. There are, however, more than one ways of (mathematically) connecting the unperturbed and the perturbed spacetimes, which is the reason behind the so-called gauge problem. The latter, which could lead to fictitious (gauge-dependent) modes in the solutions of the equations, has long been known to plague cosmological perturbation theory. Perhaps the best way of dealing with the gauge problem is by introducing gauge-invariant variables. Describing, in other words, the perturbations with quantities that are independent of the gauge choice. This is also what we did here, by exploiting the Stewart & Walker lemma.

Once the background has been set, the next step is to derive the first-order (the linear) equations that monitor the evolution of the key gauge-invariant variables. Most of the studies do that by perturbing away from the zero-order (the background) expressions. Here, instead, we have linearised the fully nonlinear equations around the chosen background model. We considered all three types of distortions, namely scalar, vector and tensor, and looked into the evolution of density perturbations (scalar), density vorticies (vector) and gravitational waves (tensor). Assuming a single fluid cosmology, we find that conventional (i.e. non-inflationary) FRW model are unstable to linear density perturbations. The latter grow with time, at least on large enough scales, where pressure support and dissipative effects are negligible. Vortices and gravitational waves, on the other hand hand, decay. The growth rate of density perturbations, however, is too slow. As a result, it is essentially impossible to explain the observed structure of the universe with purely baryonic matter. The standard solution is to introduce collisionless CDM, which can start clumping long before recombination. Then, it is fairly straightforward to show that aggregations in

the distribution of the collisionless species will accelerate the gravitational collapse of the baryonic matter after recombination. In fact, the heavier the CDM particles, the "deeper" their "potential wells" and the faster the gravitational clumping of the baryons. This is, in broad lines, the standard picture of the linear phase. Once the perturbations have grown strong, however, they no longer evolve independently and the linear approximation breaks down. This marks the beginning of the nonlinear regime.

Addressing the nonlinear phase of structure formation analytically is an essentially impossible task, due to the extreme complexity of the full problem. One then resorts to numerical methods, or employs physically motivated simplifications, like the Zeldovich approximation for example. Another option is to use averaging techniques that will encode the overall effect of inhomogeneity and anisotropy into a single "backreaction" term. Averaging means information loss, however, at least to some extent. Moreover, the intrinsic nonlinearity of general relativity makes spatial averaging a formidable task. As a result, with the exception of scalar quantities, there has been no consensus so far on how to average relativistic spacetimes. Scalar averaging methods were recently employed to explain the accelerated expansion of the universe, as the result of the increasing structure of the universe and without appealing to Dark Energy (or to any kind of new physics). These are the so-called backreaction scenarios. There has been criticism, however, mainly coming from the anticipated weakness of the backreaction effects on cosmological scales. To make things worse, the proposed models have so far failed to measure the strength of their backreaction term (even its sign remains undetermined). Nevertheless, backreaction scenarios have certain attractive aspects, such as their relative simplicity and the fact that they solve the "coincidence problem" naturally. In fact, answering the coincidence question (which within the ΛCDM framework requires an uncomfortable fine tuning) effortlessly, may prove decisive for the backrecation scenario. The future should show.

---

[1] P. A. R. Ade, N. Aghanim, C. Armitage-Caplan, M. Arnaud, M. Ashdown, F. Atrio-Barandela, J. Aumont, C. Baccigalupi, A. J. Banday, and et al. Planck 2013 results. XVI.

Cosmological parameters. *ArXiv e-prints*, March 2013.

[2] C. G. Tsagas, A. Challinor, and R. Maartens. Relativistic cosmology and large-scale structure. Phys. Rep. , 465:61–147, August 2008.

[3] S. W. Hawking and G. F. R. Ellis. *The Large Scale Structure of Space-Time*. Cambridge: Cambridge University Press, 1973.

[4] R. Maartens. Linearization instability of gravity waves? *Phys. Rev. D*, 55:463–467, January 1997.

[5] M. S. Madsen. A Note on the Equation of State of a Scalar Field. Astrophys. Space Sci. , 113:205–+, July 1985.

[6] M. S. Madsen. Scalar fields in curved spacetimes. Class. Quant. Grav. , 5:627–639, November 1988.

[7] G. F. R. Ellis and M. S. Madsen. Exact scalar field cosmologies . Class. Quant. Grav. , 8:667–676, April 1991.

[8] D. Langlois and F. Vernizzi. Non-linear perturbations of cosmological scalar fields. JCAP , 2:17–+, February 2007.

[9] F. Vernizzi. Conservation of second-order cosmological perturbations in a scalar field dominated universe. *Phys. Rev. D*, 71(6):061301–+, March 2005.

[10] P. S. Letelier. Anisotropic fluids with two-perfect-fluid components. *Phys. Rev. D*, 22:807–813, August 1980.

[11] M. Bruni, G. F. R. Ellis, and P. K. S. Dunsby. Gauge-invariant perturbations in a scalar field dominated universe . Class. Quant. Grav. , 9:921–945, April 1992.

[12] R. M. Wald. *General Relativity*. Chicago: University of Chicago Press, 1984.

[13] E. Poisson. *A Relativist's Toolkit*. Cambridge: Cambridge University Press, 2004.

[14] G. F. R. Ellis and M. A. H. MacCallum. A class of homogeneous cosmological models. *Commun. Math. Phys.* , 12:108–141, June 1969.

[15] J. Wainwright and G. F. R. Ellis. *Dynamical Systems in Cosmology*. Cambridge: Cambridge University Press, 1997.

[16] L. P. Hughston and K. C. Jacobs. Homogeneous Electromagnetic and Massive-Vector Fields in Bianchi Cosmologies. *Astrophys. J.*, 160:147–+, April 1970.

[17] V. A. Belinskij, I. M. Khalatnikov, and E. M. Lifshits. Oscillatory approach to a singular point in the relativistic cosmology. *Adv. Phys.*, 19:525–573, 1970.

[18] L. D. Landau and E. M. Lifshitz. *The Classical Theory of Fields*. Oxford: Pergamon Press, 1975.

[19] H. Stephani, D. Kramer, M. MacCallum, C. Hoenselaers, and E. Herlt. *Exact solutions of Einstein's field equations*. Cambridge: Cambridge University Press, 2003.

[20] A. A. Coley. *Dynamical Systems and Cosmology*. Berlin: Springer, 2003.

[21] C. B. Collins and S. W. Hawking. Why is the Universe Isotropic? *Astrophys. J.*, 180:317–334, March 1973.

[22] E. M. Lifshitz. . *J. Phys. (Moscow)* , 10:116, 1946.

[23] E. M. Lifshitz and I. M. Khalatnikov. Investigations in relativistic cosmology. *Adv. Phys.* , 12:185–249, April 1963.

[24] R. K. Sachs and A. M. Wolfe. Perturbations of a Cosmological Model and Angular Variations of the Microwave Background. *Astrophys. J.*, 147:73–+, January 1967.

[25] J. M. Bardeen. Gauge-invariant cosmological perturbations. *Phys. Rev. D*, 22:1882–1905, October 1980.

[26] G. F. R. Ellis and M. Bruni. Covariant and gauge-invariant approach to cosmological density fluctuations. *Phys. Rev. D*, 40:1804–1818, September 1989.

[27] J. M. Stewart and M. Walker. Perturbations of space-times in general relativity. *Proc. R. Soc. A* , 341:49–74, October 1974.

[28] J. M. Stewart. Perturbations of Friedmann-Robertson-Walker cosmological models. Class. Quant. Grav. , 7:1169–1180, July 1990.

[29] M. Bruni, S. Matarrese, S. Mollerach, and S. Sonego. Perturbations of spacetime: gauge transformations and gauge invariance at second order and beyond. Class. Quant. Grav. , 14:2585–2606, September 1997.

[30] G. F. R. Ellis, J. Hwang, and M. Bruni. Covariant and gauge-independent perfect-fluid Robertson-Walker perturbations. *Phys. Rev. D*, 40:1819–1826, September 1989.

[31] G. F. R. Ellis, M. Bruni, and J. Hwang. Density-gradient-vorticity relation in perfect-fluid Robertson-Walker perturbations. *Phys. Rev. D*, 42:1035–1046, August 1990.

[32] T. Padmanabhan. *Structure Formation in the Universe*. Cambridge: Cambridge University Press, June 1993.

[33] P. J. E. Peebles. *The Large-Scale Structure of the Universe*. Princeton: Princeton University Press, 1980.

[34] P. Coles and F. Lucchin. *Cosmology: The Origin and Evolution of Cosmic Structure*. Chichester: Wiley, 1995.

[35] C. G. Tsagas. Cosmological Perturbations. In S. Cotsakis and E. Papantonopoulos, editors, *in Cosmological Crossroads*, pages 223–261, 2002.

[36] J. Silk. Fluctuations in the Primordial Fireball. *Nature (London)*, 215:1155–+, September 1967.

[37] J. Silk. Cosmic Black-Body Radiation and Galaxy Formation. *Astrophys. J.*, 151:459–+, February 1968.

[38] P. K. S. Dunsby, M. Bruni, and G. F. R. Ellis. Covariant perturbations in a multifluid cosmological medium. *Astrophys. J.*, 395:54–74, August 1992.

[39] P. Meszaros. The behaviour of point masses in an expanding cosmological substratum. Astron. & Astrophys. , 37:225–228, December 1974.

[40] E. J. Groth and P. J. E. Peebles. Closed-form solutions for the evolution of density perturbations in some cosmological models. Astron. & Astrophys. , 41:143–145, June 1975.

[41] P. J. E. Peebles. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. Astrophys. J. Lett. , 263:L1–L5, December 1982.

[42] G. R. Blumenthal, S. M. Faber, J. R. Primack, and M. J. Rees. Formation of galaxies and large-scale structure with cold dark matter. *Nature (London)*, 311:517–525, October 1984.

[43] M. Davis, G. Efstathiou, C. S. Frenk, and S. D. M. White. The end of cold dark matter? *Nature (London)*, 356:489–494, April 1992.

[44] A. R. Liddle and D. H. Lyth. The cold dark matter density perturbation. Phys. Rep. , 231:1–2, August 1993.

[45] M. White, D. Scott, J. Silk, and M. Davis. Cold dark matter resuscitated? Mon. Not. R. Astron. Soc. , 276:L69–L75, October 1995.

[46] D. N. Spergel, L. Verde, H. V. Peiris, E. Komatsu, M. R. Nolta, C. L. Bennett, M. Halpern, G. Hinshaw, N. Jarosik, A. Kogut, M. Limon, S. S. Meyer, L. Page, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. First-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Determination of Cosmological Parameters. Astrophys. J. Supp. , 148:175–194, September 2003.

[47] P. K. S. Dunsby, B. A. C. C. Bassett, and G. F. R. Ellis. Covariant analysis of gravitational waves in a cosmological context. Class. Quant. Grav. , 14:1215–1222, May 1997.

[48] C. Caprini and R. Durrer. Gravitational wave production: A strong constraint on primordial magnetic fields. *Phys. Rev. D*, 65(2):023517–+, January 2002.

[49] M. Bruni, P. K. S. Dunsby, and G. F. R. Ellis. Cosmological perturbations and the physical meaning of gauge-invariant variables. *Astrophys. J.*, 395:34–53, August 1992.

[50] S. W. Goode. Analysis of spatially inhomogeneous perturbations of the FRW cosmologies. *Phys. Rev. D*, 39:2882–2892, May 1989.

[51] C. G. Tsagas. Gravitational waves and cosmic magnetism: a cosmological approach . Class. Quant. Grav. , 19:3709–3722, July 2002.

[52] M. S. Turner and A. G. Riess. Do Type Ia Supernovae Provide Direct Evidence for Past Deceleration of the Universe? *Astrophys. J.*, 569:18–22, April 2002.

[53] A. G. Riess, L.-G. Strolger, J. Tonry, S. Casertano, H. C. Ferguson, B. Mobasher, P. Challis, A. V. Filippenko, S. Jha, W. Li, R. Chornock, R. P. Kirshner, B. Leibundgut, M. Dickinson, M. Livio, M. Giavalisco, C. C. Steidel, T. Benítez, and Z. Tsvetanov. Type Ia Supernova Discoveries at z $\geq$ 1 from the Hubble Space Telescope: Evidence for Past Deceleration and Constraints on Dark Energy Evolution. *Astrophys. J.*, 607:665–687, June 2004.

[54] D. N. Spergel, R. Bean, O. Doré, M. R. Nolta, C. L. Bennett, J. Dunkley, G. Hinshaw, N. Jarosik, E. Komatsu, L. Page, H. V. Peiris, L. Verde, M. Halpern, R. S. Hill, A. Kogut, M. Limon, S. S. Meyer, N. Odegard, G. S. Tucker, J. L. Weiland, E. Wollack, and E. L. Wright. Three-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Implications for Cosmology. Astrophys. J. Supp. , 170:377–408, June 2007.

[55] W. J. Percival, R. C. Nichol, D. J. Eisenstein, J. A. Frieman, M. Fukugita, J. Loveday, A. C. Pope, D. P. Schneider, A. S. Szalay, M. Tegmark, M. S. Vogeley, D. H. Weinberg, I. Zehavi, N. A. Bahcall, J. Brinkmann, A. J. Connolly, and A. Meiksin. The Shape of the Sloan Digital Sky Survey Data Release 5 Galaxy Power Spectrum. *Astrophys. J.*, 657:645–663, March 2007.

[56] S. Sarkar. Is the evidence for dark energy secure? *General Relativity and Gravitation*, 40:269–284, February 2008.

[57] L. Perivolaropoulos. Six Puzzles for LCDM Cosmology. *ArXiv e-prints*, November 2008.

[58] L. Perivolaropoulos. LCDM: Triumphs, Puzzles and Remedies. *ArXiv e-prints*, April 2011.

[59] M. F. Shirokov and I. Z. Fisher. Isotropic Space with Discrete Gravitational-Field Sources. On the Theory of a Nonhomogeneous Isotropic Universe. *Sov. Astron.* , 6:699–+, April 1963.

[60] G. F. R. Ellis. Relativistic cosmology - Its nature, aims and problems. In B. Bertotti, F. de Felice, and A. Pascolini, editors, *in General Relativity and Gravitation Conference*, pages 215–288, 1984.

[61] R. A. Isaacson. Gravitational Radiation in the Limit of High Frequency. I. The Linear Approximation and Geometrical Optics. *Phys. Rev.* , 166:1263–1271, February 1968.

[62] R. M. Zalaletdinov. Averaging problem in general relativity, macroscopic gravity and using Einstein's equations in cosmology. *Bull. Astron. Soc. India* , 25:401–+, September 1997.

[63] T. Buchert and J. Ehlers. Averaging inhomogeneous Newtonian cosmologies. Astron. & Astrophys. , 320:1–7, April 1997.

[64] T. Buchert. On Average Properties of Inhomogeneous Fluids in General Relativity: Dust

Cosmologies. Gen. Rel. Grav. , 32:105–126, January 2000.

[65] T. Buchert. On Average Properties of Inhomogeneous Fluids in General Relativity: Perfect Fluid Cosmologies. Gen. Rel. Grav. , 33:1381–1405, August 2001.

[66] S. Räsänen. Dark energy from back-reaction. JCAP , 2:3–+, February 2004.

[67] S. Räsänen. Constraints on backreaction in dust universes. Class. Quant. Grav. , 23:1823–1835, March 2006.

[68] T. Buchert. Averaging hypotheses in Newtonian Cosmology. In P. Coles, V. Martinez, and M.-J. Pons-Borderia, editors, *in Mapping, Measuring, and Modelling the Universe*, ASP Conference Series Vol 94, pages 349–355, 1996.

[69] E. W. Kolb, S. Matarrese, A. Notari, and A. Riotto. Primordial inflation explains why the universe is accelerating today. *ArXiv High Energy Physics - Theory e-prints*, March 2005.

[70] T. Buchert. LETTER TO THE EDITOR: A cosmic equation of state for the inhomogeneous universe: can a global far-from-equilibrium state explain dark energy? Class. Quant. Grav. , 22:L113–L119, October 2005.

POST–LINEAR FLUCTUATION EVOLUTION
Analytical treatments of non–linear cosmological stages

Silvio A. Bonometto[1]

[1] *Physics Department, Astrophysical Unit,*
Trieste University, Via Tiepolo 11, 34143 Trieste (Italy)
I.N.F.N. - Sezione di Trieste
I.N.A.F. - Oss. astronomico di Trieste
C.I.F.S., Villa Gualino, Torino (Italy)

I. INTRODUCTION

Modern cosmology tells us that slight density discrepancies among different sites, on each scale attained by the horizon size, were enough to create the observed multiform cosmic morphologies. In order to agree with the observed anisotropies of the CMB, such discrepancies must be $\mathcal{O}(1:10^5)$, and this is also what structure formation requires.

Today, such horizons are included in a much larger horizon and we perceive them as primeval density fluctuations. The standard and simplest assumption is that they were adiabatic and Gaussian. This is a generic inflationary prediction, but one can devise *ad–hoc* inflationary models, then necessarily including several scalar fields, to violate either or both predictions.

The entry in the horizon obviously depends on scale. Scales entering the horizon when the Universe is still ionized undergo rather complex linear stages. Available linear codes are very efficient in computing them, and CMB anisotropy and polarization predictions can be directly compared with data. This is a direct proof that these stages are understood. Linear codes then provide *transfer functions* telling us how the material components evolve.

In the linear regime it is worth expanding the fluctuation field into Fourier components. Owing to linearity, each component evolves separately and final results can the be re–summed up, if needed. As soon as non–linear effects start to matter, this is no longer true.

By using Fourier components one also defines fluctuation spectra and we must clearly distinguish between the primordial and the transfered spectrum. Non linearity effects then act on the transfered spectrum and furtherly modify it. There exist approximated expressions for non–linear spectra [1], which have also been inserted into linear codes. However, only for a restricted range of models they are reasonable, i.e., involve errors $\mathcal{O}(\pm 5\,\%)$. If tentatively applied to other cosmologies, results risk to be badly misleading [2].

Even more significantly, even if fluctuations were initially Gaussian, the onset of non–linearity turns them into non–Gaussian. Of course, while non linear fluctuations $+\Delta\rho > \bar{\rho}$ are possible and allowed, fluctuations $-\Delta\rho$ with modulus $> \bar{\rho}$ are obviously impossible; but this is not the only cause of the non–Gaussian behavior.

In a non–Gaussian regime, the spectrum of density fluctuations (or, equivalently, the 2–point correlation function) no longer accounts for all the properties of inhomogeneities. N–point correlation function switch on and fluctuation statistics becomes complex. Worse than so: it becomes increasingly hard to deduce high–order functions from data, and a description of the matter distribution can hardly be based on them or the related (N-1)–spectra.

As a matter of fact, in the non–linear regime, other quantities become more significant, namely for being more easily related to data. Among them, mass functions have a peculiar role, and we shall devote a particular attention to them, also providing tools allowing reasonable predictions.

In the whole of these discussion we shall make use of a background metric reading

$$ds^2 = c^2 dt^2 - a^2(t) d\ell^2 \tag{3.1}$$

($a(t)$: scale factor, $d\ell$: comoving spatial element with 3 cartesian components x_α, t : time coordinate). Deviation from this metric due to density inhomogeneity will be studied in a synchronous gauge, so that the eq. (3.1) still represent the cosmic metric, and t is univocally the *cosmic time*. Accordingly, it will however be

$$d\ell^2 = (\gamma_{\alpha\beta} + h_{\alpha\beta}) dx^\alpha dx^\beta \quad (\alpha, \beta = 1, .., 3)$$

just with $h_{\alpha\beta}$ vanishing when fluctuations are disregarded (with cartesian x^α, it is $\gamma_{\alpha\beta} = \delta_{\alpha\beta}$).

Simulations have accustomed us to a picture of the Universe, where galaxy clusters occupy the knots (dubbed halos, in simulations) and a foam made of thin sheets and slightly denser bridges connects them. One of the most intriguing results concerns the profile of the halos. N–body simulations allowed us to find that a profile

$$\rho(r) = \frac{\rho_c}{\frac{r}{r_c}\left(1 - \frac{r}{r_c}\right)^2} \tag{3.2}$$

is a general feature. Once the halo mass is known, the radius r_c where the profile changes slope is the only parameter defining it. We call *halo concentration* the ratio $c = R_V/r_c$ (see below for the definition of the *virial radius* R_V; essentially, however, R_V approaches the cluster size). This profile, dubbed NFW (Navarro, Frenk & White), is hard to explain non–numerically and some approaches aiming to deduce it exhibit a complexity not paying for the scarse generality of results.

A NFW profile should however hold for any gravitationally stable system: for galaxy clusters as well as for systems collapsed inside them, as the galaxies themselves. Data however have been accumulating, showing that the central cusp predicted in eq. (3.2) is hard to reconcile with galaxy data.

Let us now describe the specific aims and the plan of these lectures. The aims include a study of the Press & Schechter theory, and its generalization, as well as a study of the so–called Zel'dovich approximation.

We start, in the next Section, from a presentation of the cosmological models, either phenomenologically significant, or providing a useful approach.

In Section 3, we then pick up the fluctuations, still in their linear stages, as they were soon after the Universe became neutral. We then follow their gradual evolution towards the non–linear regime. There are precise numerical codes doing so, but we are more interested in still unprecise analytical treatments, allowing us to focus on the physical aspects.

In Section 4 we discuss a classical problem in non–linear cosmology: the evolution of a spherical density enhancement. Under suitable conditions concerning the boundary velocities, this problem has analytical solutions, at least in suitable cosmologies.

The results on the spherical growth are the basis to work out Press & Schecter (PS)

expressions. This will be done in Section 5 & 6 . In Section 7 we discuss then Sheth & Tormen and Jenkins extensions of PS expressions.

Finally, sections 8 & 9 are devoted to the Zel'dovich approximation.

II. PRIMORDIAL AND TRANSFERED SPECTRUM

We describe the primeval discrepancies from homogeneity by means of the *primordial spectrum* of density fluctuations

$$P(t_i, k) = \langle |\delta(t_i, k)|^2 \rangle \ , \tag{3.3}$$

at a suitable time t_i . Here $\delta(k)$ is the Fourier transform of the density fluctuation field

$$\epsilon(\mathbf{x}) = \rho(\mathbf{x})/\bar{\rho} - 1 \tag{3.4}$$

($\rho(\mathbf{x})$: cosmic density in the site $\mathbf{x}$, $\bar{\rho}$: average cosmic density). More specifically:

$$\delta(k) = (2\pi)^{-3} \int d^3x \ e^{i\mathbf{k}\cdot\mathbf{x}} \epsilon(\mathbf{x}) \ , \tag{3.5}$$

so that the physical wavelength of each Fourier component reads $\lambda = 2\pi a/k$.

We can also associate the mass scale

$$M = \frac{4\pi}{3} \rho \lambda^3 \tag{3.6}$$

to the scale λ. Here ρ is the density of a non–relativistic component (dark matter and/or baryons), therefore evolving so that $\rho \propto a^{-3}$. As $\lambda \propto a$, the mass M – as well as the wave–number k – are redshift independent. By referring to a specific k value, we then refer to a given mass scale.

In principle, the average in eq. (3.3) is made among an ensamble of universes; but, of course, we average through space domains, much greater than the scale $L = 2\pi/k$. This clearly implies problem when the scale L approaches today's cosmic horizon. Notice also that $P(k)$ is supposed to depend on the modulus of the vector $\mathbf{k}$. This means that, although inhomogeneities make each cosmic site anisotropic, isotropy is reattained once we average among a large set of them.

The spectrum $P(k)$ evolves with time, at least because of the action of gravity, both before and after the entry of the scale $L = 2\pi/k$ in the particle horizon. Let us then choose the *initial* time t_i when all scales relevant for today's large scale structure have not yet been reached by the horizon size, and assume that, at that time, the spectrum reads

$$P(t_i, k) = A(t_i)k^{n_s} \quad : \tag{3.7}$$

a power law aims to avoid any priviledged scale. If the *spectral index* $n_s = 1$, the spectrum is said *Harrison–Zel'dovich*'s. It can be then shown that, for $n_s = 1$, the average fluctuation amplitude is equal over all scales, in the instant when they are reached by the horizon size (see Appendix A1). Such amplitude is then the only figure needed to define the spectrum.

According to observations $n_s \simeq 0.96$; $n_s = 1$ is also "excluded" at 2–σ's, unless we consider peculiar cosmological models (Komatsu et al., 2010). Let us add that 0.96 is consistent with the inflationary paradigm and with the possibility that inflation gave rise to primeval gravitational waves.

After the entry in the horizon, we must take into account the physical nature of different components (radiation, DM, baryons, etc.). The study of the linear stages of fluctuation evolution is however out of the scopes of these lectures. Here we shall consider those stages as a sort of black box, turning the *primeval* spectrum into the *transfered* spectrum.

Focusing then on matter components, assumed to be DM and baryons, the post–recombination spectrum will read

$$P(t, k) = P(t_i, k)\mathcal{T}^2(t_i, t; k) \tag{3.8}$$

and all linear physics is hidden in the *transfer function* $\mathcal{T}(t_i, t; k)$, telling us what is $\delta(t, k)$ once $\delta(t_i, k)$ is assigned. As we shall debate in more detail in the next section, the transfer spectrum is different for DM and baryon fluctuations, but differences are doomed to be almost reabsorbed before the onset of non–linear physics.

On the contrary, density fluctuations in photons, after recombination, may be neglected, when we study the shaping of large scale structure. Again, we shall be more precise on that in the sequel.

On the contrary, possible *hot* component of DM could complicate this scenario. Hot DM is made of particles initially belonging to the radiative component and becoming non–relativistic after equality. Neutrinos with a mass $\sim eV$ would could play such role. The combination of CMB and large scale data yield a constraint $M_\nu = \sum_{\nu=1}^{3} m_\nu <\sim 0.5$–$0.9$ eV on the sum ot the masses of the 3 canonical neutrino species. The exact limit however depends on the range of models considered and could be significantly softened in cosmologies with unconventional DE nature. Here we shall however neglect the possibility that neutrinos of other hot DM particles have a mass relevant in the shaping of cosmic structure.

Another possibility is that DM – or part of it – is warm, i.e. made of light particles however becoming non–relativistic before equality. Multiple DM components have been considered by several authors (see, e.g., [4], [5]), to overcome specific observational problems, that we shall briefly discuss here.

A first difficulty concerns the amount of substructure in Milky Way sized haloes [10]. Models involving cold DM overpredict their abundance by approximately one order of magnitude. A second issue concerns the density profiles of CDM haloes in simulations, exhibiting a cuspy behavior [11, 12], while the density profiles inferred from rotation curves suggest a core like structure [13]. A third issue concerns dwarf galaxies in large voids: recent studies [14] re-emphasized that they are overabundant.

Replacing cold DM with a "warmer" DM component, made of particles whose mass is $\sim$ 2–3 keV, yields better predictions. There are a number of "thermal" candidates for such warm dark matter (WDM); among them, a sterile neutrino and a gravitino [15] find a reasonable motivation in particle theory [16]. The long streaming length of such particles causes a strong suppression of the power spectrum on galactic and sub-galactic scales [17] and solves several above problems.

Moreover, the profiles of warm DM haloes, similar to cold DM haloes in the outer regions, flatten towards a constant value in the inner regions, as predicted in [18] and found in simulations [19]. However, the core size found is 30–50 pc, while the observed cores in dwarf galaxies are around the 1000 pc scale [20]. A dwarf galaxy core in this scale range would be produced by higher velocity particles, as those belonging to a thermal distribution if their mass is < 0.1–0.3 keV. Increasing the velocity, however,

yields a greater streaming length, exceeding the size of these very dwarf galaxies, in the first place [21].

In view of these difficulties, the idea that WDM is accompanied by a smaller amount of CDM is perhaps the best option [4]. The warm DM particle velocities could then be greater, while a low–mass population is however produced by CDM clustering.

In these lectures, however, the discussion of post–linear evolution will be centered on two basic classes of cosmologies: pure CDM *(pure cold–dark–matter)* and ΛCDM (Λ–*cold–dark–matter*) models. Until a decade ago, pure CDM models were also indicated as SCDM, where S stands for "standard". ΛCDM models are also dubbed *concordance* models and the "standard" denomination is being transfered on them. Both models have $\Omega_{tot} = 1$. For pure CDM

$$\Omega_{o,m} = \Omega_{o,c} + \Omega_{o,b} \simeq 1 \qquad \text{(SCDM)}$$

($\Omega_{o,c}, \Omega_{o,b}, \Omega_{o,m}$: DM, baryon, total matter today's density parameters) while, for ΛCDM,

$$\Omega_{o,m} = \Omega_{o,c} + \Omega_{o,b} \simeq 0.25 - 0.30 \; ; \qquad \text{(ΛCDM)}$$

in top of that there must be a DE component with density parameter $\Omega_{o,de} \simeq 0.7 - 0.75$.

To have a lifetime consistent with stellar observation, pure CDM require that $h \sim 0.5$ (*Hubble parameter*) ; ΛCDM, instead, have $h \sim 0.7 - 0.75$.

Pure SCDM models are now falsified by data. They however keep important for two reasons: (i) Until a redshift ~ 3 , pure CDM and ΛCDM are essentially coincident. (ii) Quite a few problems exhibit an analytic solution for pure CDM cosmologies, while, in ΛCDM, only numerical results are available.

Considering these problem in for pure CDM cosmologies eases then a qualitative understanding of quantitative ΛCDM results.

III. FROM RECOMBINATION TO NON–LINEARITY

Let us then focus on the inhomogenity scenario soon after recombination.

Radiative components have then a density $\rho_r \simeq \rho_m(a_{eq}/a)$ and, if $a_{eq} \sim 10^{-4}$, they still are $\sim 10\,\%$ of the total energy density. Neglecting them, when evaluating

the expansion rate, is therefore quantitatively mistaken, even though residual baryon–photon interactions are negligible. Such neglect, however, eases the understanding of physical evolution, and we shall do so.

DE, on the contrary, is no problem. If it behaves as a cosmological constant, its contribution to the cosmic density budget, at a redshift $\sim 10^3$, is $\sim 10^{-9}$. If DE is a scalar field, its actual significance, at most, can be a couple of orders of magnitude greater. But the picture keeps essentially safe even if we are dealing with an open (or closed) model consistent with basic observations. In such limiting case, the discrepancy of the total density parameter from unity, at $z \sim 10^3$, can hardly exceed 0.1 % .

Let us then focus on DM and baryon spectra, initially quite different, namely over scales $k > k_{hor,eq}$. Here $k_{hor,eq} = 2\pi/L_{hor,eq}$, while $L_{hor,eq}$ is the comoving length scale at the matter–radiation equality (see Figure 1).

In fact, starting from equality, at $z \sim 10^4$, DM fluctuations grow as $\delta \propto a \propto t^{2/3}$. Meanwhile, fluctuations in baryons keep bound to radiation, being part of sonic waves. Therefore, far from increasing, their amplitude suffers a dissipative damping. Quite a slight effect, initially, but becoming more and more severe as the average distance between two successive electron–photon collisions approaches the sonic wavelength. Over small length scales, typically $k >\sim 2.5\, k_{hor,rec}$, the photon m.f.p. reaches the wavelength $\lambda = 2\pi a/k$ when hydrogen has not yet recombined, so that a substantial part of the power in sonic waves is dissipated into heath. The effect is greater for greater k's (smaller length scales). In the baryon distribution, some residual of the primeval fluctuation can be traced just up to $\sim$ 15–20 $k_{hor,rec}$. Beyond that, even the small baryon fluctuations, visible in Fig. 1 at $z \simeq 800$, are a secondary effect triggered by DM fluctuations.

In general, after recombination, we shall therefore have a *small* or vanishing fluctuation in baryons and a major fluctuation in DM. It is also quite possible that the residual primary baryon fluctuation has an opposite sign, in respect to the DM fluctuation of the same wavelength. In fact, during the sonic regime, the sign of fluctuations changes over a half period. Recombination peaks up a casual sonic phase and 'cristallizes' it. Accordingly, in almost half of the cases, the residual primary fluctuation is opposite to DM. In Figure 1, owing to the logarithmic scale, relative

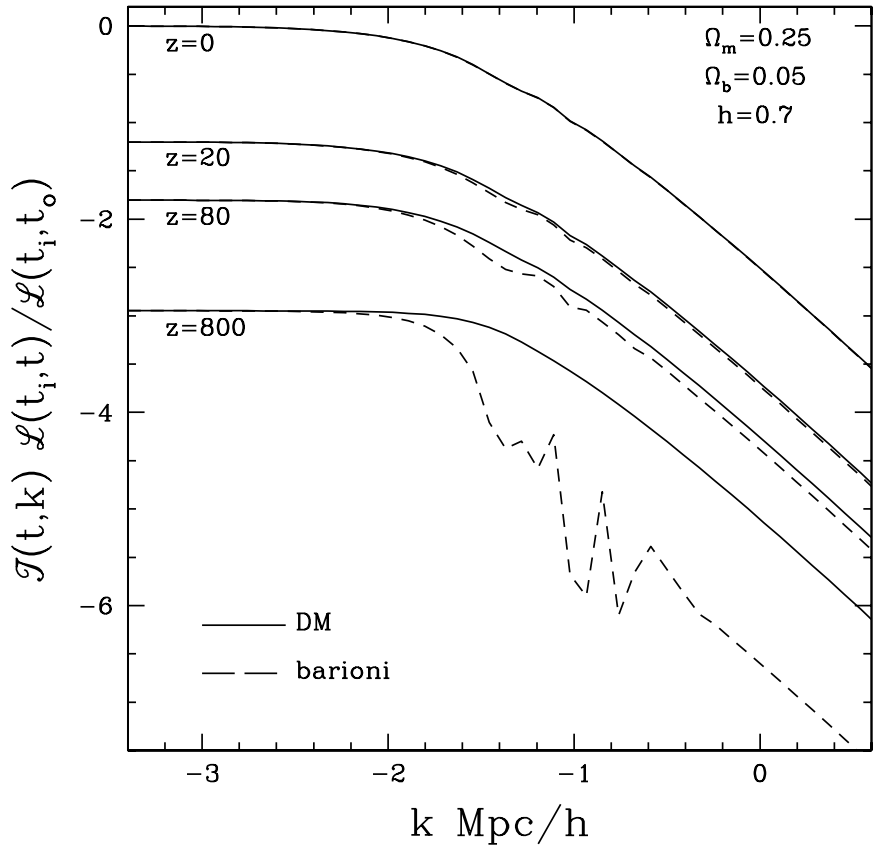


Figure 1. Transfer functions and evolution of DM and baryon fluctuations, for the model indicated within the frame. At $z \simeq 800$ and for $k > k_{hor,eq}$, baryon fluctuations –whose moduli are shown in the Figure– still visibly feel previous sonic regime. At $z = 80$, sonic wave residuals are still more evident in baryons than in DM. At $z \simeq 20$ the two spectra practically coincide and baryon oscillations (BAO) transfer also on DM, being still visible also at $z = 0$.

fluctuation signs are not exhibited.

In any case, DM fluctuation are the *seed* of future inhomogeneities, both for being mostly wider and, namely, because DM density widely exceeds baryons.

Let us then look at a specific wavenumber and consider an "initial time" t_i, now after recombination; let $\delta_{i,c}$ and $\delta_{i,b}$ be then DM and baryon fluctuations. After a reasonable time since t_i, the evolution occurs according to the laws

$$\delta_c = \alpha_+(\Omega_c\delta_{c,i} + \Omega_b\delta_{b,i})(t/t_i)^{2/3} + \zeta_+(\delta_{c,i} - \delta_{b,i}) \ ,$$

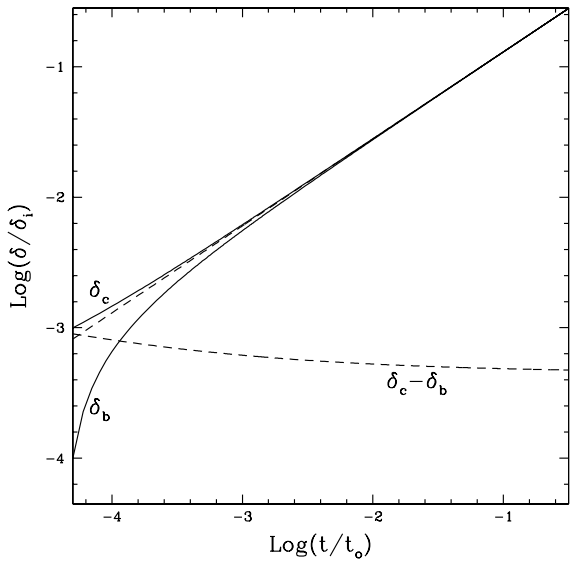


Figure 2. Evolution of fluctuations in DM and baryons, after recombination and until an epoch when DE can be dynamically relevant. In this plot any residual baryon–radiation interaction is neglected. In the example shown we assumed: $\delta_{i,c}/\delta_{i,b} = 10$, $\Omega_{o,m}/\Omega_{o,b} = 5$. Dashed lines show the decrease of the difference between δ_c and δ_b as well as the regular power law growth of the sum $\Omega_c\delta_c + \Omega_b\delta_b$.

$$\delta_b = \alpha_+(\Omega_c\delta_{c,i} + \Omega_b\delta_{b,i})(t/t_i)^{2/3} - \zeta_+(\delta_{c,i} - \delta_{b,i}) \; ; \tag{3.9}$$

here α_+ and ζ_+ are suitable numbers between 0 and 1. They account for our ignorance on the "connection" between the equation solution pre– and post–recombination. A proof of eqs. (3.9) is given in Appendix A2.

In the case shown in Figure 2 we assume $\delta_{i,b} \simeq 0.1\,\delta_{i,c}$ and $\alpha_+ = 1$, $\zeta_+ = 1/2$. The equations (3.9) show that the difference $\delta_c - \delta_b$ decreaes, finally becoming $\sim 10^{-3}$, while δ_c and δ_b separately increase. If photons are taken into account, a slight correction arises from the tenuous residual photon–baryon interaction; it furtherly slowing down the rate at which the two spectra approach, in respect to the law (3.9); this can be somehow appreciated also in Fig. 1.

What happens is that, after recombination, baryons accrete on the *potential wells* DM creates. Meanwhile eqs. (3.9) show that the overall fluctuation of non–relativistic

matter,

$$\delta = \frac{\delta\rho_c + \delta\rho_b}{\rho} = \frac{\delta_c\Omega_c\rho + \delta_b\Omega_b\rho}{\rho} = \Omega_c\delta_c + \Omega_b\delta_b \ , \tag{3.10}$$

grows $\propto (t/t_i)^{2/3}$. We verify it by multiplying the former eq. (3.9) by Ω_c, the latter by Ω_b, and summing side by side.

In Fig. 2 fluctuation amplitudes are arbitrarily normalized. Fig. 1, instead, exhibits the linear behavior of fluctuations as a function of z, for different k values, obtained from a numerical code including all components; amplitudes are normalized so that $\delta = 1$, for $k \to 0$, today. Linear equation solutions are however independent from amplitude.

Let us also outline that "mean" amplitudes at the horizon entry were $\sim 10^{-5}$; they grew by $\sim 1/2$ order of magnitude until equality ($z \sim 10^4$); since then, in a pure CDM model the growth is $\propto a$. At $z \sim 3$, e.g., the "mean" fluctuations will be $\sim 10^{-1}$. Non–linear effects are triggered only when amplitudes are ~ 0.2–0.3 . An individual fluctuation, whose amplitude is double than average, therefore, already enters non–linear stages at $z \sim 3$. Exceptional fluctuations can do so at $z \sim 10$ and, in extreme cases, at still greater redshifts.

In a ΛCDM model, until $z \sim 3$, however, the density parameter of DE is still ~ 0.03, so that the difference between the two models mostly concerns the later fluctuation evolution.

At such low redshift, far from recombination, the difference between fluctuations in DM and baryons, in Fig. 2, is $\sim \mathcal{O}(10^{-3}$–$10^{-4})$. In turn, although taking also into account the residual baryon–photon friction, Fig. 1 shows an apparent complete coherence between baryon and DM already at $z \sim 20$.

We are therefore allowed to conclude that, when non–linear stages begin, DM and baryon fluctuations are however practically equal. Residual effects of sonic waves affect both DM and baryons, and are denominated BAO's (Baryon Acoustic Oscillations), being the remnants of the acoustic waves involving baryons before recombination. The non–linear dynamics will then be able to differentiate DM and baryons again. But the starting point for non–linear evolution is the same for both matter components.

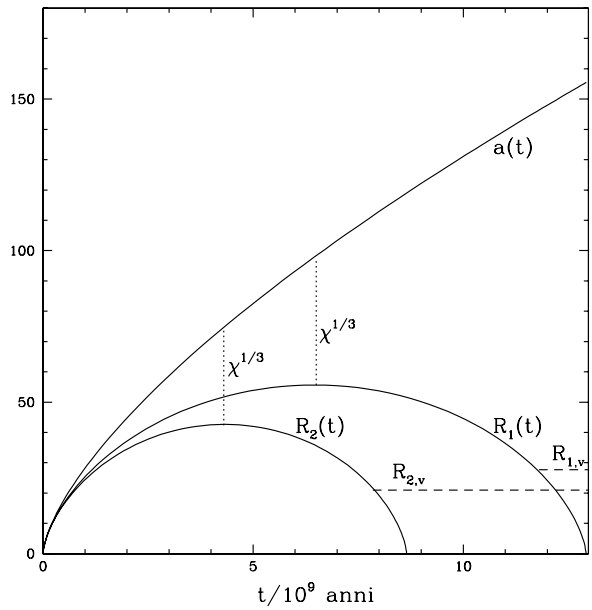


Figure 3. Evolution of spherical overdensities in a SCDM model. $a(t)$ yields the scale factor behavior; $R_1(t)$ and $R_2(t)$, instead, are the radii of overdense spheres, which would expand according to $a(t)$ if their density had been the average cosmic density. At the moment when overdense spheres reach their top expansion, the ratio between their radius and the radius of an equal sphere, expanding in accordance with the *Hubble flow*, is $\chi^{1/3}$ with $\chi \simeq 5.6$, indipendently from the initial overdensity level. Different overdensity levels, instead, correspond to different times when the top expansion is attained. The follwing recollapse stops when the overdensity radius is typically halvened: see text.

IV. POST–LINEAR EVOLUTION

Non linearity puts an end to the separate dynamics of different k components. In the linear regime, the time evolution of the density fluctuation $\epsilon(\mathbf{x}, t)$, in a site $\mathbf{x}$, is obtained by suitably summing up the time behavior of its Fourier components $\delta_{\mathbf{k}}(t)$. At a non–linear level, instead, the unknown function is $\epsilon(\mathbf{x}, t)$ itself, and the equations of motion are partial derivative equations, so that a generic analytic solution is not available.

There is, however, a problem which can be treated. If one considers a spherical top–hat overdensity, whose *initial* radius is R_i, its evolution is fully accounted by the time dependence of its radius $R(t)$. Let us then seek the equation that $R(t)$ must fulfill;

it will be a single ordinary differential equation and, in any case, it will be numerically integrable, at least. In a pure CDM model, however, we shall see that its integral can be put in a simple analytical form.

Let us soon outline that this is problem is far from describing any realistic situation. Spherical perturbations are a zero–volume set of the infinite–dimensional space of possible initial geometries. In spite of that, the solution of this *over*–simplified problem will bear a concrete relevance in predicting the statistical behavior of cosmic inhomogeneities.

Let then $\bar{\rho}_i$ be the average cosmic matter density at the *initial* time t_i. The uniform density inside the sphere, whose radius is then R_i, is then fully specified by the "density parameter", $\Omega_i = \rho_i(< R_i)/\bar{\rho}_i$.

Here we shall make a further ad–hoc hypothesis: that no velocity field initially exist: at the time t_i the radius $R(t)$ grows coherently with Hubble flow. The equation for $R(t)$ can be written also if this extra assumption is lifted, provided that deviations from the Hubble flow occur coherently, in all directions, so to preserve spherical symmetry. Considering some such cases can be a useful exercise.

Finally, let us also assume that $\Omega_i > 1$.

In Figure 3 the behavior of the sphere radius R is shown, for a pure CDM model and for two different positive values of Ω_i. In the next subsection we shall give the analytical expressions from which Figure 3 is obtained. The abscissa of the Figure covers $\sim$ 13 billion years and one of the Ω_i values is selected to cause a complete recollapse just in our epoch. The other value is slightly greater, so that the sphere reaches full recollapse when the cosmic age is 8–9 billion years, at the eve of the present epoch.

Parametric expression and maximum expansion radius

In a pure CDM model, the problem owns an analytic solution. It is worked out in Appendix A3 in the parametric form:

$$\frac{R(\theta)}{R_i} = \frac{\Omega_i}{2(\Omega_i - 1)}(1 - \cos\theta) \ , \quad H_i t(\theta) = \frac{\Omega_i}{2(\Omega_i - 1)^{3/2}}(\theta - \sin\theta) \ . \tag{3.11}$$

According to espressions (3.11), t ed R increase when θ goes from 0 to π. The maximum radius $R_m = R(\pi)$ is attained at the rime $t_m = t(\pi)$. The time t continues then to

increase with θ, while the radius R decreases and returns to 0 for $\theta = 2\pi$. At the maximum expansion, we have

$$R_m = R_i \frac{\Omega_i}{\Omega_i - 1} \ , \quad H_i^2 \Omega_i \, t_m^2 = \left(\frac{\pi}{2}\right)^2 \left(\frac{\Omega_i}{\Omega_i - 1}\right)^3 = \left(\frac{\pi}{2}\right)^2 \left(\frac{R_m}{R_i}\right)^3 \ . \tag{3.12}$$

Accordingly, it is

$$H_i^2 \Omega_i \, t_m^2 \left(\frac{R_i}{R_m}\right)^3 = \left(\frac{\pi}{2}\right)^2 \tag{3.13}$$

so that, reminding that at the initial time (similarly as at any time) $H_i^2 = (8\pi/3) G \, \bar\rho_i$, and by dubbing ρ_m the density in the sphere at the time t_m, we have

$$\frac{8\pi G}{3} \ \Omega_i \, \bar\rho_i \left(\frac{R_i}{R_m}\right)^3 \ t_m^2 = \frac{8\pi G}{3} \rho_m \, t_m^2 = \left(\frac{\pi}{2}\right)^2 \ . \tag{3.14}$$

In fact $\Omega_i \bar\rho_i = \rho_i$ and $(4\pi/3)\rho_i R_i^3$ $[(4\pi/3)\rho_m R_m^3]$ is the conserved mass within the sphere at the initial [maximum] time. Let us then compare eq. (3.14) with the equation

$$\frac{8\pi G}{3} \bar\rho(t_m) \, t_m^2 = \left(\frac{2}{3}\right)^2 \tag{3.15}$$

derived from Friedmann equations, taking into account that $H = 2/3t$. Herefrom we can estimate the *density contrast*, between inside and outside the perturbation, when the maximum radius is attained, reading

$$\chi = \frac{\rho_m}{\bar\rho(t_m)} = \left(\frac{3\pi}{4}\right)^2 \simeq 5.6 \ , \tag{3.16}$$

indipendently from the radius of the sphere considered, the time when this radius is attained and, therefore, the initial Ω_i value. In Fig. 3 the density contrast is obtainable from the ratio between R e a, which holds $\chi^{1/3}$ for both spheres, when they are at their top expansion.

An alternative way of stating the same results amounts to saying that a spherical fluctuation which has attained a density contrast 5.6 is at its top expansion. Clearly, we are quite far from a linear regime.

Virialization

Radius increase and decrease, as shown in Fig. 3 (and as often occurs for gravitational trajectories), are symmetric. Both the Figure and the solutions (3.11) however hold for

a pure CDM model only. We shall later debate what changes in a ΛCDM cosmology or, generically, in the presence of DE.

Let us then outline that, in the idealized case of an exactly top–hat spherical perturbation, the $R(t)$ behavior is faithfully followed by any point inner to the sphere. Namely, let r_i e $r(t)$ be the distances, at the initial time and at the time t, between the center O and a generic point P_i inside the sphere. The ratio $r(t)/r_i$ must then behave exactly as $R(t)/R_i$.

Since the very beginning of this treatment, however, we outlined that spherical top–hat perturbations are highly unlikely matter distributions. They keep so even if we admit that small perturbations exist to the uniform density inside the spheres.

During expansion, however, inner points get gradually farther and the ratio $r(t)/r_i$ has a trend not so far from $R(t)/R_i$. If any discrepancy occurs, due to the presence of small inhomogeneities, the expansion can only smear it out.

On the contrary, contraction bursts inhomogeneities. We would need a perfect synchronization of all $r(t)$, in order that all motions continue coherently, until R vanishes and recollapse is complete. Even small inhomogeneities, instead, are enough to perturbate the order, giving rise to transversal motions. Also in a system mostly made of non collisional DM particles, individual motions become unordered. Particle velocities are re–distributed and, when the virial of the sphere

$$Vir = 2\,E_k + E_p \tag{3.17}$$

($E_{k,p}$: kinetic, potential energy) vanishes, the spheres reaches an equilibrium configuration and no further contraction occurs.

In a pure CDM model, virialization is attained when the radius is

$$R_v = R_m/2\ . \tag{3.18}$$

In fact, let M be the mass of the sphere and let us require $Vir = 0$ in eq. (3.17); we obtain

$$E_k = -\frac{E_p}{2} = \frac{1}{2} G f \frac{M^2}{R_v}\,; \tag{3.19}$$

here f is a numerical factor taking into account the geometry in gravitational interaction among different sphere points. Henceforth, if we require that the total

energy

$$E_k + E_p = -\frac{1}{2}Gf\frac{M^2}{R_v} = -Gf\frac{M^2}{2R_v} \ , \tag{3.20}$$

coincides with energy at top expansion, which is only potential and reads

$$E_{p,m} = -Gf\frac{M^2}{R_m} \ , \tag{3.21}$$

we soon obtain $R_m = 2R_v$.

In Fig. 3 virialization radii are indicated by dashed lines.

Virialization is not instantaneous; numerical simulations show that the time to achieve it is well approximated by the time ideally needed to achieve full recollapse; we therefore estimate that the virialization time $t_v = 2t_m$.

When R reduces by a factor 1/2 and the time increases by a factor 2, the scale factor increases by $2^{2/3}$. The ratio a/R, therefore, increases from a value $\chi^{1/3}$ to a value $\chi^{1/3} \times 2^{5/3}$. Accordingly, the density contrast between the inner density ρ_v and the average cosmic density $\bar{\rho}(t_v)$ will be

$$\Delta_v = \frac{\rho_v}{\bar{\rho}(t_v)} = \chi \times 2^5 = 2(3\pi)^2 \simeq 178 \ . \tag{3.22}$$

This result holds for any sphere radius and initial density contrast Ω_i. Changing Ω_i just modifies the times t_m and $t_v = 2t_m$. Therefore, the expression (3.22) holds for a sphere recollapsing at any time, at the moment when virialization is attained.

If all that occurs before the present time ($t_v < t_o$) and, after virialization, no significant further material accretion occurs, the date of virialization (t_v) can be recognized by the density contrast Δ_o we observe at t_o. As shown in the Figure, the further expansion of the Universe after t_v makes $\Delta_o > 178$.

All that holds for pure CDM models only. The value 178, often approximated into 180 or 200, had however a great success in the analysis of simulations and observational materials, where configurations with such density contrasts are often sought.

It is also fair reminding that the value (3.22) is obtained taking into account gravitation only and neglecting any residual action of radiation. A numerical test of residual radiation effects raises the virial density contrast up to $\sim$ 180–181 .

ΛCDM models

Let us now discuss what changes in the presence of a DE component, in particular in ΛCDM cosmologies.

The very treatment of the evolution of a top–hat overdensity should then be modified. In Appendix A3 we studied its self gravity by means of the equation

$$(\dot{R}/R)^2 = (8\pi/3)G\rho_i(R_i/R)^3 - \kappa/R^2 \ , \tag{3.23}$$

formally similar to one of Friedmann's equation, but just meaning energy conservation: more specificaly, the term $(\dot{R}/R)^2$ essentially accounts for kinetic energy, the term $(4\pi/3)G\rho_i(R_i/R)^3$ essentially derives by gravitational self–interaction, and κ is proportional to the conserved total energy.

When DE is taken into account this equation cannot be extended: DE density is scale independent, while

$$\Omega_m(z) = \frac{\Omega_{o,m}}{\Omega_{o,m} + (1-\Omega_{o,m})(1+z)^{-3}} \ , \tag{3.24}$$

so that $\Omega_{de} = 1 - \Omega_m$ is the varying density parameter of DE. Moreover DE does not cluster. Then, while the sphere expands, matter dilutes and DE contribution increases, being maximum around top expansion. During recontraction, then, it is as though DE were leaking out from the overdensity. Clearly, the total energy within the sphere is not constant and no energy conservation equation, similar to eq. (3.23), can be used.

The energy integral (3.23) must then be replaced by the second order equation

$$-(\ddot{R}/R) = (4\pi/3)G\left[\rho_{m,i}(R_i/R)^3 + (\rho_{DE} + 3p_{DE})_t\right] \ , \tag{3.25}$$

similar to the other Friemann eq. and essentially meaning $\mathbf{f} = m\mathbf{a}$, although $3p$ must be added to ρ, to build the fair gravity source m. Beside of that, when DE yields a substantial contribution to expansion, the law relating t and a is modified.

If these effects are taken into account, both the density contrast at top expansion and the density contrast at virialization are different. The empirical expression [22]

$$\Delta_v = 18\pi^2 - 82\Omega_{de} - 39\Omega_{de}^2 \tag{3.26}$$

yields then the virial density contrast at any redshift, just as a function of Ω_{de} (see also Fig. 4), at any redshift, with an approximation better than 1 %.

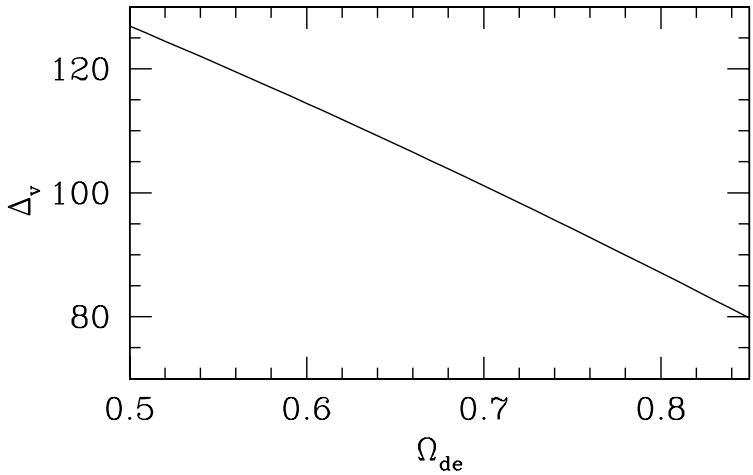


Figure 4. Virial density contrast in ΛCDM models.

Δ_v is the ratio between the density inside the sphere and the critical density ρ_{cr}, assumed to be the average cosmic density. Accordingly

$$\Delta_v = \frac{\rho_{m,in}}{\rho_{cr}} = \frac{\rho_{m,in}}{\rho_m}\Omega_m = \Delta_{v,app}\Omega_m \ , \tag{3.27}$$

$\Delta_{v,app}$ being the "apparent" density contrast, between the matter densities inside and outside the sphere. A few numerical examples are in Table I.

Table I

Ω_m	$\Delta_v = \frac{\rho_{m,in}}{\rho_{cr}}$	$\Delta_{v,app} = \frac{\rho_{m,in}}{\rho_m}$
0.25	94.2	377
0.28	98.4	351
0.31	102.5	331

They show that, when a density contrast 200 is used to define phaenomenological clusters, non virialized areas get included in the systems.

In Figure 5, we plot virial density contrasts at $z = 0$, for models with DE different also from ΛCDM, as obtained from numerical computations.

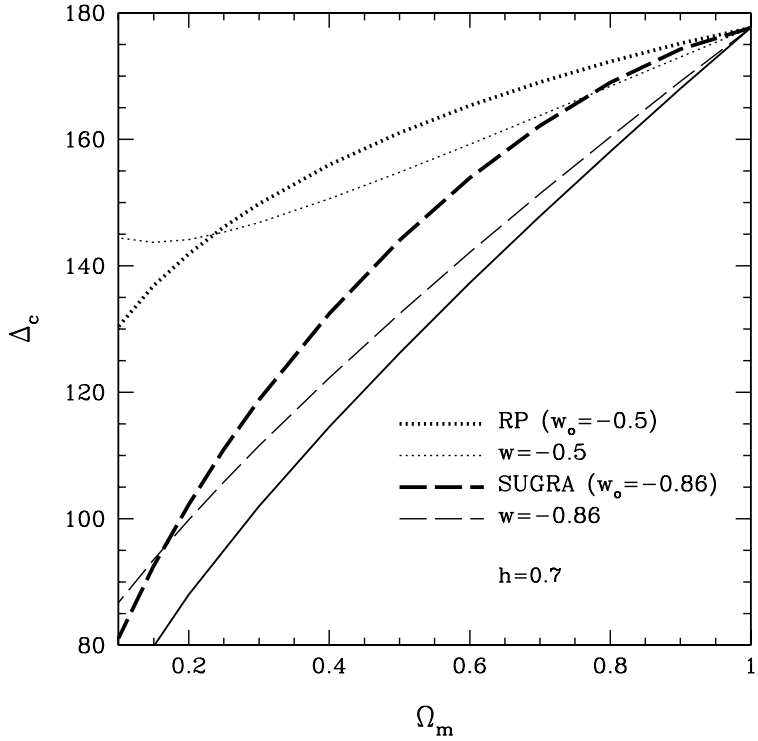


Figure 5. Virial density contrast (here dubbed Δ_c) in models with DE and $\Omega_{om} = 0.3$, $h = 0.7$. We show results for models with a state parameter of DE $w = p/\rho = \text{const.}$, as well as models with variable w, when DE is due to a scalar field self–interacting through a RP or SUGRA potential (see Cap XII); these latter models are parametrized by the value taken by the ratio $w_o = p/\rho$ at the present time. For the sake of comparison, the solid line shows also the values ΛCDM.

Indipendently of the value of Δ_v, however, aside of the *virial density contrast* there shall be a *virial radius* R_v, such that

$$M = \frac{4\pi}{3} R_v^3 \rho_{cr} \Delta_v \ , \quad \text{ovvero} \quad R_v(M) = (3M/4\pi\rho_{cr}\Delta_v)^{1/3} \ . \tag{3.28}$$

In the present cosmic conditions, this corresponds to

$$R_v(M) = 2.05 \times 10^{-5} \left(\frac{M/h^{-1}M_\odot}{\Delta_v/100} \right)^{1/3} \tag{3.29}$$

For instance, in a ΛCDM model with $\Delta_v = 100$, the virial radius of a mass $4 \times 10^4 h^{-1} M_\odot$ is $\simeq 1.5\, h^{-1}$Mpc, approximately the value of the *Abell radius* often used to define galaxy clusters from optical data.

Post–linear density contrasts and linear evolution.

The growth of the density contrast, from an intial value $\Delta_i \equiv \Omega_i = 1 + \delta_i$, up to Δ_v, can be compared with the behavior δ_i would have according to linear equations. In a pure CDM model the linear growth is

$$\delta_+(t) = \delta_{i,+}(a/a_i) = \delta_{i,+}(t/t_i)^{2/3} \,, \tag{3.30}$$

for the increasing mode. Here $\delta_{i,+} = (3/5)\delta_i$, as follows from assuming that, at the time t_i the expansion rate of the spherical overdensity coincides with Hubble flow (see Appendix A3).

Being then $\Omega_i - 1 = \delta_i$ and $H = 2/3t$, eq. (3.11) also reads

$$\frac{2}{3}\frac{t(\theta)}{t_i} = \frac{\Omega_i}{2\delta_i^{3/2}}(\theta - \sin\theta) \tag{3.31}$$

and, for $\theta = \pi/2$, i.e. at the maximum expansion, we have

$$\frac{2}{3}\frac{t_m}{t_i} = \frac{\pi}{2}\frac{\Omega_i}{\delta_i^{3/2}} \qquad \text{i.e.} \qquad \delta_i^{3/2}\frac{t_m}{t_i} = \frac{3\pi\Omega_i}{4} \tag{3.32}$$

so that, taking into account that, in the increasing mode, $\delta \propto t^{2/3}$,

$$\delta_i \left(\frac{t_m}{t_i}\right)^{2/3} = \frac{5}{3}\delta_{i,+}\left(\frac{t_m}{t_i}\right)^{2/3} = \frac{5}{3}\delta_m = \left(\frac{3\pi\Omega_i}{4}\right)^{2/3} \,. \tag{3.33}$$

Hence, neglecting the tiny difference between Ω_i and 1, we shall have that

$$\delta_m \simeq \frac{3}{5}\left(\frac{3\pi}{4}\right)^{2/3} \simeq 1.06\,. \tag{3.34}$$

Accordingly, from the physical point of view, at the time t_m the fluctuation is no longer in the linear regime; this was however known as we already showed that the density contrast at the time t_m is not $\chi_l = 1 + \delta_m \simeq 2.06$ but reads $\chi \simeq 5.6\,$.

In spite of that, linear equations still bear a formal significance. Only when $\delta(t) \ll 1$ they yield the physical value of the ratio $\delta(t)/\delta_i$; otherwise they map a variable δ on the variable t : Instead of stating that a fluctuation has attained its t_m time, e.g., we can say that its "linear amplitude" is $\delta = \delta_m \equiv 1.06$. There is a clear advantage in doing so, for this characterizes maxima at any time.

Similarly, at the time $t_v = 2t_m$, the "linear amplitude" has grown by a further factor $2^{2/3}$, becoming

$$\delta_v = \frac{3}{5}\left(\frac{3\pi}{2}\right)^{2/3} \simeq 1.68 \ . \tag{3.35}$$

This linear amplitude therefore characterizes a fluctuation which has just virialized. Any fluctuation which, at a given time, has a linear amplitude $\delta > \delta_v$, therefore, should have already settled in a virialized state.

A final point: δ_m and δ_v values have been computed for a pure CDM model. At variance from χ and Δ_v, however, their dependence on the model is marginal and these values can be safely used for a ΛCDM model as well.

V. FLUCTUATION AMPLITUDE DISTRIBUTIONS

The distribution of matter, after recombination, is surely not made of spherical overdensities. Spherical structures, in the astronomical world, arise from pressure balance when their surfaces are then isocores. Stars are almost spherical, for instance, unless fast rotators. Nothing similar at recombination, when cosmic materials had negligible pressure.

In due time, arising from the matter distribution at recombination, greater and greater systems form; among them galaxy groups and clusters. In the filamentary fabric of the present cosmic distribution, they occupy intersection sites, as a kind of knots. At *zero*–order they now approach a spherical symmetry, however acquired only recently, thanks to the so–called *violent relaxation* towards virial equilibrium, after the stop of gravitational collapse.

The previous analysis however bears a significance if interpreted backwards; starting from the present quasi–spherical configuration, we go back to antecedent times, when configurations were mostly different from spheres. However, it is not misleading to assume that the clock we established for a spherical overdensity approaches the growth schedule of the materials destined to yield a virialized structure.

For instance, from the density contrast measured in a cosmic system (a galaxy cluster, a galaxy group, or – with suitable cautions – a galaxy) we can approach

the average density of the Universe at the time it virialized and, then, deduce its approximate formation redshift.

The overall recipe could improve if we are able to achieve a better timing for the growth history, and we shall see that several authors tried to implement this program. The basic point, however, is that good results can already be achieved by using the timing estabished from the study of the spherical growth. In particular, in this way we arrive at the *so–called* Press & Schechter expression for the mass function of virialized systems [23].

Within this context, the time variable is systematically replaced by the "linear amplitude" δ, so that we need not seeking the time $t_i(\underline{t})$ yielding the virialized structures at a suitable time $\underline{t}$: in order to know how many objects of mass M can form, we rather compute the probability to have fluctuations of "linear amplitude" $\delta(\underline{t}) > \delta_v$ over the mass scale M, with no reference to any particular initial time, when the physical and linear amplitudes possibly coincided.

This procedure has a further fundamental advantage allowing us, first of all, to take into account that not all fluctuation, over a given mass scale M, have equal amplitude. Let us then first refer to a time when fluctuation amplitudes are safely in the linear regime, at an epoch like the time t_i considered above. Let us suppose that the mass variance over the mass scale M

$$\sigma_M^2 = \langle |\epsilon_{R_M}(\mathbf{r})|^2 \rangle \tag{3.36}$$

is $\ll 1$ at that time. Here $M = (4\pi/3)\rho_m R_M^3$ and $\epsilon_{R_M}(\mathbf{r})$ are the fluctuation amplitude in the site $\mathbf{r}$, *smoothed* on a scale R_M.

Here we shall keep to a simple top–hat *smoothing*: We start from the matter density field $\rho(\mathbf{r})$ and define the smoothed field $\rho_R(\mathbf{r}) = M_R(\mathbf{r})/V_R$, with $V_R = 4\pi R^3/3$, while $M_R(\mathbf{r})$ is the mass contained in a volume V_R centered in $\mathbf{r}$. Then $\epsilon_R(\mathbf{r}) = \rho_R(\mathbf{r})/\bar{\rho} - 1$.

In this case, M_R is the whole mass within at distance $< R$ from $\mathbf{r}$, while any mass above R is excluded by M_R evaluation. This abrupt inclusion–exclusion jump, in same cases, is to be replaced by a smoother mechanism, e.g. by a Gaussian weight applied to the density as the distance from $\mathbf{r}$ increases. We shall no longer debate this kind of extensions here.

The averaging in eq. (3.36) should be done at constant **r** over an universe ensamble. According to the *ergodic* assumption and more realistically, we average over a volume of size $\gg R_M$.

If the spectrum of density fluctuations is $P(k)$, it can then be shown that

$$\sigma_M^2 = \frac{1}{2\pi^2}\int_0^\infty dk\; k^2\, P(k)\, W(kR_M) \tag{3.37}$$

Here $W(x) = 3(\sin x - x\cos x)/x^3$ is obtained from a Fourier–transform of the top–hat filter (see Appendix A4).

In the average (3.36), positive and negative fluctuation have an equal weigth. If the matter distribution is Gaussian, we then also expect that the probability distribution to have a fluctuation of size δ, over a scale M, reads

$$p(M,\delta) = \frac{1}{(2\pi)^{1/2}\sigma_M} e^{-\frac{1}{2}\left(\frac{\delta}{\sigma_M}\right)^2} \;. \tag{3.38}$$

Let us then outline that the linear evolution, acting with an equal factor over all amplitudes, preserves the nature of the "initial" distribution on amplitudes: e.g., an initial Gaussian distribution keeps Gaussian until it evolves linearly. The only effect of the linear growth will be the gradual increase of the mass variance σ_M. In a pure CDM cosmologies it is simply $\sigma_M \propto a$.

As time elapses ans σ_M increases, therefore, the Gaussian distribution undergoes a progressive dilatation, as shown in Figure 6 for a scale factor growing from a_1 to $a_2 = 4\, a_1$. For graphic reason it is hard to plot a wider growth. Let us rather outline again that the growth described in Figure 6 concerns the "linear amplitude".

In Fig. 6 the abscissa value $\delta_v = 1.68$ is then also indicated. The ratio between the (quite small) area of the triangoloid between Gaussian and abscissa, at the r.h.s. of such δ_v, and the total area between Gaussian and abscissa, yields the fraction of fluctuations over the scale M which, according to a linear evolution rate, are above δ_v at the time t_2 (scale factor is a_2).

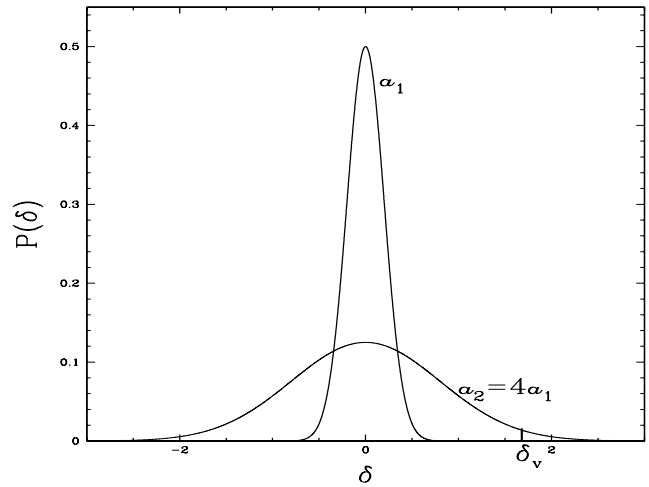


Figure 6. Cosmological evolution of amplitudes, if ruled by linear equations. When the scale factor reaches the value a_2 (time t_2), a small but significant fraction of fluctuations (linearly) overcomes the critical value $\delta_v = 1.68$. In the plot it is given by the ratio between the area of the triangoloid between curve and abscissa at the right of δ_v (quite small), and the whole are between curve and abscissa. Among the fluctuations over the scale considered, this is the fraction forming virialized objects within the time t_2.

VI. PRESS & SCHECHTER EXPRESSIONS AND THEIR GENERALIZATIONS

Let us then reconsider eq. (3.37), yielding the "mass variance" over the scale M. Clearly, when M or R_M increase, the upper limit to the integration in eq. (3.38) decreases; therefore, $\sigma_M \equiv \sigma(M)$ is a decreasing function of M, for any kind of spectrum. We shall use the inverse ot this function:

$$\textit{being} \quad \bar{\sigma} = \sigma(M) \ , \quad \textit{let also be} \quad M = \mu(\bar{\sigma}) \tag{3.39}$$

In general, the decreasing function $\mu(\sigma)$ needs to be worked out numerically. However, if we assume

$$\sigma(M) = (M/M_o)^{-\alpha} \tag{3.40}$$

the exponent α is not rapidly variable and some handable expressions can be worked out by assuming that $\alpha = const.$. For our purposes the inverse function

$$\mu(\sigma) = M_o \sigma^{-1/\alpha} \tag{3.41}$$

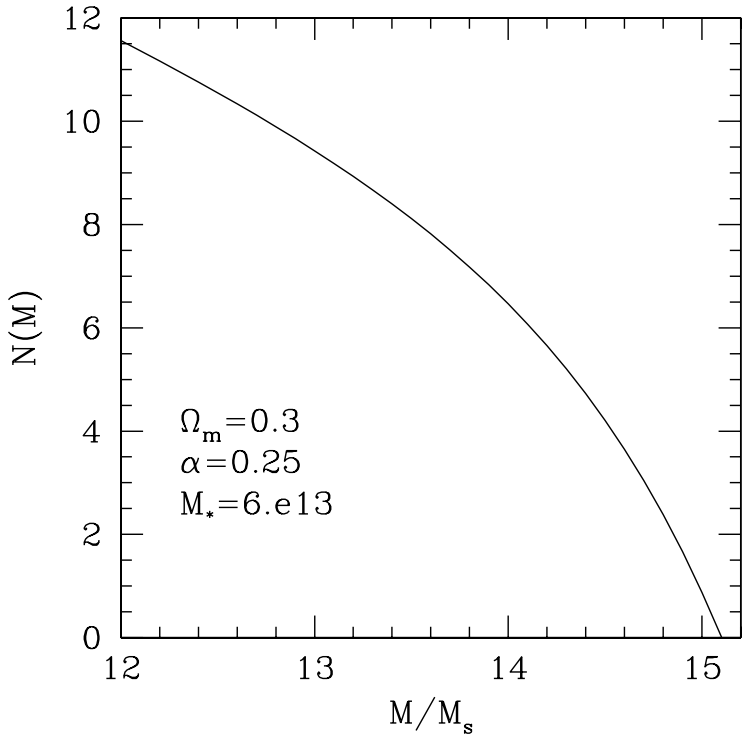


Figure 7. PS mass function.

is also useful. Given the distribution (3.38), the probability to have a fluctuation $> \delta_v$ on the scale M is obtained by differentiating

$$\Pi(\delta_v, M) = \int_{\delta_v}^{\infty} d\delta \frac{1}{(2\pi)^{1/2}\sigma(M)} e^{-[\delta^2/2\sigma^2(M)]} = \int_{\delta_v/\sigma(M)}^{\infty} d\nu \sqrt{\frac{2}{\pi}}\, e^{-\nu^2/2} \tag{3.42}$$

in respect to M. Press & Schechter added a factor 2 to improve the fit of this expression to datasets, claiming that it allowed to take into account the accretion of materials from underdense regions (half of the total) towards areas where systems are forming.

In the expression

$$f_{PS}(\nu) = \sqrt{\frac{2}{\pi}}\, e^{-\nu^2/2} \tag{3.43}$$

for the *PS excursion set*, and in forthcoming expressions, we include this correction, although its motivation is weak; indeed, improvements of the Press & Schechter approach require more extended and better motivated modifications in the excursion set.

The distribution on M of the number of objects on scale M is then obtainable by multipying the derivative $\partial\Pi(\delta_v, M)/\partial M$ by the ratio ρ_m/M, so to have

$$n(M) = 2\frac{\rho_m}{M}\frac{1}{\sqrt{2\pi}}\frac{\delta_v}{\sigma^2(M)}\frac{d\sigma(M)}{dM}e^{-\delta_v^2/2\sigma^2(M)} \tag{3.44}$$

and, by using the expression (3.40) of the mass variance as though α were constant, we obtain

$$n(M) = \frac{2\alpha}{\sqrt{\pi}}\frac{\rho_m}{M_o^2}\left(\frac{M}{M_o}\right)^{\alpha-2}\frac{\delta_v}{\sqrt{2}}e^{-(\delta_v^2/2)(M/M_o)^{2\alpha})} =$$

$$= \frac{2\alpha}{\sqrt{\pi}}\frac{\rho_m}{M_*^2}\left(\frac{M}{M_*}\right)^{\alpha-2} e^{-(M/M_*)^{2\alpha}} \quad with \quad M_* = M_o(2/\delta_v^2)^{1/2\alpha}\ . \tag{3.45}$$

Such differential expression(s), however, have scarce applications. Observational data, unavoidably, will be integral, yielding the number of systems above a fixed scale.

By integrating the expression 3.D.8 we then obtain

$$N(>M) = \sqrt{\frac{2}{\pi}}\rho_m\int_M^\infty d\mu\frac{1}{\mu}\frac{\delta_v}{\sigma^2(\mu)}\frac{d\sigma(\mu)}{d\mu}e^{-\delta_v^2/2\sigma^2(\mu)} =$$

$$= \frac{\rho_m}{M}\int_{\delta_v/\sigma(M)}^\infty d\nu\ \sqrt{\frac{2}{\pi}}\,e^{-\nu^2/2}\frac{M}{\mu(\delta_v/\nu)}\ ; \tag{3.46}$$

here $\mu(\delta_v/\nu)$ is the inverse variance function defined in eq. (3.39), while the PS excursion set is again in evidence.

The meaning of eq. (3.46) is transparent. If it were $N(>M) = \rho_m/M$, all cosmic matter would be in systems of mass $> M$. Of course it is not so, and the fraction of cosmic matter in systems $> M$ is obtainable from the integral coefficient in eq. (3.46).

The expression at the second line in eq. (3.39) is dubbed *Press & Schechter* mass function (hereafter PS). This expression, although based on simple arguments and including "naive" results from spherical collapse theory, approaches gravitational simulation results, namely in the mass range above $\sim 10^{14}M_\odot h^{-1}$.

VII. IMPROVEMENTS OF PRESS & SCHECHTER EXPRESSION

The original work on PS mass function [23] dates back to 1974. Simulation and data precision set a real challenge to PS expression validity only one decade ago and several

improvements were then proposed. A comprehensive paper on possible improvements is [24], to whom we refer also for further citations.

In any case one should compare the growth of formal and numerical complications with the real improvement in predictions. Doubtless, some peculiar aspects in the PS approach are intrinsically weak and it is worth focusing on them. But the basic issue is up to which point any such expression can really allow us to compare different cosmologies, which enter the expressions just through the model dependence of σ_R and its evolution.

We may wonder, e.g., whether in two cosmologies, yielding the same σ_R at two different times, we really expect the same mass functions. For instance, time derivatives of σ_R could bear a significant impact.

We can then anticipate that modern simulations allow us to conclude that there exist a model dependence which is not conveyed by the σ_R behavior, but that this extra dependence is a "higher order" correction. Accordingly, it makes sense, first of all, to improve the "zero–th order" expression, trying to fully exploit all the information conveyed by the direct dependence of σ_R on the cosmology.

When this is not enough, one might still seek a pattern to avoid to run a full simulation. Also because the numerical complexity of simulation algorithms often make hard to focus on the impact of different parameters.

Within the former pattern, a good result was obtained by Sheth & Tormen [27]. If, on the contray, the latter pattern is pursued, an algorithm like PINOCCHIO [25] may be the best option. Here we shall briefly discuss these patterns.

The cloud–in–cloud problem and PS–like expressions for Warm Dark Matter

Before entering in such discussions, however, let us outline that an apparent weakness of the Press-Schechter arguments lays in the so–called cloud–in–cloud problem. Its significance depends on the scale range and the model, becoming dramatic in cosmologies with CDM replaced by WDM, where it vanifies the use of any PS–like expression.

Let us debate this question by starting from the expression (3.37) that we rewrite

$$\sigma_R^2 \;=\; \int_0^\infty \frac{dk}{k}\;|\Delta(k)|^2\tilde{W}^2(kR) \quad \text{with} \quad \Delta^2(k) \;=\; \frac{1}{2\pi^2}k^3\;P(k) \tag{3.47}$$

The spectral function $\Delta^2(k)$ then accounts for the contribution of each logarithmic k–interval to σ_R.

Clearly, $\Delta^2(k)$ is an increasing function of k if the transfered spectrum $P(k)$, in the relevant scale range, declines in a way $\propto k^{-\alpha}$ with $\alpha < 3$. In such range, in fact, $\sigma_R^2 \propto R^{\alpha-3}$ and, for $\alpha < 3$, the average fluctuation $|\epsilon_R|^2$ steadily increases when R decreases, so that we expect more and more systems to form at increasingly smaller scales.

Indeed, in a ΛCDM or analogous models, it is $\alpha \sim$ 2–2.5 and this sets a scale dependence which allows to take PS–like results as a good approximation, but does not prevent the possibility that, in a specific site, when smoothing the mass distribution on a scale R, it is $\epsilon_R(\mathbf{x}) < \delta_c$, while $\epsilon_{R'}(\mathbf{x}) > \delta_c$ for some $R' > R$. In that site, then, systems of size R' collapse to form a virialized object, overwhelming the more diffuse patches within it, whose contribution to the mass function cannot then be evaluated within the PS philosophy.

Correcting for the effects of the cloud–in–cloud problem is non trivial. This is a typical point where modest improvements can only be reached at the expences of high complications. Let us rather focus on a critical consequence of the σ_R dependence on R, when CDM is replaced by WDM.

The point is that, in WDM cosmologies, primeval fluctuations on small scales are suppressed. The average momentum distribution of WDM is mostly assumed to read

$$f(p) = \frac{\mathcal{N}_s}{(2\pi)^3}\frac{1}{expp/T_w + 1}\;, \tag{3.48}$$

with a temperature parameter $T_w \sim (2/g_{dg})T_\gamma$. Here g_{dg} is the effective number of spin states belonging to the thermal background when WDM decoupled from it; the estimate of T_w is only approximate, as we should also take into account other components possibly decoupling from photons later on; one such component surely exists, being made of neutrinos.

Notice that (3.48) approaches an equilibrium distribution only when the WDM particle mass $m_w \ll T_w$. Accordingly, T_w is the order of magnitude of WDM particle

energy and they keep relativistic until $T_w > m_w$. On the scales entering the horizon when it is still $T_w > m_w$, particles rapidly stream freely out from density excesses or defaults, which are then soon erased. Only on greater scales ($k < k_{fs}$) entering the horizon later, the primeval spectrum survives.

Accordingly, WDM models predict a high–k spectral cut–off. At scales R approaching $1/k_{fs}$, therefore, $\Delta^2(k)$ declines and σ_R receives no further spectral contribution when R is lowered. Therefore (3.46) ceases to hold. Notice that WDM models do not exclude the presence of systems on scales $> 1/k_{fs}$, due to fragmentation of structures on greater scales. As we shall soon debate, most gravitational collapses are non–spherical, and a highly asymmetrical structure easily fragments, in its non–linear stages. The point is that eq. (3.46) is derived within a fully different philosophy and cannot help predicting anything in such regime.

In this case even the complex techniques, devised by Bond et al. [26] and others, to overcome the cloud–in–cloud problem, become useless.

Sheth & Tormen expression

Let us then discuss the improvement of PS expression obtainable when the non–sphericity in the gravitational growth is taken into account. The neglect of *a*–sphericity in real collapses causes scale–dependent errors and is the main reason why Press & Schechter predictions, when corrected by a factor 2 over all scales, exceed N–body results by a significant factor at low scale.

Sheth & Tormen [27] (hereafter ST) found then that simulations are better fitted by a mass function obtainable by replacing the excursion set (3.43) by

$$f_{ST} = A\sqrt{2\alpha/\pi}\left[1 + (\nu\alpha^{1/2})^{-2p}\right]\nu e^{-\alpha\nu^2/2} \ . \tag{3.49}$$

Here $\alpha = 0.707$, $p = 0.3$, and $A = 0.322$. These values are mostly obtained by fitting to simulations (in my experience, a value of A slightly different from the one proposed by ST can improve the fit to simulation mass functions, when different halo–finding algorithms are considered).

The expression (3.49) is obtained by replacing the requirement $\nu > \delta_c/\sigma$ with the

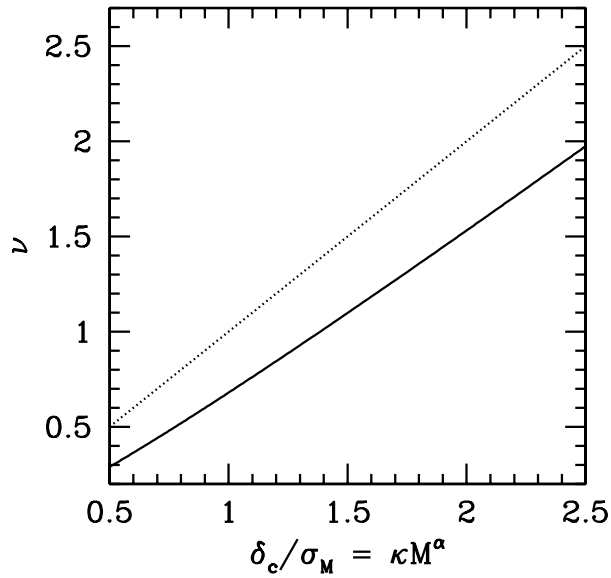


Figure 8. PS (dotted line) and ST (solid line) barriers compared.

condition

$$\nu/(1+\beta/\nu^{\gamma}) > \delta_c/\sigma \ . \tag{3.50}$$

The values $\beta \simeq 0.47$ ($\sim 1/2$) and $\gamma \simeq 0.62$ ($\sim 2/3$) are in agreement with the above α and p.

The rational is that the integration threshold (barrier) should be derived by considering the growth of ellipsoids, rather than spheres. In principle, one should then integrate over a suitable range of over ellipsoidal shapes. The problem was however simplified by considering a single ellipsoidal shape allowed to be scale dependent. The parameters which appear in the expressions (3.49), (3.50) implicitly define the preferred ellipsoidal shape and its scale dependence.

In Figure 8, the ellipsoidal collapse barrier in eq. (3.50) is compared with the PS barrier. The integration over ν is performed from the dotted or solid line upwards. Clearly, starting from the solid line, we include a wider ν interval and the integration result is unavoidably greater. Such increase, however, is smaller for smaller mass–scales as, according to eq. (3.39), $\delta_c/\sigma_M = \kappa M^{\alpha}$ with $\kappa = \delta_c/M_o^{\alpha}$.

The overall increase, somehow replacing the "artificial" extra–factor 2, is however

not so significant as the final result is subject to data fitting, yielding the constant A in eq. (3.49). What is more significant is the different mass dependence.

The physical issue, first of all, is that the average ellipsoidal collapse is however faster than a spherical one. But the key point is that, for lower–mass systems, tidal disruption is more likely. Faster collapses, i.e. greater ellipticity systems, are then more likely to avoid tidal disruption; the integration barrier is therefore to be displaced upwards, at low M.

In other words: less systems form at low mass beacuse tidal disruption is more effective. The ST expression allows to take this effect into account and therefore yields a significant improvement in respect to PS.

Jenkins' expression

ST predictions fit different simulations with residual errors $\sim$ 10 %, compatible with the discrepancies between mass functions obtainable from different halo–finding algorithms. The discrepancy is however systematically greater at large masses, where the ST expression seems slightly in excess.

A further improvement was obtained in [28]. Rather than replacing the excursion set, they directly provide the expression

$$N(>M) = A \exp\left\{-|\log \sigma_M^{-1} + c|^u\right\} \tag{3.51}$$

for the cumulative halo function. Here again, the values $A \simeq 0.315$, $c \simeq 0.61$, $u \simeq 3.8$ are the best fit to a large set of simulations. This expression is expected to hold for $\log \sigma_M^{-1}$ in the interval (-1.2–1.05) and is compatible with PS, apart of the range $\log \sigma_M^{-1} >\sim 0.8$–$0.9$, where it yields a smaller $N(>M)$, apparently closer to simulation outputs.

The PINOCCHIO algorithm

Rather than putting more and more details in analytical expression, a solid prediction on mass function can be achieved by improving the very technique that we discuss in the next section, and performing a sort of *mini*–simulation, focusing on those ingredients needed to predict a mass function.

The publicly–available code PINOCCHIO [25], developed by Pierluigi Monaco

and collaborators, predicts the mass functions by using a third order Lagrangian perturbation theory.

Being perturbative, the PINOCCHIO computation progresses in a fraction of the time needed for a full N–body simulation. It then groups mass elements together into halos using an orbit–crossing condition: when the orbits of two mass elements intersect, the Lagrangian approach breaks down, as the mapping from Eulerian position to Lagrangian position becomes multi–valued. At this point, the mass elements are grouped together into a halo or a filament, in analogy with what is done when a spherical top–hat overdensity collapses to a point.

The result is shown to be adequate to identify halos and to deduce several halo properties, such as mass, angular momentum and velocity. By comparing the perturbative result with N–body outputs, it becomes evident that measures depending on collective particle properties are reliably established. On the contrary, results more sensitive to individual particle setting, e.g. spectral properties, are not so reliably reproduced.

In the next Section we shall consider the first order Lagrangian perturbation theory, which constitutes the Zeldovich approximation. This approximation is also used to create initial conditions for N–body simulations. For higher order Lagrangian approximation we suggest to consider references in [25].

Another similar fast tool to generate mock catalogs of halos is *PTHalos*, developed by Scoccimarro and Sheth [29].

Halo properties to discriminate between cosmological models.

We wish to conclude this Section by outlining the nature of results obtainable through PS–like algorithms, and their impact on model discrimination.

As a matter of fact, the shape of the mass functions found at $z = 0$ exhibit a modest dependence on the model, at least in the mass range of galaxy clusters and groups. This is partially motivated by the fact that model predictions are usually normalized onto data at $z = 0$. Furthermore, models yielding a badly wrong spectrum $P(k)$ at $z = 0$ are not taken in consideration. Slight spectral discrepancies, consistent with observational errors, are then partially smoothed out when the spectrum itself is

integrated.

Some discrepancy between mass functions, on the contrary, can be found in simulations, namely at the top–mass end. This was appreciated only when very large volumes could be simulated with a suitable dynamical range and starting from different realizations of the same model. For instance, in boxes with a side $\mathcal{O}(200\,h^{-1}\mathrm{Mpc})$, mass functions are discrepant up to 10–20 %, in a mass range starting from groups and including large clusters. Different models, in general, exhibit smaller differences. The uncertainty can be reduced only by inspecting greater and greater volumes, so to have a sample variance below model discrepancies.

Far mode discriminatory, instead, is the redshift dependence of the mass function $N(>M,z)$. A pure CDM model, e.g., is badly excluded by data concerning the amount of large clusters at $z\sim 0.5$–1 . The evolution of the mass function is too fast in these models in respect to data, which show the presence of large clusters also at rather high z .

Data are however accumulating and discrimination between much closer models will become possible. In this context it is however worth outlining that observations yield

$$\frac{dN(>M)}{d\Omega dz}=\frac{dN(>M)}{dV}\frac{dV}{d\Omega\,dz} \tag{3.52}$$

and both simulations and PS–like expressions predict $dN(>M)/dV$, rather than $dN(>M)/d\Omega dz$. When different models are compared, one must carefully gauge both the impact of dynamics and the effects of geometry, carried by the factor $dV/d\Omega dz$. For instance, if we consider two ΛCDM models with different Ω_Λ, geometry effetcs increases the discriminatory power. On the contrary, when one considers different DE state equations, the effects of geometry and evolution can be opposite, erasing one another on mass range close to the cluster mass scale.

VIII. THE ZEL'DOVICH APPROXIMATION

When analysing the evolution of a spherical overdensity, the non–linear problem simplifies and becomes integrable (within a SCDM cosmology) thanks to the reduction

of the spatial degrees of freedom to a single variable, the overdensity radius R. The analysis of non–linear stages in inhomogeneity evolution has however gone much beyond that case, still letting apart the wide field of numerical techniques.

However, even for them, initial conditions are built through the *Zel'dovich approximation* which, anyhoww, owns also an autonomous significance.

This approximation requires the use of *lagrangian* picture, instead of the usual eulerian approach: in stead of using the density fluctuations $\epsilon(\mathbf{x}) \equiv \epsilon_{\mathbf{x}}$ as a continuous set of coordinates (consider $\mathbf{x}$ as *continuous index* labeling coordinates), one takes the *comoving initial positions* of volume elements $\mathbf{x} \equiv \mathbf{q}$ as coordinates.

This advocates the passage to a discrete coordinate set, obtained through a "sampling" of volume elements through a discrete set of points. In fact, when one starts from a (quasi)–homogenous initial conditions and samples matter distribution through a wide set of points (randomly distributed, but in average equispaced), their trajectories directly exhibit the evolution of density, the birth of concentrations, the gradual forming of structures.

This scheme holds both for the Zel'dovich approximation and numerical techniques. The former case, however, is obviously simpler, but the key difference is that simulations strike at the the heart of the non linear nature of large density fluctuation evolution; on the contrary, the Zel'dovich approximation tries to exploit completely the information contained in a perturbative approach. In doing so, one discovers that the usual eulerian approach threw away a huge deal of information.

In principle, a eulerian perturbative approach holds until density fluctuations, on the relevant scales R, keep a m.s. amplitude $\sigma_R^2 = \langle \epsilon_R^2(\mathbf{x}) \rangle \ll 1$. On the contrary, the Zel'dovich approximation is often so "smart" to reach much beyond this limit.

After incensing Zel'dovich so much, let us stress the limitations of his approach. The main limit us that it applies to non–relativistic materials only and provided they are the only source involved in fluctuation growth. It is therefore suitable to study the evolution on inhomogeneities in baryons and CDM, after recombination. On the contrary, e.g., we cannot use a Zel'dovich approximation to follow photons or neutrinos, unless the latter are massive and safely non–relativistic. Using the Zel'dovich approximation for models involving a hot component needs suitable cautions; in

particular, when it is used to set initial conditions for simulations of models with massive neutrinos, there are specific problems to be solved.

Keeping however well inside the domain for its application, let us soon outline that the critical issue, enabling Zel'dovich to extend the potentiality of the perturbative approach, is the remark that the growth of fluctuations, in any site, does not yield spheroidal structures, but strongly asymmetrical shapes. Apparently, this conflicts with the emphasis PS gave to the growth of spherical overdensities and, namely, with the success their approach achieved; the point is that they "used" spherical dynamics just to create a scale independent clock. Their results could be improved when the spherical clock was replaced by an ellipsoidal one, but just because this option gives room to the possibility that there is a tiny scale dependence.

Link between lagrangian and eulerian approaches

Let us consider a point $\mathbf{x} = a\mathbf{r}$ and suppose that, in its neighborough, there exists an initial linear overdensity. The perturbative density evolution, about $\mathbf{x}$, can be shown to be consistent with the expression

$$\rho(\mathbf{r}, a)a^3 \left[1 - \alpha_1 G(a)\right] \left[1 - \alpha_2 G(a)\right] \left[1 - \alpha_3 G(a)\right] = \text{const.} \tag{3.53}$$

In fact, if $\rho(\mathbf{r}, a)$ is the background density $\rho_B(a)$, then $\alpha_i = 0$ and eq. (3.53) means that $\rho_B(a) \propto 1/a^3$. On the contrary, if $\rho(\mathbf{r}, a) > \rho_B(a)$, cosmic expansion yields a slower diluition:

$$\rho(\mathbf{r}, a) \propto \frac{1}{a^3 \prod_{i=1}^{3} \left[1 - G(a)\alpha_i\right]} \simeq a^{-3} \left[1 + G(a) \sum_i \alpha_i\right]. \tag{3.54}$$

The last passage is exact for infinitesimal α_i, being consistent with the assumption to be in a linear regime. From here, we easily derive that

$$\epsilon(\mathbf{r}, a) = \frac{\rho(\mathbf{r}, a) - \rho_B(a)}{\rho_B(a)} = G(a) \sum_i \alpha_i \tag{3.55}$$

so that $\epsilon(\mathbf{r}, a')/\epsilon(\mathbf{r}, a) = G(a')/G(a)$, showing that $G(a)$ must be the *linear growth factor* of density perturbation.

These relations do not fix the constants α_i and just puts their constant sum in relation with the fluctuation amplitude. Eq. (3.53), however, means much more: that

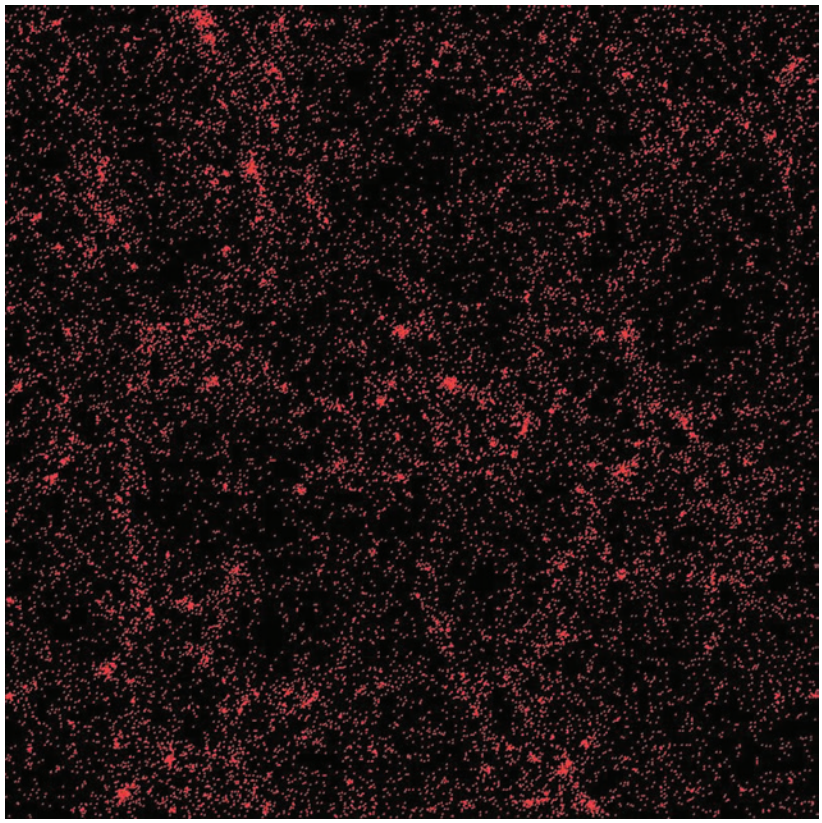


Figure 9. Filaments and knots in a bidimensional projection of a slice of a N–body simulation.

each α_i is separately constant and each factor $1 + \alpha_i G(a)$ describes the growth along one of a suitable triplet of orthogonal axes. In the next Section, we discuss how these axes are fixed.

The scale dependence of the linear growth factor is the link with the eulerian perturbative theory. Its results can be so translated into the lagrangian framework. In a pure CDM model (or when pure CDM is a fair approximation), we can set $G(a) = a$, by assuming that the present scale factor is unity. Then, eq. (3.53) reads

$$\rho(\mathbf{r}) \left[a - \alpha_1 a^2\right] \left[a - \alpha_2 a^2\right] \left[a - \alpha_3 a^2\right] = \text{const.} \tag{3.56}$$

and the discussion in the next subsection is based on this expression. There is no difficulty, however, to extend it to any case when $G(a) \neq a$ as, e.g., a ΛCDM cosmology. In the absence of an analytical expression for $G(a)$, however, results would not be so handable.

Turnaround & shell crossing

In each direction, *i.e.* for each factor, we define a *turnaround* and a *shell crossing* instant. Turnaround means that the growth of the coefficient $[a - \alpha_i a^2]$ $(i = 1, 2, 3)$ turns from increasing to decreasing. By differentiating in respect to a we see that this occurs at $a_{ta,i} = 1/2\alpha_i$. In order that $a_{ta,i} < 1$ it should be $\alpha_i > 1/2$ and this could conflict with the linearity assumption. The beauty of the lagrangian approach is that it yields reliable results even for greater α_i.

When we assumed spherical symmetry, the turnaround is also beyond the perturbative regime, the density constrast being then ~ 5.6 . Here we can consider the limiting case when $\alpha_1 \neq 0$ and $\alpha_2 = \alpha_3 = 0$; at turnaround the overall density contrast depends on the growth along the 1 axis, where we have $a - \alpha_1 a^2 = a/2$. Although a factor 2 is smaller than 5.6, also this density contrast is beyond the perturbative regime. The point is that, when results found in this way are compared with simulations, the Zel'dovich behavior is close to the numerical one.

The approximation is satisfactory even for still greater density constrasts, until we approach the *shell crossing*, i.e. the vanishing of the factor $a - \alpha_1 a^2$, occuring when the scale factor is $a_{sc} = 1/\alpha_1$; density is then divergent.

The key of success of the Zel'dovich approximation is the preliminary analysis of matter distribution, finding the simmetry axes for the deformations. The finding of such axes is a substantial extra information, of *tensor* nature, in respect to the *scalar* density fluctuation field. Gravity then indices just longitudinal actions, so that any further gravitational evolution strengthen the initially teneous deformations.

On the contrary, if we assume spherical symmetry, we implicitly make a wrong assumption; not because the density contrast behavior is roughly mistaken, but because its spatial shape is not suitably calibrated.

In the example we supposed that 2 of the 3 α_i coefficients vanish; but, in more general and less extreme cases, we however find that turnaround and shell crossing occur first along a particular axis.

Accordingly, the first structures to form will be typically flat and are therefore called *pancakes*. Pancakes often become a stable structures, whose image insists on the sites where shell crossing occurred. Before numerical techniques were so widely developped,

approximation schemes were introduced, aiming to give a quantitative basis to such picture (*e.g., the adhesion model*).

But the Zel'dovich approximation accounts also for the physics beyond the first shell–crossing. In the two other directions, after the first shell crossing, evolution proceeds, guided by the other coefficients $[a - \alpha_i a^2]$. Therefore, after a 2–dimensional surface structure has formed, evolution proceeds towards the formation of 1–dimensional filamentary structures, and eventually, towards the formation of *knots*, at the points where filaments join.

The lattice of spatial structures, observed or obtained through simulations, does not differ much from this scheme, with galaxy clusters on the sites where filaments meet and tenuous bidimensional structures between filaments. See, as an example, the 2–dimensional projection of a slice, $40\,h^{-1}$Mpc deep, obtained from a simulation in a *box* with $100\,h^{-1}$Mpc side (Figure 9).

Let us finally comment on adapting these results to a ΛCDM cosmology. In principle this just requires to use the right (numerical) $G(a)$. However, the linear growth rate, when DE becomes significant, gradually decreases. On the contrary, fluctuations which have already entered the non–linear regime do not feel such slowdown. Therefore, if one uses the ΛCDM growth rate to approximate a non linear evolution, instead of $G(a) \propto a$, it is not obvious that results will improve. The conclusion is that the performance of the Zel'dovich approximation is surely best when DE is negligible.

IX. SYMMETRY AXES & VELOCITY FIELDS IN THE ZEL'DOVICH APPROXIMATION

As previously outlined the comoving coordinates $\mathbf{r}\equiv\mathbf{q}$ are taken as lagrangian coordinates, to study the evolution of a self–gravitating system. The position of each point evolves in time according to the law

$$\mathbf{x}(t) = a(t)\,[\mathbf{q} + \delta\mathbf{q}(\mathbf{q}, t)] = a(t)\,[\mathbf{q} + G(t)\mathbf{s}(\mathbf{q})]\ , \tag{3.57}$$

by using the factorization $\delta\mathbf{q}(\mathbf{q},t) = G(t)\mathbf{s}(\mathbf{q})$, which is safe for linear $\delta\mathbf{q}$. Here a peculiar role is played by the vector function $\mathbf{s}$, whose derivatives

$$A_{ij} = \partial s_i/\partial q_j \ . \tag{3.58}$$

yield the "stress tensor". By solving then the eigenvalue equation

$$A_{ij}\hat{a}_j = \alpha^{(k)}\hat{a}_i$$

we can determine the unit eigenvectors $\hat{\mathbf{a}}^{(k)}$ and the related eigenvalues $\alpha^{(k)}$.

The vectors $\hat{\mathbf{a}}^{(k)}$ define the system of orthogonal coordinates where the expression (3.53) holds. The eigenvalues $\alpha^{(k)} \equiv \alpha_k$ $(k = 1,\ 2,\ 3)$ are the coefficients appearing in that very expression. In such reference system the matrix A_{ij} is diagonal, so that

$$A_{ij} \equiv \left|\left| \begin{matrix} \alpha_1 & & \\ & \alpha_2 & \\ & & \alpha_3 \end{matrix} \right|\right| \tag{3.59}$$

and, therefore, by dubbing q_k the components of $\mathbf{q}$ in respect to the axes $\hat{\mathbf{a}}^{(k)}$,

$$\mathbf{s} = \sum_k q_k \alpha_k \hat{\mathbf{a}}^{(k)}, \tag{3.60}$$

so that

$$\frac{\partial s_k}{\partial q_j} = \delta_{jk}\alpha_k \qquad \text{and thence} \qquad \vec{\nabla}\cdot\mathbf{s} = \sum_{i=1}^{3}\alpha_i \ . \tag{3.61}$$

Accordingly, if eq. (3.55) holds,

$$\epsilon(\mathbf{r}) = G\vec{\nabla}\cdot\mathbf{s} \ , \tag{3.62}$$

a relation that we shall use in the next subsections.

Eq. (3.57) obiously holds also when $\delta\mathbf{q}$ vanishes. This means that the density in $\mathbf{x}$ is the background density ρ_B. A deformation $\delta\mathbf{q}$ goes together with density shifting to a suitable value $\rho \neq \rho_B$. The relation between $\delta\mathbf{q}$ and $\rho - \rho_B$ is set by mass conservation, prescribing that

$$\rho_B(a)d^3q = \rho(\mathbf{r},a)d^3(q+\delta q) \ . \tag{3.63}$$

Taking into account that $q_i + \delta q_i = q_i[1 + G(t)\alpha_k]$, it is then easy to see that the determinant of the Jacobian $d(\mathbf{q} + \delta\mathbf{q})/d\mathbf{q}$ reads

$$J(\mathbf{q}, a) = \prod_i [1 + G(t)\alpha_i] \tag{3.64}$$

while $\rho(\mathbf{q}, a) = \rho_B(a)/J(\mathbf{q}, a)$. Taking then into account that $\rho_B a^3$ is constant, it follows that

$$\rho(\mathbf{q}, a) a^3 \prod_i [1 + G(t)\alpha_i] = \text{const.} \ , \tag{3.65}$$

coinciding with eq. (3.53).

Velocity fields in the Zel'dovich approximation

In the presence of density fluctuations, velocity fields will unavoidably arise.

They can be described either by $\mathbf{u} = \dot{\mathbf{r}}$ or by $\dot{\mathbf{x}}$, when assuming $\mathbf{x} = a\mathbf{r}$. It is then easy to show that the continuity equation

$$\dot{\epsilon} = \vec{\nabla} \cdot \mathbf{u} \tag{3.66}$$

yields mass conservation. In fact

$$\dot{\mathbf{x}} = H\mathbf{x} + a\dot{\mathbf{r}} \qquad \text{i.e.} \qquad a\mathbf{u} = \dot{\mathbf{x}} - H\mathbf{x} \tag{3.67}$$

while, by differentiating the relation $\mathbf{x} = a(\mathbf{q} + G\mathbf{s})$, we obtain

$$\dot{\mathbf{x}} = Ha\mathbf{x} + a\dot{G}\mathbf{s} \qquad \text{i.e.} \qquad \frac{1}{a}(\dot{\mathbf{x}} - H\mathbf{x}) = \dot{G}(t)\mathbf{s}(\mathbf{q}) \tag{3.68}$$

and

$$\mathbf{u} = \dot{G}(t)\mathbf{s}(\mathbf{q}) \qquad \text{or} \qquad \vec{\nabla} \cdot \mathbf{u} = \dot{G}(t)\vec{\nabla} \cdot \mathbf{s}(\mathbf{q}) \tag{3.69}$$

Taking then eq. (3.62) (differentiated in respect to time) into account, we obtain the continuity equation (3.66).

In a model of pure CDM, where $G \propto a$, it is then

$$\mathbf{u} = HG(t)\mathbf{s} \ . \tag{3.70}$$

If the field $\mathbf{u}$ is irrotational, then, there exist a scalar potential $\psi(q_i)$ such that

$$\mathbf{s} = \vec{\nabla}_{\mathbf{q}}\psi(\mathbf{q}) \ . \tag{3.71}$$

This potential is related to the peculiar gravitational potential of linear fluctuations, $\phi(\mathbf{q})$, trought the relation

$$\phi(\mathbf{q}) = (3/2)H^2(t)a^3(t)G(t)\psi(\mathbf{q}) \tag{3.72}$$

whose demonstration is omitted.

---

[1] R.E. Smith, J.A. Peacock, A. Jenkins, S.D.M. White, C.S. Frenk, F.R. Pearce, P.A. Thomas, G. Efstathiou and H. Couchman, 2003, MNRAS 341, 1311S

[2] Luciano Casarini, Giuseppe La Vacca, Luca Amendola, Silvio A. Bonometto, Andrea V. Macci, arXiv:1102.3877 & JCAP03(2011)026

[3] E. Komatsu et al., arXiv:1001.4538

[4] A. V. Maccio', O. Ruchayskiy, A. Boyarsky and J. C. Munoz-Cuartas, arXiv:1202.2858 [astro-ph.CO].

[5] M. Blennow, E. Fernandez-Martinez, O. Mena, J. Redondo, P. Serra, arXiv:1203.5803

[6] Damour T., Gibbons G. W. & Gundlach C., 1990, Phys.Rev., L64, 123D Damour T. & Gundlach C., 1991, Phys.Rev., D43, 3873

[7] J. -Q. Xia, Phys. Rev. D **80**, 103514 (2009) [arXiv:0911.4820 [astro-ph.CO]].

[8] L. Amendola, Phys. Rev. D **69**, 103524 (2004) [astro-ph/0311175].

[9] R. Mainini, Phys. Rev. D **72**, 083514 (2005) [astro-ph/0509318].

[10] A. Klypin, A.V. Kravtsov, O. Valenzuela, & F. Prada, ApJ, 522 (1999) 82; B. Moore, S. Ghigna, F. Governato, G. Lake, T. Quinn, J. Stadel, & P. Tozzi, ApJ 524 (1999) L19

[11] B. Moore, Nature, 370 (1994) 629

[12] R. A. Flores & J.R. Primack, ApJ 427 (1994) L1; J. Diemand, M. Zemp, B. Moore, J. Stadel, & M. Carollo, MNRAS, 364 (2005) 665; A.V. Macci'o, A.A. Dutton, F.C. van den Bosch, B. Moore, D. Potter, J. Stadel, MNRAS, 378 (2007) 55; V. Springel, J. Wang, M. Vogelsberger, A. Ludlow, A. Jenkins, A. Helmi, J.F. Navarro, C.S. Frenk, & S.D.M. White, MNRAS 391 (2008) 1685

[13] W.J.G. de Blok S.S. McGaugh, A. Bosma, A., & V.C. Rubin, ApJ (2001) 552, L23; R. Kuzio de Naray, S.S. McGaugh, & J.C. Mihos, ApJ 692 (2009) 1321; S.-H. Oh, W. J. G. de Blok, E. Brinks, F. Walter, & R.C. Kennicutt, AJ 141 (2011) 193

[14] A.V. Tikhonov, S. Gottlober, G. Yepes, & Y. Hoffman, MNRAS 399 (2009) 1611; J. Zavala, Y. P. Jing, A. Faltenbacher, G. Yepes, Y. Hoffman, S. Gottlober, & B. Catinella, ApJ 700 (2009) 1779; P.J.E. Peebles, & A. Nusser, Nature 465 (2010) 565

[15] K. Abazajian and S. M. Koushiappas, Phys. Rev. D **74**, 023527 (2006) [astro-ph/0605271]; A. Boyarsky, J. Lesgourgues, O. Ruchayskiy and M. Viel, Phys. Rev. Lett. **102**, 201304 (2009) [arXiv:0812.3256 [hep-ph]].

[16] S. Dodelson and L. M. Widrow, Phys. Rev. Lett. **72**, 17 (1994) [hep-ph/9303287]; F. Takayama and M. Yamaguchi, Phys. Lett. B **485**, 388 (2000) [hep-ph/0005214].

[17] J.R. Bond, G. Efstathiou, & J. Silk, Phys.Rev. 45 (1980) L1980; S.A. Bonometto, R. Valdarnini, Phys.Lett.A 103 (1984) 369; S. Dodelson, & L.M. Widrow, Phys.Rev. 72 (1994) L17; C.J. Hogan, & J.J. Dalcanton, Phys.Rev.D, 62 (2000) 063511 M. Viel, J. Lesgourgues, M.G. Haehnelt, S. Matarrese, & A. Riotto, Phy.Rev.D 71 (2005) 063534 K. Abazajian, & S.M. Koushiappas, Phy.Rev.D 74 (2006) 023527

[18] F. Villaescusa-Navarro, & N. Dalal, JCAP 1103 (2011) 024

[19] P. Colin, O. Valenzuela, & V. Avila-Reese, ApJ 673 (2008) 203; A.V. Macci'o, et al. , ApJ 744 (2012) L9

[20] M.G. Walker, & J. Penarrubia, ApJ 742 (2011) 20

[21] A.V. Macciò, & F. Fontanot, MNRAS 404 (2010) L16

[22] Brian G.L. & Norman M.L., ApJ 495 (1998) 80

[23] Press & Schechter, ApJ 187 (1974) 187

[24] Zentner A.R., astro–ph/061145v1 and Int.J.Mod.Phys. D16 (2007) 763-816

[25] Monaco P., Theuns T., Taffoni G., Governato F., Quinn T., Stadel J., MNRAS 364 (2002) 8

[26] Epstein R.A., MNRAS 205 (1983) 207, Peacock J.A. and Heavens A.F., MNRAS 243 (1990) 133, Bower R.J., MNRAS 248 (1991) 332

[27] Sheth R.K. and Tormen G., MNRAS 308 (1999) 119

[28] Jenkins A., Frenk C.S., White S.D.M., Colberg J.M., Cole S., Evrard A.E. , Couchman H.M.P., Yoshida N., MNRAS 321 (2001) 372

[29] Scoccimarro R. and Sheth R.K., MNRAS 329 (2002) 629

Appendix A: Zel'dovich spectrum and mass variance at horizon

Fluctuation growth (in a synchronous gauge) occurs according to the laws

$$\delta = \delta_i (a/a_i)^2 \ \text{ per } z \gg z_{eq} \ , \qquad \delta = \delta_i (a/a_i) \ \text{ per } z \ll z_{eq} \tag{3.A.1}$$

(the index $_i$ refers to the "initial" time). On the horizon, then,

$$\begin{aligned} \delta_{hor} &\simeq \delta_i (a_{hor}/a_i)^2 && \propto \delta_i a_{hor}^2 \;\; per\; z \gg z_{eq} \;, \\ \delta_{hor} &\simeq \delta_i (a_{eq}/a_i)^2 (a_{hor}/a_{eq}) && \propto \delta_i a_{hor} \;\; per\; z \ll z_{eq} \;. \end{aligned} \tag{3.A.2}$$

We assume to be known that the scale of the horizon

$$\begin{aligned} k_{hor}(a) \propto a^{-1} \;\; for\; z \gg z_{eq} \;,\; k_{hor}(a) \propto a^{-1/2} \;\; for\; z \ll z_{eq} \;, \\ and\; then: \; a_{hor}(k) \propto k^{-1} \;\; for\; z \gg z_{eq} \;,\quad a_{hor}(k) \propto k^{-2} \;\; for\; z \ll z_{eq} \;. \end{aligned} \tag{3.A.3}$$

Putting together eqs. (3.A.2) and (3.A.3), we obtain

$$\begin{aligned} \delta_{hor}(k) \;\propto\; \delta_i(k) a_{hor}^2(k) \;\propto\; \delta_i(k) k^{-2} \;\; for\; z \gg z_{eq} \;, \\ \delta_{hor}(k) \;\propto\; \delta_i(k) a_{hor}(k) \;\propto\; \delta_i(k) k^{-2} \;\; for\; z \ll z_{eq} \;; \end{aligned} \tag{3.A.4}$$

so that there is no difference between scales greater or smaller than the horizon at equality. Accordingly, the spectrum on the horizon

$$P_{hor}(k) \propto |\delta_{hor}(k)|^2 = |\delta_i(k)|^2 \, k^{-4} = A \, k^{n_s - 4} \tag{3.A.5}$$

and, by replacing this scale dependence in the expression for the mass variance

$$\sigma^2_{R_{hor}} = \frac{1}{2\pi^2} \int_0^\infty dk \; k^2 \, P_{hor}(k) \, W(kR_{hor}) \;. \tag{3.A.6}$$

we obtain

$$\sigma^2_{R_{hor}} = \frac{A}{2\pi^2} \int_0^\infty dk \; k^{n_s - 2} W(kR_{hor}) = R_{hor}^{1-n_s} \frac{A}{2\pi^2} \int_0^\infty dx \; x^{n_s - 2} W(x) \;. \tag{3.A.7}$$

Here $W(kR)$ is a *filter function*, smoothedly setting a upper limit to integration at $k \sim 1/R$. Different filters yield slightly different values for the last integral factor, but the scale dependence $\propto R_{hor}^{1-n_s}$ is however the same and, clearly, $\sigma_{R_{hor}}$ is scale invariant for $n_s = 1$.

Appendix B: DM and baryon fluctuations after recombination

Let us consider a time t_i, when recombination is over, and let then DM and baryon fluctuations have amplitudes $\delta_{c,i}$, $\delta_{b,i}$ and amplitude derivatives $\dot{\delta}_{c,i}$, $\dot{\delta}_{b,i}$, as fixed by

previous evolution. On the basis of the equations used in previous Chapter, we can easily see that fluctuations must now obey the equations

$$\ddot{\delta}_c + \frac{4}{3t}\dot{\delta}_c - \frac{2}{3t^2}(\Omega_c\delta_c + \Omega_b\delta_b) = 0$$

$$\ddot{\delta}_b + \frac{4}{3t}\dot{\delta}_b - \frac{2}{3t^2}(\Omega_c\delta_c + \Omega_b\delta_b) = 0 \ . \tag{3.B.1}$$

If we take then the difference of these equations and the sum of the former one $\times\Omega_c$ plus the latter one $\times\Omega_b$, we obtain

$$\ddot{D} + \frac{4}{3t}\dot{D} = 0 \ , \qquad \ddot{S} + \frac{4}{3t}\dot{S} - \frac{2}{3t^2}S = 0 \tag{3.B.2}$$

(remind: $\Omega_c + \Omega_b = 1$). Here

$$D = \delta_c - \delta_b \ , \qquad S = \Omega_c\delta_c + \Omega_b\delta_b \tag{3.B.3}$$

and $S = \delta\rho/\rho$ is the total fluctuation in non relativistic matter. These two equations can be integrated by assuming their solutions to be $\propto t^\alpha$ and working out α. We then have

$$S = A(t/t_i)^{1/3} + B(t/t_i)^{-1} \ , \qquad D = V(t/t_i)^{-1/3} + Z \ ; \tag{3.B.4}$$

A, B, V, Z should then be worked out from the values of $\delta_{c,i}$, $\delta_{b,i}$, $\dot{\delta}_{c,i}$, $\dot{\delta}_{b,i}$.

For generic initial conditions, after a reasonable time has elapsed, only the term $A(t/t_i)^{1/3}$ will keep relevant in S. What can then be said, about the two terms in D, is that, after a reasonable time, the former one will give no significant contribution, while the latter one is still significant and keeps $\sim \delta_{c,i} - \delta_{b,i}$. For instance, when δ_c and δ_b have increased by a factor 100, this difference keeps $\sim \mathcal{O}(1\,\%)$ or just little less than so. Its value cannot be fixed if we prescind from pre–recombination evolution, but it is a value gradually becoming more and more negligible.

Quite in general, we can however define

$$\alpha_+ = A/(\Omega_c\delta_{c,i} + \Omega_b\delta_{b,i}) \ , \quad \zeta_+ = Z/(\delta_{c,i} - \delta_{b,i}) \tag{3.B.5}$$

($0 < \alpha_+, \zeta_+ < 1$) so that the behavior of fluctuations is

$$\delta_c = \alpha_+(\Omega_c\delta_{c,i} + \Omega_b\delta_{b,i})(t/t_i)^{2/3} + \zeta_+(\delta_{c,i} - \delta_{b,i}) \ ,$$

$$\delta_b = \alpha_+(\Omega_c\delta_{c,i} + \Omega_b\delta_{b,i})(t/t_i)^{2/3} - \zeta_+(\delta_{c,i} - \delta_{b,i}) \tag{3.B.6}$$

Then, although the solution cannot prescind from details of previous evolution, our ignorance is completely hidden in the values –comprised between 0 and 1– of the parameters α_+ and ζ_+.

Appendix C: Evolution of a spherical overdensity

Let us assume that, at a time t_i ($\gg t_{rec}$), there is a spherical overdensity with radius R_i, whereinside density is $\rho_i > \bar{\rho}_i$ ($\bar{\rho}_i$: average cosmic density at the time t_i); its amplitude $\delta = \rho_i/\bar{\rho}_i - 1$ is assumed to be $\ll 1$ initially. Let us also define the *density contrast* Δ_i and the *density parameter* inside of the sphere

$$\Delta_i = \Omega_i = \rho_i/\bar{\rho}(t_i) \ . \tag{3.C.1}$$

Staring from the rime t_i the scale factor grows according to the equation

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{4\pi G}{3}\bar{\rho}_i\left(\frac{a_i}{a}\right)^3 = H_i^2\left(\frac{a_i}{a}\right)^3 \tag{3.C.2}$$

which is also an expression of the conservation of mechanical energy (vanishing, in average) inside of a self–gravitating sphere of radius a, when its initial density coincides with the average cosmic density (and it will keep so at later times).

An analogous equation holds for the radius R of the overdense sphere. In this case, however, the total energy is not zero and $\bar{\rho}_i$ must be replaced by ρ_i. All that leads to the equation

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{4\pi G}{3}\rho_i\left(\frac{R_i}{R}\right)^3 - \frac{\kappa}{R^2} \ . \tag{3.C.3}$$

Let us then add an ingredient essential to obtain simple results: We require that, at the initial time, the sphere expands with the Hubble flow, $(\dot{R}/R)_{t_i} = H_i$. At the initial time t_i, it will be

$$H_i^2 = \frac{4\pi G}{3}\bar{\rho}_i\Omega_i - \frac{\kappa}{R_i^2} = H_i^2\Omega_i - \frac{\kappa}{R_i^2}$$

and eq.(3.C.3) becomes

$$\left(\frac{\dot{R}}{R}\right)^2 = H_i^2\Omega_i\left(\frac{R_i}{R}\right)^3 + H_i^2(1-\Omega_i)\frac{R_i^2}{R^2} \tag{3.C.4}$$

or

$$\left(\frac{\dot{R}}{R_i}\right)^2 = H_i^2\left(\Omega_i\frac{R_i}{R} + 1 - \Omega_i\right) \tag{3.C.5}$$

This equation coincides with the one fulfilled by the scale factor, in a closed model of the Universe, when the cosmic "substance" has state parameter $w = 0$. This equation owns an exact integral in parametric form:

$$R(\theta) = R_i\frac{\Omega_i}{2(\Omega_i-1)}(1-\cos\theta)\ , \quad H_i t(\theta) = \frac{\Omega_i}{2(\Omega_i-1)^{3/2}}(\theta-\sin\theta)\ . \tag{3.C.6}$$

that we also gave in eq. (3.11).

PROOF

$$Infact: \qquad \frac{dR}{d\theta} = R\frac{\sin\theta}{1-\cos\theta}\ , \quad \frac{d\,(H_i t)}{d\theta} = \frac{R}{R_i(\Omega_i-1)^{1/2}}$$

$$wherefrom: \qquad \frac{\dot{R}^2}{H_i^2} = \frac{\sin^2\theta}{(1-\cos\theta)^2}R_i^2(\Omega_i-1)$$

$$and,\ being: \quad \frac{1-\cos^2\theta}{(1-\cos\theta)^2} = \frac{2}{1-\cos\theta} - 1 = \frac{R_i}{R}\frac{\Omega_i}{\Omega_i-1} - 1$$

$$it\ follows\ that: \qquad \frac{\dot{R}^2}{H_i^2} = R_i^2\left(\Omega_i\frac{R_i}{R} - \Omega_i + 1\right)$$

and this verifies that the expression (3.C.6) is the integral of eq. (3.C.5).

COMPARISON WITH THE LINEAR SOLUTION

The generic linear solution reads

$$\delta(t) = \delta_{i,+}(t/t_i)^{2/3} + \delta_{i,-}(t/t_i)^{-1} \tag{3.C.7}$$

with

$$t\,\dot{\delta}(t) = (2/3)\delta_{i,+}(t/t_i)^{2/3} - \delta_{i,-}(t/t_i)^{-1} \tag{3.C.8}$$

In our case $\dot{\delta}(t_i) = 0$, thanks to the assumption that the initial expansion occurs within the Hubble flow. It will then be $\delta_{i,-} = (2/3)\delta_{i,+}$ e

$$\delta_{i,+} = (3/5)\delta_i$$

At $t \gg t_i$, the linear solution will then be

$$\delta(t) \simeq (3/5)\delta_i (t/t_i)^{2/3} \ . \tag{3.C.9}$$

Let us compare this expression with density evolution within the spherical overdensity:

$$\delta(t) = \Delta(t) - 1 = [R(t)/a(t)]^3 - 1 \tag{3.C.10}$$

with $R(t)$ and $a(t)$ given by eqs.(3.C.6) and (3.C.2). In general such comparison cannot be achieved analytically; but, for some significant configurations it is easier; in the text of the Chapter, *e.g.*, we did it for the top expansion time.

Appendix D: Fluctuation spectrum and mass variance

The smoothing on a scale R of the fluctuation field can be formally done by performing the integral

$$\epsilon_R(\mathbf{x}) = \int_0^\infty du\ u^2 \int_{4\pi} d\Omega\ W_R(|\mathbf{u} - \mathbf{x}|)\epsilon(\mathbf{u}) \ . \tag{3.D.1}$$

Here we use a top–hat filter function, reading

$$W_R(y) = 1/V_R\ for\ y < R, \quad = 0\ \ elsewhere, \tag{3.D.2}$$

with $V_R = 4\pi R^3/3$.

In the expression (3.D.1) we can however use different filter functions. Most of the rest of this Appendix can be easily extended to other filters.

Let us now set $\mathbf{s} = \mathbf{u} - \mathbf{x}$ and re–write eq.(3.D.1) by performing a Fourier expansion of the field. We obtain that

$$\epsilon_R(\mathbf{x}) = \int_0^\infty ds\ s^2 \int_{4\pi} d\Omega\ W_R(s) \int d^3k\ \delta(k)\ e^{i\mathbf{k}\cdot(\mathbf{x}+\mathbf{s})} =$$

$$= \int d^3k\ \delta(k)\ e^{i\mathbf{k}\cdot\mathbf{x}} \int_0^\infty ds\ s^2 \int_{4\pi} d\Omega\ W_R(s)\ e^{i\mathbf{k}\cdot\mathbf{s}} = \int d^3k\ \delta_R(k)\ e^{i\mathbf{k}\cdot\mathbf{x}}\ . \quad (3.D.3)$$

Here we use the smoothed spectrum

$$\delta_R(k) = \delta(k)\tilde{W}(kR) \quad (3.D.4)$$

including the transform

$$\tilde{W}(kR) = \int_0^\infty ds\ s^2 W_R(s) \int_{4\pi} d\Omega\ e^{i\mathbf{k}\cdot\mathbf{s}} = \frac{2\pi}{V_R} \int_0^R ds\ s^2 \int_{-1}^1 d(\cos\theta)\ e^{iks\cos\theta} =$$

$$= \frac{3}{2} \int_{-1}^1 d(\cos\theta) \int_0^1 ds\ s^2\ e^{ikRs\cos\theta} = \frac{3}{(kR)^3}[\sin(kR) - kR\ \cos(kR)]\ . \quad (3.D.5)$$

Accordingly, the smoothed spectrum is obtainable by multiplying $P(k)$ by the transformed filter

$$\tilde{W}(kR) = \frac{3}{(kR)^3}[\sin(kR) - kR\ \cos(kR)]$$

which clearly tends to vanish for $kR \ll \pi$, so damping the short wavelength spectral components, with $k \ll 2\pi/R$.

Let us now try to apply the definition

$$\sigma_R^2 = \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3x\ \epsilon_R^2(\mathbf{x})\ , \quad (3.D.6)$$

of mass variance, by averaging over a large volume $\mathcal{V}$, that we shall assume to be a cube of side L.

The expansion (3.D.3) can then be written as a sum

$$\epsilon_R(\mathbf{x}) = \sum_{\mathbf{k}} \delta_R(k) e^{i\mathbf{k}\cdot\mathbf{x}} \quad with \quad \mathbf{k} = \mathbf{n}\ 2\pi/L \quad (\mathbf{n}:\ integer\ component\ vector)\ . \quad (3.D.7)$$

that shall be replaced in eq. (A4.6), so yielding

$$\sigma_R^2 = \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3r\ \sum_{\mathbf{k}} \delta_R(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \sum_{\mathbf{k}'} \delta_R^*(\mathbf{k}') e^{-i\mathbf{k}'\cdot\mathbf{r}} =$$

$$= \sum_{\mathbf{k},\mathbf{k}'} \delta_R(k)\delta_R^*(k') \frac{1}{\mathcal{V}} \int_{\mathcal{V}} d^3r\ e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}} = \sum_{\mathbf{k},\mathbf{k}'} \delta_R(k)\delta_R^*(k')\delta_{\mathbf{k}\,\mathbf{k}'} = \sum_{\mathbf{k}} |\delta_R(\mathbf{k})|^2\ .$$

We replaced here the integral $\mathcal{V}^{-1}\int_{\mathcal{V}} d^3r\ \exp\{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{r}\}$ by the Dirac distribution $\delta_{\mathbf{k}\mathbf{k}'}$, as is formally correct when $\mathcal{V}\to\infty$. This is almost exact as $\mathcal{V}$, by definition, ought to be extremely large. However, the limit must now be taken, so that the last equation yields exactly

$$\sigma_R^2 \;=\; (2\pi)^{-3}\int_0^\infty d^3k\ |\delta(k)|^2\tilde{W}^2(kR) \;=\; \frac{1}{2\pi^2}\int_0^\infty dk\ k^2\ |\delta(k)|^2\tilde{W}^2(kR)\ , \quad (3.D.8)$$

an expression we used in the text. This result is independent from the form of the filter function.

Dynamics of Dark Matter: Linear and Non-linear Effects

J. A. de Freitas Pacheco[1,2]

[1] *Universidade Federal do Espirito Santo, Vitoria*

[2] *University of Nice-Sophia Antipolis – Observatoire de la Côte d'Azur*

I. INTRODUCTION

In our days, the comprehension of the nature of dark matter is one of the greatest challenges for physicists and astrophysicists. Gravitational effects at different scales suggest the presence of some unknown form of matter in the universe, corresponding to an amount of about 23% of the so-called critical density ($\rho_{\rm crit} = 3H_0^2/8\pi G = 1.89 \times 10^{-29}$ h^2 g cm^{-3}, where h is the present value of the Hubble parameter in units of 100 km/s/Mpc). In fact, the abundance of elements like ^{2}H, ^{4}He, ^{3}He and ^{7}Li, "cooked" in the early phases of the universe, depends essentially on the baryon-to-photon ratio η. For a universe evolving adiabatically, such a ratio remains constant and the observed abundance of the primordial elements permits to estimate η. Since the present relic photon density is known (about 410 cm^{-3}), the baryon density can be computed. The resulting value corresponds to about 4% of the critical density, indicating clearly that baryons under any form (stars, cold or hot gas) cannot explain the gravitational effects (which will be analysed later) observed in scale of galaxies, groups and clusters.

The only particle of the standard model having abundance large enough to explain the observations is the neutrino (density of about 112 cm^{-3} per flavour). In this case, in order to satisfy the constraint of a cosmic mass density of about 23%, the average mass of the different neutrino flavours should be of the order of 4.3 eV. In fact, neutrino oscillation experiments, which measure mass squared differences (see [1] for a review) suggest lower values. If considered as a Majorana particle, the Heidelberg-Moscow experiment (whose data analysis is contested by some other experimentalists) on double β-decay of ^{76}Ge, indicate a mass of 0.33$\pm$0.22 eV, one order of magnitude less than that required to explain the "observed" amount of dark matter. In reality, there

is another difficulty with the neutrino as a dark matter candidate. Neutrinos decouple relativistically and if they had a mass of 1.0 eV, they stream freely until the period of matter-radiation equality ($z \approx 3300$) up to distances of about 14 Mpc, corresponding to a mass scale of about $4 \times 10^{14}\ M_{\odot}$. Thus, neutrinos erase all fluctuations in the galactic scale, favouring an up-down scenario of structure formation.

The absence of relic candidates of the standard model pushed physicists to examine extensions of that model. In particular, models including supersymmetry (SUSY), lead to new candidates like s-neutrinos, gravitinos or photinos among others. Presently, the most plausible candidate is the neutralino χ, which is the lightest supersymmetric particle. However, there is an embarrassing problem: up to now, no signal of supersymmetry has been seen in the Large Hadron Collider (LHC) of CERN. The decay of B-mesons (including $b\bar{b}$ quarks) shows no indication of "exotic" particles like the chargino or the neutralino. The situation is more delicate when direct search experiments are considered. In these experiments, a DM particle collides elastically with a nucleus, which suffers recoil and dissipates energy that is measured by the detector. The orbital velocity of the Earth combines with that of the Sun around the galactic centre and, consequently, a modulation of the signal amplitude with a period of one year should be seen. In fact, this has been claimed for more than one decade by the team of the experiment DAMA/LIBRA and now there are indications that a similar signal has been also detected by the experiments CoGeNT and CRESS-II [2–4]. However, these findings are not confirmed by the more sensitive experiment XENON100 [5].

The recent claim by two experiments (ATLAS and CMS) running at CERN, that a signal at the energy of about 125 GeV could be associated to the Higgs boson, permits to restrict better the mass range of dark matter particles. This value, introduced in SUGRA models and combined with the upper limits on the dark matter particle-nucleon cross section derived from XENON100 and the dark matter abundance derived from the seven years WMAP data, constrain the mass of the lightest particle of the "Minimal Supersymmetric Model" (MSSM) to be $m_{\chi} > 160$ GeV [6–8]. Interestingly, the non-detection of gamma rays originated from the $\chi\bar{\chi}$ annihilation in the centre of M87, either at energies of 100 MeV or 1 GeV by FERMI-LAT, implies also that

$m_\chi > 100$ GeV [9]. On the other hand, there are recent claims concerning the detection of a gamma ray line at approximately 130 GeV from the galactic centre [10, 11] and from several clusters of galaxies [12] in FERMI-LAT data. Gamma-ray lines can be produced during the annihilation of dark matter particles through two channels: either by two-photon emission by the reaction $\chi + \bar{\chi} \longrightarrow \gamma + \gamma$ or by a single photon emission by the reaction $\chi + \bar{\chi} \longrightarrow Z^0 + \gamma$ citeDeFreitasPacheco:2005et. In the former, the photon energy is $\epsilon_\gamma \approx m_\chi$, whereas in the latter the photon energy is given by $\epsilon_\gamma = (m_\chi^2 - m_{Z^0}^2/4m_\chi)$. For a photon energy of 130 GeV, the corresponding neutralino masses are 130 GeV and 144 GeV for the first and the second channel respectively, consistent with the constraints mentioned previously. A different origin for this possible γ-ray line was proposed by Bergstrom [14], who assumed the existence of right-handed neutrinos having a mass of about 135 GeV. According to his calculations, the FERMI-LAT feature at 130 GeV can be quite well explained by the following processes: i) broad internal bremsstrahlung ($\nu_R + \nu_R \longrightarrow l^+ + l^- + \gamma$), producing a bump around 120 GeV; ii) a two-photon line at 135 GeV issued from the process $\nu_R + \nu_R \longrightarrow \gamma + \gamma$ and finally, iii) a γ-line around 119.6 GeV resulting from the process $\nu_R + \nu_R \longrightarrow Z^0 + \gamma$. Further observations must decide about the reality of such a feature and its true origin.

Although the DM theory predicts correctly different aspects of the large structure of the universe and, in particular, the formation of galaxies, some problems persist. Cosmological simulations predict too much dark matter at the centre of galaxies and a large number of satellites, which are not observed. Different issues to these problems have been proposed. Some of them are quite drastic, since they claim a failure of General Relativity at cosmological scales, requiring alternative gravitational theories.

In these lectures, it is assumed that dark matter is a reality (thus, General Relativity is a "good" theory) and constituted by weakly interacting particles not issued from the standard model. In the following sections the evidence for the existence of dark matter in different scales is reviewed as well as its dynamical behaviour at different evolutionary phases of the universe.

II. EVIDENCE OF DARK MATTER IN THE UNIVERSE

A. Clusters of galaxies

Historically, since 1933 the presence of "non-luminous" matter in a gravitational system was suspected by the Swiss astronomer Fritz Zwicky. He noticed that the mean velocity of galaxies in the Coma cluster was too high for the measured (total) luminosity, under the assumption that the system was in dynamical equilibrium. Zwicky assumed a certain relation between the mass and the luminosity for the galaxies of the cluster and the virial relation to compute the expected mean galaxian velocity that was lower than the observed value.

As we will see, the basic assumption involved in mass estimates of astronomical systems is that of **dynamical equilibrium**. Most authors consider that Coma is the archetype of a "relaxed" cluster, an idea based mainly on its quasi-spherical morphology. However, substructures have been identified in the cluster [15], weakening this argument. Even so, let us assume that the cluster is relaxed and that galaxies track the **total** gravitational potential. For a spherical system in which the 1D-velocity dispersion is constant, using the Jeans equation (the derivation of this equation will be discussed later) the total mass inside the radius R is

$$M(\leq R) = -\frac{\sigma_r^2 R}{G} \left. \frac{d\ln n(r)}{d\ln r} \right|_R \; , \tag{4.1}$$

where $n(r)$ is the density of galaxies. This quantity can be obtained by inversion of an Abel integral involving the projected density of galaxies $\Sigma(a)$ at a given projected distance "a" from the centre, an observed quantity. Figure 1 shows the projected density of galaxies derived from Zwicky counts.

The de-projected density of galaxies is given by

$$n(r) = -\frac{1}{\pi} \int_r^\infty \frac{d\Sigma(a)}{da} \frac{da}{\sqrt{a^2 - r^2}} \; . \tag{4.2}$$

Combining eqs. (4.1) and (4.2) and using $\sigma_r = 1008$ km s^{-1}, one obtains for the total mass inside R = 2.5 Mpc the value $M = 1.48 \times 10^{15}\ M_\odot$. Since the luminosity of the cluster is about $L_B = 1.0 \times 10^{13}\ L_\odot$, the corresponding mass-to-luminosity

ratio is $M/L_B \approx 148$ (in solar units), indicating a substantial amount of non-baryonic matter in the cluster (it should be mentioned that the mass-to-luminosity ratio in the B filter varies from 4 for spiral galaxies up to 8 for elliptical ones). From the derivative of equation (4.1), one obtains for the central DM density a value of 2.9 $M_\odot$ pc^{-3}. Galaxies are not the only tracer of the potential. Hot gas, detected by its

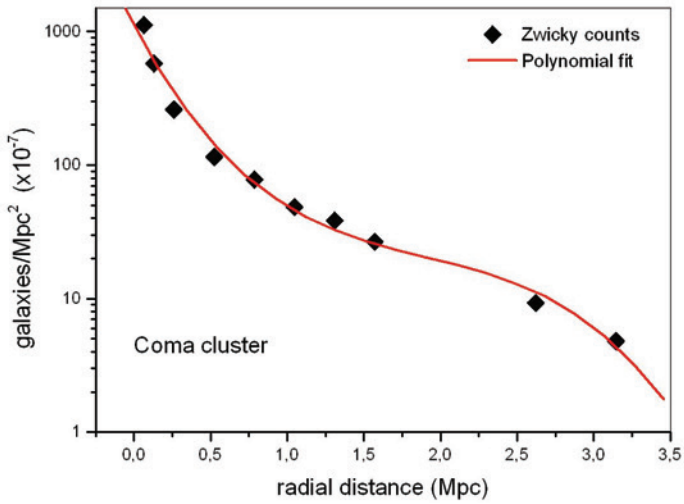


Figure 1. Projected galaxy density as a function of the radial distance. Data from Zwicky and solid (red) curve represents a polynomial fit.

X-ray emission can also be used to track dark matter. If the hot intra-cluster gas is in hydrostatic equilibrium and has a uniform temperature distribution, then the total mass inside a radius R is given by

$$M(\leq R) = -\frac{kT}{\mu m_p}\frac{R}{G}\left.\frac{d\ln n_{gas}}{d\ln r}\right|_R , \tag{4.3}$$

where μ is the mean molecular weight, m_p is the proton mass and n_{gas} is the hot gas particle density. Combining eqs. (4.1) and (4.3) gives $n_{gas} \propto n^\beta$ where $\beta = (\mu m_p/kT)\sigma_r^2$. These last relations define the so-called "β-model" for the intra-cluster gas. Using the following ansatz for the density profile of the hot gas

$$n_{gas} = \frac{n_0}{\left[1+\left(\frac{r}{r_c}\right)^2\right]^\beta} , \tag{4.4}$$

one obtains for the X-ray brightness distribution (remember that the emissivity due to bremsstrahlung is proportional to the square of the gas density)

$$S(a) = \frac{S_0}{\left[1 + \left(\frac{a}{r_c}\right)^2\right]^{2\beta - 1/2}} \ . \tag{4.5}$$

In the above equations, n_0 is the central gas density, r_c is the "core" radius, S_0 is the central X-ray brightness and a is the projected distance to the centre of the cluster.

Let us take as an example the cluster A1795. The X-ray data from **Chandra** (see Figure 2) can be fitted by eq. (4.5), permitting an estimation of the parameters β, r_c and S_0. Then, the parameters β and r_c can be inserted into eq. (4.4), which permits to compute the total mass from eq. (4.3). The measured average temperature of the hot gas is 4.0 keV, the resulting mass inside a radius of 2.0 Mpc is $5.7 \times 10^{14}\ M_\odot$ and the central DM density is about 0.035 $M_\odot\ pc^{-3}$. Again, when compared with the mass of baryons under the form of hot gas and under the form of stars, we conclude that a substantial amount is "dark", i.e., not seen in visible light or in X-rays.

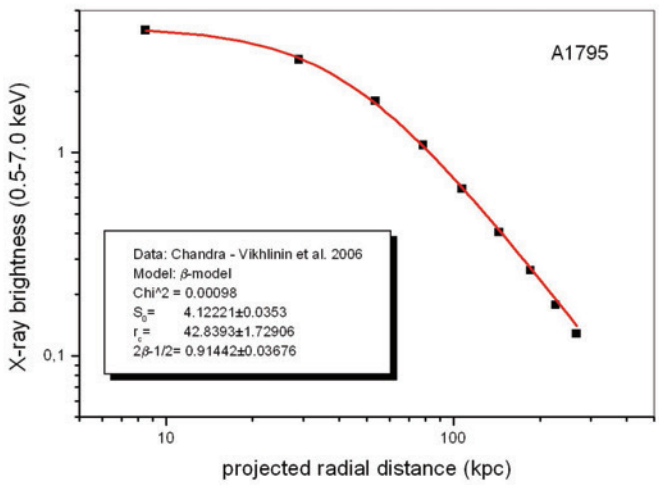


Figure 2. X-ray brightness profile of A1795.

B. Groups of Galaxies

If clusters permit to probe the presence of dark matter in scales of 1-2 Mpc, groups of galaxies, including a few tens of objects, permit to investigate dark matter effects at scales of 100-400 kpc. Groups are generally supposed to be in equilibrium in order that the virial be satisfied. Here we assume that this is not generally the case and that the global dynamics of the system is described by the Jacobi identity.

Supposing that the system is homogeneous, i.e., the density of galaxies is constant inside a fictitious spherical volume of radius R, the probability $P(a)$ of finding a galaxy at a projected distance a from the centre depends only of the effective depth at such a distance, i.e.,

$$P(a) = 2K \int_0^{\sqrt{R^2-a^2}} dz = 2K\sqrt{R^2 - a^2}\,, \tag{4.6}$$

where $K = 3/4\pi R^3$ is a normalization constant such as

$$\int_0^R 2\pi P(a) da = 1\,. \tag{4.7}$$

Then, the mean projected distance, the mean inverse projected distance and the squared projected distance, which are the observable quantities, can be computed and they are given respectively by

$$\langle a \rangle = \frac{3\pi}{16} R\,; \qquad \left\langle \frac{1}{a} \right\rangle = \frac{3\pi}{4R}\,; \qquad \langle a^2 \rangle = \frac{12}{15} R^2\,. \tag{4.8}$$

The Jacobi identity describing the dynamical state of the system is

$$\frac{1}{2}\frac{d^2 I}{dt^2} = 2T + W\,, \tag{4.9}$$

where I is the moment of inertia of the system, T is the kinetic energy and W is the gravitational potential energy. Notice that in dynamical equilibrium, the left side of eq. (4.9) is null and the virial relation is recovered. Using the relations in (4.8), the different terms of the Jacobi identity can be written as

$$W = -\frac{3GM^2}{5R} = -\frac{4}{5\pi} GM^2 \left\langle \frac{1}{a} \right\rangle\,, \tag{4.10}$$

$$T = \frac{1}{2}\sum_{i}^{N} m_i v_i^2 = \frac{3}{2} M \sigma_z^2 \, , \tag{4.11}$$

and

$$I = \frac{3}{5} M R^2 = \frac{3}{4} M \langle a^2 \rangle \, . \tag{4.12}$$

An isotropic distribution of velocities was supposed when computing the kinetic energy. Now, make the approximation

$$\frac{1}{2}\frac{d^2 I}{dt^2} \approx \pm \frac{I}{2\tau^2} \approx \pm \frac{81\pi^2}{2048} \frac{\langle a^2 \rangle}{\langle a \rangle^2} M \sigma_z^2 \, , \tag{4.13}$$

where we have introduced the crossing time defined by

$$\tau = \frac{R}{\sqrt{3}\sigma_z} = \frac{16}{3\sqrt{3}\pi} \frac{\langle a \rangle}{\sigma_z} \, . \tag{4.14}$$

Replacing all these relation into eq. (4.9) and solving for the mass M one obtains

$$M = \frac{15\pi}{4G} \sigma_z^2 \left\langle \frac{1}{a} \right\rangle^{-1} \left[1 \pm \frac{27\pi^2 \langle a^2 \rangle}{2048 \langle a \rangle^2} \right] \, . \tag{4.15}$$

In the above equation, the positive sign corresponds to an expanding group while the negative sign indicates a contracting group.

As an example, consider the group NGC 383, constituted by 9 bright galaxies whose total luminosity is $L_B = 3.6 \times 10^{11} L_{B\odot}$. The velocity dispersion of galaxies is $\sigma_z = 504$ km/s and the different averages involving projected distances are: $\langle a \rangle = 105$ kpc; $\langle a^2 \rangle = 14167$ kpc^2; $\langle 1/a \rangle = 0.022$ kpc^{-1}. From these, we derived for the effective radius of the system R = 178 kpc and, from eq. (4.13), a mass $M = 3.14 \times 10^{13} (1 \pm 0.168)\ M_\odot$. The group is probably is collapsing, so its probable mass is $M = 2.6 \times 10^{13}\ M_\odot$, implying a mass-to-luminosity ratio $M/L \approx 73$, which indicates again the presence of dark matter in the system.

C. Binary Galaxies

Binary galaxies represent an intermediate scale between clusters and groups at which dark matter can be probed. Typically, the separation of these interacting

systems is in the range 50-80 kpc. One of the major difficulties to study these systems is that catalogs are often contaminated by "optical" pairs, i.e., galaxies that are not physically connected but are close only by projection effects. When tidal distortions are not seen in the morphology of the galaxies, it is difficult to conclude if a pair is real or "optical".

Here we follow the statistical approach by de Freitas Pacheco & Junqueira [16] to analyse a given sample of binaries. The observed quantities are the velocity difference in the line-of-sight V_z and the projected separation of the pair r_p. Introduce now the dynamical quantity $q = V_z^2 r_p$. If the galaxies follow elliptical orbits, a simple exercise gives

$$q = G(M_1 + M_2)F(\theta, i, \varpi, e) \ , \tag{4.16}$$

where M_1, M_2 are the masses of the galaxies and the orbital function F depends on the phase angle θ, the inclination of the orbit i, the longitude of the periastron ϖ and on the eccentricity e. The orbital function is given explicitly by the equation

$$F = \frac{\sin^2 i \, (\sin\theta + e \sin\varpi)^2 \left(\cos^2\theta \cos^2 i + \sin^2\theta\right)^{1/2}}{1 + e\cos(\theta - \varpi)} \ . \tag{4.17}$$

We further assume that the total mass of the pair is proportional to its total luminosity, i.e., $M \propto fL$ and that the mass-to-luminosity ratio is the same for a given sample of pairs. Here we will discuss the analysis of a sample of 233 pairs prepared by Karachentsev. Under these conditions, from (4.16), we can write for each member of the catalogue

$$\log q_i = \log f + \log(GL_i F_i) \ . \tag{4.18}$$

Using a Monte Carlo procedure, the distribution of the quantity $\log(GLF)$ is calculated by assuming: i) a log-normal distribution for the luminosity whose parameters were obtained from the catalogue data (median equal to $\log L_B = 10.80 L_\odot$ and dispersion equal to $\sigma_{\log L} = 0.53$); ii) the longitude of the periastron and the inclination of the orbit are random variables, varying uniformly in their domains of validity; iii) the orbits fill the phase space, thus the distribution probability of eccentricities is $P(e)de = 2ede$;

iv) for elliptical orbits, the distribution probability of the phase angle θ is

$$P(\theta)d\theta = \frac{(1-e^2)^{3/2}}{1+e\cos(\theta-\varpi)}d\theta \; . \tag{4.19}$$

On the other hand, the distribution of the quantity $\log q$ can be derived directly from the catalogue data. The best fit of both distributions gives the mean mass-to-luminosity ratio f of the sample. This is shown in figure 3. The best fit gives

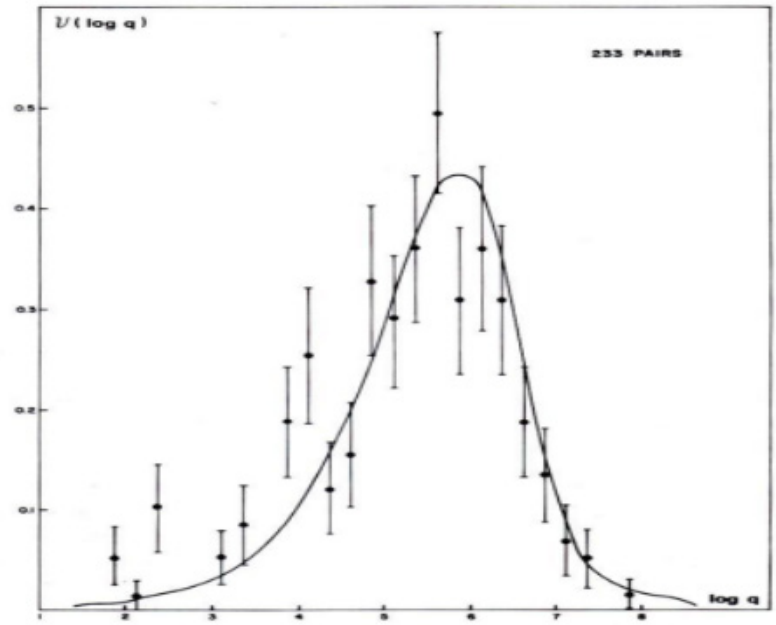


Figure 3. Simulated and observed distributions for binary systems.

$M/L = 25$ while the expected value, taking into account that 80% of the galaxies in the catalogue are early type is $M/L = 5.6$, which puts in evidence the presence of dark matter in scales of the order of 47 kpc, the mean projected separation of the pairs.

D. Galaxies – Flat Rotation Curves

If light traces matter, the rotation curves of galaxies should decline for distances beyond 10-15 kpc from the centre. However, this is not the case because the large amount of data collected by Vera Rubin (and others) since the seventies indicate that most of galaxies have flat rotation curves.

Rotation curves are generally modelled by the following procedure. From the observed brightness profile, the mass density can be derived by inversion techniques and by assuming a constant (or variable) mass-to-light ratio. Once the density is known, the solution of the Poisson equation gives the potential that introduced into Jeans equations permits an evaluation of the velocity profile. Another strategy consists to adopt a self-consistent potential-density pair whose parameters are chosen in order to fit the brightness profile of the galaxy. The latter approach was adopted by Ortega & de Freitas Pacheco [17] to describe the kinematics of Sc galaxies. Gas-rich spirals have the advantage that dark matter can be tracked not only by stellar motions but also by the gas (21cm line) that can be detected up to distances of about 25-30 kpc from the centre. The adopted potential-density pair is that derived by Satoh [18], which is free of singularities and given by the equations (in cylindrical coordinates)

$$\phi_s(r,z) = -\frac{GM_s}{\left\{r^2+z^2+a\left[a+2\left(z^2+b^2\right)^{1/2}\right]\right\}^{1/2}}\ , \tag{4.20}$$

and

$$\rho_s(r,z) = \frac{ab^2M_s}{4\pi}\frac{\left\{r^2+z^2+\left[a+3\left(z^2+b^2\right)^{1/2}\right]\left[a+2\left(z^2+b^2\right)^{1/2}\right]\right\}}{\left\{r^2+z^2+\left[a+2\left(z^2+b^2\right)^{1/2}\right]\right\}^{5/2}\left(z^2+b^2\right)^{3/2}}\ . \tag{4.21}$$

In these equations M_s is the total mass of the stellar component, b and a are the scales characterizing the distribution of mass along the vertical and radial directions respectively. These constants are free parameters of the model. The dark halo is modelled by the potential-density pair describing a non-singular isothermal halo as given by Binney & Tremaine [19], i.e.,

$$\phi_h(r,z) = \frac{V_0^2}{2}\log\left(r^2+z^2+R_c^2\right)\ , \tag{4.22}$$

and

$$\rho_h(r,z) = \frac{V_0^2}{4\pi G}\frac{r^2+z^2+3R_c^2}{\left(r^2+z^2+R_c^2\right)^2}\ . \tag{4.23}$$

In these equations, V_0 (dimension of a velocity) and R_c, the core radius of the halo, are constants and free parameters of the model. While the parameters in the first

potential-density pair are derived by fitting the light distribution, those in the second pair are derived by fitting the kinematics of the galaxy. The next step consists to solve the Jeans equation in cylindrical coordinates describing the motion of the stars, namely

$$\frac{\partial\left(\rho_s\sigma_r^2\right)}{\partial r}+\frac{\rho_s}{r}\left(\sigma_r^2-\sigma_\theta^2-\bar{V}^2\right)+\rho_s\frac{\partial\left(\phi_s+\phi_h\right)}{\partial r}=0\ . \tag{4.24}$$

In the above equation σ_r and σ_θ are respectively the radial and the tangential velocity dispersions and $\bar{V}$ is the average tangential velocity of the stars. This should not be confounded with the circular velocity defined by

$$V_c^2=r\frac{\partial\left(\phi_s+\phi_h\right)}{\partial r}\ . \tag{4.25}$$

The so-called rotation "drift" expressing the difference between the average tangential velocity and the circular velocity is derived from eqs. (4.24) and (4.25), i.e.,

$$\bar{V}^2-V_c^2=\left(\sigma_r^2-\sigma_\theta^2\right)+\frac{r}{\rho_s}\frac{\partial\left(\rho_s\sigma_r^2\right)}{\partial r}\ . \tag{4.26}$$

For stars, the observed quantities are the average tangential velocity and the velocity dispersion (profiles either along the major or the minor axes) projected in the line-of-sight. Thus, to compare the solutions of the system of equations above, it is necessary to calculate the velocity components along the line-of-sight, weighted by the luminosity density (or, equivalently, by the mass density), i.e.,

$$\left\langle\bar{V}(y)\right\rangle=\frac{1}{\mu}\int_{-\infty}^{\infty}\frac{y}{r}\rho_s(r,z)\bar{V}(r,z)\sin i\ ds\ , \tag{4.27}$$

where the integral is performed along the line-of-sight, y is the projected distance to the centre along the major (minor) axis, i is the inclination angle of the orbital plane with respect to the line of sight and

$$\mu=\int_{-\infty}^{\infty}\rho_s(r,z)ds\ , \tag{4.28}$$

is the projected mass density along the line-of-sight.

Contrary to the stars, which have a rotation "drift" as mentioned above, the gas is supposed to follow only a "Keplerian" motion, i.e., its velocity is equal to the local

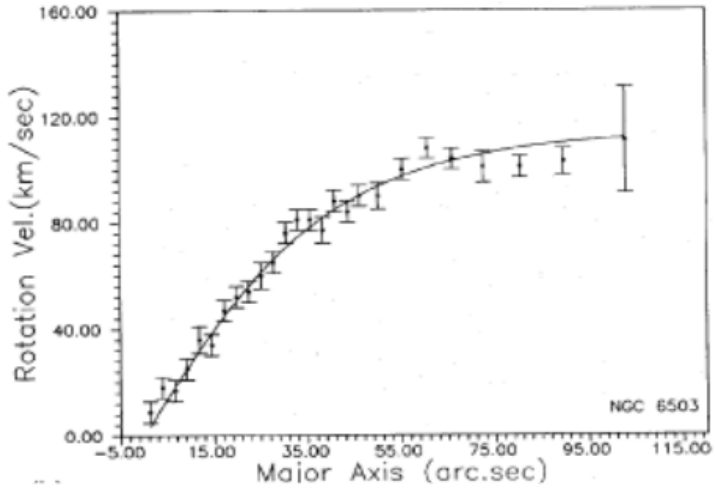


Figure 4. Theoretical rotation curve for the stellar component of NGC 6503 compared with observational data.

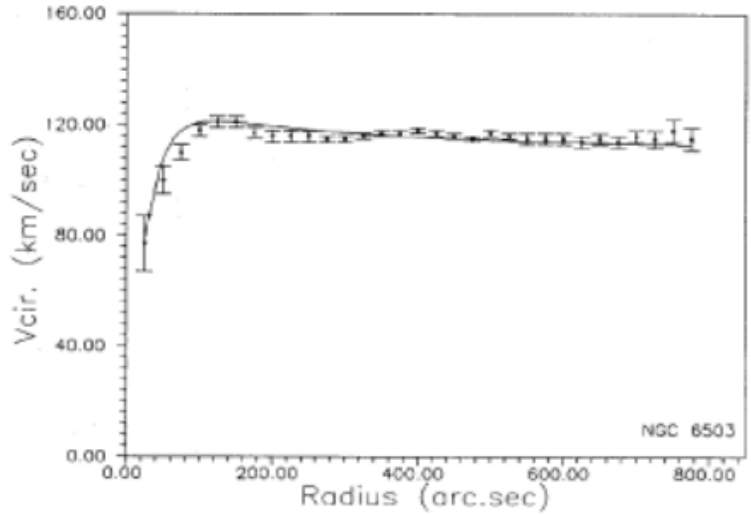


Figure 5. Circular velocity compared with 21cm data for NCG 6503.

circular velocity. As an example of such a procedure, fig. 4 shows the rotation curve (average tangential velocity) computed for the Sc galaxy NGC 6503 (inclination angle $i = 74^o$) compared with observational data. In fig. 5, the computed circular velocity is show in comparison with data derived from the 21 cm line. Notice that the gas permits to probe the potential up to distances 8 times longer than the stellar component. From

the best fit of data, the following parameters for the potentials were derived: $a = 1.2$ kpc; $b/a = 0.1$ and $R_c = 1.4$ kpc. The central density of dark matter is 0.2 $M_\odot$ pc^{-3} while that of the baryonic component is one order of magnitude higher. The contribution of dark matter increases with the distance to the centre and is dominant beyond 50 kpc. A similar result was obtained for another Sc galaxy NGC 3198, whose potential parameters are: $a = 2.0$ kpc; $b/a = 0.1$ and $R_c = 4.0$ kpc. The central density of dark matter for this galaxy is lower than the precedent case, being only 0.045 $M_\odot$ pc^{-3} and the density of the baryonic component is about 40 times higher. These results indicate a general trend for late type galaxies – the central regions are dominated by the baryonic component and not by dark matter, contrary what cosmological simulations suggest, as it will be discussed later.

Another disturbing fact concerns the measurement of rotation velocities for 47 gas dominated galaxies (or Low Surface Brightness – LSB – objects) by McGaugh [20]. For these objects, the baryonic mass is essentially under the form of gas and can be estimated more precisely than the stellar component. McGaugh found a robust correlation between the maximum rotation velocity and mass of the galaxy (supposed to be essentially the gaseous mass) of the form $M \propto V^4$. He claims that such a relation (Tully-Fisher relation) cannot be explained by the "canonical" ΛCDM model but it results naturally from MOND theory. We will return to this point later.

In the case of elliptical galaxies, the presence of "hot coronae", seen in X-rays in a large number of massive early type objects, seem to require a dark matter halo to confine gravitationally the hot gas. The dynamical analysis of stellar rotation curves and velocity dispersion profiles by [21] of a sample of these galaxies suggests that in about a half of the objects the M/L ratio increases slightly with radius, indicating the presence of dark matter. In other words, in early type galaxies dark matter manifests its presence only for distances larger than 15-20 kpc. This conclusion seems to be supported by the investigation of Romanowsky et al. [22], who studied the rotation of E-galaxies using planetary nebulae as probes of the potential. They found little if any dark matter in the objects of their sample. More recently, Wegner et al. [23] studied rotation curves of 8 E-galaxies in the A262 cluster and they found extremely low values for the mean halo density within the optical radius of the order of 0.02

$M_\odot$ pc^{-3}. However, this picture is challenged by the results of Biressa & de Freitas Pacheco [24] who estimated the central dark matter density in three E-galaxies situated respectively at redshift 0.11, 0.48 and 0.94. These galaxies act as gravitational lenses, imaging distant quasars that display an almost circular image (Einstein "ring"). For these galaxies, the authors found an average central dark matter density of 0.55 $M_\odot$ pc^{-3}, substantially higher than previous investigations. Is this result a characteristic of massive E-galaxies or simply a consequence of the evolutionary history of their halos? More data are required to elucidate this important question.

E. Dark Matter in the Solar Neighborhood

Already in the early thirties, the Dutch astronomer Jan Oort from the analysis of the stellar motions perpendicular to the galactic plane concluded that the dynamical mass density is about twice that obtained by adding the masses of the different stellar populations present in the solar neighbourhood. This problem is quite similar to the situation found by Zwicky in the Coma cluster. The presence (or not) of dark matter in our vicinity is quite important since its density is a key factor for all direct detection experiments running in the world, mentioned before.

Here also the situation is highly controversial. In the sixties Oort [25] revised his estimates of the dynamical mass density in the solar vicinity from which it is possible to derive a dark matter density of about $\rho_{dm} \approx 0.05\ M_\odot$ pc^{-3}. Crezé et al. (A&A 329, 920, 1998) using Hipparcos data found no evidence for dark matter in our neighbourhood. A subsequent analysis by Holmberg & Flynn [26] again based on Hipparcos data on A and F stars, leads to an estimated for the dynamical mass of $0.102 \pm 0.010\ M_\odot$ pc^{-3}, which should be compared with the estimate for the "visible" mass of 0.095 $M_\odot$ pc^{-3}. Thus, the authors concluded that taking into account the uncertainties, there is no compelling evidence for a significant amount of dark matter in the solar region. These "negative" results were reinforced by the recent investigation of Moni Bidin et al. [27] who analysed the kinematics of 412 red giants in the thick disk, concluding that there is no robust evidence for dark matter in the disk near the Sun. However, Bovy & Tremaine [28] contested the method of analysis by Moni Bidin et al.,

concluding for a dark matter density in the solar neighbourhood of $\rho_{dm} \approx 0.008 \pm 0.002$ $M_\odot$ pc^{-3}. A still higher density was obtained by Garbari et al. [29], who have analysed a sample of 2000 K dwarves taken from the literature and found a dark matter density of $\rho_{dm} \approx 0.025 \pm 0.014$ $M_\odot$ pc^{-3}.

Taking into account these contradictory results, we have performed a new analysis (unpublished) based on the Lindblad-Parenago approach. In cylindrical coordinates, the Jeans equation along the z-axis is

$$\frac{1}{\nu_*}\frac{\partial\left(\nu_*\sigma_z^2\right)}{\partial z} + \frac{\partial \Psi}{\partial z} = 0 \; , \tag{4.29}$$

where ν_* is the density of the considered stellar (probe) population and Ψ is the total gravitational potential, including contributions both from baryons and dark matter. The corresponding Poisson equation is

$$\frac{\partial^2 \Psi}{\partial r^2} + \frac{1}{r}\frac{\partial \Psi}{\partial r} + \frac{\partial^2 \Psi}{\partial z^2} = 4\pi G \rho_{tot} \; , \tag{4.30}$$

where ρ_{tot} is the total matter density (baryons + dark matter). Introduce now the (local) Oort's galactic rotation constants A and B as

$$A = \frac{1}{4\omega}\left(\frac{\partial^2 \Psi}{\partial r^2} - \frac{1}{r}\frac{\partial \Psi}{\partial r}\right) \; , \qquad \text{and} \qquad B = \frac{1}{4\omega}\left(\frac{\partial^2 \Psi}{\partial r^2} + \frac{3}{r}\frac{\partial \Psi}{\partial r}\right) \; , \tag{4.31}$$

where $\omega = A - B$. Combine eqs. (4.29), (4.30) and (4.31) to obtain, after some algebra

$$4\pi G \rho_{tot} = 2\left(A^2 - B^2\right) - \frac{\partial}{\partial z}\left[\frac{1}{\nu_*}\frac{\partial\left(\nu_*\sigma_z^2\right)}{\partial z}\right] \; . \tag{4.32}$$

This equation allows an estimate of the total density in the galactic plane at the solar position from the knowledge of the Oort's constants, the stellar density and the velocity dispersion profiles along the z-axis. Using the Oort's constants derived from Hipparcos data ($A = 14.8 \pm 0.8$ km/s/kpc and $B = -12.4 \pm 0.6$ km/s/kpc) and kinematical data from K-dwarves from Kuijken & Gilmore [30] one obtains for the total matter density $\rho_{tot} = 0.113 \pm 0.023$ $M_\odot$ pc^{-3} and a dark matter density $\rho_{dm} = 0.018 \pm 0.010$ $M_\odot$ pc^{-3} (or 0.69 GeV cm^{-3}). These values are consistent with Garbari et al. but indicate a dark matter density in the solar neighbourhood twice the value usually assumed.

Globular clusters (GC) can also be used as probes of the gravitational potential of the halo, since they are distributed almost spherically around the galactic centre up

to distances of 30 kpc. Following the analysis by Borges & de Freitas Pacheco [31], F-clusters were distributed in bins corresponding to the following distance intervals: 0-5 kpc, 5-11 kpc and 11-20 kpc. Then, the velocity dispersion was computed for each bin and the results are shown in fig. 6. Notice that at least for distances of the order of 20 kpc a "heating" of the system is observed, a fact also remarked by other authors. Anyway, it would be temerarious to extrapolate such a trend to higher distances since a decrease (or "cooling") of the velocity dispersion must occur at the outskirts of the halo. On the other hand, the density of GC decreases from the centre of the Galaxy

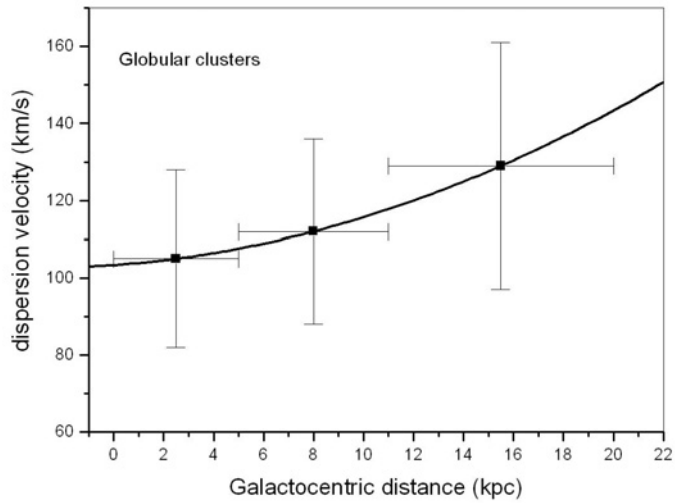


Figure 6. Velocity dispersion profile for galactic globular clusters.

according to $n_{gc} \propto r^{-10/3}$. If GC trace the halo potential, the Jeans equation can be used to compute the total mass inside the radius r, i.e.,

$$M_h(\leq r) = -\frac{r^2}{G n_{gc}} \frac{\partial \left(n_{gc}\sigma_r^2\right)}{\partial r} \ . \tag{4.33}$$

Then, the dark matter density can be estimated from the trivial relation

$$\rho_{dm}(r) = \frac{1}{4\pi r^2} \frac{dM_h(r)}{dr} \ . \tag{4.34}$$

The numerical solution of these equations permits to compute the dark matter density at $r = 8.5$ kpc, distance of the Sun to the galactic centre. The derived density is

$\rho_{dm} = 0.011\ M_\odot\ \mathrm{pc}^{-3}$, consistent with the value obtained from the analysis of K-dwarves.

F. Dark Matter at Cosmological Scales

Different methods permit to estimate the amount of dark matter at cosmological scales. At such large scales, one measures generally the density parameter Ω_{dm} that is the present ratio between the dark matter density to the critical density, i.e., $\Omega_{dm} = 8\pi G\rho_{dm}/3H_0^2$.

A relation between the random motion of galaxies not belonging to rich groups or clusters and the density parameter is given by the "cosmic virial theorem" by Peebles [32], given by

$$\left\langle \sigma_{3D}^2 \right\rangle < \frac{4}{n+7} H_0^2 \Omega_m J_2\ , \tag{4.35}$$

where $\left\langle \sigma_{3D}^2 \right\rangle$ is the square average of the velocity dispersion (random or peculiar motion) of galaxies, n is the exponent of the primordial power spectrum of the density fluctuations and J_2 is the second moment of the galaxy-galaxy correlation function. Bruno & de Freitas Pacheco [33] considered a sample of 301 E-galaxies and using a two-parameter distance indicator (effective radius and central stellar velocity of the galaxy), after correcting for the bulk motion due to the Great Attractor, they found for the mean random motion of galaxies $\sigma_{3D} = 352$ km/s. Using the value of J_2 derived from the CfA redshift survey, i.e., $J_2 = 150\ \mathrm{h}^{-2}\ \mathrm{Mpc}^2$, they obtained that $\Omega_m > 0.16$. Notice that this value is already 4 times higher than the limits imposed by the primordial nucleosynthesis.

The analysis of the galaxy-galaxy auto-correlation function provides also information on the density parameter, since it appear implicitly in the transfer function. The galaxy-galaxy auto-correlation function in the linear theory is related to the inverse Fourier transform of the power spectrum of the primordial density fluctuations, Thus, by measuring the auto-correlation function of galaxies, the power spectrum of the fluctuations can be accessed and by a fitting procedure (see Cole et al. [34]) one obtains: $\Omega_m = 0.27$ and $\Omega_b/\Omega_m = 0.18$, in excellent agreement with seven years WMAP data.

The density parameter can be also estimated from the motion of the Local Group of galaxies towards the Virgo cluster (Silk [35]). Originally, such a flow was detected in the CfA redshift survey with a velocity of about 470 km/s [36]. Further analyses indicated lower velocities, of the order of 200 km/s [21, 37–40]. Using the linear theory (see details in Section III.1), the expected velocity of the flow is

$$V_p = \frac{1}{3} H_0 D \langle \delta \rangle f(\Omega_m) \; . \tag{4.36}$$

Taking $H_0 D = 1180$ km/s as the Hubble flow velocity at the Virgo centre, $\langle \delta \rangle = 1.72$ the average density inside a sphere of radius D and $V_p = 200$ km/s, it results from (4.36) that the expected growth factor is $f(\Omega_m) = 0.295$ and, consequently, $\Omega_m = 0.12$ (corrections to the linear theory give a higher value, more compatible with the actual data), which indicates a value 4 times that resulting from the primordial nucleosynthesis.

Dark matter structures (filaments and halos of the cosmic web) are responsible for weak lensing of distance sources (see Massey et al. [41]). The analysis of the distortion of the shape of galaxies ("shear") by weak lensing permits an estimate of the mean global matter density. Lesgourgues et al. [42] obtained by such a procedure $\Omega_m = 0.247 \pm 0.016$, in good agreement with WMAP data.

III. THE DECOUPLING OF DARK MATTER

Dark matter particles decouple very early in the history of the universe from the cosmic plasma. The decoupling occurs when the interaction timescale between dark matter particles becomes equal to the expansion timescale. At this characteristic time t_* if the mass of the DM particle, in comparison with the plasma temperature, satisfies the condition $m_\chi c^2 > kT$, then particles decouple when they are non-relativistic. In the opposite sense, the decoupling occurs when particles are relativistic. It is generally accepted that DM particles are massive enough to decouple non-relativistically in order to be very "cold" when the formation of structures begins thanks to the gravitational instability. Primordial DM density fluctuations after decoupling are able to grow and to have the required amplitude observed in temperature maps of the CMB.

When thermal equilibrium disappears, particles begin to annihilate until their abundance becomes "frozen" because the expansion rate is faster than the annihilation rate. The annihilation process of particles X is governed by the equation

$$\frac{dn_X}{dt} + 3Hn_X = -n_X \bar{n}_X \langle \sigma_X v \rangle + P_X \ . \tag{4.37}$$

In this equation n_X and $\bar{n}_X$ are respectively the density of particles and anti-particles of species X, the first term on the right represents the annihilation rate and the last, the production rate. Clearly, in thermal equilibrium both terms on the right side of the above equation compensates each other. In this case, denoting n_{eq} the density at thermal equilibrium, equation (4.37) can be recast as

$$\frac{dn_X}{dt} + 3Hn_X = -\langle \sigma_X v \rangle \left(n_X^2 - n_{eq}^2 \right) \ , \tag{4.38}$$

where we have assumed an equal density of particles and anti-particles. Define the concentration of species X with respect to photons as $X = n_X/n_\gamma$. Since the comoving photon density is conserved, we get for the evolution of the concentration

$$\frac{dn_X}{dt} = -\langle \sigma_X v \rangle n_\gamma \left(X^2 - X_{eq}^2 \right) \ . \tag{4.39}$$

Consider now a small deviation from equilibrium such as $X = X_{eq} + \Delta X_{eq}$. Plug this into eq. (4.39) to obtain

$$\frac{1}{\Delta X_{eq}} \frac{d\Delta X_{eq}}{dt} = -2\langle \sigma_X v \rangle n_\gamma X_{eq}^2 \ . \tag{4.40}$$

The above equation permits to define the "freezing" timescale

$$t_{fre} = \left| \Delta X_{eq} / \left(d\Delta X_{eq}/dt \right) \right| \ ,$$

since the equilibrium is destroyed when $t_{fre} \geq H^{-1}$. The equality permits to fix the "freezing" time and to estimate the thermal average annihilation cross section. Considering the Hubble equation for a radiation dominated universe, which is valid when DM particles decouple, the thermal average annihilation cross section can be expressed as

$$\langle \sigma_X v \rangle = \frac{4\pi^3}{\sqrt{90}} \left(\frac{\hbar G^{1/3}}{c} \right)^{3/2} \frac{\sqrt{g_{eff}\left(T_{fre}\right)}}{m_X g_X} \left(\frac{kT_{fre}}{m_X c^2} \right)^{1/2} \exp\left(\frac{m_X c^2}{kT_{fre}} \right) \ . \tag{4.41}$$

In this equation, symbols have their usual meaning, $g_{eff}(T_{fre})$ is the number of degrees of freedom of particles present in the cosmic plasma at the freezing point, when the temperature is T_{fre} and g_X is the degree of freedom of particle X. Notice that $\langle \sigma_X v \rangle \propto T^q$, with $q = 0$ for a s-wave and $q = 1$ for a p-wave.

On the other hand, the expected present density parameter Ω_X related to particles X can be derived from the knowledge of the concentration of X at the freezing point (which remains constant during the remaining evolution of the universe) and the present photon density, i.e.,

$$\Omega_X = \frac{8\pi G}{3H_0^2} \left[\frac{g_{eff}(T_0)}{g_{eff}(T_{fre})} \right] m_X n_{\gamma,0} X_{eq}(T_{fre}) \ , \tag{4.42}$$

where the ratio between degrees of freedom at the freezing point and the present time corrects for the different "heating" episodes associated to annihilation of species. Since from WMAP data $\Omega_{dm} h^2 = 0.1108$, from equation (4.42) one obtains for the freezing temperature as a function of the particle mass (given in GeV)

$$T_{fre} \approx 0.048 m_X^{0.9583} \text{ GeV} \ . \tag{4.43}$$

Using this result and eq. (4.41), the thermal average annihilation cross section (in units of 10^{-26} cm^3 s^{-1}) can be expressed also as a function of the particle mass (in GeV) as

$$\langle \sigma_X v \rangle = 0.9447 + 0.11318 \log m_X \ . \tag{4.44}$$

The result above corresponds only to the s-wave contribution. The p-wave contribution amounts to about 4.6×10^{-28} cm^3 s^{-1} if the particle mass is in the range 30-900 GeV. If one assumes a DM particle of mass equal to 144 GeV, required to explain the putative 130 GeV γ-ray line the total annihilation rate is 1.25×10^{-26} cm^3 s^{-1} while the value necessary to explain the intensity of the γ-ray line is 2.3×10^{-27} cm^3 s^{-1}. This would suggest that only one fifth of the annihilations goes to the channel $\chi + \bar{\chi} \longrightarrow Z^0 + \gamma$ and the remaining goes to the channels $\chi + \bar{\chi} \longrightarrow W^+ + W^-, b^+ + b^-, \mu^+ + \mu^-, \tau^+ + \tau^-$.

A. Evolution after Decoupling

After decoupling and "freezing", dark matter particles evolve as a "collisionless" fluid and obey the Vlasov equation. For pedagogical purposes, a non-relativistic treatment will be considered but the basic conclusions are valid in a fully relativistic analysis. The one-particle distribution function $f(\overrightarrow{r}, \overrightarrow{v}, t)$ that gives the probability for a given particle in a given instant of time have a velocity $\overrightarrow{v}$ at the position $\overrightarrow{r}$, satisfies the equation

$$\frac{\partial f}{\partial t} + \overrightarrow{v} \cdot \frac{\partial f}{\partial \overrightarrow{r}} - \frac{\partial \phi}{\partial \overrightarrow{r}} \cdot \frac{\partial f}{\partial \overrightarrow{v}} = 0 \; . \tag{4.45}$$

This equation is coupled to the Poisson equation as

$$\nabla^2 \phi = 4\pi G \rho = 4\pi G \int f(\overrightarrow{r}, \overrightarrow{v}, t) d^3 \overrightarrow{v} \; . \tag{4.46}$$

At the early phases of the matter dominated era, the universe behaves as the Einstein-de Sitter cosmology. Thus, the matter density decreases as the square of the age of the universe, namely

$$\rho = \frac{1}{6\pi G t^2} \; . \tag{4.47}$$

Combining this relation with the solution of the Poisson equation, one obtains for the gradient of the gravitational potential

$$\frac{\partial \phi}{\partial r} = \frac{4\pi G}{3} \rho r = \frac{2}{9} \frac{r}{t^2} \; . \tag{4.48}$$

On the other hand, the characteristics of the Vlasov equation (4.45) are

$$dt = \frac{dx_i}{v_i} = -\frac{dv_i}{2x_i/9t^2} \; . \tag{4.49}$$

A solution of this system of equations is

$$I = v_i t^{2/3} - \frac{2}{3} x_i t^{-1/3} \; . \tag{4.50}$$

Recalling that the scale factor in the Einstein-de Sitter model varies as $a \propto t^{2/3}$, the equation above can be recast as

$$u_i = a(v_i - H x_i) \; . \tag{4.51}$$

Since solutions of the Vlasov equation can be expressed in terms of the characteristics, let us assume

$$f = A \exp\left[-\frac{a^2}{\theta^2}(\overrightarrow{v} - H\overrightarrow{r})^2\right] , \tag{4.52}$$

where A is a normalization constant that can be estimated from the condition that the integral in velocity space should be equal to the matter density (see eq. (4.46)). Thus,

$$f(\overrightarrow{r}, \overrightarrow{v}, t) = \frac{1}{6\pi^{5/2} G\theta^3 t_0^2} \exp\left[-\frac{a^2}{\theta^2}(\overrightarrow{v} - H\overrightarrow{r})^2\right] , \tag{4.53}$$

where t_0 is the present age of the universe and, without loss of generality we have put $\overrightarrow{r} = 0$.

Let us evaluate the mean square velocity using the distribution function (4.53). A simple calculation gives

$$\langle v^2 \rangle = \frac{3\theta^2}{2a^2} . \tag{4.54}$$

The above relation expresses a well known result: the kinetic energy of the particle decreases as square of the scale factor during the expansion of the universe. Define now the quantity $Q_{1D} = \rho/\langle v_z^2 \rangle^{3/2}$, which is an indicator of the phase space density. Using eq. (4.54) and supposing isotropic motion, one obtains

$$Q_{1D} = \sqrt{8}\frac{\rho a^3}{\theta^3} . \tag{4.55}$$

Since in the Einstein-de Sitter model $\rho a^3 =$ constant, the phase space density indicator is also constant. Thus, collisionless particles evolve satisfying the constancy of the phase space density, which will be violated only when structures begin to be formed. It is useful to estimate the value of Q_{1D} at the freezing point. The 1D velocity dispersion is $\sigma_{1D}^2 = Y_{fre}^{-1} c$ and the density is

$$\rho_x(T_{fre}) = \frac{g_X m_X^4}{\sqrt{8}} \left(\frac{c}{\hbar}\right)^3 Y_{fre}^{-3/2} e^{-Y_{fre}} , \tag{4.56}$$

where we have introduced $Y_{fre} = m_X c^2/kT_{fre}$. Thus, the initial phase space density indicator of dark matter particles is

$$Q_{1D} = \frac{g_X}{\sqrt{8}} \frac{m_X^4}{\hbar^3} e^{-Y_{fre}} . \tag{4.57}$$

As a consequence, if the value of Q_{1D} could be estimated from observations, the mass of the dark matter particle could be estimated. Unfortunately, as we shall see later, when structures begin to be formed, dark matter relaxes trough "violent relaxation", a process that reduces by several orders of magnitude the available phase space density.

B. The Growth of Linear Perturbations

Primordial fluctuations originated probably during the inflation era, are able to grow once dark matter particles decouple from the highly relativistic plasma. In fact, the amplitude of these fluctuations is seen is in the temperature map of the CMB, being of the order of $\Delta\rho/\rho \approx 10^{-4}$. A natural question rises: for baryons, the minimal gravitational unstable scale corresponds to the Jeans length. What is the corresponding scale for dark matter? Starting with the Vlasov equation (see eq. (4.45)), we first integrate all terms in the velocity space. The resulting equation is

$$\frac{\partial \rho}{\partial t} + \frac{\partial \rho \langle v_i \rangle}{\partial x_i} = 0 \, , \tag{4.58}$$

that simply expresses the mass conservation. Then, multiply the Vlasov equation by the velocity component v_j and integrate again over velocities to obtain the momentum conservation equation

$$\frac{\partial \rho \langle v_j \rangle}{\partial t} + \frac{\partial \rho \langle v_i v_j \rangle}{\partial x_i} + \rho \frac{\partial \phi}{\partial x_j} = 0 \, . \tag{4.59}$$

In the next step, multiply eq. (4.58) (mass conservation) by the velocity component v_j and subtract from eq. (4.59)

$$\rho \frac{\partial \langle v_j \rangle}{\partial t} - \langle v_j \rangle \frac{\partial \rho \langle v_i \rangle}{\partial x_i} + \frac{\partial \rho \langle v_i v_j \rangle}{\partial x_i} + \rho \frac{\partial \phi}{\partial x_j} = 0 \, . \tag{4.60}$$

Introduce now the velocity dispersion tensor

$$\sigma_{ij}^2 = \langle (v_i - \langle v_i \rangle)(v_j - \langle v_j \rangle) \rangle = \langle v_i v_j \rangle - \langle v_i \rangle \langle v_j \rangle \, . \tag{4.61}$$

Substitute (4.61) into (4.60) to obtain finally for the momentum conservation equation

$$\frac{\partial \langle v_j \rangle}{\partial t} + \langle v_j \rangle \frac{\partial \rho \langle v_i \rangle}{\partial x_i} + \frac{1}{\rho} \frac{\partial \rho \sigma_{ij}^2}{\partial x_i} + \frac{\partial \phi}{\partial x_j} = 0 \, . \tag{4.62}$$

Once the main conservation equations were derived, the following assumptions are made to get the linear approximation: a) the flow velocity is a combination of the Hubble flow and peculiar velocities, i.e., $\langle \overrightarrow{v} \rangle = \overrightarrow{V}_H + \overrightarrow{V}_p$; b) peculiar velocities are small in comparison with the Hubble flow, in particular in the early universe; c) the velocity dispersion of peculiar motions is isotropic, i.e., $\sigma_{ij}^2 = \sigma^2 \delta_{ij}$; d) the density contrast $\delta(\overrightarrow{r}, t) = \left[\rho_1(\overrightarrow{r}, t) - \rho_b(t)\right] / \rho_b(t)$, where $\rho_b(t)$ is the matter density of the expanding background satisfies the condition $\delta(\overrightarrow{r}, t) \ll 1$; e) in absence of strong relaxation processes, the evolution of density fluctuations occurs keeping constant the phase space density indicator Q_{1D}, which hereafter will be denoted simply by Q. This hypothesis is equivalent to say that the variations in density are adiabatic, since the entropy density is related to the phase-space density by $S \propto -\log Q$. Under these conditions, the effective pressure term $\rho\sigma^2$, appearing in (4.62), leads to a perturbed effective pressure $P_1(r,t) = \frac{5}{3} \frac{\rho_b^{5/3}(t)}{Q^{2/3}} \delta(r,t)$. Thus, the linearisation of the mass conservation equation, Poisson equation and momentum conservation equation leads to the system:

$$\frac{\partial \delta}{\partial t} + \overrightarrow{\nabla} \cdot \overrightarrow{V}_p = 0 \,, \tag{4.63}$$

$$\nabla^2 \phi = 4\pi G \rho_b \delta \,, \tag{4.64}$$

and

$$\frac{\partial \overrightarrow{V}_p}{\partial t} + H \overrightarrow{V}_p + \frac{5}{3} \left(\frac{\rho_b}{Q} \right)^{2/3} \overrightarrow{\nabla} \delta + \overrightarrow{\nabla} \phi = 0 \,, \tag{4.65}$$

where, as usually, the Hubble parameter is related to the scale factor a by $H = d \log a / dt$. Introduce now comoving coordinates $\overrightarrow{x}$ related to physical coordinates by $\overrightarrow{r} = a \overrightarrow{x}$ and take the Fourier transform (defined in the comoving frame) of the perturbed quantities. Under these conditions, the Fourier components of the linearised equations are

$$\frac{\partial \delta_k}{\partial t} + \frac{i \overrightarrow{k} \cdot \overrightarrow{V}_k}{a} = 0 \,, \tag{4.66}$$

$$-\frac{k^2}{a^2} \phi = 4\pi G \rho_b \delta_k \,, \tag{4.67}$$

and

$$\frac{\partial \overrightarrow{V}_k}{\partial t} + H\overrightarrow{V}_k + \frac{5}{3}\left(\frac{\rho_b}{Q}\right)^{2/3} \frac{i\,\overrightarrow{k}}{a}\delta_k + \frac{i\,\overrightarrow{k}}{a}\phi_k = 0 \; . \tag{4.68}$$

After some algebraic manipulation, we get for the evolution of the density contrast

$$\frac{\partial^2 \delta_k}{\partial t^2} + 2H\frac{\partial \delta_k}{\partial t} + \left[\frac{5}{3}\left(\frac{\rho_b}{Q}\right)^{2/3} \frac{k^2}{a^2} - 4\pi G\rho_b\right]\delta_k = 0 \; . \tag{4.69}$$

Gravitational instability requires that the bracketed term in (4.69) be negative, which implies the existence of a minimum wavelength λ_J above which perturbations grow and given by

$$a^2\lambda_J^2 = \frac{5\pi}{3GQ^{2/3}\rho_b^{1/3}} \; . \tag{4.70}$$

Notice that for dark matter the critical length (or Jeans length) depends inversely on the phase space density or, equivalently, on the mass of the dark matter particle. As in the baryonic case, a critical mass (or the Jeans mass) can be defined by

$$M_J = \frac{\pi}{6}\rho_b\lambda_J^3 = \sqrt{\frac{125}{972}}\pi^{5/2}\rho_b^{1/2}G^{-3/2}Q^{-1} \; . \tag{4.71}$$

The equation above can be expressed in terms of the redshift as

$$M_J = 7.91 \times 10^{-37}\sqrt{\Omega_{dm}h^2}(1+z)^{3/2}Q^{-1}\; M_\odot \; . \tag{4.72}$$

It is worth mentioning that the Jeans mass for dark matter increases with the redshift, an opposite behaviour of baryonic matter whose Jeans mass decreases with the redshift. If the mass of the DM particle is fixed, then the freezing temperature can be estimated (see eq. (4.43)) as well as the corresponding redshift and the phase space density indicator (eq. (4.57)). These quantities permit an evaluation of the Jeans mass from (4.72). For DM particles having masses in the range 100-200 GeV, the Jeans mass varies from $2.88 \times 10^4\; M_\odot$ down to $1.02 \times 10^4 M_\odot$. These values represent probably the lowest possible masses for dark matter halos.

C. Warm Dark Matter

Particles decoupling relativistically may have presently, depending on their masses, a non-negligible velocity dispersion. However, as mentioned previously, these particles

stream freely up to the matter-radiation equality, defining a scale λ_f below which perturbations are erased. The comoving free streaming scale is given by

$$\lambda_f \approx 3.15 \frac{kT_0}{m_X cA} \left[1 + \log \left(\frac{1+z_{NR}}{1+z_{eq}} \right) \right] . \tag{4.73}$$

In this equation T_0 is the present CMB temperature, m_X is the particle mass, z_{NR} is the redshift at which the particle becomes non-relativistic and z_{eq} is the redshift at matter-radiation equality. The parameter A is defined by

$$A^2 = \frac{8\pi G}{3c^2} g_{eff}(T_{fre}) u_{\gamma,0} , \tag{4.74}$$

including, besides usual constants, the effective number of degrees of freedom at the freezing point $g_{eff}(T_{fre})$ and the present energy density of CMB photons $u_{\gamma,0}$. Numerically, the free streaming scale can be written as

$$\lambda_f \approx \frac{0.46}{m_{keV} \sqrt{g_{eff}(T_{fre})}} [1 + \log(407 m_{keV})] \text{ Mpc} . \tag{4.75}$$

If the WDM particle mass is conveniently chosen, the free streaming acts as a filter, erasing masses below the characteristic scale

$$M_f = \frac{4\pi}{3} \rho_b \lambda_f^3 \approx 1.12 \times 10^{12} \Omega_m h^2 \lambda_f^3 \, M_\odot , \tag{4.76}$$

where λ_f is in Mpc. Moreover, simulations suggest that even cored dark matter profiles can be obtained (Maccio et al., [43]). A generally invoked candidate is the sterile neutrino, whose mass is expected to be around 2 keV. In this case, the free streaming scale is about 0.54 Mpc and the characteristic mass scale is about 2×10^{10} $M_\odot$. However, for these values, the simulations by Maccio et al. indicate that the core radius is too small, namely, of the order of 10 pc only. If one requires for such a mass scale a core radius of about 1.0 kpc, i.e., two orders of magnitude larger, then the mass of the putative sterile neutrino should be around 0.1 keV. Particles with such a mass have a free streaming scale of about 6.6 Mpc and a characteristic mass scale of 4.3×10^{13} $M_\odot$ and, in this case, besides dwarf galaxies even bright and massive objects are erased.

IV. RELAXATION OF COLLISIONLESS SYSTEMS

Dark matter particles interact only through gravitation and, if halos are in dynamical equilibrium (or in "quasi"-equilibrium), it is important to understand the mechanisms leading to such a relaxed state. Elliptical galaxies are old stellar systems, which can be imagined to be constituted by a collisionless (stellar) fluid. The nearly "universal" brightness profiles seen in these objects require short timescale processes and no dependence on the stellar mass, otherwise significant colour gradients would be observed (see, for instance, de Freitas Pacheco et al. [44]). Concerning dark halos, cosmological simulations indicate that these structures have peaked density profiles (Navarro et al. [45, 46]). If density fluctuations have an initial power spectrum of the form $P(k) \propto k^n$ in an Einstein-de Sitter universe, then halos are expected to have a density profile $\rho \propto r^{-(9+3n)/(4+n)}$ (Hoffman & Shaham [47]) and the behaviour of the Navarro-Frenk-White (NFW) profile in the central region of halos is recovered only for $n \gg 1$.

A short relaxation timescale cannot be obtained via two-body interactions since the expected timescale in longer than the Hubble time. Relaxation processes in the presence of a variable gravitational potential has been investigated since the sixties by Michel Hénon [48]. His simulations indicate that the system reach equilibrium in a timescale of the order of the dynamical timescale, i.e., $t \propto 1/\sqrt{G\rho_0}$ where ρ_0 is the initial mean density of the system. After the numerical experiments by Hénon, the concept of "violent" relaxation (VR) was developed by Lynden-Bell [49]. VR describes how a collisionless system relaxes from an initial chaotic state to quasi-equilibrium, when rapid changes in the gravitational potential occur during the gravitational collapse or following a merger episode. A fundamental characteristic of the Lynden-Bell theory is that the resulting particle energy distribution is "Fermi-like". These ideas were revisited by Shu [50], who concluded that collisionless relaxation via chaotic changes of the collective gravitational potential lead to no mass segregation and that stars of different masses have the **same velocity dispersion** (not the same temperature). In realistic collisionless systems the velocity distribution is probably Maxwellian and, according to the conclusions by Nakamura [51], collective effects like

phase-mixing produce a Gaussian velocity distribution independent on the particle mass. Gaussian velocity distributions were found in N-body simulations by Merrall & Henriksen [52] and by Diemand et al. [53]. The latter authors also found that the velocity distribution has a negative kurtosis, meaning a distribution with a flat topped profile.

The difficulty to describe these fully non-linear processes leading to the (quasi) equilibrium state of collisionless systems demands the use of numerical tools. A panoply of simulations on the subject has been reported in the literature in the past years. Here, we will be concerned only to some specific results.

The first point concerns the central slope $d\log\rho/d\log r$ of the density profile of halos. According to Klypin et al. [54], such a slope may depend on the merger history while Ascasibar et al. [55] suggest a dependence on the initial conditions.

The second aspect concerns the phase density indicator Q. High resolution simulations indicate that $Q \propto r^{-\beta}$, where the exponent $\beta \approx 1.87$ is the same either for galaxy-size halos (Taylor & Navarro [56]) or cluster-size halos (Rasia et al., [57]). Such a power-law profile for the phase space density could be a consequence of the hierarchical assembly of halos that preserves stratification or to be a generic feature of the violent relaxation process (Williams et al., [58]).

A. A Toy Model for Halos

In this section a self-consistent toy model for dark matter halos will be examined. This model reproduces quite well the density profile resulting from simulations, excepting for the very central regions and reproduces the power low profile observed for the phase space density indicator Q. Moreover, the velocity distribution has a negative kurtosis, i.e., the velocity distribution has a flat topped profile as seen in most simulations.

We assume that the halo is characterized by a power law distribution for the energy of dark matter particles, i.e.,

$$f(E) = K|E|^p \,, \tag{4.77}$$

where K is a normalization constant and $E = V^2/2 - \phi(r)$. Since gravitationally bound particles must have negative energies, the maximum velocity at a distance r from the halo centre is $V_{max} = 2\phi(r)$, corresponding to a particle of zero energy. The density can be derived from

$$\rho(r) = \int_0^{V_{max}} f(E) d^3V = \frac{4\pi\sqrt{2}K\Gamma(3/2)\Gamma(p+1)}{\Gamma(p+5/2)}\phi^{p+3/2} \; . \tag{4.78}$$

In other to compute the integral, we have used (4.77) and the definition of the energy. Similarly, the velocity dispersion of DM particles at the distance r from the halo centre can be calculated as

$$\rho(r)\sigma^2 = \int_0^{V_{max}} V^2 f(E) d^3V = \frac{8\pi\sqrt{2}K\Gamma(5/2)\Gamma(p+1)}{\Gamma(p+7/2)}\phi^{p+5/2} \; . \tag{4.79}$$

Assume that the system is in equilibrium, has spherical symmetry and an isotropic velocity distribution. Under these conditions, the system is described by the following Jeans equation

$$\frac{1}{3}\frac{\partial \rho\sigma^2}{\partial r} + \rho\frac{GM(r)}{r^2} = 0 \; , \tag{4.80}$$

which can be rewritten as

$$\frac{r^2}{3\rho}\frac{\partial \rho\sigma^2}{\partial r} = -GM(r) \; . \tag{4.81}$$

Then, differentiate both sides of the above equation with respect to the radial coordinate

$$\frac{\partial}{\partial r}\left[\frac{r^2}{3\rho}\frac{\partial \rho\sigma^2}{\partial r}\right] = -G\frac{\partial M(r)}{\partial r} = -4\pi G\rho r^2 \; . \tag{4.82}$$

Replace eqs. (4.78) and (4.79) into the equation above to obtain finally

$$\frac{1}{r}\frac{\partial}{\partial r}\left[r^2\frac{\partial \phi}{\partial r}\right] + A\phi^{p+3/2} = 0 \; , \tag{4.83}$$

where A is a constant including all previous constants. The equation above is a Lane-Emden equation for the gravitational potential and describes a system with a polytropic equation of state. Despite the existence of non-singular solutions for the Lane-Emden equation, which do not describe adequately the structure of halos, we

consider singular solutions of the form $\phi \propto r^{-\beta}$. These solutions exist if $p > 3/2$ and $\beta = 4/(2p+1)$. As we shall see, halos can be reasonably described by the particular case $p = 9/2$ or $\beta = 2/5$. Then, the density profile varies as $\rho \propto \phi^{p+3/2} \propto r^{-12/5}$ and the phase density indicator varies as $Q \propto \phi^p \propto r^{-9/5}$. Figures 7 and 8 compare these theoretical expectations with simulated data by Taylor & Navarro [56]. The velocity

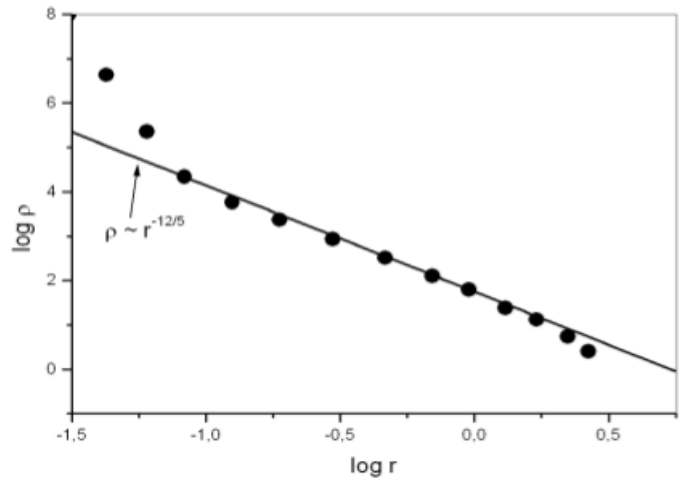


Figure 7. Predicted (solid line) and simulated density profile of a dark matter halo. Data indicate a steeper slope only in the very central regions.

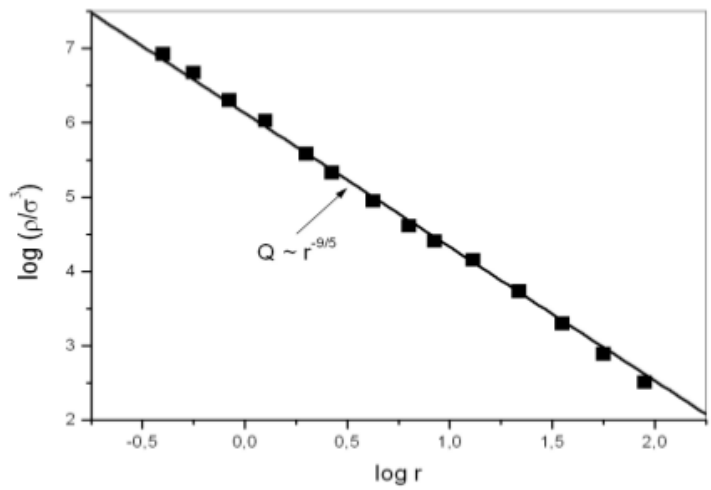


Figure 8. Predicted (solid line) and simulated phase-space density indicator Q.

distribution can be computed from the relation

$$\frac{dF(V)}{dV} = BV^2 \left(V_{max}^2(r) - V^2\right)^{9/2} \ , \tag{4.84}$$

where B is a normalization constant and the maximum velocity V_{max} was defined previously. If we consider a spherical shell in which the maximum velocity is 300 km/s, one obtains for the velocity dispersion a value of 139 km/s. On the other hand, the normalized kurtosis of the considered velocity distribution is $\kappa = (\mu_4/\mu_2) - 3 = -1.54$. This negative kurtosis indicates a flat-topped distribution as observed in simulations. Figure 9 shows data points derived from the velocity distribution (4.84) with $V_{max} =$ 300 km/s (filled squares) and the best fitted Gaussian (red curve). Notice that this Gaussian has a velocity dispersion of only 55 km/s, indicating that the distribution (4.84) is in fact broader. In conclusion, this simple self-consistent model explains quite well some basic features of dark matter halos, namely, the external density profile, the power law profile of the phase space density indicator and a flat-topped velocity distribution.

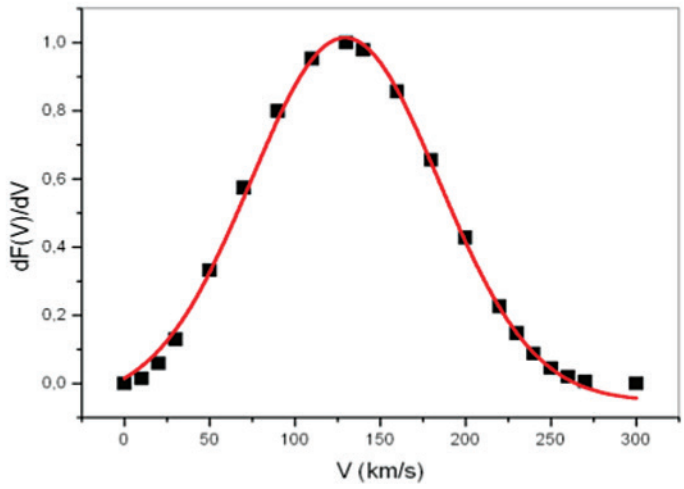


Figure 9. Velocity distribution derived from the toy model (filled squares) fitted by a Gaussian.

B. Dynamical Evolution of Halos

In the previous section we have examined the equilibrium state of halos. However, halos are in continuous evolution, since they grow either by accreting mass "quietly", via filaments constituting the cosmic web or "violently", via merger episodes. The latter process, as we shall see, reduces the phase space density and, on the average, increases the angular momentum of halos. All processes occurring during a merger event are strongly non-linear and must be studied through numerical simulations.

Here are described the results derived from cosmological simulations including only dark matter (Peirani et al., [59]). These simulations have a mass resolution of 2.05×10^8 $M_\odot$, sampling a volume of the universe of 27000 h^{-3} Mpc^3. Halos were followed from $z = 50$ up to $z = 0$. Simulations including both "cold" and "warm" dark matter were performed. The latter corresponding to particles with a mass of 0.5 keV that have a free streaming scale of 1.8 Mpc and a mass scale of 8×10^{11} $M_\odot$.

Once that halos have been identified by a percolation algorithm, unbound particles (with positive total energies) are eliminated. Thus, only gravitationally bound structures are retained. Halos were classified into two categories: those that have grown only by accretion and those that have been assembled (besides accretion) by fusions. These events are characterized by the merger of a main halo and another having at least a mass equal to 1/3 of the main object (computations modifying the criterion, i.e., considering a limit of 1/6 do not change substantially the results). The first aspect concerns the growth of halos or, in other words, how their masses evolve. For each redshift the mass distribution of halos is computed as well as the median of the distribution. Two regimes are clearly seen: an initial rapid growth followed by a more slowly gain of mass. The evolution of the median of the distribution is shown in figure 10. The evolution can be described by power laws of the form $M_{med} \propto t^\alpha$. For halos assembled essentially by accretion, $\alpha = 0.91$ for $z > 1.8$ and $\alpha = 0.58$ for $z < 1.8$ while for halos having suffered important fusions $\alpha = 1.35$ for $z > 1.5$ and $\alpha = 0.81$ for $z < 1.5$. Clearly, the later class of halos grows faster and attains mass values higher than halos of the former category. The existence of two regimes is probably a consequence of the transition observed in the total rate of fusion occurring at $z \approx 1.5$,

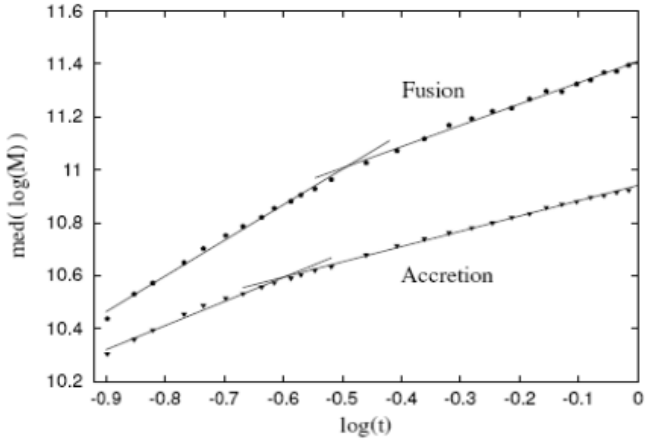


Figure 10. Evolution of the median of the mass distribution of halos.

which is drastically reduced beyond this moment. It is interesting to mention that in a hierarchical scenario, if the power spectrum of the density fluctuation is of the form $|\delta_k^2| \propto k^n$ (Peebles, [32]) then the mass of the fluctuations grows as $m \propto t^{4/(3+n)}$. Thus, for a Harrison-Zeldovich spectrum, the mass grows linearly with time, close to the results of simulations. Moreover, Toth & Ostriker [60] computed the expected growth for a CDM power spectrum and derived $m \propto t^{1.3}$, in agreement with simulations.

Similarly, the **total** energy of halos scales as a power-law with different exponents according to the redshift and according to the main process of assemblage. The evolution of the median of the (absolute values of the) energy distribution is shown in figure 11. An immediate conclusion is that halos having being assembled by different merger events are more bound than halos that have grown by accretion only.

The median of the energy distribution evolves as $E \propto t^{\beta}$, where $\beta = 1.32$ for $z > 1.8$, $\beta = 0.66$ for $z < 1.8$ for halos that have grown by accretion and $\beta = 1.93$ for $z > 1.5$, $\beta = 1.03$ for $z < 1.5$, for halos assembled by significant merger events.

Another fundamental aspect concerns the dynamical state of halos. It is generally assumed in the literature that once halos collapse they attain equilibrium. The analysis is essentially based on the spherical model or the Press-Schechter approach (see details in the next section). However, even after the collapse, halos continue to grow by accretion or mergers that destroy their equilibrium state. Thus halos are always relaxing or searching for equilibrium. On the average, deviations from equilibrium

can be quantified by the virial ratio $r_V = 2T/|W|$, which in equilibrium is equal to one. Higher values indicate "expansion" and lower values indicate "contraction". In

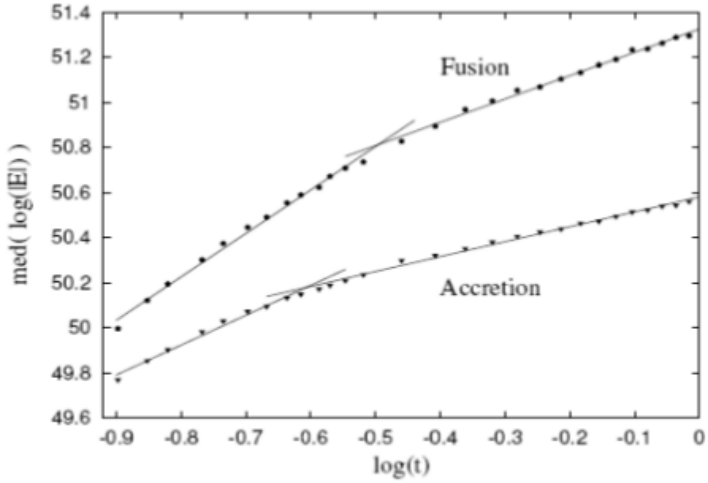


Figure 11. Evolution of the median of the total energy distribution of halos.

figure 12 it is shown the evolution of the average value of the virial ratio for CDM and WDM for halos of both classes. Notice that for CDM the average virial ratio decreases as $r_V \propto t^{-0.123}$, reaches the unity at $z \approx 0.25$ and then becomes slightly smaller. WDM halos, on the average, have not yet reached a fully equilibrium state. When considered separately, CDM halos assembled only by accretion reach $r_V = 1$ at $z \approx 0.61$ while those assembled by important fusion events have not yet reached equilibrium on the average. This confirms that fusions destroy the equilibrium state reached just after shell crossing (see next section) Instead of considering average values, the distribution of the virial ratio as a function of the halo mass can be investigated. This is done in figure 13 where it can be seen that only a small fraction of the halos have $r_V = 1$. In general, low mass halos are still contracting while massive halos have a tendency to "expand".

1. Evolution of the phase-space density indicator

The early evolution of Q can be understood if the Press-Schechter approach is adopted, namely, a spherical symmetric overdense region evolves as a closed universe.

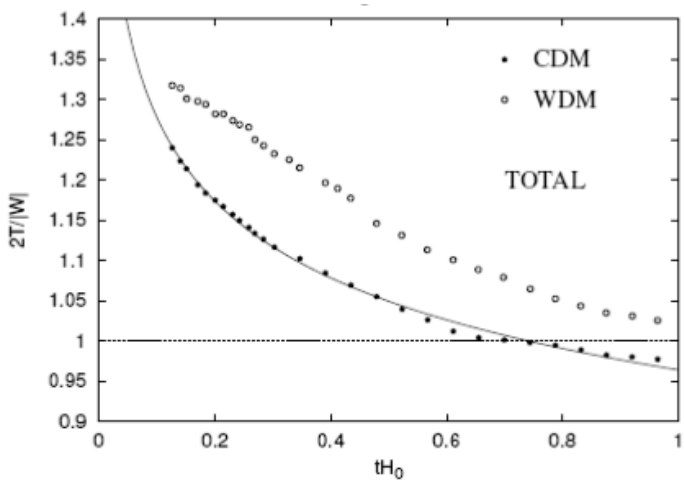


Figure 12. Evolution of the average value of the virial ratio.

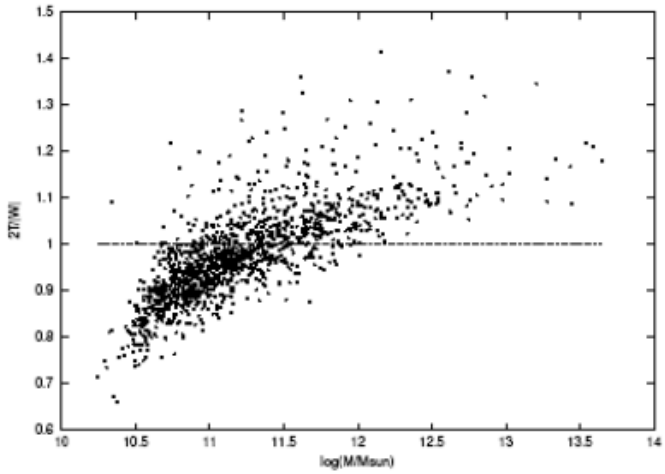


Figure 13. Distribution of the virial ratio as a function of the halo mass.

As long as the expansion of the universe is matter dominated, the external radius R enclosing a mass M obeys the equation

$$\frac{d^2R}{dt^2} = -\frac{GM}{R^2} \ . \tag{4.85}$$

This equation can be trivially integrated and one obtains

$$\frac{1}{2}\left(\frac{dR}{dt}\right)^2 = \frac{GM}{R} + E \ , \tag{4.86}$$

where E is an integration constant. The above equation can be rewritten as

$$\left(\frac{\dot{R}}{R}\right)^2 = \frac{2GM}{R^3} + \frac{2E}{R^2} \, , \Leftrightarrow \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G}{3}\rho - \frac{Kc^2}{a^2} \, , \tag{4.87}$$

which can be formally identified with the Hubble equation if we put $M = 4\pi R^3\rho/3$ and $2E = -Kc^2$. Then, the integration constant (the total energy) is related to the curvature of a closed world-model. The parametric solution of (4.86) (or (4.87)) is well known and given by

$$R = \frac{R_{\rm max}}{2}(1 - \cos\theta) \, , \qquad t = \frac{t_{\rm col}}{2\pi}(\theta - \sin\theta) \, . \tag{4.88}$$

The spherical perturbation expands initially as fast as the background, but the density contrast increases, reaches a maximum expansion at $\theta = \pi$ when the kinetic energy is zero, detaches from the expanding background and then contracts until a complete collapse at $\theta = 2\pi$. During the final contracting phase, the potential energy is comparable to the kinetic energy and the rapid variation of the potential transfer energy from bulk motions to random motions (violent relaxation) until equilibrium is reached in a timescale $t_{\rm col} = \pi^2 R_{\rm max}^2/2GM$. This last phase is characterized by the crossing of different infalling shells and is dubbed "shell crossing" phase. Simple algebra shows that the density contrast after collapse is $\rho_V/\rho_b = 18\pi^2 \approx 177.6$. Some observers use the round value $\rho_V/\rho_b = 180$ as a criterion to define "virialization". If one takes into account the effect of the cosmological constant in the expansion of the background, and consider a flat universe ($\Omega_m + \Omega_\Lambda = 1$), the density contrast is given by (Bryan & Norman, [61])

$$\Delta_V = \frac{\rho_V}{\rho_b} = 18\pi^2 + 82\left[\Omega_m(z) - 1\right] - 39\left[\Omega_m(z) - 1\right]^2 \, , \tag{4.89}$$

where the density parameter Ω_m should be taken at the virialization redshift z_V. In order to compute the phase space density indicator we need the mean density and the 1D velocity dispersion. The mean halo density at virialization is

$$\rho_V = \frac{3H_0^2\Omega_{m,0}}{8\pi G}(1 + z_V)^3\Delta_V \, . \tag{4.90}$$

On the other hand, since at virialization the 1D velocity dispersion satisfies the condition $3\sigma^2 = -2E$, one obtains (see details in Peirani & de Freitas Pacheco, [62])

$$\sigma^3 = \left[\frac{2}{3}\lambda^{-1/3} + \frac{1}{3}\lambda^{2/3}\right]^{3/2} \left(\Omega_\Lambda H_0^2\right)^{1/2} GM \ , \tag{4.91}$$

where $\lambda = \Omega_\Lambda H_0^2 R_{\rm max}^3/GM$. Using also the result derived from simulations, i.e., the virialization redshift for a given halo of mass M is approximately given by

$$M_{11} = \frac{185}{(1+z_V)^{2.86}} \ , \tag{4.92}$$

where M_{11} is the halo mass in units of $10^{11}\ M_\odot$. From all these relations, one obtains finally for the phase space density indicator

$$Q \approx \frac{3.51\times 10^{-9}}{M_{11}^{1.54}}\ M_\odot\ {\rm pc}^{-3}\ {\rm km}^{-3}\ {\rm s}^{-3} \ . \tag{4.93}$$

The relation above represents the expected value derived from the spherical model of the phase space density indicator as a function of the halo mass just after virialization. This expression is compared (solid line) in figure 14 with data derived from the aforementioned simulations. Notice that values of the phase space density indicator are

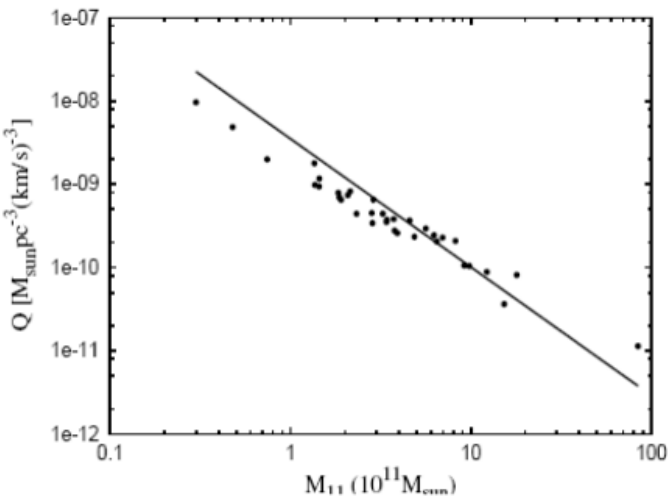


Figure 14. Predicted phase space density compared with simulated data.

several orders of magnitude smaller than those estimated at freezing of the abundance of dark matter particles. Such a reduction is due essentially to violent relaxation

and phase mixing processes (to be discussed later), which lead the system to a quasi-equilibrium state.

After this dramatic reduction of the initial phase space density at shell crossing, halos continue to be assembled either by accretion through filaments or by mergers. These processes lead to a continuous decrease of the phase space density. This is illustrated in figure 15, where the evolution of the parameter Q is shown for a typical halo.

Another important information can be obtained from the correlation between Q and the central velocity dispersion of DM particles (both Q and σ_{1D} are computed inside a spherical volume with a radius equal to one tenth of the gravitational radius of the halo). The observed slope of such a correlation is $d\log Q/d\log\sigma_{1D} = 2.65$ while

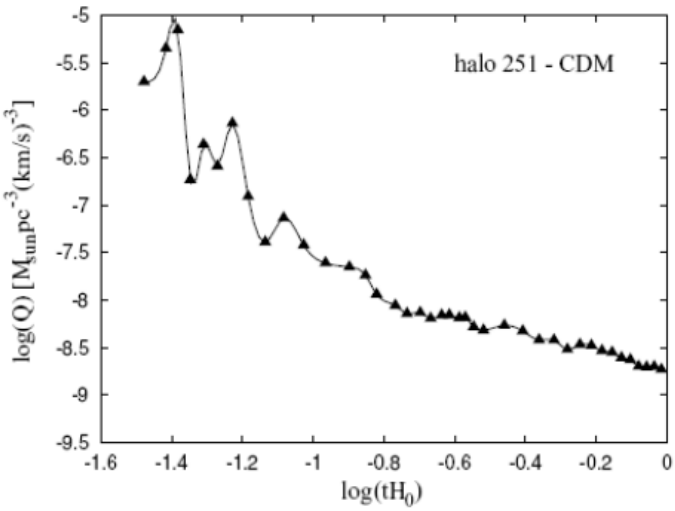


Figure 15. Evolution of the phase space density indicator Q derived from simulations.

simulations indicate a slope $d\log Q/d\log\sigma_{1D} = 2.16$. These results indicate that larger systems like clusters of galaxies have a lower phase space density than dwarf galaxies. This could be explained if clusters suffered different merger episodes (contrary of what has happened in the history of dwarf galaxies) that are able to reduce considerably the phase space density. The effect of mergers in the evolution of the density parameter can be seen in three examples shown in figure 17. The upper panels of the figure show the evolution of the derivative dQ/dt of the phase space density, while the lower panels

show the corresponding variation of the velocity dispersion. In all the cases, the label 1 indicates the moment at which the shell crossing occurs. Notice that for each halo there is a peak in the derivative of the velocity dispersion and a valley in the derivative of Q. Thus, in the shell crossing phase, the transfer of energy from bulk to random motions increases the velocity dispersion and reduces the phase space density. The other events labeled 2, 3, ..., correspond to merger episodes. In these episodes there is some heating produced by the violent relaxation mechanism (peaks in the derivative of the velocity dispersion) and a reduction of the phase space density (valleys in the derivative of the phase space density.

Despite these sudden variations of the phase space indicator during mergers, the relation Q–σ_{1D} suggests that halos evolve in quasi-equilibrium. As the halo is assembled, its mass scales on the average with the central velocity dispersion as $M \propto \sigma_{1D}^{8/3}$. If, indeed halos evolve in quasi-equilibrium, the gravitational radius of the system evolves as $R \propto M/\sigma_{1D}^2 \propto \sigma_{1D}^{-2/3}$. Thus, $Q = \rho/\sigma_{1D}^3 \propto MR^{-3}\sigma_{1D}^{-3} \propto \sigma_{1D}^{-7/3}$, which is essentially the slope derived from simulations. These simulations demonstrate that the formation of structures and their late evolution define the value of Q at the present time.

C. Phase Space Mixing

Processes of mixing in phase space are generally due to non-local large scale dynamics, occurring far from equilibrium but promoting such a state. The process of mixing can be seen during merger events when significant variations of the gravitational potential occur. Here we consider the example of a main halo whose mass is about $4.0 \times 10^{12}\ M_\odot$ and the 1D central velocity dispersion is 202 km/s, which captures two satellites in eccentric orbits that will be disrupted by tidal forces. Satellite-1 has a mass of $1.6 \times 10^{11}\ M_\odot$ and a 1D central velocity dispersion of 51 km/s and satellite-2 has a mass of $2.8 \times 10^{11}\ M_\odot$ and a 1D central velocity dispersion of 78 km/s. The first sub-halo has an orbital period of about 6 Gyr and a periastron of about 100 kpc whereas the second sub-halo has a shorter period, namely, ≈ 2 Gyr. In figure 18 are shown different snapshots of the evolution in the phase space diagram including

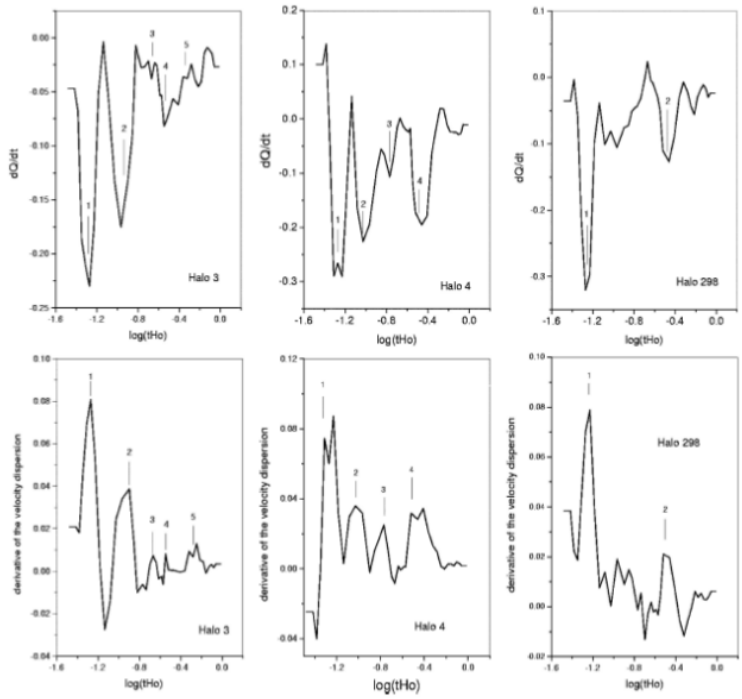


Figure 16. The evolution of the derivative of the phase space density indicator (upper panels) compared with the evolution of the derivative of the velocity dispersion (lower panels) for three simulated halos.

as coordinates the radial velocity component of the particle and its radial distance to the centre of the halo. For each panel of figure 18, the corresponding redshift is indicated. Satellite-2 is completely disrupted after two periastron passages while satellite-1 passes to far from the centre of the main halo to be disrupted, although it loses a significant amount of mass from its external regions. Notice that structures in phase space are gradually formed as satellite-2 passes by periastron. The observed streams in phase-space are similar to those derived from models of halo formation by secondary infall (Sikivie et al., [63]). However, in that case, structures are associated to "caustics" (mathematical surfaces of "infinite" density associated to cold collisionless particles). In real physical systems the density is limited by the fact that the motion of DM particles is not purely radial and that the velocity dispersion is finite and not zero. In a given instant and for a fixed value of the radial distance, a certain number of velocity peaks, corresponding to different streams can be seen in figure 18, which

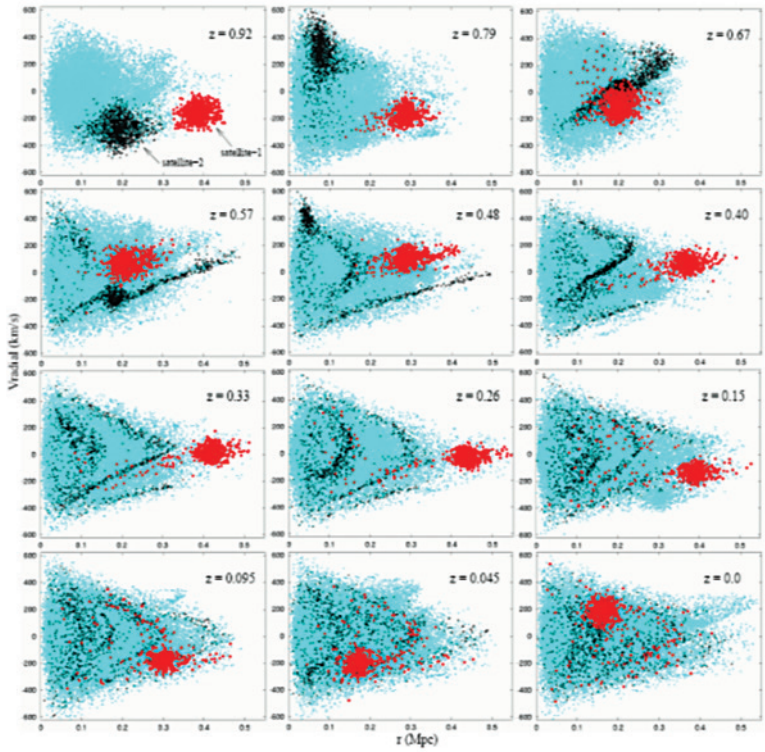


Figure 17. Phase space evolution of the capture of two sub-halos.

depend on the initial angular momentum. In the extreme limit of circular orbits, only one stream should be observed, as it occurs for satellite-1. The width of the m-velocity peak associated to the m-caustic is generally estimated from the Liouville theorem. In figure 19 are shown for the same snapshots, the evolution of the different radial velocity distributions. In the first snapshot of fig.19 ($z = 0.92$) it is possible to distinguish three distinct distributions associated to the main halo and to the two satellites. Despite satellite-1 to have preserved its identity, stripped particles from its outskirts mix in the velocity space. Notice in particular the appearance of extended tails in the velocity distribution of satellite-1, while the velocity distribution of satellite-2 coincides practically with that of the main halo after 3.5 Gyr due to an efficient mixing in phase-space. The "heating" process is probably due to tidal shocks. In spite of Gaussians be able to fit the radial velocity distribution of the main halo and sub-halos, the normalized kurtosis of these Gaussians is negative ($\kappa = -0.57$ for the main halo and $\kappa = -0.73$ for satellite-2), indicating that the radial velocity distributions

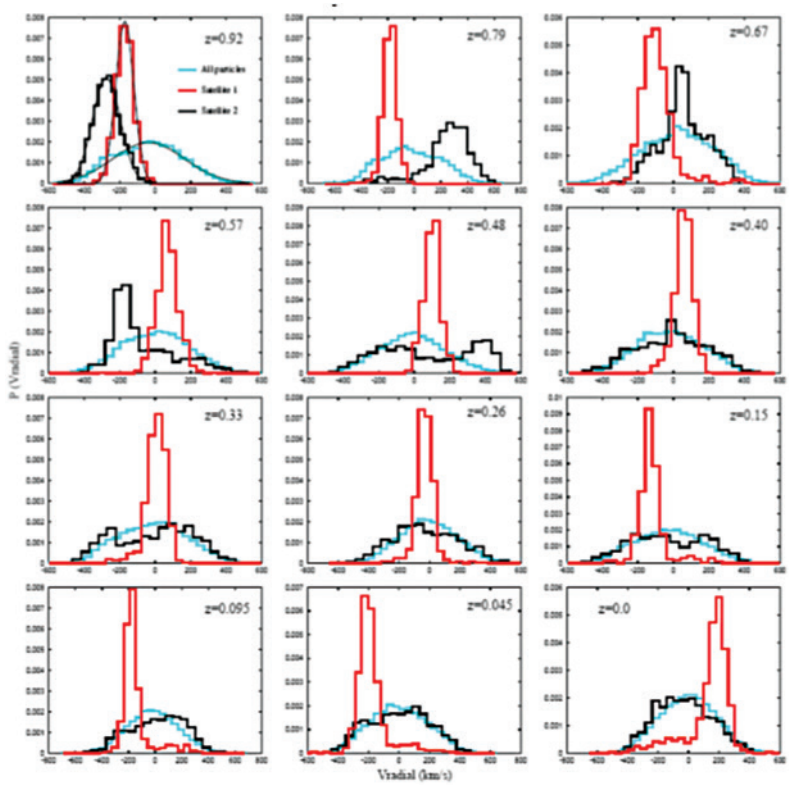


Figure 18. Evolution of the radial velocity distribution for the main halo and satellites.

have top-flatted profiles, with the effect being more accentuated in substructures.

Another interesting example concerns the fusion at $z = 0.92$ of two halos having respectively masses equal to $4.3\times10^{12}\ M_\odot$ and $1.2\times10^{12}\ M_\odot$ (see figure 20). Structures due to different periastron passages are still seen in phase-space at $z = 0.48$ but both objects are completely mixed at $z = 0$ (upper panels of figure 20). In the second row, panels show for the same redshifts the evolution of the radial velocity distribution for all particles constituting both halos and that of the sub-halo. Notice that at $z = 0.92$ the satellite is distinguishable, having a flat-topped Gaussian with a velocity dispersion of 131 km/s. At $z = 0.48$, despite the presence of structures in phase space, both velocity distributions practically coincide, having a velocity dispersion of 251 km/s. Such an increase or "heating" is probably due to tidal shocks as already mentioned. At $z = 0$, when structures have disappeared, the velocity distributions are indistinguishable from each other and a small "cooling", due to such a mixing, is noticed since the velocity dispersion of the "absorbed" halo is 237 km/s.

Panels in the third and fourth rows show respectively the radial velocity distribution for all particles and for the sub-halo only in spherical shells at different radial distances. The distributions are well fitted by Gaussians having decreasing velocity dispersions as one should expect. This result contradicts the study by Wojtak et al. [64], who concluded that out of the central regions halos do not have Gaussian velocity distributions.

D. The Angular Momentum of Halos

Since the work of Fred Hoyle [65], different studies addressed to the origin of the angular momentum of galaxies consider tidal interactions as the basic mechanism. As the protogalaxy (or protohalo) expands, its angular momentum grows almost linearly with time due to tidal interactions with their surroundings. The tidal mechanism ceases to be effective once the protohalo decouples from the expanding background, turns around and collapses. Hence, the turn around marks the instant of maximum angular momentum acquisition by the protohalo. The acquired angular momentum by this mechanism in the linear theory and under the Zeldovich approximation is

$$J_i = a^2 \frac{dD(t)}{dt} \epsilon_{ijk} Q_{jl} I_{lk} \ , \tag{4.94}$$

where a is the scale factor, $D(t)$ is the linear growth factor, Q_{jl} is the deformation tensor that depends on the nearby distribution of matter and I_{lk} is the inertia tensor of matter inside the density perturbation volume. The tidal torques are non-zero only if the main axes of inertia are not aligned with the principal axes of the deformation tensor. In an Einstein-de Sitter universe, $J \propto t$ until maximum expansion. In the hierarchical scenario, small lumps of matter collapse first but accretion and merger effects modify the epoch of the maximum expansion with respect to predictions of the spherical model. Thus, the maximum angular momentum acquired by halos depends on their growth histories. Using the tidal linear theory one obtains for the ensemble average of the angular momentum (Catelan & Theuns, [66])

$$\left\langle J^2 \right\rangle^{1/2} = 7.4 \times 10^{65} (1 + z_{\max})^{-1/2} \ M_{11}^{5/3} \ \mathrm{kg \ m^2 \ s^{-1}} \ , \tag{4.95}$$

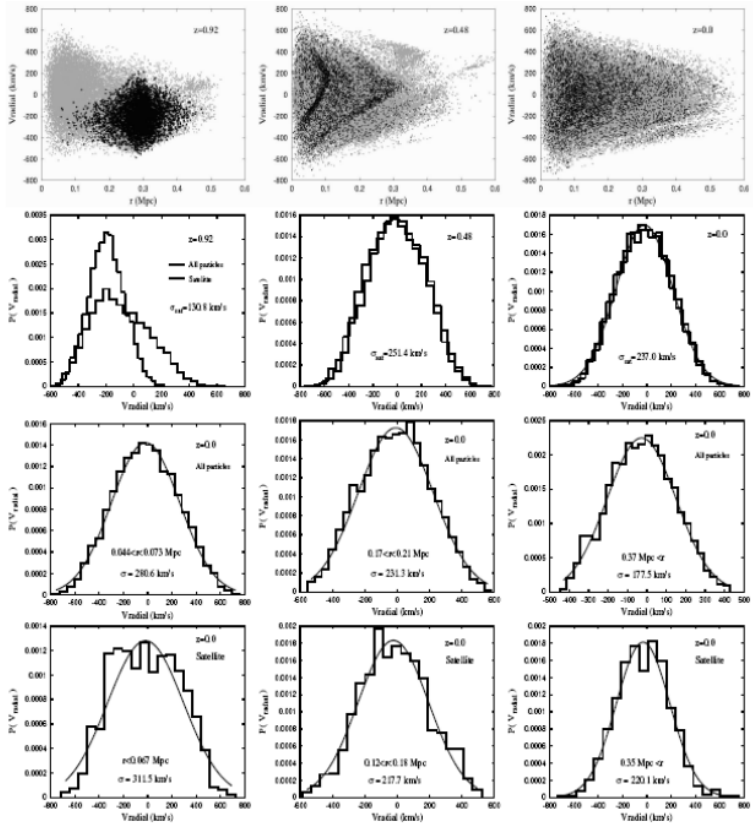


Figure 19. Evolution in phase-space and of radial velocity distributions.

where $z_{\max}$ is the redshift of maximum expansion and M_{11} is the halo mass in units of $10^{11}\ M_\odot$.

To fix the ideas, lets us first consider the case of our Galaxy, the Milky Way. The original data by Fall [67] on the angular momentum of spiral galaxies was recently revised and, from these data one obtains the fit

$$J_d = 7.3 \times 10^{49}\ M_d^{1.57}\ \mathrm{kg\ m^2\ s^{-1}}\ , \tag{4.96}$$

where the (baryonic) disk mass is in solar units. The mass of the disk of the Galaxy is not well known but a value around $4.0\times10^{10}\ M_\odot$ seems reasonable. Similarly, estimates for the mass of our halo are uncertain but dynamical estimates lead to values of the order of $1.5 \times 10^{12}\ M_\odot$ (see next section). Notice that these numbers imply a baryon-to-dark matter ratio lower than the cosmic value. Using the adopted value for the disk mass and the equation (4.96), one obtains for the specific angular momentum of the

galactic disk $J_d/M_d \approx 4.0 \times 10^{25}$ m^2 s^{-1}. The redshift of maximum expansion for our halo is $z_{\max} \approx 2.9$. Then, using equation (4.95) one obtains for the specific angular momentum of the halo $J_h/M_h = 1.2 \times 10^{25}$ m^2 s^{-1}. Most theories of galaxy formation assume that the specific angular momentum of baryons and dark matter are equal. If this is the case then, taking into account the uncertainties in the masses, the linear tidal theory predicts a value about a factor 3-4 less than observations indicate. The situation would be worse had we assumed a cosmic baryon-to-dark matter ratio. If we wish to study non-linear effects, then cosmological simulations are needed. In fact, numerical experiments on the formation of galaxies indicate difficulties to produce galaxies with disks large enough (Steinmetz & Navarro, [68]). Simulated disks have, in general, less angular momentum than halos in the inner regions and miss high J-values observed in the halo J-distribution. On the other hand, Bullock et al. [69] claim, based on their simulations, that disks may have a "universal" angular momentum distribution. However, such a distribution is unable to produce "exponential" disks.

Based on semi-analytical models, Vitvitska et al. [70], considered an alternative scenario, namely, the spin of halos is built up gradually and randomly by accretion and mergers. Peirani et al. [71] have investigated such a scenario using cosmological simulations as a tool. The follow up of individual halos has shown a strong correlation between the mass accretion rate and the variation rate of the angular momentum. The latter could be positive or negative according to the sense of the accreting flow with respect to the halo spin or, in case of mergers, according to the relative orientation of the spin and the orbital angular momentum. This is shown in figure 21. The acquisition of angular momentum by halos during accretion phases and merger episodes is also clearly demonstrated by figure 22, in which it can be seen the evolution of the median of the angular momentum distribution for halos that have grown either by accretion only or by important merger events. In figure 22, the upper curve corresponds to halos of the "merger class" while the lower curve corresponds to halos of the "accretion class". Notice that halos of the former have presently an angular momentum almost one order of magnitude higher than halos of the latter class. Contrary to the tidal linear theory that predicts a linear growth of the angular momentum up to maximum expansion, cosmological simulations indicate a more rapid growth, i.e., $J \propto t^q$ where

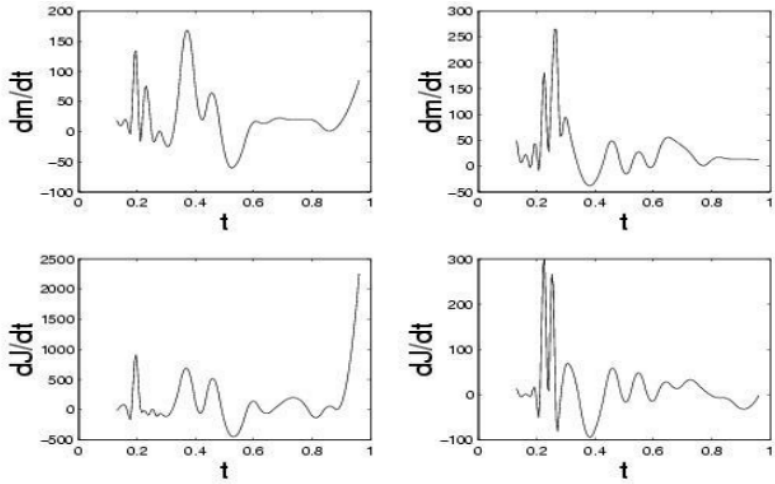


Figure 20. Upper panels – variation of the accretion rate for two halos. Lower panels – variation of the angular momentum for the same halos.

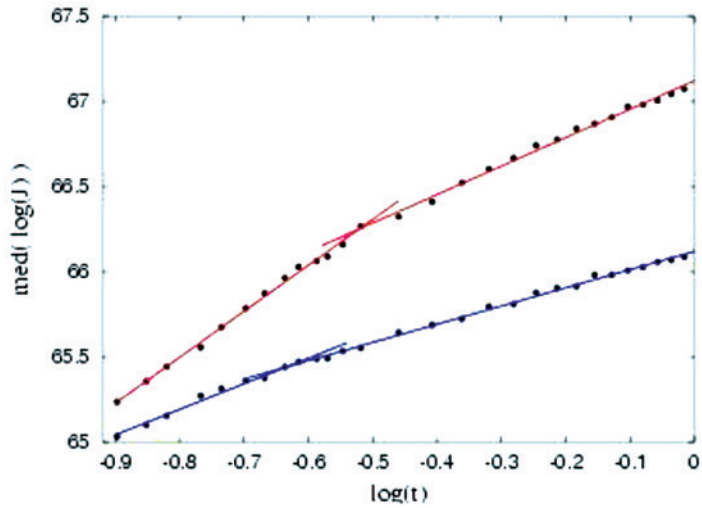


Figure 21. Evolution of the angular momentum of halos.

for halos of the "merger class", $q = 2.7$ if $z > 1.5$ and $q = 1.67$ if $z < 1.5$, whereas for halos of the "accretion class", $q = 1.5$ for $z > 1.8$ and $q = 1.07$ for $z < 1.8$ (the same regimes seen in the assembly of halos and in the total energy evolution). Moreover, the angular momentum of halos continue to grow even after the maximum expansion epoch, thanks to the transfer of the orbital angular momentum to the spin during a fusion episode or to the transfer of angular momentum of the accreted flow originated

in filaments of the cosmic web.

V. THE "COLDNESS" OF THE HUBBLE FLOW

Using data from the CfA redshift survey, Davis & Peebles [72] derived for the r.m.s. of the velocity difference in the line-of-sight of correlated pairs, separated by less than 1 Mpc the value of 340 km/s. However, from a sample of E-galaxies, Bruno & de Freitas Pacheco [33] derived for their random motions a 1D velocity dispersion of 187 km/s, in agreement with the previous estimate of Hale-Sutton et al. [73], i.e., 167 km/s. These values are significantly higher than the 1D velocity dispersion found by Sandage et al. [74] for the local Hubble flow ($R < 2 - 4$ Mpc), i.e., 60 km/s. This result motivated those authors to conclude that the local flow is relatively "quiet" and "cold". Further analyses revealed even lower values $\approx$ 30 km/s [75, 76]. The 1D velocity dispersion seems to increase slightly with the radius R of the local volume since Karachentsev et al. [77] found a velocity dispersion of about 85 km/s for $R \leq 5.5$ Mpc while Tikhonov & Kyplin [78] found 97 km/s for $R \leq 7.0$ Mpc. The latter results are consistent with the conclusions by Maccio et al. (MNRAS 359, 941, 2005), who derived the following expression for the 1D velocity dispersion

$$\sigma_H = (88 \pm 20) \times (R/7 \text{ Mpc}) \text{ km/s} . \tag{4.97}$$

Different explanations have been proposed to understand the "coldness" of the local flow and, in particular, the possible effects of the cosmological constant. However, different cosmological simulations **including** Λ, disprove such an idea: Governato et al. (New Ast. 2, 91, 1997) obtained $\sigma_H = 150 - 300$ km/s, Peirani & de Freitas Pacheco (New Ast. 11, 325, 2006) derived $\sigma_H = 73$ km/s, in good agreement with the simulations by Macciò et al. [79], who obtained a 1D velocity dispersion of 80 km/s within a local volume of radius 3 Mpc.

More recently, this question was reviewed by Martinez-Vaquero et al. [80], who performed new N-body simulations of the local universe. These authors found that world models including or not a cosmological constant have similar values of σ_H that excludes Λ as the cause or main cause of the "coldness" of the flow. However, Martinez-

Vaquero et al. obtained an interesting result from their simulations: σ_H correlates with the mean matter density of the local volume (radius of 7 Mpc in their case) normalized by the mean cosmological density. A totally different approach was considered by Aragon-Calvo et al. [81]. In their scenario, the "coldness" is a consequence of the geometry and dynamics of our local wall. The expansion of the local wall also affects the measurement of the local Hubble constant and, in the range $1 < R < 3$ Mpc, according to their simulations, the expected 1D velocity dispersion should be around 30 km/s.

In general, the 1D velocity dispersion with respect to the Hubble flow is computed from the relation

$$\sigma_H^2 = \frac{1}{N}\sum_i^N (V_i - H_0 r_i)^2 \ , \tag{4.98}$$

where V_i and r_i are respectively the radial velocity and the radial distance to the centre of mass of the Local Group of a given satellite. In reality, nearby satellites have negative velocities since they feel the gravitational effects of the MW-M31 pair and only beyond a certain distance they follow a "pure" Hubble flow. Thus, in our approach, the 1D velocity distance will be computed from the relation

$$\sigma_H^2 = \frac{1}{N}\sum_i^N [V_i(r) - V_{TL}(r)]^2 \ . \tag{4.99}$$

Here we follow Peirani & de Freitas Pacheco [82, 83], assuming that the velocity profile near the Local Group is given by the Tolman-Lemaître spherical model, including the contribution of the cosmological constant. Inside the "zero"-velocity surface satellites are collapsing and therefore have negative velocities, outside they have positive velocities but are gravitationally bound (they will reach maximum expansion in the future and then collapse) and beyond the "zero"-energy surface, they follow a pure Hubble flow. From the numerical solution of the equations of motion, they have obtained the velocity profile

$$V_{TL}(r) = -\frac{0.976}{r^n}\left(\frac{GM}{H_0^2}\right)^{(n+1)/3} + 1.377 H_0 r \ , \tag{4.100}$$

where $n = 0.627$ and M is the total mass of the Local Group (MW + M31). The "zero"-velocity surface can be estimated from the above equation from the condition

$V_{TL}(R_0) = 0$. Expressing the result in terms of the mass, one obtains

$$M = 1.89 \frac{H_0^2 R_0^3}{G} \ . \tag{4.101}$$

Thus, if the radius R_0 of the "zero"-velocity surface can be estimated from observations of the velocity profile, the mass of the system can be derived. Notice that the radius R_E of the "zero"-energy surface satisfies the condition $V_{TL}(RE) = H_0 R_E$ and one can easily show that $R_E = 2.217 R_0$. Using this approach, Peirani & de Freitas Pacheco [83] derived the 1D velocity dispersion for several systems in the local universe, listed in table 1.

Group	MW/M31	M81	NGC 253	IC 342	Cen A/M83
σ_H [km/s]	36	53	45	34	45
M [10^{12} $M_\odot$]	3.0 ± 0.8	0.92 ± 0.24	0.13 ± 0.18	0.20 ± 0.13	2.1 ± 0.5

Notice that the "coldness" is observed not only for the pair MW-M31 but also for the groups M81, NGC 253, IC 342 and CenA-M83. Suppose that these groups are represented by density fluctuations above the background having a Gaussian profile, namely

$$\delta_f(r) = A \exp\left(-\frac{r^2}{2R^2}\right) \ , \tag{4.102}$$

where A is a normalization constant, which can be obtained from the condition that the total mass of the perturbation is

$$M = \int_0^\infty \delta_f(r) \rho_b d^3 r \ . \tag{4.103}$$

From eqs. (4.102) and (4.103) one obtains

$$A = \frac{M}{(2\pi^3)\rho_b R^3} \ . \tag{4.104}$$

Compute now the Fourier transform of (4.102), i.e.,

$$\delta_f(k) = \int \delta_f(r) \exp(-i \overrightarrow{k} \cdot \overrightarrow{r}) d^3 r = \frac{M}{8\pi^{3/2}\rho_b} \exp\left(-\frac{k^2 R^2}{4}\right) \ . \tag{4.105}$$

On the other hand, in the linear approximation, the Fourier component of the peculiar velocity is given by

$$\overrightarrow{V}(k) = -\frac{Ha}{ik^2} \frac{\partial \log D}{\partial \log a} \delta_f(k) \overrightarrow{k} \ , \tag{4.106}$$

where a is the scale factor and D is the growing mode of the linear density fluctuation. From Parseval's theorem, the mean square velocity within a volume V_* is

$$\left\langle V_p^2 \right\rangle = \frac{1}{V_*} \int \left| V^2(k) \right| d^3k \; . \tag{4.107}$$

Plugging eqs. (4.105) and (4.106) into eq. (4.107) one obtains

$$\left\langle V_p^2 \right\rangle = \frac{1}{16\sqrt{2\pi^3}} \frac{H^2 a^2}{V_* R} \left(\frac{\partial \log D}{\partial \log a} \right)^2 \frac{M^2}{\rho_b^2} \; . \tag{4.108}$$

Now, fix the parameter R by imposing that 90% of the total mass is included within a sphere of radius r_* (for instance, in the case of the MW-M31 pair, this is about 1 Mpc). Thus, $R = 2r_*/5$. Recalling that the mean matter density inside the volume V_*, where the mean velocity dispersion is computed is $\bar{\rho} = M/V_* = 3M/4\pi r^3$, one obtains finally for the mean square peculiar velocity

$$\left\langle V_p^2 \right\rangle^{1/2} = \frac{\sqrt{10}}{(4608\pi)^{1/4}} Hra \left(\frac{\partial \log D}{\partial \log a} \right) \left(\frac{r}{r_*} \right)^{1/2} \frac{\bar{\rho}}{\rho_b} \; . \tag{4.109}$$

Notice that the present ($a = 1$) 1D-velocity dispersion is (case of isotropic peculiar motion) $\sigma_H = \langle V_p^2 \rangle^{1/2}/\sqrt{3}$. Thus

$$\sigma_H = 0.167 f(\Omega_m) H_0 r \left(\frac{r}{r_*} \right)^{1/2} \frac{\bar{\rho}}{\rho_b} \; . \tag{4.110}$$

Within a spherical volume of radius $r = 2.5$ Mpc (about the radius of the "zero"-energy surface for the Local Group) and $r_* = 1$ Mpc, one obtains from the equation above $\sigma_H = 23(\bar{\rho}/\rho_b)$ km/s. The matter overdensity in such a volume due to the pair MW-M31 is about 1.3, implying $\sigma_H = 30$ km/s, in good agreement with the derived value (see table 1).

VI. MAIN CONCLUSIONS

Up to the present date, there is no convincing evidence from direct or indirect experiments for the existence of dark matter. Despite negative results from the LHC concerning supersymmetry, the possible detection of the scalar Higgs boson with a mass around 125 GeV points to a minimum mass around 160 GeV for the

lightest supersymmetric particle. This is consistent with the non-detection of γ-rays originated from M87, either at energies of 100 MeV or 1 GeV by Fermi-LAT. A possible annihilation γ-ray line at 130 GeV is also consistent with these figures.

Dark matter manifests via gravitational effects at different scales: clusters of galaxies ($\approx$ 2 Mpc), groups of galaxies (100-400 kpc), binary galaxies ($\approx$ 50 kpc), galaxies themselves ($\approx$ 20 kpc) and in the solar neighbourhood ($\approx$ 1 kpc).

The structures present in the universe are formed hierarchically by fusion of small lumps of dark matter, whose minimal mass is around 104 $M_{\odot}$, if the mass of particles is around 100-200 GeV. As halos are assembled either by accretion or mergers, the core density decreases on the average, by one order of magnitude in the redshift interval $10 > z \geq 0$.

The inner regions of the lumps of matter collapse first and, in this shell crossing phase, a rapid transfer of energy from bulk to random motions occur, leading to a rapid increase of the velocity dispersion and a decrease in the phase-space density. This early evolutionary phase is followed by a late period of "slow" heating. Simulations indicate that the energy transfer from bulk to random motions produces a heating of all particles regardless their initial energies. Halos issued from simulations with particles of different masses relax according to predictions of the violent relaxation mechanism, i.e., at the end all particles have the very same velocity dispersion. This relaxation process leads to a flat-topped velocity distribution.

In the process of accretion and/or during merger episodes, halos acquire angular momentum. The growth of angular momentum during the assembly of halos is faster than that due to tidal torques, which is operative only up to the phase of maximum expansion of lumps.

---

[1] S. F. King, [arXiv:0712.1750 [physics.pop-ph]].

[2] C. Kelso, D. Hooper and M. R. Buckley, Phys. Rev. D **85** (2012) 043515

[3] C. Arina, J. Hamann, R. Trotta and Y. Y. Y. Wong, JCAP **1203** (2012) 008

[4] D. Hooper, Phys. Dark Univ. **1** (2012) 1

[5] E. Aprile *et al.* [XENON100 Collaboration], Phys. Rev. Lett. **105** (2010) 131302

[6] C. Beskidt, W. de Boer, D. I. Kazakov and F. Ratnikov, JHEP **1205** (2012) 094

[7] H. Baer, V. Barger and A. Mustafayev, Phys. Rev. D **85** (2012) 075010

[8] S. Akula, B. Altunkaynak, D. Feldman, P. Nath and G. Peim, Phys. Rev. D **85** (2012) 075001

[9] S. Peirani, R. Mohayaee and J. A. de Freitas Pacheco, Phys. Rev. D **70** (2004) 043503

[10] E. Tempel, A. Hektor and M. Raidal, JCAP **1209** (2012) 032

[11] M. Su and D. P. Finkbeiner, arXiv:1206.1616 [astro-ph.HE].

[12] A. Hektor, M. Raidal and E. Tempel, Astrophys. J. **762** (2013) L22

[13] J. A. De Freitas Pacheco and S. Peirani, Grav. Cosmol. **11** (2005) 169

[14] L. Bergstrom, Phys. Rev. D **86** (2012) 103514

[15] E. Escalera, E. Slezak, and A. Mazure, 1992, A&A, 264, 379

[16] de Freitas Pacheco, J. A., & Junqueira, S. 1988, Ap&ss, 149, 141

[17] Ortega, V. G., & de Freitas Pacheco, J. A. 1993, AJ, 106, 899

[18] Satoh, C. 1980, PASJ, 32, 41

[19] Binney, James J. Galactic dynamics. Princeton university press, 1987.

[20] S. S. McGaugh, Phys. Rev. Lett. **106** (2011) 121303 [Erratum-ibid. **107** (2011) 229901]

[21] Ortega, V. G., & de Freitas Pacheco, J. A. 1989, AJ, 97, 1000

[22] A. J. Romanowsky, N. D. Douglas, M. Arnaboldi, K. Kuijken, M. R. Merrifield, N. R. Napolitano, M. Capaccioli and K. C. Freeman, Science **301** (2003) 1696

[23] G. A. Wegner, E. M. Corsini, J. Thomas, R. P. Saglia, R. Bender and S. B. Pu, arXiv:1206.5768 [astro-ph.CO].

[24] T. Biressa and J. A. de Freitas Pacheco, Gen. Rel. Grav. **43** (2011) 2649

[25] Oort, J. H. 1960, BAN, 15, 45

[26] J. Holmberg and C. Flynn, Mon. Not. Roy. Astron. Soc. **313** (2000) 209

[27] C. M. Bidin, G. Carraro, R. A. Mendez and R. Smith, Astrophys. J. **751** (2012) 30

[28] J. Bovy and S. Tremaine, Astrophys. J. **756** (2012) 89

[29] S. Garbari, C. Liu, J. I. Read and G. Lake, Mon. Not. Roy. Astron. Soc. **425** (2012) 1445

[30] K. Kuijken and G. Gilmore, Mon. Not. Roy. Astron. Soc. **239** (1989) 605.

[31] Borges, A. C. A., & de Freitas Pacheco, J. A. 1989, APSS, 153, 67

[32] Peebles, Phillip JE. The large-scale structure of the universe. Princeton university press, 1980.

[33] Bruno, R., & de Freitas Pacheco, J. A. 1992, Cosmology and Large-Scale Structure in the Universe, 24, 173

[34] S. Cole *et al.* [2dFGRS Collaboration], Mon. Not. Roy. Astron. Soc. **362** (2005) 505

[35] Silk, J. 1974, ApJ, 193, 525

[36] Tonry, J. L., & Davis, M. 1981, ApJ, 246, 680

[37] Dressler, A. 1984, ApJ, 281, 512

[38] Tammann, G. A., & Sandage, A. 1985, ApJ, 294, 81

[39] Federspiel, M., Tammann, G. A., & Sandage, A. 1998, ApJ, 495, 115

[40] S. Peirani and J. A. de Freitas Pacheco, New Astron. **11** (2006) 325

[41] R. Massey, T. Kitching and J. Richard, Rept. Prog. Phys. **73** (2010) 086901

[42] J. Lesgourgues, M. Viel, M. G. Haehnelt and R. Massey, JCAP **0711** (2007) 008

[43] A. V. Maccio, S. Paduroiu, D. Anderhalden, A. Schneider and B. Moore, arXiv:1202.1282 [astro-ph.CO].

[44] Pacheco et al., Recent Research Developments in Astronomy & Astrophysics, 1, 271, 2003

[45] J. F. Navarro, C. S. Frenk and S. D. M. White, Mon. Not. Roy. Astron. Soc. **275** (1995) 56

[46] J. F. Navarro, C. S. Frenk and S. D. M. White, Astrophys. J. **462** (1996) 563

[47] Y. Hoffman and J. Shaham, Astrophys. J. **297** (1985) 16.

[48] Hénon, M. 1964, Annales d'Astrophysique, 27, 83

[49] D. Lynden-Bell, Mon. Not. Roy. Astron. Soc. **136** (1967) 101.

[50] Shu, F. H. 1978, ApJ, 225, 83

[51] Nakamura, T. K. 2000, ApJ, 531, 739

[52] T. E. C. Merrall and R. N. Henriksen, Astrophys. J. **595** (2003) 43

[53] Diemand, J., Moore, B., & Stadel, J. 2004, MNRAS, 352, 535

[54] A. Klypin, A. V. Kravtsov, J. Bullock and J. Primack, Astrophys. J. **554** (2001) 903

[55] Ascasibar, Y., Yepes, G., Gottlöber, S., Müller, V. 2004, MNRAS, 352, 1109

[56] J. E. Taylor and J. F. Navarro, Astrophys. J. **563** (2001) 483

[57] Rasia, E., Tormen, G., & Moscardini, L. 2004, MNRAS, 351, 237

[58] L. L. R. Williams, C. Austin, E. Barnes, A. Babul and J. Dalcanton, PoS BDMH **2004** (2004) 020 [astro-ph/0412442].

[59] Peirani, S., Durier, F., & de Freitas Pacheco, J. A. 2006, MNRAS, 367, 1011

[60] Toth, G., & Ostriker, J. P. 1992, ApJ, 389, 5

[61] Bryan, G. L., & Norman, M. L. 1998, ApJ, 495, 80

[62] S. Peirani and J. A. de Freitas Pacheco, [astro-ph/0701292].

[63] P. Sikivie, I. I. Tkachev and Y. Wang, Phys. Rev. D **56** (1997) 1863

[64] Wojtak, R., Łokas, E. L., Gottlöber, S., & Mamon, G. A. 2005, MNRAS, 361, L1

[65] Hoyle, F. 1949, MNRAS, 109, 365

[66] Catelan, P., & Theuns, T. 1996, MNRAS, 282, 455

[67] Fall, IAU Symposium 71, 1983, p.391

[68] M. Steinmetz and J. Navarro, Astrophys. J. **513** (1999) 555

[69] J. S. Bullock, A. Dekel, T. S. Kolatt, A. V. Kravtsov, A. A. Klypin, C. Porciani and J. R. Primack, Astrophys. J. **555** (2001) 240

[70] M. Vitvitska, A. A. Klypin, A. V. Kravtsov, R. H. Wechsler, J. R. Primack and J. S. Bullock, Astrophys. J. **581** (2002) 799

[71] S. Peirani, R. Mohayaee and J. A. de Freitas Pacheco, Mon. Not. Roy. Astron. Soc. **348** (2004) 921

[72] Davis, M., & Peebles, P. J. E. 1983, ApJ, 267, 465

[73] Hale-Sutton, D., Fong, R., Metcalfe, N., & Shanks, T. 1989, MNRAS, 237, 569

[74] Sandage, A., Tammann, G. A., & Hardy, E. 1972, ApJ, 172, 253

[75] Ekholm, T., Baryshev, Y., Teerikorpi, P., Hanski, M. O., & Paturel, G. 2001, A&A, 368, L17

[76] Karachentsev, I. D., Sharina, M. E., Makarov, D. I., et al. 2002, A&A, 389, 812

[77] Karachentsev, I. D., Makarov, D. I., Sharina, M. E., et al. 2003, A&A, 398, 479

[78] Tikhonov, A. V., & Klypin, A. 2009, MNRAS, 395, 1915

[79] Macciò, A. V., Governato, F., & Horellou, C. 2005, MNRAS, 359, 941

[80] Martinez-Vaquero, L. A., Yepes, G., Hoffman, Y., Gottlöber, S., & Sivan, M. 2009, MNRAS, 397, 2070

[81] Aragon-Calvo, M. A., Silk, J., & Szalay, A. S. 2011, MNRAS, 415, L16

[82] Peirani, S., & de Freitas Pacheco, J. A. 2006, New Astronomy, 11, 325

[83] Peirani, S., & de Freitas Pacheco, J. A. 2008, A&A, 488, 845

NUMERICAL SIMULATIONS OF THE DARK UNIVERSE

Hermano Velten[a1] and Luciano Casarini [b2]

[1] *Universitaetstrasse 25,*
Bielefeld, 33615, Germany
[2] *Universidade Federal do Espírito Santo,*
Departamento de Física
Av. Fernando Ferrari,
514, Campus de Goiabeiras,
CEP 29075-910, Vitória,
Espírito Santo, Brazil

This short review on numerical simulations of structure formation of the universe, partially follows the lecture notes of Jean-Michel Alimi at JPB School of Cosmology in UBU - Anchieta, Brazil, October 14-19, 2012. We revise the techniques envolved in cosmological simulations, paying particular attention to the recent results of the DEUS collaboration which is the first project able to simulate the dynamics of dark matter particles in box sizes compared to the Hubble volume.

I. INTRODUCTION: WHY ARE WE TALKING ABOUT A DARK UNIVERSE?

For more than 80 years, astronomy has accumulated evidence that the dominant form of matter in the universe is dark and non-baryonic. Only 15% of the observed matter would be in the form of known standard particles. In a scheme where particles are classified according to their masses, starting with zero mass particles like photons and passing through leptons and mesons, baryons represent the most massive particles in nature. Therefore, such 15% known matter is usually called "baryons".

If we are really interested in the dynamics of the matter in the universe it is necessary to fully understand its dark part. Dark matter came to the attention of astrophysicists due to discrepancies between the mass of large astronomical objects determined from their gravitational effects, and the mass calculated from the luminous

[a] E-mail: velten@physik.uni-bielefeld.de
[b] E-mail: casarini.astro@gmail.com

matter they contain such as stars, gas and dust. Most spiral galaxies show flat rotation curves out as far as we can trace them, even where no more stars are visible. We conclude that the gravitational mass is more than 10 times more massive than the luminous mass.

With the help of lensing measures in galaxy cluster, by measuring how the background galaxies are distorted by the foreground clusters, astronomers can measure the mass of the gravitational lens. The mass in the cluster is more than five times larger than the inferred mass in visible stars, gas and dust. The observation of the Bullet cluster is one of the most famous evidences in favor of the existence of dark matter in the universe. The comparison of images of the gas obtained by the Chandra X-ray telescope to maps of the gravitational field deduced from weak lensing observations shows a clear separation of dark matter and gas clouds.

Apart from other several astronomical observations which confirm the dark matter paradigm, since more than a decade ago we have solid observational evidence of cosmic acceleration, requiring yet another mysterious contribution to the total energy budget of the universe, the dark energy. In order to obtain a concordance cosmological model which is based on the existence of two new dark components (matter and energy) the different data-sets have to be fitted simultaneously. The best fit region in the free parameter space of modern cosmologies are fully consistent with a universe where 70% of its today's cosmic energy density is in the form of dark energy whilst the matter contribution, baryons plus dark matter, covers the remaining 30%.

The current challenge for cosmologists is to understand the dark matter clustering process and the subsequent accretion of baryons when the presence of dark energy is taken into account as a background component that acts against structure formation. The standard view for structure formation assumes a hierarchical scenario where small structure form first. The merger events of small structures give rise to larger clumps.

The new generation of cosmological simulations, like the DEUS project, are fully exploring the consequences of dark energy on the clustering process at very large volumes. Differently to the the classical scenario where dark matter is described by a non-interacting pressureless fluid, it is also important for the next years to simulate different types of dark matter as for example warm dark matter.

In some sense, the ΛCDM scenario depends on the confirmation that alternative models fail in describing the observations properties of the universe. Since numerical simulations represent a very important tool for modern cosmology, it is timely to expand the number of simulated models.

II. NUMERICAL SIMULATIONS OF THE DARK UNIVERSE: DARK MATTER

A crucial question is what is the nature of the Dark Matter? As the signatures of the nature of DM and DE depend on the detailed distribution of DM (maybe of DE too) throughout the universe, the content of the dark universe and the origin of the cosmic structure are deeply related.

The origin of structures is a fundamental question in cosmology. The universe is spatially homogenous and isotropic and it is described having the General Relativity as the fundamental theory for gravity. Moreover, the dynamics of the expansion is given only by the scale factor $a(t)$ related in the FLRW metric.

A summary of the structure formation picture is given as follows. Primordial perturbations are seeded by quantum fluctuations and are stretched to cosmological scales by inflation. They leave an observable imprint on the CMB sky. This enables us to set the initial conditions for cosmology, i.e. the shape and the amplitude of these perturbations. Their growth, via gravitational instability in an expanding universe, leads to the cosmic structures we observe today. The gravitational interactions are modelled using quasi-Newtonian laws (taking into account the cosmic expansion). Indeed, the Newtonian approximation is appropriate when we are dealing with sub-horizon modes.

Structure formation models attempt to reduce cosmology to an initial value problem. The purpose of cosmological simulations is to model the growth of structures in the universe and is a very powerful theoretical tool to interpret and to support present and future cosmological observations.

In order to shortly review the status of numerical simulations this contribution is organized around three main questions:

1. Why do we need to perform numerical simulations in cosmology and what do we simulate?

2. How do we perform numerical simulations of cosmic structure formation?

3. What have we learned from numerical simulations on cosmic structure formation?

Many lectures on cosmic structure formation are available. Current reviews can be found in [1, 2]. We shall discuss the above aspects but focusing on the logic of the development of numerous methods for cosmological numerical simulations. To show up the links between these methods we shall give also some remarks on optimization process, supercomputer architecture, grand Challenge Project (DEUS) and some results on large scale structure, halo profile, Non Linear Gravity/Cosmology, Universality/Non-Universality.

A. Gravitational instability in an expanding universe and N-body collisionless gravitational dynamics

Let us compare some time scales. A system of N particles interacting gravitationally with total mass M and a given size R reaches a dynamic equilibrium state on a timescale comparable to a few times the typical time t_{cr} needed for a particle to cross the system

$$t_{cr} \approx \frac{1}{\sqrt{GM/R^3}}. \tag{5.1}$$

The quantity t_{cr} is the response time needed to settle down to equilibrium. If the system is initially out of equilibrium this is reached through mixing in phase space due to fluctuations of the gravitational potential, a process called violent relaxation [3].

Once the system is in dynamical equilibrium a long term evolution is possible. This process would be driven by two-body relaxation, where the energy is slowly exchanged between particles, and the system tends to evolve towards energy equipartition. The time scale t_{rel} for this process depends on the number of particles and on the geometry

of the system [4],

$$t_{rel} \approx \left[\frac{N}{10Log(N)}\right] t_{cr}. \tag{5.2}$$

Assuming, for example, $N = 10^{12}$, which is the typical number of objects in a galaxy like the Milk Way, a simple estimation of the above quantity gives $t_{rel} \sim 1.2 \times 10^{12} t_{cr}$. N-body systems such as galaxies and dark matter halos have therefore a relaxation time much longer than the life of the universe and are thus considered collisionless systems.

A system of N particles interacting gravitationally defines a $6N + 1$ dimensional phase space given by the N position and velocity vectors associated to each particle at each time t. The solution of the N-body problem defines a trajectory in this phase space.

As the number of dark matter particles in cosmic structure is large enough i.e., if the two body relaxation time is long compared to the time-frame one is interested in, they just respond to their collective gravitational field and a statistical description of them is possible. This allows us to pass from a $6N + 1$ to a $6 + 1$ dimension phase space.

B. Mean Field approximation: Vlasov-Poisson equations and N-body system

We adopt a mean field description of the dynamical system in terms of a single particle distribution function $f(r, \mathbf{v}, t)$ where $f(r, \mathbf{v}, t)drd\mathbf{v}$ is proportional to the probability of finding a particle in a $6D$ element of volume $drd\mathbf{v}$ centered around a position r and a velocity $\mathbf{v}$ at time t.

Within this simplified framework the knowlodge of the distribution function uniquely defines all the properties of the system. The dynamics is then described by the Vlasov-Poisson system

$$\frac{Df}{Dt} = \frac{\partial f}{\partial t} + \vec{v}.\frac{\partial f}{\partial \vec{r}} - \frac{\partial \phi}{\partial \vec{r}}.\frac{\partial f}{\partial \vec{v}} = 0, \tag{5.3}$$

$$\nabla^2 \phi(\vec{r}, t) = 4\pi G \rho(\vec{r}, t), \tag{5.4}$$

where the density of finding a particle is

$$\rho(\vec{r},t) = \int f(\vec{r},\vec{v},t)d\vec{v}. \tag{5.5}$$

Equation (5.3) is also known as the collisionless Boltzmann equation. Note also that a phase space is conserved along each characteristic (i.e. particle trajectory).

The system of partial differential equations is very difficult (impossible) to solve directly in non-trivial cases. Given its high dimensionality $(6+1)$, Vlasov equation is usually solved by sampling the initial distribution function and then by evolving the resulting N-body system by means of numerical methods that suppresses two body interactions at small scales. In other words, the interaction is softened.

Since the Vlasov equation states that f is constant along any trajectory $\{r(t), \vec{v}(t)\}$, the trajectories obtained from the time integration of N points $\{r_i, \vec{v}_i\}$ sampled from the distribution function f at time $t = t_{initial}$ forms a representative sample of f at each time t. Hence the problem reduces to solving Poisson's equation (5.4) for a set of N particles and advancing them forwarding in time according to their equations of motion.

In the N-body approach, the initial phase space is splitted into small domains, each domain is represented by a (macro-) particle, and (macro-) particles are evolved self-consistently using equations of (macro-) particle motion. The most important equation of motion is the velocity/acceleration equation

$$\frac{d\vec{r}}{dt} = \vec{v} \qquad \rightarrow \qquad \frac{d\vec{v}}{dt} = -\nabla\phi. \tag{5.6}$$

It is assumed then that particles are tracers of the density field.

The potential ϕ is given by,

$$\phi(\vec{r}) = -\sum_{j\neq i} G\frac{m_j}{\sqrt{|\vec{r}-\vec{r}_j|^2+\epsilon^2}} \qquad \text{or} \qquad \nabla^2\phi = 4\pi G\rho_T. \tag{5.7}$$

The form used in left expression (5.7) for the potential is called Plummer Softening and ϵ is the gravitational softening length that is a typical distance below which the gravitational interaction is "suppressed". It avoids any divergence for close particles, however it limits the spatial resolution to 2ϵ. The second way to calculate the potential,

right form in (5.7), is via the Poisson equation which calculates ϕ directly from the total density ρ_T.

In some sense, the N-body approach can be seen as a Monte Carlo technique for random sampling of the characteristics. Equations of motion are used in each step of N-body codes. The following list summarizes the sequence of actions (steps) used in N-body codes:

1) compute the mass density on the mesh using an interpolation scheme from the position of particles. A Cloud-In-Cell (CIC) scheme is used;

2) compute the potential on the mesh using the field equation. In gravitational case, it is the Poisson equation which is solved by a multigrid method;

3) compute the acceleration on the mesh. A standard finite-difference approximation of the gradient is used;

4) compute each particle acceleration using the inverse interpolation scheme used in the first step;

5) update each particle velocity according to its acceleration;

6) update each particle position according to its velocity;

7) check energy conservation equation and modification of the time-step.

8) return to step 1).

Remember also that the above procedure follows the growth of primordial fluctuations in an expanding universe. Then, it is convenient to re-write the equations in comoving variables $\mathbf{x} = a(t)^{-1}\mathbf{r}$. The scale factor $a(t)$ satisfies the Friedmann equations (homogeneous and isotropic solutions of Einstein equations). If we assume the cosmic matter content to consist of pressureless matter (composed of ordinary baryonic matter and cold dark matter), relativistic matter (photons and neutrinos) and a dark energy component, the expansion rate of the universe, defined in terms of the scale factor by the Hubble parameter $H = \frac{\dot{a}}{a}$ (a dot denotes a derivative with respect to the cosmic time), is given by the Friedmann equation

$$H^2(a) = H_0^2 \left[\frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} + \Omega_{DE}(a) \right] \tag{5.8}$$

with H_0 being the Hubble constant and $\Omega_i = \rho_i/\rho_c$ is the present value of the density

parameters, where $\rho_c = 3H_0^2/(8\pi G)$ is the critical density and i indicates the density contribution of pressureless matter, radiation and Dark Energy, respectively.

After this short introduction about the basic steps concerning the N-body approach in the cosmological scenario, several questions come up now:

- How do we compute the gravitational forces efficiently and accurately? (Numerical techniques by using or not using a Mesh)
- How do we generate appropriate initial conditions? (Primordial Perturbations)

Besides these issues, at any given time, halos exist on a large range of mass-scales and lenght-scales. Then, two conflicting requirements complicate strongly the study of hierarchical structure formation:

1) If the particle mass is too large (or particle number N is too low), gravitational physics can not be correctly resolved and smallest structures (dwarf galaxies) will be missed.

Solution: We want the smallest particle mass and the largest number of particles to resolve internal structure of halos.

2) If the computing volume is too small, fundamental mode goes non-linear soon after halos form and computation can not be meaningfully continued beyound this point. Moreover, rare events, rare objects like very massive objects (rich clusters) or the first halos could be missed.

Solution: We want a large volume to get a representative sample of the universe. In particular, physics/Cosmology impose the largest N and the largest volume.

In the following sections we discuss in more detail such issues.

C. Why are approximate solutions of the non linear gravitational dynamic not sufficient?

Cosmic structure formation can be described using approximations of non-linear gravitational dynamics. Semi-analytical methods have been developed to help solving the non-linear equations. But why do not use such approximations?

In fact, there are three kinds of approximations:

- Approximation of the dynamics of the collapse (Zeldovich and perturbative theory)
- Approximation on the geometry of the collapse (spherical and ellipsoidal collapse)
- Approximation on the statistics (Press Schechter formalism and Excursion Set Theory)

Each approximation is efficient to describe a specific aspect. The analytical approximations are propably the only way to deeply understand the origin of the mass function and the distribution of dark matter. But none approximation is able to calculate and to describe correctly all observables and the complexity of the cosmic structure formation process.

The process of structure formation is really complex. The gravitational collapse in hierarchical models produces filaments, structures at the intersection of these filaments, clumps, etc. By hierarchical we mean that smaller structures form first. Larger structures are formed from the merger process. See Fig. 1 for an illustration. Structures appear at all scales and merge forming larger clumps. The key issue is that the power spectrum only is not sufficient to describe the matter distribution. Higher order description and other statistics are oken needed. We also want to know the number of objects formed, i.e., the halo mass function, the structure of these objects, the halo profile, etc. Only a cross-analysis of numerous observables help us to better understand the formation of cosmic structures and allow to deduce the cosmological model.

Although numerical simulations are not an exact reproduction of the gravitational dynamics, it does not exist any approximate solution which is able to evaluate the evolution of the dark matter, its distribution on large range of scales, the structure of DM halos and the number of DM halos.

In fact, numerical simulations are the only, or at least an effective, way to compute any cosmological observable. Since we evaluate the accuracy with which we are able to calculate these observables, numerical simulations are a powerful way to constrain and to discriminate between cosmological models.

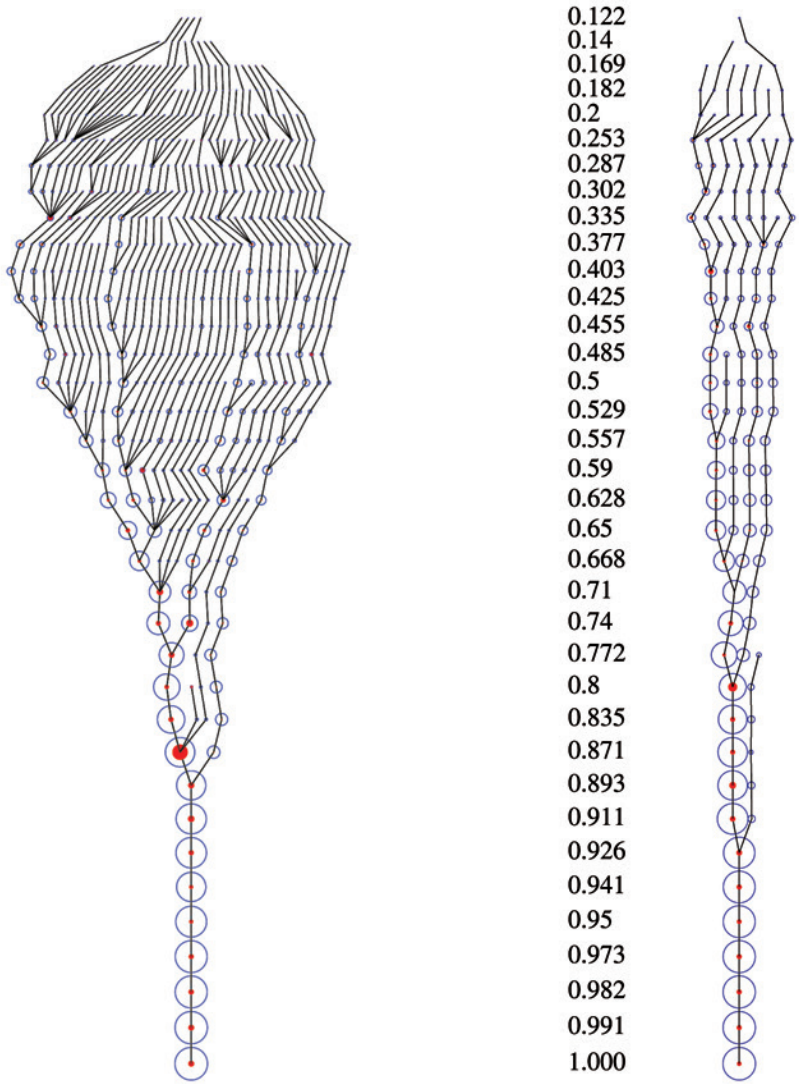


Figure 1. Merging history illustrated for two halos (on the left: cluster-mass $M_{\rm vir} = 2.8 \times 10^{14} h^{-1} \mathrm{M}_{\odot}$; on the right: galaxy-mass $M_{\rm vir} = 2.9 \times 10^{12} h^{-1} \mathrm{M}_{\odot}$ at $a = 1$) as function of the scale factor a listed in the center of the plot. Lines connect halos with their progenitor halos. The radii of the outer and inner (filled) circles are respectively proportional to the virial and inner NFW [11] radii, scaled such that the two halos have equal sizes at $a = 1$. Figure taken from [12]

.

III. HOW CAN WE PERFORM NUMERICAL SIMULATION OF LARGE SCALE STRUCTURE FORMATION

In order to perform simulations it is crucial to discuss some kind of "Gravity Solver" method.

The easiest step is to sum the gravitational forces by direct particle-to-particle interaction.

$$\vec{F}_i = -\sum_{j \neq i} G \frac{m_i m_j (\vec{r}_i - \vec{r}_j)}{|\vec{r}_i - \vec{r}_j|^3}. \tag{5.9}$$

This summation is the essence of PP (particle-particle) algorithms.

The problem is thus a set of non-linear second order ordinary differential equations relating the acceleration $d^2 r_i/dt^2 = F_i/m_i$ with the position of all the (macro) particles in the system. Note also that the calculation is a pairwise operation, and therefore the total time needed to compute all the forces scales as N^2. In general, it is a hard work to integrate systems with particle numbers larger than 10^4.

The gravitational force presents a singularity when the distance of two particles approaches 0, which can lead to arbitrarily large relative velocities. In addition, given the non-linear nature of the equations, the singularities are movable, i.e., they depend on the specific choice of initial conditions (In contrast, all singularities in linear ordinary differential equations are independent of initial conditions). As said before, the singularity may be avoided by introducing the smoothing length that is by modifying the gravitational interaction at small scales such as

$$\vec{F}_i = -\sum_{j \neq i} G \frac{m_i m_j (\vec{r}_i - \vec{r}_j)}{\left(|\vec{r}_i - \vec{r}_j|^3 + \epsilon^2\right)^{3/2}}, \tag{5.10}$$

where the introduction of a $\epsilon > 0$ effectively suppresses binary formation and strong gravitational interactions [5].

All N-body numerical methods in the following introduce in one way or another way such smoothing length.

A state of art, publicy available, serial direct N-body integrator is Aarseth's NBODY6. On one hand, such method is very intuitive, with a very simple first version

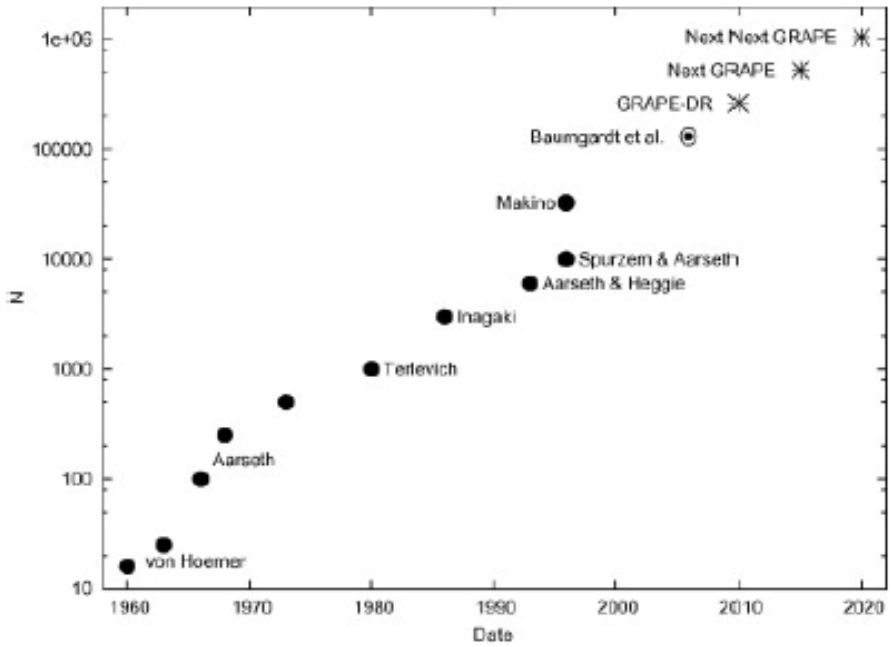


Figure 2. Number of particles used in N-body simulations up to and beyond core collapse, as a function of year of publication. The dot labeled Makino refers to calculations using the GRAPE4. The circled dot denotes work in progress, using the GRAPE-6. The stars indicate predicted years of publication for runs using computers that are currently under development or planned to be built.

and exact forces are calculated on each particle. On the other hand, there are some disadvantages. It takes N^2 summations to calculate forces. Then, doable for $N \sim 10^6$, but impossible for $N \sim 10^9$. It is also difficult to apply periodic boundary conditions.

There is an amazing example oriented to study stellar dynamics: a computer which has been specifically developped and designed for computing the gravitational term is the Gravity Piep (abbreviated GRAPE) that is a project using hardware acceleration to perform gravitational computations. Other versions are GRAPE-2, GRAPE-4, GRAPE-6, ...).

Figure 2 shows the hardware performance evolution, in terms of N, that some projects related to stellar dynamics have reached in the last decades.

Cosmological N-body simulations have grown rapidly in size over the last three decades. It is known that computers double their speed every 18 months (Moore's

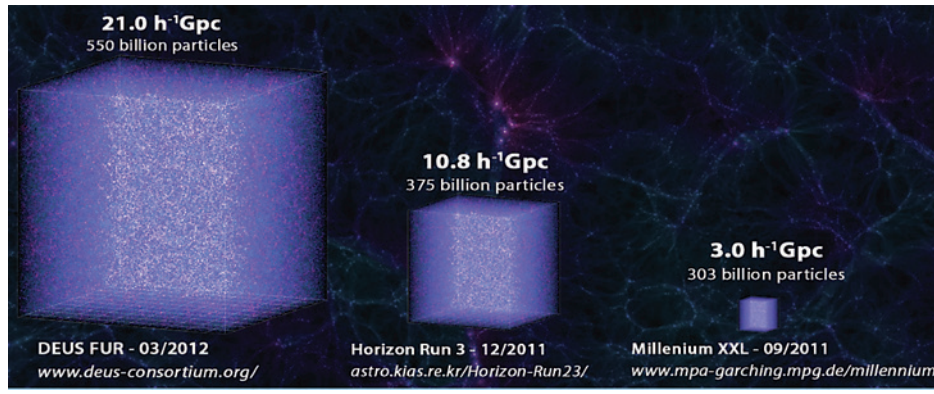


Figure 3. Comparison of some N-Body simulations in terms of particles and volume sizes.

law). N-body simulations have doubled their size every 16 months. Recently, growth has accelerated further. Figure 3 shows a comparison of the volume simulated by some of the most known N-body simulations. In section IV we discuss in more detail the today's largest N-Body simulations with particular attention to the DEUS consortium.

Let us now discuss two usual approximations on the gravitational potential, namely, the Multipole Approximation (the Tree method) and the Harmonic Approximation (Particle Mesh method).

A. Tree Method

This is the first approximation for treating with systems with a large ($> 10^4$) number of particles. It was originally proposed by Barnes and Hut [6].

The simulation volume is usually divided up into cubic cells via an octree (in a three dimensional space), see Fig. 4 for a two-dimensional scheme, so that only particles from nearby cells have to be treated individually (PP interaction), and particles in distant cells can be treated as a single large particle centered at its center of mass (or as low-order multipole expansion). This reduces strongly the number of particle-particle interactions that must be computed.

The space is divided recursively into a hierarchy of cells. The particles are organized in a tree-like structure based upon a cubical decomposition of the computational domain. Consequently, for each particle we "walk the tree" and add the forces from

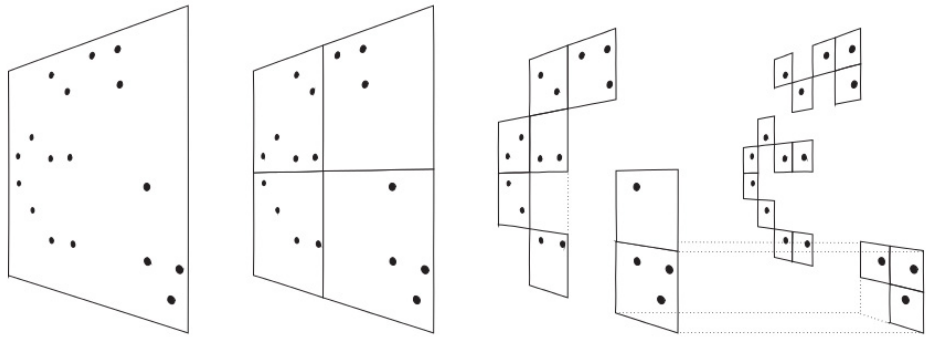


Figure 4. Oct-tree scheme in two dimensions. Figure taken from [7].

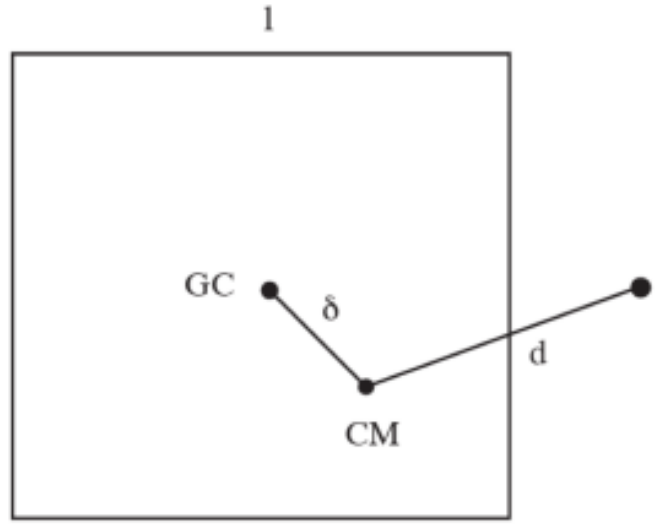


Figure 5. Cell opening criterion.

branchings that need no further unfolding into finer branches according to some pre-selected opening criteria.

The cell opening criterion is used to determine whether a cell is sufficiently distant for a force evaluation via multipoles. The simplest criterion is based on an opening angle parameter θ. We consider a cell where GC is the geometric center of the cell and CM the center of mass of the cell, see Fig. 5. If the size of the cell is l and the distance of the particle from the cell center of mass is d, the algorithm accepts the cell for a force evaluation if $l/d < \theta$. Smaller values of θ lead to more cell openings and

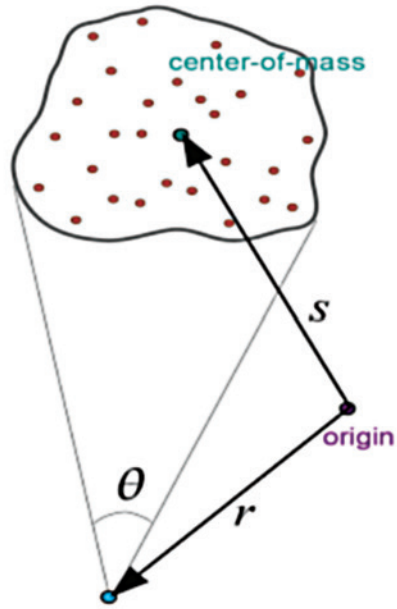


Figure 6. Geometry used in the multipole expansion.

more accurate forces.

One uses hierarchical multipole expansion to account for distant particle groups. Starting with the expression for the potential

$$\phi(r) = -\sum G\frac{m_i}{|r - x_i|}, \tag{5.11}$$

and rewriting it following the geometry shown in Fig. 6,

$$\frac{1}{|r - x_i|} = \frac{1}{|(r - s) - (x_i - s)|}, \tag{5.12}$$

for $|x_i - s| << |r - s|$, we define $y = r - s$ and obtain [1]

$$\frac{1}{|r - x_i|} = \frac{1}{|y|} - \frac{\vec{y}.(s - x_i)}{|y|^3} + \frac{1}{2}\frac{y^T\left[3(s - x)(s - x)^T - I(s - x_i)^2\right]}{|y|^5} + ... \tag{5.13}$$

The multipole moments are computed for each node of the tree. We define the monopole moment $M \sim m_i$, the quadrupole tensor

$$Q_{ij} = \sum m_k\left[3(s - x_k)_i(s - x_k)_j - \delta_{ij}(s - x_k)^2\right], \tag{5.14}$$

[1] The dipole term vanishes when summed over all particles in the group.

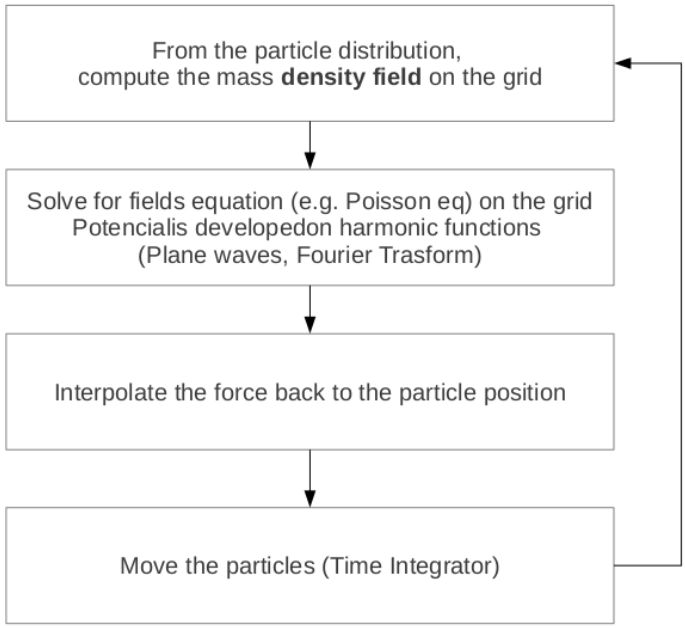


Figure 7. Four main steps involved in PM codes.

resulting in the potential/force approximation:

$$\phi(\vec{r}) = -G\left[\frac{M}{|\vec{y}| + \frac{1}{2}\frac{y^T Q Y}{|y|^5}}\right] : \tag{5.15}$$

For a single force evaluation it is not necessary to compute N single-particle forces, but rather only the order of $LogN$ multipoles, i.e. $NLogN$ operations. Bouchet and Hernquist [8] and Suginohara et al [9] extended the tree code for the periodical boundary conditions adapted for cosmological simulations.

B. Particle-Mesh Methods

With this approach we are able to distinguish two different spaces. The first one is a continuous space where the particles evolve. At the same time, we have a discrete space (grid) where the fields like density, potential and forces are computed. Let us use Fig. 7 to summarize the main steps used in particle-mesh codes.

From the particle distribution we compute the mass density field on the grid. We assign to each particle a shape $S(x)$. Then to each mesh cell we assign the fraction of

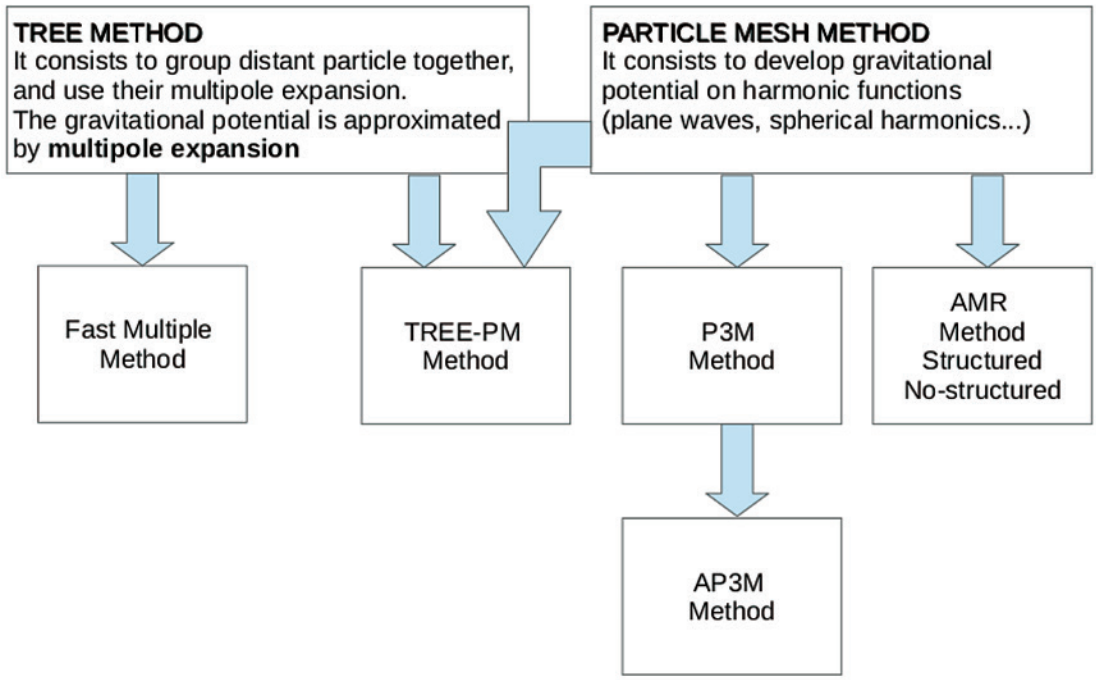


Figure 8. Diagram showing the steps involved in both the Tree and Particle-Mesh methods.

mass that falls into this cell. The overlap for a cell is given by

$$W^{\rho}(x_p - x_i) = \int_{x_i - \Delta x/2}^{x_i + \Delta x/2} S(x_p - x)dx, \tag{5.16}$$

where x_p is the set of discret mesh centres. From the function W the total charge in the cell can be calculated as

$$\rho_i = \frac{1}{\Delta x} \sum m_p W^{\rho}(x_p - x_i). \tag{5.17}$$

The gravitational potential can be then calculated from the density field via the Poisson equation.

This method has been extensively tested and a detailed theoretical analysis of this method is provided in the literature [10]. The evaluation of numerical effects due to the presence of the grid can be accurately performed. Apart from this, there are other advantages for this methods. In general, it is a very fast code since one needs only $O(N_p) + O(N_g log N_g)$ - (N_p is the particle number while N_g is the grid number) -

operations. It is a simple code to write and to parallelize. Consequently, periodic boundary conditions are trivial to implement.

Fig. 8 gives a more complete picture about the procedures envolved in both the PP and PM methods.

C. Particle-Partice/Particle-Mesh Methods

The P^3M approach combines PP and PM methods. At large distances it uses Particle-Mesh to calculate gravitational potentials. At small distances ($< 2dx$) uses tree to calculate local contributions by direct summation (N^2 summation). It is much faster than other codes and has a efficient and high dynamic range.

The Adaptative $P^3M(AP^3M)$ improves the P^3M method by using additional subgrids at regions where clustering occurs (like APM, but with additional particle-particle summation) [13]. See also [14]. An example of such code is the N-body part of Gadget2 [15].

All the previous combined methods (P^3M, AP^3M, TreePM) are considered to be part of high resolution numerical simulations.

D. Generation of Initial Conditions

A typical question that appears in simulations is related to the initial conditions which are used.

Basically, we need to answer four questions:

- What is the shape of the initial fluctuations? This question is related to the physics of the primordial universe and the cosmic microwave background physics.
- What is the amplitude of the initial fluctuations fixed in a cosmological simulation? We could answer to this question by using linear theory and CMB observations.
- How does one generate the initial background of particles in a cosmological simulation?

- How does one generate the initial positions and velocities of particles in a cosmological simulation in agreement with the cosmological initial conditions fixed by observations or theory?

The formation of cosmic structures is seen as a stochastic process. In general, the distribution of cosmic fields is supposed Gaussian. As long as the fluctuations are linear, the power spectrum is only necessary to generate a possible realization of such a process (the extension to non-Gaussian case is possible). The accurate computation of the power spectrum is obtained by taking into account the physics of the early universe and the linear growth of perturbations during the era dominated by the radiation. Usually codes like CAMB or CLASS are used. A fixed set of cosmological parameters determine the shape of the power spectrum $P(k)$.

To determine the power spectrum amplitude we normalize the spectrum to observations of clustering (CMB, galaxy clusters, ...). The filtered density field is

$$\sigma^2(R,z) = D_+^2(z) \int \frac{dk}{2\pi^2} k^2 P(k) \left[\frac{3j_1(kR)}{kR}\right]^2, \tag{5.18}$$

where for a sphere with radius $R = 8h^{-1}$ Mpc the observational input is $\sigma_8 = 0.66-0.9$. The extrapolation back to the starting redshift z_i of the simulation occurs via the linear growth factor $D(z)$ that depends on the cosmology adopted.

To create one realization of the perturbation spectrum we need a model for an unperturbed density field (homogeneous) which introduces some statistical noise (white noise or k-linear noise) on cell scale.

Using the Zeldovich approximation [16] the density field fluctuations are converted to displacements of the unperturbed particle load. This is the case of a first-order (linear) Lagrangian collapse model. This can be expressed by

$$\begin{aligned} \vec{x}(\vec{q},t) &= \vec{q} + D_+(t)\vec{p}(\vec{q}) \\ \vec{v}(\vec{q},t) &= \dot{\vec{x}}(\vec{q},t) = \dot{D}_+(t)\vec{p}(\vec{q}), \end{aligned} \tag{5.19}$$

where q is the initial position of a particle in N-body simulations and $x(q,t)$ is its position at time t and $p(q)$ is a time-independent displacement vector.

In the linear regime the Zeldovich approximation reduces to

$$\delta(t) \approx D_+(t)\vec{\nabla}.\vec{p}(\vec{q}) \rightarrow \vec{\nabla}.\vec{v} = -\dot{\delta}(t). \tag{5.20}$$

The Zeldovich approximation[2] works well up to $\delta \sim 1$, where Eulerian linear perturbation theory breaks down.

We generate the Fourier transform of the overdensity fields

$$\delta_k = |\delta_k|\, exp(i\theta_k), \rightarrow \delta_k = D_+(z)\sqrt{-2P(k, z=0)lnx}; \theta_k = 2\pi y \tag{5.21}$$

where x and y are random numbers uniformely distributed in $[0, 1]$. The inverse Fourier transform is then used to get the real space density fluctuation $\delta_n = \delta(\vec{x}_n)$. To get the velocity field and consequently the displacement field, and then positions and velocities for the particles, we use the linear approximation for the (non-rotational) velocity field

$$v_k = \frac{i\vec{k}}{k^2}\frac{\dot{D}_+}{D_+}\delta_k \tag{5.22}$$

then, FFT^{-1} it to get $\vec{v}_n = \vec{v}(\vec{x}_n)$.

IV. THE DEUS CONSORTIUM

A. High performance numerical simulations

At this point a simple question can arise. Why should we simulate larger and larger volumes. The answer is quite simple. The larger is the volume, smaller is the statistical error (simple variance) in the measured power spectrum. The error σ_k is given by

$$\sigma_k = \sqrt{\frac{1}{N_k}} \propto \sqrt{\left(\frac{\Delta K}{k}\right)^2} \propto \frac{1}{kL_{Box}}, \tag{5.23}$$

where N_k is the number of independent Fourier modes used for estimating the power spectrum.

[2] The Zeldovich approximation has incorrect second- and higher-order growth rates. It is worth considerig higer-order approximation methods if you are doing something particularly sensitive to this [17].

Several groups have pushed to the limits both size box and resolution of cosmological N-body simulations in the last years. The Millenium Simulation in 2005 has run a 2.2 billion light-years simulation box with 10 billion particles [18]. Since then, the performance of cosmological simulations has rapidly increased. More recently, the Millenium-XXL simulation has evolved 303 billion particles in a 13 billion light-years box [19], while the Horizon Run 3 has reached the evolution of 374 billion particles in a 49 billion light-years box [20].

The Dark Energy Universe Simulation (DEUS) [21] have performed the first ever numerical N-body simulations of the full observable universe ($\sim$ 95 billion light years box) for three different cosmological models (two dynamical dark energy scenarions and the standard ΛCDM) to be in agreement with the CMB measurements [22] and SNIa data [23].

For this purpose, a likelihood analysis of the combined Supernova Ia UNION data set and Wilkinson Microwave Anisotropy Probe 5yr data was developed to identify realistic cosmologic models. In this way, all cosmologies reproduce current constraints at the homogeneous and linear level. In order to break degeneracies between dark energy models, we would like to find a new probe of dark energy properties in the non-linear regime of structure formation.

A new version of the Adaptive Mesh Refinement grid cosmological code RAMSES [24] was performed, including: a strongly improvement of the memory occupancy, an optimization of some aspects of parallelization and the addition of the quintessence background evolution.

The simulations have evolved 550 billion particles with more than two and half trillion computing points along the entire evolutionary history of the universe and across 6 orders of magnitudes length scales, from the size of the Milky Way (mass of one particle) to that of the whole observable universe.

Such runs provide unique information on the formation and evolution of the largest structures in the universe and an exceptional support to future observational programs dedicated to the mapping of the distribution of matter and galaxies in the universe. Each simulation has run on 4752 (of 5040) thin nodes of GENCIs supercomputer CURIE, using more than 300 TB of memory for 5 million hours of computing time.

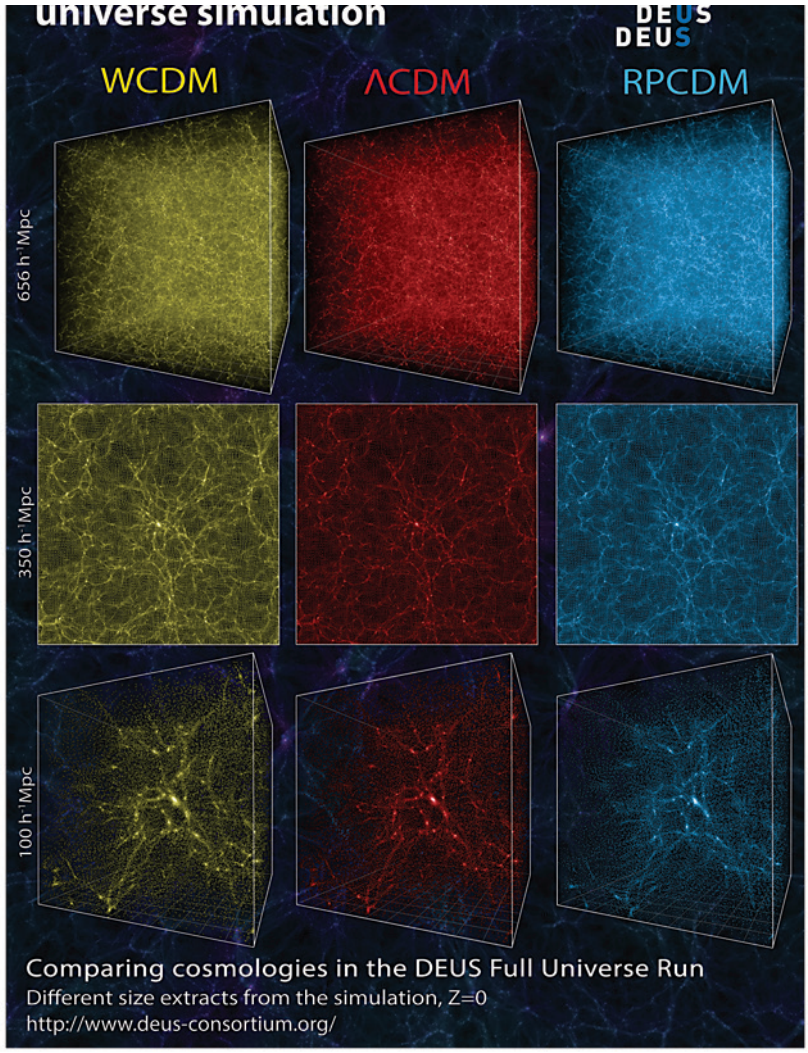


Figure 9. ΛCDM and Dark Energy (Ratra-Peebles and $w = -1.2$) realizations of DEUS simulation.

About 50 PBytes of rough data were generated throughout each run. Using an advanced and innovative reduction workflow the amount of useful stored data has been reduced almost one the fly for each to 500 TBytes. Overall the realization of such large simulations required the development of a global application which integrated all aspects of the physical computational problem: initial conditions, dynamical computing, data validation, reduction and storage. Hence, it required optimizing not only the efficiency of the numerical dynamical solver, but also the memory usage, communications and I/O at the same time.

B. Non linear Power Spectrum

Quintessence has been proposed to account for dark energy in the universe. This component causes a typical modification of the background cosmic expansion, which in addition to its clustering properties, can leave a potentially distinctive signature on large scale structures.

Many previous studies have investigated this topic, particularly in relation to the non-linear regime of structure formation. However, no careful pre-selection of viable quintessence models with high precision cosmological data was performed.

In Alimi et al. [25] it is investigated the non-linear matter power spectrum. The realistic quintessence models allow, for relevant differences of the dark matter distribution with the respect to the ΛCDM scenario well into the non-linear regime, deviations up to 40% in the non-linear power spectrum. Such differences are shown to depend on the nature of DE, as well as the scale and epoch considered. At small scales ($k = 1 - 5hMpc^{-1}$, depending on the redshift) the structure formation process is about 20% more efficient than in ΛCDM. These imprints are a specific record of the cosmic structure formation history in dark energy cosmologies and therefore can not be accounted in standard fitting functions of the non-linear matter power spectrum based on the ΛCDM model only (Fig. 10).

The data generated have also allowed to evaluate with high accuracy the imprint of the BAO (Baryon Acoustic Oscillations) in the density matter power spectrum fluctuations. The position and height of the peaks measured in the simulation with unprecedented precision allows to constrain the cosmological model, especially the amount and the nature of Dark Energy. At small scales (large k-values) we observe the deformation due to non-linear dynamics (Fig. 11).

C. Non Universality of the Mass Function

The universality of the halo mass function is studied in the context of dark energy cosmologies. This widely used approximation assumes that the mass function can be expressed as a function of the matter density Ω_m and the root-mean-square linear

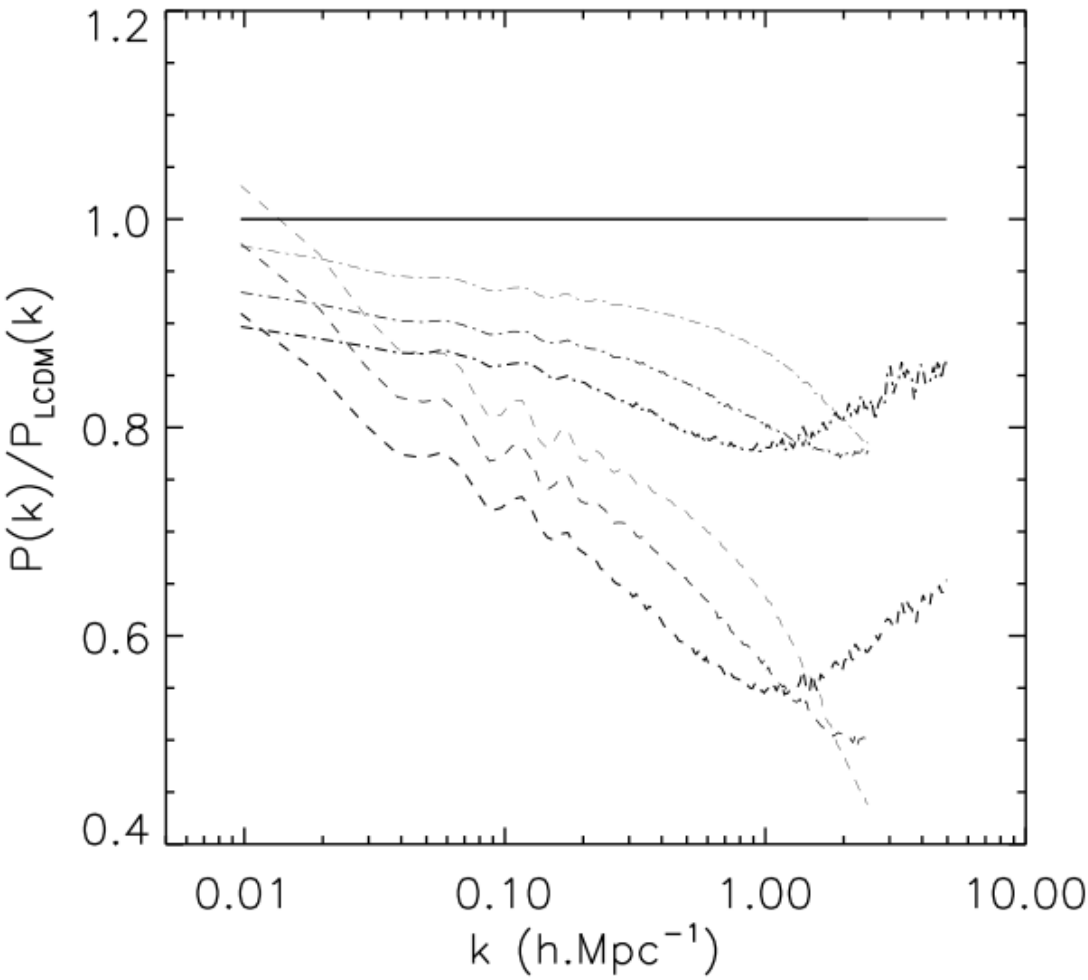


Figure 10. Evolution of the ratio of the power spectrum between RPCDM (dash line) or SUCDM (dash-dot line) and ΛCDM for a=0.3,0.5,1 (from top to bottom). Figure taken from [25].

density fluctuation σ_8 only, with no explicit dependence on the properties of dark energy or redshift.

Courtin et al. [26] show a redshift-dependent deviations from a universal behavior, which are correlated with the linear growth factor of the investigated cosmologies. Using spherical collapse as guiding for the virialization process, it explains that such deviations are caused by the cosmology dependence of the non-linear collapse.

In order to test the universality of the mass function it is preferable to work with

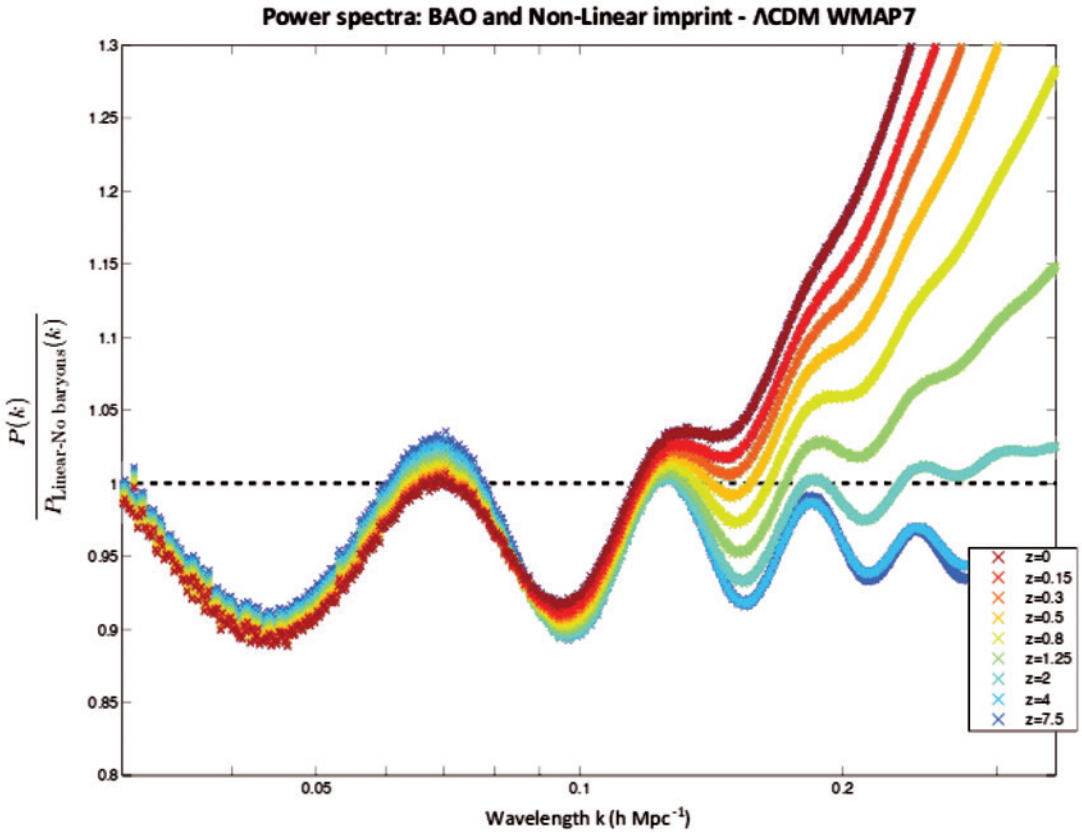


Figure 11. Analysis of the matter power spectrum reveals with unprecedented accuracy the deformations of the primordial plasmas acoustic oscillations at different redshift. Figure taken from [21].

the function

$$f(\sigma) = \frac{M}{\bar{\rho}_0} \frac{dn}{d\ln(\sigma^{-1})}, \tag{5.24}$$

where $dn/d\ln(\sigma^{-1})$ is the comoving number of halos per unit of natural logarithm of σ^{-1}. Fig 12 shows the residuals respect to this parameter. The amplitude of the residuals is a signature of the different dark energy equation of state evolution.

To recover the universality, a fitting formula of the mass function is calibrated with accuracy of 5-10 percents over the all range of investigated cosmologies, using the well know functional form of Sheth and Tormen [27]:

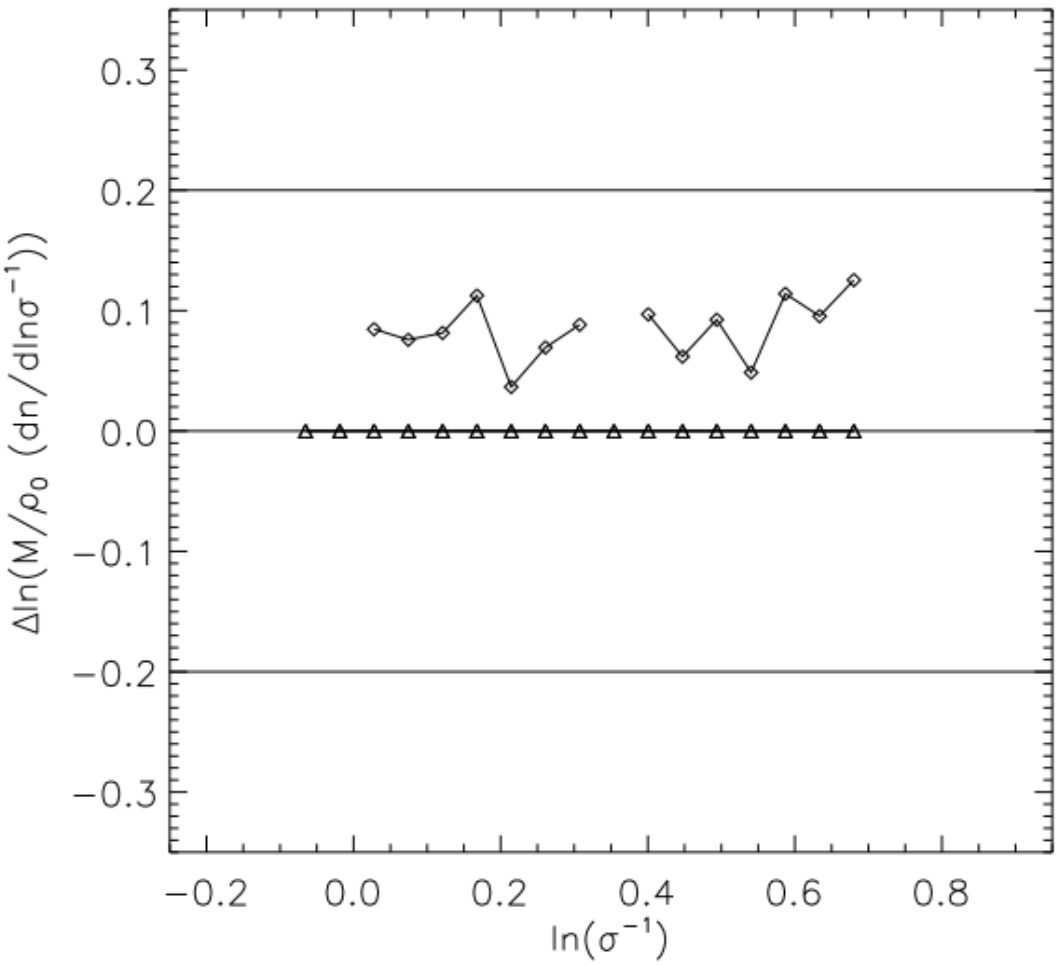


Figure 12. The residuals between the mass functions of RPCDM (diamonds) and the mass functions for ΛCDM (tri-angles) cosmology at z=0, present deviations from universality clearly above the numerical errors and correlated with linear growth history. Figure taken from [26].

$$f(\sigma) = \tilde{A}\left(\frac{2\tilde{a}}{\pi}\right)^{1/2}\frac{\tilde{\delta}_c}{\sigma}\left[1+\left(\frac{\tilde{\delta}_c}{\sigma\sqrt{\tilde{a}}}\right)^{-2\tilde{p}}\right]e^{-\tilde{\delta}_c^2\tilde{a}/(2\sigma^2)}, \tag{5.25}$$

with $\tilde{A} = 0.348$, $\tilde{a} = 0.695$ and $\tilde{p} = 0.1$. The esplicit dependence of δ_c, predicted by the cosmology, leads to different velocity in the clustering during the structure formation.

In addition, it has provided an empirical relation between the FoF (Friends of Friends) linking parameter and the virial overdensity which can account for most of the

deviations from an exact universal behavior. In this way the FoF algorithm permits to recover the universal form of mass fuction and the virial overdensity at a given redshift for each cosmology.

This suggest that, probably the numerical precision of DEUS simulations will allow to discover other hidden non-linear mechanisms, contributing to further shaping of the mass function.

V. CONCLUSIONS

- The numerical simulation activity is rapidly evolving, as computational resources continue to grow, enabling larger, more complicated, and more realistic simulations to be performed.

- As dark matter dynamics can be studied with high precision in cosmological N-body simulations. The simulation allows high-precision measurements of the dark matter clustering, like halo mass function, power spectrum, bulk flow, halo profile, abundance of dark matter substructure, etc.

- The simulation is a powerful theoretical tool to probe the nature of dark matter and dark energy.

- The simulation is a powerful theoretical tool to support present and future cosmological observations (redshift space).

- The next decade promises to be filled with exciting challenges and the potential for great discoveries. It is safe to say that the field of numerical simulations of the dark universe will not run out of things to do.

ACKNOWLEDGMENTS

We are grateful to Jean-Michel Alimi for making his notes available to us. Hermano Velten and Luciano Casarini acknowledge the Brazilian research institutions FAPES

and CNPq for their financial support.

---

[1] M. Kuhlen, M. Vogelsberger and R. Angulo, Physics of the Dark Universe, **1**, 1-2, 50 (2012).

[2] M. Baldi, Physics of the Dark Universe, **1**, 1-2, 162 (2012).

[3] D. Lynden-Bell, MNRAS, **136**, 101 (1967).

[4] L. S. Spitzer, *Dynamical Evolution of Globular Clusters*, Princeton University Press (1987).

[5] S. J. Aaserth, Bulletin of the Astronomical Society, **4**, 417 (1972).

[6] J. Barnes and P. Hut, Nature, **324**, 446 (1986).

[7] V. Springel, N. Yoshida, & S. D. M. White, New Astron., **6**, 79 (2001)

[8] F. R. Bouchet and L. Hernquist, ApJS **68**, 521 (1988).

[9] S. Suginohara et al, Astrophys. J. Suppl., **75**, 631 (1991).

[10] R.W Hockney and J.W Eastwood, *Computer Simulation Using Particles*, CRC Press (1988).

[11] F. J. Navarro, C. S. White, S. D. M. White, ApJ **490**, 493 (1997).

[12] R. H. Wechsler et al., ApJ **568**, 52 (2002).

[13] H. M. P. Couchman, Astrophys. J., **368**, 23 (1991).

[14] G. Xu, Astrophys. J. Suppl., **98**, 355 (1995).

[15] V. Springel, MNRAS, **364**, 1105 (2005).

[16] Y. B. Zeldovich, A&A, **5**, 84 (1970).

[17] M. Crocce et al., MNRAS **373**, 369 (2006).

[18] V. Springel et al., Nature, **435**, 629 (2005).

[19] R.E. Angulo, et al., MNRAS, **426** 2046(2012).

[20] J. Kim et al, Journal of The Korean Astronomical Society, 44.6.217 (2011).

[21] J.-M. Alimi et al., IEEEComputer Soc. Press, CA, USA, SC, 73 (2012).

[22] E. Komatsu et al., Astrophys.J.Suppl., **180**, 330 (2009).

[23] M. Kowalski et al., Astrophys.J. **686**, 749 (2008).

[24] R. Teyssier, Astronomy and Astrophysics, **385**, 337 (2002).

[25] J,-M. Alimi et al, MNRAS, **401**, 775 (2010).

[26] J. Courtin et al, MNRAS **410**, 1911, (2011).

[27] R. K. Sheth and G. Tormen, MNRAS, **308**, 119 (1999).

Seminars

CMB and the large-scale structure formation in the Universe

Armando Bernui[1]

[1] *Observatório Nacional*
Rua General José Cristino 77, São Cristóvão,
20921-400, Rio de Janeiro – RJ, Brazil

Current astrophysical observations confirms the concordance cosmological model ΛCDM. In this model, the primordial universe underwent an early inflationary phase that generate primordial cosmological perturbations which are regarded as the seeds for structure formation in the universe. A primordial phenomenon, termed Baryon Acoustic Oscillations (BAO), left imprints both in the cosmic microwave background (CMB) and -later on- in the large-scale distribution of galaxies. They originated in primordial matter density fluctuations that do not collapse completely by gravity but instead oscillate due to radiation pressure, and spherical sound waves were driven into the surrounding medium. At the baryon drag epoch (soon after decoupling era) the photons stream away from the baryons and the sound waves freeze out leaving the baryons in the spherical shell, with a radius that in comoving distance is around $105\,h^{-1}$ Mpc. In this scenario, the large-scale structure of the universe evolves *via* gravitational instability from a mixture superposition of these spherical shells, with galaxies forming preferentially in regions of dark matter overdensities where baryons fall, but also with a slight probability to form galaxies in shell's regions where baryons overdensities accumulate and dark matter falls.

I. INTRODUCTION: THE EARLY UNIVERSE

In the current cosmological scenario, the large-scale structure formation is one of the most interesting and complicated issues. At early times, matter and radiation were in thermal equilibrium and due to the tight coupling between baryons and photons they perform acoustic oscillations as a single fluid. We shall describe the role played by the cosmic background radiation in this process, and emphasize the importance of such oscillations in the structure formation process occurred by gravitational instability after matter-radiation decoupling [1–3].

A. Cosmological perturbations in the early Universe

In the standard Friedmann-Lemaître-Robertson-Walker (FLRW) cosmological model, the primordial universe underwent an early cosmological inflation phase where the primordial comoving curvature perturbation ζ, generated *via* quantum fluctuations, seeded the curvature perturbation Φ of the background spacetime, that is $\Phi \sim \zeta$ [4],[1]

$$ds^2 \,=\, a^2(\eta)\,\left[\,-(1+2\Psi)\,d\eta^2 \,+\, (1-2\Phi)\,d\bar{x}^2\,\right]\,, \tag{7.1}$$

where a is the scale factor, η is the conformal time, Ψ is the gravitational potential and Φ is also termed the Newtonian curvature. In the absence of sources of anisotropic stress or viscosity one has $\Phi = \Psi$. In turn, these cosmological perturbations of the background spacetime, $\partial g_{\mu\,\nu}$, originated the matter seeds ($\delta\rho$) for structure formation in the universe. This set of relationships, not all simultaneous in time, can be summarized as follows

$$\zeta \sim \Psi \sim \frac{\delta\rho}{\bar{\rho}} \sim \frac{\Delta T}{\bar{T}}\,, \tag{7.2}$$

where the last relationship concerns the features where baryonic matter and radiation were tightly coupled and in thermal equilibrium, a phase that we shall describe below.

[1] $\zeta = (5/3)\Phi$, in the matter dominated era, where Φ is called the Bardeen's curvature perturbation.

The important thing here is that, at the end of inflation, density matter perturbations $\delta\rho$ were produced by curvature perturbations of the spacetime. In fact, originally introduced as a possible solution to the cosmological puzzles such as the entropy, flatness and horizon problems, by far the most useful property of inflation is that it generates the spectrum of matter density perturbations.

B. After Inflation

According to the Friedmann equations, after the inflationary phase the universe was dominated by the radiation species, this epoch is termed radiation dominated universe or RD, and the scale factor behaves like $a(t) \sim t^{1/2}$. Afterwards the universe continuous to expand, the temperature decreases and then the matter density, ρ_m, (corresponding to non-relativistic species) equals the radiation density, ρ_r at the so-called equipartition time: $\rho_m(t_{eq}) = \rho_r(t_{eq})$. In sequence of this, matter started to dominate over the other components, this era is known as matter dominated universe or MD, where the scale factor behaves like $a(t) \sim t^{2/3}$. One can observe these features in figure 1.

From other side, the Hubble radius (also termed Hubble horizon), c/H, behaves like $\sim t$, where $H \equiv \dot{a}/a$ is the Hubble function. Since the physical scales in the universe, λ, are proportional to the scale factor a, we then have $\lambda \sim a$. Therefore, there was a period in the universe where physical scales were out of the horizon $\lambda \sim a \sim t^{1/2} > t \sim c/H$, these are called the *super-horizon scales*. An schematic representation of the evolution of these quantities can be seen in figure 2.

In fact, during inflation the scale factor grows quasi-exponentially, while the Hubble radius remains almost constant. Consequently the wavelengths of the perturbations [2] soon exceed the Hubble radius, and perturbation modes are out of the horizon. Once inflation has ended, the Hubble radius increases faster than the scale factor $a \sim t^{1/2}$, so the fluctuations reenter the Hubble radius during the radiation or matter dominated eras, that is before or after the equipartition time t_{eq}. According to

[2] Each wavelength corresponds to a perturbation mode in harmonic space.

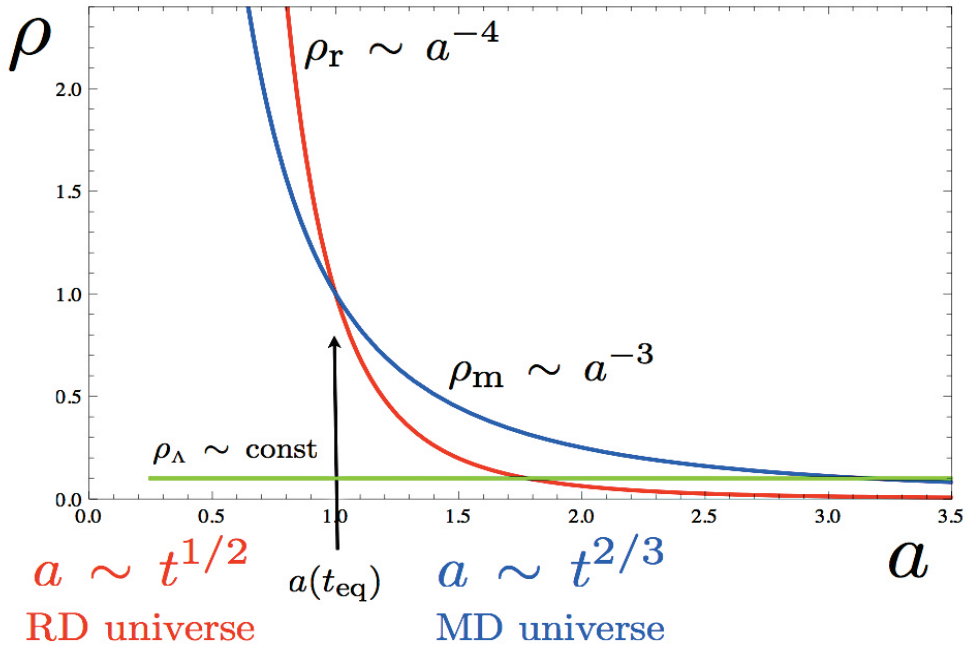


Figure 1. Evolution of the density components in the universe, namely ρ_r, ρ_m, and ρ_Λ, for the radiation, the (baryonic plus dark) matter, and the cosmological constant, respectively.

current cosmological model, the fluctuations that exit around 60 e-foldings or so before reheating reenter with suitable physical wavelengths [4, 5].

C. The Universe before the matter-radiation decoupling

The primordial universe was hot, dense and ionized. At that epoch, photons and baryonic matter are tightly coupled by Thomson elastic scattering. Therefore, the short mean free path, τ (mfp), of the photons allows a fluid approximation and baryonic matter and radiation can be well described as a *matter-radiation* single fluid.

Since matter and radiation were in thermal equilibrium, the cosmic background radiation (CMB) should have a black-body spectral distribution, with $\bar{T} = 2.725 \pm 0.002$ K as predicted by Tolman in 1934 and confirmed by the COBE satellite in 1992.

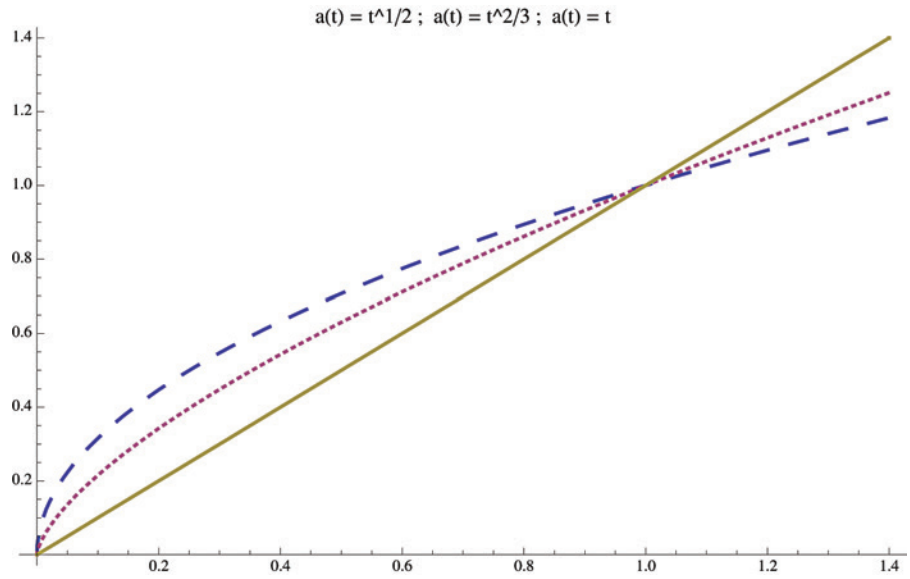


Figure 2. The dashed-line corresponds to the case $a(t) \sim t^{1/2}$, the dotted-line to the case $a(t) \sim t^{2/3}$, and the continuous line to the Hubble horizon (also termed Hubble distance) $c/H \sim t$.

If deviations are created, then free-free interactions (also known as bremsstrahlung radiation) provided thermalization at $z > 3\times10^7$. Afterwards, the free-free interaction time scale becomes longer than the expansion timescale H^{-1} and this process ends. Elastic Thomson scattering interaction has a mean free path $\lambda = 8.3\, H^{-1}/[x_e\,(1+z)]$, where x_e is the ratio of the number of free protons (hydrogen ions) over the number of hydrogen atoms, in this way the value x_e is quantifying how much baryonic matter is still coupled to radiation. Thus, as long as the plasma is ionized, $x_e = 1$, which happens at $z > 1100$, thus the universe is opaque; from the other side, when the value $x_e = 0$ means that there are no free protons because they combined with free electrons to form neutral hydrogen. At $t \lesssim 3\,000$ K, x_e falls quickly ($\Delta z \sim 80$) being $x_e < 10^{-2}$ at $z \lesssim 1\,100$, thereafter the universe becomes transparent, till reionization by galaxies. The Thomson optical depth, τ, from $z \sim 1\,100$ until reionization $z \sim 20$ is rather weak, i.e., $\tau \sim 0.1$, in this way one expects only small secondary distortions.

Let us now briefly introduce some basic properties of the cosmic microwave background radiation [6].

II. THE COSMIC BACKGROUND RADIATION

A. Black-body radiation: a brief introduction

A simple look at a night sky shows that the universe is plenty of luminous stars, that are systems in hydrodynamical equilibrium where the nuclear fusion of the hydrogen atoms produce heat, i.e. photons. In other words, the universe seems to be full of common matter (i.e. that appearing in the Periodic Table of elements) plus radiation (photons). The present knowledge of the origin of light elements and photons starts in the 40's when the first attempts to describe an evolving universe come out (since it expands, after the Edwin Hubble discoveries, was clear that it evolves). Between 1942 and 1948 George Gamow, Ralph Alpher, and Robert Herman proposed models to explain the abundance of light elements, H and He, in an expanding universe. This model also predicts the existence of a background radiation, of primordial origin and isotropic across the sky, and having a black-body spectrum with temperature between $5 - 50$ K. In 1964-1965 an excess of temperature antenna radiation, isotropic (up to the sensibility of the instruments), of 3.5 ± 1 K was detected in the receptor antenna of 4 080 MHz of the Bell Telephone Laboratories Inc. by A. Penzias e R. Wilson. This result was correctly interpreted by R. Dicke, P. Peebles, P. Roll, and D. T. Wilkinson, as a relic cosmic microwave background radiation previously predicted by R. Alpher and R. Herman in 1948 [7].

When a system is in thermodynamic equilibrium, at temperature T, the energy density of photons in the system is a function of photons energy $h\,\nu$, and it depends only on the value of T. Thus, the energy density of photons is given by

$$\varepsilon(\nu) \;=\; \frac{8\pi h}{c^3}\,\frac{\nu^3}{e^{h\,\nu/k_{\rm B}\,T}\;-\;1}\;. \tag{7.3}$$

Integrating over all the frequencies: $\int \varepsilon(\nu)\,d\nu \equiv \rho_r,\ \nu \in [0,\infty\rangle$, one gets

$$\rho_r = \alpha\,T^4\,, \tag{7.4}$$

known as the Stefan-Boltzmann law, where $\alpha = (\pi^2/15)\,(k_{\rm B}^4/\hbar^3 c^3)$. We can also

calculate the photons number density n_γ using

$$n_\gamma = \int \frac{\varepsilon(\nu)}{h\,\nu}\, d\nu = \beta\, T^3\,, \tag{7.5}$$

where $\beta = (2.4/\pi^2)(k_{\rm B}^3/\hbar^3 c^3)$.

Consider a photon gas in thermodynamic equilibrium, at temperature T, in a FLRW universe [6, 8]. It can be shown that the temperature T of this system, with adiabatic conditions, behaves like $T \sim a^{-1}$, where a is the scale factor, that is the gas temperature decreases as the universe expands. Moreover, interpreting the light emitted by receding galaxies, first measured by E. Hubble in 1929, as Doppler shifted one obtains $a \sim 1/(1+z)$, where $1 + z \equiv \lambda_{obs}/\lambda_{emi}$ is the redshift that measures the universe expansion. Additionally, the physical wavelenghts (frequencies) behaves like $\lambda \sim a$ ($\nu \sim a^{-1}$).

From the equation (7.4) and using $T \sim a^{-1}$, one concludes that the radiation density is $\rho_r \sim a^{-4}$. From other side, using the fact that matter density ρ_m is inversely proportional to the 3-space volume, then $\rho_m \sim a^{-3}$.

Using this result, $T \sim a^{-1}$, in equation (7.5) one concludes that the photon number density $n_\gamma \sim a^{-3}$. From other side, the baryon number density $n_b \sim a^{-3}$ because, as a density quantity is inversely proportional to the 3-space volume, which is $\sim a^3$. Hence the baryon-photon number density ratio, $\eta \equiv n_b/n_\gamma$, is constant with the approximate value $\eta \simeq 10^{-9}$. This means that, for each baryon there are 10^9 photons, in other words, the baryon-photon fluid can be well approximated by a photon gas considering the baryons as a small -*but non null*- perturbation. In this way we justify, a posteriori, that one can consider the tightly coupled primordial photon-baryon fluid as a photon gas in thermodynamic equilibrium in an expanding FLRW universe. It is worth to notice that *baryon f luid* means electrons (which are not baryons), protons and neutrons. The physics of the photon-baryon single fluid (see, e.g., [1] and references therein) is: photons are tightly coupled to the electrons by Thomson scattering which in turn are tightly coupled to the baryons by Coulomb interactions. Thus, for instance, any extra energy of the electrons is redistributed to the energy of the photons (CMB) *via* Thomson scattering.

Here a brief summary of the main discoveries regarding the CMB:

- 1934 Richard Tolman shows that CMB radiation in an expanding universe cools down but keep its thermal distribution and remains as a blackbody.
- 1941 A. McKellar was attempting to measure the average temperature of the interstellar medium, and reported the observation of an average bolometric temperature of 2.3 K based on the study of interstellar absorption lines.
- 1946 R. Dicke predicts "... radiation from cosmic matter" at < 20 K but did not refer to background radiation.
- 1948 G. Gamow calculates a temperature of 50 K (assuming a 3-billion-year old Universe), commenting it "... is in reasonable agreement with the actual temperature of interstellar space", but does not mention background radiation.
- 1948 R. Alpher and R. Herman estimate "the temperature in the Universe" at 5 K. Although they do not specifically mention microwave background radiation, it may be inferred.
- 1956 G. Gamow estimates 6 K.
- 1964-1965 A. Penzias and R. Wilson measure the temperature to be 3.5 ± 1 K. R. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson interpret this radiation as a signature of the hot big bang.
- 1966 R. Sachs and A. Wolfe predict amplitude fluctuations in the CMB created by variations of the gravitational potential between the last scattering surface and the observer
- 1968 M. Rees and D. Sciama predict amplitude fluctuations in the CMB caused by photons crossing time-dependent gravitational potential barriers.
- 1969 R. A. Sunyaev and Y. B. Zel'dovich study the inverse Compton scattering of CMB photons caused by hot electrons, defining the basis of the so-called Sunyaev-Z'eldovich effect.

- 1990 FIRAS/COBE measures the black-body CMB spectrum with superb precision. The current value of this black-body radiation spectrum is $\bar{T} = 2.725 \pm 0.002$ K.
- 1999 First measurements of acoustic oscillations in the CMB angular power spectrum from the BOOMERANG experiment.

III. BARYON ACOUSTIC OSCILLATIONS

A primordial phenomenon, termed Baryon Acoustic Oscillations (BAO), left imprints both in the CMB and -later on- in the large-scale distribution of galaxies. They originated in primordial matter density fluctuations that do not collapse completely by gravity but instead oscillate due to radiation pressure, as a result spherical sound waves were driven into the surrounding medium. At the baryon drag epoch (soon after decoupling era) the photons stream away from the baryons and the sound waves freeze out leaving the baryons in the spherical shell, with a radius that in comoving distance is around $105\,h^{-1}$ Mpc. In this scenario, the large-scale structure of the universe evolves from a mixture superposition of these spherical shells, with galaxies forming preferentially in regions of dark matter overdensities where baryons fall, but also with a slight probability to form galaxies in shell's regions where baryons overdensities accumulate and dark matter falls. The correlated origin of the overdensities in the center and in the shell provide a tiny excess of probability to find pairs of galaxies separated by a distance equal to such comoving radius. In fact, a distance relationship corresponding to the characteristic radius was revealed in the two-point correlation function of galaxy catalogs where it appears as a peak (see, e.g., [9] and references therein).

A. Continuity, Euler, and Poisson equations

Consider a generic instability scenario in which perturbations are generated in the early universe and start to grow under gravity when non-relativistic matter begins to

dominate (MD). The fluid is assumed pressureless and ideal, where particles do not cross. The equations describing the fluid motion are the continuity, Euler, and Poisson equations

$$\dot{\delta} + \nabla \cdot \vec{v} \, = \, 0 \, , \tag{7.6}$$

$$\frac{\partial \vec{v}}{\partial t} + \frac{\dot{a}}{a} \, \vec{v} = -\nabla \phi - c_s^2 \, \nabla \delta \, , \tag{7.7}$$

$$\nabla^2 \Phi = 4\pi G \, \bar{\rho} \, a^2 \, \delta \, , \tag{7.8}$$

where $\bar{\rho}$ is the mean baryonic+dark matter density (is the homogeneous background density, i.e., it depends only on time, $\bar{\rho} = \bar{\rho}(t)$), c_s is the sound horizon velocity of the fluid, and $\delta \equiv \delta\rho/\bar{\rho}$ is the baryonic+dark matter density contrast. Combining these equations one gets

$$\ddot{\delta} \, + \, 2\frac{\dot{a}}{a}\dot{\delta} \, - \, c_s^2 \, a^{-2} \, \nabla^2 \delta \, - \, 4\pi G \bar{\rho} \, \delta \, = \, 0 \, . \tag{7.9}$$

The last term of this equation corresponds to the gravitational attraction which acts to compress the fluid, while the third term accounts for the *radiation pressure* repulsion avoiding the full collapse. The behavior of the solutions of this equation results from the predominance of one of these terms, or both.

It is interesting to look for solutions of the form $\delta(t, \vec{x}) \, = \, \delta_{\vec{k}}(t) \, e^{i \, \vec{k} \cdot \vec{x}}$ for this equation, where $k \equiv |\vec{k}| \, = \, 2\pi a(t)/\lambda$ is the wave number of the perturbation mode and λ is the wavelength of such perturbation. With this ansatz the equation (7.9) converts into an ordinary differential equation with non-constant coefficients

$$\ddot{\delta}_{\vec{k}} \, + \, 2\frac{\dot{a}}{a}\dot{\delta}_{\vec{k}} \, + \, (c_s^2 \, a^{-2} \, k^2 - 4\pi G \, \bar{\rho}) \, \delta_{\vec{k}} \, = \, 0 \, , \tag{7.10}$$

and using the definition of the Jeans wave number

$$k_J^2 \, \equiv \, \frac{4\pi G \, \bar{\rho} \, a^2}{c_s^2} \, , \tag{7.11}$$

then we obtain

$$\ddot{\delta}_{\vec{k}} + 2\frac{\dot{a}}{a}\dot{\delta}_{\vec{k}} + (k^2 - k_J^2)\, c_s^2\, a^{-2}\, \delta_{\vec{k}} = 0\,. \tag{7.12}$$

Equivalently one can also use the Jeans wavelength: $\lambda_J \equiv 2\pi/k_J$. Now we shall look for the time evolution of the solutions $\delta_{\vec{k}}$, which for simplicity we simply denote by δ, satisfying this equation in different epochs of the universe, like RD and MD eras. We are also interested in the behavior of different density modes, like those corresponding to super-horizon scales, $k < k_J$ or $\lambda > c/H \gtrsim \lambda_J$, and sub-horizon scales, $k > k_J$ or $\lambda < \lambda_J$.

1. Super-horizon solutions in the RD universe

Consider the early universe phase when radiation dominates, RD, $a \sim t^{1/2}$, when almost all the modes are out of the horizon $\lambda \gg c/H \gtrsim \lambda_J$, and baryonic (together with radiation) and dark matter components are all coupled and evolve together. In this case, $k \ll k_J$ and the third term of the equation (7.12) is negative and of the form $\sim a^{-4}$ because $\bar{\rho}_r \gg \bar{\rho}_m$, that is the radiation components dominates (RD universe). With this information we look for numerical solutions for $\delta = \delta(t)$ and find out that the density contrast behaves like $\delta \sim t$ as shown in the figure 3.

2. Sub-horizon solutions in the RD universe for dark matter

After the modes enter the horizon it happens that the dark matter decouples from the baryon-photon fluid, and therefore they follow different evolution equations. This is because, for dark matter the radiation pressure term in eq. (7.12) is absent, then for sub-horizon modes $k > k_J$ or $\lambda < \lambda_J$ the dark matter density equation is

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} - 4\pi G\, \bar{\rho}_{\rm DM}\, \delta = 0\,. \tag{7.13}$$

In this way, the third term is negative and proportional to $\bar{\rho}_{\rm DM} \sim a^{-3}$, for this it behaves like $t^{-3/2}$ in the RD universe. Thus, solving numerically this evolution equation for

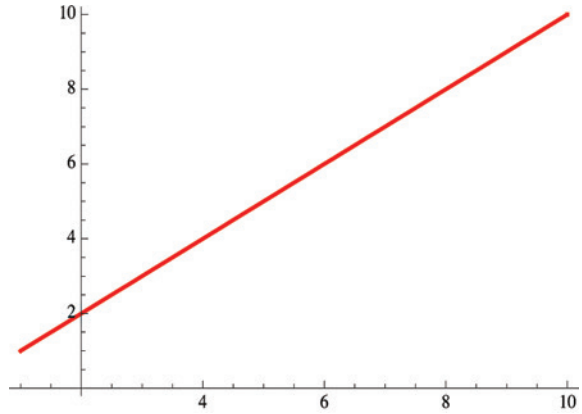


Figure 3. Solving numerically eq. (7.12) in the RD universe epoch, $a(t) \sim t^{1/2}$, and for super-horizon modes $k < k_J$ or $\lambda > c/H \sim \lambda_J$ we find that density contrast perturbations grows like $\delta \sim t$.

the dark matter density contrast we find that it grows like $\delta_{\rm DM} \sim \ln t$ (see figure 4), for all sub-horizon modes $k > k_J$. Observe that the third term in this equation is independent of the k variable.

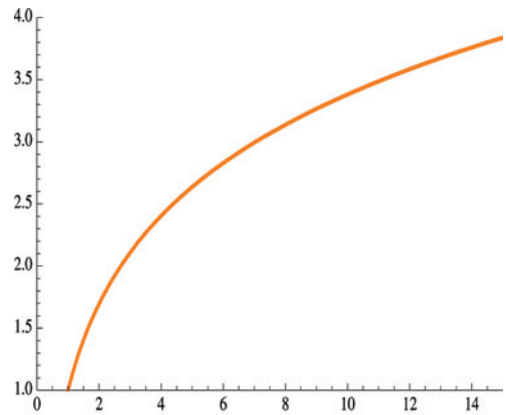


Figure 4. We solve numerically eq. (7.12) in the RD epoch, $a(t) \sim t^{1/2}$ and for the dark matter decouples from the baryonic matter (which is tightly coupled with radiation), then the evolution equation (7.12) has not the radiation pressure term and therefore the third term is independent of the k variable. Solving numerically this evolution equation for the dark matter density contrast, for all sub-horizon modes $k > k_J$, we find $\delta_{\rm DM} \sim \ln t$.

3. Sub-horizon solutions in the RD and MD universes for the photon-baryon fluid: acoustic oscillations

As mentioned above, after the modes enter the horizon it happens that the dark matter decouples from the baryon-photon fluid, and therefore they follow different evolution equations. The evolution equation for the baryon-photon fluid

$$\ddot{\delta}_{b\gamma} + 2\frac{\dot{a}}{a}\dot{\delta}_{b\gamma} + (k^2 - k_J^2)\, c_s^2\, a^{-2}\, \delta_{b\gamma} = 0\,, \tag{7.14}$$

contains the radiation pressure term, then for sub-horizon modes $k > k_J$ or $\lambda < \lambda_J$ the photon-baryon density fluid satisfies $\delta_{b\gamma} \sim \cos(c_s k\, t)$, as can be seen in figure 5. These are just the acoustic oscillations performed by the baryon-photon fluid.

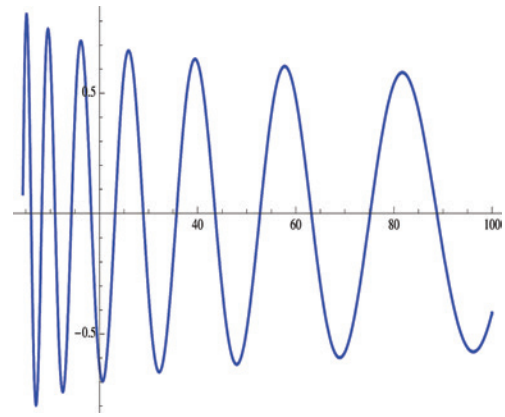


Figure 5. Solving numerically eq. (7.12) in the RD and MD epochs, $a(t) \sim t^{1/2}, t^{2/3}$ where now all modes are inside the horizon, that is $\lambda < c/H \sim \lambda_J$ (equivalently $k > k_J$) we find that density contrast for the baryon-photon fluid decouples from the dark matter density and evolves performing acoustic oscillations $\delta_{b\,\gamma} \sim \cos(c_s k\, t)$ with variable frequency because $c_s^2 \sim a^{-1}(t)$.

4. Sub-horizon solutions in the MD universe for dark matter

Now we investigate the solutions for the dark matter density fluctuations in a MD universe ($a(t) \sim t^{2/3}$). The evolution equation for this case is

$$\ddot{\delta}_{\rm DM} + 2\frac{\dot{a}}{a}\dot{\delta}_{\rm DM} - 4\pi G\, \bar{\rho}\, \delta_{\rm DM} = 0\,, \tag{7.15}$$

because there is no radiation pressure term, and $\bar{\rho} \sim a^{-3} \sim t^{-2}$. Solving numerically this equation we find that the dark matter density grows like $\delta_{\rm DM} \sim t^{2/3}$, as can be verified in figure 6.

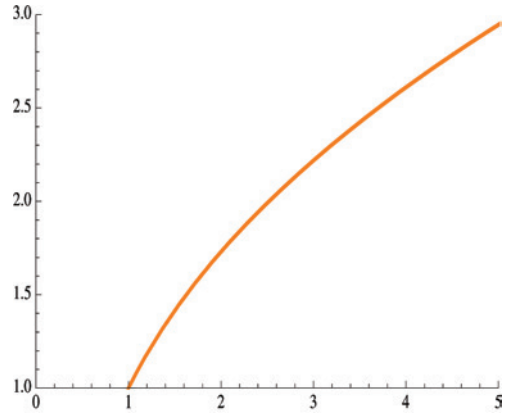


Figure 6. Consider the dark matter density perturbations in the MD universe. The evolution equation of these perturbations is eq. (7.15). Solving numerically we find that the dark matter density fluctuations grow like $\delta_{\rm DM} \sim t^{2/3}$, for sub-horizon scales.

IV. CONCLUSIONS AND FINAL REMARKS

The seeds for structure formation in the universe were generated as small fluctuations during cosmic inflation. Consequently the wavelengths modes of these perturbations soon exceed the Hubble radius, and perturbation modes are out of the horizon (as shown in figure 2). After these modes reenter the horizon it happens that the dark matter decouples from the baryon-photon fluid, and starts to form large overdensities.

As we have shown in the above section, the dark matter density perturbations grows like

$$\delta_{\rm DM} \sim \ln t \qquad \text{RD universe}$$
$$\delta_{\rm DM} \sim t^{2/3} \qquad \text{MD universe}$$

which is of fundamental importance for the structure formation in the universe. In fact, this grow of dark matter perturbations means that overdensities form large potential

wells very early, just after the perturbations reenter the horizon, and increases their rate of growth after the equipartition time, when the universe became MD. Later on, the matter-radiation fluid decouples and the large gravitational attraction lead the baryons to fall into the dark matter overdensities to form large structures.

Concomitantly, the baryon acoustic oscillations originate also when the baryon-photon fluid perturbations reenter the horizon. As we have shown, primordial matter-radiation density fluctuations do not collapse completely by gravity but instead oscillate due to radiation pressure, and spherical sound waves were driven into the surrounding medium. Then at the baryon drag epoch, soon after the matter-radiation decoupling era, the photons stream away from the baryons and the sound waves freeze out leaving the baryons in the spherical shell, with a radius that in comoving distance is around $105\,h^{-1}$ Mpc. In this scenario, the large-scale structure of the universe evolves *via* gravitational instability from a mixture superposition of these spherical shells, with galaxies forming preferentially in regions of dark matter overdensities where baryons fall, but also with a slight probability to form galaxies in shell's regions where baryons overdensities accumulate and dark matter falls. As predicted, this primordial phenomenon left imprints both in the CMB and in the large-scale distribution of galaxies that have been confirmed in recent surveys [1, 9].

ACKNOWLEDGMENTS

I would like to thank the organizers of the José Plinio Baptista School of Cosmology, in particular Davi Rodrigues and Júlio Fabris, for their support and for the opportunity to participate in this stimulating meeting. I also acknowledge the CNPq fellowship.

[1] W. Hu, arXiv:0802.3688; Lecture Notes on CMB Theory: From Nucleosynthesis to Recombination, XIX Canary Island Winter School of Astrophysics

[2] S. Dodelson, *Modern Cosmology*, Academic Press, 2003

[3] Gastão B. Lima Neto, *Astronomia Extragaláctica*, IAG–USP, 2010 www.astro.iag.usp.br/∼gastao

[4] A. Riotto, Lecture Notes, II Jayme Tiomno School of Cosmology:
http://www.cbpf.br/∼iijtsc/lecture.php
A. Riotto, arXiv:hep-ph/0210162

[5] T. Padmanabhan, *Structure formation in the universe*, Cambridge University Press, Cambridge, 1995

[6] B. Ryden, *Introduction to Cosmology*, Addison Wesley, San Francisco, 2003

[7] R. Alpher and R. Herman, *Evolution of the universe*, Nature **162** (1948) 774

[8] A. Liddle, *An Introduction to Modern Cosmology*, John Wiley & Sons, West Sussex, England, 2003

[9] D. J. Eisenstein, *Observing dark energy*, Class. Quantum Grav. **25** (2008) 114001
http://scholar.harvard.edu/deisenstein/book/baryon-acoustic-oscillations

Modified Gravity and Structure Formation

Marcos Lima[1]

[1] *Departamento de Física Matemática,*
Instituto de Física,
Universidade de São Paulo,
São Paulo SP, Brazil

(Dated: 28th January 2014)

Models of modified gravity offer an explanation of cosmic acceleration with unique observational features and theoretical consequences. We highlight some of the modified gravity signatures in structure formation which appear even for models with an expansion history that is identical to that of a dark energy in the form of a cosmological constant.

I. INTRODUCTION

The acceleration of the background expansion of the Universe has now been observed by a number of complementary cosmological probes. Within the context of General Relativity (GR), this phenomenon can be explained by adding to the energy-momentum tensor of the Einstein equations an energy component with negative pressure, called dark energy. An alternative explanation that has received attention lately is the assumption that gravity itself must be modified on cosmologically relevant scales. Models of modified gravity offer unique observational features on various scales, which may be probed in future galaxy surveys and, in conjunction with Cosmic Microwave Background (CMB) experiments and solar system tests, may shed light into the new fundamental physics required to explain cosmic acceleration.

A number of modified gravity models have been proposed in the last years. Those include a class of phenomenological modifications that change the Einstein-Hilbert action, the so-called $f(R)$ models [1–3]. Other examples include braneworld models, such as the one proposed by Dvali, Gabadadze and Porrati (DGP, [4]), models of degravitation [5], and models of massive gravity [6–8], where the graviton is allowed to have a non-zero mass. All these possibilities add extra degrees of freedom which couple to matter and propagate extra forces, generating interesting modifications on

cosmological observables. However, viable models also contain screening mechanisms that allow them to pass local gravity tests [9]. The screening causes the extra force to disappear locally, reducing the model to GR on e.g. solar system scales, where significant modifications are not expected.

Here, we highlight a few cosmologically relevant observational features of modified gravity models, as inferred from dark-matter N-body simulations [5, 10–14]. These results represent the starting point to understand the non-linear cosmological effects induced by viable models of modified gravity. These effects must be parameterized in detail in order to constrain specific models from observations in present and future galaxy surveys.

II. MODIFIED GRAVITY

A. $f(R)$ Gravity

In $f(R)$ gravity [1–3], we phenomenologically modify the Einstein-Hilbert action in the Jordan frame by adding to it a general function $f(R)$ of the Ricci scalar R:

$$S_{\rm JF} = \frac{1}{2\kappa^2}\int d^4x\sqrt{-g}\,[R + f(R)] + \int d^4x\sqrt{-g}\mathcal{L}_{\rm m}(\chi_i, g_{\mu\nu})\,, \tag{7.1}$$

where $\kappa^2 = 8\pi G$ and $\mathcal{L}_{\rm m}$ is the Lagrangian density describing the χ_i matter fields, minimally coupled to the metric $g_{\mu\nu}$. Variation of this action with respect to the metric $g_{\mu\nu}$ produces the modified Einstein equations

$$G_{\mu\nu} + f_R R_{\mu\nu} + \left(\Box f_R - \frac{f}{2}\right) g_{\mu\nu} - \nabla_\mu\nabla_\nu f_R = 8\pi G T_{\mu\nu}\,. \tag{7.2}$$

Such class of modifications to GR must be consistent with a number of current observations, including local gravity tests, the observed expansion history as well as the growth of structure induced by linear and non-linear perturbations. A model that satisfies all such conditions has been proposed by Hu & Sawicki [2]:

$$f(R) \propto -\frac{R}{AR+1}\,. \tag{7.3}$$

Working in the high-curvature regime, ($AR \gg 1$) we may expand this relation and match to an equivalent ΛCDM model to zeroth order, obtaining [10, 11]

$$f(R) = -16\pi G\rho_\Lambda - f_{R0}\frac{\bar{R}_0^2}{R}\,, \tag{7.4}$$

where $\bar{R}_0$ is the present value of the background Ricci scalar, ρ_Λ is the energy density of the equivalent ΛCDM model and $f_R = df(R)/dR$ plays the role of an extra scalar degree of freedom. Deviations from GR are parameterized by the present value of the scalar $f_{R0} = f_R(\bar{R}_0)$, whose absolute value must be small so that the expansion history does not differ significantly from that of ΛCDM. Stability conditions impose extra general constraints on the $f(R)$ function and its parameters. On sub-horizon scales and under the quasi-static limit, perturbations in the scalar f_R evolve according to a modified non-linear Poisson equation given by [10, 11]

$$\nabla^2\delta f_R = \frac{1}{3}\left[\delta R(f_R) - 8\pi G\delta\rho_{\rm m}\right]\,, \tag{7.5}$$

where δR represents perturbations on the Ricci scalar and $\delta\rho_{\rm m}$ are perturbations on the matter field. An important distance scale in $f(R)$ models is the Compton wavelength λ_C, which for the model above is given by [10, 11]

$$\lambda_C^2 \equiv 3\frac{df_R}{dR} \approx \sqrt{\frac{6|f_{R0}|\bar{R}_0^2}{R^3}}\,, \tag{7.6}$$

and represents the range of the extra force induced by the scalar f_R. Even though the present value of λ_C can be significant on cosmological scales, it becomes small at early times (not altering the physics of the early Universe) as well as in regions of high curvature R. Such behavior, known as the *chameleon mechanism* [15], allows these models to be consistent with local gravity tests. For models that pass these tests and have a similar expansion history to GR, there still remains to be evaluated their effects on linear and non-linear perturbations. The latter can be studied via N-body simulations [10–13].

Finally, by performing a conformal transformation $\tilde{g}_{\mu\nu} = e^{\beta\kappa\phi}g_{\mu\nu}$ with $e^{\beta\kappa\phi} = 1 + f_R$, we obtain the Einstein-frame action, with a scalar coupling to matter [3]

$$\begin{aligned} S_{\rm EF} =& \frac{1}{2\kappa^2}\int d^4x\sqrt{-\tilde{g}}\tilde{R} + \int d^4x\sqrt{-\tilde{g}}\left[-\frac{\tilde{g}^{\mu\nu}}{2}\nabla_\mu\phi\nabla_\nu\phi - V(\phi)\right] + \\ &+ \int d^4x\sqrt{-\tilde{g}}\; e^{-2\beta\kappa\phi}\mathcal{L}_{\rm m}(\chi_i, e^{-\beta\kappa\phi}\tilde{g}_{\mu\nu})\,, \end{aligned} \tag{7.7}$$

where the potential

$$V(\phi) = \frac{1}{2\kappa^2}\frac{Rf_R - f}{(1+f_R)^2}\,. \tag{7.8}$$

Therefore, $f(R)$ models are classically equivalent to a particular class of scalar-tensor models. Since the matter fields are now also coupled to ϕ, it is similar to models of dark-matter and dark-energy interactions [16], or mass-varying neutrinos via coupling to dark-energy [17], but with the interaction extending to all matter fields.

B. Galileons, Extra Dimensions and Massive Gravity

Another general class of modifications to GR that add extra scalar degrees of freedom has been considered lately. These models typically attempt to associate the modification to a concrete physical mechanism. As a result, these models generally have more complicated equations of motion, compared to e.g. $f(R)$ models, and exhibit a distinct screening mechanism in order to satisfy solar system tests.

In this case, the extra scalar φ is called *galileon*, and satisfy a general evolution equation given by

$$\nabla^2\varphi + \frac{r_c^2}{3}\left[(\nabla^2\varphi - (\nabla_i\nabla_j\varphi)(\nabla^i\nabla^j\varphi))\right] + ... = \frac{8\pi G}{3}T\,, \tag{7.9}$$

where the parameter r_c controls deviations from GR, and the ellipsis represents further complications that may appear in specific models. Examples in this class include braneworld models, such as the DGP model [4] with an extra spatial dimension, and theories of massive gravity [6–8] in the so-called decoupling limit. In the case of DGP, r_c represents the scale of transition from 4 to 5-dimensional gravity, whereas in massive gravity r_c is related to the graviton mass.

In this class of models, the non-linearities induced by the gradients in the φ field cause the modifications to GR to be suppressed on scales smaller than a characteristic scale r_*, known as the Vainshtein radius, via the so-called *Vainshtein mechanism* [18].

III. OBSERVABLE SIGNATURES FROM MODIFIED GRAVITY SIMULATIONS

We now consider signatures of modified gravity that can be constrained by observations that probe the linear and specially the non-linear structure of these models. These include the galaxy power spectrum in real and redshift space, as well as the abundance and clustering of galaxy clusters. The results were extracted by comparing dark-matter N-body simulations of the modified gravity models to ΛCDM simulations with the same initial conditions and same expansion history.

A. Real Space Power Spectrum

Since the power spectrum is the fundamental quantity that characterizes the clustering properties of galaxies, we start by showing the results on the dark matter power spectrum as obtained by simulations. From the effects on the power spectrum we can roughly trace the effects on other observables presented later.

In Fig. 1 we show the deviations (relative to ΛCDM) in the real-space dark matter power spectrum at $z = 0$. Results are shown for $f(R)$ models [12], for different values of $|f_{R0}|$ (left) and for galileon models [14] for different values of r_c (right). Deviations become specially important once $|f_{R0}|$ becomes of the order of the typical gravitational potentials, i.e. 10^{-5}. Notice, however, that for this value of $|f_{R0}|$, the expansion history is completely indistinguishable from ΛCDM [2], so measurements of the growth of structure provide an interesting test for the model. Similar effects are seen for galileon models, in which case effects increase as r_c decreases. The chameleon and Vainshtein mechanisms are both important to quantify the effects and need to be well characterized for precision cosmology, though they are likely partially degenerate with non-linear baryonic effects. One should be able to measure these deviations e.g. in the lensing power spectrum with future observations.

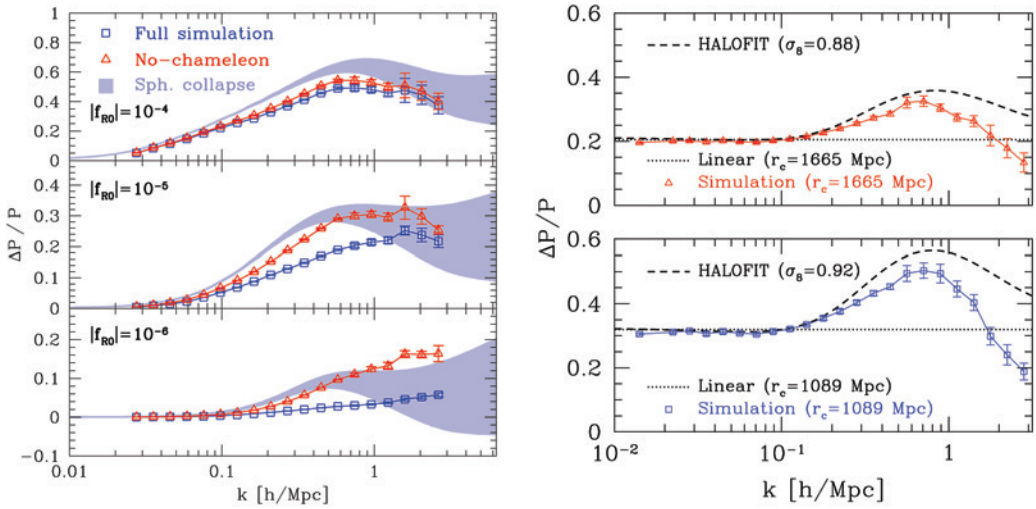


Figure 1. (*Left*): Deviations in the real-space power spectrum at $z = 0$ relative to ΛCDM for $f(R)$ models (from [12]). The effect of the chameleon mechanism is important for lower values of $|f_{R0}|$ and is not trivial to model it from ΛCDM prescriptions. (*Right*): Same results for galileon models (from [14]), displaying now the effect of the Vainshtein screening.

B. Redshift-Space Distortions

Another interesting effect of modified gravity is on redshift-space distortions induced by the peculiar velocities of galaxies. In fact, the velocity field is typically more affected than the density field in modified gravity models, so we expect redshift distortions to be a relevant probe of modified gravity.

In Fig. 2 it is shown a comparison of the redshift-space power spectrum and correlation function to their real-space counterparts in GR and in modified gravity models [13, 14] . It is clear that a careful characterization of these effects can help not only distinguish modified gravity from GR, but also distinguish modified gravity models among themselves. For instance, in both $f(R)$ and galileon modified gravity, there is a suppression of the ratio of the redshift-space and real-space power spectrum on small scales. However, since the extra scalar in $f(R)$ gravity is limited by its Compton scale, it does not generate differences on large scales. In contrast, for galileon models, long-range deviations from GR remain present on large scales. Spectroscopic measurements of galaxy redshifts will therefore play an important role for constraining

modified gravity models.

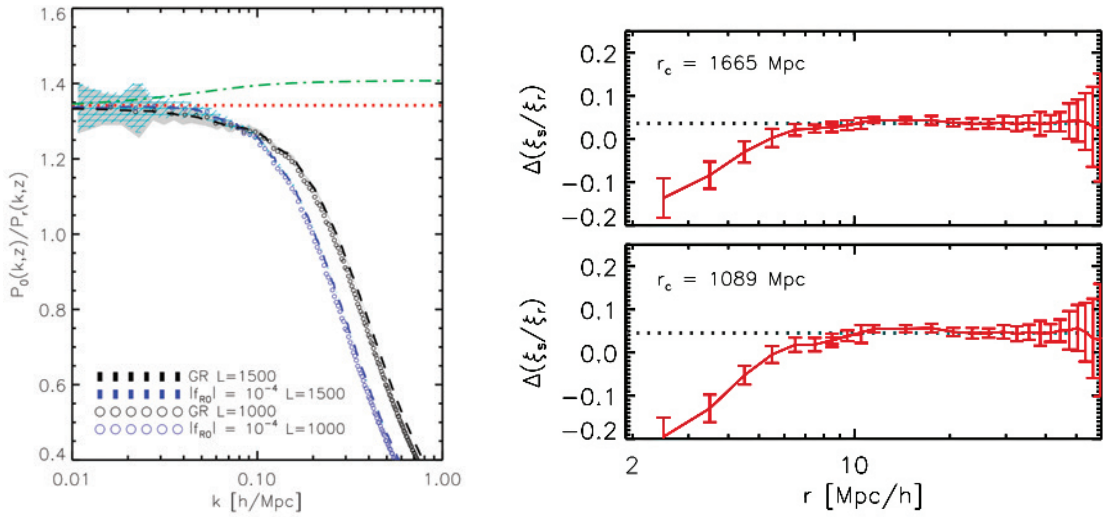


Figure 2. (*Left*): Ratio of the redshift-space power spectrum monopole to the real-space power spectrum at $z = 0$ for GR and $f(R)$ models with $|f_{R0}| = 10^{-4}$ (from [13]). The dotted line is the linear prediction for GR and the dot-dashed line is the linear prediction for $f(R)$. (*Right*): The relative difference in the ratio of the correlation function in redshift space to that in real space, in the galileon model with $r_c = 1665$ Mpc ($r_c = 1089$ Mpc) compared to ΛCDM is shown on the top (bottom) panel. The dotted line shows the Kayser linear prediction (from [14]).

C. Halo Mass-function

It has been long known that the abundance of galaxy clusters are an important probe of cosmology [19]. Being exponentially sensitive to the density fluctuations (e.g. [20]), clusters are naturally powerful probes of dark energy and modified gravity.

In Fig. 3 we show the deviations (relative to ΛCDM) in the dark matter halos detected at $z = 0$ for $f(R)$ [12] and galileon [14] models. In the case of $f(R)$ models, deviations are larger for massive halos, but the difference is suppressed by the chameleon mechanism for low values of $|f_{R0}|$. As a result the mass-function universality breaks down, and fits calibrated from ΛCDM simulations using the linear

power spectrum are not able to describe the modified gravity mass-function on all masses. For galileon models, we also notice a large deviation at high masses but, at least for the parameter values shown here, fits calibrated in ΛCDM simulations [20] can describe the deviations reasonably well.

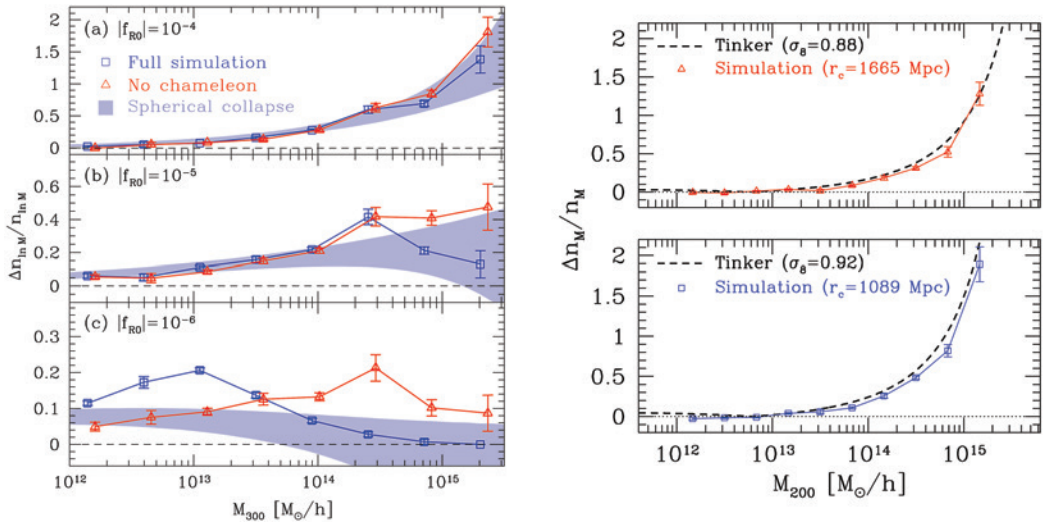


Figure 3. (*Left*): Relative deviations in the halo mass-function at $z = 0$ relative to ΛCDM for $f(R)$ models (from [12]). The effect of the chameleon mechanism is important for lower values of $|f_{R0}|$. Again we notice that it is not trivial to model this effect from ΛCDM prescriptions. (*Right*): Same results for galileon models (from [14]). Also shown is the deviation of the Tinker [20] mass-function in ΛCDM with higher values of σ_8.

D. Linear Halo Bias

Finally, in addition to the abundance of clusters, their intrinsic clustering has further information that can be combined to constrain modified gravity models.

In Fig. 4 we show the deviations (relative to ΛCDM) in the linear bias of dark matter halos detected at $z = 0$ for $f(R)$ and galileon models. The linear bias is computed as the ratio of the halo-matter cross-spectrum and the matter power spectrum on large scales. This measurement is much noisier due to the shot noise induced by the much lower number of halos relative to dark matter particles. In both modified gravity cases, the linear bias decreases relative to GR, as a result of the increase in the number of

halos, which become less rare objects.

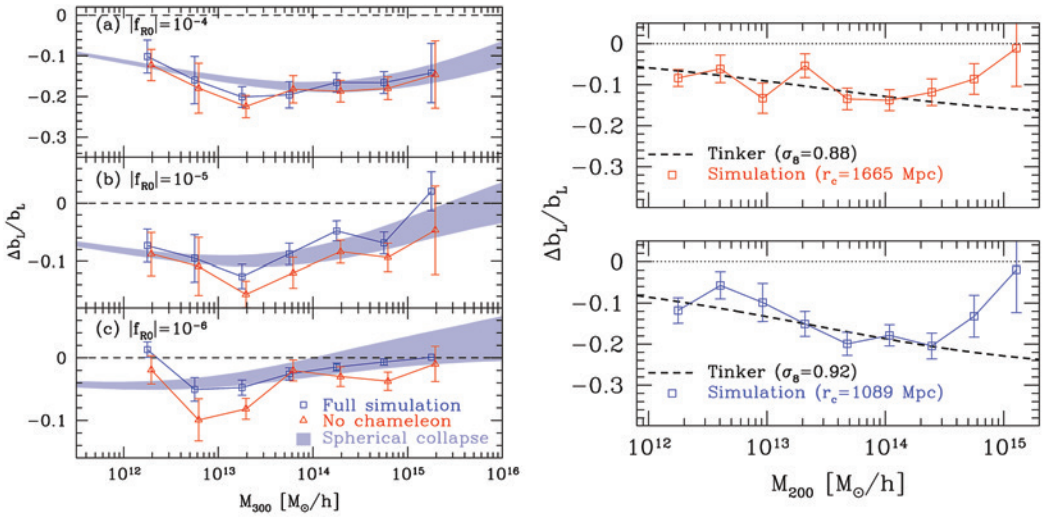


Figure 4. Deviations in the linear halo bias at $z = 0$ for the same cases shown in Fig. 3. The results for the halo bias are noisier than those for the mass function, but similarly display the degeneracy between $f(R)$/galileon modified gravity and ΛCDM with larger σ_8 at single redshifts.

E. Discussion

We have described $f(R)$ and galileon models of modified gravity and shown results from N-body simulations that indicate how the non-linear effects of these models can be measured by future observations. Results in the power spectrum, redshift distortions and halo properties show that it should be possible to distinguish modified gravity from GR with cosmological observations of future galaxy surveys.

If future cosmological constraints point to modified gravity parameters that render these models indistinguishable from GR, and local gravity tests simultaneously continue to show no indication of modifications to GR, these models obviously shall become less attracting. Until that happens, they offer an alternative to explain cosmic acceleration without invoking extra fields in the standard model solely for that purpose.

Nonetheless, these models themselves introduce extra degrees of freedom which may be interpreted as dynamical dark energy fields. So it is not clear if it will ever be

possible to distinguish modified gravity and dark energy from classical cosmological observations alone. Hopefully along the way, motivations based on micro-physics and first principles may shed light on these dark issues and tell us whether dark energy is a new *aether* that we must get rid of, or if it must be somehow accommodated within the standard model.

ACKNOWLEDGMENTS

I thank the organizers and the participants of the JPBCosmo for a pleasant meeting. This work is partially supported by FAPESP and CNPq.

[1] S. M. Carroll, V. Duvvuri, M. Trodden, and M. S. Turner, Phys.Rev. **D70**, 043528 (2004), arXiv:astro-ph/0306438 [astro-ph].

[2] W. Hu and I. Sawicki, Phys.Rev. **D76**, 064004 (2007), arXiv:0705.1158 [astro-ph].

[3] T. P. Sotiriou and V. Faraoni, Rev.Mod.Phys. **82**, 451 (2010), arXiv:0805.1726 [gr-qc].

[4] G. Dvali, G. Gabadadze, and M. Porrati, Phys.Lett. **B485**, 208 (2000), arXiv:hep-th/0005016 [hep-th].

[5] J. Khoury and M. Wyman, Phys.Rev. **D80**, 064023 (2009), arXiv:0903.1292 [astro-ph.CO].

[6] G. Gabadadze, Phys.Lett. **B681**, 89 (2009), arXiv:0908.1112 [hep-th].

[7] C. de Rham and L. Heisenberg, Phys.Rev. **D84**, 043503 (2011), arXiv:1106.3312 [hep-th].

[8] P. Gratia, W. Hu, and M. Wyman, Phys.Rev. **D86**, 061504 (2012), arXiv:1205.4241 [hep-th].

[9] C. M. Will, Living Reviews in Relativity **9** (2006).

[10] H. Oyaizu, Phys.Rev. **D78**, 123523 (2008), arXiv:0807.2449 [astro-ph].

[11] H. Oyaizu, M. Lima, and W. Hu, Phys.Rev. **D78**, 123524 (2008), arXiv:0807.2462 [astro-ph].

[12] F. Schmidt, M. Lima, H. Oyaizu, and W. Hu, Phys.Rev. **D79**, 083518 (2009), arXiv:0812.0545 [astro-ph].

[13] E. Jennings, C. M. Baugh, B. Li, G.-B. Zhao, and K. Koyama, Mon. Not. R. Astron. Soc. **425**, 2128 (2012), arXiv:1205.2698 [astro-ph.CO].

[14] M. Wyman, E. Jennings, and M. Lima, (2013), arXiv:1303.6630 [astro-ph.CO].

[15] J. Khoury and A. Weltman, Phys.Rev.Lett. **93**, 171104 (2004), arXiv:astro-ph/0309300 [astro-ph].

[16] J.-H. He, B. Wang, and E. Abdalla, Phys.Rev. **D84**, 123526 (2011), arXiv:1109.1730 [gr-qc].

[17] A. Brookfield, C. van de Bruck, D. Mota, and D. Tocchini-Valentini, Phys.Rev.Lett. **96**, 061301 (2006), arXiv:astro-ph/0503349 [astro-ph].

[18] A. Vainshtein, Phys.Lett. **B39**, 393 (1972).

[19] L.-M. Wang and P. J. Steinhardt, Astrophys.J. **508**, 483 (1998), arXiv:astro-ph/9804015 [astro-ph].

[20] J. L. Tinker, A. V. Kravtsov, A. Klypin, K. Abazajian, M. S. Warren, *et al.*, Astrophys.J. **688**, 709 (2008), arXiv:0803.2706 [astro-ph].

Tests and extensions of the Cosmological Principle

Thiago Pereira[1]

[1] *Departamento de Física – Universidade Estadual de Londrina*
Rodovia Celso Garcia Cid, Km 380,
86051-980, Londrina – Brasil

This article is an extended version of a talk given at *José Plinio Baptista School of Cosmology*, held at the city of Anchieta, in the state of Espirito Santo, Brazil, during October 2012. It briefly highlights a few recent attempts to explore theoretical extensions and observational tests of the Cosmological and Copernican Principle, with possible consequences to the physics of dark energy and CMB.

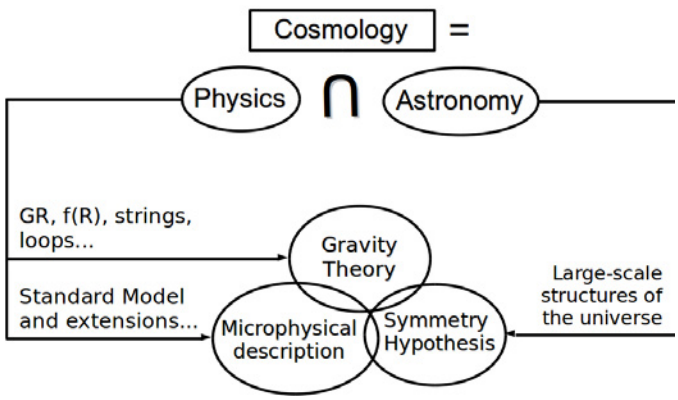
I. INTRODUCTION

One of the oldest questions of mankind is related to our physical position in the universe. Since the early Mesopotamian cosmology of a flat and round Earth standing in an infinite cosmic ocean, following Aristotle's conception of a spatially finite and ever existing geocentric cosmos and beyond, our conceptions of the geographic configuration of the universe has faced a fascinating history of scientific development. Along this road, a very important moment was Copernicus's realization that the Earth cannot be in the center of the solar system. This radical change of perception of our conceptual ideas, putting aside its religious and philosophical consequences, has introduced a sort of geographic "mediocrity principle" in cosmology in which, given the absence of evidence for exotic speculations, we should assume that our configuration in the universe is the simplest possible. The *Copernican Principle* has since then played a central role in physical cosmology. Nowadays, backed up by a plethora of astronomical observations, modern cosmologists have further generalized this idea to take into consideration that not only our position is typical, but so is our overall orientation in the cosmos. This principle is now known as the *Cosmological Principle*, and it's a central idea of modern cosmological models. In fact, it is so central to the hot big bang model that its observational plausibility should be exhaustively tested, specially when observations don't seem to have an obvious interpretation in the standard model.

This is in fact the situation with our current understanding of the cosmos. Observations dating to nearly a decade ago have suggested that about 70% of the cosmic energetic budget is unknown to us, meaning that the 70% of the universe is made up of a sort of energy not composed of known elementary particles. Moreover, this mysterious component seem to be causing galaxies to recede apart at an increasing rate, just like a repulsive force between them would do. This problem has been named the *dark energy* problem of the universe, and is perhaps one of the most important problems of modern physics (see [1] for a recent review). We will show in this brief article that, among recent attempts to model the dark energy, either by modifying gravity or by postulating the existence of a new exotic particle, one real possibility lies, ironically enough, in violating the Cosmological and/or Copernican principles.

II. STANDARD MODEL INGREDIENTS

To understand how the standard model of cosmology is made up, we should first understand what is really meant by physical cosmology. In this article we will follow the perspective that cosmology is not a fundamental physical theory, like quantum mechanics or thermodynamics, but rather an arena in which modern physical theories can be tested and applied[1]. Briefly, modern cosmology can be seen as the intersection of physics and astronomy, as illustrated by the chart bellow.



[1] The reason the universe is now seen as a "natural laboratory" for testing physical theories should be evident, given the eminent material limitations of modern high energy physics experiments (the best known example being Cern's Large Hadron Collider).

In this perspective, physics provide fundamental theories for the microphysical description of matter and the behavior of gravity at large scales, while astronomy is responsible for providing observational support (if any) of these theories and, most importantly, of the symmetry hypothesis necessary to apply these very theories to the universe.

The problem with this perspective is its large degeneracy, since there are many combinations of gravitational and particle theories which could be used to explain the observed universe. So we can ask what is the simplest (i.e. with minimum number of free parameters) combination of theories compatible with current data. As of today, the simplest possible combination is known as the ΛCDM model (Lambda Cold Dark Matter), and is constructed out of Einstein's general relativity, plus the standard model of particle physics, plus the hypothesis that the universe is geometrically homogeneous and isotropic[2]:

$$\Lambda\text{CDM} = \text{Gen. Relativity} + \text{Standard Model} + \text{Cosmological Principle}\,. \tag{8.1}$$

When taken together, different astrophysical observations seem to converge, with an excellent agreement, to a model in which the universe figures as a 13.7 billion years old, spatially flat, homogeneous and isotropic structure, composed of nearly 70% of dark energy (i.e. Einstein's cosmological constant Λ), 25% of cold dark matter and only 5% of standard (i.e. baryonic) matter [3].

[2] The universe could be topologically anisotropic and still compatible with the Cosmological Principle. This happens because while general relativity constrains the universe geometry, but not its topology. See [2] for a discussion of this point.

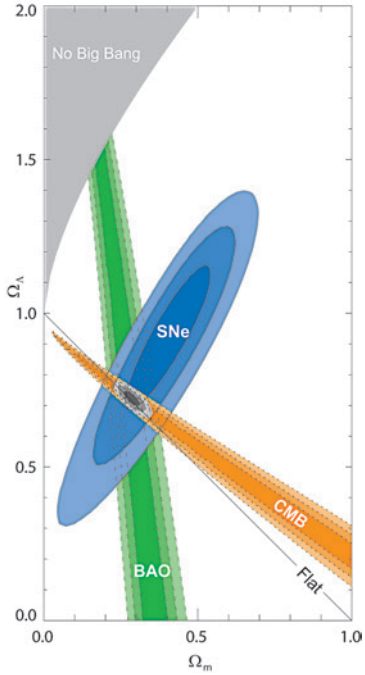


Figure 1. Joint statistical constraint on total matter (Ω_m) and dark energy (Ω_Λ) parameters using Supernovae (SNe), Cosmic Microwave Background radiation (CMB) and Baryonic Acoustic Oscillations (BAO). From [4].

Albeit extremely successful the ΛCDM model has, not surprisingly, many drawbacks. One of them is the already mentioned dark energy problem – the fact that we don't know why the universe seems to be expanding at an accelerated rate, nor what is the nature of 70% of its mass. There are many other important problems that we are not going to present here, namely the dark matter problem (a still mysterious substance required to form gravitationally bound structures and which comprises nearly 25% of the cosmic mass), the real mechanisms of inflation, baryogenesis and so on (see [5] for a thorough perspective on the present status of cosmology). The important point is that any extension of the ΛCDM model necessary to solve the dark energy problems necessarily requires the modification of one of the three key ingredients in (8.1). We could argue, for example, that the accelerated expansion of the universe reflects a failure of general relativity (GR) at cosmological scales. This is in fact a reasonable hypothesis, given that most observational tests of GR refer to small (astronomical) scales. Another very popular route to explore the nature of dark energy is to

postulate the existence of a new matter component that could act similarly as a repulsive fluid. Given the difficulty to construct particle physics model from first principles that could mimic dark energy, one very popular approach is to phenomenologically postulate the existence of a new field and use observations to constrain its fine details. The simplest field we can think of is a canonical scalar field, and this is by far the most studied phenomenological model of dark energy [6]. The third approach, which is also the least explored of the three, is the possibility that the expansion of the universe is a result of a failure of the Cosmological Principle. If we suppose for example that we live in the center of a spherical void of the universe (that is, a region with matter density slightly less than the cosmic average) than the gravitational attraction of the outer region could give us the impression that our patch of the universe is expanding. Although placing ourselves at the center of the (observed) universe does not seem to be a good philosophical approach, it is also true that every hypothesis of the model has to be confronted against empirical data. In this vein, the hypothesis of the Cosmological Principle are the most critical ones: while it's technically possible to look around and ask whether our universe is isotropic, we are still very far from conducting experiments in other points of the universe so as to verify its homogeneity. This explains why the Cosmological Principle, although a fundamental pillar of the standard model, is also one which is very hard to deal with. In the next pages we are going to give a glance of some recent theoretical attempts to test the symmetries principles of the universe, both from the phenomenological and fundamental point of view.

III. SYMMETRY PRINCIPLES

In order to discuss tests of symmetry principles, it's important to classify them into distinct classes. The ΛCDM model assumes that our universe is, on average, homogeneous and isotropic. In mathematical language this means that the universe is invariant under rotations (isotropic) and spatial translations (homogeneous). Now, since these are two different concepts, it's possible to construct models in which none, one or both of these symmetries are violated. We are therefore going to make the following distinction:

- The Copernican Principle: our position in the universe is not special.
- The Cosmological Principle: our position and overall orientation in the universe is not special.

Note that the Copernican Principle allows the universe to be anisotropic, as long as all observes see the same pattern of anisotropies. The Cosmological Principle on the other hand is more restrictive, and requires every observer to see the same features in all directions. One example of a geometry which satisfies the Cosmological Principle is the Friedmann-Lemaître-Robertson-Walker (FLRW) metric. The Copernican Principle, on the other hand, could be described by any anisotropic geometry of the Bianchi family (see [7] for a mathematical description). The Bianchi I metric is the simplest example where the universe expands in different directions with different speeds, and is the simplest extension of the FLRW metric we can think of.

Before we move on we should mention a crucial point here: the above symmetries are expected to hold only on average. The universe is certainly anisotropic and inhomogeneous at planetary and even astronomical scales. In fact, current data suggest that the universe is (approximately) homogeneous and isotropic only at scales above 100 Mpc. Moreover, current cosmological data is unable to probe the geometry of the universe at scales above 1Gpc. For a comparison, note that the radius of the observable universe, in the ΛCDM model, is of order of 10 Gpc. This leaves us with a good fraction of the observable universe yet to be explored. In the next sections we are going to discuss some ways of testing both of the above principles.

A. Tests of Copernican Principle

The very low temperature fluctuations ($\sim 0.001\%$) of the Cosmic Microwave Background radiation (CMB) sounds like a very restrictive constraint on the amount of anisotropy of the universe. If taken at its face value, we conclude that the only way to violate the Copernican Principle and yet obey this observational constraint is by placing us at the center of a large spherical void. The need for an underdense region comes from the known fact (but still debatable) that spherically symmetric void

models of the universe, described for example by a Lemaître-Tolman-Bondi metric (LTB), seems to be in good agreement with the distance-redshift supernovae data (see [8] and references therein), thus possibly challenging quintessence or modified gravity models as explanations for dark energy.

The question then is: how to test this hypothesis? In order to answer this question, Caldwell and Stebbins have developed a test based on deviations of the black body spectrum of CMB [9]. As is well known, CMB has one of the most perfect Planck spectral shape ever measured [10]. Based on this consideration, the authors argued that a large anisotropic distribution of electrons in the reionized universe could produce spectral distortions in the CMB photons through inhomogeneous Thompson scattering. To see how this can happen, imagine that we live in the center of a spherical empty region (call it region A) encompassed by a homogeneous spherical ring of ionized gas (region B), which is itself adjacent to a neutral spherical zone of constant gravitational potential and infinite extent (region C). Regions B and C are supposed to have the same gravitational potential Φ, so that Φ can be seem as a step function which is zero in A and constant outside it. Superposed to this geometry are CMB photons which define the CMB rest frame. Because of the spherical geometry, free electrons will move radially outward with respect to this frame. Photons coming from regions B and C which scatter inside the void before reaching us will experience different Doppler shifts, but no temperature shift from gravitational potential (Sachs-Wolfe effect), since they probe the same value of Φ. On the other hand, photons from regions C and A which scatter (say, in region B) before reaching us will probe different values of Φ, therefore experiencing a different Sachs-Wolfe effect. This difference in the temperature of scattered photons produce a spectral distortion u of the CMB given by [11]

$$u(\hat{n}) \propto \int_0^\infty \dot{\tau} dz \int d^2\hat{n} \left[1 + \left(\hat{n} \cdot \hat{n}'\right)^2\right] \left[\Delta T(\hat{n}, \hat{n}, z) - \Delta T(\hat{n}', \hat{n}, z)\right]^2$$

where τ is the optical depth due to Thompson scattering, z is the redshift, and $\hat{n}'$ and $\hat{n}$ are the directions of photons before scattering and at detection, respectively.

In order to model the void, Caldwell and Stebbins have used a two-parameter gravitational potential given by $\Phi(r) = \Phi_0 \left[1 - (r/r_v)^2\right] \Theta(r_v - r)$, where r_v is the comoving radius of the void and Φ_0 its amplitude. These parameters can in turn

be related to directly measurable quantities like the total matter density Ω_0 and the present Hubble rate H_0[3]. By using current bounds on the spectral distortion u, they were able to put severe constraints on these parameters. Their main conclusion is that this test rules out large and deep void models at 95% confidence level; the larger the void, the shallower it has to be in order to pass their tests. When combined with Supernovae Ia (SN-Ia) data compiled in [12, 13], the analysis also rule out most of the void models which can explain the accelerated expansion. Although one can argue that this analysis is restricted to one single void model, the simplicity of the model is likely to retain most of the main features of a spherically inhomogeneous universe. This result poses a considerable limitation to void models as dark energy alternatives.

More recently, another serious challenge to LTB models was proposed by Zhang and Stebbins [14]. Their approach is also based on the fact that a radially inhomogeneous distribution of matter will in general produce a net velocity between the average matter frame and the CMB rest frame. The physical mechanism employed by the authors to test this hypothesis is the kine ct Sunyaev-Zel'dovich effect (kSZ). The kSZ effect arises when the relative motion between high energy electrons and CMB photons induces a shift in the radiation temperature distribution through inverse Compton scattering. Typically, for a cloud of free electrons with mean velocity $\vec{v}_{\rm bulk}$ and mean electron number density δ_e, the temperature fluctuation will be given by

$$\Delta T_{\rm bulk}(\hat{n}) = T_{\rm CMB} \int [1 + \delta_e(\hat{n}, z)] \, \frac{\vec{v}_{\rm bulk}(\hat{n}, z) \cdot \hat{n}}{c} d\tau \, . \tag{8.2}$$

Thus, the kSZ effect is a powerful tool that uses CMB anisotropies as a tracer of the local matter distribution in the universe. In a perfectly symmetric universe, $\vec{v}_{\rm bulk}$ will average to zero at distances above the homogeneity scale. However, if we happen to live in the center of a large spherical region, gradients in the gravitational potential will induce a radial velocity component $\vec{v}_{\rm void}$ in free electrons which would not average to zero at these scales. This velocity will lead to a temperature fluctuation $\Delta T_{\rm bulk}(\hat{n})$ identical to Eq. (8.2) and to an effective temperature fluctuation given by

$$\Delta T(\hat{n}) = \Delta T_{\rm bulk}(\hat{n}) + \Delta T_{\rm void}(\hat{n}) \, .$$

[3] The reader can check the original reference [9] for more details. Here we are mainly interested in qualitative results.

In order to proceed, we should ask how big $\vec{v}_{\rm void}$ has to be in order to explain the observed redshift-distance relation of the ΛCDM model. The kSZ effect in a LTB model was studied in [15], where it was found that typical values for $\vec{v}_{\rm void}$ are unacceptably large given current observational constraints. As an example, for a typical void of radius $\gtrsim 1h^{-1}$Gpc, the required velocity is $|\vec{v}_{\rm void}| \gtrsim 10^4\,$km/s. Using this value as a crude estimate, Zhang and Stebbins found $\Delta T^2_{\rm void} \gtrsim 80\,\mu{\rm K}^2$ at a multipolar scale $\ell = 3000$. This should be compared with the much lower observational bound on the total temperature fluctuation, $\Delta T^2 < 6.5\mu{\rm K}^2$, at a confidence level of 95%, found by the South-Pole telescope team [16].

The two tests outlined above are in excellent agreement with the Copernican Principle, at least for the simple class of models considered. While future data is expected to improve observational constraints on these tests, or else to verify the dark energy hypothesis, theoretical developments can help us with the formulation of more general scenarios. One interesting proposal in this direction appeared in [17], where a completely geometrical and model-independent test of the Copernican Principle was constructed. The test relies on the equivalence between luminosity distance $d_L(z)$ and the Hubble rate $H(z)$ in a purely FLRW universe:

$$d_L(z) = \frac{(1+z)}{H_0\sqrt{|\Omega_k|}}\sin\left(\sqrt{|\Omega_k|}\int_0^z du\frac{H_0}{H(u)}\right)\,.$$

where Ω_k is the curvature parameter *evaluated today* and $H_0 = H(z=0)$. From this expression one can invert Ω_k to show that

$$\Omega_k = \frac{\left[H(z)D(z)'\right]^2 - 1}{\left[H_0 D(z)\right]^2}\,,$$

where $D(z) = d_L(z)/(1+z)$. Given that Ω_k is constant, the derivative of this expression will be zero provided that the "Copernican function"

$$\mathrm{Cop}(z) \equiv 1 + H^2\left[DD'' - D'^2\right] + HH'DD'$$

is zero. Note that z, $H(z)$ and $d_L(z)$ are directly measurable quantities. Therefore, if $\mathrm{Cop}(z)$ is measured to be non-zero within error bars, the underlying geometry is not FLRW and the Copernican Principle is violated. Unfortunately, this is not an "if

and only if" statement. In other words, a null measurement measurement of $\mathrm{Cop}(z)$ does not imply that the FLRW geometry holds. Another drawback of the proposal is the observational difficulty to measure both $H(z)$ and $d_L(z)$ at different redshifts and with different data sets, so as to assure that the measurement is not biased in any way. However, the test has the desirable feature of being completely model independent; it does not rely on any (metric) gravitational theory nor on any model for the cosmic matter content.

B. Tests of Cosmological Principle

In the last section we saw a few examples of how to test the hypothesis that our position in the universe is not special. We will now discuss the possibility that our position is typical (the Copernican Principle holds) but our overall orientation in the universe is not, thus violating the (stronger) Cosmological Principle. The main difficulty with tests of global anisotropy is to evade the primordial constraints imposed by CMB which, as we know, implies that the universe was isotropic to 1 part in 10^5 at $z \sim 1100$. In the simplest anisotropic scenario the metric of the universe is given by

$$\mathrm{d}s^2 = -\mathrm{d}t^2 + a^2(t)\mathrm{d}x^2 + b^2(t)\mathrm{d}y^2 + c^2(t)\mathrm{d}z^2,$$

where a, b and c are different functions of the cosmic time t. One measure of anisotropy is given by the spatial shear, whose amplitude scale as

$$\sigma = \frac{\sigma_0}{S^3}\,,$$

where $S = \sqrt{abc}$ is the mean scale factor [18]. It's then not hard to see that an observational bound $\sigma \lesssim 10^{-5}$ from CMB would imply an extremely high and fine tuned value for σ_0 during inflation [19]. However, it is still plausible that, by some unknown mechanism, the universe developed late time anisotropies which could be detected through large scale structure observations [20].

In face of the restrictive bounds of CMB one can ask: is the isotropy of the background radiation a proof that our universe is in accordance with the Cosmological Principle? Or, to put it differently, does the assumption of homogeneity plus the

isotropy of CMB imply that our metric is the FLRW one? The answer to this question is, interestingly, in the negative [21]. Indeed, a well known theorem ensures that the CMB will be isotropic if and only if the universe is described by a conformally stationary metric (see for example [22]). Of course, if fundamental observers happen to measure only dust as the matter content, then the spacetime is necessarily FLRW (this is known as the EGS theorem [23]). But it is always possible to postulate the existence of some anisotropic source of matter which would feed the spatial anisotropy while preserving the isotropy of the CMB.

Consider for example the conformally static spacetime metric of the form

$$\mathrm{d}s^2 = a^2\left(-\mathrm{d}\eta^2 + \mathrm{d}r^2 + \mathrm{d}\theta^2 + f(\theta)^2\mathrm{d}\phi^2\right) . \tag{8.3}$$

For $f(\theta) = \sinh\theta$, this represents the Kantowski-Sachs geometry, which is a spacetime whose spatial section is of the form $R \times S^2$. For $f(\theta) = \sin\theta$ this is the Bianchi III metric, which is a hyperbolic metric with spatial section given by $R \times H^2$ (H^2 being the two-dimensional pseudo-sphere). Eq. (8.3) is a solution of Einstein's equations with an imperfect matter content of the form

$$T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + p g_{\mu\nu} + \pi_{\mu\nu}$$

where, as usual, ρ, p and $\pi_{\mu\nu}$ are the homogeneous energy density, pressure and anisotropic stress measured by a fundamental observer with 4-velocity u_μ. Besides being globally anisotropic[4] and admitting an isotropic CMB, these two models have another peculiarity. To see that, let us write down Einstein field equations explicitly in the coordinate system of Eq. (8.3). We have (using $8\pi G = c = 1$):

$$\begin{aligned} 3\mathcal{H}^2 &= a^2\rho - 2k \\ \mathcal{H}^2 + 2\mathcal{H}' &= -a^2 p - \pi_\perp \\ \mathcal{H}^2 + 2\mathcal{H}' &= -a^2 p - \pi_\perp - k \end{aligned}$$

where $\mathcal{H} = a'/a$, $' = d/d\eta$ and the parameter k represents the Kantowski-Sachs ($k = +1$) and Bianchi III ($k = -1$) cases. Note that $\pi_\parallel$ and $\pi_\perp$, which are the

[4] Actually, these models have a residual rotational symmetry around one axis, so that they are also known as locally rotationally symmetric (LRS) spacetimes.

only two non-zero degrees of freedom of the stress tensor in this coordinate basis, are completely free functions of η. As it turns out, if we *choose* these two functions such that

$$\pi_{\perp} = -2\pi_{\parallel} = 2k/3a^2$$

and redefine $k \to 3k/2$, the resulting equations become

$$3\mathcal{H}^2 = a^2\rho - 3k$$
$$\mathcal{H}^2 + 2\mathcal{H}' = -a^2 p - k$$

which are formally identical to the Friedmann equations with curvature [24]. This is an interesting result, and it was first shown to happen whenever $\pi_{\mu\nu}$ is proportional to the electric part of the Weyl tensor [25].

The above discussion is an explicit example, albeit lacking some further physical intuition, of a model which violates the Cosmological Principle, despite the fact that we measure an incredibly isotropic CMB. Moreover, the background dynamics is formally identical to the one found in a FLRW universe with a perfect fluid, which could be another source of cosmic confusion. Evidently, there must be intrinsic features hidden in this model which could be used to distinguish it from the standard FLRW scenario. At the background level, it was shown in [26] that the luminosity distance in these models present a stronger angular dependence on position than the one found in isotropic models – a feature which can be detected with future supernovae experiments. Another way of constraining these models is by looking to its perturbative sector. Recently, the linear perturbation theory for the Bianchi III spacetime was constructed [27], where it was shown that perturbative dynamics also look remarkably close to the FLRW case. However, these spacetimes are not symmetric under parity transformations. This asymmetry will lead to off-diagonal couplings in the CMB covariance matrix of the form $\langle a_{\ell m} a_{\ell\pm1,m}\rangle$, where the $a_{\ell m}$s are the multipole temperature components ($\Delta T(\hat{n}) = \sum_{\ell m} a_{\ell m} Y_{\ell m}(\hat{n})$), which can in principle be detected with forthcoming CMB experiments such as Planck [28].

IV. FINAL REMARKS

The success of the standard ΛCDM model in describing global features of the universe is undeniable. Right now, cosmology faces a second order challenge in which not only the chart of the cosmic pie is precisely divided, but also carefully explained with well grounded microphysical laws. Dark energy is the main protagonist of this plot and, whether it is due to a modification of gravity, a new field or a modification of the symmetry principles, it's likely to change our understanding of the cosmos in a profound way. Among these tree routes, the possibility of dark energy arising as a geographical confusion of a specially located Earth is, justifiably, the least explored scenario. This situation is likely to change in the next decade with more and better calibrated data from large scale structure, such as Euclid [29], LSST [30] and J-PAS [31], just to mention a few. Meanwhile, existing data is already capable of ruling out void models of the universe with 2σ or even 3σ.

In this article we have presented two specific tests from the literature in which simple models are ruled out by basic physical principles. We have also discussed the possibility that the Copernican Principle can be tested in a completely geometric and model independent way, although a greater observational effort might be necessary to follow this route. One alternative perspective in the search for deviations of the Cosmological/Copernican Principle is to look up for models which can mimic existing cosmological observations. Specifically, we have shown that tests of the Cosmological Principle can be pursued with anisotropic models which preserves the isotropy of the CMB. One promising example is given by the Bianchi III and Kantowski-Sachs spacetimes, which can be further constrained with more powerful CMB experiments, as for example the Planck satellite [32].

Acknowledgements

I would like to thank the organizers of the José Plinio Baptista School of Cosmology, specially Júlio Fabris and Davi Rodrigues, for their support and kind invitation to

participate in this event.

---

[1] Joshua Frieman, Michael Turner, and Dragan Huterer. Dark Energy and the Accelerating Universe. *Ann.Rev.Astron.Astrophys.*, 46:385–432, 2008.

[2] Marcelo J. Reboucas and G.I. Gomero. Cosmic topology: A Brief overview. *Braz.J.Phys.*, 34:1358–1366, 2004.

[3] E. Komatsu et al. Seven-Year Wilkinson Microwave Anisotropy Probe (WMAP) Observations: Cosmological Interpretation. *Astrophys. J. Suppl.*, 192:18, 2011.

[4] M. Kowalski et al. Improved Cosmological Constraints from New, Old and Combined Supernova Datasets. *Astrophys.J.*, 686:749–778, 2008.

[5] G. F. R. Ellis, R. Maartens, and M. A. H. MacCallum. *Relativistic Cosmology*. Cambride University Press, England, 2012.

[6] Edmund J. Copeland, M. Sami, and Shinji Tsujikawa. Dynamics of dark energy. *Int.J.Mod.Phys.*, D15:1753–1936, 2006.

[7] G.F.R. Ellis and Malcolm A.H. MacCallum. A Class of homogeneous cosmological models. *Commun.Math.Phys.*, 12:108–141, 1969.

[8] R. Ali Vanderveld, Eanna E. Flanagan, and Ira Wasserman. Mimicking dark energy with Lemaitre-Tolman-Bondi models: Weak central singularities and critical points. *Phys.Rev.*, D74:023506, 2006.

[9] R.R. Caldwell and A. Stebbins. A Test of the Copernican Principle. *Phys.Rev.Lett.*, 100:191302, 2008.

[10] John C. Mather, E.S. Cheng, D.A. Cottingham, R.E. Eplee, D.J. Fixsen, et al. Measurement of the Cosmic Microwave Background spectrum by the COBE FIRAS instrument. *Astrophys.J.*, 420:439–444, 1994.

[11] Albert Stebbins. CMB Spectral Distortions from the Scattering of Temperature Anisotropies. *Phys.Rev.D*, 2007.

[12] Adam G. Riess, Louis-Gregory Strolger, Stefano Casertano, Henry C. Ferguson, Bahram Mobasher, et al. New Hubble Space Telescope Discoveries of Type Ia Supernovae at z¿=1: Narrowing Constraints on the Early Behavior of Dark Energy. *Astrophys.J.*, 659:98–121, 2007.

[13] W. Michael Wood-Vasey et al. Observational Constraints on the Nature of the Dark Energy: First Cosmological Results from the ESSENCE Supernova Survey. *Astrophys.J.*, 666:694–715, 2007.

[14] Pengjie Zhang and Albert Stebbins. Confirmation of the Copernican principle at Gpc radial scale and above from the kinetic Sunyaev Zel'dovich effect power spectrum. *Phys.Rev.Lett.*, 107:041301, 2011.

[15] Chul-Moon Yoo, Ken-ichi Nakao, and Misao Sasaki. CMB observations in LTB universes: Part II – the kSZ effect in an LTB universe. *JCAP*, 1010:011, 2010.

[16] N.R. Hall, L. Knox, C.L. Reichardt, P.A.R. Ade, K.A. Aird, et al. Angular Power Spectra of the Millimeter Wavelength Background Light from Dusty Star-forming Galaxies with the South Pole Telescope. *Astrophys.J.*, 718:632–646, 2010.

[17] Chris Clarkson, Bruce Bassett, and Teresa Hui-Ching Lu. A general test of the Copernican Principle. *Phys.Rev.Lett.*, 101:011301, 2008.

[18] Thiago S. Pereira, Cyril Pitrou, and Jean-Philippe Uzan. Theory of cosmological perturbations in an anisotropic universe. *JCAP*, 0709:006, 2007.

[19] Cyril Pitrou, Thiago S. Pereira, and Jean-Philippe Uzan. Predictions from an anisotropic inflationary era. *JCAP*, 0804:004, 2008.

[20] Tomi Koivisto and David F. Mota. Dark energy anisotropic stress and large scale structure formation. *Phys.Rev.*, D73:083502, 2006.

[21] Christopher A. Clarkson, A.A. Coley, E.S.D. O'Neill, R.A. Sussman, and R.K. Barrett. Inhomogeneous cosmologies, the Copernican principle and the cosmic microwave background: More on the EGS theorem. *Gen.Rel.Grav.*, 35:969–990, 2003.

[22] Christopher A. Clarkson and Richard Barrett. Does the isotropy of the CMB imply a homogeneous universe? Some generalized EGS theorems. *Class.Quant.Grav.*, 16:3781–3794, 1999.

[23] J. Ehlers, P. Geren, and R.K. Sachs. Isotropic solutions of the Einstein-Liouville equations. *J.Math.Phys.*, 9:1344–1349, 1968.

[24] Saulo Carneiro and Guillermo A. Mena Marugan. Anisotropic cosmologies containing isotropic background radiation. *Phys.Rev.*, D64:083502, 2001.

[25] J.P. Mimoso and P. Crawford. Shear - free anisotropic cosmological models. *Class.Quant.Grav.*, 10:315–326, 1993.

[26] Tomi S. Koivisto, David F. Mota, Miguel Quartin, and Tom G. Zlosnik. On the Possibility of Anisotropic Curvature in Cosmology. *Phys.Rev.*, D83:023509, 2011.

[27] Thiago S. Pereira, Saulo Carneiro, and Guillermo A. Mena Marugan. Inflationary Perturbations in Anisotropic, Shear-Free Universes. *JCAP*, 1205:040, 2012.

[28] S. Carneiro, T. S. Pereira, and G. Mena-Marugan. *To appear.*

[29] R. Laureijs, J. Amiaux, S. Arduini, J.-L. Augueres, J. Brinchmann, et al. Euclid Definition Study Report. 2011.

[30] Z. Ivezic, J.A. Tyson, R. Allsman, J. Andrew, and R. Angel. LSST: from Science Drivers to Reference Design and Anticipated Data Products. 2008.

[31] N. Benitez, E. Gaztanaga, R. Miquel, F. Castander, M. Moles, et al. Measuring Baryon Acoustic Oscillations along the line of sight with photometric redshifs: the PAU survey. *Astrophys.J.*, 691:241–260, 2009.

[32] P.A.R. Ade et al. Planck Early Results. I. The Planck mission. *Astron.Astrophys.*, 536:16464, 2011.

Structure in galaxy clusters

G.B. Lima Neto,[1] T.F. Laganá,[2] F. Andrade-Santos,[3] and R.E.G. Machado[1]

[1] *IAG/University of São Paulo, Brazil*

[2] *NAT/Unicsul, São Paulo, Brazil*

[3] *Harvard-Smithsonian CfA, Cambridge, MA, USA*

We will discuss here how structures observed in clusters of galaxies can provide us insight on the formation and evolution of these objects. We will focus primarily on X-ray observations and results from hydrodynamical N-body simulations. This paper is based on a talk given at the School of Cosmology *José Plínio Baptista* – "Cosmological perturbations and Structure Formation" in Ubu/ES, Brazil.

I. INTRODUCTION: CLUSTER FORMATION

In the current accepted cosmological scenario, where the matter content of the Universe is dominated by "cold" collisionless dark matter, clusters of galaxies are the last collapsed structures to form. N-body cosmological simulations [e.g., 1, 20, 37] show clearly that, in a cold dark matter dominated Universe, halos of 10^{14} to a few $10^{15} M_{\odot}$ form at the nodes or intersections of large scale cosmic filaments. Large halos are built hierarchically by accreting smaller objects, as shown in cosmological simulations (Fig. 1) and represented as a "merging tree" (Fig. 2). Therefore, even though clusters are usually taken as already formed structures, we should expect matter infall and accretion along the cosmic filaments, which may still impact cluster evolution.

The hierarchical formation of collapsed structures will determine the halo mass function through the density fluctuation power spectrum. The first cosmological analysis of the mass function in a dark matter dominated Universe was made by Press & Schechter [32], where they assumed that primordial density fluctuations were Gaussian, and that regions at any given time which exceeds a certain density constant are collapsed,.

Even with its shortcomings (an *ad hoc* factor 2 that had to be inserted in the mass function), their mass function was remarkably accurate compared to N-body

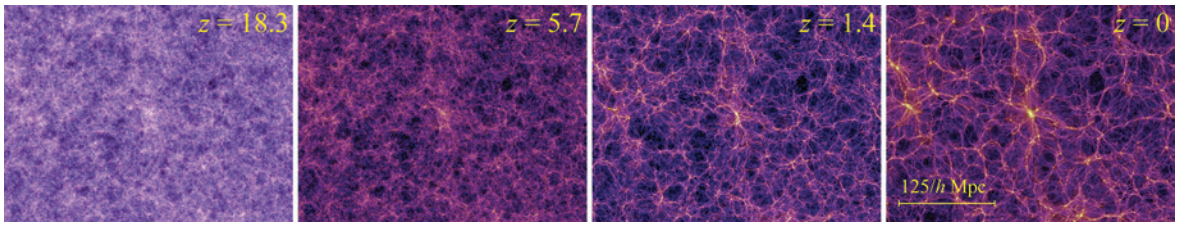


Figure 1. Snapshots from the Millennium Simulation [37] showing the time evolution of a slice of the Universe. Large haloes are formed at the intersection of cosmic dark matter filaments by accreting mass.

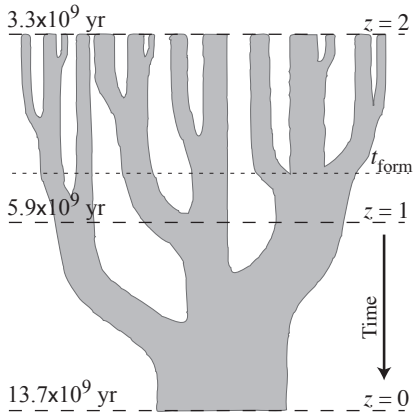


Figure 2. Merging tree showing the events leading to the formation of a massive dark matter halo at $z = 0$. $t_{\rm form}$ corresponds to the instant that half of the present halo mass is already assembled. Figure adapted from Lacey & Cole [24].

simulations. During the last 15 years, the halo mass function was modified using the results of cosmological simulations [e.g. 18, 40].

The mass function is an important cosmological probe, in particular for the mean mass density of the Universe, Ω_M, and the amplitude of the density fluctuation power spectrum. Fig. 3 illustrates how the mass function is different for different cosmological parameters. Notice that the abundance of rich clusters with $M > 10^{14} M_\odot$ gives strong constraints for the halo mass function.

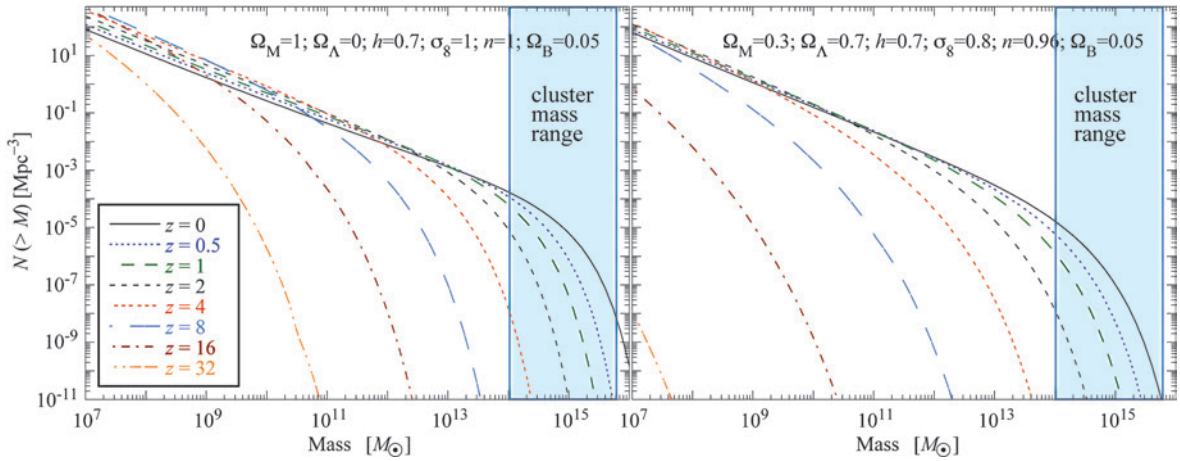


Figure 3. Halo mass function as a function of mass at different redshifts. Left: Mass function for an Einstein-de Sitter Universe (flat, $\Omega_M = 1$). Right: Mass function for a standard ΛCDM scenario. The curves are computed using the Press-Schechter mass function and an approximate power spectrum similar to the one given by Tegmark et al. [39].

II. CLUSTERS OF GALAXIES

Clusters of galaxies are the largest collapsed structures in (quasi) equilibrium in the Universe. Rich clusters have thousands of galaxies but, since they are relatively rare only about 5–10% of galaxies live inside clusters [see 4, for a review].

It is usual to refer to collapsed structures with mass greater than about $10^{14} M_\odot$ as clusters, while for objects with masses between $\sim 10^{13}$ and 10^{14} we call them groups. The typical radii of clusters are 1–3 Mpc (a value close to the Abell radius, $1.5h^{-1}$Mpc or $R_A = (1.72/z)$ arcmin). The mean density is ~ 200 times the critical density of the Universe at the cluster redshift.

Galaxy clusters are also characterized by their galaxy population, largely dominated by early-type galaxies (elliptical and lenticular), contrary to the field where spiral galaxies are more common.

The galaxies in clusters correspond only to $\sim$ 2–5% of the total mass. Most of the baryonic matter is in the form of a hot diffuse plasma, the intracluster gas, amounting to about 12–15% of the total mass, i.e., there is roughly 6 times more baryons in the gas than in all galaxies together. However, most of the mass, about 85% is in the form

of dark matter.

The dark component cannot be directly observed; its presence is inferred by gravitational effects. Under the hypothesis that clusters are in equilibrium, the total mass may be obtained using either the galaxies as tracer (by the virial theorem) or the intracluster gas, by hydrostatic equilibrium, which results in the well know formula for the total mass inside the radius r:

$$\frac{1}{\rho}\vec{\nabla}P = -\vec{\nabla}\Phi \quad \Rightarrow \quad M(r) = -\frac{kT}{G\,\mu\,m_{\rm H}}\,r\left(\frac{d\ln\rho}{d\ln r} + \frac{d\ln T}{d\ln r}\right), \tag{9.1}$$

where we have assumed spherical symmetry.

A third evidence of dark matter comes without the equilibrium hypothesis from gravitational lens effect – the cluster mass acts as a gravitational lens on the background galaxies and the total projected mass along the line-of-sight may be estimated [see 6, 21]

A. The intracluster gas

The intracluster gas has typical temperature ranging from 10^7 to 10^8 K, that are generally given in the literature in energy units: $2 < kT < 12$ keV, approximately. It is a very rarefied plasma, with typical central number densities of $n_0 \approx 10^{-2}$–10^{-3} cm^{-3}. Such a plasma is a strong X-ray source emitting through thermal bremsstrahlung, i.e. free–free scattering process of electrons by ions [see, e.g., 7, 35].

The intracluster plasma is metal enriched by processed gas in galaxies, having a typical mean metallicity of a third of the solar metallicity. The heavier elements are responsible for producing the recombination lines observed in the X-ray spectrum of clusters. The expected emission from a 5 keV cluster is shown in Fig. 4.

The gas looses energy by bremsstrahlung emission, which is approximately proportional to $n^2T^{1/2}$, so we can define a cooling-time as:

$$t_{\rm cool} \approx \frac{E}{dE/dt} \quad \Rightarrow \quad 9.3\times10^9\frac{(kT_{\rm keV})^{1/2}}{n_3}{\rm years}\,, \tag{9.2}$$

where kT is given in keV and n_3 is the gas density in units of 10^{-3} cm^{-3}, and E and dE/dt are the thermal energy and the cooling ratio of the gas.

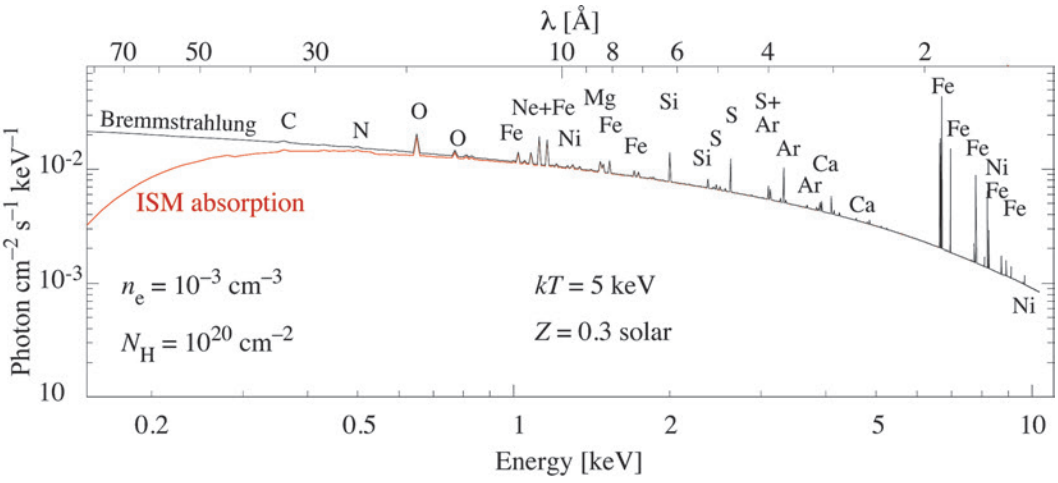


Figure 4. Theoretical emission from the intracluster gas, bremsstrahlung continuum plus emission lines. The red curve corresponds to the absorbed spectrum, mainly from the interstellar HI from our Galaxy.

The relatively short cooling times at the core of clusters have led to the scenario where cool gas would flow to the cluster center at rates up to $\sim 1000 M_\odot$/yr, a phenomenon know as cooling-flow [13, 14]. The lack of great amounts of cool gas, no evidence of ongoing star formation, and, observations of X-ray spectra led to the actual scenario where the gas is heated and the cooling-flow is greatly reduced. The most popular hypothesis is that the activity of the central AGN is regulated by the infall of gas producing a feedback mechanism that keeps the gas temperature at roughly 1/3 of the virial temperature value [e.g., 29].

Still, due to its short relaxation time, the intracluster gas is an important probe of cluster dynamics. The X-ray surface brightness of clusters (Fig. 5), which is the projection of n^2 on the plane of the sky, reveals their dynamical state, the presence of substructures and on-going collisions. Temperature maps are even more revealing, showing the presence of shocks and cold fronts moving through clusters.

With very deep imaging one can perform detailed morphological analysis, specially of nearby clusters, that shows a wealth of substructures: bubbles (cavities), filaments, and even acoustic waves. Some of them are driven by non-gravitational process linked to the central AGN activity. These structures impact on the global cluster properties and affect their scaling relations (such as T_X–L_X) currently used, for example, for

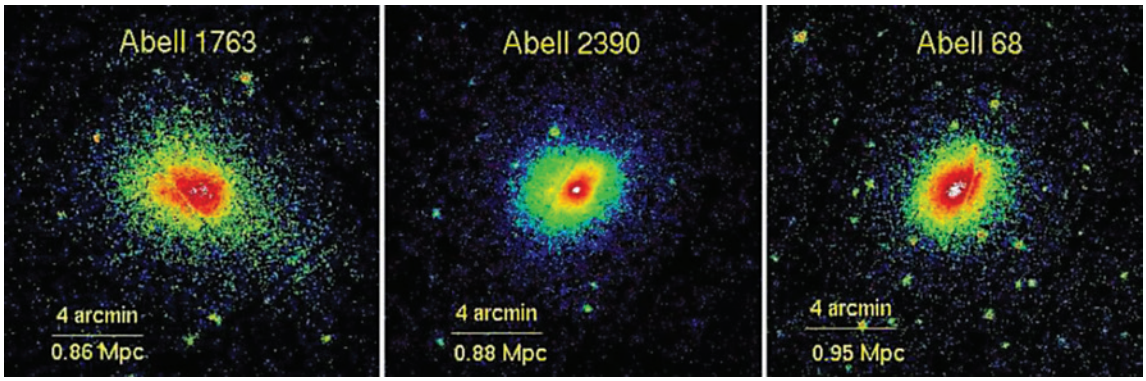


Figure 5. Example of three clusters observed with XMM-*Newton*. The images show the X-ray surface brightness. Even for relaxed clusters it is possible to see that they are not perfectly regular (i.e., axial symmetric), indicating some previous dynamical activity.

cosmological studies. Figure 6 shows an example of the kind of bias that may be introduced when substructures are ignored when selecting a sample. This motivates the detailed study of individual clusters both from the morphological point of view and dynamically, helped by N-body simulations.

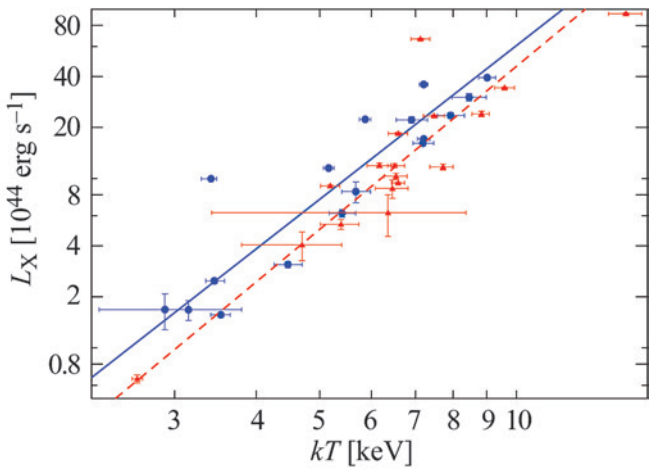


Figure 6. L_X–T_X relation fitted separately for clusters showing a high level of substructures (red dashed line) and a low level of substructures (blue continuous line). For a given luminosity, substructured clusters are hotter then more relaxed clusters. Figure taken from Andrade-Santos et al. [2].

B. Search for substructures in clusters

Motivated by the dynamical information we can infer from the analysis of irregular morphology in clusters, many authors have searched and quantified substructures in clusters. Focusing on X-ray observations, the first systematic study was done by Jones & Forman [17] based on *Einstein* data.

Observation of the surface brightness allows us to derive quantitatively the degree of substructure of clusters. Various tests were proposed in the literature: centroid and ellipticity level variation [31], moments of the expansion in Fourier series [8], power-ratio method [19], ratio of fluxes between residual and original X-ray images [2], among others.

Laganá et al. [25] chose a different approach, using a qualitative analysis based on the residual image after subtraction of a smooth elliptically symmetrical model. Selecting a sample of 15 nearby clusters, $z < 0.06$, i.e., within $250h_{70}^{-1}$ Mpc from us, observed by *Chandra* for more than 30 ks, they examined the surface brightness and temperature map morphologies. Figure 7 shows the large scale substructures in three of these clusters. The most striking feature is the spiral arm structure.

In half of their sample a spiral-like arm substructure is observed, all of them in cool-core clusters. Such a substructure is more easily seen in the residual map than in the hardness ratio map. All clusters show, in some degree, substructures (not necessarily an arm structure).

In order to verify if the arm feature was not an artifact of the image processing algorithm, Andrade-Santos et al. [2] has performed several tests with simulated cluster images (cf. an example in Fig. 8). Gas hydrodynamics is essential to be well understood if we want to use X-ray observables as mass estimators.

The arm substructures have a scale of hundreds of kiloparsecs, larger than the region where cavities are produced by the AGN activity. The mechanical work, $P\,dV$, from the AGN may not be enough to push the gas and generate these observed arms. The work required to form an arm with ~ 200 kpc is roughly $\sim 10^{61}$ erg or 5×10^{45} erg s^{-1} for 10^8 years. Minor collisions (that may lead to a merger) may provide an alternative source of energy through tidal forces, by "shaking" the cool-core. A small cluster/rich

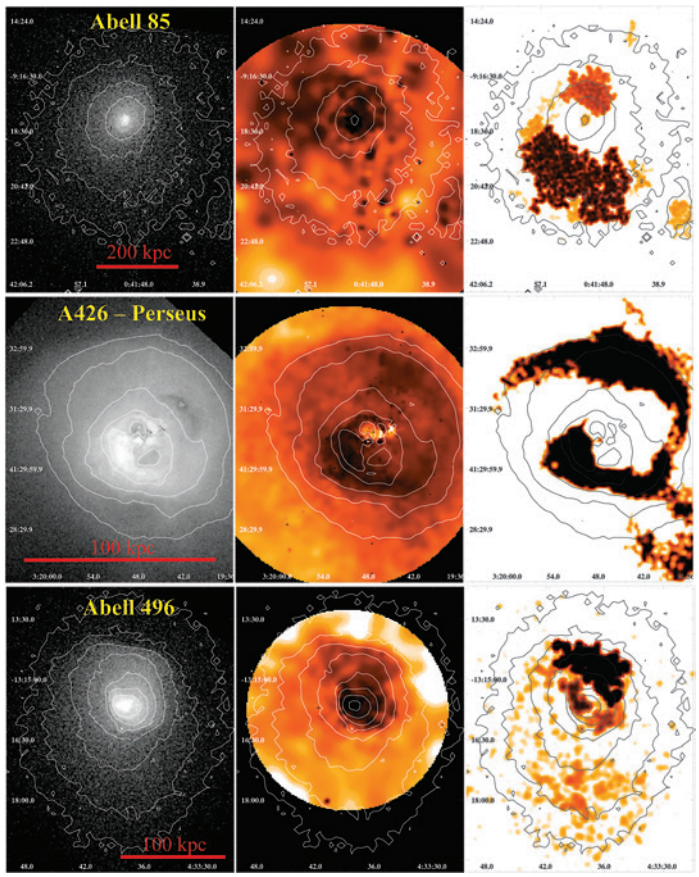


Figure 7. From left to right: surface brightness image, hardness ratio (temperature) map, and residual substructure distribution, showing the most prominent arm substructure (from Laganá et al. [25]), seem in both temperature and substructure maps.

group with $10^{14} M_{\odot}$ moving with a modest velocity of 500 km s^{-1} has 5×10^{62} erg of kinetic energy; only a small fraction of this energy would be enough to generate a large scale spiral arm on the intracluster gas.

Ascasibar & Markevitch [3], ZuHone et al. [42] have shown with N-body simulations that off-center collisions of a cluster with a galaxy group – mass ratio between 1/2 to 1/10 – may slosh the cool central gas and produce the arms we observe. Notice that this mechanism is efficient to move the cool gas from the center but it is not a heating mechanism to stop or regulate the cooling-flow.

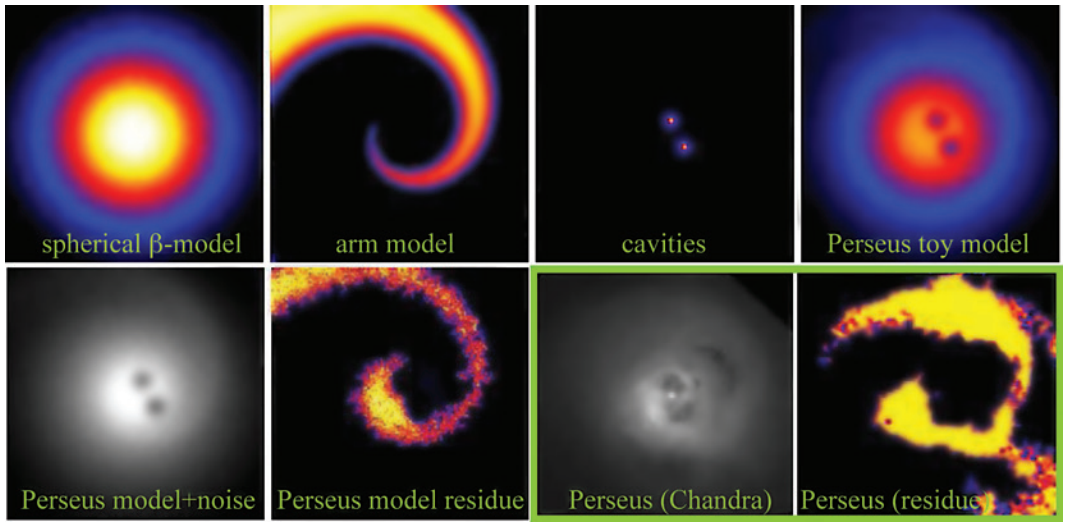


Figure 8. Test of detectability of the spiral feature (arm) in clusters. Top: construction of a toy model of the Perseus X-ray surface brightness. Bottom: application of the same method for the toy model and *Chandra* Perseus data. Figure from Laganá et al. [25]

1. Cluster–AGN connection

Even though cluster AGNs may not be powerful enough to generate large scale substructures, the radio jets – relativistic particles ejected from near the central massive black hole – interacts with the intracluster gas and have important consequences. The favored mechanism to suppress the cooling-flow is AGN feedback [e.g., review by 16, 29].

Most, if not all, Brightest Cluster Galaxies (BCGs) have radio emission and some present strong jets (e.g., Cygnus A, Hydra A, M 84, NGC 1275 [Perseus A]) that blow "cavities" (also called bubbles or holes) in the intracluster gas (see, e.g., Fig. 9), observed in high resolution X-ray images. The evidence for this lies in the spatial coincidence between the observed synchrotron radio emission and the regions where there is a decrease in X-ray flux. Those cavities are also buoyant and probably migrate outwards from the center, towards lower density regions.

Very deep imaging of the Perseus Cluster [15] confirms earlier finding that the AGN activity through their jets produces acoustic (pressure) waves that propagate outwards throughout the intracluster plasma. A challenging task would be to detect these waves

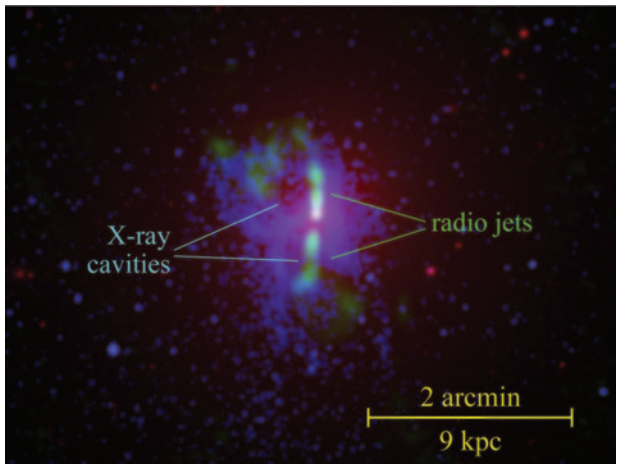


Figure 9. Composite image of M84. Red: SDSS r band image; Green: VLA 1.4 First Survey; Blue: Chandra 0.3–7.0 keV band. The X-ray "cavities" and radio jets are indicated.

in other clusters. Therefore, it seems plausible that the AGN energy is converted to mechanical and thermal energy, heating the gas in a feedback process.

III. CLUSTER COLLISIONS

Massive clusters are built by way of major cluster collisions leading to a hierarchical merging process. Not only does the study of merging clusters provide a way to understand large scale structure formation in the Universe, but it also helps us understand galaxy evolution, the interplay between the intracluster medium and galaxies, and the nature of dark matter.

One way to approach this problem is to follow the growth of clusters with cosmological simulations, with a minimum of *ad hoc* assumptions. This requires a great amount of computational power, that is still somewhat limited by mass/spatial resolution in particular if the simulation is hydrodynamical – usually only collisionless dark matter is simulated [see, e.g., 22].

Another option is to simulate individual systems with self-consistent N-body hydrodynamical simulations, sometimes even with sub-grid physics, i.e., star formation, feedback, cooling, metal enrichment, and other physical process usually called "gastrophysics" by the community, and are dealt with semi-analytical models [e.g.

23, 36, 41] The main setback here is to ignore the continuous matter accretion in clusters in a self-consistent way, focusing only in a given event, a cluster-cluster collision.

A. The Bullet Cluster 1E 0657-558

The most studied colliding cluster today is doubtless 1E 0657-558, also known as the *Bullet Cluster* [27]. This cluster, located at redshift $z = 0.296$, acquired its nickname because of a prominent bow shock observed with high resolution X-ray imaging, suggesting that a smaller cluster is passing through a lager cluster at high speed: the bow shock appears as a consequence of the gas collision at Mach number ~ 3 almost on the plane of the sky.

Deep optical imaging shows two mass concentrations detected from their weak lensing signal, spatially coincident with the galaxy distribution of both colliding clusters (the "bullet" and the main cluster), but displaced from the two gas concentrations (detected in X-rays). N-body hydrodynamical simulations [e.g., 28, 38] confirm the collision scenario of the Bullet Cluster, where the relative velocity of the clusters halos is ~ 2700 km s^{-1} (actually smaller than the shock speed), and reproduce the observed separation of the bulk of mass from the intracluster gas. The displacement of the intracluster plasma with respect to the dominating halo dark matter is due to the gas dissipative nature and provides strong constraints for the hypothetical dark matter particles cross section [in the simplest scenario, dark matter is collisionless, 9].

There are a few other clusters that were recently discovered to have the gas component dissociated from the dark matter, for instance, Abell 2744 [30], Abell 1758 [33], and DLSCL J0916.2 [10].

B. Abell 3376 – a nearby bullet

The cluster Abell 3376 is a nearby cluster, $z = 0.046$, that shows signs of having passed recently through a major collision [5]. The XMM-*Newton* X-ray surface brightness map shows a cometary-like morphology, similar to 1E 0657-586. Moreover,

the brightest cluster galaxy (BCG) is displaced from the central X-ray emission peak, the gas temperature map shows alternating hot and cold gas, and deep VLA observation shows a huge ($\sim 2 \times 1.6$ Mpc on the plane of the sky) diffuse synchrotron radio emission that suggests a structure in an ellipsoidal distribution around the cluster (Fig. 10). This radio emission is possibly generated by electrons accelerated by Fermi process due to the shock of cluster collision. Analyses of the galaxy distribution and dynamics [11, 12, 34] show that Abell 3376 has an elongated, substructured distribution along the putative collision path.

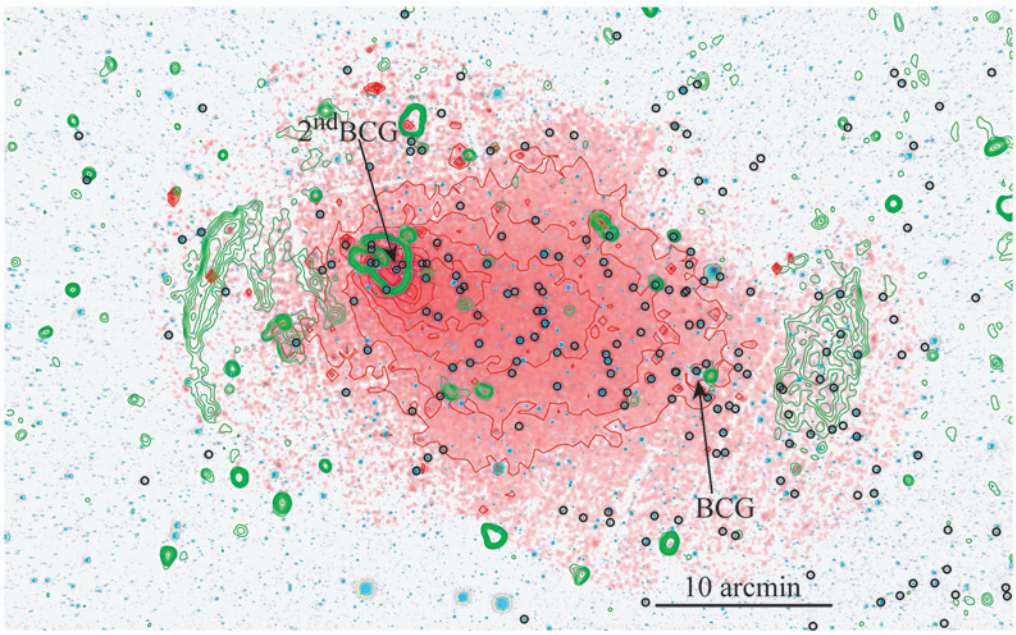


Figure 10. XMM-*Newton* image of Abell 3376 in the [0.3–8.0] keV band (red shades and contours) with radio VLA 1.4 GHz contours (green) and Blanco/Mosaic II B-band image (blue). The black circles represent galaxies in the $0.036 < z < 0.056$ redshift range. The two brightest galaxies are indicated.

Recently, Machado & Lima Neto [26] carried out a large set of high-resolution adiabatic hydrodynamical N-body simulations, in order to verify and model the collision event of Abell 3376. They were able to propose a specific scenario for the collision: the most successful model – a 1:6 mass ration, head-on collision at Mach ~ 3.5 viewed with a 40° with respect to the plane of the sky – accounts for several of the features observed in the cluster (gas morphology, temperature, virial mass, total

X-ray luminosity). Interestingly, a testable theoretical prediction emerges from this scenario, concerning the distribution of dark matter. In the resulting configuration, the dark matter (and therefore the bulk of the total mass) is predicted to be concentrated around two density peaks where the cluster brightest galaxy (BCG) and the second BCG are located (Fig. 11).

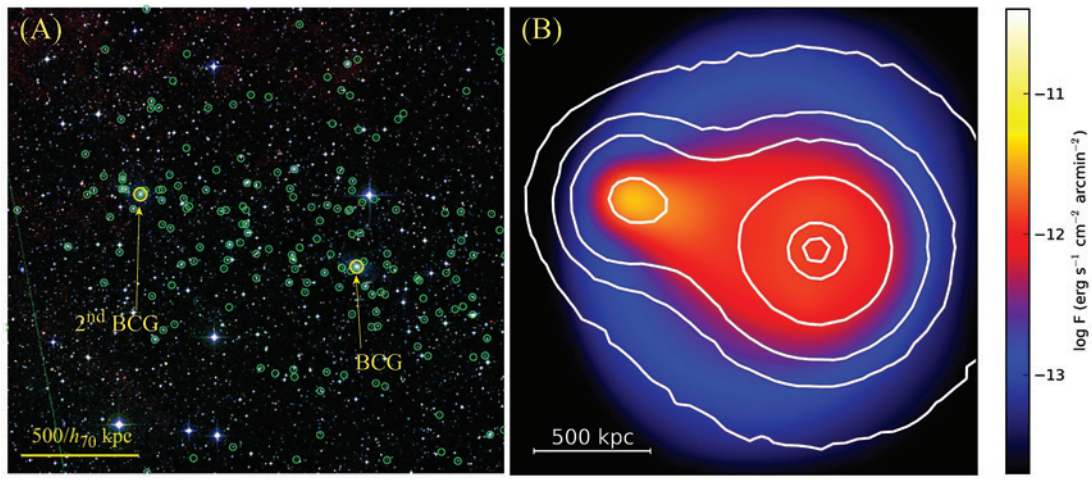


Figure 11. (A) DSS optical image in which the circles mark positions of the galaxies at the cluster redshift. The first BCG (right) and the second BCG (left) are highlighted. **(B)** Simulated X-ray surface brightness after the cluster collision, where the white contour lines show the total projected mass from Machado & Lima Neto [26] best model.

It is intriguing to notice that, contrary to the Bullet Cluster, the *simulated* mass distribution seems to follow the *simulated* gas distribution. The *real* mass distribution of Abell 3376 is, as of yet, unknown. If the mass map approximately matches the simulation results, this would lend more weight to our scenario, and might even set tighter constraints. If, on the other hand, the mass map reveals some complicated or unexpected dark matter structure, then the current models will be shown to be insufficient, and alternative models will have to be explored. Hence, mapping the total mass distribution through gravitational weak lensing effect on background galaxies is an important step for our understanding of this system.

IV. CONCLUDING REMARKS

Galaxy clusters are fundamental cosmological probes and giant astrophysical laboratories. In order to fully use these objects for precision cosmology and astrophysics, we must well grasp the physical process occurring with them. We briefly explored here the possibilities open by the X-ray astronomy and numerical simulations of cluster interactions, discussing a few recent results on substructure quantitative determination, AGN feedback, the consequences of off-center collisions and sloshing, and head-on supersonic collisions generating bow shocks.

Acknowledgments: GBLN thanks the Brazilian agencies CNPq and CAPES for financial support and the hospitality during the *José Plínio Baptista* School. TFL and REGM thank FAPESP (fellowships 2012/00578-0 and 2010/12277-9, respectively) We also acknowledge enjoyable and fruitful discussions with Florence Durret and Melville Ulmer.

[1] Alimi J.-M., Bouillot V., Rasera Y., et al., 2012, arXiv:1206.2838
[2] Andrade-Santos F., Lima Neto G.B., Lagan T.F., 2012, ApJ 746, id. 139
[3] Ascasibar Y., & Markevitch M. 2006, ApJ 650, 102
[4] Bahcall N., 1999, in: "Formation of structure in the universe", p. 135, A. Dekel, J.P. Ostriker, Eds., Cambridge University Press
[5] Bagchi J., Durret F., Lima Neto G.B., Surajit P., 2006, Science 314, 791
[6] Bartelmann M., Schneider P., 2001, Phys. Rep. 340, 291
[7] Böhringer H., Werner N., 2010, Astr. Astrophys. Rev. 18, 127
[8] Buote D.A., Tsai J.C., 1995, ApJ 452, 522
[9] Clowe D., Bradač M., Gonzalez A.H., 2006, ApJLett 648, L109
[10] Dawson W.A., 2013, ApJ 772, 131
[11] Durret F., Perrot C., Lima Neto G.B., et al., 2013, A&A *in press* (arXiv:1310.7493)
[12] Escalera E., Biviano A., Girardi M., et al., 1994, ApJ 423, 539
[13] Fabian A.C., Nulsen P.E.J., Canizares C.R., 1984, Nature 310, 733
[14] Fabian A.C., 1994, ARA&A 32, 277
[15] Fabian A.C., Sanders J.S., Allen S.W., et al., MNRAS 418, 2154
[16] Fabian A.C., 2012, ARA&A 50, 455
[17] Jones C., Forman W., 1984, ApJ 276, 38

[18] Jenkins A., Frenk C.S., White S.D.M., et al., 2001,MNRAS 321, 372
[19] Jeltema T.E., Canizares C.R., Bautz M.W., Buote D.A., 2005, ApJ 624, 606
[20] Klypin A.A., Trujillo-Gomez S., Primack J., 2011, ApJ 740, 102
[21] Kneib J.-P., Natarajan P., 2011, A&A Rev. 19, id #47
[22] Kuhlen M., Vogelsberger M., Angulo R., 2012, Phys. Dark Universe, 1, 50
[23] Lacey C., Silk J., 1991, ApJ 381, 14
[24] Lacey C., Cole S., 1993, MNRAS 262, 627
[25] Laganá, T.F., Andrade-Santos F., Lima Neto G.B., 2010, A&A, 511, A15
[26] Machado R.E.G., Lima Neto G.B., 2013, MNRAS 430, 3249
[27] Markevitch M., Gonzales A.H., David L., et al., 2002, ApJLett 567, L27
[28] Mastropietro C., Burkert A., 2008, MNRAS 389, 967
[29] McNamara B., Nulsen P., ARA&A, 45, 117M, 2007
[30] Merten J., Coe D., Dupke R., et al., 2011, MNRAS 417, 333
[31] Mohr J.J., Evrard A.E., Fabricant D.G., Geller M.J., 1995, ApJ 447, 8
[32] Press W.H., Schechter P.L., 1974, ApJ 187, 425
[33] Ragozzine B., Clowe D., Markevitch M., Gonzalez A.H., Bradač M., 2012, ApJ 744, 94
[34] Ramella M., Biviano A., Pisani A., et al., 2007, A&A 470, 39
[35] Sarazin C., 1988, "X-ray emission from clusters of galaxies", Cambridge University Press
[36] Somerville R.S., Hopkins P.F., Cox T.J., et al., 2008, MNRAS 391, 481
[37] Springel V., White S.D.M., Jenkins A., et al., 2005, Nature 435, 629
[38] Springel V., Farrar G.R., 2007, MNRAS 380, 911
[39] Tegmark M., Blanton M.R., Strauss M.A., et al., 2004, ApJ 606, 702
[40] Tinker J., Kravtsov A.V., Klypin A., Abazajian K., et al., ApJ 688, 709
[41] White S.D.M., Frenk C.S., 1991, ApJ 379, 52
[42] ZuHone J.A., Markevitch M., Lee D., 2011, ApJ 743, id 16

Este impresso foi composto utilizando-se a família tipográfica Euclid. Sua capa foi impressa em papel Supremo 300g/m² e seu miolo em papel Offset 90g/m², na cor branca, medindo 19 x 26,9 cm, com uma tiragem de 300 exemplares.

É permitida a reprodução parcial desta obra, desde que citada a fonte e que não seja para qualquer fim comercial.

EDUFES