ISAS - INTERNATIONAL SCHOOL FOR ADVANCED STUDIES

Clustering Properties of Dark Matter Haloes in Hierarchical Models for Structure Formation in the Universe

Thesis submitted for the degree of
Doctor Philosophiæ

Astrophysics Sector

Candidate:

Cristiano Porciani

Supervisors:

Dennis Sciama

Sabino Matarrese

SISSA - SCUOLA INTERNAZIONALE SUPERIORE DI STUDI AVANZATI

TRIESTE
Via Beirut 2-4

April 1999

TRIESTE

To Miriam and Sergio

Contents

Acknowledgements

I would like to take this opportunity to express my immeasurable gratitude to Sabino Matarrese for all that I learned from him over my four years at SISSA. His expertise, knowledge, friendship and willingness to help were always available and greatly appreciated. I was fortunate to work closely with him, and I do hope that our collaboration will last for long. Many thanks to him also for carefully reading the manuscript version of this thesis.

It is a pleasure to thank my supervisor Dennis Sciama for his support and help. He has been a fine mentor and I am privileged to have studied with him. His passion for science will always represent a valuable example for me.

I would like to thank the rest of faculty, staff and students at SISSA for providing such a stimulating and enjoyable environment during my years in Trieste. In particular, I am indebted to John Miller who read part of this thesis and to Alex, Andrea and Riccardo for their constant help.

I would like to thank Piero Madau for the support and encouragement he gave me during the wonderful nine months I spent at STScI. My collaboration with him extended my scientific horizon.

A special thank you goes to Paolo Catelan. I benefited greatly from the many hours spent working with Paolo during my stay in Copenhagen. Our subsequent collaboration, that unfortunately developed mainly by phone and e-mail, not only led to several joint publications but was also most enjoyable.

I am indebted to Francesco Lucchin and Lauro Moscardini for many discussions and for their constant support and help.

There are many other excellent scientists I have met and benefited from during my graduate studies: Andrea Ferrara, Federico Ferrini, Cedric Lacey, Colin Norman, Nino Panagia and Massimo Stiavelli. In particular Colin is thanked, together with Caleb Scharf and Marcella Carollo, for the help they kindly provided me during the preparation of my post-doc applications.

Many thanks to Scilla degl'Innocenti and the whole faculty of the Astronomy Department of Pisa University who supplied me with a computer every time I was visiting my parents.

Alejandra, Bruce, Carlo, Cecilia, Elena, Enrica, Francesca, Fu Kun, Giovanni, Giuliano, Laura, Marco, Massimo, Paola, Pierluigi, Stefano, Valentina, Vanessa and You Hong touched me with their special brand of friendship. Guys, I do hope to meet you all again in the near future! Benedetta, Luciano, Micol and Stefano, have been my companions during a wonderful spring in Copenhagen. Thanks for the enjoyable time we spent together. Claudia, Duilia, Elena, Gys, Hermine, Marek, Marco, Martino, Pat, Paul, Riccardo, Rosa, Sandra, Stefania, Tim and Tommaso formed the pre-post-doc team of STScI. They are all thanked. A special thank you to my flatmates, Giuseppe, Luciano and Paolo, for providing a pleasant and sociable home environment.

Finally, I would like to thank my parents, Miriam and Sergio, who were absolutely terrific in assuring their constant moral (and financial) support and encouragement. This thesis is dedicated to them.

My last thought goes to Valeria, thank you for being like you are.

Introduction

Astronomical observations show that galaxies are not sprinkled uniformly throughout space (in a Poisson sense) but form complex and coherent patterns. In particular, deep redshift surveys reveal a foam-like textured galaxy distribution, characterized by the presence of large underdense regions (the so-called "voids", with typical dimension $\sim 50 - 100$ Mpc) surrounded by thin overdense "walls" (with thickness $\sim$ few Mpc). Galaxies present various levels of clustering: many of them occur in pairs or small groups which in turn are parts of larger associations (the galaxy clusters). A filamentary pattern that joins galaxy clusters into a complex web is also manifest. All these features that characterize the departure from randomness of the galaxy distribution are generally indicated with the term "large-scale structure" of the universe (hereafter LSS).

From the theoretical point of view, it is widely believed that the structure we see formed through gravitational instability of small density perturbations generated by quantum effects in the early universe (see chapters 1 and 2).

Observations give also strong evidence for the presence of dark matter: a material component that cannot be directly detected by gathering photons of any wavelength (see chapter 1 for a review). A popular speculation is that the dark matter is constituted by a "sea" of weakly interacting particles, left over from the earliest phases of the universe (see chapter 1). In this framework, for the most popular dark matter candidates (with the exception of massive neutrinos, see chapter 2), the growth of structure develops hierarchically: gravitational instability first forms relatively small objects, while larger clumps are generated through accretion and merging of smaller units (see chapter 4). The inner parts of these gravitationally bound objects, which relax to a quasi-equilibrium state, are usually called dark matter haloes (see chapter 4). Galaxies are expected to form from the gas that cools and condenses within the potential wells generated by these haloes.

Since galaxy formation can happen only in dense environments (where the gas can cool), the distribution of galaxies is likely to be different from the overall distribution of dark matter (see chapter 6). Understanding this difference (the so-called galaxy bias) is crucial to compare observational data with the predictions of models for structure formation, especially in order to estimate the value of some cosmological parameters (see chapters 1, 2 and 6). Unfortunately, at present, we know little about the details of star and galaxy formation which, involving complex fluid-dynamical and radiative processes, cannot be modelled reliably. From a theoretical perspective, a first step towards understanding the clustering properties of galaxies is then achieved by studying how the dark haloes are distributed in space. For instance, this can be done numerically, by simulating the process of gravitational clustering. However, even though N-body simulations give very accurate results on a limited range of scales, they do not provide direct understanding of the physical processes involved. Moreover, they are limited both in resolution and in dynamical range. Analytical methods, can be alternatively used to describe clustering growth. There are, of course, some details which such an approach cannot bring out, but the analytical treatment has the advantage of providing deep insight into the physics of gravitational clustering.

In this thesis, we present original results dealing with the formation and evolution of large-scale

structure. In particular, we develop some analytical models to follow the evolution of both the mass and the halo distribution. The models here presented are tested against N-body simulations and some of their astrophysical applications are also discussed.

The layout of this thesis is as follows. In chapter 1 we introduce some basic elements of modern cosmology, describing the main features of the standard hot big-bang model and the concept of inflation. Special attention is devoted to show that recent observational data are consistent with the theoretical framework. In chapter 2 we review the theory of gravitational instability in an expanding universe. Both the Eulerian and the Lagrangian perturbative approaches to the evolution of density perturbations are discussed in detail. The spherical top-hat model and a series of dynamical approximations are also presented. Chapter 3 deals with scaling solutions to the problem of clustering growth. After a brief review, we present some original results regarding the evolution of the autocorrelation function of the mass density field. In particular, we test the predictions of some empirically calibrated scaling Ansätze against the analytical solutions obtained by using the Zel'dovich approximation. Chapter 4 is devoted to the issue of hierarchical clustering. First, we describe the Press-Schechter theory for the abundance and mass distribution of dark matter haloes, and its excursion-set extension. A new model for the clustering of dark haloes in Lagrangian space is then discussed. In chapter 5 we present an astrophysical application of the Press-Schechter formalism. In particular, we investigate the lensing effect of background supernovae due to mass condensations in three popular CDM cosmologies. Our results suggest that it is not inconceivable that new and existing search teams will soon be able to detect magnified supernovae by conducting deep pencil beam surveys. In chapter 6 we review the theory of biased galaxy formation. Both the original motivations that led to its formulation and recent developments are discussed. In chapter 7 we present a new stochastic approach to the clustering evolution of dark matter halos in Eulerian space. Our results clearly point to a characterization of the halo-to-mass biasing as a highly non-linear and non-local process. This chapter contains some of the most important results of this thesis. We devote chapter 8 to compare the predictions of the models introduced in chapters 4, 6 and 7 with N-body simulations. In appendix A we introduce the basic concepts of the theory of random fields, while in appendix B we briefly summarize the main results of classical kinetic theory. The definition of the n-point correlation functions for a population of discrete objects (e.g. galaxies) is also given.

Conclusions are presented in the last section of those chapters which contain original results (namely, chapters 3, 4, 5, 7, 8).

The material presented in this thesis represents only a fraction of the work I carried out during my graduate studies. In particular, a numerical study of the percolating phase characterizing the process of inhomogeneous reionisation (in collaboration with Piero Madau and Avery Meiksin), and the application of stochastic optimization and pattern recognition algorithms to identify coherent features in the cosmic web (in collaboration with Manolis Plionis) are not discussed here.

Chapter 1

The standard hot big-bang model

Since the dawn of civilization, cosmological investigation (comprising the study of the origin and the structure of the observable universe) has been intimately connected to the philosophical systems in vogue, and its evolution coincided with a progressive removal of anthropocentric theories. Somehow, this evolutionary path diverted the cosmological research from the tracks of myth and religion to those of science.

Modern cosmology developed in the last 80 years, after Einstein's presentation of the first mathematical model for the universe based on general relativity. Since 1917, both observational results and theoretical speculation contributed to the development of what is now (almost universally) considered the standard description of cosmic evolution: the so called "hot big bang" model which represents the basis for cosmological investigation. This chapter is devoted to illustrate its main features and is organized as follows.

Section 1.1 briefly describes the main kinematical and dynamical properties of the model, introducing the notation that will be adopted in the whole thesis. A sketch of the main stages of the thermal history of the universe is also included. For more details (and for many of the calculations) we refer the reader to classical general-relativity books like Weinberg (1972), Misner, Thorne & Wheeler (1973), Rindler (1977) and Wald (1984).

The hot big bang model relies on three observational cornerstones: the confirmed expansion of the universe, the detection of a relic microwave background and the agreement between predicted and measured primordial chemical abundances. In section 1.2 we quickly review how these astronomical results support the standard model. We then list up-to-date estimates of the cosmological parameters, showing that a coherent description of the observable universe in the framework of the hot big bang model is available.

Section 1.3 deals with some extensions of the standard model that recently gained a lot of popularity because of their attractive features. In particular, we discuss the inflationary paradigm for the early universe, showing how it can solve some "problems" of the standard model and give birth to primordial density perturbations. Inflation is a scientifically falsifiable theory and, till now, has not been contradicted by experiments. However, the present lack of direct observational support, does not permit to include it into the standard picture.

1.1 Model overview

1.1.1 Robertson-Walker metric

The first problem to face for the construction of a mathematical model of the universe is the choice of a coherent dynamics for all the components (matter, radiation, geometry) that will appear in

it. The standard cosmological model adopts Einstein's theory of General Relativity (hereafter GR) and describes the universe as being permeated by a continuous mass-energy distribution, called the "cosmological fluid" (like in fluid-dynamics, this description will be meaningful only on scales much larger than the typical separation between the single fluid units, i.e. the galaxies). It is then assumed that the universe is spatially homogeneous and isotropic on large-scales (cosmological principle). In GR jargon, the universe is said to be spatially isotropic about every point if, at any event, an observer who is moving with the cosmological fluid cannot distinguish one of his space directions from the others by any local physical measurement. Spatial homogeneity, instead, means that through each event passes a hypersurface of homogeneity on which the physical conditions are the same everywhere. These definitions can be more precisely formulated using differential geometry (see e.g. chapter 13 in Weinberg 1972 or section 5.1 in Wald 1984). Note that isotropy about every point implies homogeneity.

Historically the cosmological principle was introduced as a simplifying hypothesis. Nowadays, even though astronomical observations reveal the existence of huge inhomogeneities on galaxy scales, there is indeed strong evidence for the cosmological principle to hold on much larger length scales (for a recent review see Wu, Lahav & Rees 1998):

1. The temperature of the cosmic microwave background (hereafter CMB) is isotropic by about one part in 10^5 on angular scales ranging from some degrees to fractions of a degree (which roughly correspond to length scales $\sim 300 - 1000$ Mpc);

2. The distribution of radio-sources and the X-ray background are isotropic to a few percent (Webster 1976; Treyer et al. 1998). Recent theoretical analyses of these data suggests that the corresponding sources are nearly homogeneously distributed on scales ~ 600 Mpc (Lahav, Piran & Treyer 1997; Baleisis et al. 1998).

3. Studies of the clustering properties of galaxies give (tenuous) evidence for a homogeneous distribution on the largest scales surveyed, i.e. ~ 100 Mpc (Baugh & Efstathiou 1993; Peacock & Dodds 1994; Bharadwaj, Gupta & Seshadri 1999; for a completely discrepant point of view see, however, Sylos-Labini, Montuori & Pietronero 1998).

Standard cosmology takes this large-scale viewpoint as a first approximation, and treats the fine-scale structure as a perturbation on the smooth background. This working hypothesis can be disproved by observations.

The cosmological principle allows the construction of a comoving and synchronous coordinate system where:

- the cosmological fluid is always at rest relative to space coordinates (in other words the coordinates are merely labels for the world lines of the fluid);

- the time coordinate, t, is proper time along the fluid world lines;

- the geometry of the spatial hypersurfaces at $t = \text{const}$ is invariant under rotations and translations. As shown independently by Robertson and Walker, these symmetries determine the spatial metric tensor up to coordinate transformations and a global scale-factor.

With an appropriate choice of the spatial coordinates $\mathbf{x} \equiv (x_1, x_2, x_3) = (r, \theta, \phi)$, the space-time metric, known as Robertson-Walker metric, then reads

$$ds^2 = g_{\mu\nu} dx^\mu dx^\nu = c^2 dt^2 - a^2(t) \left[\frac{dr^2}{1 - kr^2} + r^2 (d\theta^2 + \sin^2 \theta \, d\phi^2) \right] . \tag{1.1}$$

Here,

1. $k = +1, 0, -1$ gives the sign of the (constant) curvature of the spatial hypersurfaces at $t =$ const;

2. the scale-factor, $a(t)$, is a multiplicative term that transforms dimensionless distances computed using coordinates on homogeneity hypersurfaces into physical ones. The quantity k/a^2 is the curvature parameter of the spatial sections at $t =$ const.

Although the demand for homogeneity and isotropy determines completely the local geometric properties of a hypersurface of homogeneity up to two adjustable factors (k and a), it leaves the global topology of the hypersurface undetermined. A very common assumption is to take the 3−spaces $t =$ const to be simply connected. In this case, the coordinate r runs from 0 to ∞ if $k = 0$ or $k = -1$, and from 0 to π if $k = +1$ (while, always, $0 \leq \theta < \pi$ and $0 \leq \phi < 2\pi$). However, there is no a priori reason to prefer simply-connected spaceforms, and different (multi-connected) topologies are possible. For $k = -1$ and $k = 0$ there exist infinite as well as finite spaceforms, while all locally spherical 3−spaces are finite. Different projects to detect potentially observable signatures of the topological structure of the universe either on the CMB (Lachièze-Rey & Luminet 1995; Bond, Pogosyan & Souradeep 1998; Cornish, Spergel & Starkman 1998a,b; Roukema & Blanloeil 1998; Weeks 1998; Scannapieco, Levin & Silk 1999) or on deep three dimensional catalogues of cosmic objects (Uzan, Lehoucq & Luminet 1999 and references therein) are now being proposed to the scientific community.

1.1.2 Friedmann models

The dynamics for the scale-factor $a(t)$ is determined by the Einstein equations of GR (see e.g. Misner, Thorne & Wheeler 1973), assuming that the stress-energy tensor of the universe can be idealized as a perfect fluid with 4-velocity $u^\mu = dx^\mu/ds$, density of mass-energy ρ and pressure p:

$$T_{\mu\nu} = -pg_{\mu\nu} + (p + \rho c^2)u_\mu u_\nu \ . \tag{1.2}$$

The system of evolutionary equations is closed (and then cosmic dynamics is completely determined) by specifying an equation of state for the cosmological fluid $p = p(\rho)$. Substituting eq. (1.2) into Einstein equations, one obtains the so called Friedmann equations

$$\frac{\ddot{a}}{a} = -\frac{4\pi G}{3}\left(\rho + \frac{3p}{c^2}\right) + \frac{\Lambda c^2}{3} \ , \tag{1.3}$$

$$\left(\frac{\dot{a}}{a}\right)^2 + \frac{kc^2}{a^2} = \frac{8\pi G}{3}\rho + \frac{\Lambda c^2}{3} \ , \tag{1.4}$$

where G is Newton's gravitational constant and Λ is the cosmological constant. [1] The time evolution of the energy density follows:

$$d(\rho c^2 a^3) = -pda^3 \ . \tag{1.5}$$

[1] Historically Λ has been introduced as a fundamental constant of nature. Note that, if $\Lambda \neq 0$, GR does not reduce to Newton's theory in the weak field slow motion limit (see section 2.1). However, the values of Λ permitted by astronomical data are small enough that the deviations from Newtonian theory would not be noticed in known systems. Therefore only observations on cosmological length scales can help constraining the value of Λ.

Following a suggestion by Zel'dovich, the Λ term in Einstein equations can be regarded as part of the gravitational sources rather than part of the geometry. The criterion of covariance implies that vacuum expectation of quantum fluctuations are described as a perfect fluid with energy density $\rho_{\rm vac} = \Lambda c^2/8\pi G$ and pressure $p_{\rm vac} = -\Lambda c^4/8\pi G$. We will discuss some interesting application of this interpretation introducing the inflationary paradigm in Section 1.3.

When energy exchanges between different cosmological components are negligible (as we will see, barring some very early phases of the life of the universe, this is always the case), eq. (1.5) can be applied to each component separately. For ultra-relativistic components ($p = \rho c^2/3$) one obtains $\rho \propto a^{-4}$, for non-relativistic matter ($p \simeq 0$) $\rho \propto a^{-3}$, and for vacuum energy density ($p = -\rho c^2$) $\rho = \text{const}$. In general, for an equation of state $p = w\rho$ with $w = \text{const}$, the energy density evolves as $\rho \propto a^{-3(1+w)}$.

In GR, the density of mass-energy determines space-time geometry. Introducing the Hubble parameter $H(t) = \dot{a}/a$, the critical density $\rho_{\rm crit}(t) = 3H^2/8\pi G$, the density parameter $\Omega(t) = \rho/\rho_{\rm crit}$ and the dimensionless cosmological constant $\lambda(t) = \Lambda c^2/3H^2$, eq. (1.4) reduces to

$$\frac{kc^2}{a^2H^2} = \Omega + \lambda - 1 \,. \tag{1.6}$$

There is then a direct correspondence between the energy density content of the universe and the curvature of its homogeneity hypersurfaces (note that, for vanishing Λ, $\rho_{\rm crit}$ is the density of a flat universe). It is convenient to parametrize the present value of the Hubble parameter through the dimensionless quantity h defined imposing $H_0 = 100h\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. In this case $\rho_{\rm crit0} = 1.88 \times 10^{-29}\, h^2\,\mathrm{g\,cm^{-3}} = 2.778 \times 10^{11}\, h^2\,\mathrm{M_\odot\,Mpc^{-3}}$. [2]

The second Friedmann equation can be also used to predict the future of the Universe. Only a particular combination of Ω and λ gives rise to a stationary universe, i.e. $a = \text{const}$ (Einstein Universe). For vanishing λ, models with $k \leq 0$ expand to infinity, while all the others collapse to a singularity after a preliminary expansion phase. In the general case, the classification of the solutions becomes more complex (see e.g. Peacock 1999). It is worth mentioning that the expansion parameter (or the corresponding curvature radius) and the "Hubble radius" $R_H = c/H$ (i.e. the distance covered by a light-ray during the characteristic expansion time of the universe) are the only characteristic scales in the model.

Expansion and contraction of the universe are expected to give strong observational signatures. Let us consider an observer and a light source in his past ligth-cone, separated by a cosmologically relevant distance and both comoving with the cosmological fluid. Because of the universal expansion (contraction), the wavelength of the photons detected by the observer is expected to be redshifted (blueshifted) with respect to its value at emission. This effect can be quantified introducing the redshift parameter, z, through

$$z \equiv \frac{\lambda_r - \lambda_e}{\lambda_e} = \frac{a(t_r)}{a(t_e)} - 1 \,, \tag{1.7}$$

where the subscripts e and r denote, respectively emission and reception times. Characteristic features in the observed spectrum of far away galaxies appear indeed redshifted with respect to local samples. Even though some alternative explanations of the effect have been proposed, this is generally considered a proof of the expansion of the universe. The recent detection of time dilation in the light curve and in the evolution of spectral features of some high-z supernovae (Leibundgut et al. 1996; Riess et al. 1997) gives a striking proof of the conventional interpretation of redshift. In fact, as a consequence of light propagation in an expanding universe, distant objects will appear to age at a slower rate than nearby objects. In total analogy with photons, the motion of freely propagating massive particles with respect to the comoving frame is damped by the universal expansion

$$\frac{v}{[1-(v/c)^2]^{1/2}} \propto \frac{1}{a(t)} \,, \tag{1.8}$$

[2] Hereafter, we adopt the common convention to denote with the subscript 0 the present value of a cosmological parameter.

with v the modulus of the 3-velocity $d\mathbf{x}/dt$. This implies that, in the absence of density fluctuations, every observer is destined to come to rest in the comoving frame.

A varying expansion parameter, affects also the apparent luminosity of cosmic objects. Let us consider a point-source of light that emits isotropically with specific luminosity $\mathcal{L}(\nu)$. It can be shown that the specific flux measured at the frequency ν_0 by an observer at $z = 0$ is

$$\mathcal{S}(\nu_0) = \frac{(1+z)\mathcal{L}[(1+z)\nu_0]}{4\pi D_L^2} , \tag{1.9}$$

where $D_L = D_L(z, \Omega_0, \lambda_0, H_0)$ is the luminosity distance (see also §5.2). Integrating over a (finite or infinite) frequency range, gives the observed flux $\mathcal{S}_0 = \mathcal{L}/4\pi D_L^2$ in terms of the luminosity $\mathcal{L}$. For sources at small distances from the Earth (i.e. $D_L \ll R_H$, $z \ll 1$), redshift and lumnosity distance are related as follows:

$$\frac{H_0 D_L}{c} = z + \frac{1}{2}(1 - q_0)z^2 + \mathcal{O}(z^3) , \tag{1.10}$$

with $q \equiv -\ddot{a}a/\dot{a}^2$ the deceleration parameter. Keeping only the first order term in the right hand side, one obtains the Hubble's law which is observationally well established and represents one of the cornerstones supporting the standard cosmological model.

It is useful to define also the angular diameter distance to a light source, $D_A = D_A(z, \Omega_0, \lambda_0, H_0)$, as the ratio between the proper diameter of the source and its observed angular extension. In full generality, $D_L = (1+z)^2 D_A$.

Friedmann models allow two kinds of solutions for $a(t)$; in the first, the universe starts contracting from arbitrarily low density and then expands after reaching a minimum radius. This requires $\Lambda > 0$ and $k = +1$. Observational limits on the value of the redshift at bounce, rule this solution out. In the second possible picture, the universe starts expanding from a singularity (the big-bang) and, if $k = +1$, eventually recollapses. However, at the extreme conditions very near the initial singularity one expects quantum effects to become important, and the predictions of classical general relativity are expected to break down (see §1.3).

Even though the universe was very small in the stages immediately following the bang, its rapid expansion prevented a causal connection between all the events on a hypersurface of homogeneity. Denoting by $t = 0$ the initial singularity, the proper distance covered by a photon at time t is

$$R_{\rm hor}(t) = a(t)\int_0^t \frac{cdt'}{a(t')} = \frac{ct}{1-n} = \frac{n}{1-n}R_H(t), \qquad a(t) \propto t^n, \quad n < 1 . \tag{1.11}$$

Friedmann models with a big bang then present particle horizons with characteristic dimensions comparable with the Hubble radius.

1.1.3 Thermal history

According to the standard model, the universe has evolved from a hot and dense early phase to its present state. In this section we will briefly describe the main stages that characterize this evolution. Available data on chemical abundances in high-z objects and the CMB provide strong evidence for this picture.

Our main aim here is just to label some precise epochs in the history of the photon-baryon interactions, since they will be important for the discussion of structure formation in the universe. For this reason, we will not give a detailed description of the physical processes characterizing the very first moments after the singularity (for this subject we refer the reader to classical cosmology textboos like Kolb & Turner 1990, Coles & Lucchin 1995, and Peacock 1999). These early phases

are described using the framework of Grand Unified Theories (GUT) developed in modern quantum field physics. In a few words, they are characterized by symmetry breaking phase transitions that give rise to a large variety of particles, most of which disappear when the temperature drops below a given threshold. Baryogenesis occurs during these phases too.

The thermal history of the universe in the standard big-bang model, is the description of the progressive cooling of the universal contents caused by the Hubble espansion. An important issue is whether a typical interaction between given particle species proceed on a rapid enough pace for thermalization to occur locally (i.e. within the particle horizon). It can be shown (e.g. Kolb & Turner 1990), that this is the case if the time scale for interaction $t_I \sim 1/n\sigma v \ll 1/H$, with n the number density of interacting particles, v their typical relative velocity and σ the corresponding cross section. Let us consider an equilibrium reaction which converts particles of type A into B and viceversa. When, as the universe expands and cools, the reaction rate becomes smaller than the expansion rate, the ratio between the A and B number densities "freezes out". For instance, when $t \sim 1$ s (corresponding to a typical energy per particle $E \sim 1$ MeV), the interactions of massless neutrinos have become sufficiently weak that they decouple from the rest of the matter. As a remnant of these stages, we expect today a thermal neutrino background with temperature $T_{\nu 0} = 1.96$ K. Its detection is far beyond present capabilities.

When the typical energy per particle becomes lower than 0.5 MeV, the production rate of electron-positron pairs drops below the annihilation rate, and shortly after this time all the positrons disappear, leaving a relatively small population of residual electrons. This is the beginning of the so-called "radiative era", when the energy content of the universe is dominated by the photonic component. At this stage, the universe is composed by a mixture of non-relativistic matter (protons, neutrons, electrons and, maybe, some particles left over from the previous universal phases) and radiative components (photons and non interacting massless neutrinos). Ordinary matter is fully ionised and the photons are tightly coupled to the baryons and electrons. During this epoch simple atomic nuclei are formed (see §1.2.1).

It is customary to distinguish four important moments in the following thermal evolution of the universe.

- **Matter-radiation equality** ($t_{\rm eq} \sim 1.4 \times 10^3 (\Omega_0 h^2)^{-2}$ yr, $z_{\rm eq} \sim 2.3 \times 10^4\, \Omega_0 h^2$)
  This is the epoch when the energy density of non-relativistic matter comes to dominate over its relativistic (mainly photonic) counterpart. Bound-free and free-free interactions still couple the baryons to the radiation field.

- **(Re)-Combination** ($t_{\rm rec} \sim 1.4 \times 10^5 (\Omega_0 h^2)^{-1/2}$ yr, $z_{\rm rec} \sim 1300$)
  The decreasing matter-radiation temperature changes the ionisation equilibrium: matter starts becoming neutral, greatly reducing its coupling with photons.

- **Decoupling: the last scattering surface** ($t_{\rm dec} \sim 1.8 \times 10^5 (\Omega_0 h^2)^{-1/2}$ yr, $z_{\rm dec} \sim 1100$) At this stage interactions between matter and radiation become so rare that a typical photon we observe today will not have scattered since then (the mean free path increases rapidly, making this a rather sharply defined time in cosmic history). The photons decouple completely from the baryons and subsequently cool with the expansion of the universe. The present universe should then be filled with a low temperature blackbody radiation whose photons last interacted with matter at the decoupling time. Such a radiation background at temperature $T \sim 2.73$ K was discovered by Penzias & Wilson (1965).

  Before decoupling, the pressure provided by the radiation inhibited the gravitational growth of density perturbations. After $t_{\rm dec}$, the radiation pressure has no effect on the matter and

gravitational instability becomes an effective way to form the large scale structure we observe today.

- **Reionisation** ($z_{\rm ion} \sim 5 - 50$) The absence of Compton y-distortion in the cosmic microwave background indicates that the intergalactic medium was neutral before redshift $z \sim 50$. The lack of Gunn-Peterson absorption along the lines-of-sight to distant quasars shows that the diffuse helium and hydrogen components of the IGM must have been highly ionised by $z \sim 2.5$ and $z \sim 5$ respectively. the inter-galactic medium (IGM) was reionised. Several sources of radiation could have contributed. If these were continuously and uniformly distributed, like radiatively decaying particles or an abundant population of early star clumps, the transition from a neutral to a ionised IGM is expected to have been homogeneous. Alternatively, very luminous and widely separated sources, like quasars, would have generated expanding HII regions in the neutral medium. In this case, the epoch of reionisation would have been determined by the percolation of the ionised phase (e.g. Meiksin & Madau 1993). It is possible that this happened sufficiently early to move the last scattering surface forward in time. Thus, new features in the cosmic microwave background anisotropies, potentially observable by MAP and PLANCK, could have been imprinted at this epoch. Whether or not this happened depends on details of star formation and evolution which are, at present, poorly understood. It is probably through the next generation of CMB experiments that we will learn the most about the ionisation history.

1.2 Do observations support the standard model?

This section deals with astronomical observations. We want to check if present data consistently support the standard cosmological model from the time of nucleosynthesis to the present day. In asking this question, we are concerned with the theoretical concepts introduced in the previous section. In fact, Robertson-Walker metric and Friedmann equations (which characterize the so called Friedmann–Robertson–Walker, hereafter FRW, models) represent an unavoidable framework to discuss the universe in its globality. The standard method in cosmology is to parametrize the present status of a FRW model in terms of a few fundamental quantities (H_0, λ_0, $\Omega_0 = \sum_i \Omega_i$ where the index i denotes the present fraction of the critical density contributed by the different cosmic components). Observations are then used to constrain these parameters. A consistency check is then performed by looking for a region of parameter space in agreement with all the observational constraints (e.g. Krauss & Turner 1995; Ostriker & Steinhardt 1995; Wang et al. 1999).

1.2.1 The primordial nucleosynthesis

In the standard hot big bang model, nuclear reactions are expected to produce light-elements a few seconds after the bang. The outcome of this primordial nucleosynthesis can be computed by solving a system of equations for the nuclear reaction rates in an expanding environment filled by matter and radiation. Since all input microphysics is well determined (cross-sections at the relevant energies, 0.1 - 1 MeV, are commonly determined in nuclear-physics laboratories), the numerical results show very small uncertainties. Irrespective of the baryon density, Ω_b, about 24% of the baryonic mass is expected to be converted into ^{4}He, while protons constitute almost all the remaining fraction. Other light isotopes (i.e. D, ^{3}He and ^{7}Li) are produced in much smaller amounts, and their final abundance strongly depends on $\Omega_b h^2$. The great success of the big bang nucleosynthesis model is that it can account for the observed primordial abundances (relative to hydrogen) of all four light elements (for a recent review see Schramm & Turner 1998) by tuning a single free parameter,

i.e. the product $\Omega_b h^2$. Recent technological developments (e.g. the construction of the Keck telescopes and of their HiRes spectrograph) opened an high precision era in constraining $\Omega_b h^2$. In fact, Burles & Tytler (1998a,b) recently measured the deuterium abundance in high-z hydrogen clouds, obtaining D/H = $[3.4 \pm 0.3\,(\mathrm{stat})] \times 10^{-5}$. If this value reflects the primordial abundance, then nucleosynthesis models with 3 light neutrino families, at 95% confidence level give (Burles et al. 1999)

$$\Omega_b h^2 = 0.0190 \pm 0.0024\ . \tag{1.12}$$

Note that the main uncertainty originates in estimating how much the observed high-z abundances have been contaminated by nuclear activity in stars.

In brief, the agreement between data and models on primordial chemical abundances successfully tests the standard cosmological models at early times (roughly from 0.01 to 200 seconds after the singularity) and allows a very accurate determination of the density of ordinary matter in the universe.

1.2.2 The Cosmic Microwave Background

The CMB has been serendipitously discovered by Penzias and Wilson in 1965. The Far Infrared Absolute Spectrophotometer (FIRAS) instrument onboard the COsmic Background Explorer (COBE) satellite has recently confirmed that the CMB is constituted by thermally distributed photons with $T = 2.728 \pm 0.004$ K (95 % confidence level; Fixsen et al. 1996). The temperature of the CMB is very uniform across the sky but three forms of temperature anisotropies have now been detected (see e.g. Smoot 1999 for a review).

- A dipole anisotropy of about a part in 10^3, generally believed to be due to the peculiar motion of our galaxy relative to the cosmic rest frame, at a speed of $369.0 \pm 2.5\,\mathrm{km\,s^{-1}}$ (e.g. Lineweaver et al. 1996).

- An annual modulation of the temperature in a given direction of the sky of about a part in 10^4, due to the orbital motion of the Earth at $30\,\mathrm{km\,s^{-1}}$ (e.g. Smoot 1999).

- The temperature anisotropies of about a part in 10^5 detected for the first time by the Differential Microwave Radiometer (DMR) onboard COBE, corresponding to a rms quadrupole amplitude $(\Delta T)_{Q_{\mathrm{rms}}} = 16 \pm 12\mu\mathrm{K}$ at the 95 % confidence level (e.g. Smoot 1999).

The most widely accepted explanation for this background is that these are "frozen out" photons which were once in equilibrium with baryonic matter. In fact the cosmic expansion preserves a thermal distribution even after the freeze out with $T \propto 1/a$. Three principal arguments can be given against a non-cosmological origin of the CMB:

- Its energy density $\rho_\gamma = (\pi^2/15)(k_{\mathrm{b}}^4/\hbar^3 c^3)T^4 \simeq 5 \times 10^{-34}\,\mathrm{g\,cm^{-3}}$ (or $\Omega_\gamma h^2 \simeq 3 \times 10^{-5}$) represents some five times the total electromagnetic energy density emitted by the observed galaxies but appearing at microwave wavelengths. Energetically, this would be extremely difficult to explain from local sources.

- The high degree of isotropy and the absence of relevant spectral distortions over a broad frequency range (FIRAS was built to measure the spectrum of diffuse emission in the wavelength interval ranging from 0.01 cm to 1 cm) are hard to reproduce using local sources;

- Spectral distortions of the CMB in the direction of the cluster of galaxies CL0016+16 at $z = 0.5455$ have been observed. The measured spectral features are characteristic of a cold black body radiation that passes through a hot electronic plasma (Sunyaev-Zel'dovich effect, hereafter SZ effect, see e.g. Sunyaev & Zel'dovich 1972; Birkinshaw 1999), thus indicating that the CMB originates at $z > 0.55$.

The COBE detection of CMB anisotropies, confirmed by a series of balloon experiments, then provides strong evidence for primeval density inhomogeneities of amplitude $\sim 10^{-5}$.

1.2.3 The Hubble law

The Hubble law is the natural outcome of an universal expansion that preserves homogeneity. However the real universe is inhomogeneous on galaxy scales and peculiar motions of galaxies in groups, clusters and in the field are expected on top of the isotropic expansion. The typical magnitude of these peculiar velocities is $v \sim 100 - 1000\,\mathrm{km\,s^{-1}}$ corresponding to a Doppler (red or blue) shift $\Delta\lambda/\lambda = v/c \sim 0.0003 - 0.003$. Objects at $z \gg v/c$ ($D_L \gg 1 - 10\,h^{-1}$ Mpc) are therefore needed to minimize the impact of peculiar motions on measured redshifts. In other words, to test eq. (1.10) and obtain an unbiased value for H_0 it is important to analyse a "fair-sample" of the universe, i.e. the test particles that we use to measure the Hubble flow should lie over regions larger than the scale of homogeneity.

Classical determinations of H_0 rely upon the cosmic distance ladder [3] and are therefore subject to the uncertainties inherent in the calibration of each step in the ladder. At $D_L \gg 10$ Mpc, however, we do not dispose of a primary distance indicators (Cepheids work to ~ 40 Mpc), and we must rely on secondary indicators (like Type Ia SNa explosions, planetary nebula or globular cluster brigthness distribution, brightest supergiant stars). The ongoing *Hubble Space Telescope* (*HST*) *Key Project on the Extragalactic distance scale* (Kennicutt, Freedman & Mould 1995) aims at measuring the Hubble constant to 10% accuracy by obtaining distances based on Cepheid variables stars to eighteen different galaxies which will provide a firm basis for the calibration of several secondary distance indicators. The most recent compilation of results (combining data out to a distance in excess of 500 Mpc) gives $H_0 = 72 \pm 5$(statistical) $\pm$ 7(systematic) $\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Madore et al. 1999). Also galaxy dynamics (Tully-Fisher and Faber-Jackson relations respectively for spiral and elliptical galaxies) can be used to establish the cosmic distance scale. Tully (1999) recently obtained $H_0 = 77 \pm 4$ (stat) $\pm$ 8 (syst) $\mathrm{km\,s^{-1}\,Mpc^{-1}}$, while Giovanelli et al. (1997) give $69 \pm 5\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. Both studies consider data within $cz \sim 9000\,\mathrm{km\,s^{-1}}$.

Supernovae Ia are considered extremely good standard candles (in the sense that the shape of their ligth curve is strongly correlated with their absolute luminosity) and are observable to very large distances. In a series of papers, Saha et al. (1997 and references therein) used the Cepheid distance of different supernovae to calibrate the Hubble diagram (i.e. the D_L vs. z relationship) for Type Ia SNe at $v > 10,000\,\mathrm{km\,s^{-1}}$, obtaining $H_0 = 58^{+7}_{-8}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$. On the other hand, combining data relative to the 5 Ia SNe for which we have both Cepheid distance and reliable light curves, Suntzeff et al. (1999) find $H_0 = 64.0 \pm 2.2$ (stat) $\pm$ 3.5 (syst).

Besides classical determinations of H_0 based on the Hubble diagram, new "direct" methods have been recently developed. They do not rely on the distance ladder, thus providing an extremely

[3] The distance ladder is an iterative algorithm to measure the distance of cosmic objects. The fundamental operation to iterate consists in using certain techniques over short distance scales to calibrate other methods to adopt over longer length scales. Each rung in the ladder is based on the selection of a particular class of cosmic objects as "distance indicator". Indicators that are directly calibrated by high-precision local measurements (e.g. via parallax) are called primary and are used to fix the zero-point for secondary indicators.

useful consistency check for it. However, they often require some a priori modelization of the physics therefore including systematic errors to be understood and allowed for.

Measurement of the time delay between brightness variations of multiple images of a gravitationally lensed object can be used to derive an estimate for the Hubble constant (Refsdal 1964). In this case no distance has to be measured, while knowledge of both source and lens redshifts, respectively z_s and z_l, in conjuction with modelling the mass distribution of the lensing object, is required. At present, the following systems has been analysed.

- The double quasar B0957+561 at $z_s = 1.41$ lensed by an elliptical galaxy at $z_l = 0.36$ which resides in a cluster of galaxies that also contribute to the lensing. Assuming $\Omega_0 = 1$, Kundić et al. (1997a; see also Grogin & Narayan 1996) derive $H_0 = 64 \pm 13\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$; $\Omega_0 = 0.1$ would increase the H_0 estimate by 7% while $\Omega_0 = 0.25$ with $\lambda_0 = 0.75$ by 4%.

- The quadruply imaged quasar PG 1115+080 at $z_s = 1.722$, lensed by the combined action of an elliptical galaxy and a small group at $z_l = 0.310$. The quasar host galaxy is lensed into an Einstein ring connecting the four quasar images. Regrettably, degeneracies in the allowed models for the mass distribution of the lens resulted in a large spread for H_0 estimates (Schechter et al. 1997; Keeton & Kochanek 1997). More recent studies give:

  - $H_0 = 53^{+10}_{-7}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Courbin et al. 1997)
  - $H_0 = 52 \pm 14\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Kundič et al. 1997b),
  - $H_0 = 44 \pm 8\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Impey et al. 1998, assuming a flat rotation curve for the lens), $H_0 = 65 \pm 10\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Impey et al. 1998, assuming a constant mass to ligth ratio for the lens),
  - $H_0 = 77^{+29}_{-24}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ to 95 % confidence level (Bernstein & Fischer 1999).

- The radio source PKS 1830-211 which forms an Einstein ring through the probable lensing action of two intervening galaxies at $z_{l1} = 0.89$ and $z_{l2} = 0.19$. The recent identification of the double imaged core with a quasar at $z_s = 2.507$ gives (Lidman et al. 1999; see also Lovell et al. 1998) $H_0 = 65^{+18}_{-9}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (assuming $\Omega_0 = 1$) and $H_0 = 76^{+18}_{-10}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (assuming $\Omega_0 = 0.3$).

- The lens system B0218+357 is represented by an Einstein ring and two images of a compact radio source at $z_s = 0.96$. The deflecting mass is probably a single galaxy at $z_l = 0.6847$ implying $H_0 = 69^{+18}_{-19}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (Biggs et al. 1999).

Data regarding the quadruple image system B1608+656, with a background galaxy at $z = 1.39$ lensed by a galaxy at $z = 0.63$, are now under analysis (Fassnacht et al. in preparation; Koopmans & Fassnacht in preparation).

The SZ effect produces a characteristic signature in the CMB that can be used to measure the Hubble constant. In fact, analyses of the SZ signal, combined with $X-$ray observations of the cluster that generates it, permit to estimate the linear core radius of the cluster that, combined with the observed angular core radius, yields the angular diameter distance (see e.g. Cavaliere, Danese & De Zotti 1979). Combining the results for 4 clusters at $z \leq 0.09$, Myers et al. (1997) obtain $H_0 = 54 \pm 14\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ (assuming an Einstein–de Sitter cosmology; for low redshift sources, different values of Ω_0 and λ can alter the H_0 estimates by $\Delta H_0/H_0 \sim \Delta q_0 z/2 \sim 2.5\%$). Recent studies of Abell 1413 at $z = 0.143$ (Grainge et al. 1999) and Abell 773 at $z = 0.217$ (Saunders et al. 1999) gave, respectively, $H_0 = 57^{+23}_{-16}\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ and $H_0 = 46^{+14}_{-11}$ (both for $\Omega_0 = 1$ and $\lambda = 0$). This method requires an a priori modelization of the cluster density profile and of the temperature

of the electron gas, which introduces some systematic errors. In particular, substructure in cluster atmospheres, elongation of clusters along the line of sight and the presence of cooling cores can bias H_0 estimates.

It is encouraging that estimates coming from such a plethora of techniques agree to within 20%. However, measures of H_0 based on high-z sources (i.e SNe and lensing) seem to indicate a slightly lower value than obtained from local surveys. This could be related to the fact that local analyses are not deep enough to reach the scale of universal homogeneity. For example, from studies of the peculiar velocity field of 44 SNe Ia at distances between 20 and 300 Mpc, Zehavi et al. (1999) find (weak) evidence for a local underdense bubble ($\delta\rho/\rho = -3\delta H_0/H_0 \sim -0.2$ inside a sphere of $70\,h^{-1}\,$Mpc). This suggests that the Hubble constant as determined within $70\,h^{-1}\,$Mpc could be overestimated by $\sim 6\%$. On the other hand, analysing the motion of 76 clusters, Giovanelli (1998) finds that H_0 does not exhibits significant changes from $35\,h^{-1}$Mpc to $200\,h^{-1}$Mpc.

In summary, different methods for measuring H_0 has not yet converged but for the first time, after years of violent debate, most of the results are consistent within the quoted errors. A conservative estimate emerging from this collection of data is

$$H_0 = 65 \pm 15\,\mathrm{km\,s^{-1}\,Mpc^{-1}} \tag{1.13}$$

at the 2σ level (Wang et al. 1999).

1.2.4 Is there a standard model?

In this section we briefly review current knowledge about the values of the cosmological parameters and show that there exists a coherent description of the observed universe in the framework of the standard big-bang model.

The age of the universe

The expansion age of the universe (the extrapolated time back to the initial singularity), t_0, must necessarily be greater than the age of any object within it. A lower limit for t_0 can be determined combining observations of astronomical objects with models for their evolution (for recent reviews see Chaboyer 1998 and Jimenez 1998). Ages of old galactic globular clusters are estimated to be 11.5 ± 1.3 Gyr (Chaboyer et al. 1998), implying a minimum age of the universe of $t_0 \geq 9.5$ Gyr (95 % confidence level). It is also likely that an additional Gyr or so elapsed before the formation of their stars. A different estimate by Jimenez (1998) gives $10.5\,\mathrm{Gyr} < t_0 < 16\,\mathrm{Gyr}$ with 99 % confidence. Note that nucleochronology and studies of the cooling of white dwarfs stars in the disk of the Galaxy lead to the less robust estimates of 15.2 ± 3.7 Gyr (Sneden et al. 1996) and 9.3 ± 2 Gyr (Oswalt et al. 1996) respectively. Considerably improved constraint from cooling curves are expected when the advanced camera becomes operational on the *HST* (scheduled to occur in the year 2000).

From the theoretical point of view t_0 can be computed as a function of H_0, Ω_0 and λ_0. Assuming $\lambda_0 = 0$ and using our best estimate for H_0 given in eq. (1.13), one obtains $t_0 = 2/3H_0 = 10 \pm 2$ Gyr for $\Omega_0 = 1$ and $t_0 = 1/H_0 = 15 \pm 3$ Gyr for $\Omega_0 = 0$ at the 2σ level. The effect of a positive cosmological constant is to increase t_0 for a given Ω_0 and H_0. In this case an approximated formula for t_0 (precise to 5 %) has been given by Carroll et al. (1992):

$$t_0 \simeq \frac{2}{3H_0}\frac{\sinh^{-1}\left[(|1-\Omega_a|/\Omega_a)^{1/2}\right]}{|1-\Omega_a|^{1/2}}\,, \tag{1.14}$$

where $\Omega_a \equiv \Omega_0 - 0.3(\Omega_0 + \lambda_0) + 0.3$ and $\sinh^{-1}$ must replace $\sin^{-1}$ if $\Omega_a > 1$.

Comparing observational data to theoretical results, we can conclude that even though, in the past, an age problem has been invoked to reconcile globular cluster and expansion ages of Friedmann models with $\lambda_0 = 0$, recent data nicely fit into a large set of cosmological models. Anyway, low-density models (with positive or vanishing λ_0) are slightly favoured on statistical grounds. On the other hand, ages of high-z objects can give stronger constraints on cosmological parameters. For example, using stellar population synthesis models to fit the spectral energy distribution of a red elliptical galaxy at $z = 1.43$, Jimenez (1998; see also Dunlop et al. 1996) finds a minimum age for the galaxy of 3.2 Gyr (corresponding to $t_0 = 13 \pm 2$ Gyr, in agreement with globular cluster determinations). This high-redshift constraint strongly reduce the allowed volume in the H_0, Ω_0, λ_0 space. For example a flat, matter dominated universe is ruled out unless $H_0 \leq 45\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$ while $\lambda_0 \geq 0.4$ is required for $H_0 > 60\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

Dark matter

Let us consider now the mass density of the universe. Summing up all the contributions of the astronomical sources we can detect through their electromagnetic emission (stars and diffuse gas), one obtains $\Omega_\star \sim 0.003$ (Persic & Salucci 1992). On the other hand, combining the most recent nucleosynthesis constraints (1.12) and the most probable range for H_0 (1.13), one obtains $0.026 \leq \Omega_b \leq 0.086$. Therefore, trusting the standard model for nucleosynthesis, one has to conclude that most of the baryons are dark (i.e. in the form of brown or white dwarfs, neutron stars, Jupiters or black holes) or, at least, not detectable by present instruments (i.e. low surface brightness galaxies or a diffuse ionised background).

At present, there is indeed a lot of observational evidence for the presence of dark matter:

- Rotation curves in spiral galaxies keep flat out to scales much greater (by a factor 5 or more) than the typical length of their luminosity distribution (e.g. Rubin, Ford & Thonnard 1980; Persic & Salucci 1995).

- X-ray observations reveal extended halos of hot gas around elliptical galaxies (Forman, Jones & Tucker 1985; Fabian et al. 1986; Matsushita et al. 1998). It is widely believed that the halos are created through gas heating and ejection by supernovae. A large-scale halo of dark matter is then needed to gravitationally retain the gas inside the galaxies.

- Satellite dynamics in galactic systems (e.g. Zaritsky et al. 1997a,b) suggests that galaxies are surrounded by very large dark halos (having characteristic radii $\gtrsim 200$ kpc) with very low matter density, $\rho_{\mathrm{DM}} \sim 6 \times 10^{-5} \mathrm{M_\odot\,pc^{-3}}$, with respect to the stellar component $\rho_\star \sim 0.3\,\mathrm{M_\odot\,pc^{-3}}$. The total mass of the dark component comes out to be roughly an order of magnitude larger than that of the luminous matter.

- Gravitational lensing also offers a way to directly probe the mass distributions of distant galaxies, and to compare them with the corresponding light profiles. Recent results (e.g. Maoz & Rix 1993; Kochanek 1995, 1996; Grogin & Narayan 1996) suggest that lens galaxies require dark (roughly isothermal) halos, in agreement with results from stellar dynamics and X-ray observations.

- The matter content of clusters of galaxies has been largely investigated through numerous different techniques. Recent results from dynamics, gravitational lensing and X-ray analyses generally indicate very large mass to light ratios (e.g. Lewis et al. 1999 and references therein).

Several attempts have been done to measure Ω_0, but none of them can be considered totally convincing. On the one hand, it has not yet been possible to analyse a fair sample of the universe

down to the scale of homogeneity (if any). On the other hand, important assumptions always underlie the estimate of Ω_0. Recently, X-ray measurements of gas in clusters of galaxies have been used (together with some theoretical modelling) to estimate the baryon mass fraction (White et al. 1993; White & Fabian 1995; Mohr, Mathiensen & Evrard 1999; Lewis et al. 1999). Results generally are compatible within a small factor of order unity even though they do not agree within the quoted errorbars. For example Mohr et al. (1999) find $f_{\rm gas} = \Omega_{\rm gas}/\Omega_0 = (0.075 \pm 0.002)h^{-3/2}$ to 90 % confidence level while Lewis et al. (1999) obtain $f_{\rm gas} = (0.047 \pm 0.002)h^{-3/2}$. From the measurement of the SZ effect, Myers et al. (1997) independently give $f_b = (0.061 \pm 0.011)h^{-1}$. One can then take $f_b \sim (0.060 \pm 0.015)h^{-3/2}$ as a rough estimate and assume that the $f_{\rm gas}$ found in galaxy clusters is representative of the whole universe. Alternatively one can rely on gas-dynamical simulations showing that, on average, the X-ray method overestimate f_b by $\sim 12\%$ (Mohr et al. 1999; note however that Lubin et al. 1996 claim that in hydro simulations the gas fraction in clusters is $\sim 20\%$ lower than the cosmological value). Combining these data with the Ω_b estimate coming from primordial nucleosynthesis, one obtains

$$0.2 \lesssim \Omega_0 \lesssim 0.6 \, . \tag{1.15}$$

Large-scale data on galaxy clustering consistently suggest $\Omega_0 h \simeq 0.25$ (Peacock & Dodds 1994), but this result depends on crucial assumptions about the unknown biasing between galaxy and mass distributions as a function of scale (see chapters 5 and 6). Similar problems flaw large-scale velocity field studies (e.g. Dekel 1994; Strauss & Willick 1995), that are only able to determine the ratio $\beta = \Omega_0^{0.6}/b$. The "linear bias parameter" b (probably of order unity) quantifies the different statistical properties of the galaxy and mass density fields both smoothed on a suitably large scale (see chapter 5). It is reasonable to think that the degeneracy between b and Ω_0 will soon be resolved by combining velocity studies with high quality data coming from the next generation of galaxy surveys (e.g. Fry 1994; Matarrese, Verde & Heavens 1997). Results are still controversial, Willick & Strauss (1998) find $\beta = 0.50 \pm 0.04$ (one sigma error) at $300\,{\rm km\,s^{-1}}$ smoothing of the velocity field traced by the galaxies detected with the Infrared Astronomical Satellite (IRAS), while Sigad et al. (1998) obtain $\beta = 0.89 \pm 0.12$ within a volume of effective radius $40\,h^{-1}$Mpc, where the biasing parameter refers to the IRAS galaxy distribution smoothed at $12\,h^{-1}$Mpc. There is, however, general agreement that the data imply $\Omega_0 \gtrsim 0.3$ (Dekel 1994; Dekel & Rees 1994; Sigad et al. 1998).

These results imply that the dominant massive component of the universe has to be non-baryonic. A widely accepted possibility is a "sea" of relic, stable particles, not coupled to the electromagnetic field, left over from the earliest phases of the universe. Theoretical particle physics is able to provide at least some viable candidates: a light massive neutrino (with mass ~ 10 eV), a very light axion (with mass $\sim 10^{-5}$ eV) and the lightest supersymmetric particle (probably the neutralino with mass in the range 30 GeV - 500 GeV). Data on CMB anisotropy and models of structure formation in the universe give further support for the non-baryonic origin of dark matter.

Evidence for a non vanishing cosmological constant

Programs are now underway to determine the distance-redshift relation from SNe Ia which are supposed to be less vulnerable to evolutionary effects than galaxies. The key assumption is that the maximum luminosity of these supernovae does not change as a function of cosmic age. This is, in fact, sufficient to determine their distance independently from their redshift. Two groups, the High-z Supernova Search Team (Riess et al. 1998; Schmidt et al. 1998) and the Supernova Cosmology Project (Perlmutter et al. 1999), independently found evidence for accelerated expansion of the universe. Since, in Friedmann models, $q_0 = \Omega_0/2 - \lambda_0$, these results are fully consistent with

the existence of a non vanishing cosmological constant. For example, Perlmutter et al. (1999), analysing 42 SNe at redshifts between 0.18 and 0.83, find that the cosmological constant is non-zero and positive with a confidence level of 99% (including the identified systematic uncertainties). Their result can be synthesized through the relation

$$0.8\,\Omega_0 - 0.6\,\lambda_0 \simeq -0.2 \pm 0.1\ . \tag{1.16}$$

For a flat cosmology the data give $\Omega_0^{\rm flat} = 0.28^{+0.09}_{-0.08}({\rm stat})^{+0.05}_{-0.04}({\rm identified\ syst})$, corresponding to an age of $t_0^{\rm flat} = 14.4^{+1.4}_{-1.1}\,(0.65/h)$ Gyr. Riess et al. (1998) find consistent results from a smaller sample of SNe.

The statistics of gravitationally lensed quasars and radio sources represents another (potentially powerful) test to constrain the value of λ_0. Kochanek (1993; 1996) and Maoz & Rix (1993) showed that the observed statistics of lensed sources (numbers, redshifts, magnitudes and angular separations) is consistent with the hypothesis that lensing is caused by the observed population of $E/S0$ galaxies, assuming that their comoving number density and internal structure do not evolve between $0 \leq z \lesssim 1$ (see e.g. Lilly et al. 1995 and Keeton, Kochanek & Falco 1998 for some observational tests of these assumptions). In this case, the number of lensed systems detected in a survey can be used to constrain the comoving volume of our past light-cone as a function of z. Fitting the data on lensing can then put stringent constraints on the cosmological parameters. Unfortunately, present-day results, probably affected by unidentified systematic errors, are controversial. For example, the 2σ limits on the cosmological constant in flat models deriving from the radio, optical, and joint samples analysed by Falco, Kochanek & Muñoz (1998, see also Kochanek 1996) are $\lambda_0^{\rm flat} < 0.73, 0.69,$ and 0.62, respectively. Assuming $\lambda_0 = 0$ one instead obtains $\Omega_0 > 0.24, 0.30,$ and 0.51, respectively (Falco, Kochanek & Muñoz 1998). Known systematics regard the role of spirals, dust extinction and mass modelling. On the other hand, the statistical uncertainties are dominated by the Poisson errors from the small number of lensed sources and the uncertainties in the determination of the local galaxy luminosity functions for different Hubble types. The second point is emphasized by Chiba & Yoshi (1999) who, combinig recent estimates of the type specific luminosity functions with the lensed quasar sample used by Kochanek (1996), find, in agreement with SN Ia results, that a flat universe with $\Omega_0^{\rm flat} = 0.3^{+0.2}_{-0.1}$ gives the best fit. The likelihood for a universe with $\lambda_0 = 0$, always smaller than in the flat case, monotonically increases with decreasing Ω_0.

The study of CMB anisotropies is nowadays considered the most promising tool for obtaining a precise estimate of the cosmological parameters. In particular, the position of the first acoustic peak in the CMB power spectrum is sensitive to the combination $\Omega_0 + \lambda_0$. Recent analyses considering ~ 35 CMB data points obtained at different angular resolutions, while not conclusive, indicate that $0.5 \lesssim \Omega_0 + \lambda_0 \lesssim 1.2$ (Lineweaver 1998; see also Tegmark 1999). The intersection of these results with a series of other observational constraints including the SN Hubble diagram gives $\lambda_0 = 0.62 \pm 0.16$ and $\Omega_0 = 0.24 \pm 0.10$ (Lineweaver 1998). This complementarity between the CMB and other observations rules out $\lambda_0 = 0$ models at more than the 99% confidence level. [4]

[4]The observational data we reported suggest that the universe is presently dominated by an energy component with effective negative pressure. Non-vanishing vacuum energy (or a cosmological constant) does not represent the only viable candidate. For instance, another possibility is given by quintessence: the energy associated with a slowly evolving scalar field with positive potential energy, similar to the inflaton field in inflationary cosmology (see §1.3). Unlike a cosmological constant, a dynamically evolving field can support long-wavelength fluctuations which leave an inprint on the CMB and large-scale structure. A prime motivation for considering quintessence models is to address the "coincidence" (or "why now?") problem: since the energy density associated with the negative pressure component and the matter density decrease at different rates, it appears that their ratio must be set to a specific, infinitesimal value in the very early universe in order for the two densities to nearly coincide today, some 15 billion

Summary

In the last few years, observational cosmology registered a period of rapid growth thanks to the construction of dedicated instruments:

- *HST* and the Keck telescopes opened new windows on the high redshift universe;
- COBE, with the first detection of CMB anisotropies, offered a totally new research field to the scientific community.

It is then reassuring that the parameters of the standard cosmological scenario can be adjusted to accommodate all observations, singling out a preferred "concordance" model where $\Omega_0 \sim 0.35$, $\lambda_0 \sim 0.60$, $\Omega_b \sim 0.05$, $h = 0.65$ (e.g. Krauss & Turner 1995; Ostriker & Steinhardt 1995; Wang et al. 1999). The current constraints must be, however, considered as work in progress rather than final. The tremendous amount of cosmologically relevant data that will become available in the coming decade [5] is expected to revolutionize the field, opening the era of high-precision cosmology. For example, we are confident that the forthcoming satellite missions Microwave Anisotropy Probe (MAP; expected to be launched by NASA during late 2000) and Planck surveyor (scheduled by ESA for 2007) will be able to determine a host of cosmological parameters at the unprecedented accuracy of a few percent (Jungmann et al. 1996; Bond, Efstathiou & Tegmark 1997; Zaldarriaga, Spergel, & Seljak 1997).

1.3 The inflationary paradigm

The hot big bang model provides a reliable physical description of the universe from about 10^{-2} seconds after the intial singularity onward. However it leaves open a number of important questions related to the initial conditions.

- **The flatness problem**
  Astronomical observations reveal that Ω_0 is of order unity. Using the Friedmann equations to trace the value of Ω back in time, one gets results closer and closer to one as $t \to 0$. For example, when nucleosynthesis starts, $|\Omega(t_{\rm nuc}) - 1| \lesssim 10^{-16}$. Such finely-tuned initial conditions seem extremely unlikely. In fact, one expects the typical curvature radius at early times to roughly coincide with the fundamental scale of gravity, i.e. the Planck length $L_{\rm Pl} = (G\hbar/c^3)^{1/2} \simeq 1.7 \times 10^{-33}$ cm. If this would be the case, the universe would either immediately recollapse, or expand and cool below 2.7 K in a fraction of a second. Then, why is our universe so flat?

- **The horizon problem**
  Even though, in all realistic Friedmann models, the particle horizon at decoupling subtends about a degree on the celestial sphere, we detect the same CMB temperature in the whole sky. How is this possible? Following Andrei Linde, we can give an even more dramatic formulation of this problem: why did all the causally disconnected regions of the universe start their expansion simultaneously at $t = 0$?

years later. In fact, fine tuning is avoided in those quintessence models which contain a "tracker field" whose evolution rapidly approach a common evolutionary track starting from a wide range of initial conditions (see Zlatek, Wang & Steinhardt 1999 for a nice review).

[5] For information regarding ongoing and forthcoming ESA and NASA missions ckeck the web page *http://universe.gsfc.nasa.gov/new/missions.html*

- **The problem of cosmic relics**
  Modern particle physics predicts that peculiar forms of energy have to be produced by phase transitions at some early epochs and can survive till to the present. These cosmic relics can be either particles (like gravitinos) or stable topological defects (like magnetic monopoles, cosmic strings, textures and domain walls). Generally, their presence at late times would totally upset the hot big bang picture. Sometimes their influence is less dramatic but it is anyway expected to lead to strong observational signatures that have not been detected. In both cases, a mechanism to get rid of the unwanted cosmic relics would be required.

- **The fluctuation problem**
  Why is the universe so homogeneous on large length scales? What is the origin of its small scale clumpiness?

The inflationary scenario, mainly developed by Guth, Linde and Steinhardt (see e.g. Linde 1990 for a historical perspective), offers answers to these questions. The idea is that, at some time $t_{\rm i}$ after the Planck epoch, a causally connected patch of the universe, undergoing a finite phase of accelerated expansion (called inflation) driven by the zero-point energy of a homogeneous scalar field, gives rise to a region that nowadays encompasses the present Hubble radius. During the inflationary stage Ω is driven towards one, and the flatness problem is solved if the ratio $a(t_{\rm f})/a(t_{\rm i}) \gtrsim {\rm e}^{70}$ ($t_{\rm f}$ is the time at which inflation ends). This automatically solves also the horizon and relics problems; in fact, in the inflationary scenario, i) all the presently observable universe derives from a casually connected region, and ii) the density of the unwanted relics is extremely diluted by the enormous amount of expansion, provided they have been formed before the inflationary phase. In principle, barring the issue about relics, inflation represents only a temporary fix: it does not guarantee that in the very distant future some monsters (i.e. wild inhomogeneities) will enter our past light-cone.

Many different versions of inflation (with the most incredible fancy names) have been proposed, and it is obviously beyond the scope of this thesis to review them all. We then refer the interested reader to specialistic textbooks (e.g. Linde 1990; Kolb & Turner 1990) and reviews (e.g. Liddle & Lyth 1993; Turner 1993; Liddle 1999). At present no inflationary model is universally preferred. This is not surprising since inflation involves very speculative theories beyond the standard model of particle physics. However, even though there are many details that are model-dependent, it is still possible to give a synthetic description of the inflationary paradigm. For this purpose, we briefly list here the main characteristics of one of the simplest models (Linde 1983).

- The universe emerging from the Planck era is expectd to be chaotic, i.e. divided into many, causally disconnected domains with typical dimension $\gtrsim 10^{-32}$ cm. [6]

- Let us imagine that the energy content of one of these domains is dominated by the zero-point contribution of a scalar field ψ. It can be shown (see e.g. Linde 1990) that spatial gradients of ψ do not generally play any crucial role since, after a short transient phase, they are quickly erased. Therefore, we can consider that ψ is nearly homogeneous in our domain of space-time, which can then be described by the Robertson-Walker metric. We further assume that GR applies to the classical mean value of the field, $\psi_0(t)$, while quantum effects deal only with its fluctuating part $\delta\psi(\mathbf{x},t) \equiv \psi(\mathbf{x},t) - \psi_0(t)$. This is reasonable when ψ_0 is energetically and

[6] At length scales smaller than the Planck length, quantum fluctuations of the metric are so large that it becomes meaningless to speak about classical space-time. Rather, one is dealing with something like a space-time foam. Therefore, when one speaks about initial conditions in some patch of a classical space-time, one automatically implies that this domain has a size greater than 10^{-33} cm and a density lower than the Planck energy density, $\rho_{\rm Pl} = 5.16 \times 10^{93}\,{\rm g\,cm^{-3}}$.

then dynamically dominant (i.e. after the Planck era). The evolutionary equations for ψ_0 and a are, respectively,

$$\ddot{\psi}_0 + 3H\dot{\psi}_0 = -V'(\psi_0);, \tag{1.17}$$

$$H^2 = \left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi}{3m_{\rm Pl}^2}\left[\frac{\dot{\psi}_0^2}{2} + V(\psi_0)\right] - \frac{k}{a^2};, \tag{1.18}$$

where $V(\psi_0)$ is the effective potential of the classical scalar field ψ_0 and a prime denotes derivatives with respect to ψ_0 (natural units with $\hbar = c = 1$ and $G = 1/m_{\rm Pl}^2$ are used throughout this section).

- An inflationary solution is obtained when ψ_0 is initially displaced from, but slowly approaching, the value that minimize V. In the slow-roll approximation ($\dot{\psi}^2 \ll V$, $|\ddot{\psi}| \ll H\dot{\psi}$), skipping the terms $\ddot{\psi}_0$ in eq. (1.17) and $\dot{\psi}_0^2$ in eq. (1.18), one easily gets

$$a(t) \sim \exp(Ht), \qquad H(t) = \left(\frac{8\pi}{3}\frac{V}{m_{\rm Pl}^2}\right)^{1/2} \sim \text{const}. \tag{1.19}$$

In other words, when the inflaton ψ_0 rolls down its potential-energy curve so slowly that V keeps nearly constant, a quasi exponential expansion of the universe is obtained. The slow-roll approximation is valid when V is sufficiently flat, i.e. when

$$\frac{m_{\rm Pl}^2}{16\pi}\left(\frac{V'}{V}\right)^2 \ll 1, \tag{1.20}$$

$$\frac{m_{\rm Pl}^2}{8\pi}\left(\frac{V''}{V}\right)^2 \ll 1. \tag{1.21}$$

The inflationary stage ends when these conditions are violated and the inflaton approaches the minimum of V.

- Once inflation is over, the scalar field rapidly and coherently oscillates about the minimum of its potential. During this phase the energy density of our patch of space-time decreases as a^{-3}, i.e. the universe is matter dominated.

- The decay of this oscillations eventually leads to the productions of relativistic particles and produces a radiation dominated universe. This process is called reheating. In the simplest inflationary scenario, radiation dominates until the onset of the present matter dominated era. Therefore, when the universe reaches thermal equilibrium, it recovers the regime described by the standard hot big bang model.

The tremendous expansion of the universe during the slow-rolling phase is very efficient in wiping out any pre-existing inhomogeneities and anisotropies. However, besides this very attractive feature, the inflationary scenario gives us a valuable extra bonus: a plausible explanation for the origin of density fluctuations needed to seed structure formation. In the simplest case, density perturbations emerge from quantum fluctuations of the inflaton field. The exponential growth of the scale factor during the slow-rolling phase is in fact able to transform microscopical length scales into cosmologically relevant distances. The importantant phenomenon for structure formation is that the comoving Hubble length decreases during the inflationary phase (while the physical Hubble

length keeps nearly constant). It can be shown that, in these conditions, vacuum fluctuations with wavelength stretched to scales well larger than the Hubble radius (the so called super-horizon sized modes) stop behaving like quantum variables and give rise to a stochastic classic field with power spectrum (see appendix A for its mathematical definition)

$$\frac{k^3 P_{\delta\psi}(k)}{2\pi^2} = \left(\frac{H}{2\pi}\right)^2 \Bigg|_{k=aH} . \tag{1.22}$$

These fluctuations then induce perturbations in the curvature of the homogeneity surfaces of the (unperturbed) Robertson-Walker geometry. After the inflationary phase, the comoving Hubble length starts growing again (like before inflation), progressively encompassing more and more length scales that inflation drove out of the horizon. Sub-horizon sized fluctuations then manifest themselves as adiabatic density perturbations to comoving observers. The term adiabatic here means that for each universal component i

$$\frac{\delta\rho_i}{\rho_i + p_i/c^2} = \frac{\delta\rho}{\rho + p/c^2} , \tag{1.23}$$

where $\delta\rho_i \equiv \delta\rho_i(\mathbf{x}, t)$ denotes the density fluctuation field and the right hand side refers to the total density and pressure. The statistical properties of these fluctuations depends on the characteristic features of the inflationary phase. In general,

- the resulting density fluctuations form a Gaussian random field (see appendix A for its precise definition). This happens whenever the scalar field that drives inflation is very weakly coupled and behaves like a free field.

- The power spectrum of the density fluctuations at horizon crossing (i.e. evaluated at different epochs for different modes, to satisfy $k = aH$) is nearly scale invariant. This happens beacuse all the cosmologically relevant scales become super-horizon sized during a time interval that is negligible with respect to the duration of the slow-rolling phase. Therefore H cannot change much in this short lapse of time and a flat spectrum of perturbations is generated (roughly speaking $\delta\rho \propto \delta\psi$ and eq. (1.22) shows that $\delta\psi \propto H$).

- A nearly scale-invariant spectrum of gravitational waves arises from quantum mechanical fluctuations excited in the space-time metric during inflation.

Gravitational instability is expected to be responsible of transforming these density fluctuations into the large scale structure we observe. This process will be deeply analysed in the next chapter.

Even though, in this thesis we will only discuss structure formation deriving from the simplest inflationary scenario (i.e. from Gaussian, adiabatic and nearly scale invariant density perturbations), we think it is worth mentioning here a few viable alternatives to the picture we will adopt.

- Quantum fluctuations of a non-inflaton field during an inflationary phase can generate the so-called isocurvature density perturbations. In this case, fluctuations in the energy densities of two (or more) species add up to zero and there is no net spatial curvature of comoving slices. Isocurvature perturbations give rise to entropy perturbations at horizon entry and follow a different evolution from their adiabatic counterparts (see e.g. Liddle & Lyth 1993). The most generic density fluctuation just before horizon entry can be decomposed in the linear superposition of an isocurvature perturbation and an adiabatic fluctuation.

- "Non-Gaussian" density perturbations can be generated from vacuum fluctuations of the inflaton when the latter is strongly coupled to some other field. The detailed statistical properties of the resulting fluctuations obviously depends on the specific form of the interactions between the fields.

- Topological defects (formed during a symmetry breaking phase transition in the early universe) provide an alternative way to seed structure formation. For example, moving cosmic strings are expected to produce wakes of dark matter behind them which define wall-like structures. In addition, wiggles along the strings can lead to filamentary aggregation of dark matter by simple gravitational interaction. Theoretical predictions however are not yet that detailed since the evolution of a string network is a very delicate and numerically heavy problem. Moreover, it is not clear if the defect scenario is consistent with the inflationary solution of the flatness problem. Since inflation dilutes all the defects formed before its beginning, new topological defects should be created after its end. However, to lead to the observed large scale structure, topological defects need to be formed during a phase transition at an energy scale of $\sim 10^{16}$ GeV (Hindmarsh & Kibble 1995) and it appears that the reheating of the universe can hardly be so efficient.

The future satellite missions MAP and PLANCK are expected to give a definitive answer to the problem of the initial conditions for structure formation.

Chapter 2

Gravitational instability in the universe

Understanding the formation of LSS has become one of the most important tasks for modern cosmology. Many fundamental questions, ranging from the physical nature of dark matter to the temperature anisotropies of the CMB and the determination of the epoch of galaxy formation, join together here. It is clearly beyond the scope of this thesis to explore all the proposed mechanisms for structure formation. We will only give a brief presentation of the most accredited scenario, introducing a number of techniques that we will intensively use in the following chapters. A much more complete discussion of these issues (exploring also less standard possibilities) can be found in the review article by Sahni & Coles (1996) and in some cosmology textbooks (e.g. Peebles 1980; Zel'dovich & Novikov 1983; Padmanabhan 1993; Coles & Lucchin 1995).

Current popular structure formation scenarios are based on three assumptions:

1. the matter content of the universe is dominated by a weakly interacting component having negligible velocity dispersion; [1]

2. the main mechanism for structure formation is the gravitational instability of primordial density fluctuations;

3. the initial perturbations constitute a Gaussian random field (see appendix A for a precise definition).

The Newtonian limit of GR (reviewed in §2.1) is, in general, sufficiently accurate to describe the evolution of density perturbation in the cosmological setting. From the mathematical point of view, structure formation can be faced with a fluidodynamic approach, i.e. by solving a complex system of coupled partial differential equations. The corresponding dynamics may in principle be described either from an Eulerian or a Lagrangian point of view. In both cases, the early phases of gravitational instability, involving a slightly perturbed FRW universe, can be dealt with a perturbative approach. In §2.2 and §2.3 we will summarize the methods and the results of these techniques. On the other hand, when the density field shows a high degree of symmetry, it is possible to solve exactly the evolution equations up to the formation of singularities (§2.4). Alternatively, for more complex (and realistic) configurations, one can look for approximate solutions to the instability problem by symplifying the underlying dynamics. We will quickly review some of these attempts in the last sections of this chapter.

[1] Notice that this does not exclude the possibility that the cosmological constant is dynamically dominant on cosmological scales.

2.1 Particle dynamics

2.1.1 Newtonian limit

In the limit of weak gravitational fields, low velocities, small pressures and vanishing cosmological constant, GR reduces to Newton's theory of gravity. Let us consider an isolated system in a asymptotically flat space-time and suppose that there exists a coordinate system $(t, \mathbf{r})$ in which the stress-energy tensor satisfies $T^{00} \gg |T^{0j}| \gg |T^{ij}|$, the velocities of every material component are $v \ll c$, the gravitational potential is weak (i.e. $|\Phi| \ll c^2$), and the metric deviates only slightly from that of a Minkowsky space. The Newtonian equation of motion

$$\frac{d^2\mathbf{r}}{dt^2} = -\nabla_{\mathbf{r}}\Phi \, . \tag{2.1}$$

is obtained from the geodesic equation after keeping only first order terms in $|\Phi|/c^2$ and imposing $g_{00} = 1 + 2\Phi/c^2$. [2] Here, $\Phi(\mathbf{r}, t) < 0$ is Newton's gravitational potential (normalized in such a way that, for an isolated system, $\Phi(r \to \infty) = 0$). In the Newtonian limit, rest mass density is the only source of the gravitational field as described by the Poisson equation

$$\nabla^2\Phi = 4\pi G\rho \, . \tag{2.2}$$

A similar approach can be adopted in cosmology. In this case, it is first necessary to choose locally inertial coordinates and then perform the weak-field, slow motion limit (e.g. Peebles 1980). The result account for the presence of a smooth radiation background and a non-vanishing Λ, and is characterized by the potential

$$\nabla^2\Phi = 4\pi G\left(\rho + \frac{3p}{c^2}\right) - \Lambda \, . \tag{2.3}$$

In the following we will use the term Newtonian gravity to identify the dynamics described by equations (2.1) and (2.3).

2.1.2 Particle dynamics

We are interested in using Newtonian gravity to describe the growth of mass density perturbations in a FRW background model. This analysis will be valid for a wide range of spatial scales ranging from some multiple of the typical Schwarzschild radius of the collapsed systems to some fraction of the particle horizon of the FRW reference universe. Hereafter a subscript b will denote the hydrodynamic variables of the unperturbed cosmological fluid.

In analogy with the analysis presented in chapter 1, it is useful to introduce comoving coordinates $\mathbf{x}$ through $\mathbf{r} = a(t)\mathbf{x}$, with $a(t)$ the expansion factor of the background model. However, in contrast with our previous choice, it is convenient here to re-define the expansion parameter as a dimensionless quantity normalized so that $a_0 = 1$. This implies that, in the relativistic case, we can continue to use the Robertson–Walker metric provided we allow the curvature parameter k to assume a continuous spectrum of values. In practice, we are simply choosing to define the comoving coordinates through distances measured on our present hypersurface of homogeneity instead of using their ratio with the actual value of the expansion parameter. This is convenient because we have no clue about the present value of the universal curvature radius, while we are able to measure

[2] Note that, while to recover Newtonian theory is sufficient to define the time-time component of the metric tensor, all the metric coefficients are determined up to errors of $\mathcal{O}(\Phi v)$, $\mathcal{O}(\Phi^2)$ and $\mathcal{O}(\Phi|T_{jk}|/T_{00})$ only in the so-called *linearized theory of gravity* where the line element is $ds^2 = (1 + 2\Phi/c^2)c^2dt^2 - (1 - 2\Phi/c^2)(dr_1^2 + dr_2^2 + dr_3^2)$.

physical distances in the cosmic neighbourhood of the Earth and redshifts of far away objects (note that in the new notation, simply, $1 + z = 1/a$).

The Lagrangian for a test particle of mass m is $\tilde{\mathcal{L}}(\mathbf{x}, \dot{\mathbf{x}}, t) = (m/2)(a\dot{\mathbf{x}} + \dot{a}\mathbf{x})^2 - m\Phi(\mathbf{x}, t)$ that, after subtracting the total time derivative $d\psi/dt$ with $\psi = m\, a\, \dot{a}\, x^2/2$, is equivalent to

$$\mathcal{L} = \frac{1}{2} m\, a^2 \dot{\mathbf{x}}^2 - m\, \phi(\mathbf{x}, t) \; . \tag{2.4}$$

The potential $\phi(\mathbf{x}, t) = \Phi(\mathbf{x}, t) + a\, \ddot{a}\, x^2/2$ is generated by density fluctuations according to the Poisson equation

$$\nabla^2 \phi = 4\pi G a^2 [\rho(\mathbf{x}, t) - \rho_b(t)] \; , \tag{2.5}$$

where the gradient operator ∇ denotes differentiation with respect to the comoving coordinates $\mathbf{x}$. The general solution to eq. (2.5) is:

$$\phi(\mathbf{x}, t) = -Ga^2 \int \frac{\rho(\mathbf{x}', t) - \rho_b(t)}{|\mathbf{x} - \mathbf{x}'|} d^3x' + \Gamma(\mathbf{x}, t) \; , \tag{2.6}$$

where Γ represents the contribution of possible source-free terms (i.e. gravitational waves) that we will neglect in the following discussion. Self-consistency of the theory is assured if and only if the integral that defines $\phi(\mathbf{x}, t)$ converges to a definite value at separations $|\mathbf{x} - \mathbf{x}'| \ll R_H/a$. It is customary to model the density fluctuation field,

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \rho_b(t)}{\rho_b(t)} \; , \tag{2.7}$$

with a spatially homogeneous and isotropic random field (see appendix A); in this case, the convergence of ϕ is assured if $x_\delta \ll R_H/a$, with x_δ the comoving correlation length of δ (Peebles 1980).

The Hamiltonian for particle motion is

$$\mathcal{H}(\mathbf{x}, \mathbf{p}, t) = \frac{p^2}{2ma} + m\phi(\mathbf{x}, t) \tag{2.8}$$

where $\mathbf{p} = ma^2\dot{\mathbf{x}}$ denotes the canonical conjugate momentum to $\mathbf{x}$. The resulting equations of motion are

$$\dot{\mathbf{x}} = \frac{\mathbf{p}}{ma^2} \tag{2.9}$$

$$\dot{\mathbf{p}} = -m\nabla\phi(\mathbf{x}, t) \; . \tag{2.10}$$

It is convenient to describe particle motion through the proper peculiar velocity, $\mathbf{v} = a\dot{\mathbf{x}}$, that evolves according to

$$\dot{\mathbf{v}} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{\nabla\phi}{a} \; . \tag{2.11}$$

The corresponding gravitational acceleration, $\mathbf{g}(\mathbf{x}, t) \equiv -\nabla\phi/a$, can be expressed in term of the total mass density ρ,

$$\mathbf{g}(\mathbf{x}, t) = Ga \int \rho(\mathbf{x}, t) \frac{\mathbf{x} - \mathbf{x}'}{|\mathbf{x} - \mathbf{x}'|^3} d^3x' \; , \tag{2.12}$$

after conventing that the volume integral has to be performed in 2 steps: first over angles, and then over the radial displacement. The gravitational acceleration generated by a set of point particles (of mass m_j) is then given by

$$\mathbf{g}(\mathbf{x},t) = \frac{G}{a^2}\sum_j m_j \frac{\mathbf{x}_j - \mathbf{x}}{|\mathbf{x}_j - \mathbf{x}|^3} \,. \tag{2.13}$$

The exact equations describing the evolution of a set of gravitationally interacting particles (the so-called BBGKY hierarchy) are obtained in appendix B, where their relation to the fluid limit we will adopt in the forthcoming discussion is also clarified.

2.2 Eulerian perturbation theory

We adopt macroscopic (fluid) variables to describe the matter content of the universe: the mass density fluctuation field $\delta(\mathbf{x},t)$ and the peculiar velocity field $\mathbf{v}(\mathbf{x},t)$, which are functions of the comoving spatial coordinate $\mathbf{x}$ and the time t. We also assume that the cosmic fluid is in the single-stream regime, i.e. that at any location in space one can associate a unique peculiar velocity. As discussed in appendix B, it is possible to obtain a closed system of evolutionary equations for this physical system by taking velocity moments of the kinetic equation for the cosmic fluid. Considering a non-relativistic, collisionless medium characterized by negligible thermal motions (i.e. assuming that the weakly interacting dark matter dominates the dynamics on the scales of interest) and the mean-field approximation (i.e. starting from the Vlasov equation) one gets the so-called fluid equations for pressureless dust

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla\cdot(1+\delta)\mathbf{v} = 0 \tag{2.14}$$

$$\frac{\partial\mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v}\cdot\nabla)\mathbf{v} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a}\nabla\phi \tag{2.15}$$

$$\nabla^2\phi = 4\pi G\rho_b a^2\delta \tag{2.16}$$

It is possible to combine the Euler equation (2.15) and the continuity equation (2.14) into a single differential equation for δ, that, using the Poisson equation (2.16), can be recast in the form

$$\frac{\partial^2\delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta}{\partial t} - 4\pi G\rho_b\delta = 4\pi G\rho_b\delta^2 + \frac{1}{a^2}\nabla\delta\cdot\nabla\phi + \frac{1}{a^2}\frac{\partial^2}{\partial x_\alpha\partial x_\beta}(1+\delta)v_\alpha v_\beta \,, \tag{2.17}$$

where summation over repeated indices is understood. The non-local terms (i.e. those involving spatial derivatives) in eq. (2.17) describe tidal and shear effects due to density fluctuations at points in the neighbourhood of $\mathbf{x}$. We are interested in solving the partial differential equation (2.17) ahead in time, starting from stochastic initial conditions. Also the inverse problem of recovering the initial conditions from an evolved snapshot of the density and velocity field (extracted from observational data) is of great interest for cosmology (even though "time-machine" algorithms are, except in a few trivial cases, mathematically more troublesome than initial value problems).

2.2.1 Linear theory

In the homogeneous background $\delta(\mathbf{x},t)$, $\mathbf{v}(\mathbf{x},t)$ and $\phi(\mathbf{x},t)$ vanish. We assume that the structure we observe in the real universe formed from small primordial density fluctuations. Therefore, at early times, we expect that δ, $\mathbf{v}$ and ϕ are all $\mathcal{O}(\epsilon)$ with $\epsilon \ll 1$ so that it is possible to ignore all terms $\mathcal{O}(\epsilon^n)$ with $n > 1$ in the evolutionary equations. We will show in §3.2 that this perturbative method is reliable for a wide class of initial conditions (e.g. Zel'dovich 1965; Peebles 1974; Peebles and Groth 1976).

Growing and decaying modes

Applying this procedure of linearization to eq. (2.17) one gets a completely local description of perturbation growth [3]

$$\frac{\partial^2\delta^{(1)}}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta^{(1)}}{\partial t} - 4\pi G\rho_b\delta^{(1)} = 0\ . \tag{2.18}$$

In full generality, the linear solution $\delta^{(1)}$ is given by the linear superposition of two modes, one (hereafter denoted by the subscript $+$) describing an overdensity that grows with time and the other (subscript $-$) that decays,

$$\delta^{(1)}(\mathbf{x},t) = \delta_+(\mathbf{x})D_+(t) + \delta_-(\mathbf{x})D_-(t)\ . \tag{2.19}$$

The functions D_+ and D_- are the independent solutions of the second order, ordinary differential equation $\ddot{D}+2(\dot{a}/a)\dot{D}-4\pi G\rho_b D = 0$, and take different functional forms depending of the assumed background cosmology. For instance, in a matter dominated Einstein-de Sitter universe, where $a \propto t^{2/3}$ and $4\pi G\rho_b = 2/3t^2$, one finds $D_+(t) \propto t^{2/3}$ and $D_-(t) \propto t^{-1}$. Analytical solutions for different cosmological models have been given, among others, by Peebles (1980) and Eisenstein (1997). On the other hand, approximated formulae for D_+ showing simple dependencies on cosmological parameters can be found in Lahav et al. (1991) and Carroll et al. (1992). In full generality, considering a matter dominated universe and choosing a as time coordinate, one finds (Heath 1977)

$$D_+(a) \propto \Theta(a)\int_0^a \frac{da'}{a'^3\Theta(a')^3} \qquad\qquad D_-(a) \propto \Theta(a) \tag{2.20}$$

$$\Theta(a) \equiv \frac{H(a)}{H_0} = \sqrt{\frac{\Omega_0}{a^3} + \frac{1-\Omega_0-\lambda_0}{a^2} + \lambda_0}\ . \tag{2.21}$$

The normalization of these functions is arbitrary (since any multiplicative factor can be inserted in the definition of δ_+ and δ_-) and from now on we set $D_+(t_0) = 1$. This means that $\delta_+(\mathbf{x})$ will denote the growing mode of the density field linearly extrapolated to the present.

[3] A collisional medium follows a different linear evolution. In fact, the conservation equations deriving from the Boltzmann kinetic equation are obtained by simply adding the term $-\nabla p/a\rho$ to the right hand side of eq. (2.15), while keeping eq. (2.14) and eq. (2.16) as they appear in the main text. This implies the appearance of the term $\nabla^2 p^{(1)}/a^2\rho_b$ in the right hand side of eq. (2.18). The competitive action of the gravitational driving term $4\pi G\rho_b\delta^{(1)}$ and of the restoring force given by the pressure gradient, introduces a characteristic scale, λ_J (the Jeans length), in the system. This is best understood by reasoning in Fourier space (e.g. Coles & Lucchin 1995). On the one hand, density perturbations with wavelength $\lambda < \lambda_J \simeq c_s(\pi/G\rho_b)^{1/2}$ (with $c_s = (\partial p/\partial\rho)^{1/2}$ the adiabatic sound speed) propagate as acoustic waves and are slowly damped by the Hubble espansion. On the other hand, perturbations with larger characteristic scales are gravitationally unstable. Fourier modes with $\lambda \gg \lambda_J$ follow the evolution discussed in the main text.

Solutions for ϕ, $\mathbf{g}$ and $\mathbf{v}$ to first order in ϵ are obtained by linearizing the continuity and Euler equations,

$$\frac{\nabla \cdot \mathbf{v}^{(1)}}{a} = -\frac{\partial \delta^{(1)}}{\partial t} \tag{2.22}$$

$$\frac{\partial \mathbf{v}^{(1)}}{\partial t} + \frac{\dot{a}}{a}\mathbf{v}^{(1)} = -\frac{1}{a}\nabla\phi^{(1)};, \tag{2.23}$$

and by solving the Poisson equation

$$\phi^{(1)}(\mathbf{x},t) = -G\rho_b(t)a^2(t)\int \frac{\delta^{(1)}(\mathbf{x}',t)}{|\mathbf{x}-\mathbf{x}'|}d^3x' = \phi_g(\mathbf{x},t) + \phi_d(\mathbf{x},t) = \tag{2.24}$$

$$= \rho_b(t)a^2(t)[\phi_+(\mathbf{x})D_+(t) + \phi_-(\mathbf{x})D_-(t)]\,. \tag{2.25}$$

The final result can be cast in the form

$$\mathbf{g}^{(1)}(\mathbf{x},t) = -\frac{\nabla\phi^{(1)}(\mathbf{x},t)}{a} \tag{2.26}$$

$$\mathbf{v}^{(1)}(\mathbf{x},t) = a\frac{\partial}{\partial t}\left(\frac{\mathbf{g}^{(1)}}{4\pi G\rho_b a}\right) + \frac{\mathbf{F}(\mathbf{x})}{a}\,, \tag{2.27}$$

where the term $\mathbf{F}/a$ denotes the rotational part of the peculiar velocity ($\nabla \cdot \mathbf{F} = 0$) which is not coupled to density perturbations and is damped by the cosmological expansion as a result of angular momentum conservation. In total analogy with density and potential fluctuations, the peculiar acceleration, $\mathbf{g}^{(1)}$, and the irrotational part of $\mathbf{v}^{(1)}$ can be decomposed into the sum of growing and decaying modes; i.e.

$$\mathbf{g}^{(1)}(\mathbf{x},t) = \rho_b(t)a(t)[\mathbf{g}_+(\mathbf{x})D_+(t) + \mathbf{g}_-(\mathbf{x})D_-(t)] \tag{2.28}$$

$$\mathbf{v}^{(1)}(\mathbf{x},t) = \frac{2}{3H(t)\Omega(t)}\rho_b(t)a(t)\left[\mathbf{g}_+(\mathbf{x})f_+(t)D_+(t) + \mathbf{g}_-(\mathbf{x})f_-(t)D_-(t)\right] + \frac{\mathbf{F}(\mathbf{x})}{a}\,, \tag{2.29}$$

where $f_i(t) = d\ln D_i/d\ln a$. A common analytic approximation is $f_+(\Omega,\lambda) \simeq \Omega^{0.6}$ (Peebles 1980) but more accurate expressions are available in the literature (Lahav et al. 1991; Carroll et al. 1992). In the matter dominated Einstein-de Sitter universe $\rho_b \propto a^{-3}$ and $D_+ \propto a$, so that the linearized gravitational potential associated with the density growing mode does not evolve in time. In this case $f_+ = 1$ and $f_- = -3/2$, thus the growing and decaying modes of the peculiar velocity field evolve proportionally to $t^{1/3}$ and $t^{-4/3}$, respectively. Note that, for the growing (decaying) mode, the gravitational acceleration is always parallel (antiparallel) to the irrotational part of the peculiar velocity. This result is emphasized by defining a peculiar velocity potential $\phi_v(\mathbf{x},t) = \phi_{vg}(\mathbf{x},t) + \phi_{vd}(\mathbf{x},t) = (2f_+/3\Omega H)\phi_g(\mathbf{x},t) + (2f_-/3\Omega H)\phi_d(\mathbf{x},t)$ so that $\mathbf{v}^{(1)} = -\nabla\phi_v/a + \mathbf{F}/a$. Here, we adopted the notation introduced in eq. (2.25) using the subscripts g and d to denote, respectively, the growing and the decaying modes with their full spatio-temporal dependence.

Using the equations (2.19) and (2.26) we can express δ_+ and δ_- in terms of the initial conditions at $t = t_i$,

$$\delta_+(\mathbf{x}) = \frac{\dot{D}_-(t_i)\,\delta(\mathbf{x},t_i)}{D_+(t_i)\dot{D}_-(t_i) - D_-(t_i)\dot{D}_+(t_i)} + \frac{D_-(t_i)\,\nabla\cdot\mathbf{v}^{(1)}(\mathbf{x},t_i)}{a(t_i)[D_+(t_i)\dot{D}_-(t_i) - D_-(t_i)\dot{D}_+(t_i)]} \tag{2.30}$$

$$\delta_-(\mathbf{x}) = \frac{1}{D_-(t_i)}[\delta(\mathbf{x},t_i) - \delta_+(\mathbf{x})D_+(t_i)]\,. \tag{2.31}$$

For example, considering the Einstein-de Sitter universe and requiring that $\mathbf{v}^{(1)}(\mathbf{x},t) = 0$, one obtains $\delta_+(\mathbf{x})D_+(t_i) = (3/5)\delta(\mathbf{x},t_i)$ and $\delta_-(\mathbf{x})D_-(t_i) = (2/5)\delta(\mathbf{x},t_i)$. It is then reasonable to expect that, at every point $\mathbf{x}$, growing and decaying modes have similar characteristic amplitudes at early times. This implies that the decaying mode soon becomes negligible with respect to its growing counterpart and $\delta^{(1)}$ has the same spatial dependence as δ_+ at any time. In this sense, we can say that, in the linear regime, the developing structure retains a memory of its origin.

The transfer function

In the previous section we considered the linear evolution of a collection of identical, collisionless particles that we supposed to dominate the matter-energy content of the universe. However, a realistic model for structure formation should account for the parallel evolution of all the material and radiative components of the universe. In particular one has to consider both the physical interactions between different particle species and the dissipative effects associated to transport phenomena.

As far as density and velocity fluctuations keep small, it is possible to deal with these processes in the framework of the linear perturbative regime. In practice, one assumes that small perturbations are somehow generated during an early phase of the universal history (e.g. inflation) and follows their linear evolution. This is generally done in Fourier space since, in linear theory, Fourier modes evolve independently of each other. [4] It is then convenient to parametrize the physical evolution of the density field for each universal component (i.e. baryons, photons and non-baryonic matter; here denoted by a generic index i) by introducing the transfer function, $T_i(k)$, that relates the amplitude of a primordial (growing) mode, $\tilde{\delta}_i^{(1)}(\mathbf{k},t_{\rm in})$ (where $t_{\rm in}$ characterizes an epoch at which all the scales of interest are still super-horizon sized), to its counterpart at some later stage (in general at recombination, when matter fluctuation start growing), $\tilde{\delta}_i^{(1)}(\mathbf{k},t_{\rm out})$, scaling out the temporal dependence of the linear growing mode at small $|\mathbf{k}|$

$$T_i(\mathbf{k}) \equiv \frac{D_+(t_{\rm in})}{D_+(t_{\rm out})} \frac{\tilde{\delta}_i^{(1)}(\mathbf{k},t_{\rm out})}{\tilde{\delta}_i^{(1)}(\mathbf{k},t_{\rm in})} \,. \tag{2.32}$$

It is often assumed that the primordial density fluctuations form a Gaussian random field. In the linear regime, the density field retains its initial Gaussian character. The statistical properties of the field $\delta(\mathbf{x})$ are then completely specified by its two-point correlation function in Fourier space, which is related to the power-spectrum $P(k)$ by $\langle\tilde{\delta}(\mathbf{k}_1,t)\,\tilde{\delta}(\mathbf{k}_2,t)\rangle = (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2)P(k_1,t)$. Here, the symbol δ_D represents the three-dimensional Dirac delta distribution, and the brackets $\langle\cdot\rangle$ denote ensemble averaging. Then, in linear theory,

$$P(k,t_{\rm out}) = \left[\frac{D_+(t_{\rm out})}{D_+(t_{\rm in})}\right]^2 T^2(k)\; P(k,t_{\rm in}) \,. \tag{2.33}$$

The transfer function is computed by (numerically) solving the scalar components (in linear theory vector and tensor modes are not coupled to the density field) of the dynamical system obtained by linearizing the Einstein equations coupled to a number of kinetic equations (one for each universal energy component). This obviously requires the selection of a cosmological background

[4] Our convention for the Fourier transform of the field $A(\mathbf{x},t)$ is

$$\tilde{A}(\mathbf{k},t) = \int d\mathbf{x}\, A(\mathbf{x},t)\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}$$

and the detailed specification of the physical characteristics of all the cosmic components (i.e. their interactions, stability, possible decaying products, velocity dispersion, etc.).

Even though the resulting shape of $T_i(k)$ depends on many inputs, the net effect of linear evolution is to introduce some characteristic scales in the spectrum of each cosmic component. A detailed discussion of the photon and baryon transfer function is beyond the scope of this thesis and we refer the interested reader to the papers by Hu & White (1997), Eisenstein & Hu (1998) and Hu & Eisenstein (1999). Since our work is aimed at discussing structure formation in a universe dominated by non-baryonic dark matter, we describe here only the characteristic features appearing in the spectrum of this collisionless medium.

- At early times the wavelengths of all the fluctuations relevant for structure formation are larger than the Hubble radius. Density perturbations (a not well defined concept in this conditions, e.g. Bardeen 1980), subject to kynematic evolution, are given by the superposition of two physical modes: one proportional to $D_+(t)$ and the other to $D_-(t)$. As far as only the growing mode is considered, the primordial spectral shape is preserved during all this evolutionary phase.

- As time passes the Hubble radius grows, becoming larger than the wavelength of more and more perturbations that start interacting. Since sub-horizon sized perturbations remain frozen (i.e. do not grow sensibly) up to the equivalence epoch while super-horizon sized fluctuations keep growing proportionally to $D_+(t)$, the power spectrum flattens on scales smaller than the Hubble radius at $t_{\rm eq}$.

- Collisionless phase mixing severely damps the amplitude of all the Fourier modes with wavelength smaller than the free-streaming scale, $\lambda_{\rm fs}$, which is the characteristic distance covered by non-baryonic particles since they decouple from the other cosmological components until they become non-relativistic. Roughly speaking $\lambda_{\rm fs} \sim 1\,{\rm Mpc}\,(m_X/{\rm keV})^{-1}(T_X/T_{\rm CMB})$ with m_X the mass of the particle and $T_X/T_{\rm CMB}$ the actual ratio of its temperature to the photon temperature (Kolb & Turner 1990).

- During the radiation dominated epoch, photons and baryons are tightly coupled. Because of the support coming from photon pressure, the amplitude of the baryonic perturbations cannot grow and undergo harmonic motion with a slightly decaying amplitude. Modes that, at recombination, reach a maximum or a minimum of the oscillatory phase generate peaks in the baryon and radiation power spectrum. In a universe with a substantial baryon fraction, gravitational coupling is able to transfer this oscillating behaviour to the spectrum of dark matter (Eisenstein et al. 1998; Eisenstein & Hu 1998; Meiksin, White & Peacock 1999).

Structure formation scenarios are almost totally driven by the assumed form of dark matter as well as its overall contribution to the mass density of the universe. We can then consider two limiting cases of structure formation that arise under the gravitational instability paradigm. Each of these cases assumes that a single kind of particle dominates the matter content of the universe.

In the hot dark matter (HDM) scenario, the universe is dominated by massive neutrinos with mass around $10 - 30$ eV. The copious free-streaming of these light particles smooths the matter fluctuations on scales $\sim 40\,(m_\nu/30{\rm eV})^{-1}$ Mpc. Collapsed structures of these sizes, corresponding to objects with $10^{15} - 10^{16}{\rm M}_\odot$, would then be the first to form (Centrella & Melott 1983; White, Frenk & Davis 1983). The formation of clumps on smaller scales would occur by gas fragmentation within the dark matter potential wells. While this top-down scenario is very efficient in producing structures on large scales, there are serious difficulties in obtaining the early galaxy formation that observations suggest.

In the cold dark matter (CDM) model (e.g. Blumenthal et al. 1984), the dominant particles due to their much higher mass (e.g. supersymmetric particles) or non-thermal distribution (e.g. axions) undergo negligible free-streaming. In this case, structure in the universe is built by the hierarchical gravitational clustering of sub-units (see chapter 4). This naturally leads to an early galaxy formation phase. The "standard" CDM model ($\Omega_0 = 1$, $\lambda_0 = 0$, $h = 0.5$, and primordial scale invariant spectrum with index $n = 1$) was the favourite model for structure formation during the 1980s, since it could qualitatively (and in some case also quantitatively) reproduce many properties of the observed galaxy population (see e.g. Ostriker 1993 for a review). However, it has been realized that this bottom-up scenario is unable to produce a sufficient amount of large-scale structure to be consistent with observations (e.g. Peacock & Dodds 1994). Nevertheless, because SCDM represents such a successful first approximation, it has not been abandoned entirely: a common approach is to relax some of its assumptions in order to get better agreement with observations. One possibility to generate more large scale structure is to consider the class of "open" CDM (OCDM) models where the density parameter is reduced to $\Omega_0 \sim 0.2 - 0.3$. Also a non-vanishing cosmological constant can be added to account for the "cosmic concordance" discussed in chapter 1 (ΛCDM models). Alternatively, it is possible to assume that a mixture of hot and cold dark matter (CHDM or Mixed Dark Matter) is responsible of the matter content of the universe. For a fixed large-scale normalization, adding a hot component has the effect of suppressing the power spectrum amplitude at small wavelengths. Many other variants of the CDM paradigm are available (see e.g. Gawiser & Silk 1998).

2.2.2 Non-linear theory

We are interested in discussing the dynamics of density and velocity perturbations in a self-gravitating collisionless fluid beyond their linear regime. A significative simplification of the evolutionary equations is obtained by introducing a new velocity field. We already noted that any initial vorticity in the cosmic fluid decays proportionally to a^{-1} as a consequence of angular momentum conservation. Moreover, Kelvin's circulation theorem states that vorticity cannot be generated in a potential flow. This holds true until multi-streaming regions (or shocks in a collisional medium) form. Therefore, limiting our analysis to the single-stream regime and to growing modes, we take the velocity field to be irrotational so that it can be completely described by its divergence $\theta(\mathbf{x}, t) \equiv \nabla \cdot \mathbf{v}(\mathbf{x}, t)$.

When dealing with higher orders in perturbation theory, it is often convenient to work in Fourier space. In this way, spatial integrals and derivatives become simple algebraic operations while simple products become convolutions. In fact, indicating by FT the Fourier transform operation,

$$\begin{aligned} &\mathrm{FT}\,[A_1(\mathbf{x})A_2(\mathbf{x})\dots A_n(\mathbf{x})] = \\ &\quad = \int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\cdots\frac{d^3k_n}{(2\pi)^3}(2\pi)^3\delta_D(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2-\cdots-\mathbf{k}_n)\tilde{A}_1(\mathbf{k})\tilde{A}_2(\mathbf{k})\dots\tilde{A}_n(\mathbf{k})\ . \end{aligned} \tag{2.34}$$

In $\mathbf{k}$-space, continuity and Euler equations become, respectively

$$\frac{\partial\tilde{\delta}(\mathbf{k},t)}{\partial t} + \frac{1}{a}\tilde{\theta}(\mathbf{k},t) = -\frac{1}{a}\int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\delta_D(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\alpha(\mathbf{k},\mathbf{k}_1)\tilde{\theta}(\mathbf{k}_1,t)\tilde{\delta}(\mathbf{k}_2,t)\ , \tag{2.35}$$

$$\begin{aligned} &\frac{\partial\tilde{\theta}(\mathbf{k},t)}{\partial t} + \frac{\dot{a}}{a}\tilde{\theta}(\mathbf{k},t) + 4\pi G a\rho_b\tilde{\delta}(\mathbf{k},t) = \\ &\qquad = -\frac{1}{a}\int \frac{d^3k_1}{(2\pi)^3}\frac{d^3k_2}{(2\pi)^3}\delta_D(\mathbf{k}-\mathbf{k}_1-\mathbf{k}_2)\beta(\mathbf{k},\mathbf{k}_1)\tilde{\theta}(\mathbf{k}_1,t)\tilde{\theta}(\mathbf{k}_2,t)\ . \end{aligned} \tag{2.36}$$

where $\alpha(\mathbf{k},\mathbf{k}_1) = \mathbf{k}\cdot\mathbf{k}_1/k_1^2$ and $\beta(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2) = k^2(\mathbf{k}_1\cdot\mathbf{k}_2)/2k_1^2k_2^2$. It is evident that, in Fourier space, non-linear dynamics is determined by the mode coupling of the fields at all pairs of wavectors $\mathbf{k}_1$ and $\mathbf{k}_2$ whose sum is $\mathbf{k}$. It can be shown that this is required by translational invariance in a spatially homogeneous universe (e.g. Jain & Bertschinger 1994).

A full perturbative approach to the problem of gravitational instability is obtained by assuming

$$\tilde{\delta}(\mathbf{x},t) = \sum_{n=1}^{\infty} \tilde{\delta}^{(n)}(\mathbf{k},t) \ , \quad \tilde{\theta}(\mathbf{x},t) = \sum_{n=1}^{\infty} \tilde{\theta}^{(n)}(\mathbf{k},t) \ , \tag{2.37}$$

where, at each order n, only the fastest growing mode is taken into account and, for example, $\tilde{\delta}^{(n)}$ is $\mathcal{O}(\tilde{\delta}^{(1)n})$. Formally, one can think to compute perturbations at every order and sum the whole series. This is obviously impossible in practice (and probably it would be also useless). In general perturbative expansions are asymptotic series and little is now about their convergence. Comparison with N-body simulations shows that perturbative results are generally reliable when the rms density fluctuation $\sigma = \langle\delta^2\rangle^{1/2} < 1$ (e.g. Juszkiewicz, Bouchet & Colombi 1993; Baugh, Gaztañaga & Efstathiou 1995). Beyond this regime, the expansions (2.37) does not make sense and all the terms in the series would be of the same order.

Since lower order results provide source terms for higher orders, the equation of motion determine $\tilde{\delta}^{(n)}(\mathbf{k},t)$ and $\tilde{\theta}^{(n)}(\mathbf{k},t)$ in terms of the corresponding linear fluctuations:

$$\tilde{\delta}^{(n)}(\mathbf{k},t) = \int d^3q_1 \ldots d^3q_n \ \delta_D(\mathbf{k}-\mathbf{q}_1-\ldots\mathbf{q}_n)\ F_n^{(S)}(\mathbf{q}_1,\ldots,\mathbf{q}_n)\ \tilde{\delta}^{(1)}(\mathbf{q}_1,t)\ldots\tilde{\delta}^{(1)}(\mathbf{q}_n,t) \ , \tag{2.38}$$

$$\tilde{\theta}^{(n)}(\mathbf{k}) = \int d^3q_1 \ldots d^3q_n \ \delta_D(\mathbf{k}-\mathbf{q}_1-\ldots\mathbf{q}_n)\ G_n^{(S)}(\mathbf{q}_1,\ldots,\mathbf{q}_n)\ \tilde{\theta}^{(1)}(\mathbf{q}_1,t)\ldots\tilde{\theta}^{(1)}(\mathbf{q}_n,t) \ , \tag{2.39}$$

where the kernels $F_n^{(S)}$ and $G_n^{(S)}$ are symmetric homogeneous functions of the wavevectors $(\mathbf{q}_1,\ldots,\mathbf{q}_n)$ with degree zero. In the Einstein-de Sitter model, they are constructed from the fundamental mode coupling functions $\alpha(\mathbf{k},\mathbf{k}_1)$ and $\beta(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2)$ according to the recursion relations ($n \geq 2$; Goroff et al. 1986, Jain & Bertschinger 1994):

$$F_n(\mathbf{q}_1,\ldots,\mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1,\ldots,\mathbf{q}_m)}{(2n+3)(n-1)}[(2n+1)\alpha(\mathbf{k},\mathbf{k}_1)F_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n) + \\ + 2\beta(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2)G_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n)] \ , \tag{2.40}$$

$$G_n(\mathbf{q}_1,\ldots,\mathbf{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\mathbf{q}_1,\ldots,\mathbf{q}_m)}{(2n+3)(n-1)}[3\alpha(\mathbf{k},\mathbf{k}_1)F_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n) + \\ + 2n\beta(\mathbf{k},\mathbf{k}_1,\mathbf{k}_2)G_{n-m}(\mathbf{q}_{m+1},\ldots,\mathbf{q}_n)] \ , \tag{2.41}$$

where $\mathbf{k}_1 = \mathbf{q}_1+\cdots+\mathbf{q}_m$, $\mathbf{k}_2 = \mathbf{q}_{m+1}+\ldots,+\mathbf{q}_n$, $\mathbf{k} = \mathbf{k}_1+\mathbf{k}_2$, $F_1 = G_1 = 1$), and the symmetrization procedure:

$$F_n^{(S)}(\mathbf{q}_1,\ldots,\mathbf{q}_n) = \frac{1}{n!}\sum_{\pi} F_n[\mathbf{q}_{\pi(1)},\ldots,\mathbf{q}_{\pi(n)}] \ , \tag{2.42}$$

$$G_n^{(S)}(\mathbf{q}_1,\ldots,\mathbf{q}_n) = \frac{1}{n!}\sum_{\pi} G_n[\mathbf{q}_{\pi(1)},\ldots,\mathbf{q}_{\pi(n)}] \ , \tag{2.43}$$

where the sum is taken over all the permutations π of the set $\{1,\ldots,n\}$. For example, for $n=2$ one has

$$F_2^{(S)}(\mathbf{q}_1,\mathbf{q}_2)=\frac{5}{7}+\frac{1}{2}\frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right)+\frac{2}{7}\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2}\,, \tag{2.44}$$

$$G_2^{(S)}(\mathbf{q}_1,\mathbf{q}_2)=\frac{3}{7}+\frac{1}{2}\frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right)+\frac{4}{7}\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2}\,. \tag{2.45}$$

Explicit expressions for the non-symmetrized kernels F_3 and F_4 are given in Goroff et al. (1986). Note that the complexity of the symmetrized kernels increases rapidly with n: third and fourth order kernels have 134 and 8523 terms, respectively.

In a generic FRW model, equations (2.38) and (2.39) are not valid since non-linear solutions are non-separable in the $\mathbf{x}$ and t dependence (Catelan et al. 1995). Therefore $\delta^{(n)}(\mathbf{x},t)$ does not come out simply proportional to D_+^n and the corresponding kernels do depend on time. In this case, adopting $D\equiv D_+$ as time variable, the expression for $F_2^{(S)}$ and $G_2^{(S)}$ become

$$F_2^{(S)}(\mathbf{q}_1,\mathbf{q}_2)=\frac{1}{2}\left(1-\frac{E}{D^2}\right)+\frac{1}{2}\frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right)+\frac{1}{2}\left(1+\frac{E}{D^2}\right)\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2}\,, \tag{2.46}$$

$$G_2^{(S)}(\mathbf{q}_1,\mathbf{q}_2)=-\frac{dE}{dD}+\frac{1}{2}\frac{\mathbf{q}_1\cdot\mathbf{q}_2}{q_1q_2}\left(\frac{q_1}{q_2}+\frac{q_2}{q_1}\right)+\left(1+\frac{dE}{dD}\right)\frac{(\mathbf{q}_1\cdot\mathbf{q}_2)^2}{q_1^2q_2^2}\,, \tag{2.47}$$

where $E/D^2\to -3/7$ and $(dE/dD)/D\to -6/7$ when $D\to 0$. In general these corrections to the Einstein-de Sitter terms appear small: for vanishing λ_0 $E/D^2\simeq -(3/7)\Omega^{-2/63}$ (Bouchet et al. 1992) while, if $\Omega_0+\lambda_0=1$, $E/D^2\sim -(3/7)\Omega^{-1/140}$ (Kamionkowski & Buchalter 1998).

2.2.3 The weakly non-linear regime

Gravitational instability is expected to deeply modify the statistical properties of the density fluctuation field as time goes on. Roughly speaking, because of the attracting nature of gravity, regions in which $\delta>0$ will tend to become more and more overdense by accumulating matter flowing away from zones where $\delta<0$. While δ cannot assume values smaller than -1 (since mass is a semi-positive defined quantity), positive density fluctuations can easily grow to values much larger than unity. Moreover, from the spatial point of view, overdense regions tend to contract and collapse, while their underdense counterparts tend to expand becoming less and less dense. After some time, one then expects rare events with a large and positive density contrast separated by large volumes with a small and negative δ. Therefore, as time passes, the shape of the one-point probability density function for δ is expected to significantly depart from a Gaussian distribution, being skewed towards large overdensities. This evolution can be characterized through the emergence of non-zero cumulants (of order greater than two) for this distribution. The computation of this characteristic signature, in the weakly non-linear regime, represents the major achievement of the Eulerian perturbation theory (beyond the linear approximation).

Assuming Gaussian initial conditions and using the perturbative machinery, it is possible to devise a diagrammatic technique to compute statistical (i.e. obtained after averaging over the ensemble which is equivalent to spatial averaging if a suitable ergodic theorem holds; cf. appendix A) quantities for the density and the velocity fields (Goroff et al. 1986; Fry 1984). Tree diagrams are associated to leading order contributions, one-loop diagrams to next-to-leading order and so on. We briefly summarize here the main results achieved so far.

One-point statistics

Let us consider the density fluctuation in a given point of space and compute its connected moments $\langle\delta^p\rangle_c$. Since $\langle\delta\rangle = 0$, the simplest non-trivial moment to analyse is the variance $\sigma^2 = \langle\delta^2\rangle$. Eulerian perturbation theory gives $\sigma^2 = \sigma_\ell^2[1 + s^{(1)}\sigma_\ell^2 + s^{(2)}\sigma_\ell^4 + \ldots]$, where $s^{(i)}$ denotes the contribution coming from graphs containing i-loops. The linear term, which is related to the linear power spectrum, $P_\ell(k)$, through

$$\sigma_\ell^2(t) = \frac{D_+^2(t)}{2\pi^2}\int_0^\infty k^2 P_\ell(k)dk\ , \tag{2.48}$$

acts as a sort of coupling constant. Note that $P_\ell(k) = AT^2(k)P_{\rm in}(k)$, where $P_{\rm in}$ denotes the primordial power spectrum of growing-mode density fluctuations and A is its normalization constant that has to be fixed empirically as described in the next section. For scale-free power spectra, $P_\ell(k) \propto k^n$ with $-2 \leq n \leq 2$, $s^{(1)} \simeq 1.82$ and weakly depends on the spectral index (Scoccimarro & Frieman 1996a; note that for $n > -1$ a cutoff in the spectrum at small scales is needed to assure the convergence of $s^{(1)}$). Thus, loop corrections begin to dominate over tree level contributions when $\sigma_\ell^2 \sim 0.5$ and perturbations theory breaks down. For the variance of θ this happens only for $\sigma_\ell^2 \simeq 1$ (Scoccimarro & Frieman 1996a).

A convenient way to measure the departure from Gaussianity due to dynamical evolution is to introduce the S_p and the T_p parameters ($p \geq 3$) defined by rescaling the one-point cumulants (Goroff et al. 1986) [5]

$$S_p \equiv \frac{\langle\delta^p\rangle_c}{\langle\delta^2\rangle_c^{p-1}}\ , \qquad T_p \equiv \frac{\langle\theta^p\rangle_c}{\langle\theta^2\rangle_c^{p-1}} \tag{2.49}$$

Since the S_p values are constructed directly from one-point cumulants, each of these parameters contains statistical information independent of the others. For example, S_3 is a measure of the skewness of the probability distribution (i.e. its asymmetry with respect to the mean value), whereas S_4 is a measure of the kurtosis (i.e. the extent to which the distribution is concentrated or "peaked" about the mean).

Perturbation theory gives $S_p = S_p^{(0)} + S_p^{(1)}\sigma_\ell^2 + S_p^{(2)}\sigma_\ell^4 + \ldots$, where the superscript (i) again denotes the contibution from graphs with i-loops. In the Einstein-de Sitter model, to lowest order the S_p parameters are constants and indepedent of the normalization of the linear power spectrum. In particular, one gets: $S_3^{(0)} = 34/7 \simeq 4.86$, $S_4^{(0)} = 60712/1323 \simeq 45.89$, $T_3^{(0)} = 26/7 \simeq 3.71$ and $T_4^{(0)} = 12088/441 \simeq 27.41$ (Peebles 1980; Goroff et al. 1986; Bernardeau 1994a). These results have been generalized to different cosmological models by Bouchet et al. (1992; who used a Lagrangian approach), Catelan et al. (1995) and Kamionkowski & Buchalter (1998). In general the dependence of the $S_p^{(0)}$ parameters on the value of cosmological parameters is extremely weak. For example, if $\Lambda = 0$, $S_3^{(0)} = (34/7) + (6/7)(\Omega^{-2/63} - 1)$ (Bouchet et al. 1992). Conversely, the T_p parameters and the peculiar velocity PDF show a stronger temporal dependence $T_p^{(0)} \propto [f_+(\Omega,\Lambda)]^{2-p}\,T_p^{(0)}(\Lambda = 0, \Omega = 1)$ (Bernardeau 1994).

The whole hierarchy of $S_p^{(0)}$ and $T_p^{(0)}$ parameters in a flat matter dominated model have been derived by Bernardeau (1992).

[5] Note that the definition of the S_p parameters given in the main text differs from the standard one in statistics where the skewness and kurtosis factors are given by $s_p = \langle\delta^p\rangle_c/\langle\delta^2\rangle_c^{p/2}$ (e.g. Kendall & Stuart 1977).

The coarse-grained density field

Astronomical observations and N-body simulations deal with discrete units (e.g. galaxies or particles) instead of a continuum. Smoothing over a finite volume is therefore required in order to compare data and theoretical results based on a fluid-type description of the cosmological mass distribution. It is then of extreme importance to study the statistical properties of the density fluctuation field at some resolution scale R_f. This is introduced by convolving $\delta(\mathbf{x},t)$ by some filter function $W(|\mathbf{x}'-\mathbf{x}|,R_f)$,

$$\delta(\mathbf{x},t;R_f) = \int d\mathbf{x}'\, W(|\mathbf{x}'-\mathbf{x}|,R_f)\,\delta(\mathbf{x}',t) = \frac{1}{(2\pi)^3}\int d\mathbf{k}\,\widetilde{W}(kR_f)\,\tilde{\delta}(\mathbf{k},t)\,\mathrm{e}^{-i\mathbf{k}\cdot\mathbf{x}}\ , \tag{2.50}$$

where $\widetilde{W}$ is the Fourier transform of the filter. At each point $\mathbf{x}$, the smoothed field represents the weighted average of $\delta(\mathbf{y},t)$ over a spherical region of characteristic dimension R_f centred in $\mathbf{x}$. The most commonly used smoothing kernels are the top-hat filter $W_{\rm TH}(|\mathbf{x}|,R_f) = 3\,\Theta(R_f-|\mathbf{x}|)/4\pi R_f^3$, where $\Theta(x)$ is the Heaviside step function, and the Gaussian one $W_{\rm G}(x,R_f) = (2\pi R_f^2)^{-3/2}\,\exp(-x^2/2R_f^2)$. [6] It can be easily shown that the power-spectrum of the coarse-grained density field is given by the product $\widetilde{W}^2(kR_f)P(k,t)$, so that the corresponding variance is

$$\sigma^2(t,R_f) \equiv \langle\delta^2(\mathbf{x},t;R_f)\rangle = \frac{1}{2\pi^2}\int k^2\widetilde{W}^2(kR_f)P(k,t)dk\ . \tag{2.51}$$

It is commonly assumed that even when the density and velocity fields are highly non-linear on small-scales (i.e. $\sigma^2(t,R_{\rm small}) \gg 1$), one can still use the perturbative approach to accurately describe the evolution of these fields on large scales (i.e. $\sigma^2(t,R_{\rm large}) \ll 1$, with $R_{\rm large} \gg R_{\rm small}$). This Ansatz has not been proven theoretically but N-body simulations seem to confirm its validity for a restricted set of linear power spectra (e.g. Jain & Bertschinger 1994; Baugh, Gaztañaga & Efstathiou 1995; Heavens, Matarrese & Verde 1998). We will discuss extensively this issue in §3.2, showing that the assumption is expected to hold for the most commonly used spectra like those that characterize the various CDM variants. Note that this Ansatz has enormous consequences. For instance, linear theory which we have developed to treat the evolution of density fluctuations at early times now can be applied at any time provided that the density field is smoothed on a sufficiently large scale. Similarly, the transition between linear and non-linear regimes can be associated to a particular length scale instead than to a phase of cosmic history. Since the perturbative approach is expected to break down when fluctuations are of the same order of their background value, it is customary to define the scale of non-linearity at time t, $R_{\rm nl}(t)$, by $\sigma(t,R_{\rm nl}) = 1$.

The statistical properties of the smoothed density field are sensitive to the filtering kernel used. The whole series of the S_p and T_p parameters to leading order in perturbation theory has been analytically derived by Bernardeau (1994b) adopting a top-hat window function. His results can be expressed in terms of the logarithmic derivatives of the linear variance with the smoothing radius, $\gamma_p(R_f) = d^p\log\sigma_\ell^2(R_f)/d(\log R_f)^p$; for example one has $S_3^{(0)}(R_f) = (34/7)+\gamma_1$, $S_4^{(0)}(R_f) = (60712/1323)+(62/3)\gamma_1+(7/3)\gamma_1^2+(2/3)\gamma_2$. Comparison with numerical simulations of clustering growth showed that these perturbative results are robust on scales corresponding to $\sigma \lesssim 1$ (e.g. Juszkiewicz, Bouchet & Colombi 1993; Bernardeau 1994a; Baugh, Gaztañaga &

[6] We then have

$$\widetilde{W}_{\rm TH}(y) = 3\frac{\sin(y)-y\cos(y)}{y^3}\ ,$$

$$\widetilde{W}_{\rm G}(y) = \exp\left(-\frac{y^2}{2}\right)\ .$$

Efstathiou 1995). Bernardeau (1994b) also showed that the $S_p^{(0)}$ and $T_p^{(0)}$ parameters assume the value corresponding to the unsmoothed case when the filtering is done at fixed mass instead that at fixed radius. Therefore the γ_p terms appear because the filtering process in Eulerian space mixes different mass scales. Knowledge of the whole hierarchy of cumulants allows the computation of the one-point probability density function (PDF) for δ and θ in good agreement with N-body simulation even when $\sigma \sim 1.5$ (Bernardeau & Kofman 1995). A rough analytic approximation for the density PDF is a lognormal distribution (Coles & Jones 1991). Gaussian smoothing has instead been considered by Catelan & Moscardini (1994) and Lokas et al. (1995): in this case the $S_p^{(0)}$ parameters can be expressed in terms of hypergeometric functions.

Normalization of the linear power spectrum

Even though inflationary models are able to specify the shape of the primordial power spectrum of density fluctuations, they give extremely model dependent predictions for its overall amplitude, A. An empirical approach, based on observational data, is then required to fix A in the framework of selected models for structure formation. Different possibilities are available.

- Analysing the Harvard-Smithsonian Center for Astrophysics (CfA) galaxy redshift survey, Davis & Peebles (1983) found that the rms fluctuation in galaxy counts measured in spheres of radius $8\,h^{-1}$Mpc is equal to one. One can then directly relate this result to the mass density fluctuations and fix the amplitude of the linear power spectrum by requiring that $\sigma_8 \equiv \sigma_\ell(t_0, R_f = 8\,h^{-1}\mathrm{Mpc}) = 1$ where top-hat smoothing is adopted. However, as we will show in chapters 4 and 6, the statistical properties of the galaxy distribution are not necessarily equivalent to those of the underlying matter distribution. Therefore this normalization method is somewhat arbitrary and, at present, has mainly historical value.

- The discovery of temperature anisotropies in the CMB has furnished an alternative procedure: assuming these fluctuations are due to scalar perturbations in the metric (i.e. to irregularities in the gravitational potential), it is then possible to adjust the spectrum normalization by using linear theory to match the observed data (e.g. Bunn & White 1997). The one ambiguity of this approach is that part of the temperature anisotropies may actually be caused by long-wavelength tensor perturbations instead (e.g. Krauss & White 1992; Lucchin, Matarrese & Mollerach 1992). Even though this method constrains the amplitude of the linear power spectrum on scales ~ 1000 Mpc, it is customary to express its results in terms of σ_8.

- A third observational result that can be used to normalize the power spectrum concerns the observed number density of rich clusters of galaxies in the local universe. Analytical methods (Press & Schechter 1974; see chapter 4), cross-checked with N-body simulations, are usually adopted to relate the primordial amplitude of density fluctuations (again in the form of σ_8) to the mass distribution of present day clusters (e.g. Eke, Cole & Frenk 1996).

A reliable spectrum has to match both the CMB and the cluster normalizations (since the two methods test amplitudes on very different length scales, this can be considered as a rough constraint on the spectral shape).

Two-point statistics

The evolution of the simplest two-point statistics of the mass density field in real space (the correlation function $\xi(x_{12}, t) \equiv \langle \delta(\mathbf{x}_1, t)\delta(\mathbf{x}_2, t)\rangle$ with $x_{12} = |\mathbf{x}_2 - \mathbf{x}_1|$) and in Fourier space

$(\langle\tilde{\delta}(\mathbf{k},t)\tilde{\delta}(\mathbf{k}',t)\rangle = (2\pi)^3\delta_D(\mathbf{k}+\mathbf{k}')P(k,t)$ with $P(k)$ the power spectrum) has been widely discussed in the literature. Perturbative analyses give $P(k,t) = P^{(0)}(k,t) + P^{(1)}(k,t) + \ldots$ with $P^{(0)}(k,t) \equiv D_+^2(t)P_\ell(k)$ and

$$P^{(1)}(k,t) = 2\int [F_2^{(s)}(\mathbf{k}-\mathbf{q},\mathbf{q})]^2\, P^{(0)}(|\mathbf{k}-\mathbf{q}|,t)\, P^{(0)}(q,t)\, d^3q + 6\int F_3^{(s)}(\mathbf{k},\mathbf{q},-\mathbf{q})\, P^{(0)}(k,t)\, P^{(0)}(q,t)\, d^3q\,. \tag{2.52}$$

Non-linear corrections including the full contributions up to two loops and the most important terms at small k in graphs with three and four loops have been computed by Fry (1984). Note that, since $P^{(0)}(k,t)$ keeps its shape and simply grows by an overall amplitude as time goes on, higher order terms are necessary to study how power is transferred between large and small scales. The evolution of the mass autocorrelation function is then obtained by Fourier transform. The effect of mode-coupling on the two-point statistics has been studied by many authors using higher-order than linear terms in the perturbation expansion. Juskiewicz, Sonoda & Barrow (1984) computed the second-order contribution to $\xi(r)$ for an exponentially smoothed linear spectrum $P(k) \propto k^2$, finding that non-linear interactions among long wavelength modes act as a source for short λ perturbations. As a matter of fact, they found a substantial decrease of the characteristic scale of clustering with the evolution. However Suto & Sasaki (1991) and Makino, Sasaki & Suto (1992), analysing exponentially filtered scale-free spectra, found that second-order effects can either suppress or enhance the growth of perturbations on large scales, depending on the shape and the amplitude of the fluctuation spectrum. In particular Makino, Sasaki & Suto (1992), modelling a CDM spectrum with two different power laws, concluded that the effects of mode-coupling are generally very small and completely negligible on scales $r \gtrsim 20\,h^{-1}$Mpc. The second-order correction to the "true" linear CDM spectrum has been calculated by Coles (1990) who computed also the respective correlation function. The results show that, for moderate evolution, the large-scale distortions are of no importance, while later (for $\sigma_8 \gtrsim 1$) non-linear effects can increase the clustering strength on scales $r > 35\,h^{-1}$Mpc; for example, the first zero crossing of $\xi(r)$ can be significantly shifted with respect to linear predictions. Similar results were obtained by Baugh & Efstathiou (1994) who also found good agreement with the output of numerical simulations. However, Jain & Bertschinger (1994) pointed out that the perturbative approach is able to reproduce the N-body outcomes only at early times ($\sigma_8 \lesssim 0.5 - 1$). We refer the reader to chapter 3 for a detailed analysis of these issues performed using a different approach.

Three-point statistics

To leading order in Eulerian perturbation theory, the three-point correlation function, $\xi_3(x_{12}, x_{13}, x_{23}, t)$ (see appendix A for a precise definition), has a hierarchical behaviour, i.e. can be expressed through a linear combination of binary products of the two-point function as follows (see also appendix A)

$$\xi_3(x_{12},x_{13},x_{23}) = \mathcal{Q}(\mathbf{x}_{12},\mathbf{x}_{13})\xi(x_{12})\xi(x_{13}) + \mathcal{Q}(\mathbf{x}_{12},\mathbf{x}_{23})\xi(x_{12})\xi(x_{23}) + \mathcal{Q}(\mathbf{x}_{13},\mathbf{x}_{23})\xi(x_{13})\xi(x_{23})\,. \tag{2.53}$$

To remove the dependence on the overall size of the triangle $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3$ and examine the residual shape dependence, it is convenient to look at

$$\mathcal{Q}_{\rm eff} = \frac{\xi_3(x_{12},x_{13},x_{23})}{\xi(x_{12})\xi(x_{13}) + \xi(x_{12})\xi(x_{23}) + \xi(x_{13})\xi(x_{23})}\,, \tag{2.54}$$

(actually, the size dependence is totally cancelled only in the case of scale free clustering). Note that $\mathcal{Q}_{\rm eff} \to S_3/3$ when $x_{12}, x_{13}, x_{23} \to 0$.

Switching to Fourier space, we introduce the bispectrum $B(k_1, k_2, k_3, t)$ through the relation

$$\langle \tilde{\delta}(\mathbf{k}_2, t)\, \tilde{\delta}(\mathbf{k}_2, t)\, \tilde{\delta}(\mathbf{k}_3, t) \rangle \equiv (2\pi)^3 \delta_D(\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3)\, B(k_1, k_2, k_3, t) \; . \tag{2.55}$$

Note that, as a consequence of translational and rotational invariance, B is defined only for configurations that form closed triangles (i.e. $\mathbf{k}_1 + \mathbf{k}_2 + \mathbf{k}_3 = 0$) and can be parametrized by any three quantities that define the shape of these triangles (such as, for example, the moduli k_1, k_2 and k_3).

The perturbative expansion for the bispectrum reads $B(k_1, k_2, k_3, t) = B^{(0)}(k_1, k_2, k_3, t) + B^{(1)}(k_1, k_2, k_3, t) + \ldots$. where the tree level term is

$$\begin{aligned} B^{(0)}(k_1, k_2, k_3, t) = 2[F_2^{(S)}(\mathbf{k}_1, \mathbf{k}_2) P^{(0)}(k_1, t) P^{(0)}(k_2, t) + F_2^{(S)}(\mathbf{k}_1, \mathbf{k}_3) P^{(0)}(k_1, t) P^{(0)}(k_3, t) \\ + F_2^{(S)}(\mathbf{k}_2, \mathbf{k}_3) P^{(0)}(k_2, t) P^{(0)}(k_3, t)] \; . \end{aligned} \tag{2.56}$$

The reader is referred to Scoccimarro et al. (1998) for an analytic representation of $B^{(1)}$. In analogy with the real space case, it is convenient to remove the dependence of the bispectrum on the size of the triangle $\mathbf{k}_1, \mathbf{k}_2, \mathbf{k}_3$ by defining the effective bispectrum amplitude (again, this works perfectly only for scale free power spectra)

$$Q = \frac{B(k_1, k_2, k_3, t)}{P(k_1, t)P(k_2, t) + P(k_1, t)P(k_3, t) + P(k_2, t)P(k_3, t)} \; . \tag{2.57}$$

It follows that its lowest order term in perturbation theory, $Q^{(0)}$, is independent of the normalization of the linear power spectrum and shows a very weak temporal dependence (Fry 1984). For the particular case of equilateral configurations ($k_1 = k_2 = k_3$ and $\mathbf{k}_i \cdot \mathbf{k}_j / k_i k_j = -1/2$ for all pairs), $Q^{(0)} = (1/4)[1 - 3E/D^2]$ independently of the spectral index. This reduces to $Q^{(0)} = 4/7$ in the Einstein-de Sitter universe.

Expressions for the four-point functions at tree level are given by Fry (1984).

On scales larger than the correlation length of the density field x_δ, tree level perturbation theory has shown to be sufficiently successful in reproducing the power spectrum and the bispectrum extracted from numerical simulations (Fry, Melott & Shandarin 1992, 1993; Jain & Bertschinger 1994; Scoccimarro et al. 1998). It is then interesting to check if higher-order corrections can extend the validity domain of perturbative calculations. The results show that (Scoccimarro & Friemann 1996b; Scoccimarro 1997; Scoccimarro et al. 1998)

- for scale-free linear spectra, the integrals that define the corrective terms converge to a finite value only if $-3 < n < -1$. Since for $n \geq -1$ the relative power on small scales is large, this probably is a signature of the breakdown of the assumption that the large-scale evolution can be calculated perturbatively even when there are non-linear fluctuations on small scales.

- When the one-loop terms are convergent they improve the agreement with N-body simulations on scales larger than x_δ with respect to tree-level computations.

- There is some evidence for the existence of a critical power index $n_c \simeq -1.4$. In fact, in one-loop calculations,

  - for $n \gtrsim -1.4$ power grows at a slower rate than predicted by linear theory and the configuration dependence of the bispectrum is reduced with respect to tree level calculations;
  - for $n \lesssim -1.4$ power grows at a faster rate than predicted by linear theory and the configuration dependence of the bispectrum is enhanced with respect to tree level calculations.

In practice, n_c behaves like a fixed point: the power index of density perturbation in the weakly non-linear regime always approaches n_c. Similar results, with $n_c \simeq -1$, have been found numerically by Bagla & Padmanabhan (1997) who investigated the non-linear regime $1 < \delta(R_f) < 200$. According to their interpretation of the phenomenon, power cascades from larger to small scales (which is dominant over inverse cascades in gravitational clustering) leads to equipartition of energy at small scales thus giving $n \sim -1$. Alternatively one can think to a previrialization effect, i.e. by increasing the spectral index small-scale power becomes more effective in slowing down the collapse of high density regions and to disrupt coherent structures and flows. All these results have anyway to be considered as preliminary and the issue about a fixed point in gravitational clustering certainly deserves additional investigation.

- Considering smaller and smaller scales than the correlation length, non-linear interactions in numerical simulations progressively reduce the shape dependence of the effective bispectrum parameter so that, in the deeply non-linear regime, $Q \simeq \text{const} = 3/(3+n)$ with n the spectral index (Fry, Melott & Shandarin 1992). One-loop calculations are not able to reproduce this flattening but continue to give reliable results in the mildly non-linear regime for collinear configurations ($\mathbf{k}_1 \cdot \mathbf{k}_2/k_1k_2 = \pm 1$). The saturation at small scales in the exact dynamics is probably due to randomization of the density and velocity gradients in multi-streaming regions.

2.3 Lagrangian perturbation theory

Let us again consider a fluidodynamic approach to describe the evolution of a set of collisionless, self-gravitating particles in an expanding universe. This time, however, we adopt a Lagrangian description of the system; i.e. we follow the motion of each fluid element by writing its actual (Eulerian) comoving position, $\mathbf{x}$, at time t as the sum of its initial (Lagrangian) comoving position, $\mathbf{q}$, plus a displacement:

$$\mathbf{x}(\mathbf{q}, t) = \mathbf{q} + \mathbf{S}(\mathbf{q}, t). \tag{2.58}$$

The independent variable $\mathbf{q}$ is simply a label that identifies a given patch of fluid and the displacement vector field $\mathbf{S}(\mathbf{q}, t)$ describes the effect of density perturbations on its trajectory. When fluctuations are small, one expects that the mapping $\mathbf{q} \to \mathbf{x}$ is one-to-one. From the mathematical point of view, this means that the Jacobian determinant $J \equiv \|\partial \mathbf{x}/\partial \mathbf{q}\|$ does not vanish and the mapping is reversible. On the other hand, while structure forms, it is possible that many fluids elements (obviously coming from different initial positions) reach the same Eulerian position, $\mathbf{x}$, at a given time. When this happens, the mapping of Lagrangian space onto Eulerian space becomes singular and regions with infinite density (caustics) form at $\mathbf{x}$ (e.g. Shandarin & Zel'dovich 1989). In Eulerian space, caustics have complex geometrical structures (Arnold, Shandarin & Zel'dovich 1982) which can be studied in the framework of the catastrophe theory (Arnold 1986). Soon after orbit crossing, extended multi-stream regions develop.

It is clear that Eulerian perturbation theory, which is based on the assumption $\delta \ll 1$, has to break down well before caustic formation. The main advantage of the Lagrangian framework over its Eulerian counterpart is that small perturbations of the particle paths carry a large amount of non-linear information about the corresponding Eulerian variables. In other words: the Lagrangian picture is intrinsecally non-linear in the density field.

Following Zel'dovich (1970), the evolution of the whole system can be discussed in terms of the only dynamical field $\mathbf{S}$ and its derivatives. For example, the Euler equation corresponds to

$$\frac{d^2\mathbf{S}(\mathbf{q},t)}{dt^2} = \frac{\mathbf{g}(\mathbf{q},t)}{a} - 2\frac{\dot{a}}{a}\frac{d\mathbf{S}(\mathbf{q},t)}{dt} , \tag{2.59}$$

where the time derivative is taken at fixed $\mathbf{q}$; i.e. $d/dt \equiv \partial/\partial t|_x + d\mathbf{x}/dt \cdot \nabla_x$ is the usual convective derivative which follows the fluid element. On the other hand, assuming that the mass distribution in Lagrangian space is uniform, mass conservation can be recast in terms of J as

$$1 + \delta[\mathbf{x}(\mathbf{q},t),t] = \frac{1}{J(\mathbf{q},t)} . \tag{2.60}$$

This form of the continuity equation can be easily generalized to the case of multi-streaming by writing $1 + \delta(\mathbf{y},t) = \sum_i \frac{1}{J(\mathbf{q}_i,t)}$, where the sum is taken over all the fluid elements for which $\mathbf{x}(\mathbf{q}_i,t) = \mathbf{y}$. It is often useful to single out the effect of perturbations in J by defining the deformation tensor $\Delta_{\alpha\beta} = \partial S_\alpha/\partial q_\beta$; obviously $J = I + \Delta$ where I denotes the (3×3) identity matrix. Note that Δ is symmetric if and only if the displacement field $\mathbf{S}$ is irrotational in Lagrangian space. The mathematical description of the fluid is completed by adding the Poisson equation to the system formed by equations (2.59) and (2.60); however a fourth equation is sometimes considered to require that $\mathbf{S}$ is irrotational in Eulerian space. Since the translation of ∇_x into a combination of operators that differentiate with respect to $\mathbf{q}$ introduces rather cumbersome formulae that we will not use in the following, we do not report here the Lagrangian form of these last two equations. The Lagrangian perturbative approach looks for solutions of the fluid equations assuming that the displacement field $\mathbf{S}(\mathbf{q},t) = \mathbf{S}^{(1)}(\mathbf{q},t) + \mathbf{S}^{(2)}(\mathbf{q},t) + \ldots$ and the determinant of the Jacobian determinant $J(\mathbf{q},t) = 1 + J^{(1)}(\mathbf{q},t) + J^{(2)}(\mathbf{q},t) + \ldots$, where $\mathbf{S}^{(n)}$ and $J^{(n)}$ are both $\mathcal{O}(\epsilon^n)$ with $\epsilon \ll 1$. Solutions up to the third order have been obtained (Buchert 1989; Moutarde et al. 1991; Bouchet et al. 1992; Buchert 1993; Catelan 1995)

2.4 Non-perturbative methods

When the fluctuations become strongly non-linear, perturbation theory breaks down, and one has to resort to N-body simulations to study the subsequent evolution. On the other hand, several non-linear approximation schemes have been proposed in the literature which allow analytic or numeric calculations beyond the domain of linear perturbation theory. These are approximations in the sense that they replace the Poisson equation by a given Ansatz which is true only in linear theory for the exact dynamics and therefore they are neither exact nor asymptotic to the exact solution.

2.4.1 The spherical top-hat model

The simplest configuration for which we are able to analytically describe the complete evolution is a spherically symmetric density perturbation in an otherwise uniform background:

$$\rho(r,t) = \rho_b(t)\,[1 + \delta(r,t)] \tag{2.61}$$

where the initial density contrast $\delta(r, t_{\rm in})$ is a non–increasing function of r (note that in this section we use physical coordinates instead of comoving ones).

Let us consider the motion of a thin shell of matter labelled by its actual radial coordinate r, the Gauss law gives: [7]

$$\frac{d^2r}{dt^2} = -\frac{GM}{r^2} \tag{2.62}$$

where M denotes the mass enclosed by the shell,

$$M = \frac{4\pi}{3} r^3 \rho_b (1+\bar{\delta}) \qquad \bar{\delta}(r,t) = \frac{3}{r^3}\int_0^r \delta(y) y^2 dy\,. \tag{2.63}$$

Assuming that different shells cannot cross each other during the evolution, M becomes a constant of motion. In this case the first integral of equation (2.62) express the conservation of energy:

$$\frac{1}{2}\dot{r}^2 - \frac{GM}{r} = E\,. \tag{2.64}$$

Moreover, if we admit that initially the perturbation is expanding just like the background (i.e. no peculiar velocity), we have $\dot{r}_{\rm in} = H_{\rm in} r_{\rm in}$ and $GM/r_{\rm in} = \Omega_{\rm in} H_{\rm in}^2 r_{\rm in}^2 (1+\bar{\delta}_{\rm in})/2$ which give:

$$E = \frac{1}{2} H_{\rm in}^2 r_{\rm in}^2 \Omega_{\rm in} \left[\frac{1}{\Omega_{\rm in}} - (1+\bar{\delta}_{\rm in})\right]\,. \tag{2.65}$$

Equations (2.64) and (2.65) show that the shell under analysis can reverse its expanding motion and start contracting only if $E < 0$, i.e. if

$$\bar{\delta}_{\rm in} > \frac{1}{\Omega_{\rm in}} - 1\,. \tag{2.66}$$

Thus, while in a flat or close cosmological model every positive density perturbation can collapse, in "open" models $\bar{\delta}_{\rm in}$ must exceed a threshold in order to be able to form a bound object.

The time evolution of a collapsing shell is obtained integrating equation (2.62), in parametric form we get: [8]

$$r = A(1-\cos\theta) \qquad t = B(\theta - \sin\theta) \qquad A^3 = GMB^2\,. \tag{2.67}$$

The corresponding $\delta(t)$ can be computed by combining this result with the formulae describing the evolution of the mass density in the background universe (e.g. Peebles 1980). In brief, the spherical perturbation starts evolving linearly

$$\delta(t) \simeq \frac{3}{5}\delta_{\rm in}\frac{D_+(t)}{D_+(t_i)} \qquad t \gtrsim t_{\rm in} \tag{2.68}$$

(we assume, for simplicity, that $\Omega_{\rm in} \simeq 1$ and the universe is matter dominated at $t_{\rm in}$) and its expansion rate is progressively reduced with respect to the background until (at $\theta = \pi/2$) it reaches its largest radius $r_{\rm ta}$. The value of $r_{\rm ta}$ as a function of the initial conditions can be obtained from

[7] To simplify the discussion we assume here $\Lambda = 0$. In the most general case, an additional term $\Lambda r/3$ appears in the rhs of eq. (2.61).

[8] Shells that expand for ever instead follow

$$r = A(\cosh\theta - 1) \qquad t = B(\sinh\theta - \theta) \qquad A^3 = GMB^2\,.$$

equations (2.64) and (2.65) by imposing that at the moment of maximum expansion (turnaround) the shell has vanishing kinetic energy:

$$\frac{r_{\rm ta}}{r_{\rm in}} = \frac{1+\bar{\delta}_{\rm in}}{\bar{\delta}_{\rm in} - (\Omega_{\rm in}^{-1} - 1)} \,. \tag{2.69}$$

This result can be used to express the constants A and B as functions of $\bar{\delta}_{\rm in}$ and $\Omega_{\rm in}$ (Peebles 1980; Padmanabhan 1993). After the turnaround the shell starts contracting at an ever increasing rate and will finally collapse to a singularity (at $t = t_{\rm ta}$, corresponding to $\theta = 2\pi$). However, in a real situation, the approximation of assuming radial orbits and negligible random velocities breaks down soon after the beginning of the contracting phase. We expect large fluctuations in the gravitational potential, acting in a dynamical time scale, to lead the collisionless material component to quasi-equilibrium (violent relaxation, Lynden-Bell 1964). At the same time, shocks can reheat the gaseous (baryonic) component.

The final state is constituted by a virialized lump of matter (neither expanding nor contracting). The radius of this collisionless "halo" can be extimated starting again from energy conservation. When virial equilibrium is settled, the total energy of a given shell equals half of its potential energy while at turnaround all the energy is potential. Since for a sphere $U \propto -GM/r$, we get:

$$r_{\rm vir} = \frac{r_{\rm ta}}{2} \qquad \rho_{\rm vir} = 8\rho_{\rm ta} \,. \tag{2.70}$$

In the matter dominated Einstein-de Sitter universe, assuming that virialization happens at $t = 2t_{\rm ta}$, one then has $1 + \delta_{\rm vir} = 18\pi^2 \simeq 178$.

It is possible to test the accuracy of the linear perturbative approach by comparing its predictions for the evolution of δ in a homogeneus spherical distribution (top-hat profile) to the exact solution. Using equation (2.68), it can be shown (e.g. Peebles 1980; Padmanabhan 1993) that at turnaround $\delta_{\rm lin}(t_{\rm ta}) = (3/5)(3\pi/4)^{2/3} \simeq 1.062$ which strongly underestimates the exact value $\delta = (3\pi/4)^2 - 1 \simeq 4.552$. At the final collapse $\delta_{\rm lin} = (3/5)(6\pi/4)^{2/3} \simeq 1.686$ with a formally diverging error. In chapter 4 we will discuss how this results can be used to build an approximate algorithm for computing the mass distribution of collapsed objects in a bottom-up cosmological scenario (see also Press & Schechter 1974 and Bond et al. 1991).

The evolution of a spherical perturbation in Friedmann models with a positive cosmological constant has been studied by Peebles (1984), Lahav et al. (1991), Martel (1991). Their results are briefly summarized in §5.2.2.

The spherical model has been extensively used in the literature to discuss the secondary infall of material onto a collapsed core (Gott 1975; Gunn 1977; Fillmore & Goldreich 1984; Bertschinger 1985; Hoffmann & Shaham 1985; White & Zaritsky 1992). If the binding energy of each infalling shell is due to the bound object, the resulting density profile of the final halo is $\rho \propto r^{-9/4}$. In general, when the binding energy depends on the mass of the shell through $E(M) \propto M^{\frac{2}{3}-\epsilon}$, the density distribution of the halo comes out to be $\rho \propto r^{-9\epsilon/(1+3\epsilon)}$.

It is important to stress here that one does not expect the primordial perturbations to be highly symmetric, so the spherical model should be considered only on a qualitative level. Indeed, non linear gravitational dynamics of collisionless matter is known to amplify any asymmetry in the density distribution (Lin, Mestel & Shu 1965; Zel'dovich 1970). For example, an homogeneous ellipsoidal perturbation in a otherwise uniform background evolves through uniform ellipsoidal configurations of increasing eccentricity (strictly speaking this is true only assuming that the density distribution outside the ellipsoid remains homogeneous; numerical simulations show that this approximation is sufficiently accurate, e.g. White 1996). Finally, when the length of its smallest principal axis is vanishing, the structure collapses into a flat elliptical pancake (White & Silk 1979; More, Heavens & Peacock 1986).

2.4.2 The Zel'dovich approximation

Let us adopt $D \equiv D_+(t)$ as time variable and consider the comoving peculiar "velocity" $\mathbf{w} \equiv d\mathbf{x}/dD = \mathbf{v}/a\dot{D}$. The Euler equation (2.15) can then be recast in the form

$$\frac{d\mathbf{w}}{dD} = -\frac{3}{2f_+^2 D}\left[\mathbf{w} + \frac{2f_+}{3\Omega H}\frac{1}{\dot{D}a}\frac{\nabla\phi}{a}\right] . \tag{2.71}$$

The Zel'dovich approximation (Zel'dovich 1970) is obtained by assuming that the relation $\phi_v = (2f_+/3\Omega H)\phi$, valid for the growing mode of linear velocity perturbations, holds true at every epoch. In this case, eq. (2.71) simply gives

$$\frac{d\mathbf{w}}{dD} = 0 , \tag{2.72}$$

which is easily solved in Lagrangian form: $\mathbf{w}(\mathbf{q},t) = \mathbf{w}_\circ(\mathbf{q})$. This means that fluid elements move along straight paths, $\mathbf{S}(\mathbf{q},D) = \mathbf{S}_\circ(\mathbf{q}) + D\,\mathbf{w}_\circ(\mathbf{q})$, keeping the same $\mathbf{w}$-velocity. Assuming that mass is evenly distributed when $D \to 0$, we fix the initial displacement to $\mathbf{S}_\circ(\mathbf{q}) = 0$. Moreover, in the same limit, we can approximate $\phi(\mathbf{q},t)$ with its linear version $\phi^{(1)}(\mathbf{q},t)$; considering only the growing mode one eventually gets

$$\mathbf{S}(\mathbf{q},t) = -D_+(t)\frac{\nabla\phi_+|_{\mathbf{q}}}{4\pi G} \equiv -D_+(t)\nabla\phi_S|_{\mathbf{q}} . \tag{2.73}$$

The Zel'dovich approximation (hereafter ZA) can be also extracted from the perturbative Lagrangian approach to the evolution of density fluctuations. In this case, ZA corresponds to the first order solution provided that the initial velocity field is irrotational and the initial peculiar velocity and acceleration fields are everywhere parallel.

It is often stated that ZA is kinematic, in the sense that does not take into account gravity. Actually, this is not true: in ZA the right hand side of eq. (2.72) vanishes because the gravitational term generated by density inhomogeneities exactly balances the Hubble drag term due to the uniform expansion of the universe. This generally underestimates the gravitational pull in overdense regions.

Since $\mathbf{S}$ is a potential field, the corresponding deformation tensor $\mathcal{D}_{\alpha\beta}$ is symmetric and can be locally diagonalized by going to the frame formed by its principal axes. Eq. (2.60) can then be recast in the form

$$1 + \delta_{\rm ZA}[\mathbf{x}(\mathbf{q},t),t] = \frac{1}{[1+D_+(t)\lambda_1][1+D_+(t)\lambda_2][1+D_+(t)\lambda_3]} , \tag{2.74}$$

where $\lambda_1(\mathbf{q})$, $\lambda_2(\mathbf{q})$ and $\lambda_3(\mathbf{q})$ denote the eigenvalues of Δ when $D_+ = 1$ (we assume $\lambda_1 \le \lambda_2 \le \lambda_3$).

It is evident that all the fluid elements initially located at points in which Δ has at least one negative eigenvalue will develop caustics as time goes on. Assuming that ϕ_v (and thus δ) is a Gaussian random field, Doroshkevich (1970) computed the probability distribution of λ_1, λ_2 and λ_3

$$\mathcal{P}(\lambda_1,\lambda_2,\lambda_3) = \frac{5^{5/2}27}{8\pi\sigma_\ell^6}(\lambda_1-\lambda_2)(\lambda_1-\lambda_3)(\lambda_2-\lambda_3)\exp\left[-\frac{1}{\sigma_\ell^2}\left(3\mu_1^2 - \frac{15}{2}\mu_2\right)\right] , \tag{2.75}$$

with $\mu_1 = \lambda_1+\lambda_2+\lambda_3$, $\mu_2 = \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3$ and $\mu_3 = \lambda_1\lambda_2\lambda_3$ the three invariants of the deformation tensor. It follows that:

- 42% of the Lagrangian volume (i.e. of the mass) is occupied by fluid elements that, having one negative and two positive eigenvalues, will collapse along one axis to a pancake (a flat surface).
- The same volume fraction is covered by fluid patches whose deformation tensor has one positive and two negative eigenvalues. These eventually collapse to a filament (a one-dimensional structure).
- 8% of the mass has three negative eigenvalues and will be progressively concentrated in knots (point-like configurations).
- The remaining 8% is given by points in which $\mathcal{D}$ has only positive eigenvalues. The corresponding fluid elements will be subject to a continuos expansion and rarefaction.

The Zel'dovich approximation uses the exact Euler and continuity equations, but replaces the Poisson equation with a relation between ϕ_v and ϕ that is valid only in the linear regime. It is not surprising then that ZA gives the correct evolution of $\delta(\mathbf{x},t)$ at early times (when $D_+ \ll 1$): $1+\delta[\mathbf{x}(\mathbf{q},t),t] \simeq 1 - D_+(t)(\lambda_1+\lambda_2+\lambda_3) = 1 - \nabla\cdot\mathbf{S} = 1 + D_+(t)\delta_+(\mathbf{x})$. Later on, however, the Zel'dovich solution progressively departs from the exact one. This discrepancy can be quantified by inserting the Zel'dovich Ansatz for ϕ into the Poisson equation, obtaining $1+\delta_{\rm dyn} = 1/[1 + D_+(t)(\lambda_1+\lambda_2+\lambda_3)]$. This relation shows that ZA is exact for one-dimensional perturbations, since, in this case, it coincides with eq. (2.74). Nusser et al. (1991) showed that $\delta_{\rm dyn}$ tends to overestimate the exact result while $\delta_{\rm ZA}$ generally underestimates it.

N-body simulations show that ZA gives a reliable description of the density field before caustic formation, whereas its accuracy gets worse and worse later. This is because the "Zel'dovich fluid" is a system with infinite memory: even after the intersection of two trajectories, the motion of the particles is determined by their initial conditions according to equation (2.73). The lack of self–gravity between intersecting streams causes the forming structure to be rapidly washed out. This is a severe problem especially in hierarchical models of structure formation, where caustics appear early on small scales causing ZA to become soon inaccurate. Nevertheless Coles, Melott & Shandarin (1993) showed that a modified version of ZA, the "truncated" ZA, obtained by smoothing the initial conditions, is able to reproduce with good accuracy the density distributions obtained from numerical simulations. Melott, Pellman & Shandarin (1994) found that the optimal version of the truncation procedure is accomplished by using a Gaussian window to smooth the linearly extrapolated power spectrum of the density fluctuation field $D_+^2(t)P_\ell(k)$:

$$P_T(k,t) = D_+^2(t)P_\ell(k)\exp\left[-k^2R_f^2(t)\right] \tag{2.76}$$

where the filtering radius $R_f(t)$ increases with time being related to the typical scale going non–linear. The success of this approximation can be justified by noticing that the non–linearly evolved gravitational potential resembles its smoothed linear counterpart (Pauls & Melott 1995).

Another way for improving ZA, thereby stabilizing the thickness of pancakes, consists in using artificial viscosity to mimic non-linear gravitational dynamics in the multi-stream flow regions. This is the basic idea of the adhesion approximation introduced by Gurbatov, Saichev & Shandarin (1989). From the mathematical point of view this is obtained by replacing eq. (2.72) with

$$\frac{d\mathbf{w}}{dD} = \nu\nabla^2\mathbf{w}\ . \tag{2.77}$$

This is the Burgers' equation which, when $\nu \to 0$, can be solved analytically (through an elegant geometrical construction). In this limit, the right hand side of eq. (2.77) remains finite only in

those regions where large gradients in the velocity field are present (namely, inside the multi-stream regions) and vanishes elsewhere. Therefore the adhesion approximation reduces to ZA before caustic formation. The adhesion approximation was thoroughly compared with N-body simulations and proved to be quite accurate in describing the non-linear stage. The density distribution it predicts looks much clumpier than that generated by the truncated ZA but it is still less clumpy than the output of a numerical simulation (e.g. Sathyaprakash et al. 1995).

2.4.3 The frozen flow and frozen potential approximations

In the frozen flow approximation (Matarrese et al. 1992), one takes $\phi(\mathbf{x},t) = D_+(t)\phi_+(\mathbf{x})$ and evolves the density field in the linear potential using the full continuity equation and the linearized Eulerian equation (2.23). This corresponds to assume

$$\frac{\partial \mathbf{w}}{\partial D} = 0 \,, \tag{2.78}$$

with the initial condition $\mathbf{w}(\mathbf{q}, D=0) = -D\nabla\phi_{v+}/\dot{D}a$. The idea behind this approximation is that, on a given length scale, the gravitational potential is expected to obey linear evolution much longer than the density field. This is because $\tilde{\phi}(\mathbf{k}) \propto \tilde{\delta}(\mathbf{k})/k^2$ (i.e the gravitational potential weights the long-wavelengh modes much more strongly than the density field) and in a hierarchical scenario for structure formation (e.g. CDM) the longer wavelength modes go non-linear at a later time than their shorter wavelength counterparts. Indeed, N-body simulations show that ϕ evolves much more slowly than the density field (Brainerd et al. 1993). Note, however, that only $\nabla\phi$ appears in the equations of motion. This weakens a bit the previous argument since $\nabla\phi$ scales as one higher power of k than does ϕ, thus being more sensitive to non-linear modes.

Equation (2.78) means that each fluid element upgrades istantaneously its velocity to the vector associated by linear theory to its actual position. Thus, particles have no memory of their previous state of motion. In a sense this is the opposite situation with respect to ZA, that, in $(D, \mathbf{w})$ coordinates is dominated by inertia. In the frozen flow approximation, streamlines are then kept frozen to their initial configuration, and the dynamics remains forever in the single-stream regime. As time passes matter is progressively collected near points where $\mathbf{w}(\mathbf{q}) = 0$ (that are often improperly called pancakes), that correspond to the local minima of the initial gravitational potential. However, during the formation of these structures, trajectories approach one another asymptotically. This is because particles approach their stability points more and more slowly and need infinite time to reach them. For this reason pancake-like structures always remain thin. The frozen flow approximation obviously breaks down when the scale $R_{\rm pk} = \sqrt{2}\sigma_1/\sigma_2$ (see appendix A for the definition of the spectral moments σ_i), corresponding to the average distance between the peaks of the ϕ_+ enters the non-linear regime.

Taking into account the full Euler equation instead of its linear version but still considering the linear gravitational potential, one obtains the linear (or frozen) potential approximation (Brainerd et al. 1993; Bagla & Padmanhabhan 1994). In this case

$$\frac{d\mathbf{w}}{dD} = -\frac{3}{2f_+^2 D}\left[\mathbf{w} + \frac{2f_+}{\Omega H}\frac{1}{\dot{D}a}\frac{\nabla\phi_+^{(1)}}{a}\right] \,, \tag{2.79}$$

i.e. at each time, the acceleration of a fluid element is computed by using its instantaneous velocity and the linearly extrapolated initial potential at its current position. In the linear potential approximation, shell-crossing occurs and, in the multi-streaming regime, particles oscillate around the location of the singularity. Therefore pancakes remain thin because fluid elements are confined

by the local acceleration field. Particles are then able to move along the pancakes, towards regions of lower potential, to end up in a few clumps (again located at the minima of ϕ_+).

The assumption of linear scaling of the gravitational potential is expected to break down in the regions where ϕ assumes its largest positive and negative values. In fact, for negative peaks, linear evolution predicts too shallow potential wells. On the other hand, for positive peaks, linear theory violates the mass-positivity constraint $\delta \geq -1$, thereby predicting too high amplitudes for ϕ.

At the end of this section, it is useful to compare the predictions of the various approximations we introduced so far. For example, we can use the evolution of a spherical top-hat perturbation in the matter dominated Einstein-de Sitter universe as a simple (but, because of the highly symmetric configuration, a little bit peculiar) benchmark. Taking $t_{\rm in} \to 0$ and considering $\delta_{\rm ta} \equiv D(t_{\rm ta})\delta_+$ as the relevant quantity, one gets $\delta^{(1)}_{\rm ta} = 1.5$ in ZA, $\delta^{(1)}_{\rm ta} = 3$ for the frozen flow approximation and $\delta^{(1)}_{\rm ta} = 2.06$ in the linear potential approximation. This numbers must be compared with the exact result $\delta^{(1)}_{\rm ta} = 1.06$.

Eulerian perturbation theory gives a second chance for comparing the various approximations. In practice, one computes perturbative kernels for each approximated dynamics (with the exception of the adhesion approximation) and compares resulting statistical properties to their exact counterparts in the weakly non-linear regime (all the approximations are built to give the exact linear evolution). The results show that ZA gives the most accurate results especially if the slope of the power spectrum of density perturbations is $-3 < n \leq -1$ (Grinstein & Wise 1987; Munshi & Starobinsky 1994; Scoccimarro & Frieman 1996).

Comparison with N-body simulations is used to test the approximations in the fully non-linear regime. A number of studies showed that, soon after caustics formation, the linear potential, the frozen flow and the adhesion approximations give a better description than the simple ZA which starts erasing structure through free-streaming (e.g. Brainerd et al. 1993; Coles et al. 1993). However, as expected, linear potential and frozen flow algorithms break down when the non-linear length scale, $R_{\rm nl}$, exceeds $R_{\rm pk}$. On the other hand, if the initial potential contains sufficient power on large scales to affect particle motion, the truncated Zel'dovich and the adhesion approximations give the most reliable dynamics and reproduce more closely the exact mass distribution at late epochs, when $R_{\rm pk} < R_{\rm nl} < R_\phi$, with $R_\phi = \sqrt{2}\sigma_0/\sigma_1$ the correlation length of the linear potential (Coles et al. 1993; Sathyaprakash et al. 1995). For scale free spectra this happens when $-3 \leq n \leq 1$, which is the range of interest for physically interesting spectra like the CDM one (for which $R_{\rm pk} \lesssim 1$ Mpc and $R_\phi \sim 50$ Mpc). In general, the adhesion model is more effective in dividing the mass into clumps, since the truncation procedure in ZA prevents the presence of the smallest objects. Anyway, the truncated ZA gives the right mean density both in voids and in overdense regions and more accurately predicts the location of large condensations at late times, when the scale of non-linearity is larger than R_ϕ.

2.4.4 The local Lagrangian approximation

Protogeros & Scherrer (1997) proposed a new class of dynamical approximations particularly aimed at reproducing the mildly non-linear evolution of the density PDF. The idea is that the density at a Lagrangian point $\mathbf{q}$ at time t can be approximated as a function only of time and the initial density at the same Lagrangian point, i.e.

$$1 + \delta[\mathbf{x}(\mathbf{q}, t), t] = f[\delta(\mathbf{q}, t_0), t] \,, \tag{2.80}$$

where t_0 denotes the initial time and f is an arbitrary function. In particular, inspired by ZA, they consider the following functional form

$$1 + \delta[\mathbf{x}(\mathbf{q}, t), t] = \frac{1 + \delta(\mathbf{q}, t_0)}{[1 - D_+(t)\delta(\mathbf{q}, t_0)/\alpha]^\alpha} \,. \tag{2.81}$$

The constant α takes the value $\alpha = 1$ for the planar approximation (which gives the exact gravitational evolution in one dimension), while $\alpha = 3$ corresponds to spherical collapse in ZA (i.e. with $\lambda_1 = \lambda_2 = \lambda_3$). This spherical approximation has been also used by Padmanabhan & Subramanian (1993) to investigate the evolution of the density PDF. It is expected that the actual evolution of the density field lies somewhere in between spherical and planar collapse. The choice of $\alpha = 3/2$, which has no particular physical meaning, closely mimics the hierarchical amplitudes $S_p^{(0)}$ of the tree-level perturbation theory for the exact dynamics (Bernardeau 1992). It is possible to show that eq. (2.81) automatically conserve mass, but, unless $\alpha = 1$, violates the conservation of Eulerian volume (Protogeros & Scherrer 1997). This inconsistency is related to the appearance of multi-streaming regions and can be corrected by simply re-normalizing the Eulerian density field, i.e. by defining

$$1 + \delta[\mathbf{x}(\mathbf{q}, t), t] = \frac{\langle [1 - D_+(t)\delta(\mathbf{q}, t_0)/\alpha]^\alpha \rangle}{[1 - D_+(t)\delta(\mathbf{q}, t_0)/\alpha]^\alpha} \,, \tag{2.82}$$

where the ensemble average is obviously taken in Lagrangian space. The predictions of the local Lagrangian Ansatz with $\alpha = 3/2$ have been tested against N-body simulations by Protogeros, Melott & Scherrer (1997, see also Protogeros & Scherrer 1997). Their results show that for $-3 \leq n \leq -1$ power spectra and in the weakly non-linear regime ($\sigma_\ell < 0.72$) the approximation correctly reproduces the evolved density PDF. On the other hand, for $n \geq 0$ the agreement is acceptable only in the linear regime. Since the local Lagrangian approximation reproduces, nearly exactly, all of the hierarchical amplitudes (at tree-level) generated by gravitational instability, its lack of accuracy for $n > -1$ should imply the simultaneous breakdown of the exact perturbative approach (Protogeros, Melott & Scherrer 1997). These results are, in fact, in agreement with those of Scoccimarro & Frieman (1996a), who found that one-loop corrections to the density variance diverge for $n \geq -1$ (with a logarithmic divergence in the limiting case $n = -1$).

2.4.5 Gravitational instability via local Lagrangian fluid equations

Till now, we showed how the evolution of a self-gravitating pressureless fluid can be described either by Euler equations or by geodesic equations for individual mass elements. In a series of papers, Bertschinger and collaborators proposed to analyse gravitational instability in a cold medium by using Lagrangian fluid equations (Bertschinger & Jain 1994; Bertschinger & Hamilton 1994; Hui & Bertschinger 1996). Using the conformal time τ ($d\tau = dt/a$) as time variable and converting the time derivatives from Eulerian to Lagrangian, the continuity equation (2.14) becomes

$$\frac{d\delta}{d\tau} + (1 + \delta)\theta = 0 \,, \tag{2.83}$$

where θ denotes the divergence of the velocity $\mathbf{v} = d\mathbf{x}/d\tau$. It is convenient to decompose the velocity gradient tensor into its trace and traceless symmetric and antisymmetric parts

$$\nabla_i v_j = \frac{1}{3}\theta\delta_{ij} + \sigma_{ij} + \omega_{ij} \,, \quad \sigma_{ij} = \sigma_{ji} \,, \quad \omega_{ij} = \epsilon_{ijk}\omega^k = -\omega_{ji} \,, \tag{2.84}$$

where $2\omega = \nabla \times \mathbf{v}$ is the vorticity, ϵ_{ijk} is the three dimensional fundamental antisymmetric (Levi-Civita) tensor, and dots denote derivatives with respect to τ. Taking the divergence of the Euler equation (2.15), using the Poisson relation (2.16) and eq. (2.84), and converting the time derivative from Eulerian to Lagrangian, one obtains the Raychaudhuri equation

$$\frac{d\theta}{d\tau} + \frac{\dot{a}}{a}\theta + \frac{1}{3}\theta^2 + \sigma^{ij}\sigma_{ij} - 2\omega^2 = -4\pi G\rho_b a^2\delta \,, \tag{2.85}$$

where we defined $\omega^2 \equiv \omega^i\omega_i$. Equations (2.83) and (2.85) can be combined into a second-order differential equation for the density contrast

$$\ddot{\delta} + \frac{\dot{a}}{a}\dot{\delta} = \frac{4}{3}\frac{\dot{\delta}^2}{1+\delta} + (1+\delta)\left(\frac{2}{3}\sigma^2 - 2\omega^2 + 4\pi G\rho_b a^2\delta\right) \,, \tag{2.86}$$

where the shear scalar, σ, is defined by $\sigma_{ij}\sigma^{ij} = (2/3)\sigma^2$. Note that if $\omega = 0$ the growth rate of δ is minimized for $\sigma = 0$. This has a simple geometrical interpretation: the shear increases the rate of growth of the convergence of fluid streamlines $(-\theta)$. When $\sigma = \omega = 0$, eq. (2.86) reduces to the exact equation for the evolution of the mean density in the spherical model. From these considerations Bertschinger & Jain (1994) state a *collapse theorem* for overdense growing fluctuations: a pressureless, irrotational mass element with growing mode fluctuation $\delta_{\rm in} > (3/5)(\Omega_{\rm in}^{-1} - 1)$ collapses at least as fast as a uniform spherical perturbation with the same initial δ and $\dot{\delta}$, unless it first collide with another mass element (in analogy with §2.4.1 we assume that $\Omega_i \simeq 1$ and the universe is matter dominated at $t_{\rm in}$). The fact that the non-linear density growth rate depends on the shear has two other consequences: i) initial density maxima do not strictly identify the mass elements to collapse first; ii) initially underdense mass elements may collapse. In order to obtain a complete description of gravitational collapse it is then necessary to follow the temporal behaviour of the shear.

Evolution equations for the shear σ_{ij} and vorticity follow from the traceless symmetric and antisymmetric parts of the gradient of eq. (2.15),

$$\frac{d\omega^i}{d\tau} + \frac{\dot{a}}{a}\omega^i + \frac{2}{3}\theta\omega^i - \sigma^i_j\omega^j = 0 \,, \tag{2.87}$$

$$\frac{d\sigma_{ij}}{d\tau} + \frac{\dot{a}}{a}\sigma_{ij} + \frac{2}{3}\theta\sigma_{ij} + \sigma_{ik}\sigma^k_j + \omega_i\omega_j - \frac{1}{3}\delta_{ij}(\sigma^{kl}\sigma_{kl} + \omega^2) = -E_{ij} \,, \tag{2.88}$$

where $E_{ij} \equiv \nabla_i\nabla_j\phi - (1/3)\delta_{ij}\nabla^2\phi$ is the gravitational tidal field. It is evident that to close the system of Lagrangian fluid equations the evolution of E_{ij} must be known. From the Euler and continuity equations Bertschinger & Hamilton (1994) derive

$$\frac{dE_{ij}}{d\tau} + \frac{\dot{a}}{a}E_{ij} + M_{ij} = -4\pi G\rho_b a^2\sigma_{ij} \,, \tag{2.89}$$

where parentheses around a pair of subscripts indicates symmetrization. e.g., $\sigma^k_{(i}E_{j)k} = (\sigma^k_i E_{jk} + \sigma^k_j E_{ik})/2$ and the tensor M_{ij} is defined as follows. Let us consider the mass current $\mathbf{j} \equiv \rho\mathbf{v}$ and decompose it into

$$\mathbf{j}_{\parallel} = -\frac{1}{4\pi G a^3}\nabla\left(\frac{\partial a\phi}{\partial\tau}\right) \,, \quad \mathbf{j}_{\perp} = \mathbf{j} - \mathbf{j}_{\parallel} \,. \tag{2.90}$$

Introducing a new vector $\mathbf{H}$

$$\nabla \times \mathbf{H} = -16\pi G a^2\mathbf{j}_{\perp} \,, \quad \nabla \cdot \mathbf{H} = 0 \tag{2.91}$$

and a new tensor H_{ij}

$$H_{ij} \equiv -12\nabla_{(i}H_{j)} - 2v_k\epsilon^{kl}_{(i}E_{j)l} \,, \tag{2.92}$$

we can define M_{ij} through

$$\begin{aligned} M_{ij} = -\nabla_k\epsilon^{kl}_{(i}H_{j)l} + \theta E_{ij} + \delta_{ij}\sigma^{kl}E_{kl} - 3\sigma^k_{(i}E_{j)k} - \omega^k_{(i}E_{j)k} = \\ = -4\pi G a^2\rho\nabla_{(i}v_{j)} - \frac{1}{a}\frac{d}{d\tau}(\nabla_j\nabla_j a\phi) = \\ = -4\pi G a^2\nabla_{(i}\mathbf{j}_{\perp j} - v_k\nabla^k\nabla_i\nabla_j\phi + v_{(i}\nabla_{j)}\nabla^2\phi \,. \end{aligned} \tag{2.93}$$

A new equation for H_{ij} is then needed to close the hierarchy. However, Bertschinger and collaborators are interested in local dynamical approximations, where there is no coupling between different mass elements, aside from those implied by the initial conditions. From the mathematical point of view, this means that no spatial gradients must appear in the Lagrangian fluid equations. Therefore a local description is obtained either by finding an approximation for $-\nabla_k\epsilon^{kl}_{(i}H_{j)l}$ or by truncating the hierarchy in a way that eliminates the need to determine it. For instance, ZA is a local Lagrangian approximation since it derives from assuming $E_{ij} = -(3\Omega H^2/2\dot{a}f_+)\sigma_{ij}$ and $\nabla^2\phi \propto \theta$ as required by Eulerian linear theory (Hui & Bertschinger 1996). In this case, the evolution of each fluid element can be followed till to orbit crossing without considering eq. (2.89). Trying to generalize ZA, which gives the exact solution for plane parallel flows (before shell-crossing occurs) but overestimates the collapse time for a uniform spherical top-hat configuration, Bertschinger & Jain (1994) and Hui & Bertschinger (1996) proposed two new dynamical approximations by integrating the exact Raychaudhuri and shear evolution equations and approximating the tidal evolution equation.

- Setting $H_{ij} = 0$ one obtains the the non-magnetic approximation (Bertschinger & Jain 1994) which is exact for plane parallel and spherical [$E_{ij} = 0$, $\sigma_{ij} = 0$] configurations. However, in disagreement with numerical simulations, it generally predicts two-dimensional collapses leading to strongly prolate filaments rather than Zel'dovich pancakes.

- Setting $M_{ij} = 0$ one obtains the local tidal approximation (Hui & Bertschinger 1996), which is exact for plane parallel, spherical and cylindrical [$\nabla_i v_j = (1/2)\theta\,\mathrm{diag}(1,1,0)$, $\nabla_i\nabla_j\phi = (1/2)\nabla^2\phi\,\mathrm{diag}(1,1,0)$] flows and, more generally, for any growing mode perturbations whose gravitational equipotential surfaces have constant shape with time. The local tidal approximations generically produces pancake-like collapsed structures and it is much more accurate than ZA in describing the evolution of an ellipsoidal homogeneous perturbation (for a configuration with initial axis ratios 1 : 1.25 : 1.5 it overestimates the value of a at collapse by 3% compared with 52% for ZA). To the best of our knowledge, the local tidal approximation has not been tested for more general (i.e. irregular) initial conditions.

The problem with the Lagrangian fluidodynamic approach is that only the size and the internal state (density, expansion rate, shear) of a given mass element can be computed. No information is given about the Eulerian spatial position of the fluid element.

Chapter 3

Evolution of the mass auto-correlation function

3.1 Introduction

In the last two decades, redshift surveys provided a wealth of information about the spatial distribution of local galaxies, revealing the existence of large-scale structures. The most widely used statistical tool to quantify the degree of clustering has been the galaxy two-point correlation function, both in its angular $w_g(\theta)$ and spatial $\xi_g(r)$ versions (see, e.g., Peebles 1980). As previous studies were confined to the nearby universe, nowadays new observational resources permit to extend the correlation analysis to deeper samples. In fact, the Canada-France Redshift Survey (Le Fèvre et al. 1996), the Canadian Network for Observational Cosmology (CNOC) field galaxy redshift survey (Shepherd et al. 1997; Carlberg et al. 1998), and the Hawaii Keck K-band survey (Cowie et al. 1996; Carlberg et al. 1997) have recently provided the new opportunity to investigate the clustering properties of galaxies out to redshifts $z \sim 1.5$. Moreover, only lately it has been possible to analyse the angular distribution of faint galaxies by using the Hawaii Keck K-band survey (Cowie et al. 1996) and the Hubble Deep Field data (Villumsen, Freudling & da Costa 1997; Connolly, Szalay & Brunner 1998; Arnouts et al. 1999 and references therein). An exciting progress in this field has been achieved with the recent discovery of a large number of galaxies at $z \sim 3$ through the Lyman-break technique (Steidel & Hamilton 1993; see §6.4 for a brief description of this method). The clustering strength of this galaxy population has been determined with great precision (Steidel et al. 1998; Giavalisco et al. 1998; Adelberger et al. 1998).

Therefore, we are now able to detect a direct signature of redshift dependence in the observed correlation function. For this reason, the theoretical analysis of the evolution of the mass two-point correlation function, $\xi(r)$, has become a fundamental topic of modern cosmology. However, it is worth stressing that the interpretation of the observational data is not immediate: before obtaining the "real" change of the large-scale structure one has to consider the possible evolution of the galaxy population (as well as the related selection effects) and of the bias factor that formally relates ξ_g to ξ (see, e.g., Matarrese et al. 1997).

The observational results should then be compared to the predictions of the existing models for structure formation. One of the several issues involved in this comparison is represented by the lack of a standard description of clustering evolution: analytical treatments are generally unable to manage this fully non-linear problem while numerical simulations are limited in resolution. However, new light has been recently shed on this argument. Hamilton et al. (1991) suggested that the correlation function obtained through N-body simulations of an Einstein-de Sitter universe,

in which the structure develops hierarchically, can be easily reproduced by applying a non-local and non-linear transformation to the linear $\xi(r)$. This Ansatz has been refined and extended to more general cosmological scenarios by a number of authors (Peacock & Dodds 1994; Jain, Mo & White 1995, hereafter JMW; Peacock & Dodds 1996). Moreover it is possible to give theoretical arguments that account for the scaling hypothesis (Nityananda & Padmanabhan 1994).

This chapter is composed of two parts. In §3.2 we discuss possible scaling solutions to the growth of mass perturbations through gravitational instability. Special emphasis is given to the problem of self-similar clustering and to the scaling Ansatz by Hamilton et al. (1991). The main purpose of this chapter is, however, to compare the predictions of the Zel'dovich approximation (Zel'dovich 1970) with the scaling Ansatz formulated in the version of Jain, Mo & White (1995, hereafter JMW). This analysis is performed in the second part of the chapter (starting from §3.3). Actually, it would be very interesting to obtain all the details of the semi-empirical scaling relationship in the framework of the gravitational instability scenario. However, in the absence of a model for the advanced phases of clustering evolution, we are forced to analyse only the onset of non-linear dynamics. The original results here presented have been published in Porciani (1997).

3.2 Scaling solutions to gravitational clustering

3.2.1 Self-similar clustering

A physical system with no preferred scale in the dynamics and in the initial conditions is expected to display self-similar evolution. In particular cases, the equations that describe clustering growth in the universe do not present characteristic lengths or times and indeed admit solutions that are invariant under a similarity transformation. This happens when the lack of characteristic scales equally involves the cosmological background, particle interactions and the initial conditions of the evolutionary problem, i.e. when

- the universe is well described by an Einstein-de Sitter model ($\Lambda = 0$, $p = 0$, $\Omega = 1$);
- non-gravitational forces are negligible;
- the mean interparticle distance is negligibly small, i.e. the fluid approximation holds on every scale;
- the initial power spectrum of fluctuations is a simple power law evolving like pure growing mode perturbations [1].

In this case, the dimensional analysis of the BBGKY hierarchy of equations (see appendix B) simply gives

$$\begin{aligned} F_1(\mathbf{x},\mathbf{p},t) &= t^{-3\beta}\,\mathcal{F}_1(\mathbf{x}/t^{\alpha},\mathbf{p}/t^{\beta}) \\ F_2(r,\mathbf{p}_1,\mathbf{p}_2,t) &= t^{-6\beta}\,\mathcal{F}_2(r/t^{\alpha},\mathbf{p}_1/t^{\beta},\mathbf{p}_2/t^{\beta}) \\ &\dots \end{aligned} \tag{3.1}$$

where F_1 and F_2 are the phase-space distributions defined in appendix B, α and β are constants related by $\beta = \alpha + 1/3$, $r = |\mathbf{x}_2 - \mathbf{x}_1|$, and the symbols $\mathcal{F}_i$ denote a set of unknown functions (Davis & Peebles 1977). Note that equations (3.1) mean that self-similarity holds only on a statistical sense (i.e. after averaging over the ensemble and not for every realization of the process).

[1] Other self-similar solutions involving the decaying mode exist but are not physically acceptable because give rise to divergent correlations at large separations (Davis & Peebles 1977)

The assumptions that lead to self-similarity are very restrictive and the reader might think that they define a problem of merely academic interest. However, this is not true. Because of their intrinsic simplicity, scale-free systems can give precious insight in modelling non-linear clustering growth and are still widely studied in the literature. Their analysis is useful also for understanding the evolution of realistic cosmological spectra, like the CDM one, which can be piecewise approximated by power-laws.

For the simplest two-point statistical measures, by integrating eq. (3.1) over momenta, one gets

$$\xi_2(x,t) = \hat{\xi}(s), \qquad s \equiv x/t^{\alpha}; \qquad\qquad P(k,t) = a^{3\alpha}\,\hat{P}(u), \qquad u \equiv k\,t^{\alpha}, \tag{3.2}$$

with $\hat{\xi}$ and $\hat{P}$ arbitrary functions. The index α is unspecified, a scale-invariant solution of the form above exists for any value of α. Davis and Peebles (1977) suggested a method to pick out one unique member from the class of scale-invariant solutions. This is obtained by constraining the behaviour of the two-point correlation function in two limiting regimes:

- At large separation x, where $\xi \ll 1$, the correlation function has to match the linear perturbative result, $\xi \propto t^{4/3}\,x^{-(n+3)} = s^{-(n+3)}$. This implies that $\alpha = 4/(9+3n)$.
- On small scales, it seems reasonable to expect gravity to form stable, bound systems. If one ignores the effect of mergers, these systems should mantain their densities and sizes in proper coordinates, i.e. the clustering should be "stable". The particle pairs at small separation x, where $\xi \gg 1$, are then expected to be contained in the same gravitationally bound clump. The guess is that the mean rate of proper separation of the pairs is close to zero. In other words, the mean peculiar velocity of the pairs at separation x has to nearly cancel the corresponding Hubble flow, i.e. $v(x,t) = -H(t)a(t)x = -(2/3t)a(t)x$. Pair conservation then gives $\xi \propto s^{-\gamma}$ with $\gamma = 2/(\alpha + 2/3)$.
- The global result is then $\gamma = (9+3n)/(5+n)$.

The stable clustering argument can be generalized as follows. Let us define the dimensionless pair velocity $h(x,t) \equiv -(v/\dot{a}x)$, so that the assumption of stability corresponds to $h = 1$. If the collapsed systems have reached stationarity in the statistical sense, the function h should reach some constant value at small x (Padmanabhan 1996). This would require the correlation function to have the form $\xi \propto a^{(3-\gamma)h}\,x^{-\gamma}$ ($\xi \gg 1$) with $\gamma = 3h(n+3)/[2+h(n+3)]$.

One might wonder whether the non-linear regime is really so strongly coupled to the initial conditions that the spectral index n on large scales is determining γ. One might feel that highly non-linear physics should be independent of the initial conditions. Padmanabhan (1995) showed that this would be the case if the asymptotic value of h satisfies the scaling $h(n+3) = \text{const}$. However, it appears strange to invoke a condition involving both linear and non-linear quantities in order to explain the independence of the correlation function from the linear power-spectrum.

Testing the stability of the possible self-similar solutions with respect to small perturbations is a viable method to check which kind of scaling (if any) is suitable to describe the highly non-linear regime. Unfortunately, it appears that stability arguments cannot give any hints about the value of γ: all the possible solutions come out to be marginally stable, in the sense that perturbations do not grow or decay (Ruamsuwan & Fry 1992; Yano & Gouda 1998). However, in order to close the BBGKY hierarchy and perform the stability analysis, it is necessary to make some a priori assumptions about the behaviour of the n-point correlations with $n > 2$. The influence of these Ansätze on the final results is not yet clear.

Nonetheless, although allowed by the equations, there is no guarantee that self-similarity is necessarily required by all solutions. The first question to answer is about the range of linear spectral

indices that indeed can support hierarchical and self-similar evolution. This can be investigated by considering the evolutionary equations for the smoothed density perturbations.

- For $n < -3$, the rms density contrast diverges as $k \to 0$. Therefore the condition $n > -3$ is required to ensure that clustering proceeds from small to large length scales.

- The linear, single-particle velocity dispersion shows the same kind of divergence when $n < -1$. This fact might be responsible for a breakdown of the self-similar scaling, since it introduces a new scale in the system. This is obtained by imposing that the average velocity on the scale x does not exceed the velocity of light, implying $x \propto a$ for $n < -1$. On the other hand, a divergent contribution from small scales results if $n > 1$.

- The fact that, from the "microscopic" point of view, the cosmological fluid is formed by fundamental discrete units (whatever they are, ranging from galaxies to elementary particles) of characteristic dimension x_0 gives an upper limit to n. Momentum conservation, in fact, implies $n < 4$ for $k \to 0$. The same limit is required to prevent the evolution of perturbations with $k \ll x_0^{-1}$ being dominated by non-linear interactions with $k \gtrsim x_0^{-1}$ modes (Zel'dovich 1965; Peebles 1974; Peebles and Groth 1976).

The resulting formal bounds for the self-similar solution to hold then are $-1 < n < 1$. However, only high resolution numerical simulations can, at present, test these conjectures about self-similar scaling. Efstathiou et al. (1988) and Bertschinger & Gelb (1991) used N-body simulations of scale free spectra with $n = -2, -1, 0, 1$ to examine the evolution of the two-point correlation function and of the mass distribution of gravitationally bound objects. Their results confirmed the picture of hierarchical formation of non-linear structure on increasingly large length scales. Moreover, barring the case $n = -2$ for which they lacked resolution, they verified the validity of the scaling (3.1). More recently, Lacey & Cole (1994), JMW, Padmanabhan et al. (1996) and, in particular, Jain & Bertschinger (1998) showed that in high-resolution simulations also the $n = -2$ spectrum displays self-similar scaling. In this case, the bulk velocity is found to drive the evolution of the phases of the Fourier modes of the density field at late times. Jain & Bertschinger (1994) showed analytically that this corresponds to a kinematical effect which does not affect the growth of perturbation (its action being similar to a global translation). These results seem to suggest that the spectral range for self-similarity can be extended downward to $-3 < n < 1$. However, Lucchin et al. (1994) and Colombi, Bouchet & Hernquist (1996) testing self-similarity for higher moments of the density, found clear departures from the scaling solution for $n > -1$ in the strongly non-linear regime.

N-body simulations can be also used to directly test the stable clustering hypothesis. This issue has been addressed by Padmanabhan et al. (1996) and by Jain (1996) who, however, reached opposite conclusions. In fact, while Jain found that the onset of stable clustering occurs (in the sense that $h \sim 1$) at $x/x_{\rm nl} = 0.07$ ($\xi \sim 200 - 2000$) almost independently of the linear power spectrum, Padmanabhan and collaborators claimed that the relation between h and ξ in the deeply non-linear regime is spectrum dependent and it is not guaranteed that $h \to 1$ at small scales. They also showed that the stable clustering hypothesis is better suited for cosmologies with $\Omega < 1$ where structure freezes out when $z \sim \Omega^{-1} - 2$. Models with $\Omega = 1$ are instead characterized by a steady merging phase. It is important to stress here, that we are dealing with very small scales, where correlations are closely connected to the inner structure of dark halos. Present day simulations do not possess enough resolution to explore much smaller scales. Padmanabhan et al. (1996) and Jain (1996) then agree that it is premature to draw conclusions about a possible asymptotic regime for h. Only a future generation of high resolution simulations migth clarify the reliability of the stable clustering hypothesis.

3.2.2 The scaling Ansatz

The analysis of a large set of numerical simulations suggests that, in hierarchical models, the non-linear two-point correlation function, $\xi(r,z)$, can be related to the linear one, $\xi_{\rm L}(r,z)$, through a simple scaling relation (Hamilton et al. 1991, Peacock & Dodds 1994, Jain, Mo & White 1995, Peacock & Dodds 1996). The main idea is that the action of gravity can be represented as a continuous change of scale or, better, that the 'flow of information' about clustering propagates along the curves of equation:

$$r_0 = [1+\bar{\xi}(r,z)]^{1/3} r \ , \tag{3.3}$$

where $\bar{\xi}(r,z)$ represents the average correlation function within a sphere of radius r

$$\bar{\xi}(r,z) = \frac{3}{r^3}\int_0^r y^2 \xi(y,z) dy \tag{3.4}$$

and r_0 is a sort of Lagrangian coordinate determining a 'conserved pair surface' (Hamilton et al. 1991, Nityananda & Padmanabhan 1994). In fact, by definition, the average number of neighbours of a particle contained within a spherical volume of radius r_0 at the linear stage (when $\bar{\xi} \ll 1$) equals the average number of neighbours inside a sphere of radius r in the evolved field. Note that, as clustering develops, $\bar{\xi}$ increases and r becomes cosiderably smaller than r_0. The scaling Ansatz, then, consists in assuming that the non-linear mean correlation function at r can be expressed as a function of its linear counterpart at r_0, i.e.

$$\bar{\xi}(r,z) = F\left[\bar{\xi}_{\rm L}(r_0,z)\right] \ . \tag{3.5}$$

Nityananda & Padmanabhan (1994) showed that a scaling solution of the form (3.5) is obtained directly from the pair conservation equation (see appendix B, eq. B.17) after assuming the the mean pair velocity on a given length scale is a function only of the mean correlation function on the same scale;

The pair conservation equation is exact, but not closed, in the sense that the mean pair velocity is unknown in general. The equation for v is derived from higher order equations of the hierarchy. If, on the other hand, one wants to deal with $\bar{\xi}$ only, it is necessary to express $h(r,z)$ as a functional of $\bar{\xi}(r,z)$. The simplest possibility is a local function,

$$h(x,a) = H[\bar{\xi}(x,a)] \ . \tag{3.6}$$

This is consistent with both the linear limit $h \simeq (2/3)\bar{\xi}$ (Peebles 1980), and the stable clustering regime, where $h \simeq 1$.

Their final result is that, on the curves where $\ln[r^3(1+\bar{\xi})] = \text{const}$,

$$\bar{\xi}_L(r_0,z) = \exp\left[\frac{2}{3}\int^{\bar{\xi}(r,z)} \frac{d\mu}{H(\mu)(1+\mu)}\right] \ , \tag{3.7}$$

where the lower limit of the integral is chosen to give $\ln\bar{\xi}$ for small values of $\bar{\xi}$ in the linear regime. The non-linear mean correlation $\bar{\xi}(r,z)$ can then be expressed in terms of $\bar{\xi}_L(r_0,z)$ by a function that makes a smooth transition from the linear regime to the stable-clustering regime. This function can be determined knowing the relation that links the scaled pair velocity to the mean mass correlation.

In a series of papers, Padmanabhan and collaborators outlined how simple analytical computations can be used to extract the analytic form of the scaling Ansatz in good agreement with numerical simulations. The paradigm of their calculations is based on the spherical collapse model we already discussed in §2.4.1. In practice they distinguish three regimes in the growing phase of gravitational clustering (Padmanabhan 1996):

- The **linear regime** ($\bar{\xi} \ll 1$) during which the leading order perturbative result $\bar{\xi}(r,z) = \bar{\xi}_L(r_0,z) \propto a^2$ with $r \simeq r_0$ holds.

- The **quasi-linear regime** ($1 < \bar{\xi} < 200$, $1 < \bar{\xi}_L < 5.85$, where 200 roughly gives the mean overdensity of a collapsed structure at virialization) which is characterized by self-similar, spherical accretion onto collapsed structures corresponding to the highest density peaks of the initial fluctuation field. Assuming that different shells contribute to $\bar{\xi}$ with an effective radius that is proportional to their turn-around radius (Bertschinger 1995), one obtains $\bar{\xi}(r,z) \simeq [\bar{\xi}_L(r_0,z)]^3$. Note that, for a linear correlation $\bar{\xi}_L \propto r^{-\alpha}$, one gets $\bar{\xi} \propto r^{-3\alpha/(1+\alpha)}$. It is also assumed that, during this phase, mergers of structures have not yet begun to play an important role.

- The **non-linear regime** ($\bar{\xi} > 200$) is described through the stable clustering assumption (Padmanabhan 1996). Alternatively it can be handled in the framework of the spherical accretion model, by assuming that the accreting objects derive from typical (1σ) fluctuations of the linear density field. In both cases one obtains: $\bar{\xi}(r,z) = 14.14[\bar{\xi}_L(r_0,z)]^{3/2}$, where the numerical coefficient has been defined to assure analytical continuity to the scaling formula.

Numerical simulations confirm this general picture, showing however that that the scales at which the transitions between the various clustering regimes take place depend on the linear power spectrum assumed (JMW; Peacock & Dodds 1996; Padmanabhan et al. 1996). This spectral dependence of the scaling Ansatz arises naturally in theoretical models when averaging over the whole distribution of peak heights is taken into account (Munshi & Padmanabhan 1997).

3.3 Evolution of ξ in the Zel'dovich approximation

In this section, we study the non-linear evolution of the mass autocorrelation function by describing the growth of density fluctuations through the Zel'dovich approximation. In effect, Eulerian second-order perturbation theory may break down once the mass variance becomes sufficiently large. On the other hand, we know that ZA, especially in its "truncated" form, is able to reproduce fairly well the outcomes of N-body simulations even in the mildly non-linear regime (Melott, Pellman & Shandarin 1994). The main advantage of ZA over other dynamical approximations (e.g. Sahni & Coles 1995) is that it permits analytical investigations while at the same time ensuring good accuracy, at least for quasi-linear scales. The pioneering analysis by Bond & Couchmann (1988) showed that ZA is able to predict the shifting of the first zero crossing of the correlation function. In §3.3.1 we will give a detailed quantitative description of this effect. Other features of the mass two-point correlation function in ZA have been discussed by Mann, Heavens & Peacock (1993, hereafter MHP). Moreover, the related evolution of the power spectrum has been studied by Taylor (1993), Schneider & Bartelmann (1995) and Taylor & Hamilton (1996). These authors showed that ZA is able to describe the generation of small-scale power through mode coupling, at least at early times. Fisher & Nusser (1996) and Taylor & Hamilton (1996) succeeded in computing the power spectrum also in redshift space.

This chapter is organized as follows. In section 3.3.1 we compute the cross correlation function between the mass density field evaluated at two different times. The usual two-point correlation function is obtained as a particular case of this more general quantity. The redshift evolution of $\xi(r)$ in a CDM model is the last subject of section 3.3.1. In section 3.3.2 we compare the predictions of ZA with the scaling Ansatz of JMW. In section 3.3.3 we use our results to evaluate the correlation function of a collection of objects sampled by an observer in a wide redshift interval of his past light

cone. We then propose a simplified scheme with which to compute this quantity so as to improve another approximation presented in the literature. A brief summary is given in section 3.3.4.

3.3.1 The two-point correlation function in the Zel'dovich approximation

Assuming that initially the mass is evenly distributed in Lagrangian space, this implies that the Eulerian density field is related to the Lagrangian displacement field via the relation:

$$\rho(\mathbf{x},t) = \rho_b(t) \int d^3q\, \delta_D\left[\mathbf{x} - \mathbf{q} - \mathbf{S}(\mathbf{q},t)\right], \tag{3.8}$$

where $\delta_D(\mathbf{x})$ denotes the three-dimensional Dirac delta distribution. For purposes that will be clarified in section 3.3.3, we are interested in computing the cross-correlation function of the density contrast field evaluated at two different times:

$$\langle \delta(\mathbf{x}_1,t_1)\delta(\mathbf{x}_2,t_2)\rangle = \left\langle \int d^3q_1\, d^3q_2\; \delta_D\left[\mathbf{x}_1 - \mathbf{q}_1 - \mathbf{S}(\mathbf{q}_1,t_1)\right] \delta_D\left[\mathbf{x}_2 - \mathbf{q}_2 - \mathbf{S}(\mathbf{q}_2,t_2)\right] \right\rangle - 1 \tag{3.9}$$

where $\langle \dots \rangle$ represents the average over an ensemble of realizations (see appendix A). Before going any further, it is convenient to Fourier transform the Dirac delta functions in equation (3.9) obtaining:

$$\begin{aligned} 1 + \langle \delta(\mathbf{x}_1,t_1)\delta(\mathbf{x}_2,t_2)\rangle = \int d^3q_1\, d^3q_2\, \frac{d^3w_1}{(2\pi)^3}\frac{d^3w_2}{(2\pi)^3} \exp\left[i \sum_{j=1}^{2} \mathbf{w}_j \cdot (\mathbf{x}_j - \mathbf{q}_j)\right] \times \\ \times \left\langle \exp\left[-i \sum_{\ell=1}^{2} \mathbf{w}_\ell \cdot \mathbf{S}(\mathbf{q}_\ell, t_\ell)\right]\right\rangle . \end{aligned} \tag{3.10}$$

We then use equation (2.73) to introduce ZA into equation (3.10). In such a way, by assuming, as usual, that $\phi_S(\mathbf{q})$ is a statistically homogeneous and isotropic Gaussian field, uniquely specified by its power spectrum $P_{\phi_S}(k) \propto P(k)/k^4$, the ensemble average contained in equation (3.10) can be written as a functional integral (see appendix A):

$$\begin{aligned} &\left\langle \exp\left[-i \sum_{\ell=1}^{2} \mathbf{w}_\ell \cdot \mathbf{S}(\mathbf{q}_\ell, t_\ell)\right]\right\rangle = \\ &= (\det K)^{1/2} \int \mathcal{D}[\phi_S] \exp\left[-\frac{1}{2}\int \phi_S(\mathbf{q}) K(\mathbf{q},\mathbf{q}') \phi_S(\mathbf{q}') d^3q\, d^3q' + i \sum_{\ell=1}^{2} b(t_\ell)\mathbf{w}_\ell \cdot \nabla\phi_S\big|_{\mathbf{q}_\ell}\right], \end{aligned} \tag{3.11}$$

where the kernel $K(\mathbf{q},\mathbf{q}')$ represents the functional inverse of the two-point correlation function of the field $\phi_S(\mathbf{q})$. By defining a six-dimensional vector $\mathbf{c}^t = (\mathbf{w}_1, \mathbf{w}_2)$ and choosing the z-axis of our reference frame in the direction of the vector $\mathbf{q} = \mathbf{q}_1 - \mathbf{q}_2$, we can reduce equation (3.11) to the form:

$$\left\langle \exp\left[-i \sum_{\ell=1}^{2} \mathbf{w}_\ell \cdot \mathbf{S}(\mathbf{q}_\ell, t_\ell)\right]\right\rangle = \exp\left[-\frac{1}{2}\mathbf{c}^t \mathbf{M} \mathbf{c}\right] \tag{3.12}$$

where the matrix $\mathbf{M}$ has the structure

$$\mathbf{M} = \gamma \begin{pmatrix} b_1^2 & 0 & 0 & b_1 b_2 \psi_\perp & 0 & 0 \\ 0 & b_1^2 & 0 & 0 & b_1 b_2 \psi_\perp & 0 \\ 0 & 0 & b_1^2 & 0 & 0 & b_1 b_2 \psi_{||} \\ b_1 b_2 \psi_\perp & 0 & 0 & b_2^2 & 0 & 0 \\ 0 & b_1 b_2 \psi_\perp & 0 & 0 & b_2^2 & 0 \\ 0 & 0 & b_1 b_2 \psi_{||} & 0 & 0 & b_2^2 \end{pmatrix} \tag{3.12}$$

with $b_i = D_+(t_i)$ and

$$\gamma = \frac{1}{6\pi^2}\int_0^\infty P(k)dk\ ,$$
$$\gamma\psi_{||}(q) = \frac{1}{2\pi^2}\int_0^\infty P(k)\left[j_0(kq) - \frac{2}{kq}j_1(kq)\right]dk\ , \qquad \gamma\psi_\perp(q) = \frac{1}{2\pi^2}\int_0^\infty P(k)\frac{1}{kq}j_1(kq)dk\ , \tag{3.13}$$

having denoted by $j_\ell(x)$ the spherical Bessel function of order ℓ. By substituting this result into equation (3.10) we can easily solve the Gaussian integration over the $\mathbf{w}_i$. In order to perform the remaining integrations, it is convenient to introduce the new variables $\mathbf{q}$ and $\mathbf{Q} = \mathbf{q}_1 + \mathbf{q}_2$. In this way, after some algebra, we finally obtain:

$$\begin{aligned} 1+\xi(r,t_1,t_2) &\equiv 1 + \langle\delta(\mathbf{x}_1,t_1)\delta(\mathbf{x}_2,t_2)\rangle = \\ &= \frac{1}{(2\pi)^{1/2}r}\int_0^\infty \frac{q^2dq}{(b_1b_2)^{1/2}\gamma(\psi_\perp-\psi_{||})^{1/2}(b_1^2+b_2^2-2b_1b_2\psi_\perp)^{1/2}} \times \\ &\times \left\{W(u_+)\exp\left[-\frac{(q-r)^2}{2\gamma(b_1^2+b_2^2-2b_1b_2\psi_{||})}\right] - W(u_-)\exp\left[-\frac{(q+r)^2}{2\gamma(b_1^2+b_2^2-2b_1b_2\psi_{||})}\right]\right\} \end{aligned} \tag{3.14}$$

where $r = |\mathbf{x}_1 - \mathbf{x}_2|$,

$$u_\pm = \left[\frac{b_1b_2(\psi_\perp-\psi_{||})}{\gamma(b_1^2+b_2^2-2b_1b_2\psi_\perp)(b_1^2+b_2^2-2b_1b_2\psi_{||})}\right]^{1/2}\left[\frac{b_1^2+b_2^2-2b_1b_2\psi_\perp}{2b_1b_2(\psi_\perp-\psi_{||})}q \pm r\right] \tag{3.15}$$

and $W(x)$ represents the Dawson's integral [2] (see, e.g., Abramowitz & Stegun 1968). It is straightforward to show that for $t_1 = t_2$ the previous formula reduces to the usual expression for the mass two-point correlation function in ZA (Bond & Couchman 1988; Mann, Heavens & Peacock 1993; Schneider & Bartelmann 1995).

We numerically evaluated the two-point correlation function $\xi(r,t) \equiv \xi(r,t,t)$ employing a *COBE*-normalized standard CDM linear power spectrum (with density parameter $\Omega = 1$ and $h = 0.5$). We used the transfer function of Bardeen et al. (1986) while the normalization to the four-year *COBE* DMR data is given in Bunn & White (1997) and corresponds to $\sigma_8 = 1.22$. As already noted by MHP, the small scale behaviour of the resulting correlation function depends on the value assigned to the truncation radius, R_f, defined in equation (2.76) (see Fig. 3.1). If R_f is very small, then shell crossing will not be suppressed and $\xi(r)$ will show an unusually flat behaviour. On the contrary, if R_f is too large, the smoothing procedure will remove an important contribution to the power spectrum, causing again too low a correlation. Therefore we need a criterion with which to select R_f. Since our main purpose is to compare the clustering amplitudes predicted by ZA with those extracted from the scaling Ansatz of JMW, we can choose R_f so as to optimize the agreement between the respective correlation functions. In any case, we find that this method conforms quite well to a simpler one already used by MHP: the best R_f is the one that maximizes $\xi(r, R_\mathrm{f})$ on small scales. Strictly speaking, the optimal smoothing radius depends on the scale selected for maximizing the correlation: the smaller r is, the larger R_f comes out (we find that the difference between the smoothing lengths obtained by maximizing ξ at $r = 0.1\,h^{-1}$Mpc and at $r = 1\,h^{-1}$Mpc roughly amounts to $0.2\,h^{-1}$Mpc and remains nearly constant by varying

[2] It is worth stressing that when $\psi_\perp < \psi_{||}$, in order to avoid a complex argument for the Dawson's integral, it is convenient to express the integrand in equation (3.14) in terms of exponentials and error functions (see also the discussion in Schneider & Bartelmann 1995). However, since for the CDM spectrum (the only one considered in our analysis) $\psi_{||}$ is never larger than $\psi_\perp$, we preferred to write the solution using $W(x)$.

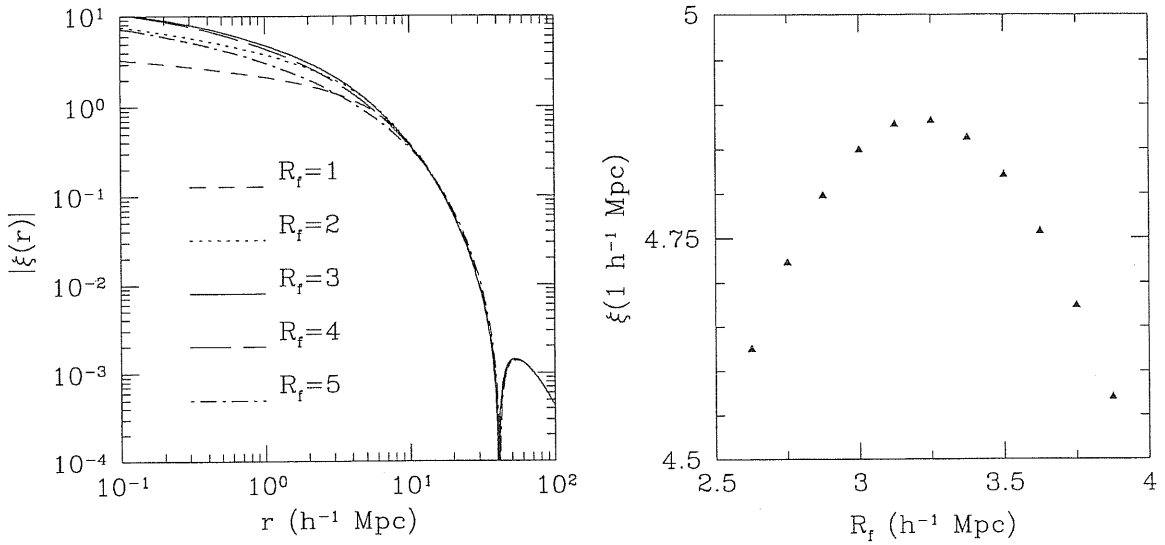


Figure 3.1: (Left panel) The mass autocorrelation function, obtained using ZA, for a *COBE*-normalized CDM linear spectrum is plotted for different values of the truncation radius $R_{\rm f}$ (in h^{-1}Mpc). (Right panel) Dependence of the correlation function evaluated at $r = 1\,h^{-1}$Mpc on $R_{\rm f}$.

σ_8). However, the effect of this discrepancy on the correlation evaluated on larger scales is indeed minimal. Following Schneider & Bartelmann (1995), we select $r = 1\,h^{-1}$Mpc as the scale at which we require $\xi(R_{\rm f})$ to be maximal. As previously stated, the optimum filtering length increases as the field evolves; the dependence of the best $R_{\rm f}$ on σ_8 is almost linear and for $\sigma_8 > 0.3$ (which in our model corresponds to $z \sim 3$) it can be approximated by:

$$R_{\rm f}(\sigma_8) = (3.16\,\sigma_8 - 0.65)\,h^{-1}\mathrm{Mpc}\,. \tag{3.16}$$

The redshift evolution of the correlation function is shown in Fig. 3.2. As expected, on scales that are not affected by shell crossing ($r > R_{\rm f}$), $\xi(r,z)$ steepens with decreasing z. Moreover, we note that the first zero crossing radius of $\xi(r,z)$ increases as time goes on (see also Bond & Couchmann 1988). A similar pattern has been noticed by Coles (1990) and by Baugh & Efstathiou (1994) in the context of second-order Eulerian perturbation theory. The displacement of the first zero crossing of ξ as a function of time is plotted in Fig. 3.2. Measuring the degree of dynamical evolution of the density field through σ_8, this shifting can be described with good approximation by the function:

$$r_{0\rm C}(\sigma_8) - r_{0\rm C}^{\rm lin} \simeq 5.3\,\sigma_8^{(1.5+0.1/\sigma_8)}\,h^{-1}\mathrm{Mpc} \tag{3.17}$$

where we denote by $r_{0\rm C}$ the scale at which the correlation function crosses the zero-level for the first time and by $r_{0\rm C}^{\rm lin}$ its linear counterpart. It would be interesting to compare this result with the predictions of second-order Eulerian (and Lagrangian) perturbation theory and of other dynamical approximations.

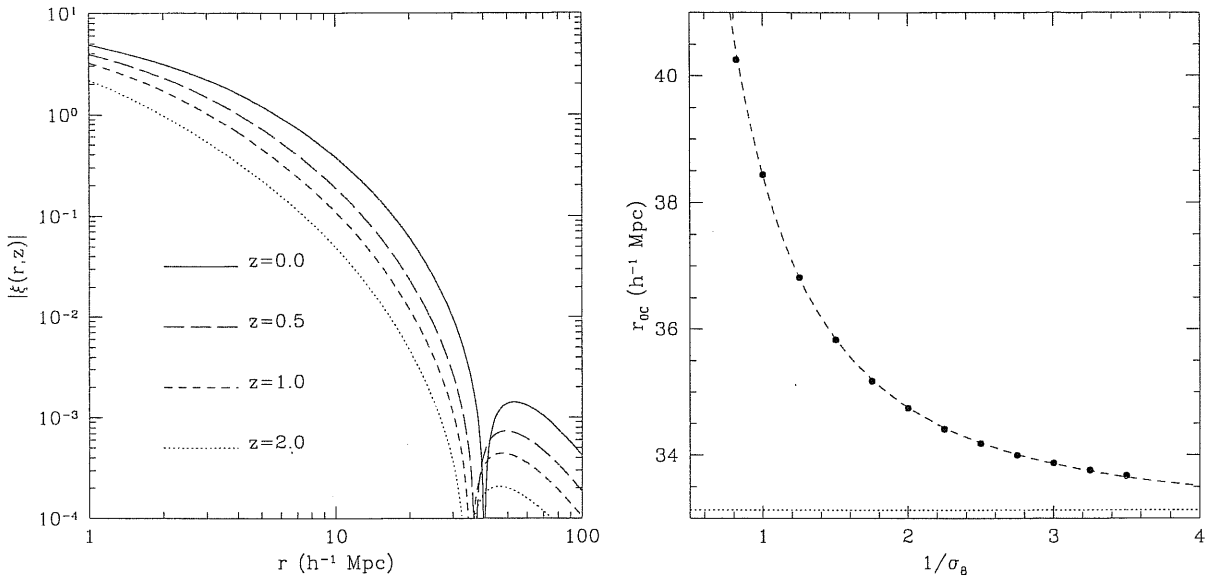


Figure 3.2: (Left panel) Redshift evolution of the mass two-point correlation function obtained using ZA to evolve a linear CDM spectrum. (Right panel) The first zero crossing radius of the correlation function is plotted against $1/\sigma_8$. The circles represent the results obtained using ZA, the dashed line is the fitting function given in the text while the dotted line shows the prediction of Eulerian linear theory.

3.3.2 Comparison with the scaling hypothesis

Here, we want to compare the results obtained in the previous section, using ZA, with the predictions of the scaling Ansatz (hereafter SA) formulated in JMW as:

$$\bar{\xi}(r,z) = B(n_{\rm eff})F\left[\frac{\bar{\xi}_{\rm L}(r_0,z)}{B(n_{\rm eff})}\right] \tag{3.18}$$

with

$$F(x) = \frac{x + 0.45x^2 - 0.02x^5 + 0.05x^6}{1 + 0.02x^3 + 0.003x^{9/2}}\,, \qquad B(n_{\rm eff}) = \left(\frac{3+n_{\rm eff}}{3}\right)^{0.8}\,, \qquad n_{\rm eff}(z) = \left.\frac{d\ln P(k)}{d\ln k}\right|_{k_{\rm NL}(z)} \tag{3.19}$$

where $k_{\rm NL}^{-1}$ denotes the radius of the top-hat window function in which the rms linear mass fluctuation is unity. However, it would be useless to perform the comparison between the spherically averaged correlation functions since, on small scales, $\xi(r)$ obtained using ZA is seriously affected by shell crossing and the computation of $\bar{\xi}$ requires an integration starting from $r = 0$. For this reason we prefer to use $\xi(r)$ directly. The two-point correlation function deriving from the Ansatz of JMW can be obtained performing a simple differentiation:

$$\xi(r,z) = \frac{\left[1 + B(n_{\rm eff})F(X)\right]F'(X)\Delta\xi_{\rm L}(r_0,z)}{1 + B(n_{\rm eff})F(X) - F'(X)\Delta\xi_{\rm L}(r_0,z)} + B(n_{\rm eff})F(X)\,, \qquad X = \frac{\bar{\xi}_{\rm L}(r_0,z)}{B(n_{\rm eff})}\,, \tag{3.20}$$

with $F'(x) = dF/dx$ and

$$\Delta\xi_{\rm L}(r_0,z) \equiv \xi_{\rm L}(r_0,z) - \bar{\xi}_{\rm L}(r_0,z) = \frac{D_+^2(z)}{2\pi^2}\int_0^\infty k^2 P(k)\left[j_0(kr_0) - \frac{3}{kr_0}j_1(kr_0)\right]dk\,. \tag{3.21}$$

We evaluated the correlation function given in equation (3.20) using a *COBE*-normalized, linear CDM spectrum. In the left panel of Fig. 3.3 we plot the result obtained at $z = 0$ with the corresponding one achieved by using ZA. For comparison we also show the prediction of Eulerian linear theory. The agreement between ZA and SA is remarkable on mildly non-linear scales ($4 \lesssim r \lesssim 20\,h^{-1}$Mpc) and on completely linear scales ($r > 50\,h^{-1}$Mpc). For example, at $r = 5\,h^{-1}$Mpc, linear theory overestimates the correlation of JMW by 82 per cent, ZA underestimates it by 2 per cent while the accuracy of the JMW fit is about 15 – 20 per cent. However, we find that in the interval $20 \lesssim r \lesssim 50\,h^{-1}$Mpc, ZA predicts more non-linear evolution than does SA (for example the r_{0C} obtained by using ZA is larger than the one determined through SA). In order to consider a less evolved field, in the right panel of Fig. 3.3 we repeat the comparison using the correlation functions evaluated at $z = 1$. Now, the main item to note is that the JMW result matches the linear solution on scales ($r \sim 10\,h^{-1}$Mpc) that, according to ZA, are already involved in non-linear phenomena.

In any case, we do not know the accuracy of the scaling hypothesis on large scales. In fact, the function $F(x)$ is obtained by requiring the resulting $\xi(r)$ to reproduce the linear behaviour where $\bar{\xi}_{\rm L} \to 0$ and, simultaneously, to approximate properly the correlation function extracted from N-body simulations. However, in order to achieve a detailed description of non-linear scales, JMW used a relatively small box to perform their simulations. Therefore, imposing the match to linear theory on large scales, without having any constraint from numerical data on quasi-linear scales, could seriously alter the accuracy of $F(x)$. This probably implies that the JMW fitting function could be improved on large scales. Our conclusion is shared by Baugh & Gaztañaga (1996, hereafter BG), who tested the scaling Ansatz for the evolution of the power spectrum against the results of 5 N-body simulations performed within a $378\,h^{-1}$Mpc box. Indeed, they found that the JMW formula gives a relatively poor description of the large-scale behaviour even though the agreement between the spectra remains always within the quoted 20 per cent accuracy.

By using the output of their simulations, BG proposed a new scaling formula calibrated on large scales. As initially suggested by Peacock & Dodds (1994), the analytic expression of this SA concerns the dimensionless power spectrum, $\Delta^2(k,z) = k^3P(k,z)/2\pi^2$ (i.e. the contribution to the the variance of the density contrast per bin of $\ln k$), while, following JMW, it takes account of a spectral dependence of the transformation:

$$\Delta^2(k,z) = \beta(n_{\rm eff}) f\left[\frac{\Delta^2_{\rm L}(k_{\rm L},z)}{\beta(n_{\rm eff})}\right], \qquad k_{\rm L} = \left[1+\Delta^2(k,z)\right]^{-1/3} k \tag{3.22}$$

where

$$f(x) = x\left(\frac{1+0.598x-2.39x^2+8.36x^3-9.01x^{3.5}+2.895x^4}{1-0.424x+[2.895/(11.68)^2]x^3}\right)^{1/2}, \tag{3.23}$$

$$\beta(n_{\rm eff}) = 1.16\left(\frac{3+n_{\rm eff}}{3}\right)^{1/2}, \tag{3.24}$$

and the subscript L marks linear quantities. The function $f(x)$ has been obtained by matching the power spectrum in the simulations at $\sigma_8 = 1$, with an accuracy of 5 per cent, over the range $0.02 < k < 1.0\,h\,{\rm Mpc}^{-1}$ and by forcing the fit to have the asymptotic form $f(x) \to 11.68\,x^{3/2}$ when $x \to \infty$ (Hamilton *et al.* 1991). The two-point correlation function is related to $\Delta^2(k,z)$ through the Fourier relation:

$$\xi(r,z) = \int_0^\infty \Delta^2(k,z) j_0(kr)\frac{dk}{k}. \tag{3.25}$$

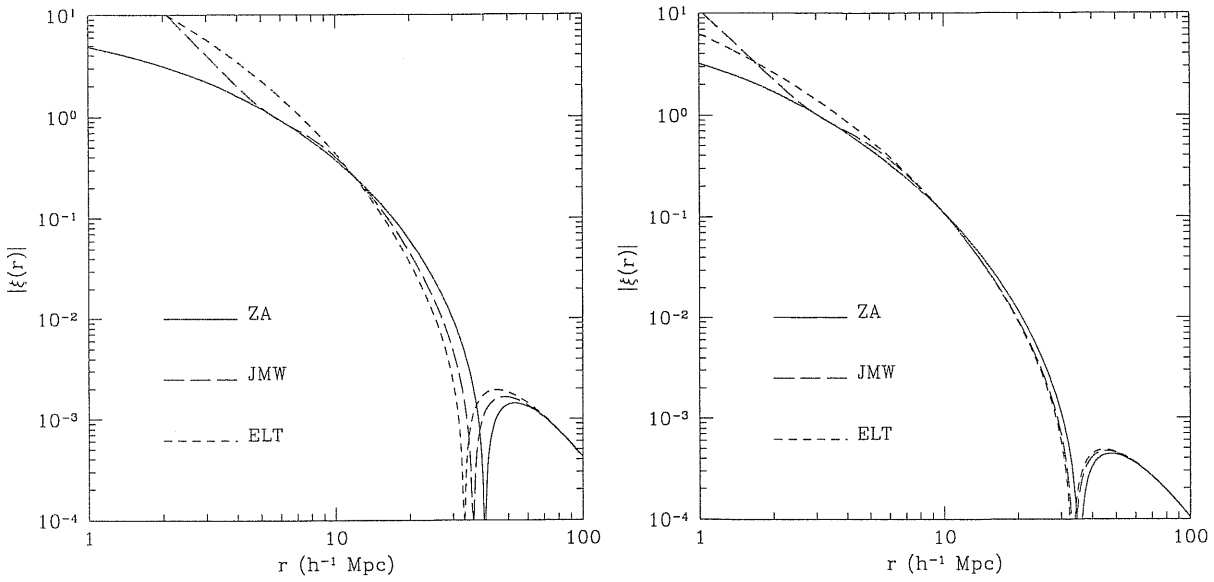


Figure 3.3: (Left panel) Comparison between the mass autocorrelation functions computed for a CDM model by using: the Zel'dovich approximation (ZA), the scaling Ansatz of Jain, Mo & White (JMW) and Eulerian linear theory (ELT). The linear power spectrum extrapolated to the present epoch ($z = 0$) is normalized to match the *COBE* DMR data ($\sigma_8 = 1.22$). (Right panel) As in the left panel, but at $z = 1$ ($\sigma_8 = 0.61$).

In Fig. 3.4 we compare the correlations obtained by using ZA and the JMW formula with the results of the scaling Ansatz by BG: we are considering a standard CDM linear spectrum at the epoch in which $\sigma_8 = 1.22$. We immediately note that using larger simulation boxes to calibrate the SA allows a better determination of the correlation function for $r \gtrsim 20\,h^{-1}$Mpc. In fact, we find that the correlations obtained with ZA and with the BG formula agree by better than 20 per cent for $r > 4.6\,h^{-1}$Mpc (with the exception of a very small r-interval centred in the first zero crossing of ξ) while the discrepancy between ZA and the JMW Ansatz is less than 20 per cent over the ranges $4.1 < r < 18.3\,h^{-1}$Mpc and $r > 49.6\,h^{-1}$Mpc. Similar patterns are obtained considering different values of σ_8. This shows that the BG fit, which has been calibrated against large box CDM simulations, gives also a very good description of the mass clustering predicted by ZA on intermediate scales. In any case, as expected, the JMW formula is sensibly more accurate for $5 \lesssim r \lesssim 15\,h^{-1}$Mpc where the BG predictions grow worse as σ_8 assumes values significantly larger than 1.

On the other hand, it would be interesting to check the reliability of ZA and second-order Eulerian perturbation theory by directly comparing their predictions on these scales. Bond & Couchmann (1988), studying the weakly non-linear evolution of the CDM power spectrum, found remarkable agreement between the two approximations. Moreover, Baugh & Efstathiou (1994) showed that second-order Eulerian perturbation theory can reproduce, at least qualitatively, the evolution of the power spectrum predicted by numerical simulations. However, Jain & Bertschinger (1994) found that the agreement between perturbation theory and N-body outcomes gets worse as the density field evolves. Besides, their results are inconsistent with the low-k behaviour of the second-order Eulerian correction to the CDM power spectrum computed by Bond & Couchmann (1988), again raising the issue about the compatibility between ZA and perturbation theory. In a recent work concerning the evolution of scale invariant spectra, Scoccimarro & Friemann (1996b)

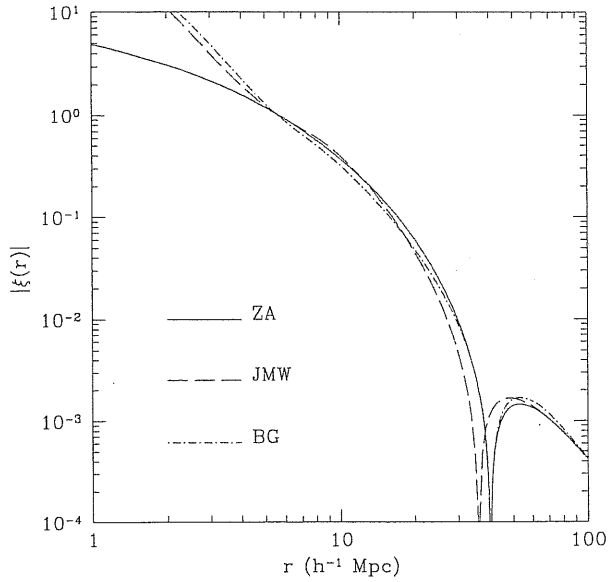


Figure 3.4: Mass two-point correlation functions at the epoch in which $\sigma_8 = 1$ obtained from a linear CDM spectrum evolved through the Zel'dovich approximation (ZA) and through the scaling Ansätze by Jain, Mo & White (1995) and by Baugh & Gaztañaga (1996).

showed that, if the spectral index n satisfies $-3 < n < -1$, Eulerian perturbation theory is able to reproduce the power spectrum obtained though the scaling Ansatz fairly well, while the one-loop perturbative version of ZA gives worse results. Anyway, Bharadwaj (1996a,b) pointed out that the effects of multistreaming on the correlation function cannot be studied perturbatively. This fact implies that our result, obtained considering the full Zel'dovich approximation, should be more reliable than any other achieved by adopting a perturbative version of ZA. In any case, it would be interesting to clarify the extent to which ZA and Eulerian perturbation theory agree on large scales.

3.3.3 The correlation of high-redshift objects

In this section, we study the evolution of the cross-correlation function of the mass density contrast evaluated at two different times as defined in equation (3.14). This quantity could play an important role in comparing the clustering properties extracted from deep redshift surveys to the predictions of theoretical models for structure formation. In practice, one always collects data on correlations in a finite redshift strip of the past light cone while the quantity $\xi(r,t)$, normally used in theoretical works, refers to objects selected on an hypersurface of constant cosmic time. Therefore, as far as one is considering a deep sample of cosmic objects, it is not correct to relate the observed clustering properties to $\xi(r,t)$. This issue is addressed in detail by Matarrese et al. (1997, hereafter MCLM; see also Yamamoto & Suto 1999) who build a theoretical quantity that allows a direct comparison of model predictions to the observed correlations. MCML's approach can be divided into three steps: first they compute the redshift evolution of mass correlations, then they relate the clustering properties of cosmic objects to the matter distribution by means of a linear bias relationship (see chapter 6), and finally they convolve the result with the observed redshift distribution of the class of objects under analysis. By assuming that the effects of redshift distortions and of the magnification bias due to weak gravitational lensing are negligible, and by considering isotropic selection functions,

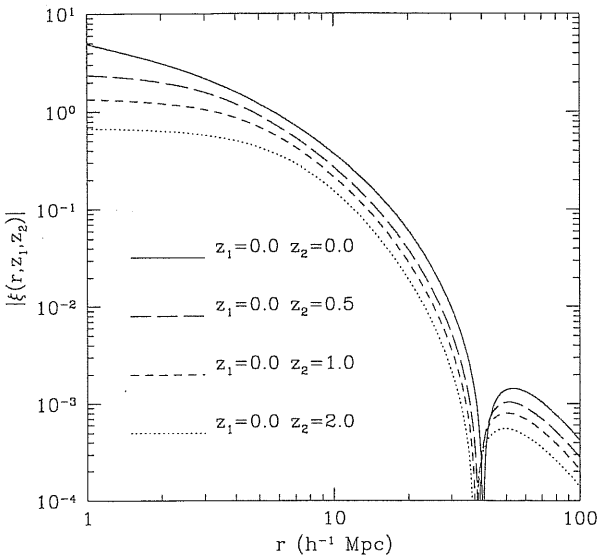


Figure 3.5: Cross-correlation between the density contrast field evaluated at two different redshifts versus comoving separation.

MCLM showed that the theoretical estimate for the observed two-point correlation function can be formally expressed as an integral over z_1 and z_2 of the function $\xi(r, z_1, z_2)$ weighted by geometrical factors and effective bias parameters (all dependent on z_1 and z_2). Different classes of objects are selected by changing the amplitude and the redshift dependence of the effective bias. However, in the absence of a model for the evolution of the cross-correlation, only assuming that the above-mentioned integral is dominated by the contribution of objects whose redshifts are nearly the same, can one estimate the observed correlation function deriving from a particular scenario of structure formation. In this way, one is allowed to replace $\xi(r, z_1, z_2)$ with $\xi(r, \bar{z})$, where $\bar{z}$ is a suitably defined average between z_1 and z_2 that, for simplicity, MCLM identify with $\bar{z} = (z_1 + z_2)/2$. This is a crucial approximation, as it allows MCLM to use the JMW Ansatz to compute the non-linear mass correlation function [there is no known scaling Ansatz for $\xi(r, z_1, z_2)$]. However, as shown in the previous paragraphs, ZA allows the computation of $\xi(r, z_1, z_2)$ so that we are able to compute the theoretical estimate for the observed correlation function by using both the complete and the approximated formulae given by MCLM (respectively their equations 15 and 18). Therefore we can check here, within the validity of ZA, the reliability of the approximation introduced by MCLM. Large discrepancies between the exact and the approximated correlations would obviously invalidate their whole analysis and consequently also their complete formula for $\xi_{\rm obs}$ would be unusable. On the other hand, if the approximated correlation function turns out to reproduce accurately the complete one, MCLM formulae could represent an important tool with which to disprove cosmological models in the light of present and future observations.

In order to compute $\xi(r, z_1, z_2)$ using equation (3.14), we truncated the linearly extrapolated power spectrum $D_+(z_1)D_+(z_2)P(k)$ according to the prescription:

$$P_T(k, z_1, z_2) = D_+(z_1)D_+(z_2)P(k)\exp\left[-k^2 R_{\rm f}(z_1)R_{\rm f}(z_2)\right] \qquad (3.26)$$

where $R_{\rm f}(z)$ represents the optimum filtering length for the density field at redshift z, determined by following the method described in §3.3.1. On small scales, the correlation functions that we obtain opting for this truncation procedure appear much more flattened than those computed at a

single time. The evolution of $\xi(r, z_1, z_2)$ as z_2 changes is shown, for a CDM model, in Fig. 3.5. It is evident that even though the correlation decreases as z_2 grows, its decay is very slow. Actually, the ratios between the correlations computed at the same r, for different pairs of redshifts, are very similar to the predictions of linear theory. We find that the redshift evolution of the cross-correlation function can be approximately described by the relation:

$$\xi(s, z_1, z_2) \simeq [\xi(s, z_1)\xi(s, z_2)]^{1/2} [1 - 2\Theta(s - 1)] \tag{3.27}$$

where the quantity $s = r/r_{0C}(z)$ is introduced in order to take into account the shifting of the first zero crossing of $\xi(r, z)$ and $\Theta(x)$ is the Heaviside step function. Moreover, the first zero crossing radius of $\xi(r, z_1, z_2)$ is nearly given by the geometric average of $r_{0C}(z_1)$ and $r_{0C}(z_2)$. For $s > 0.1$ equation (3.27), which is meaningful up to the scale at which the first of the two $\xi(s, z)$ reaches its second zero crossing, reproduces $\xi(s, z_1, z_2)$ with an accuracy of ~ 5 per cent. Anyway, for $s \gtrsim 2$, the usual relation $\xi(r, z_1, z_2) \simeq [\xi(r, z_1)\xi(r, z_2)]^{1/2}\mathrm{sign}[\xi(r, z_1)]$ deriving from linear theory is preferable.

We can now check the accuracy of the approximation introduced by MCLM that consists of computing the theoretical estimate for the observed correlation function by replacing $\xi(r, z_1, z_2)$ with $\xi(r, \bar{z})$, where $\bar{z} = (z_1 + z_2)/2$, in the appropriate formula. For simplicity (and in order to isolate the phenomenon of clustering evolution) we will assume no bias, no selection effects and a constant comoving number density in an Einstein-de Sitter universe. In this case, equation 15 of MCLM reduces to:

$$\xi_{\mathrm{obs}}(r, z_{\mathrm{min}}, z_{\mathrm{max}}) = \frac{\displaystyle\int_{z_{\mathrm{min}}}^{z_{\mathrm{max}}} \frac{2 + z_1 - 2(1 + z_1)^{1/2}}{(1 + z_1)^{5/2}} \frac{2 + z_2 - 2(1 + z_2)^{1/2}}{(1 + z_2)^{5/2}} \xi(r, z_1, z_2)\, dz_1 dz_2}{\left[\displaystyle\int_{z_{\mathrm{min}}}^{z_{\mathrm{max}}} \frac{2 + z - 2(1 + z)^{1/2}}{(1 + z)^{5/2}} dz\right]^2} \tag{3.28}$$

where we denoted by $\xi_{\mathrm{obs}}(r, z_{\mathrm{min}}, z_{\mathrm{max}})$ the (ensemble averaged) theoretical estimate for the two-point correlation function measured by an observer who acquires data from the region of his past light cone corresponding to the redshift interval $[z_{\mathrm{min}}, z_{\mathrm{max}}]$.

Considering only the linear evolution of density fluctuations, $\xi(r, z_1, z_2) = \xi(r, 0, 0)/[(1+z_1)(1+z_2)]$, the integrals contained in equation (3.28) can be analytically performed. In this case, the quantity $\xi_{\mathrm{obs}}(r, z_1, z_2)/\xi(r, 0, 0)$ does not depend on r; for example we obtain $\xi_{\mathrm{obs}}(r, 0, 2)/\xi(r, 0, 0) \simeq 0.224$ and $\xi_{\mathrm{obs}}(r, 0, 1)/\xi(r, 0, 0) \simeq 0.375$. In this regime, we find that the approximation for ξ_{obs} introduced by MCLM is accurate to $2 - 3$ per cent.

In order to extend our analysis also to mildly non-linear evolution, we numerically computed ξ_{obs} by using the cross-correlation given in equation (3.14). The result obtained for $[z_{\mathrm{min}}, z_{\mathrm{max}}] = [0, 2]$ is shown in Fig. 3.6. In this case as well, ξ_{obs} looks like the usual correlation function evaluated at some intermediate redshift. We then tested the accuracy of the above-mentioned simplified scheme for the computation of ξ_{obs}, finding good agreement between the exact and the rough estimates (excluding a small neighbourhood of the zero-crossing radius of $\xi(r, z_1, z_2)$, where the approximated method breaks down, we find a maximum discrepancy of 6 per cent for $[z_{\mathrm{min}}, z_{\mathrm{max}}] = [0, 2]$ and of 3 per cent for $[z_{\mathrm{min}}, z_{\mathrm{max}}] = [0, 1]$). The simplified procedure to compute ξ_{obs} can be further improved: adopting a different way of performing the average between redshifts, namely $1 + \bar{z} = [(1+z_1)(1+z_2)]^{1/2}$, ensures more accurate predictions (in this case the maximum error is always of the order of 1 per cent). This higher precision is probably due to the fact that we are considering mildly non-linear scales and the latter approximation gives exact results for linear evolution.

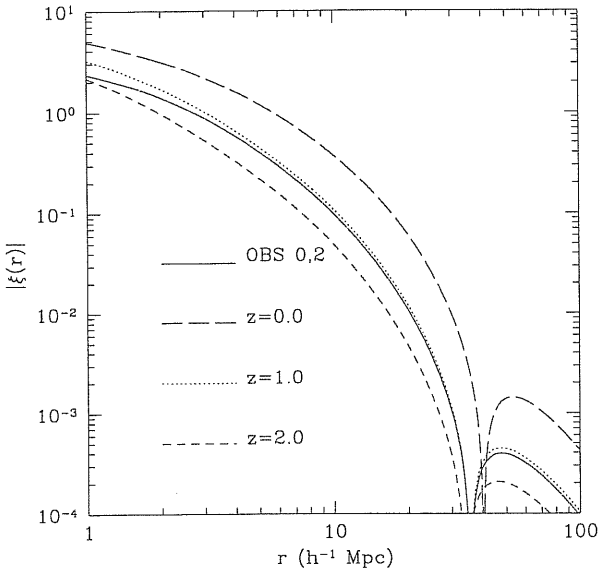


Figure 3.6: The observed two-point correlation function computed using equation (3.28) for a CDM model with $[z_{\min}, z_{\max}] = [0, 2]$. For comparison, the corresponding $\xi(r, z)$ evaluated for $z = 0, 1, 2$ are plotted.

3.3.4 Summary and conclusions

In this chapter, we have studied in detail the evolution of the mass two-point correlation function by describing the growth of density perturbations through ZA. Our motivations were originated by the well-known ability of ZA to reproduce the weakly non-linear regime of gravitational dynamics. On scales that are not affected by shell-crossing, we found that the correlation function steepens as the clustering amplitude increases. Moreover, we showed that non-linear interactions are able to move the first zero crossing of $\xi(r)$ and we gave a quantitative description of this shifting for a CDM linear spectrum.

We then compared our results with the predictions of the scaling Ansatz for clustering evolution formulated by JMW, obtaining remarkable agreement between the correlations on mildly non-linear scales and on completely linear scales. However, between these two regimes, the JMW prescription, which has been obtained requiring the resulting correlation to reproduce the linear behaviour on large scales, predicts smaller clustering amplitudes than does ZA. We think that this disagreement is caused by the smallness of the box used by JMW to perform their N-body simulations. Actually, imposing the requirement of having to match the linear solution where $\bar{\xi}_{\rm L} \to 0$, without having any constraint from numerical data on quasi-linear scales, could alter the accuracy of the fitting function that embodies the scaling Ansatz. In connection with this hypothesis, we compared ZA predictions on correlations with the output of a different scaling Ansatz calibrated against large box simulations by BG. In effect, on large scales, the BG formula agrees better with ZA, keeping the same accuracy as the JMW fit on intermediate scales.

On the other hand, the reliability of ZA on these scales and for dynamically evolved fields ($\sigma_8 \gtrsim 1$) should be verified by directly comparing its predictions with the results of other approximations and numerical simulations.

Finally, we studied the evolution of the cross-correlation between evaluations of the density field at two different epochs and, adopting the method introduced by MCLM, we used our results to compute the theoretical prediction for the observed correlation function deriving from a deep

catalogue of objects. In this context, we proposed a simplified procedure for the computation of $\xi_{\rm obs}$ that, at least for quasi-linear scales, significantly improves another approximation previously introduced by MCLM. This result confirms that the MCLM method can be used to make quantitative predictions about clustering evolution that find a direct observational counterpart in the analysis of deep surveys.

Chapter 4

Abundance and clustering of dark matter haloes

> "Can larger mass objects form from the non-linear interaction of smaller masses? We start with two ideas suggested by Peebles (1965; 1972): that the development of larger scales of condensation may have been sequential; and specifically, that statistical randomness in an 'incoherent dust' model induces the growth of larger instabilities with increasing time. We outline ... a model which is initially grainy and remains so at all times, starting with an initial spectrum of 'small' masses which are supposed to have been formed by other processes. The initial time is taken as recombination or soon after. As the expanding Friedmann cosmology evolves, the mass points condense into aggregates which (when they are themselves sufficiently bound) we identify as single particles of a larger mass."
>
> Press & Schechter 1974

> "With much of the matter in the Universe now concentrated in these discrete clouds, there is at any point a gravitational field, of the general order of GM_c/L, where L is the mean distance between the clouds (and M_c the typical mass of the clouds). A cloud thus tends to move toward nearby, higher density regions, leading to clustering of the clouds."
>
> Peebles 1965

> "In this way, the condensation process proceeds to larger scales."
>
> Press & Schechter 1974

> "In smaller subsystems this process occurred earlier, and the subsequent collapse, or adjustement to equilibrium is already well advanced. This process should be pictured as a roughly continuous progression along a hierarchy in the departure from the general expansion."
>
> Peebles 1965

This process of hierarchical clustering represents a useful mathematical idealization to describe the development of structure in a wide class of cosmological models. [1] In analogy with the

[1] Notice, however, that the discussion of structure formation in terms of gravitationally bound units and sub-units is somewhat naive since it neglects all those phenomena that describe the coherence of structure on large scales (e.g. the complex network of filaments that develops in numerical simulations of clustering growth).

self-similar behaviour of scale-free models that we discussed in §2.1, we expect this bottom-up evolution to manifest itself whenever the contribution to the variance of δ coming from modes that correspond to a logarithmic interval of k is a monotonically decreasing function of the scale $1/k$. This corresponds to the condition $n_{\rm eff} > -3$, with $n_{\rm eff} \equiv d\ln P(k)/d\ln k$ the effective spectral index. On the other hand, $n_{\rm eff} < 1$ is required to avoid that the large-scale dynamics is dominated by non-linear coupling of perturbations with small wavelength (e.g. Williams et al. 1991). Moreover, the clustering hierarchy is (formally) infinite if $\sigma^2 \to \infty$ as $R_{\rm f} \to 0$. Notice that the popular CDM models satisfy all these requirements.

An important question to clarify is whether the bound units in the hierarchy preserve the internal substructure they inherit at the moment of formation, rather than relax to monolithic systems. In the first case, the clustering hierarchy keeps stable down to arbitrarily small scales (e.g. Peebles 1978), while, in the second picture, it is washed out below the scale corresponding to the typical radius of the most recently collapsed objects (e.g. White & Rees 1978). Numerical simulations support this second model in which relaxation processes lead to the formation of virialized clumps with smooth density profiles (e.g. Navarro, Frenk & White 1995, 1996, 1997; Tormen, Bouchet & White 1997; Tormen, Diaferio, Syer 1998; Ghigna et al. 1998). Hereafter we will use the term dark matter haloes to refer to these objects. In general, the process of halo formation is comprised of a merging phase (during which two or more clumps blend to form a larger object) and a subsequent relaxing phase (during which substructure is erased and a quasi-equilibrium configuration is created). Only those subunits which contain less than a few per cent of the mass of the main halo may preserve their identity for cosmologically relevant times (e.g. Tormen, Diaferio & Syer 1998).

We expect that the baryonic gas, cooling and condensing within the potential wells generated by these haloes, may eventually turn into stars thus leading to galaxy formation (White & Rees 1978). A detailed understanding of the hierarchical formation of dark matter haloes should therefore provide the fundamental skeleton in order to develop models of galaxy formation (e.g. Kauffmann, White & Guiderdoni 1993). For this purpose, a minimal description of the halo population should then provide detailed information about their

1. abundance and mass spectrum;

2. spatial distribution;

3. formation history through mass accretion and mergers;

4. internal structure;

5. angular momentum distribution.

at any redshift. In this chapter, we address the first three points of the previous list. In particular, we first review the Press-Schechter (Press & Schechter 1974) theory for the mean mass distribution of the halo population, together with its excursion set extension (e.g. Bond et al. 1991). An original algorithm for calculating the halo correlations in Lagrangian space is then presented. This model, which is fully described in Porciani et al. (1998), is part of a two-step procedure which aims at computing the halo correlations in Eulerian space (Catelan et al. 1998; Catelan, Matarrese & Porciani 1998) The complete theory will be introduced in chapter 7. The remaining issues about the halo density profile and angular momentum distribution are widely discussed in Navarro, Frenk & White (1995, 1996, 1997) and Cole & Lacey (1996; see also Catelan & Theuns 1996a,b for an analytical approach), respectively.

4.1 The Press-Schechter model for the mass function

Press & Schechter (1974, hereafter PS) proposed a simple model to compute the comoving number density of collapsed haloes directly from the statistical properties of the linear density field, assumed to be Gaussian (see also Doroshkevich 1967). According to the PS theory, a patch of fluid is part of a collapsed region of scale larger than $M \propto R_{\rm f}^3$ if the value of the smoothed linear density contrast on the same scale exceeds a suitable threshold $t_{\rm f}$. The idea is to use a global threshold in order to mimic non-linear dynamical effects ending up with halo collapse and virialization. An exact value for $t_{\rm f}$ can be obtained by describing the evolution of density perturbations according to the spherical top-hat model. In this case, a fluctuation of amplitude δ will collapse at a redshift $z_{\rm f}$ such that $\delta(\mathbf{x}) = t_{\rm f} \equiv \delta_c/D(z_{\rm f})$. [2] In the Einstein–de Sitter universe and during the matter-dominated era the critical value δ_c does not depend on any cosmological parameter and is given by $\delta_c = 3(12\pi)^{2/3}/20 \simeq 1.686$, whereas for general cosmologies it shows a weak dependence on the value of the density parameter, the cosmological constant, the Hubble constant, thus on redshift (e.g. Lacey & Cole 1993).

The PS argument proceeds in probabilistic terms. For a given $R_{\rm f}$, the probability that a given mass element lies in a region with $\delta > t_{\rm f}$ is

$$p(\delta > t_{\rm f}|M) = \frac{1}{2}\,{\rm erfc}\left(\frac{t_{\rm f}}{\sqrt{2}\sigma(M)}\right)\,, \tag{4.1}$$

where $\sigma(M)$ is the linear rms value of the density field smoothed on the scale $R_{\rm f}$. According to the PS interpretation, $p(\delta > t_{\rm f}|M)$ corresponds to the mass fraction contained in haloes more massive than M. Notice that, when $t_{\rm f}/\sigma(M) \to 0^+$ (i.e. for infinitesimally small masses), $p(\delta > t_{\rm f}|M) \to 1/2$. This is because, as we discussed in §2.4.1, spherical perturbations can collapse only if they are overdense (i.e. $t_{\rm f} > 0$), and in a Gaussian field the PDF of δ is symmetric around $\delta = 0$ for every smoothing length. However, in a hierarchical scenario where $\sigma(R_{\rm f}) \to \infty$ as $R_{\rm f} \to 0$, one expects to find all the mass collapsed in objects of some scale. In fact, at any point, $\delta(R_{\rm f})$ must be larger than $t_{\rm f}$ for a sufficiently small value of $R_{\rm f}$. PS faced the problem by assuming that the mass contained in the initially underdense regions accretes onto the collapsed lumps. Basing their analysis on a simple spherical model, they found that the net effect is to double the mass of every halo without altering the shape of the mass distribution. Thus, the mass function $n(M, z_{\rm f})$ [where $n(M, z_{\rm f})dM$ denotes the comoving number density of haloes with mass between M and $M + dM$ collapsed at redshift $z_{\rm f}$] is obtained by differentiating the cumulative probability $2\,p(\delta > t_{\rm f}|M)$ with respect to M and multiplying the resulting expression by the factor ρ_b/M [3]

$$\begin{aligned} n(M, z_{\rm f})\,dM &= \frac{\rho_b\, t_{\rm f}(z_{\rm f})}{\sqrt{2\pi}}\,\frac{1}{M^2\sigma(M)}\left|\frac{d\ln\sigma^2(M)}{d\ln M}\right|\,\exp\left(-\frac{t_{\rm f}^2(z_{\rm f})}{2\,\sigma^2(M)}\right)\,dM \\ &= \frac{1}{\sqrt{2\pi}}\,\frac{\rho_b}{M}\,\frac{\delta_c(z)}{\sigma^3(M)}\,\exp\left[-\frac{\delta_c^2(z)}{2\sigma^2(M)}\right]\left|\frac{d\sigma^2(M)}{dM}\right|\,. \end{aligned} \tag{4.2}$$

For scale-free spectra, $\sigma^2(M) = (M/M_0)^{-(n+3)/3}$ and one obtains

$$n(M, z_{\rm f}) \propto M^{\frac{n-9}{6}}\,\exp\left[-\frac{1}{2}\left(\frac{M}{M_*}\right)^{\frac{n+3}{3}}\right]\,, \tag{4.3}$$

[2] In this chapter, we deal with the linear density field only. Since no ambiguity is possible, in order to simplify the notation, the symbol δ will be used here to denote the quantity δ_ℓ introduced in chapter 2.

[3] It is worth stressing that PS did not give a convincing argument for their procedure of re-normalization of the mass function, and the factor of two they introduced almost arbitrarily has come to be called the "fudge factor" of the PS theory.

i.e. the mass function decreases as a power-law with increasing M and is exponentially damped at the scale $M_* = t_{\rm f}^{-6/(n+3)} M_0$ such that $\sigma(M_*) = t_{\rm f}$. In this simple case, the mass function displays self-similar evolution. Notice that $n(M, z_{\rm f})$ diverges as $M \to 0$, but the total mass contained by haloes which are smaller than any given reference value is finite.

The PS formula has been thoroughly compared with the outcomes of N-body simulations, finding general good agreement (e.g. Efstathiou et al. 1988; Gelb & Bertschinger 1994; Lacey & Cole 1994). Recent analyses, however, emphasized that the halo mass function derived by the PS argument over-estimates the number density of haloes with mass $M < M_*$ by a factor of $1.5 - 2$ (Lacey & Cole 1994; Gross et al. 1998; Lemson & Kauffmann 1999; Kauffmann et al. 1999a).

4.2 Excursion set approach

The original PS model is intrinsically flawed by the cloud-in-cloud problem, namely the fact that a fluctuation on a given scale can contain substructures of smaller scales, and the same fluid elements can be assigned, according to the PS collapse criterion, to haloes of different mass. The problem was later solved by several authors (Peacock & Heavens 1990; Bond et al. 1991, hereafter BCEK; Cole 1991) according to the so-called 'excursion set' approach, by calculating the distribution of first-passage 'times' across the collapse threshold for suitably defined random walks. Lacey & Cole (1993, 1994) implemented these ideas to study the merger rates of virialized haloes in hierarchical models of structure formation.

In this section we present a straightforward extension of the excursion set approach and review the standard application of this method which allows to determine the mean mass function of DM haloes.

4.2.1 Langevin equation

Let us assume that the linear density fluctuations form a homogeneous and isotropic Gaussian random field $\delta(\mathbf{x}, z)$, with $\mathbf{x}$ the Lagrangian comoving position and z the redshift. In linear theory one can factor out the redshift dependence as $D(z)\delta(\mathbf{x})$, where $D(z)$ is the linear growth factor. With the normalization $D(z = 0) = 1$, $\delta(\mathbf{x})$ becomes the mass density fluctuation linearly extrapolated to the present epoch.

We want now to study the statistical properties of the density fluctuation field at some resolution scale $R_{\rm f}$. This is introduced as in eq. (2.50). The detailed properties of $\delta(\mathbf{x}, R_{\rm f})$ clearly depend upon the specific choice of window function. The most commonly used smoothing kernels are the top-hat filter Heaviside step function, and the Gaussian one. Recently, for convenience of analysis, top-hat filtering has been also applied in momentum space $\widetilde{W}_{SKS}(k, R_{\rm f}) = \Theta(k_{\rm f} - k)$, where $k_{\rm f} = 1/R_{\rm f}$ and $k_{\rm f} = |\mathbf{k}_{\rm f}|$. This kernel is generally called sharp k-space filter. While it is easy to associate a mass to real space top-hat filtering $M_{TH}(R_{\rm f}) = 4\pi\rho_b R_{\rm f}^3/3$, there is always a bit of arbitrariness in assigning a mass to the other window functions. The most common procedure is to multiply the average density by the volume enclosed by the filter, obtaining $M_G(R_{\rm f}) = (2\pi)^{3/2}\rho_b R_{\rm f}^3$ and $M_{SKS}(R_{\rm f}) = 6\pi^2\rho_b k_{\rm f}^{-3}$ (Lacey & Cole 1993). An alternative procedure, originally introduced by Bardeen et al. (1986), corresponds to the choice $M_{SKS}(R_{\rm f}) = 4\pi\rho_b R_{TH}^3/3$, where R_{TH} is chosen by requiring $\sigma^2_{SKS}(R_{\rm f}) = \sigma^2_{TH}(R_{TH})$, and similarly for the Gaussian filter. In this way one obtains good agreement with numerical simulations of clustering growth (Lacey & Cole 1994).

As discussed in the introduction, in hierarchical bottom-up models of structure formation at each epoch mass fluctuations have typically undergone non-linear collapse up to some scale $R_{\rm f}$, and the first objects that form are those of lower mass. Haloes of larger mass arise by the merging and accretion of subunits. The mass distribution deriving from this hierarchical halo formation process

has been successfully modeled by Peacock & Heavens (1990), BCEK, Cole (1991) and Lacey & Cole (1993). In order to mimic the accretion of matter, all these models consider a full hierarchy of decreasing resolution scales $R_{\rm f}$. The effect of varying $R_{\rm f}$ can be obtained by differentiating eq. (2.50)

$$\frac{\partial\delta(\mathbf{x},R_{\rm f})}{\partial R_{\rm f}} = \frac{1}{(2\pi)^3}\int d^3k\,\tilde{\delta}(\mathbf{k})\,\frac{\partial\widetilde{W}(kR_{\rm f})}{\partial R_{\rm f}}\,{\rm e}^{-i\mathbf{k}\cdot\mathbf{x}} \equiv \eta(\mathbf{x},R_{\rm f})\ . \tag{4.4}$$

This has the form of a Langevin equation, which shows how an infinitesimal change of the resolution scale $R_{\rm f}$ affects the value of the density fluctuation field $\delta(\mathbf{x},R_{\rm f})$ in the given position $\mathbf{x}$ through the action of the stochastic force $\eta(\mathbf{x},R_{\rm f})$. In the limit $R_{\rm f}\to\infty$ one has $\delta(\mathbf{x};R_{\rm f})\to 0$, which can be adopted as initial condition for our first-order stochastic differential equation. Thus, by solving it, we can associate to each point $\mathbf{x}$ a trajectory $\delta(\mathbf{x},R_{\rm f})$ obtained by varying the resolution scale $R_{\rm f}$. Notice that, since eq. (4.4) is linear, $\eta(\mathbf{x},R_{\rm f})$ is also a zero mean Gaussian random field, uniquely specified by its auto-correlation function

$$\langle\eta(\mathbf{x}_1,R_{f1})\,\eta(\mathbf{x}_2,R_{f2})\rangle = \frac{1}{2\pi^2}\int_0^\infty dk_1\,k_1^2\,P(k_1)\,\frac{\partial\widetilde{W}(k_1R_{f1})}{\partial R_{f1}}\,\frac{\partial\widetilde{W}(k_1R_{f2})}{\partial R_{f2}}\,j_0(k_1r)\ , \tag{4.5}$$

where $r=|\mathbf{x}_1-\mathbf{x}_2|$, and $j_0(x)$ is the zeroth order spherical Bessel function. Trajectories associated to different neighbouring points will be statistically influenced by the correlation properties of the force $\eta(\mathbf{x},R_{\rm f})$, i.e. of the underlying Gaussian field $\delta(\mathbf{x})$. On the other hand, the coherence of each trajectory along the $R_{\rm f}$ direction depends exclusively on the analytic form of the filter function, and vanishes for the sharp k-space window (BCEK). With such a filter, by decreasing the smoothing length one adds up a new set of Fourier modes of the unsmoothed distribution to $\delta(\mathbf{x},R_{\rm f})$. For a Gaussian field this is completely independent of the previous increments, and each trajectory $\delta(\mathbf{x},R_{\rm f})$ becomes a Brownian random walk.

In the case of sharp k-space filtering the notation greatly simplifies if we use as time variable the variance of the filtered density field, $\Lambda\equiv\sigma^2(k_{\rm f})=\langle\delta(k_{\rm f})^2\rangle=(2\pi^2)^{-1}\int_0^{k_{\rm f}}dk\,k^2\,P(k)$. In such a case the stochastic process reduces to a Wiener one, namely

$$\frac{\partial\delta(\mathbf{x},\Lambda)}{\partial\Lambda} = \zeta(\mathbf{x},\Lambda)\ , \tag{4.6}$$

with $\langle\zeta(\mathbf{x},\Lambda)\rangle=0$ and

$$\langle\zeta(\mathbf{x},\Lambda_1)\,\zeta(\mathbf{x},\Lambda_2)\rangle = \delta_{_D}(\Lambda_1-\Lambda_2) \tag{4.7}$$

(see equation 4.19 for the spatial correlation). In the following we will adopt Λ as time variable, unless explicitly stated. The solution of the Langevin equation (4.6) in an arbitrary point of space (the position $\mathbf{x}$ is here understood), with the initial condition $\delta(\Lambda=0)=0$, is simply $\delta(\Lambda)=\int_0^\Lambda d\Lambda'\zeta(\Lambda')$. By ensemble averaging this expression one obtains $\langle\delta(\Lambda)\rangle=0$ and $\langle\delta(\Lambda_1)\delta(\Lambda_2)\rangle=\min(\Lambda_1,\Lambda_2)$, which uniquely determine the Gaussian distribution of $\delta(\Lambda)$. More details can be found in Porciani et al. (1996).

4.2.2 The mass function

The solution of the cloud-in-cloud problem has been given by Peacock & Heavens (1990), BCEK and Cole (1991). Their approach consists in considering at any given point the trajectory $\delta(R_{\rm f})$ as a function of the filtering radius, and then determining the *largest* $R_{\rm f}$ at which $\delta(R_{\rm f})$ upcrosses

the threshold $t_{\rm f}(z_{\rm f})$ corresponding to the formation redshift $z_{\rm f}$. This determines the largest mass collapsed at that point, all sub-structures having been erased. So the computation of the mass function is equivalent to calculating the fraction of trajectories that first upcross the threshold $t_{\rm f}$ as the scale M decreases. The solution of the problem is enormously simplified for Brownian trajectories, that is for sharp k-space filtered density fields. In such a case one only has to solve the Fokker-Planck equation for the probability density $\mathcal{W}(\delta,\Lambda)\,d\delta$ that the stochastic process at Λ assumes a value in the interval $\delta,\delta+d\delta$,

$$\frac{\partial\mathcal{W}(\delta,\Lambda)}{\partial\Lambda}=\frac{1}{2}\,\frac{\partial^2\mathcal{W}(\delta,\Lambda)}{\partial\delta^2}\,, \tag{4.8}$$

with the absorbing boundary condition $\mathcal{W}(t_{\rm f},\Lambda)=0$ and initial condition $\mathcal{W}(\delta,0)=\delta_D(\delta)$. The solution is well known (Chandrasekhar 1943)

$$\mathcal{W}(\delta,\Lambda;t_{\rm f})\,d\delta=\frac{1}{\sqrt{2\pi\Lambda}}\left[\exp\left(-\frac{\delta^2}{2\Lambda}\right)-\exp\left(-\frac{(\delta-2t_{\rm f})^2}{2\Lambda}\right)\right]d\delta\,. \tag{4.9}$$

Defining $S(\Lambda,t_{\rm f})=\int_{-\infty}^{t_{\rm f}}d\delta\,\mathcal{W}(\delta,\Lambda,t_{\rm f})$ as survival probability of the trajectories, one obtains the density probability distribution of first-crossing variances by differentiation

$$\begin{aligned}\mathcal{P}_1(\Lambda)=-\frac{\partial S(\Lambda,t_{\rm f})}{\partial\Lambda}=-\frac{\partial}{\partial\Lambda}\int_{-\infty}^{t_{\rm f}}d\delta\,\mathcal{W}(\delta,\Lambda;t_{\rm f})=\\ =\left[-\frac{1}{2}\frac{\partial\mathcal{W}(\delta,\Lambda,t_{\rm f})}{\partial\delta}\right]_{-\infty}^{t_{\rm f}}=\frac{t_{\rm f}}{\sqrt{2\pi\Lambda^3}}\exp\left(-\frac{t_{\rm f}^2}{2\Lambda}\right)\,.\end{aligned} \tag{4.10}$$

The function $\mathcal{P}_1(\Lambda)\,d\Lambda$ yields the probability that a realization of the random walk is absorbed by the barrier in the interval $(\Lambda,\Lambda+d\Lambda)$ or, by the ergodic theorem, the probability that a fluid element belongs to a structure with mass in the range $[M(\Lambda+d\Lambda),M(\Lambda)]$. Finally, the comoving number density of haloes with mass in the interval $[M,M+dM]$ collapsed at redshift $z_{\rm f}$ is

$$n(M,z_{\rm f})\,dM=\frac{\rho_b}{M}\,\mathcal{P}_1(\Lambda)\left|\frac{d\Lambda}{dM}\right|dM\,. \tag{4.11}$$

Inserting the expression of $\mathcal{P}_1(\Lambda)$ of eq. (4.10) in the latter equation one obtains the well-known PS expression for the mass function

$$n(M,z_{\rm f})\,dM=\frac{\rho_b\,t_{\rm f}(z_{\rm f})}{\sqrt{2\pi}}\,\frac{1}{M^2\sqrt{\Lambda(M)}}\left|\frac{d\ln\Lambda}{d\ln M}\right|\exp\left(-\frac{t_{\rm f}(z_{\rm f})^2}{2\,\Lambda(M)}\right)dM\,. \tag{4.12}$$

The original fudge factor of 2 of the PS approach is now naturally justified. A generalization of this formalism to simple cases of non-Gaussian initial conditions has been given by Porciani et al. (1996).

Previous investigations (e.g. Peacock & Heavens 1990; BCEK) have shown that only for sharp k-space filtering it is possible to write an analytic formula for the mass function obtained from the excursion set approach. Numerical solutions of the cloud-in-cloud problem with physically more acceptable smoothing kernels like Gaussian and top-hat result in mass functions that are a factor of two lower in the high-mass tail and have different small-mass slopes compared with the PS formula. The standard interpretation of this result is that the excursion set method is not reliable for $M\ll M_*$. We remember that M_*, defined by $\Lambda(M_*)=t_{\rm f}^2$, is the typical mass collapsing at $z_{\rm f}$.

4.3 Lagrangian clustering

4.3.1 Introduction

In the following sections we provide an exact treatment of the halo correlation properties in the Lagrangian world, which extends and puts on sounder bases the results obtained by Mo & White (1996; see chapter 6), as also sketched in Catelan et al. (1998; hereafter CLMP).

Our formalism has to be considered in every respect as the natural extension of the BCEK model to include the spatial clustering properties of the haloes. Once a cosmological model and a power-spectrum of density fluctuations are selected the present formalism allows to construct halo correlations of any order in Lagrangian space. Here we only give explicit results for the two-point function, but the generalization to higher order would be straightforward.

An obvious and important application of this type of study aims at modeling galaxy clustering at different redshifts. The key idea being that the spatial distribution of DM haloes can be a clue for understanding the clustering properties of luminous objects like galaxies with different physical properties and at various epochs (e.g. Kauffmann, Nusser & Steinmetz 1997; Matarrese et al. 1997; Moscardini et al. 1998; Kauffmann et al. 1999b). The actual relation between the parent DM haloes and the galaxies, however, depends on the way haloes accrete mass and on how they grow by the process of merging with the surrounding haloes. These processes are usually studied in terms of the so-called 'halo merger trees', obtained either by Monte Carlo methods (e.g. Cole 1991; Lacey & Cole 1993; Kauffmann & White 1993; Tozzi, Governato & Cavaliere 1996; Somerville & Kolatt 1997; Salvador-Solé, Solanes & Manrique 1997 and references therein) or by directly looking at the dynamics of particles in cosmological N-body simulations (e.g. Kauffmann et al. 1997; Roukema et al. 1997; Governato et al. 1997). Realizations of halo merger trees can be in fact obtained by Monte Carlo techniques in the frame of the excursion set theory, where halo accretion histories are associated to random walks. A merger tree is then obtained by suitably connecting the accretion histories of all the progenitors of a given halo (which, in the tree-jargon, are usually called 'branches').

In the latter approach, however, spatial correlations between different branches of the tree as well as correlations between different trees are completely ignored. The existence of this shortcoming has been also noticed by Rodrigues & Thomas (1996) in connection with the so-called 'block model' (Cole & Kaiser 1988; Cole 1991), which can be considered as a rudimentary form of merging tree, where haloes can only grow by doubling their mass. Yano, Nagashima & Gouda (1996) and Nagashima & Gouda (1997) analysed the possible changes induced by spatial correlations on the PS mass function.

The problem within the standard Monte Carlo approach is that each trajectory is treated as being independent of all the others. In a general Gaussian field, however, this situation is not realized and, for power-spectra of cosmological interest, the density fluctuations turn out to be strongly correlated on small scales. This statistical property can be clearly detected by looking at a pair of trajectories associated to two neighbouring points. In Fig. 4.1 we show two trajectories obtained by smoothing a correlated field with a sharp k-space filter. When the smoothing length is much larger than the separation between the points in which the trajectories are computed the density contrast assumes values that are practically coincident. On the other hand, for filtering radii much smaller than the lag that separates the points, the trajectories become indeed uncorrelated. The transition from the completely correlated situation to the totally uncorrelated one is continuous, i.e. the trajectories become less and less correlated as the smoothing length decreases. For comparison, in Fig. 4.1 we also show the same trajectories modified by artificially removing any correlation between them. It is clear that the joint distribution of first upcrossings at different

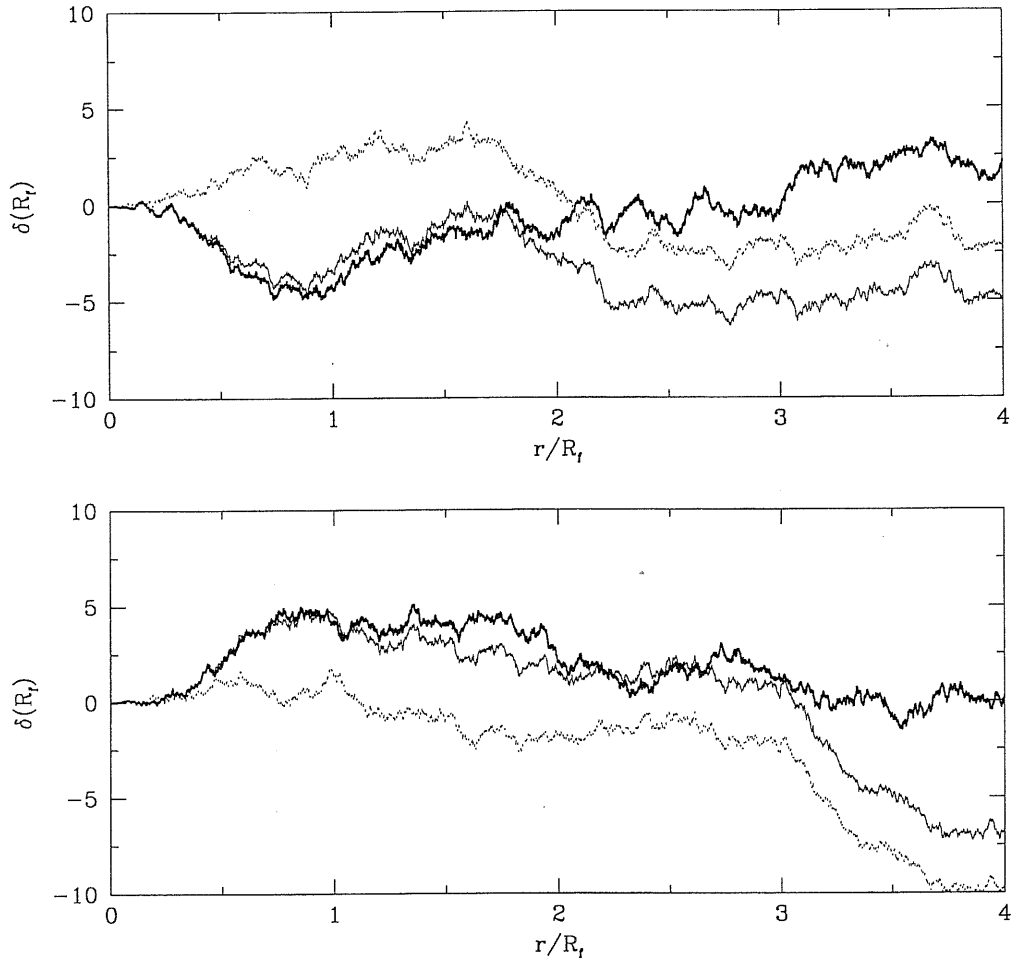


Figure 4.1: Examples of trajectories obtained by smoothing the linear density fluctuation field with a series of sharp k-space filters (with decreasing resolution scales $R_{\rm f} = 1/k_{\rm f}$) at two points separated by the distance r. Here we consider $r = 4\,{\rm Mpc}$ and a CDM power spectrum (with density parameter $\Omega = 1$ and present Hubble constant $H_{\rm o} = 50\,{\rm km\,s^{-1}\,Mpc^{-1}}$) linearly extrapolated to $\sigma_8 = 1$. The *heavy continuous* line represents the trajectory associated to a point $\mathbf{x}$ in Lagrangian space. The trajectory associated to the point $\mathbf{y}$ such that $|\mathbf{y} - \mathbf{x}| = r$ is plotted with a *light continuous* line. The *dotted* line is obtained by: *i*) considering the trajectory associated to $\mathbf{y}$, *ii*) artificially removing any correlation with the trajectory at $\mathbf{x}$ and *iii*) suitably rescaling the result. In practice, the dotted line represents a trajectory which is completely independent of the one associated to $\mathbf{x}$, but which has the same statistical one-point properties. This is what is generally used in building up merger trees of dark matter haloes. However, the actual trajectories associated to neighbouring points are strongly correlated when the smoothing length $R_{\rm f}$ is comparable to the physical separation between the points r. The same trajectories tend to become less and less correlated as $R_{\rm f}$ decreases. It is interesting to note that a memory of the strong correlation existing for $r/R_{\rm f} \ll 1$ remains even when $r/R_{\rm f} \gg 1$. In fact, when $R_{\rm f}$ is small compared to r, the real trajectory associated to $\mathbf{y}$ can be accurately approximated by adding a constant value to the independent trajectory so that it matches the random walk at $\mathbf{x}$ when $R_{\rm f} \gtrsim r$. In other words, since the trajectories are continuous functions of $R_{\rm f}$, the value assumed at $R_{\rm f} \sim r$ (i.e. at the end of the strongly correlated regime) represents the initial condition for the random walk at small $R_{\rm f}$ (i.e. in the independent regime). This derives from a sort of natural peak-background split: by evaluating the density fluctuation field in two points separated by r one samples the same background value (obtained by summing up all the Fourier components corresponding to wavelenghts larger than r) but different high frequency modes.

points is heavily influenced by the correlations of the underlying density fluctuation field.

In this section we will i) compute the joint two-point distribution of the first upcrossing events and use the solution to obtain an approximate analytical expression for the halo-halo correlation function, ii) numerically solve suitable sets of spatially correlated Langevin equations to directly compute the halo two-point function at different lags. In a forthcoming paper we will consider the effects of correlations on the merging histories of dark matter haloes.

The plan of the section is as follows. In section 4.3 we study the Lagrangian clustering of haloes identified by our scheme and obtain the halo-halo correlation function both from an approximate analytic treatment and from a numerical integration of two spatially correlated Langevin equations. Section 4.4 contains a discussion of our results and hints at some applications of our algorithm to generate spatially correlated halo merger trees.

4.3.2 Two-point correlation function from joint upcrossing distribution

In this section we show how the excursion set approach can be extended to derive the clustering properties of dark matter haloes. Specifically, we study the evolution of a set of density fluctuation processes at different spatial locations, as the smoothing scale progressively shrinks. The trajectories turn out to be intimately correlated, and the joint distribution of the first upcrossing filtering radii is used to extract the Lagrangian halo-halo correlation function.

Let us select two points in Lagrangian space $\mathbf{x}_1$ e $\mathbf{x}_2 = \mathbf{x}_1 + \mathbf{r}$. We want to study the evolution of the stochastic processes $\delta_1(\Lambda) = \delta(\mathbf{x}_1, \Lambda)$ and $\delta_2(\Lambda) = \delta(\mathbf{x}_2, \Lambda)$ as Λ varies. In particular, let us suppose that we know the joint probability distribution $\mathcal{P}_2(\Lambda_1, \Lambda_2; \mathbf{x}_1, \mathbf{r})$ of those pairs of variances (Λ_1, Λ_2) that correspond to the first upcrossing scales of the threshold $t_{\rm f}$ by the two processes $\delta_1(\Lambda)$ and $\delta_2(\Lambda)$. Because of the underlying homogeneity and isotropy, the probability density $\mathcal{P}_2$ cannot depend on the vector $\mathbf{x}_1$ and on the orientation of $\mathbf{r}$, i.e. $\mathcal{P}_2(\Lambda_1, \Lambda_2; \mathbf{x}_1, \mathbf{r}) = \mathcal{P}_2(\Lambda_1, \Lambda_2; r)$. Moreover, by the ergodic theorem one can identify $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$ with the probability distribution of the pairs (Λ_1, Λ_2), obtained by randomly selecting points spatially separated by the distance r, within a given realization of the density field. Finally, following the arguments given in the previous section, we can interpret $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)\, d\Lambda_1 d\Lambda_2$ as the probability of finding two points separated by r within two haloes with mass in the intervals $(M_1 - dM_1, M_1)$ and $(M_2 - dM_2, M_2)$, as fixed by the corresponding variance ranges $(\Lambda_1, \Lambda_1 + d\Lambda_1)$ and $(\Lambda_2, \Lambda_2 + d\Lambda_2)$. As we will discuss in detail in the next sections, the probability density $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$ can be obtained by integrating the system of correlated Langevin equations that describe the evolution of the processes $\delta_1(\Lambda)$ and $\delta_2(\Lambda)$.

In order to compute the halo-halo correlation function we adopt the following procedure. A class of objects is selected by the mass interval corresponding to the Λ-range $I \equiv [\Lambda_{\rm min}, \Lambda_{\rm max}]$. The probability of determining two points separated by r contained within collapsed objects of class I is

$$\mathcal{P}_{II}(r) = \int_I \int_I d\Lambda_1\, d\Lambda_2\, \mathcal{P}_2(\Lambda_1, \Lambda_2; r)\ . \tag{4.13}$$

Similarly, the probability of finding a point contained in an object of type I is $\mathcal{P}_I = \int_I d\Lambda\, \mathcal{P}_1(\Lambda)$. From the definition of correlation function we then obtain

$$\xi^{pts}_{II}(r) = \frac{\mathcal{P}_{II}}{\mathcal{P}_I^2} - 1 = \frac{\int_I \int_I d\Lambda_1\, d\Lambda_2\, \mathcal{P}_2(\Lambda_1, \Lambda_2; r)}{\left[\int_I d\Lambda\, \mathcal{P}_1(\Lambda)\right]^2} - 1\ . \tag{4.14}$$

In a similar fashion, considering disjoint classes I_1 and I_2,

$$\xi^{pts}_{I_1 I_2}(r) = \frac{\int_{I_1} \int_{I_2} d\Lambda_1\, d\Lambda_2\, \mathcal{P}_2(\Lambda_1, \Lambda_2; r)}{\int_{I_1} d\Lambda\, \mathcal{P}_1(\Lambda)\ \int_{I_2} d\Lambda\, \mathcal{P}_1(\Lambda)} - 1\ . \tag{4.15}$$

We stress that the quantities $\xi_{II}^{pts}(r)$ and $\xi_{I_1I_2}^{pts}(r)$ are the Lagrangian correlations of the mass elements contained in the collapsed haloes. This quantity has been recently used by Bagla (1998) to study the evolution of galaxy clustering. However, we are ultimately concerned with the calculation of the halo correlations $\xi_{II}^{hh}(r)$, so we have to weigh properly the statistical contribution for each extended halo. The problem shows particularly simple if we adopt the PS theory. Indeed, in this case, the sets of points where the first upcrossings occur at the same Λ are point-like disconnected regions (see Bond & Myers 1996 for a different approach). Only in a statistical sense they originate collapsed haloes, each contributing by $1/V(\Lambda)$, where $V(\Lambda) \equiv M/\rho_b$ is the typical Lagrangian volume of an object of mass M associated to the variance Λ. Therefore, adopting the same procedure used to calculate the mass function, the mean number density of collapsed objects of scale Λ becomes $n(\Lambda) = \mathcal{P}_1(\Lambda)/V(\Lambda)$. Similarly, for the distribution of pairs at distance r, we define

$$n_2(\Lambda_1, \Lambda_2; r) = \frac{\mathcal{P}_2(\Lambda_1, \Lambda_2; r)}{V(\Lambda_1)V(\Lambda_2)} \,. \tag{4.16}$$

The idea is, once again, to allow for the finite size of the haloes. The halo-halo correlations become, respectively,

$$\xi_{II}^{hh}(r) = \frac{n_{II}}{n_I^2} - 1 \equiv \frac{\int_I \int_I d\Lambda_1 d\Lambda_2 \, n_2(\Lambda_1, \Lambda_2; r)}{\left[\int_I d\Lambda \, n(\Lambda)\right]^2} - 1 \,, \tag{4.17}$$

and

$$\xi_{I_1I_2}^{hh}(r) = \frac{\int_{I_1} \int_{I_2} d\Lambda_1 d\Lambda_2 \, n_2(\Lambda_1, \Lambda_2, r)}{\int_{I_1} d\Lambda \, n(\Lambda) \int_{I_2} d\Lambda \, n(\Lambda)} - 1 \,. \tag{4.18}$$

4.3.3 Correlated Langevin equations: sharp k-space filtering

In section 4.3.1 we showed that in order to obtain the halo correlation function, it is crucial to know the joint distribution $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$. This quantity can be obtained by solving the system of equations governing the evolution of the pair of correlated processes $\delta_1(\Lambda)$ and $\delta_2(\Lambda)$:

$$\begin{cases} \dfrac{\partial \delta_1(\Lambda)}{\partial \Lambda} = \zeta_1(\Lambda) \,, & \delta_1(0) = 0 \,, \\ \dfrac{\partial \delta_2(\Lambda)}{\partial \Lambda} = \zeta_2(\Lambda) \,, & \delta_2(0) = 0 \,, \\ \langle \zeta_1(\Lambda) \rangle = \langle \zeta_2(\Lambda) \rangle = 0 \,, & \zeta_1 \text{ and } \zeta_2 \text{ Gaussian processes} \,, \\ \langle \zeta_1(\Lambda) \, \zeta_1(\Lambda') \rangle = \langle \zeta_2(\Lambda) \, \zeta_2(\Lambda') \rangle = \delta_D(\Lambda - \Lambda') \,, \\ \langle \zeta_1(\Lambda) \, \zeta_2(\Lambda') \rangle = \dfrac{\partial \xi(r;\Lambda)}{\partial \Lambda} \, \delta_D(\Lambda - \Lambda') \,. \end{cases} \tag{4.19}$$

The last equation, obtained by introducing sharp k-space filtering in equation (4.5), completes the definition of the stochastic field $\zeta(\mathbf{x}, \Lambda)$ given in equations (4.6) and (4.7). Here $\xi(r;\Lambda)$ is the linear two-point correlation function for the mass density fluctuations smoothed on the scale $R_f \equiv 1/k_f$ associated with the variance Λ

$$\xi(r;\Lambda) \equiv \frac{1}{2\pi^2} \int_0^{k_f(\Lambda)} dk \, k^2 \, P(k) \, j_0(kr) \tag{4.20}$$

and one has

$$\frac{\partial \xi(r;\Lambda)}{\partial \Lambda} = j_0[k_f(\Lambda)\, r] \,. \tag{4.21}$$

By integrating the above differential equations and averaging over the ensemble one obtains the unconstrained probability distribution for δ_1 and δ_2 at a given Λ

$$\mathcal{W}_r(\delta_1, \delta_2; \Lambda) = \frac{1}{2\pi\sqrt{\Lambda^2 - \xi(r;\Lambda)^2}} \exp\left[-\frac{\Lambda(\delta_1^2 + \delta_2^2) - 2\,\xi(r;\Lambda)\,\delta_1\delta_2}{2\,[\Lambda^2 - \xi(r;\Lambda)^2]}\right] . \tag{4.22}$$

This function solves the two-dimensional Fokker-Planck equation associated to the Langevin equation (4.19), namely

$$\frac{\partial \mathcal{W}_r(\delta_1, \delta_2; \Lambda)}{\partial \Lambda} = \frac{1}{2}\left[\frac{\partial^2}{\partial \delta_1^2} + \frac{\partial^2}{\partial \delta_2^2} + 2\,\frac{\partial \xi(r;\Lambda)}{\partial \Lambda}\,\frac{\partial^2}{\partial \delta_1 \partial \delta_2}\right] \mathcal{W}_r(\delta_1, \delta_2; \Lambda) \,, \tag{4.23}$$

with initial condition $\mathcal{W}_r(\delta_1, \delta_2; \Lambda = 0) = \delta_{_D}(\delta_1)\,\delta_{_D}(\delta_2)$ and natural boundary conditions (i.e. vanishing at infinity). The problem of finding the distribution of the *first* upcrossings of the threshold t_f by the binary process $\{\delta_1, \delta_2\}$ reduces to that of imposing proper boundary conditions to the equation (4.23). As done for the one-dimensional Fokker-Planck equation, we adopt the absorbing barrier approach. However, in the two-dimensional case which we are considering, the distribution $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$ cannot be eventually obtained from $\mathcal{W}_r(\delta_1, \delta_2; \Lambda)$ simply by differentiation. This is because the whole binary system automatically disappears as soon as when one 'Brownian particle' is first absorbed. None the less, the joint distribution $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$ can in principle be calculated by a two-step procedure as follows.

Assuming the same initial condition, one solves the Fokker-Planck equation with absorbing barriers at $\delta_1 = t_f$ and $\delta_2 = t_f$, thus finding the survival probability density for the pairs which have never crossed the thresholds. Having found this quantity one can compute the probability current through each point,

$$\mathbf{J}(\delta_1, \delta_2; \Lambda) = -\frac{1}{2}\left(\frac{\partial \mathcal{W}_r}{\partial \delta_1} + \frac{\partial \xi(r;\Lambda)}{\partial \Lambda}\,\frac{\partial \mathcal{W}_r}{\partial \delta_2} \,,\quad \frac{\partial \xi(r;\Lambda)}{\partial \Lambda}\,\frac{\partial \mathcal{W}_r}{\partial \delta_1} + \frac{\partial \mathcal{W}_r}{\partial \delta_2}\right) . \tag{4.24}$$

On a boundary wall, e.g. $\delta_1 = t_f$, where $\mathcal{W}_r(t_f, \delta_2; \Lambda) = 0$ (implying $\partial \mathcal{W}_r / \partial \delta_2 = 0$), this reduces to

$$\mathbf{J}(t_f, \delta_2; \Lambda) = -\frac{1}{2}\left(\left.\frac{\partial \mathcal{W}_r}{\partial \delta_1}\right|_{t_f} \,,\quad \frac{\partial \xi(r;\Lambda)}{\partial \Lambda}\,\left.\frac{\partial \mathcal{W}_r}{\partial \delta_1}\right|_{t_f}\right) . \tag{4.25}$$

The flux through any point of the barrier $\delta_1 = t_f$ is then given by the scalar product $\mathbf{J} \cdot \mathbf{n}$, where $\mathbf{n} \equiv (1, 0)$ is the unit vector perpendicular to the absorbing wall,

$$\mathcal{F}_r(t_f, \delta_2; \Lambda) = -\frac{1}{2}\left.\frac{\partial \mathcal{W}_r}{\partial \delta_1}\right|_{t_f} . \tag{4.26}$$

The quantity $\mathcal{F}_r(t_f, \delta_2; \Lambda)\, d\delta_2$ represents the probability that the pair of processes (δ_1,δ_2) leave the permitted region passing through the 'gate' $[(t_f, \delta_2), (t_f, \delta_2 + d\delta_2)]$ at the time Λ. This flux contains all the information we need for the computation of $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$. In fact, once δ_1 has crossed the barrier at Λ_1, for $\Lambda > \Lambda_1$ we are interested in studying only the evolution of the surviving process δ_2 up to its first upcrossing through the boundary $\delta_2 = t_f$. Therefore, since we are considering Brownian trajectories, free of correlations along the Λ axis, for $\Lambda > \Lambda_1$ the evolution of δ_2 is simply

governed by its own Langevin equation, and its probability distribution can be calculated from the one-dimensional Fokker-Planck equation (4.8), with absorbing boundary $\delta_2 = t_{\rm f}$, *but* with initial condition (at $\Lambda = \Lambda_1$) $\delta_{2*} \equiv \delta_2(\Lambda_1|\delta_1 = t_{\rm f})$. Thus, by simply modifying eq. (4.10), we find that the distribution of the variances Λ_2 associated to first upcrossing events of the threshold by the process δ_2, given that δ_1 upcrossed the critical level at Λ_1, is

$$\mathcal{P}_1(\Lambda_2 - \Lambda_1, t_{\rm f} - \delta_{2*}) = \frac{(t_{\rm f} - \delta_{2*})}{\sqrt{2\pi}\,(\Lambda_2 - \Lambda_1)^{3/2}} \exp\left[-\frac{(t_{\rm f} - \delta_{2*})^2}{2(\Lambda_2 - \Lambda_1)}\right] . \tag{4.27}$$

The joint distribution $\mathcal{P}_2(\Lambda_1, \Lambda_2; r)$ is eventually obtained by a convolution

$$\begin{aligned}\mathcal{P}_2(\Lambda_1, \Lambda_2; r) = \int_{-\infty}^{t_{\rm f}} d\delta_2\, \mathcal{F}_r(t_{\rm f}, \delta_2; \Lambda_1)\, \mathcal{P}_1(\Lambda_2 - \Lambda_1, t_{\rm f} - \delta_2) + \\ + \int_{-\infty}^{t_{\rm f}} d\delta_1\, \mathcal{F}_r(\delta_1, t_{\rm f}; \Lambda_2)\, \mathcal{P}_1(\Lambda_1 - \Lambda_2, t_{\rm f} - \delta_1) ,\end{aligned} \tag{4.28}$$

where the first and second integrals on the r.h.s. represent, respectively, the contributions of those pairs for which $\Lambda_2 \geq \Lambda_1$ and $\Lambda_2 < \Lambda_1$.

However, this formal expression is useless unless we solve for the probability density $\mathcal{W}_r$. There are two cases in which the calculation of $\mathcal{W}_r$ is trivial: one for $r \to \infty$ and the other for $r \to 0$. The general case of finite non-zero lag r is far more complex, as we will discuss.

Perfectly uncorrelated processes

At infinite lag ($r \to \infty$) the two processes become independent and the solution of the Fokker-Planck equation is

$$\mathcal{W}_\infty(\delta_1, \delta_2; \Lambda) = \mathcal{W}(\delta_1, \Lambda; t_{\rm f})\, \mathcal{W}(\delta_2, \Lambda; t_{\rm f}) , \tag{4.29}$$

where $\mathcal{W}(\delta, \Lambda; t_{\rm f})$ denotes the probability distribution for a one-dimensional process with absorbing boundary at $t_{\rm f}$ given in equation (4.9). This solution consists of a linear superposition of four unconstrained independent density distributions deriving from different initial conditions. In practice, one has to consider the 'real' initial distribution $\delta_{_D}(0,0)$, an image source $\delta_{_D}(t_{\rm f}, t_{\rm f})$ and two image sinks $\delta_{_D}(t_{\rm f}, 0)$ and $\delta_{_D}(0, t_{\rm f})$, that is to say

$$\begin{aligned}\mathcal{W}_\infty(\delta_1, \delta_2; \Lambda) = \mathcal{G}_\infty(\delta_1, \delta_2; \Lambda) - \mathcal{G}_\infty(\delta_1 - 2t_{\rm f}, \delta_2; \Lambda) - \\ - \mathcal{G}_\infty(\delta_1, \delta_2 - 2t_{\rm f}; \Lambda) + \mathcal{G}_\infty(\delta_1 - 2t_{\rm f}, \delta_2 - 2t_{\rm f}; \Lambda) ,\end{aligned} \tag{4.30}$$

where $\mathcal{G}(\delta_1, \delta_2; \Lambda) \equiv [2\pi\Lambda]^{-1} \exp[-(\delta_1^2 + \delta_2^2)/2\Lambda]$ is the solution of the two-dimensional Fokker–Planck equation with natural boundary conditions: $\lim_{\delta_i \to \infty} \mathcal{G}(\delta_1, \delta_2; \Lambda) = 0, i = 1, 2$. Obviously, using eq. (4.28) we obtain

$$\mathcal{P}_2(\Lambda_1, \Lambda_2; r \to \infty) = \mathcal{P}_1(\Lambda_1, t_{\rm f}) \mathcal{P}_1(\Lambda_2, t_{\rm f}) , \tag{4.31}$$

that, inserted in equations (4.14) and (4.15) or in equations (4.17) and (4.18), gives, as expected for infinite lag, $\xi^{pts} = \xi^{hh} = 0$.

Perfectly correlated processes

When $r \to 0$ the two processes become more and more correlated so that, in the limit, we end up with a one-dimensional problem. In this case the solution of the Fokker-Planck equation is

$$\mathcal{W}_0(\delta_1, \delta_2; \Lambda) = \mathcal{W}(\delta_1 + \delta_2, 4\Lambda; t_{\rm f})\, \delta_{\scriptscriptstyle D}(\delta_1 - \delta_2) \;, \tag{4.32}$$

where $\mathcal{W}(\delta, \Lambda; t_{\rm f})$ denotes the probability distribution given in equation (4.9). Expanding this expression as a superposition of Green's functions, by using the method of image sources, we obtain

$$\mathcal{W}_0(\delta_1, \delta_2; \Lambda) = \mathcal{G}_0(\delta_1, \delta_2; \Lambda) - \mathcal{G}_0(\delta_1 - 2t_{\rm f}, \delta_2 - 2t_{\rm f}; \Lambda) \;, \tag{4.33}$$

where $\mathcal{G}_0(x, y; \Lambda) \equiv \mathcal{W}(x, \Lambda; t_{\rm f})\delta_{\scriptscriptstyle D}(x - y)$. In this case, for the joint distribution of first upcrossing variances, we find

$$\mathcal{P}_2(\Lambda_1, \Lambda_2; r \to 0) = \mathcal{P}_1(\Lambda_1, t_{\rm f})\, \delta_{\scriptscriptstyle D}(\Lambda_1 - \Lambda_2) \;. \tag{4.34}$$

Consequently, we obtain $\xi^{pts}_{II} \to [1/\int_I d\Lambda\, \mathcal{P}(\Lambda)] - 1$ and $\xi^{hh}_{II} \to \{\int_I d\Lambda\, P(\Lambda)/V(\Lambda)^2/[\int_I d\Lambda\, P(\Lambda)/V(\Lambda)]^2\} - 1$. Similarly, $\xi^{pts}_{I_1 I_2} \to 0$ and $\xi^{hh}_{I_1 I_2} = 0$.

Approximate general solution

In the limiting cases just discussed we were able to account for the boundary conditions by writing the solution in terms of a superposition of image distributions, each one solving the Fokker–Planck equation. Regrettably, the position and the sign of the image sources of probability come out dependent on the correlation between the processes (i.e., on r). This fact suggests that we cannot write a general analytical solution by simply applying the image method. However, we will build here a simple function that satisfies the absorbing boundary conditions being also an accurate approximation for the solution of the correlated diffusion equation. In the next section we will test the accuracy of this solution against the numerical one.

It is evident that, for small separation $r \ll R_{\rm f}$ (we remind the reader that we change the smoothing radius $R_{\rm f}$ at fixed separation r), the perfectly correlated solution will represent a very good approximation to the true one. This can be easily deduced by the following argument. The two Gaussian stochastic processes $\Sigma(\Lambda) = \delta_1(\Lambda) + \delta_2(\Lambda)$ and $\Delta(\Lambda) = \delta_1(\Lambda) - \delta_2(\Lambda)$ are statistically independent, i.e. $\langle \Sigma(\Lambda)\Delta(\Lambda) \rangle = 0$. The variances of their unconstrained probability distributions are, respectively, $\sigma^2_\Sigma = 2\,[\Lambda + \xi(r; \Lambda)]$ and $\sigma^2_\Delta = 2\,[\Lambda - \xi(r; \Lambda)]$. Therefore, for $r \ll R_{\rm f}$ [i.e. for $\Lambda \ll \sigma^2(r)$, where $\sigma^2(r)$ denotes the variance of the mass density fluctuations smoothed on the scale r] where $\xi(r; \Lambda) \simeq \Lambda$, we have $\sigma^2_\Delta \simeq 0$ and the probability distribution of the variable Δ is practically a Dirac delta function centred on the zero value. This corresponds to the perfectly correlated situation. However, this regime is not interesting for computing the halo correlation function since the condition $r \ll R_{\rm f}$ implies that the points in which we follow the trajectories that upcross the threshold are involved in the collapse of the same halo.

On the other hand, for $r \gg R_{\rm f}$ (i.e. for $\Lambda \gg \sigma^2(r)$), we can replace in eq. (4.23) $\xi(r; \Lambda)$ with the *unsmoothed* linear mass-correlation $\xi_m(r)$. Therefore, eq. (4.23) simply becomes the uncorrelated two-dimensional diffusion equation that can be easily solved using the image method. Our Ansatz for the full $\mathcal{W}_r(\delta_1, \delta_2; \Lambda)$ is then obtained by keeping the analytic form of the solution just obtained for $r \gg R_{\rm f}$ but inserting in it the correlation function $\xi(r; \Lambda)$ to replace its large lag limit $\xi_m(r)$. Therefore, we have

$$\begin{aligned} \mathcal{W}_r(\delta_1, \delta_2; \Lambda) = \mathcal{G}^+_r(\delta_1, \delta_2; \Lambda) - \mathcal{G}^-_r(\delta_1 - 2t_{\rm f}, \delta_2; \Lambda) - \\ - \mathcal{G}^-_r(\delta_1, \delta_2 - 2t_{\rm f}; \Lambda) + \mathcal{G}^+_r(\delta_1 - 2t_{\rm f}, \delta_2 - 2t_{\rm f}; \Lambda) \;, \end{aligned} \tag{4.35}$$

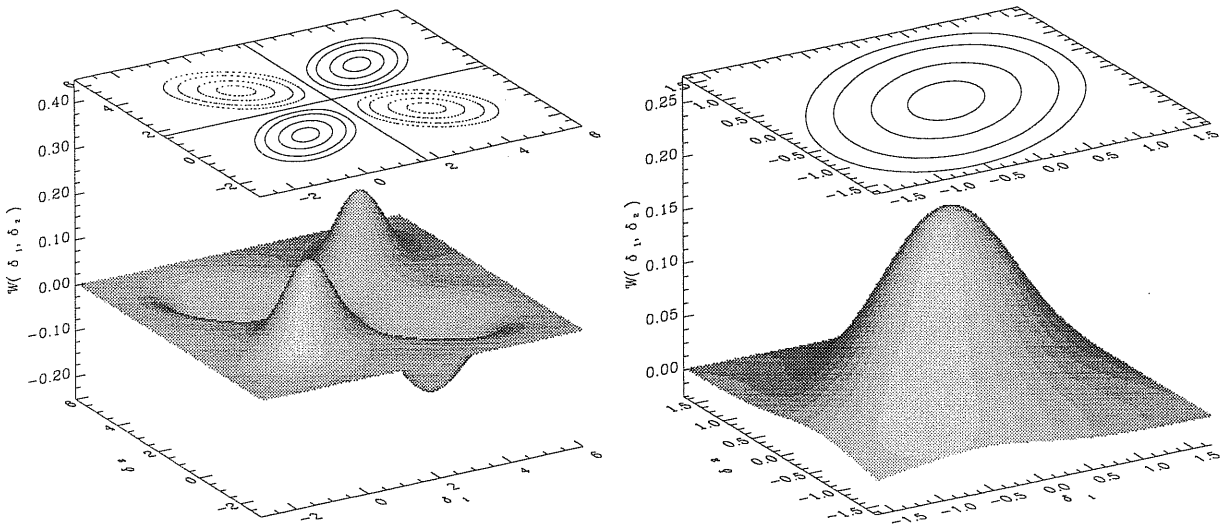


Figure 4.2: The approximate solution of the Fokker-Planck equation given in eq. (4.35) is plotted in the left panel (here $\Lambda = 1$ and $\xi = 0.4$). The probability density in the physical region is blown up in the right panel. In the contour plots the continuous lines correspond to the levels 0, 0.025, 0.05, 0.1, 0.15 (starting from outside), while the dotted lines refer to the same values multiplied by -1. The figure clearly shows that the proposed solution satisfies the absorbing boundary conditions for $\delta_1 = t_f$ and $\delta_2 = t_f$ (here $t_f = 1.686$). Moreover, the orientation of the iso-probability curves in the allowed region (these are practically ellipses with their major axes lying on the line $\delta_1 = \delta_2$) indicates that the approximate $\mathcal{W}(\delta_1, \delta_2; \Lambda)$ has the right correlation properties.

where

$$\mathcal{G}_r^{\pm} = \frac{1}{2\pi\sqrt{\Lambda^2 - \xi(r;\Lambda)^2}} \exp\left[-\frac{\Lambda(\delta_1^2 + \delta_2^2) - 2\left[\pm\xi(r;\Lambda)\right]\delta_1\delta_2}{2\left[\Lambda^2 - \xi(r)^2\right]}\right]. \tag{4.36}$$

By using the symbol $\mathcal{G}_r^{\pm}$ we want to emphasize the correlation properties of the adopted Green's function: $\mathcal{G}_r^{+}$ is a correct solution of the diffusion equation, while $\mathcal{G}_r^{-}$ does not solve it. However, to satisfy the boundary conditions, we need to insert it twice in the solution. The probability distribution given in eq. (4.35) will be a valid approximation to the proper one provided that the term

$$\frac{\partial\xi(r;\Lambda)}{\partial\Lambda}\frac{\partial^2}{\partial\delta_1\partial\delta_2}\left[\mathcal{G}_r^{-}(\delta_1 - 2t_f, \delta_2; \Lambda) + \mathcal{G}_r^{-}(\delta_1, \delta_2 - 2t_f; \Lambda)\right] \tag{4.37}$$

can be neglected compared to the Λ-derivative of the expression in eq. (4.35). A three-dimensional representation of this approximate solution and of its contour levels is given in Fig. 4.2. Inserting eq. (4.35) into equations (4.26) and (4.28) we obtain

$$\begin{aligned}\mathcal{P}_2(\Lambda_1, \Lambda_2; r) =& \frac{t_f^2\Lambda_1\Lambda_2 + \left[\Lambda_1\Lambda_2 - t_f^2(\Lambda_1 + \Lambda_2)\right]\xi(r;\Lambda_m) + t_f^2\,\xi(r;\Lambda_m)^2 - \xi(r;\Lambda_m)^3}{2\pi\left[\Lambda_1\Lambda_2 - \xi(r;\Lambda_m)^2\right]^{5/2}} \times \\ & \times \exp\left[-\frac{t_f^2}{2}\frac{\Lambda_1 + \Lambda_2 - 2\,\xi(r;\Lambda_m)}{\Lambda_1\Lambda_2 - \xi(r;\Lambda_m)^2}\right],\end{aligned} \tag{4.38}$$

where $\Lambda_m = \min(\Lambda_1, \Lambda_2)$. By using this expression to compute the halo-halo correlation function between objects selected in infinitesimal mass ranges, we obtain

$$\xi^{obj}(r) \equiv \xi^{pts}(r) = \frac{\mathcal{P}_2(\Lambda_1, \Lambda_2; r)}{\mathcal{P}_1(\Lambda_1)\mathcal{P}_1(\Lambda_2)} - 1\,, \tag{4.39}$$

whose explicit expression is

$$1+\xi^{hh}(r) = \frac{t_{\rm f}^2\Lambda_1\Lambda_2+\left[\Lambda_1\Lambda_2-t_{\rm f}^2(\Lambda_1+\Lambda_2)\right]\xi(r;\Lambda_{\rm m})+t_{\rm f}^2\,\xi(r;\Lambda_{\rm m})^2-\xi(r;\Lambda_{\rm m})^3}{\Lambda_1^{-3/2}\Lambda_2^{-3/2}\left[\Lambda_1\Lambda_2-\xi(r;\Lambda_{\rm m})^2\right]^{5/2}}\times$$
$$\times\ \exp\left[-\frac{t_{\rm f}^2}{2}\,\frac{(\Lambda_1+\Lambda_2)\,\xi(r;\Lambda_{\rm m})^2-2\,\Lambda_1\Lambda_2\,\xi(r;\Lambda_{\rm m})}{\Lambda_1\Lambda_2\left[\Lambda_1\Lambda_2-\xi(r;\Lambda_{\rm m})^2\right]}\right]. \tag{4.40}$$

The expansion of the halo two-point function in powers of the filtered mass autocorrelation function will be discussed in chapter 6. It is important to stress that eq. (4.39), with the ansatz for $\mathcal{P}_2$ provided by eq. (4.38) and $\mathcal{P}_1$ given by eq. (4.10), represents an approximation to the exact form of the halo-halo correlation function which can be obtained by numerically integrating the correlated Langevin equations, as discussed in the next section. Other approximations have been suggested by CLMP following two different approaches: i) defining a local halo counting operator acting on the underlying Gaussian density field, ii) using the peak-background split to define the halo overdensity free of the cloud-in-cloud problem down to the background scale. Their equation (14), once it has been recast in the present notation, gives

$$1+\xi^{hh}_{\rm CLMP}(r) = \frac{1}{\sqrt{1-\omega^2}}\left\{1+\frac{\sigma_2^2}{(1-\omega^2)}\left(\frac{1}{\sigma_1}-\frac{\omega}{\sigma_2}\right)\frac{\partial\omega}{\partial\sigma_2}+\frac{\sigma_1^2}{(1-\omega^2)}\left(\frac{1}{\sigma_2}-\frac{\omega}{\sigma_1}\right)\frac{\partial\omega}{\partial\sigma_1}+\right.$$
$$+\frac{\sigma_1^2\sigma_2^2}{t_{\rm f}^2}\,\frac{\partial^2\omega}{\partial\sigma_1\partial\sigma_2}+\frac{\sigma_1^2\sigma_2^2}{t_{\rm f}^2(1-\omega^2)^2}\left[\omega(1-\omega^2)+(1+\omega^2)\,\frac{t_{\rm f}^2}{\sigma_1\sigma_2}-\omega\,t_{\rm f}^2\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)\right]\times$$
$$\left.\times\,\frac{\partial\omega}{\partial\sigma_1}\,\frac{\partial\omega}{\partial\sigma_2}\right\}\times\exp\left[-\frac{t_{\rm f}^2}{2}\,\frac{\omega^2\left(\frac{1}{\sigma_1^2}+\frac{1}{\sigma_2^2}\right)-2\,\frac{\omega}{\sigma_1\sigma_2}}{(1-\omega^2)}\right]. \tag{4.41}$$

Here $\sigma_i=\Lambda_i^{1/2}$ and $\omega=\xi(\Lambda_1,\Lambda_2;r)/\sigma_1\sigma_2$ with $\xi(\Lambda_1,\Lambda_2;r)$ the correlation function between the linear mass density fluctuation field smoothed with two different resolutions Λ_1 and Λ_2. For sharp k-space filtering $\xi(r;\Lambda_1,\Lambda_2)=\xi(r;\Lambda_{\rm m})$. It is worth mentioning that for separations larger than the smoothing lengths, when the derivatives of $\xi(r;\Lambda_1,\Lambda_2)$ with respect to Λ_1 and Λ_2 are negligible, eq. (7.14) reduces to eq. (4.40). Moreover, as we will show in chapter 6, both approximations asymptotically reach the linear bias regime studied by Mo & White (1996).

4.3.4 Monte Carlo simulations

In order to check the validity of the approximate solution introduced in the previous section, we solved numerically for the joint distribution of first upcrossing variances by integrating our spatially correlated Langevin equations. We stress that this method gives the exact halo-halo correlation function in the excursion set approach, consistently completing the PS analysis of the mass function.

The stochastic differential equations (4.19) are equivalent to the integral equations

$$\begin{cases}\delta_1(\Lambda+\gamma)-\delta_1(\Lambda)=\displaystyle\int_\Lambda^{\Lambda+\gamma}d\Lambda'\,\zeta_1(\Lambda')\,, & \delta_1(0)=0\,,\\[2ex] \delta_2(\Lambda+\gamma)-\delta_2(\Lambda)=\displaystyle\int_\Lambda^{\Lambda+\gamma}d\Lambda'\,\zeta_2(\Lambda')\,, & \delta_2(0)=0\,,\end{cases} \tag{4.42}$$

where the statistical properties of the Gaussian processes ζ_1 and ζ_2 are given in eq. (4.19). The general procedure used to solve numerically a stochastic differential equation replaces the equivalent

integral equation by its expansion in power series of $\sqrt{\gamma}$, truncates the series after a selected number of terms, and gives a rule for computing each term that is considered. To control the effect of the temporal discretization an extrapolation of the results for $\gamma \to 0$ is often required (Greiner, Strittmatter & Honerkamp 1988). Fortunately, in the case of a set of Wiener processes this procedure can be greatly simplified. In fact, by integrating over a finite timestep γ, equations (4.42) simply give

$$\begin{cases} \delta_1(\Lambda+\gamma) - \delta_1(\Lambda) = a_{11}(\gamma,\Lambda)\, G_1 \,, & \delta_1(0) = 0 \,, \\ \delta_2(\Lambda+\gamma) - \delta_2(\Lambda) = a_{21}(\gamma,\Lambda)\, G_1 + a_{22}(\gamma,\Lambda)\, G_2 \,, & \delta_2(0) = 0 \,, \\ a_{11}(\gamma,\Lambda)^2 = a_{21}(\gamma,\Lambda)^2 + a_{22}(\gamma,\Lambda)^2 = \gamma \,, \\ a_{11}(\gamma,\Lambda)\, a_{21}(\gamma,\Lambda) = \xi(r;\Lambda+\gamma) - \xi(r;\Lambda) \,, \end{cases} \tag{4.43}$$

where the G_i are independent Gaussian variables with zero mean and unit variance. This set of equations gives the fundamental recipe to produce trajectories that are obtained iterating eq. (4.43) at each time-step by modelling the G_i terms with Gaussian pseudo-random numbers. To generate these normally distributed deviates in first passage problems, where rare fluctuations are crucial, it is of fundamental importance to adopt a method that is accurate even for less probable events. For this reason we adopted the Box–Muller algorithm (e.g. Press et al. 1992), which is rather slow but produces an unbiased Gaussian distribution (in practice, the limited precision of numerical computation only affects the extreme tail behaviour). A moderate speeding up (roughly 20 per cent) is obtained by modifying the algorithm to use a rejection technique (Knuth 1981; Press et al. 1992).

To estimate the first-passage time distribution one first solves the discretized stochastic equation starting at the initial point and terminates the simulation of a trajectory as soon as the boundary is reached. In order to avoid the resulting distribution being influenced by the temporal discretization, one has to account for possible intra-step crossings. In fact, the conditions $\delta(\Lambda) < t_\mathrm{f}$ and $\delta(\Lambda+\gamma) < t_\mathrm{f}$ do not guarantee that the process δ has ever crossed the threshold during the time interval γ.[4] For this reason, the simple algorithm of choosing as first-crossing time that corresponding to the first step at which $\delta > t_\mathrm{f}$ is very inaccurate, unless one uses very small time-steps. This problem can be solved by performing a small Monte Carlo test at each time step as shown by Strittmatter (1987). In this way we obtain high precision even using larger time-steps, therefore reducing the CPU time.

For a given power-spectrum, once the value for the critical threshold t_f and the lag r are selected, the algorithm just introduced gives a pair of first upcrossing variances for each realization of the processes δ_1 and δ_2. Therefore, the joint probability $\mathcal{P}_2(\Lambda_1,\Lambda_2;r)$ and the halo-halo correlation function can be obtained by considering a large number of realizations.

4.3.5 Results

We present here the results obtained by considering two different scale-free power-spectra $P(k) \propto k^n$ with $n = -1$ and $n = -2$ in an Einstein–de Sitter universe. In these cases the evolution of clustering is self-similar and the results obtained at a particular epoch are representative of the whole history. These two values of the spectral index can be regarded as typical of any physically reliable power-spectrum on scales relevant for galaxy formation in a hierarchical scenario. Adopting a standard

[4] This has a striking analogy with the solution of the cloud-in-cloud problem given by BCEK.

procedure, we normalize the power-spectrum so that the linear mass variance as measured in $8\,h^{-1}$Mpc spheres is equal to 1 and we impose $\delta_c = 1.686$.

Concerning the selection of the mass ranges that identify different classes of haloes, we optimized their broadness in order to balance the CPU time requirement (too narrow ranges turn out to be poorly statistically populated) with an accurate description of clustering. For these reasons we selected three different classes of objects for each power-spectrum (the first contains objects with $M \gg M_*$, the second has $M \simeq M_*$ and the third $M \ll M_*$) by requiring that they are roughly equi-populated (in terms of first upcrossing events, not number of haloes). In this way we can explore all the expected regimes of Lagrangian clustering (a biased halo distribution for $M \gg M_*$, an almost unclustered distribution for $M \simeq M_*$ and an antibiased distribution for $M \ll M_*$) with approximately equal numerical accuracy.

Table 1 gives the parameters that define the three classes of haloes we selected and the probability of occurrence of first upcrossing events in each of them [this is obtained by integrating eq. (4.10) over the corresponding interval of Λ]. A technical problem one has to deal with is the occurrence of spurious oscillations in the correlation function induced by the sharp k-space filter. In fact, the sharpness of the smoothing kernel unavoidably gives rise to oscillations in the mass correlations computed in the Fourier-conjugate space. To exemplify, by convolving the linear density field with $\widetilde{W}_{SKS}(k, R_\mathrm{f})$ one obtains $\xi_m(r;\Lambda) \propto [1 - \cos(k_\mathrm{f}(\Lambda)r)]/r^2$ for $n = -1$ and $\xi_m(r;\Lambda) \propto \mathrm{Si}(\mathrm{k_f}(\Lambda)\mathrm{r})/r$ for $n = -2$, where $\mathrm{Si}(z) \equiv \int_0^z dy\, j_0(y)$. Unfortunately, this oscillating behaviour affects also ξ^{hh}. This can be easily understood in the Langevin formalism, where the oscillations are generated by the unsmoothed Bessel function appearing in eq. (4.19), which is unavoidable in the Wiener process approach. One can try to overcome this problem by replacing in eq. (4.43) the sharp k-space filtered correlation $\xi(r;\Lambda)$ with the one obtained with top-hat smoothing. This is analogous to the technique generally applied to compare the mass function predicted by the excursion set approach to the outputs of N-body simulations (e.g. Lacey & Cole 1993, 1994). This method strongly reduces the oscillations but, for $n > -2$, is not able to erase them completely at separations comparable to the halo size. In fact, for $n = -1$ and top-hat filtering, the term a_{21} defined in eq. (4.43) is positive for $\Lambda \lesssim \sigma^2(r)$, becomes negative for $\Lambda \gtrsim \sigma^2(r)$ and rapidly approaches zero for $\Lambda \gg \sigma^2(r)$, after assuming a minimum negative value. This term plays a fundamental role in the computation of the halo correlation function, causing the appearance of oscillations when r is comparable to the Lagrangian radius of the given halo. This problem could be totally avoided by coherently adopting from the beginning a more realistic (i.e. non-sharp k-space) window function. The price one has to pay, however, is that of dealing with a space-correlated set of coloured stochastic processes, i.e.

$$
\begin{cases}
\dfrac{d\delta_1(R_\mathrm{f})}{dR_\mathrm{f}} = \eta_1(R_\mathrm{f})\,, & \lim\limits_{R_\mathrm{f}\to\infty} \delta_1(R_\mathrm{f}) = 0\,, \\[2ex]
\dfrac{d\delta_2(R_\mathrm{f})}{dR_\mathrm{f}} = \eta_2(R_\mathrm{f})\,, & \lim\limits_{R_\mathrm{f}\to\infty} \delta_2(R_\mathrm{f}) = 0\,, \\[2ex]
\langle \eta_1(R_\mathrm{f}) \rangle = \langle \eta_2(R_\mathrm{f}) \rangle = 0\,, & \eta_1 \text{ and } \eta_2 \text{ Gaussian processes}\,, \\[2ex]
\langle \eta_1(R_\mathrm{f})\, \eta_1(R'_\mathrm{f}) \rangle = \langle \eta_2(R_\mathrm{f})\, \eta_2(R'_\mathrm{f}) \rangle = \dfrac{1}{2\pi^2} \displaystyle\int_0^\infty dk\, k^2\, P(k)\, \frac{\partial \widetilde{W}(kR_\mathrm{f})}{\partial R_\mathrm{f}}\, \frac{\partial \widetilde{W}(kR'_\mathrm{f})}{\partial R'_\mathrm{f}}\,, \\[2ex]
\langle \eta_1(R_\mathrm{f})\, \eta_2(R'_\mathrm{f}) \rangle = \dfrac{1}{2\pi^2} \displaystyle\int_0^\infty dk\, k^2\, P(k)\, \frac{\partial \widetilde{W}(kR_\mathrm{f})}{\partial R_\mathrm{f}}\, \frac{\partial \widetilde{W}(kR'_\mathrm{f})}{\partial R'_\mathrm{f}}\, j_0(kr)\,,
\end{cases}
\tag{4.44}
$$

where we used the smoothing radius R_f as time variable for the trajectories. In this way we

emphasize the role played by the filter in determining the statistical properties along and between the trajectories. For sharp k-space filtering, the quantity $\partial\widetilde{W}(kR_{\rm f})/\partial R_{\rm f}$ reduces to a Dirac delta function, which leads to the set (4.19) of the main text. In the following we will always use eq. (4.43) and top-hat smoothing to obtain ξ^{hh}.

In order to compute the correlation function, for each physical separation r we followed the evolution of many realizations of the stochastic processes δ_1 and δ_2, until 10^6 pairs of trajectories crossed the threshold at resolutions Λ_1 and Λ_2, both contained in one of the three selected intervals. The lag r is taken in the range $1 \leq r/R_* < 12$ for $n = -1$ and $1 \leq r/R_* < 40$ for $n = -2$, where R_* is the Lagrangian radius associated to the characteristic halo mass M_*.

Each simulation has been repeated several times (20 for $n = -1$, and 8 for $n = -2$) using different sequences of pseudo-random numbers to build up the trajectories. The values for $\xi^{hh}(r)$ shown in Fig. 4.3 are obtained by averaging over the different simulations, while the error bars represent the standard deviation of the mean.

In Fig. 4.3 we also plot the analytic expressions predicted by two different models: our approximated solution of the Fokker-Planck equation in eq. (4.38) and the 'counting field' model described in eq. (14) of CLMP [see also equations (4.41) and (7.14)]. Let us recall that these two models give rise to the same Lagrangian bias factors, i.e. to the same clustering regime, provided the halo separation is a few times larger than their Lagrangian size. Asymptotically, both the two analytic forms and the numerical integrations tend to the lowest non-vanishing term of the series expansion in eq. (6.15), which, except for the mass range centered on M_*, coincides with the Mo & White (1996) prediction. In general, the agreement of both models with the results of the Monte Carlo simulations is remarkably good, except for lags of order the halo Lagrangian size. Here the discrepancy between analytical and numerical results becomes larger and larger as the ratio M/M_* decreases. In particular, neither of the two models is able to follow the detailed features of the numerical solution at small separations, where, at least for $n = -1$, the spurious oscillations induced by the adoption of top-hat smoothing in eq. (4.43) – which has been instead derived after sharp k-space filtering – play a relevant role.

We can try to give a more detailed description of the numerical outcomes by introducing in eq. (14) of CLMP an extra modulation induced by a decaying sinusoidal term. We can then introduce the following 'best-fitting models'

$$\frac{\xi^{hh}(r) - \xi_{\rm CLMP}(r)}{1 + \xi_{\rm CLMP}(r)} = C_1 \cos\left(C_2 \frac{r}{R_*} + C_3\right) \exp\left[-C_4 \left(\frac{r}{R_*}\right)^2\right] , \tag{4.45}$$

$$\frac{\xi^{hh}(r) - \xi_{\rm CLMP}(r)}{1 + \xi_{\rm CLMP}(r)} = C_1 \cos\left(C_2 \frac{r}{R_*} + C_3\right) \exp\left[-C_4 \frac{r}{R_*}\right] , \tag{4.46}$$

Table 4.1: Parameters that identify the classes of haloes.

Class	$\Lambda_{\rm min}$	$\Lambda_{\rm max}$	$M_{\rm min}/M_*$ $(n=-1)$	$M_{\rm max}/M_*$ $(n=-1)$	$M_{\rm min}/M_*$ $(n=-2)$	$M_{\rm max}/M_*$ $(n=-2)$	$\int_{\Lambda_{\rm min}}^{\Lambda_{\rm max}} \mathcal{P}_1(\Lambda)\, d\Lambda$
I_1	0.45	1.79	2	16	4	256	0.20
I_2	1.79	4.51	1/2	2	1/4	4	0.22
I_3	4.51	11.37	1/8	1/2	1/64	1/4	0.19

Table 4.2: Parameters for the 'best-fitting models'. In the last row the value of C_3 is set to π since $C_2 \simeq 0$ causes a degeneracy between C_1 and $\cos(C_3)$. This is removed by assigning an arbitrary value to C_3.

n	Class	C_1	C_2	C_3	C_4
-1	I_1	0.24 ± 0.02	2.47 ± 0.16	-2.7 ± 0.3	0.570 ± 0.016
-1	I_2	1.69 ± 0.10	1.34 ± 0.07	0.00 ± 0.08	0.880 ± 0.008
-1	I_3	21 ± 6	0.65 ± 0.13	1.1 ± 0.2	2.49 ± 0.13
-2	I_1	7.0 ± 1.5	0.35 ± 0.07	6.8 ± 0.2	1.050 ± 0.012
-2	I_2	57 ± 11	0.08 ± 0.16	1.49 ± 0.14	1.90 ± 0.02
-2	I_3	0.86 ± 0.12	0.0 ± 0.2	π	1.475 ± 0.011

respectively for $n = -1$ and $n = -2$. The coefficients C_α ($\alpha = 1, \ldots, 4$), which are found using the Levenberg–Marquardt non-linear least-squares method (e.g. Press et al. 1992) in each mass range, are given in Table 4.2. The main limitation of this 'best-fit' approach is clearly that the coefficients C_α depend both on the shape of the power-spectrum and on the halo masses. Although we did not attempt any such approach, it ought to be possible to parametrize these two dependences.

4.4 Conclusions

In this chapter we have proposed a model for the clustering of dark matter haloes in Lagrangian space. Our model is based on a natural extension of the excursion set approach to the PS theory (e.g. BCEK), namely we accounted for the spatial correlations of the linear mass density fluctuation field.

In particular, the two-point halo correlation function has been obtained by integrating the system of Langevin equations governing the evolution of pairs of correlated density processes, or, equivalently for a Gaussian mass distribution, the bivariate Fokker–Planck equation for the density probability distribution of the same processes, once appropriate boundary conditions have been imposed. We believe that the numerical integration of the correlated Langevin equations allowed for the most reliable determination of the Lagrangian halo correlation, complementing the Press–Schechter inspired analyses of the halo mass function.

Although we gave explicit results only for the halo two-point function, a generalization to higher order statistics would be straightforward. The halo correlation function obtained with the present approach is fully consistent with the recent results obtained by CLMP, and its form at large separation reproduces the linear bias relation by Mo & White (1996; see chapter 6), which has been shown to be in good agreement with the clustering of synthetic haloes in N-body simulations.

As stressed by CLMP, the issue of transforming the halo distribution from the Lagrangian to the Eulerian world is a non-trivial one, especially on smaller scales and for smaller mass systems, which are most affected by the intrinsic non-linearity of the evolved mass density field and by the occurrence of multi-streaming. Nevertheless, one might speculate that a stochastic approach analogous to the one considered so far would be viable also in Eulerian space. The main modification induced by the Lagrangian-to-Eulerian map would in fact be a *local* modulation of the halo formation threshold, namely $t_{\rm f}(z_{\rm f}) \longrightarrow t'_{\rm f}(\mathbf{x}, z_{\rm f}) \equiv t_{\rm f}(z_{\rm f}) - \delta_{M_{\rm o}}(\mathbf{x}, z_{\rm f})$, with $\delta_{M_{\rm o}}(\mathbf{x}, z_{\rm f})$ the Eulerian mass density contrast smoothed on some background mass scale $M_{\rm o}$ much larger than the halo one.

The results presented in this chapter (see also Porciani et al. 1998) represent a relevant step

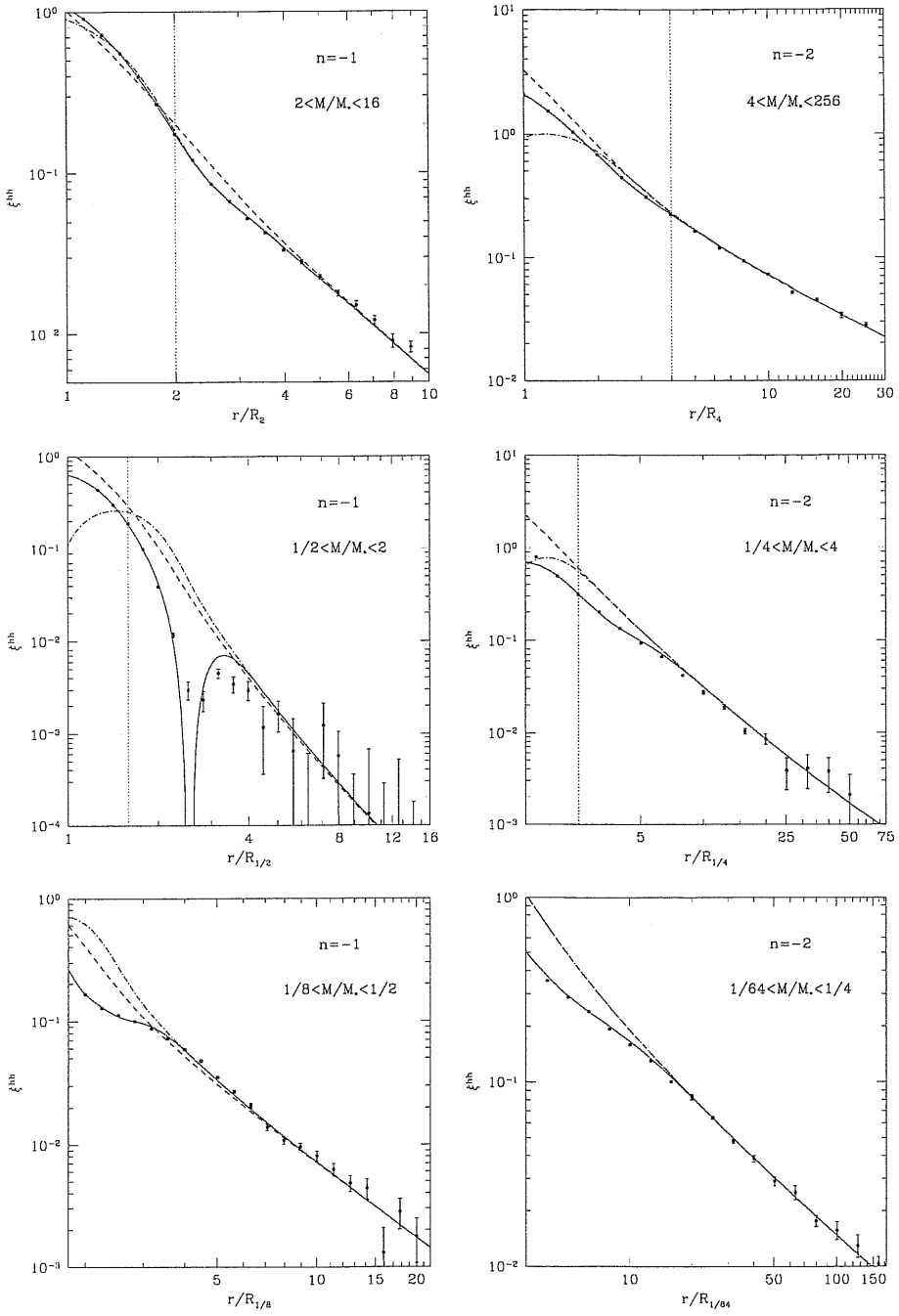


Figure 4.3: The Lagrangian halo correlation function ξ^{hh} in an Einstein-de Sitter universe with two different scale-free power-spectra, $n = -1$ (*left* column) and $n = -2$ (*right* column), is shown for three halo mass ranges. The object separation r is scaled to the Lagrangian radius of the least massive halo in each range. The vertical dotted lines, where shown, are placed at the Lagrangian radius of the most massive halo in each range. The points represent the mean value of different realizations obtained by numerically solving our correlated Langevin equations, while the error bars represent the scatter of the mean. The *continuous* lines refer to the 'best-fit models' of eqs. (4.45) and (4.46). The *dashed* lines are obtained from our approximated solution of the Fokker-Planck equation in eq. (4.38), while the *dot-dashed* lines show the predictions of the 'counting field' model of eq. (14) in CLMP. Top-hat filtering is used in all cases.

towards the construction of a local identification algorithm for halo formation sites in Lagrangian space, which would allow to depict halo maps starting from low-resolution simulations of the evolved dark matter distribution. As a straightforward application of this general idea one can apply the Zel'dovich approximation (Zel'dovich 1970) to follow the dark matter dynamics on mildly non-linear scales and use it to reconstruct the local halo number density at various epochs (Catelan, Matarrese & Porciani 1998).

One of the most interesting direct applications of our technique is the construction of spatially correlated halo merging trees. State-of-the-art algorithms to follow mass accretion histories are entirely based either on the mean one-point distribution of the first upcrossing events $\mathcal{P}_1$ (Lacey & Cole 1993; Kauffmann & White 1993; Somerville & Kolatt 1997) or on Monte Carlo realizations of the one-point Wiener process $\delta(\Lambda)$ (e.g. Tozzi et al. 1996). In both cases one has to extrapolate the succession of merging events involving a large number of haloes coming from different Lagrangian regions from the knowledge of the average properties of a single mass accretion history. Obviously some *ad hoc* assumptions need to be made in the tree reconstruction process, to supply for the lack of statistical and spatial information; examples are the assumption of binary merging in Lacey & Cole (1993) or the recent efforts aimed at distinguishing between mass accretion and merging events (Salvador-Solé et al. 1997; Somerville & Kolatt 1997).

In a more realistic approach, however, each branch of a merger tree should be associated to a different Lagrangian position. Thus spatial correlations of the initial density field would manifest themselves as statistical correlations both between different branches of the same tree and between different trees. The importance of this issue was stressed already in Lacey & Cole (1993). Somerville & Kolatt (1997) argued that accounting for the complex correlated structure of the density field would permit the construction of merger trees starting at high redshift and propagating forward in time.

A first attempt to include spatial correlations in modelling the hierarchical growth of dark matter haloes was made by Yano, Nagashima & Gouda (1996) and by Rodrigues & Thomas (1996). The main effort of these authors was, however, devoted at solving the cloud-in-cloud problem when the Lagrangian regions that ultimately collapse into haloes are properly treated as extended ones. Their predictions for the mass function imply more high-mass objects and less low-mass haloes than the PS expression.

However, the issue we want to focus on is a very different and more ambitious one. By adopting the excursion set theory, we aim at building realizations of merger histories by following the fate of the correlated trajectories associated to each branch of the same tree. Since we are not modifying the one-point distribution of first upcrossing events compared with the standard solution of the cloud-in-cloud problem (BCEK) our merging histories are *a priori* consistent with the PS mass function.

Concerning other quantities that characterize the ensemble of merging histories (such as halo survival and formation times) we expect that accounting for spatial correlations will lead to some modifications. In general, compared with the uncorrelated case, we should obtain corrections whose importance depends on the mass and the epoch under consideration. Intuitively, we would expect the larger effects on those objects whose mass accretion histories are dominated by merging of halos with similar masses. In fact, as shown in Fig. 4.1, the correlation between nearby trajectories is important only when the smoothing length is larger than or comparable to the physical separation between the Lagrangian points to which the trajectories are associated. In a merger tree the physical separation between two branches should be roughly given by the sum of the Lagrangian radii of the haloes measured just before they merge. Therefore, the effect of spatial correlations of the density field will be completely negligible if the two Lagrangian halo sizes are much different. These and related issues deserve further investigation.

Chapter 5

Gravitational lensing of distant supernovae in CDM models

5.1 Introduction

The potential of gravitational lensing as a tool for the determination of cosmological parameters or as a probe of cosmogonic models has long been recognized (see Blandford & Narayan 1992 for a review). Frequently discussed techniques include detailed models of specific lens systems such as multiple quasars (e.g., Falco, Gorenstein, & Shapiro 1991; Grogin & Narayan 1996; Kundic et al. 1997), arcs (Lynds & Petrosian 1989; Tyson & Fisher 1995; Tyson, Kochanski, & Dell' Antonio 1998), and radio rings (Kochanek et al. 1989), statistical studies of lensed QSOs (Narayan & White 1988; Fukugita et al. 1992; Kochanek 1996; Wambsganss et al. 1995), and measurements of the correlated ellipticities induced by large-scale structure in the images of background galaxies (Babul & Lee 1991; Blandford et al. 1991; Kaiser 1992; Villumsen 1996).

It has been recently realized by many authors (Kolatt & Bartelmann 1998; Metcalf 1998; Holz 1998; Marri & Ferrara 1998; Wang 1998; see also Linder, Schneider, & Wagoner 1988) that gravitational lensing of distant supernovae (SNe) may provide a new tool for doing cosmology. When corrected for light and color curve shapes, Type Ia SNe can be calibrated to be excellent standard candles, with a brightness at maximum B-light of $M_V = -19.4 \pm 0.15$ mag (Riess, Press, & Kirshner 1996; Hamuy et al. 1996). Systematic searches for SNe at high redshifts (Perlmutter et al. 1998; Garnavich et al. 1998) have already yielded a plethora of events. At this rate of accumulation, several hundreds of high-z SNe will be known within the next few years. The uniformity of Type Ia's makes them unique sources for measuring the redshift-luminosity distance relation.

In the observable (clumpy) universe, gravitational lensing will magnify and demagnify high redshift supernovae relative to the predictions of ideal (homogeneous) reference cosmological models. The magnitude and frequency of the effect depend on the power spectrum of density fluctuations, the abundance and the clustering properties of virialized clumps, the mass distribution within individual lenses, and the underlying world model. The degree to which weak lensing increases the level of noise in the Hubble diagram of Type Ia SNe, thereby decreasing the precision of cosmological parameter determinations, has been discussed by Kantowski, Vaughan, & Branch (1995), Frieman (1996), Wambsganss et al. (1997), Holz (1998), and Metcalf (1998). In this chapter we examine instead the high magnification tail of the lensing distribution. The likelihood of a SN at $z_s \lesssim 2$ being magnified by $\Delta m > 0.3$ (> 0.1) mag by a galaxy, group, or cluster is small but non-negligible, $\mathcal{P} \sim$ few $\times 10^{-2}$ (10%). One can then use recent estimates of the global history of star formation to compute the expected frequencies of lensed SNe as a function of cosmic time. The detection

rate of Type Ia's with magnification ≥ 0.3 mag is estimated to be of order a few SNe yr^{-1} deg^{-2} at $I_{\rm AB} \leq 25$ mag. Strong lensing ($\Delta m \geq 0.75$ mag) with multiple images is seven times less frequent (unless these events are specifically targeted, Kolatt & Bartelmann 1998). At bright magnitudes, the effect of magnification bias on the apparent frequency of lensed SNe is huge.

On the face of their rarity, which makes a measurement of the magnification probability density (hence of the distribution of mass inhomogeneities in the universe) a very difficult task, the modeling of individual lensed SN Ia might have a few advantages over the study of multiple quasars or galaxy correlated ellipticities. Lensed SNe will appear unusually bright for their redshift and will be easily identifiable by the proximity of a foreground galaxy, group, or cluster. Their magnification, which is related to the integrated mass density along the line-of-sight, is directly measurable to a good precision. A supernova that goes off within the Einstein ring of a foreground mass concentration may generate multiple images at different positions on the sky. The time delay between the images, inversely proportional to the Hubble constant, is longer than a month in about $60 - 70\%$ of the cases, and could be measured by comparing the light curves of the two events during a reasonable monitoring period. If all the relevant deflection angles to constrain the lens system were known, it could be possible to estimate H_0. Small-separation events are associated with galaxy lenses, and will remain undetected because of the presence of such a luminous foreground object.

The plan of the chapter is as follows. In §5.2 we compute the probability that a supernova undergoes a lensing event caused by an intervening dark matter halo. We investigate three popular hierarchical cosmogonies, two flat models with $\Omega_0 = 0.35$ and 1, and an open model with $\Omega_0 = 0.3$. The population of gravitational lenses is modeled as a collection of singular isothermal spheres, distributed in mass according to the PS theory. The rates of SN explosions as a function of cosmic time are estimated in §5.3, together with the expected number counts of magnified events, distributions of time delays and image separations, and selection effects. As the rest-frame flux of Type Ia SNe falls rather steeply below 3500 Å and we are mostly interested in redshifts $z \gtrsim 1$, the counts are computed in the I-band. We briefly summarize our results in §5.4. The original results presented in this chapter are extracted from Porciani & Madau (1999).

5.2 Basic theory

If the geometry of the universe is well approximated on large scales by the FRW metric, the probability that the light emitted by a point source at redshift z_s is received at $z = 0$ through a lensing event is (Turner, Ostriker, & Gott 1984; Fukugita et al. 1992)

$$\mathcal{P}(z_s) = \int_0^{z_s} dz(1+z)^3 \frac{dl}{dz} \int_0^{\infty} dM\, \sigma(M, z, z_s) n_L(M, z) \ , \tag{5.1}$$

where $dl/dz = cH_0^{-1}(1+z)^{-1}[\Omega_0(1+z)^3 + \Omega_K(1+z)^2 + \lambda_0]^{-1/2}$ is the cosmological line element, $\Omega_K = 1 - \Omega_0 - \lambda_0$ is the curvature contribution to the present density parameter, $\sigma(M, z, z_s)$ is the cross section for lensing as measured on the lens-plane, and $n_L(M, z)$ denotes the comoving differential distribution of lenses at redshift z with respect to the variable M (i.e. mass or velocity dispersion). Equation (5.1) assumes that each bundle of light rays encounters only one lens, the lens population is randomly distributed, and the resulting $\mathcal{P} \ll 1$.

The mass distribution in a single lens and the geometry of the source-lens-observer system completely determine the lensing cross section. The geometry of the lens system generally affects $\sigma(M, z, z_s)$ through the ratio $D(0, z_l)D(z_l, z_s)/D(0, z_s)$, with $D(z_1, z_2)$ the angular diameter

distance between z_1 and z_2. In a smooth (homogeneous) FRW universe, one has:

$$D(z_1, z_2) = \frac{c}{H_0} r_{\rm h}(z_1, z_2) = \frac{c}{H_0} \frac{S(\chi_2 - \chi_1)}{\sqrt{\kappa}(1+z_2)} , \tag{5.2}$$

where $S(x) = \sinh(x)$ if $\Omega_K > 0$, $S(x) = x$ if $\Omega_K = 0$ while $S(x) = \sin(x)$ if $\Omega_K < 0$, and

$$\chi_2 - \chi_1 = \sqrt{\kappa} \int_{z_1}^{z_2} \frac{dz}{[\Omega_0(1+z)^3 + \Omega_K(1+z)^2 + \lambda_0]^{1/2}} \tag{5.3}$$

with $\kappa = 1$ if $\Omega_K = 0$ and $\kappa = |\Omega_K|$ in all other cases. Here, the density of matter within a cone of light rays is always the same as the mean density in the universe (filled beam approach). The observed universe, however, is clumpy on galaxy scales, and local mass concentrations are known to produce significant changes in the redshift-distance relation (Zel'dovich 1964). The Dyer-Roeder formula (Dyer & Roeder 1973; hereafter DR) gives the angular diameter distance in a universe where only a constant fraction $\alpha \leq 1$ of the mass is uniformly distributed, the remaining part being strongly clustered into randomly distributed clumps (see Linder 1998 for calculations including an evolving clumping factor). This is obtained in the "average-path" approximation, i.e. assuming that the typical behavior of a light bundle can be represented by propagating each ray in a smoothed (i.e. averaged) background. In the extreme case $\alpha = 0$ and for bundles of light rays that do not cross any condensation save for the lens (empty beam approach), one obtains

$$D(z_1, z_2) = \frac{c}{H_0} r_{\rm c}(z_1, z_2) = \frac{c}{H_0}(1+z_1) \int_{z_1}^{z_2} \frac{dz}{(1+z)^2[\Omega_0(1+z)^3 + \Omega_K(1+z)^2 + \lambda_0]^{1/2}} . \tag{5.4}$$

This formula is only acceptable for light beams that subtend a small solid angle, so that the probability of intercepting a clump is small. In fact, for very large apertures, the mass density in a beam approaches the average one and the angular diameter distance reduces, practically, to the FRW case.[1]

The distance ambiguity also influences statistical lensing methodology. Equation (5.1) is obtained by integrating the differential probability for lensing along the unperturbed line-of-sight to the source, and $\mathcal{P}(z_s)$ express the probability obtained by considering an ensemble of random lines-of-sight. Ehlers & Schneider (1986) argued that, if the DR treatment of light propagation is a valuable hypothesis, the line-of-sight to the source cannot be considered as a random variable in a clumpy universe, since only those bundle of light rays that travels far away from all clumps are consistent with the DR formula. Therefore, they proposed a different formula for the lensing probability by considering the position of the source on the sphere at $z = z_s$ as a random variable. The lensing optical depth is then given by the fraction of the source sphere that undergoes a strong-lensing event. When this probability is much smaller than one, one has

$$\mathcal{P}(z_s) = \frac{r_{\rm c}^2(z_s)}{r_{\rm h}^2(z_s)} \int_0^{z_s} dz(1+z)^3 \frac{dl}{dz} \frac{r_{\rm h}^2(z)}{r_{\rm c}^2(z)} \int_0^{\infty} dM\, \sigma(M, z, z_s) n_L(M, z) , \tag{5.5}$$

where $r_{\rm h}(z) = r_{\rm h}(0, z)$ and $r_{\rm c}(z) = r_{\rm c}(0, z)$ are adimensional angular diameter distances. Both FRW and DR distances enter equation (5.5) to account, respectively, for large-scale geometry and light propagation.

In the following, we will consider both the filled beam and the empty beam approximations (in the "classic" and Ehlers-Schneider variant) in order to bracket the uncertainties related to the

[1]Numerical studies based on ray-shooting techniques (Kasai, Futamase & Takahara 1990; Watanabe & Tomita 1990; Fukushige & Makino 1994; Tomita 1998) have shown, however, that a definite distance-redshift relation does not exist in a clumpy universe.

lack of a reliable model for light propagation in a clumpy universe. Depending on the underlying cosmology, the discrepancies between the lensing probabilities obtained within the various schemes become significant only for $z \gtrsim 2$.

5.2.1 Singular isothermal lens

The standard minimal model to describe the mass distribution of an extended lens (galaxy or cluster) is the singular isothermal sphere (SIS),

$$\rho(r) = \frac{\sigma_v^2}{2\pi G}\frac{1}{r^2}\,, \tag{5.6}$$

where the mass density ρ as a function of the radial coordinate r is parametrized through σ_v, the one-dimensional velocity dispersion. By projecting the SIS mass distribution along the line-of-sight, one obtains

$$\Sigma(\xi) = \frac{\sigma_v^2}{2G}\frac{1}{\xi}\,, \tag{5.7}$$

where ξ is the distance from the center of the two-dimensional profile. In the thin lens approximation (e.g., Schneider, Ehlers & Falco 1992; Narayan & Bartelmann 1997), this corresponds to a constant deflection angle for incident light rays, $\beta = 4\pi(\sigma_v/c)^2 = 1.4'' \ (\sigma_v/220 \text{ km s}^{-1})^2$, always pointing towards the lens center of symmetry. For a source that, in the absence of lensing, would be seen at an angular distance θ from this center, the lens equation leads to a solution (image) at $\theta_+ = \theta + \theta_E$ with magnification $\mu_+ = \theta_E/\theta + 1$, where $\theta_E = \beta D_{ls}/D_s$ is the Einstein radius [hereafter $D_{ls} \equiv D(z_l, z_s)$, $D_s \equiv D(z_s)$ and $D_l \equiv D(z_l)$]. If the alignment is close enough for multiple imaging, $\theta \leq \theta_E$ (strong lensing), a second image is produced at $\theta_- = \theta - \theta_E$ with magnification $\mu_- = \theta_E/\theta - 1$. The separation between the two images is $2\theta_E$, and their time delay is $c\Delta t = 2\theta\theta_E(1+z_l)D_l D_s/D_{ls}$. At a given angle θ, the total magnification of the two images is proportional to their separation. A third image, with zero magnification, is located in correspondence of the lens center, and acquires a finite flux if the singularity that characterizes the lens profile is replaced by a core region with finite density.

A cross section for strong lensing events, $\sigma_{\rm SIS}$, can be easily associated to each SIS lens. Measuring $\sigma_{\rm SIS}$ on the lens plane, we obtain

$$\sigma_{\rm SIS}(\sigma_v, z_s, z_l) = \pi(\theta_E D_l)^2 = 16\pi^3 \left(\frac{\sigma_v}{c}\right)^4 \left(\frac{D_l D_{ls}}{D_s}\right)^2\,. \tag{5.8}$$

The cross section for the image outside the Einstein ring (including both multiple and singly imaged objects) being magnified by a factor $\mu > 1$ can be easily related to $\sigma_{\rm SIS}$ through

$$\sigma(\mu_+ > \mu) = \frac{\sigma_{\rm SIS}}{(\mu-1)^2}\,. \tag{5.9}$$

Hence the cross section associated with magnification (say) $\mu = 1.316$ (corresponding to a brightening of 0.3 magnitudes) is $\sigma(\mu_+ > 1.316) = 10\,\sigma_{\rm SIS}$. Weakly magnified events are much more common, $\sigma(\mu_+ > 1.1) = 100\,\sigma_{\rm SIS}$. In the strong lensing regime, the corresponding value for the second (dimmest) image ($\mu > 0$) is

$$\sigma(\mu_- > \mu) = \frac{\sigma_{\rm SIS}}{(\mu+1)^2}\,. \tag{5.10}$$

The cross section for a total magnification $\mu_{\rm tot} = \mu_+ + \mu_- = 2\theta_E/\theta > \mu$ (with $\mu \geq 2$) is

$$\sigma(\mu_{\rm tot} > \mu) = \frac{4\sigma_{\rm SIS}}{\mu^2} \,. \tag{5.11}$$

Therefore, given a population of sources at z_s, the distribution of the corresponding magnifications (μ_+, μ_- and $\mu_{\rm tot}$) observed at $z = 0$ is directly proportional to the probability $\mathcal{P}(z_s)$ when the latter is computed introducing $\sigma_{\rm SIS}$ in equation (5.1).

5.2.2 Halo mass function

We will follow the approach of Narayan & White (1988) (see also Kochanek 1995b), and use the PS theory for our analysis of the abundance of gravitational lenses in a cold dark matter (CDM) hierarchical universe. The mass distribution of collapsed halos, assembled through accretion and merging processes, follows then directly from the statistical properties of the linear cosmological density fluctuation field, assumed to be Gaussian. The halo abundance given in equation (4.2) is, in fact, fully determined by the linearly extrapolated (to $z = 0$) variance of the mass-density field smoothed on the scale M, σ_M^2, and by the redshift-dependent critical overdensity δ_c. The latter quantity is obtained instead by solving the spherical top-hat collapse and by computing the linear overdensity corresponding to the collapse time. Details of these calculations can be found in Lacey & Cole (1993) for vanishing cosmological constant, and Eke, Cole & Frenk (1996) for a general flat cosmology. We adopt their analytical approach; alternatively one can use N-body simulations of gravitational clustering to estimate $\delta_c(z)$ as a best fitting parameter (e.g. Lacey & Cole 1994).

Assuming that every halo virializes to form a (truncated) singular isothermal sphere of velocity dispersion σ_v, mass conservation implies

$$\sigma_v(M,z) = \frac{1}{2} H_0 r_0 \, \Omega_0^{1/3} \Delta^{1/6} \left[\frac{\Omega_0}{\Omega(z)}\right]^{1/6} (1+z)^{1/2} \,, \tag{5.12}$$

with $\Omega(z) = \Omega_0(1+z)^3/[\Omega_0(1+z)^3 + \Omega_K(1+z)^2 + \lambda_0]$. Here $r_0 = (3M/4\pi\rho_b)^{1/3}$ denotes the comoving lagrangian (i.e. initial, formally when $z \to \infty$) radius of the collapsing perturbation, and $\Delta(z)$ is the ratio between the actual density at virialization and the critical density at redshift z. This is determined by assuming that the virialization time is twice the turn-around time (in the limit $\Omega_0 = 1$, one finds the standard result $\Delta = 18\pi^2$). While in a universe with $\Omega_\Lambda = 0$ the radius of a virialized spherical perturbation, $R_{\rm vir}$, is half its turnaround radius, $R_{\rm to}$, in the presence of a cosmological constant the radius at virial equilibrium depends on the vacuum contribution to the background energy density, and is always smaller that $R_{\rm to}/2$ (Lahav et al. 1991). A detailed fitting formula for the ratio $R_{\rm vir}/R_{\rm to}$ as a function of λ_0 and z is given in Kochanek (1995b).

Equation (5.12) relates the PS mass function to the SIS lens profile we discussed in the previous section, therefore allowing the computation of the probabilities given in equations (5.1) and (5.5).

5.2.3 Lensing optical depth in CDM cosmologies

We analyse the frequency of lensing events in three different cosmologies with parameters suggested by a variety of recent observations: an open model (OCDM, $\Omega_0 = 0.3$, $\lambda_0 = 0$), a flat model (ΛCDM, $\Omega_0 = 0.35$, $\lambda_0 = 0.65$), and an Einstein-de Sitter model (SCDM, $\Omega_0 = 1$, $\lambda_0 = 0$). In all cases the amplitude of the power spectrum or, equivalently, the value of the *rms* mass fluctuation in a 8 h^{-1} Mpc sphere, σ_8, has been fixed in order to reproduce the observed abundance of rich galaxy clusters in the local universe (e.g. Eke, Cole & Frenk 1996; Jenkins et al. 1998). The model parameters are shown in Table 5.1. The "cosmic concordance" model proposed by Ostriker & Steinhardt

Table 5.1: Cosmological parameters

Model	Ω_0	λ_0	h	n	σ_8
SCDM	1.00	0.00	0.50	1.00	0.50
OCDM	0.30	0.00	0.70	1.00	0.85
ΛCDM	0.35	0.65	0.65	0.96	0.87

(1995) has been selected to represent the whole class of Λ-dominated cosmogonies (ΛCDM). While this model is able to account for most known observational constraints, ranging from globular clusters ages to CMB anistropies to recent analyses of the Type Ia SNe Hubble diagram, it may be marginally in conflict with the statistics of gravitational lenses (Kochanek 1996). Our choice of an open CDM (OCDM) cosmology is guided by the growing evidence in favour of a low value of Ω_0 from cluster studies (White & Fabian 1995; Carlberg, Yee, & Ellingson 1997). For reference, and as instructive counterpart to the previous variants, we also consider a standard CDM (SCDM) scenario. Figure 5.1 shows the probability of a strong-lensing ($\sigma = \sigma_{\rm SIS}$) event for a source at redshift $z_s \leq 3$. The lensing frequency is the highest in SCDM and the lowest (by about a factor of 2 at $z_s = 3$) in OCDM cosmologies, with ΛCDM models in between. The relative amplitudes depend on a combination of effects: (*a*) The geometry of the lens system affects the probability through the ratio $(D_l D_{ls}/D_s)^2 \equiv (c/H_0)^2 g(z_l, z_s)$ and only halos in a narrow redshift interval can be efficient lenses for a source at redshift z_s. If $z_s = 2$ ($z_s = 1$), for example, the maximum contribution to $\mathcal{P}(z_s)$ comes from halos at $z \sim 0.5$ ($z \sim 0.3$). In the filled beam approximation with $z_l = 0.5$ and $z_s = 2$ (say), the ratios between the geometric factors in the various cosmologies are $g_{\rm SCDM}/g_{\Lambda{\rm CDM}} = 0.58$ and $g_{\rm SCDM}/g_{\rm OCDM} = 0.85$. (*b*) The path-length to a source at $z_s = 2$ is (in units of the Hubble radius) 0.71, 0.61, and 0.54 in ΛCDM, OCDM, and SCDM universes, respectively. (*c*) For a fixed mass, the halo velocity dispersion shows only a weak dependence on the world model. At $z_l = 0.5$, for example, σ_v^4 in ΛCDM is smaller than in SCDM by 23%. On the other hand, OCDM shows the largest velocity dispersions, with σ_v^4 larger than in SCDM by 6%. (*d*) The number density of halos obviously influences the lensing frequency. Roughly speaking, for objects with mass $M \ll M_*$, the number of halos contained in a Hubble volume is proportional to Ω_0, i.e. there are approximately three times more lenses in SCDM than in low-density universes. By contrast, halos with $M \gtrsim M_*$ form earlier in low-density cosmogonies. In SCDM, the formation of the massive (and exponentially rare) objects which are responsible for large separation multiple images is delayed to such low redshifts that they fail to be efficient lenses for sources at $z_s > 1$. This is graphically shown in Figure 5.2, where the function $1/M\sigma_{\rm SIS}\, n_L(M, z_l)$, the mean-free-path per logarithmic mass interval of a beam to encountering a lensing event, is plotted as a function of mass for different values of z_l and a source at $z_s = 2$. The lensing optical depth is dominated by masses in the range 10^{12} to 10^{14} $M_\odot$, i.e. by massive galaxies, groups, and poor clusters. The most efficient foreground lenses are at a redshift less than unity. The larger number of lenses in SCDM more than compensate for the longer path lengths to a given redshift and earlier collapse times that characterize open and Λ models. It is fair at this stage to point out that, while it has become customary to model (as we have done) virialized dark matter halos by isothermal spheres, high-resolution N-body simulations have recently shown that the halos that form in a hierarchically clustering universe have a universal density profile which is shallower (but still diverging like r^{-1}) than isothermal near the center of a halo, and steeper than isothermal [with $\rho(r) \propto r^{-3}$] in its outer regions (Navarro, Frenk, & White 1997). Depending on the halo mass (halos of increasing mass being less centrally concentrated, Navarro et al. 1997), a SIS approximation will overestimate

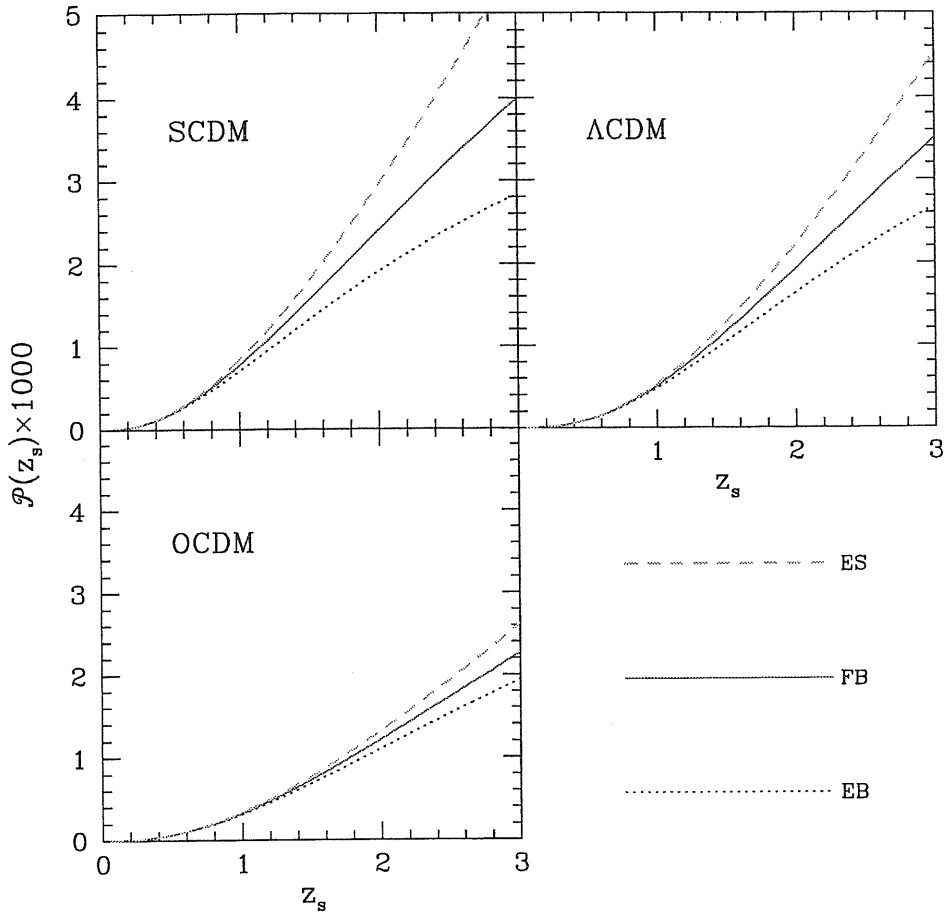


Figure 5.1: Strong lensing probabilities for a point source at z_s in three different hierarchical cosmogonies. The lens distribution is described by the Press-Schechter theory, and each lens is modeled by a singular isothermal sphere. *Solid lines:* filled beam (FB) approximation. *Short-dashed lines:* empty beam (EB) scheme. Both frequencies are computed using eq. (5.1). *Long-dashed lines:* Ehlers-Schneider (ES) probability from eq. (5.5). The lensing optical depth for magnifications $\Delta m > 0.3$ (> 0.1) mag is 10 (100) times larger than plotted here.

the lensing probability and underestimate the total magnification. On the other hand, photometric analyses of E/S0 galaxies performed with the *Hubble Space Telescope* (*HST*) (Tremaine et al. 1994), and a number of gravitational lensing studies (e.g. Wallington & Narayan 1993; Kochanek 1995a) have shown that a singular isothermal profile is a good approximation for lensing galaxies. The giant arcs associated with the lensing of background galaxies by galaxy clusters also seem require cluster "core" radii (where the density is approximately constant) to be rather small (Grossman & Narayan 1988; Soucail & Mellier 1994).

Another simplifying assumption we made involves the spatial distribution of the lens population, here modeled with a Poisson process. Indeed, the clustering properties of dark halos are known to slightly alter the lensing statistics. In particular, the presence of rich aggregates and large voids will increase the frequency of events in the low- and high-magnification tails of the distribution (e.g., Jaroszyński 1992).

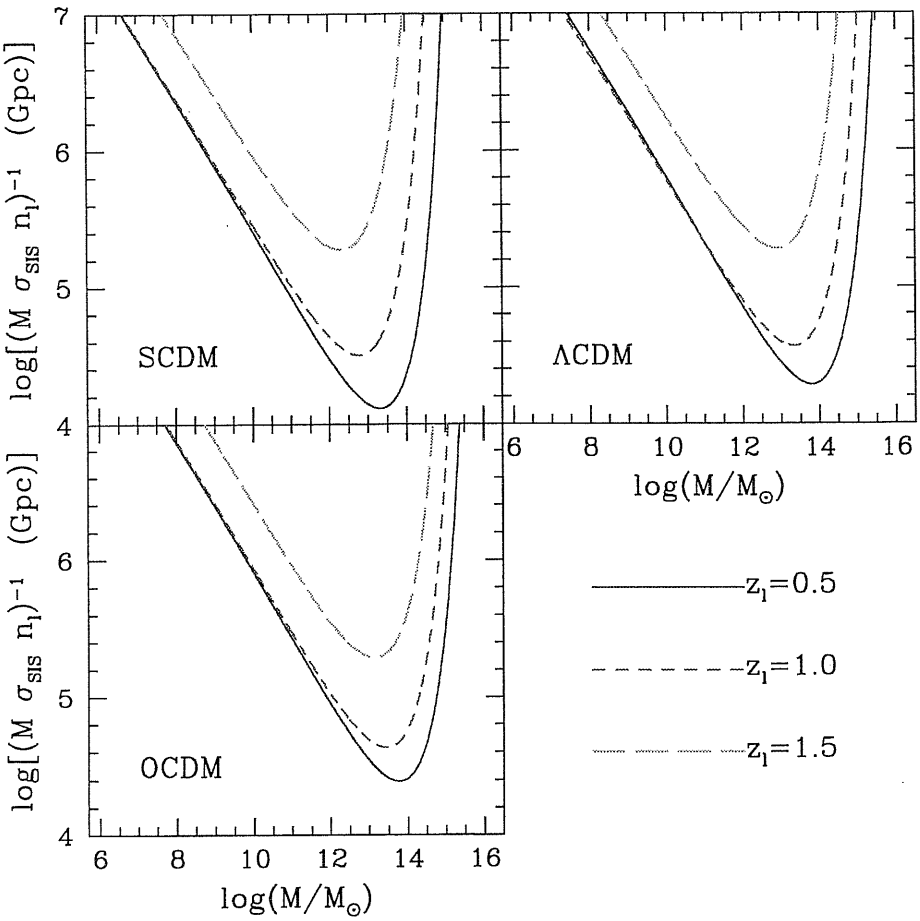


Figure 5.2: Mean-free-path per logarithmic mass interval of a (filled) beam to encountering a strong lensing event, as a function of mass for different values of the lens redshift z_l and for different cosmological models. The source is assumed to be at $z_s = 2$.

5.2.4 Galaxy lensing

As a way to gauge the possible shortcomings of the PS method discussed above [e.g. the intrinsic lack of accuracy at the low-mass tail, Lacey & Cole 1994, Gross et al. 1998 Lemson & Kauffmann 1999, as well as uncertainties in $\delta_c(z)$ or in the conversion between mass and velocity dispersion], it is interesting to consider in some detail the contribution of known galaxies only to the magnification frequency. We model the lensing galaxies again as SIS. Their luminosity function,

$$n(L) = \frac{n_*}{L_*}\left(\frac{L}{L_*}\right)^{\alpha} \exp\left(-\frac{L}{L_*}\right) , \tag{5.13}$$

includes just early E/S0 types, as late spirals and irregulars are known to make a small contribution ($< 20\%$) to the lensing statistics (Kochanek 1996; Maoz & Rix 1993). The velocity dispersion of the dark matter in each galaxy halo is related to the luminosity through a Faber-Jackson law

$$L = L_*\left(\frac{\sigma_v}{\sigma_*}\right)^4 . \tag{5.14}$$

The assumption that the local properties of E/S0 galaxies persist out to $z \sim 1$ has some observational basis (e.g. Lilly et al. 1995; van Dokkum et al. 1998). It also appears that, by and large, the statistics of multiple imaged QSOs is consistent with the locally observed galaxy population extrapolated to higher redshifts (Kochanek 1993; Maoz & Rix 1993). If the galaxy comoving number density does not change with cosmic time, and one neglects all evolutionary corrections to equation (5.14), equation (5.1) can be simplified, and the lensing probability becomes proportional to a non-dimensional constant factor F (Turner, Ostriker, & Gott 1984)

$$F = \left(\frac{c}{H_0}\right)^3 \left(\frac{D_l D_{ls}}{D_s}\right)^{-2} \int_0^\infty dM\, \sigma_{\rm SIS}(M, z, z_s) n_L(M, z) = \\ = 16\pi^3 n_* \left(\frac{c}{H_0}\right)^3 \left(\frac{\sigma_*}{c}\right)^4 \Gamma(2+\alpha)\ , \tag{5.15}$$

where $\Gamma(x)$ is Euler's gamma function. We use the best-fit E+S0 Schechter parameters from the SSRS2 samples ($z \leq 0.05$, Marzke et al. 1998): $\alpha = -1.00 \pm 0.09$, $n_* = (4.4 \pm 0.8) \times 10^{-3} h^3$ Mpc^{-3}, and $M_* = -19.37 \pm 0.1 + 5 \log h$ (the absolute blue magnitude associated with a luminosity L_*). With a fiducial value of $\sigma_* = 220$ km s^{-1} (Kochanek 1994), this model yields $F = 0.017$ (cf. Fukugita & Turner 1991). For comparison, the all-type luminosity function from the Autofib survey ($0.02 < z < 0.15$, Ellis et al. 1996), when scaled by the Postman & Geller (1984) fraction (31%) of E/S0 galaxies, gives $\alpha = -1.16 \pm 0.05$, $n_* = (7.5 \pm 1.3) \times 10^{-3} h^3$ Mpc^{-3}, $M_* = -19.30 \pm 0.13 + 5 \log h$, and $F = 0.033$. With these assumptions, cosmology only affects the lensing probability through geometric effects, and the number of lensed sources is the highest in a ΛCDM model. How does the probability of galaxy lensing compare then with the PS frequency shown in Figure 5.1? For a source at (say) $z_s = 2$, the former (with Ellis et al. parameters) is 1.27 times smaller than the latter in OCDM, 1.13 times smaller in ΛCDM, and 3.65 times smaller in SCDM. The number of galaxy lenses is two times smaller with Marzke et al. parameters. For a source at $z_s = 1.5$, about 70% (40%) of the PS optical depth in SCDM is due to halos with velocity dispersion $\sigma_v > 150$ km s^{-1} (> 250 km s^{-1}). In order to include the effects of galaxy groups and clusters, in the following sections we shall estimate the number of magnified SNe from the PS formalism. Unless otherwise stated, we will assume a SCDM cosmology in the filled beam approximation. It is clear from the above discussion, however, that the uncertainties in the lensing probability are non-negligible. One should also keep in mind that the PS mass function, $n(M) \propto M^\zeta$ with $\zeta \approx -2$, is significantly steeper at the low-mass end than that derived from equations (5.6), (5.13), and (5.14), $n(M) \propto M^{2\alpha+1}$. The statistics of gravitational lensing within the PS theory will tend to overestimate the frequency of small-separation images.

5.3 Supernovae

5.3.1 Rates

The rates of SNe as a function of cosmic time depend on the star formation history of the universe, $\psi(t)$, the initial mass function (IMF) of stars, $\phi(m)$, and the nature of the binary companion in Type Ia events. The frequency of core-collapse supernovae, SN II and possibly SN Ib/c, which have short-lived ($\lesssim 50$ Myr) progenitors (e.g. Wheeler & Swartz 1993), is essentially proportional to the instantaneous stellar birthrate of stars with mass $> 8\,M_\odot$. For a Salpeter IMF (with lower and upper mass cutoffs of 0.1 and 125 $M_\odot$), the core-collapse supernova rate can be related to the star

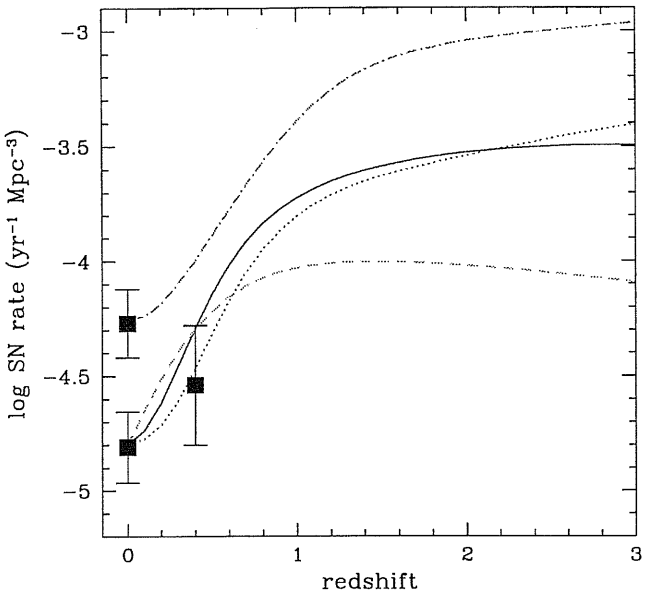


Figure 5.3: Estimated Type Ia and II (rest-frame) frequencies as a function of redshift. *Dot-dashed line:* SN II rate. *Dotted line:* SN Ia rate with $\tau = 0.3$ Gyr. *Solid line:* SN Ia rate with $\tau = 1$ Gyr. *Dashed line:* SN Ia rate with $\tau = 3$ Gyr. The model assumes a SCDM cosmology. The data points with error bars have been derived from the measurements of Cappellaro et al. (1997), Tammann et al. (1994), Evans et al. (1989), and Pain et al. (1997). See MDP for details.

formation rate per unit comoving volume according to

$$R_{\rm II}(t) = \psi(t)\frac{\int_8^{125} dm\,\phi(m)}{\int_{0.1}^{125} dm\,m\,\phi(m)} = 0.0074 \times \left[\frac{\psi(t)}{\mathrm{M}_\odot\,\mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}}\right]\ \mathrm{yr}^{-1}\,\mathrm{Mpc}^{-3}. \qquad (5.16)$$

By contrast, the specific evolutionary history leading to a Type Ia event is poorly known. SN Ia are believed to result from the explosion of C-O white dwarfs (WDs) triggered by the accretion of material from a companion, possibly another WD or a non-degenerate, evolved star (e.g. Ruiz-Lapuente, Canal, & Burkert 1997). The clock for the explosion is set by the lifetime of the primary star and, e.g., by the time taken for a non-degenerate companion to evolve and fill its Roche lobe (or, in the case the companion is another WD, the orbital decay time following gravitational wave emission). The evolution of the rate may depend then, among other things, on the unknown mass distribution of the secondary binary components and on the distribution of initial orbital separations. In this sense Type Ia SNe follow a slower evolutionary clock than Type II's, and their rate at any given time is sensitive to the past history of star formation in galaxies. Here we follow Madau, Della Valle, & Panagia (1998, hereafter MDP), and parameterize the rate of Type Ia's in terms of a characteristic explosion timescale, τ (which defines an explosion probability per WD assumed to be independent of time), and an explosion efficiency, η. The former accounts for the time elapsed in the various scenarios from a newly born (primary) WD to the SN explosion itself, the latter for the fraction of stars in binary systems that will never undergo a SN Ia explosion because of unfavorable initial conditions. The rate of Type Ia events at any one time is then given

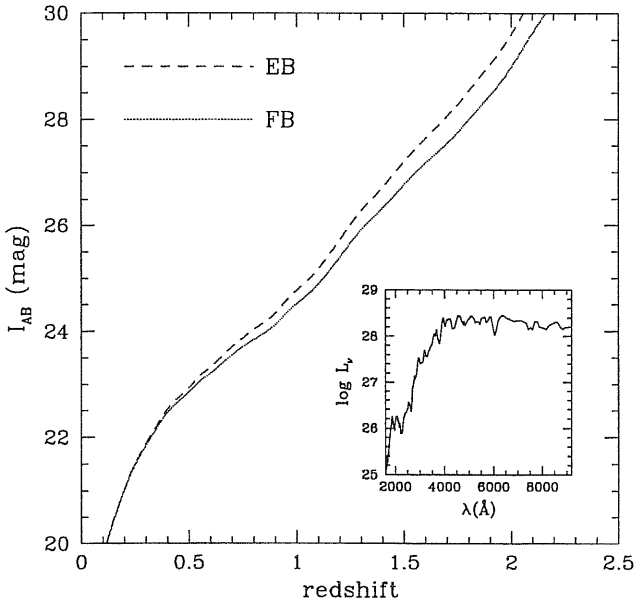


Figure 5.4: I-band magnitude-redshift relation (SCDM) for Type Ia SNe. The apparent AB magnitude has been computed by redshifting and dimming the rest-frame *HST*–FOS spectrum of SN 1992A, and renormalizing it to $M_V = -19.4$. (The renormalized spectrum is shown in the inset, see Kirshner et al. 1993). The emitted power L_ν is in ergs s^{-1} Hz^{-1}.

by the explosions of all the binary WDs produced in the past that have not gone off already, i.e.

$$R_{\rm Ia}(t) = \frac{\eta \int_0^t dt' \, \psi(t') \int_{m_c}^8 dm \, \exp(-\frac{t-t'-t_m}{\tau}) \phi(m)}{\tau \int dm \, \phi(m)}, \tag{5.17}$$

where $m_c \equiv \max[3, m(t')]$, $m(t') = (10\,{\rm Gyr}/t')^{0.4}$ is the minimum mass of a star that reaches the WD phase at time t', and $t_m = 10\,{\rm Gyr}/m^{2.5}$ is the standard lifetime of a star of mass m (all stellar masses are expressed in solar units). Figure 5.3 shows the predicted Type Ia and II rates per unit proper time and unit comoving volume for a cosmic star formation history that traces the rise from $z = 0$ to $z \approx 1.5$ of the galaxy UV emissivity (Madau, Pozzetti, & Dickinson 1998). The stellar evolution model (depicted in Figure 1*b* of MDP) assumes a Salpeter IMF and a weakly increasing star formation density above $z \approx 2$, and is able to account for the entire optical background light recorded in the galaxy counts: half of the present-day stars – the fraction contained in spheroidal systems – formed at $z > 2.5$. Because of the uncertainties in the amount of starlight that is absorbed by dust and re-radiated in the far-IR as a function of cosmic time, these numbers are only meant to be indicative. Our estimates may be on the conservative side, as the assumed stellar evolution model produces only a fraction $\sim 50\%$ of the IR background detected by COBE (Dwek et al. 1998). On the other hand, dust along the line-of-sight may affect the detectability of Type II SNe (and also Type Ia's if the explosion delay timescale is short enough that dust is retained in the star-forming regions.) No attempt has been made to include the possibility of optically hidden SNe in our calculations.

The Type Ia rates plotted in the figure assume characteristic "delay" timescales after the collapse

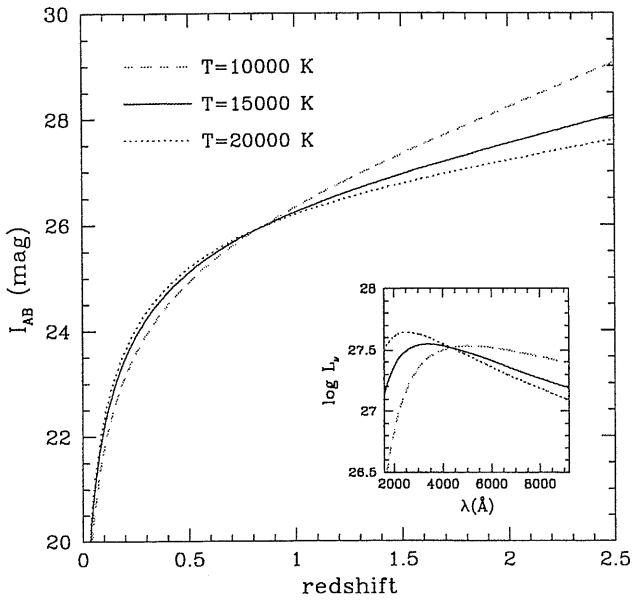


Figure 5.5: Magnitude-redshift relation (SCDM, filled beam) for Type II SNe. The apparent I magnitude has been computed by assuming a blackbody spectrum at different temperatures. In all cases the emitted power has been renormalized to $M_B = -17.2$. *Dashed line*: $T = 10000$ K. *Solid line*: $T = 15000$ K. *Dotted line*: $T = 20000$ K.

of the primary star to a WD equal to $\tau = 0.3, 1$, and 3 Gyr, which virtually encompass most relevant possibilities. For a fixed initial mass m, the frequency of Type Ia events peaks at an epoch that reflect an "effective" delay $\Delta t = \tau + t_m$ from stellar birth. *A prompter (smaller τ) explosion results in a higher SN Ia rate at early epochs.* The explosion efficiency ($5 \lesssim \eta \lesssim 10\%$) was adjusted to reproduce the observed ratio of Type II to Type Ia rates in the local universe. When normalized to the emitted blue luminosity density, the predicted frequencies match rather well the data at $0 \leq z \leq 0.4$ (MDP).

5.3.2 Number counts of SN Ia

In this section we will derive the observationally relevant quantity, i.e. the rate of significantly magnified SNe in a flux-limited survey, given a cosmological model, lens population, and star formation history. An apparent magnitude-redshift relation for unlensed SNe is needed to estimate the number counts of distant Type Ia events. Traditional (i.e. from observed B- to emitted B-band light) K-corrections out to $z \approx 0.2$ have been calculated by Hamuy et al. (1993) using available spectra of nearby SN Ia. Multi-band photometry has been used to extend these results out to $z \approx 0.6$ (Kim, Goobar & Perlmutter 1996). No accurate K-corrections have been determined, however, for supernovae at $z \geq 0.8$, as very few ultraviolet spectra from the local samples with good signal-to-noise ratio are currently available. To circumvent this problem, we have constructed an (unlensed) magnitude-redshift relation for Type Ia's by using as a template the *HST*–FOS spectrum of SN 1992A taken 5 days after maximum blue light (Kirshner et al. 1993). No correction for foreground and intrinsic extinction is apparently needed for this event (Burnstein & Heiles 1984;

Kirshner et al. 1993). The observed spectrum has been renormalized to a brightness (at maximum B-light) of $M_V = -19.4$. The $m - z$ relation in the I-band is shown (both in the empty and filled beam approximations) in Figure 4 for $z < 2.5$. Note that, because of the steep optical/UV spectrum of 1992A, our template Type Ia SN at $z = 2.5$ will be much brighter in the near-IR ($J_{\rm AB} = 26.9$) than in the optical ($I_{\rm AB} = 31.2$, filled beam, SCDM). The K-correction significantly weakens the observed I flux relative to the luminosity distance D_L^{-2} effect.

For the intrinsic luminosity function of SN Ia, $\Phi_0(I, z)$, we have adopted a Gaussian with standard deviation 0.15 mag (cf. Van den Bergh & McClure 1994). The lensed distribution, $\Phi(I)$, can be obtained by convolving $\Phi_0(I, z)$ with the magnification probability density (Vietri & Ostriker 1983)

$$\Phi(I, z) = \int dI' \, \Phi_0(I', z) \mathcal{P}(\Delta m, z) \; , \tag{5.18}$$

where $\mathcal{P}(\Delta m, z)$ is the probability of having magnification $\Delta m = I' - I$ for a source at redshfit z. The *observed* rate on the sky of Type Ia events is then

$$\frac{dN}{dt}(< I_{\rm lim}) = \int dz \, \frac{dV(z)}{dz} \frac{R_{\rm Ia}(z)}{1+z} \int_{-\infty}^{I_{\rm lim}} dI \, \Phi(I, z) \; , \tag{5.19}$$

where $V(z)$ is the comoving volume element surveyed, and the factor $(1+z)^{-1}$ accounts for the cosmological time dilation.[2] From equation (5.18) we find

$$\frac{dN}{dt}(< I_{\rm lim}) = \int dz \, \frac{dV(z)}{dz} \frac{R_{\rm Ia}(z)}{1+z} \int_{-\infty}^{+\infty} dI \, \Phi_0(I, z) \int_{\mu_{\rm min}(I)}^{+\infty} d\mu \, \mathcal{P}(\mu, z) \; , \tag{5.20}$$

where $\log[\mu_{\rm min}(I)] = (I - I_{\rm lim})/2.5$ is the minimum magnification needed to detect a source in a flux-limited survey. The above equation relates the number counts to the probability distribution of magnification discussed in §5.2. To compute the total number of supernovae brighter than $I_{\rm lim}$ that have been gravitationally magnified by more than Δm magnitudes, the function $\mathcal{P}(\mu)$ in equation (5.20) must be integrated from $\max(\mu_{\rm min}, 10^{\Delta m/2.5})$ to infinity. Our treatment will be limited to magnifications $\Delta m \geq 0.3$ mag, i.e. to images that are more than 2σ away from the mean intrinsic luminosity of SN Ia. In this regime, the weak (with typical magnifications of the order of a few per cent) lensing due to large scale structure can safely be neglected.

Figure 5 shows the counts of SNe Ia at maximum B-light for a SCDM cosmology (filled beam). The rate of detection in the absence of gravitational lensing effects (which is very close to the total expected number of events, since the probability of lensing is small) is compared with the estimated rates of double-imaged events [obtained by requiring that the dimmest image is still sufficiently bright to enter the flux-limited sample, see eq. (5.10)], and with the rates of individual SNe (brightest image only) magnified by more than 0.3 and 0.75 magnitudes [in this case $\mathcal{P}(\mu)$ is computed using the cross section in eq. (5.9)]. In a one square degree field, and for an effective observation duration of 1 yr, the total expected number of Type Ia SNe down to $I_{\rm AB} \leq 25$ mag is (assuming $\tau = 1$ Gyr) ~ 400. About 3 events will be gravitationally lensed by more than 0.3 mag, 0.4 by more than 0.75 mag, and only 0.15 will have detectable double images. These numbers are sensitive to the assumed explosion timescale. Given a well-determined cosmic star formation history, it may be possible to infer something about the nature of Type Ia SNe from the observed number counts. It is also interesting to note that, in general, gravitational magnification facilitates

[2]Note that, as in a snapshot survey SNe are not necessarily detected at maximum light, the total number of events will actually depend on the effective timescale over which a SN can be observed above threshold.

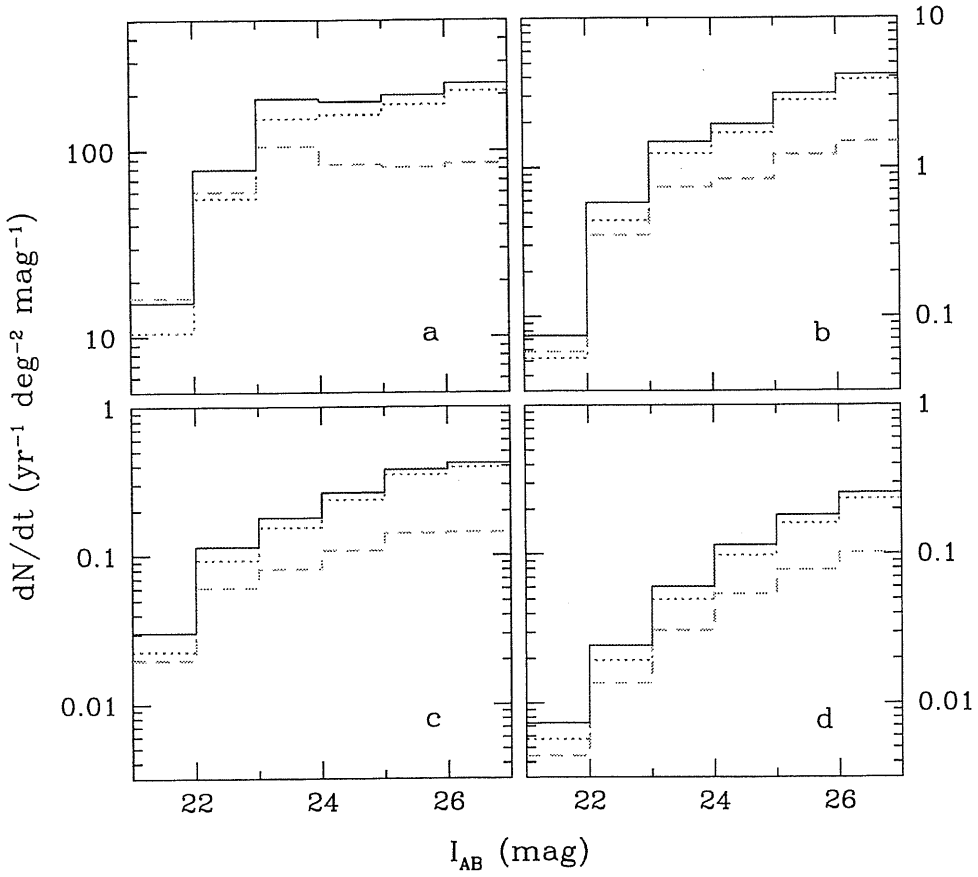


Figure 5.6: Differential number counts of Type Ia SNe at maximum B-light versus limiting I-band magnitude of the survey (SCDM universe, filled beam). *a*) SNe in the absence of lensing; *b*) lensed SNe with magnification greater than 0.3 magnitudes (brightest image only); *c*) lensed SNe with magnification greater than 0.75 mag (brightest image only); *d*) lensed SNe with double images (dimmest image above the detection threshold). *Dotted line:* rate assumes a time delay $\tau = 0.3$ Gyr. *Solid line:* $\tau = 1$ Gyr. *Dashed line:* $\tau = 3$ Gyr. The effect of dust extinction on the detectability of Type Ia's has not been included in the models.

the detection of SNe at higher redshift in a flux-limited sample, and the fraction of strongly-lensed SNe tends to be higher at fainter magnitudes if the rate of events increases with look-back time and the gain in available volume is not balanced by a large K-correction. This is true up to $I_{\rm AB} \approx 23$. For fainter magnitudes, the increase in physical volume is smaller and the SN rate saturates (cf. Fig. 3); deeper searches will not lead then to an increased fraction of magnified sources. This defines an optimal survey sensitivity in the I-band for Type Ia's magnified by $\Delta m \geq 0.75$ mag of $I_{\rm AB} \approx 23 - 24$ (cf. Linder et al. 1988). By contrast, there is no optimal flux threshold for detecting double images (Fig. 5d) in succession.

5.3.3 Number counts of SNe II

Type II SNe are a couple of magnitudes fainter than SN Ia in the blue band. Contrary to Type Ia's, they have a very wide luminosity function, which is well represented by a broad Gaussian with

$\langle M_B(\max)\rangle = -17.2$ ($h = 0.5$) and dispersion 1.2 mag (Tammann & Schröder 1990).[3] The spectra of most SN II at maximum light approximate well a single-temperature Planck function of about 15000 K from the UV through IR – by contrast, Type Ia's exhibit a prominent UV deficit relative to the blackbody fit at optical wavelengths (e.g. Filippenko 1997). The difference in behavior of the K-correction is clearly visible in Figure 5, where the I-band $m - z$ relation for Type II's is plotted for three different photospheric temperatures at peak. The intrinsically fainter SN II appear brighter than SN Ia already at $z \approx 1.5$. This has a strong effect on the expected number counts, shown in Figure 6. At $I_{\rm AB} = 24$ mag, there is one SN II every 2.3 Type Ia's. The detection rate increases to one Type II every 1.2 SN Ia at $I_{\rm AB} = 25$, and to one every 0.56 at $I_{\rm AB} = 26$. Because of the flat K-correction and wide luminosity function, the fraction of lensed events is higher for Type II's, as only small magnifications are needed to bring high redshift events above the flux threshold (see also Linder et al. 1988).

5.3.4 Multiple images and time delays

Although exceedingly rare in the field, multiple imaged Type Ia SNe are of particular interest for cosmography, as the propagation time from the source to the observer varies from one image to another, and can be accurately measured if the source has a well defined light curve. The time delay Δt is proportional to the absolute scale of the lens geometry and hence inversely proportional to the Hubble constant (Refsdal 1964). Given a reliable time delay measurement, a well-constrained lensing mass distribution (with a measured redshift and velocity dispersion), and a world model, one could estimate H_0.

The cumulative probability functions of image separations $\Delta\theta$ and time delays Δt for a SN at $z_s = 2$ in the field are shown in Figure (5.8). About 40% (SCDM) of double images will have angular separations $< 1''$, and will therefore be flagged as multiple events in ground-based surveys only if the time delay is comparable or larger than the SN decay time. The fraction of images with $\Delta\theta > 4''$ is of order 20%. Barring selection effects (see below), approximately 20% of mirror events will be observed with an arrival time difference of < 3 days, 30% with < 10 days, 40% with < 30 days. The frequency distributions have been computed (in the filled beam approximation) by modifying equation (5.1) to account for that fact that, at any redshift, only halos having mass larger than a threshold value can generate two images with a given angular separation or time delay. Normalizing the total probability to unity, one gets

$$\mathcal{P}(>\Delta\theta|z_s) = \mathcal{P}^{-1}(z_s)\int_0^{z_s} dz\,(1+z)^3\frac{dl}{dz}\int_{\mathcal{M}(\Delta\theta,z,z_s)}^{\infty} dM\,\sigma\,n_L\ , \tag{5.21}$$

where $\mathcal{M}$ is the mass that, for given z and z_s, produces a separation $\Delta\theta = 8\pi[\sigma_v(\mathcal{M})/c]^2\ (D_{ls}/D_s)$. Similarly, one obtains

$$\mathcal{P}(>\Delta t|z_s) = \mathcal{P}^{-1}(z_s)\int_0^{z_s} dz\,(1+z)^3\frac{dl}{dz}\int_{\mathcal{M}(\Delta t,z,z_s)}^{\infty} dM\left[1-\frac{\Delta t^2}{\Delta t^2_{\rm max}(M,z,z_s)}\right]\sigma\,n_L\ , \tag{5.22}$$

where $\Delta t_{\rm max}$ is the time delay at $\theta = \theta_E$ and $\mathcal{M}$ solves $\Delta t_{\rm max}(\mathcal{M},z,z_s) = \Delta t$. The differences between low-Ω_0 and $\Omega_0 = 1$ cosmological models are mainly caused by the different mass distributions at $z \sim 0.5$. The presence of very large dark matter halos in OCDM and ΛCDM universes is responsible of an extended tail towards large separations and long delays.

[3]For the present purposes, we shall ignore the possibility of an extended tail of Type II events at bright (Patat et al. 1994) and faint (Woltjer 1997) magnitudes.

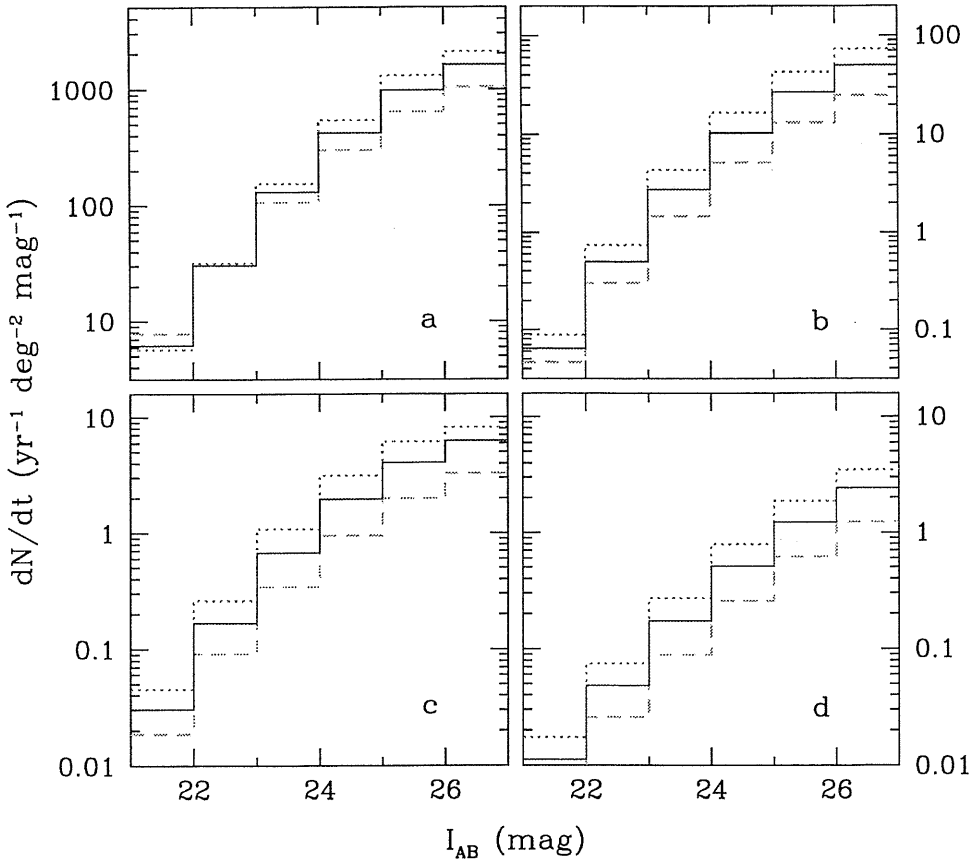


Figure 5.7: Differential number counts of Type II SNe at maximum B-light versus limiting I-band magnitude of the survey (SCDM universe, $z_{\max} = 5$, filled beam). *a*) SNe in the absence of lensing; *b*) lensed SNe with magnification greater than 0.3 magnitudes (brightest image only); *c*) lensed SNe with magnification greater than 0.75 mag (brightest image only); *d*) lensed SNe with double images (dimmest image above the detection threshold). *Dashed line:* rate assumes a photospheric temperature of $T = 10000$ K. *Solid line:* $T = 15000$ K. *Dotted line:* $T = 20000$ K.

5.3.5 Selection effects

A peculiar feature of SNe Ia relative to QSOs is the relatively short lifetime of their emission. While, for example, the finite angular resolution $\theta_{\rm res}$ of the observations may not allow to distinguish multiple images with separation on the sky $\Delta\theta < \theta_{\rm res}$, if the time delay between different images is comparable or longer than the typical decay time of a SN light curve [$t_{\rm SN} \sim 20(1+z_s)$ days], two mirror copies will appear roughly in the same position of the sky but at different times. In this sense, even two images with $\Delta\theta < \theta_{\rm res}$ could in practice flag a strong-lensing event if $\Delta t \gtrsim t_{\rm SN}$: the same SN would go off in two (or more) incarnations of the same explosion, and the different images will be seen in succession. Because of the geometry of the source-lens-observer system, longer time delays will be obtained when the deflector is closer to the radiation source. Consider, e.g., a SN at $z_s = 2$ in a SCDM cosmology, lensed by a massive halo ($M = 1.2 \times 10^{13}\, M_\odot$, $\sigma_v = 250$ km s^{-1}) at $z_l = 1.2$. Two images will result, with $\Delta\theta = 0.83'' < \theta_{\rm res} \sim 1''$ (say). For a source-lens alignment $\theta = \theta_E/2$, the magnifications of the images will be $\mu_+ = 3$ and $\mu_- = 1$, and the arrival time difference $\Delta t = 81$

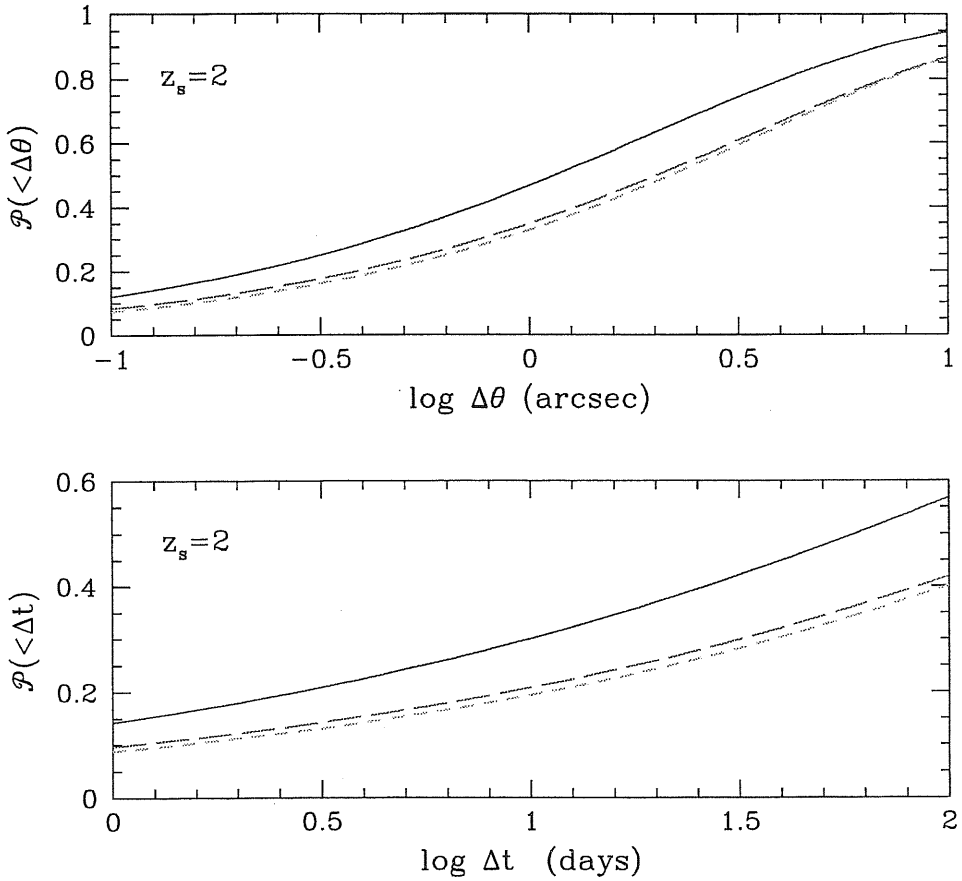


Figure 5.8: Normalized cumulative probability distributions of angular separations and time delays between the images of a SN at $z_s = 2$ (filled beam approximation). *Solid line:* SCDM. *Long-dashed line:* ΛCDM. *Short-dashed line:* OCDM. Events with $\Delta t < 100$ days are dominated by halos with velocity dispersion $\sigma_v < 200\ \mathrm{km\,s^{-1}}$.

days. Unresolved images with shorter delays may still be recognizable due to the large discrepancy between the apparent expected magnitude and the one observed, and a characteristic *double-humped* light curve. At fixed z_l and z_s, the angular separation between the lensed images depends only on the mass of the deflector, while the time delay is anti-correlated with the magnification: smaller delays correspond to more perfect source-lens alignments and thus to larger magnifications. Nearly synchronous images will then have rather similar brightnesses. A SN at $z_s = 2$, lensed by a halo at $z_l = 0.5$ with mass $8.2 \times 10^{12}\, M_\odot$ ($\sigma_v = 220\ \mathrm{km\,s^{-1}}$, SCDM), may produce (when $\theta = 0.15\,\theta_E$) two well resolved images ($\Delta\theta = 1.6''$), with magnifications $\mu_+ = 7.8$ and $\mu_- = 5.8$. The arrival time difference is only 20 days in this example. The maximum time delay occurs for $\theta = \theta_E$ ($\Delta t_{\max} \propto \theta_E^2$), hence less massive deflectors will always produce small-separation events and relatively short delays. A serious bias against the detection of small-separation images is due to the luminosity of the lensing galaxy. Galaxy-type halos with velocity dispersion $\sigma_v < 200\ \mathrm{km\,s^{-1}}$ contribute to about 40–50% of the total lensing optical depth at $z_s \sim 1-1.5$, but completely dominate the rate of small-separation events. These must then be detected in the background of a lens elliptical galaxy, most probably at $z_l \sim 0.5$. Figure 9 compares the observed surface brightness of a foreground L_* elliptical with a

background lensed Type Ia SN as a function of angular distance from the galaxy center. The effect of atmospheric seeing has been approximated by convolving the true source flux with a Gaussian of dispersion $\sigma = \mathrm{FWHM}/\sqrt{8\ln 2}$. Even in the case of excellent FWHM= $0.6''$ seeing, the lens will greatly outshine the faintest SN image. Double events will only be detectable for nearly perfect lens-source alignments, an extremely rare case. In more modest seeing conditions (FWHM= $1''$), even the brightest image may remain undetected. From the ground, multiple SN events will be more easily observed in the background of a galaxy group or cluster.

In a magnitude-limited sample, the sources that are observed to be gravitationally lensed include not only the lensed objects that are intrinsically brighter than the flux limit, but also sources that are intrinsically fainter and are brought into the sample by the magnification. Relative to the lensing probabilities shown in Figure 1, lensed systems will then be overrepresented among SNe of a given apparent magnitude. We have quantified this effect from the Type Ia and Type II number counts, by defining the magnification bias $B(< I_{\rm lim})$ as the ratio between the actual flux-limited counts of sources magnified by more than Δm and the counts of lensed SNe that are intrinsically brighter than the flux limit,

$$B(< I_{\rm lim}) = \frac{\displaystyle\int dz \, \frac{dV(z)}{dz} \frac{R(z)}{1+z} \int_{-\infty}^{+\infty} dI \, \Phi_0(I,z) \int_{\bar{\mu}}^{+\infty} d\mu \, \mathcal{P}(\mu,z)}{\displaystyle\int dz \, \frac{dV(z)}{dz} \frac{R(z)}{1+z} \int_{-\infty}^{I_{\rm lim}} dI \, \Phi_0(I,z) \int_{10^{\Delta m/2.5}}^{+\infty} d\mu \, \mathcal{P}(\mu,z)} \, , \tag{5.23}$$

where $\bar{\mu} = \max(\mu_{\rm min}(I), 10^{\Delta m/2.5})$. This definition generalizes that given in Turner, Ostriker, & Gott (1984) and Fukugita & Turner (1991) by taking into account the redshift dependence of the lensing probability.[4] The magnification bias corresponding to $\Delta m = 0.3$ and 0.75 mag is plotted in Figure (10) for Type Ia and Type II SNe. At $I_{\rm AB} \leq 22$ the effect is huge, as the apparent frequency of strongly lensed events is boosted by a factor of 40. Note that at these magnitude levels the counts are still nearly Euclidean: the large bias is actually due to the rapidly increasing lensing frequency with lookback time. In the standard treatment with non-evolving optical depth, a large magnification bias is typically associated with integral number counts that increase faster than f^{-2}, where f is the flux (e.g. Schneider, Ehlers & Falco 1992).

5.3.6 Microlensing

A modulation of the SN light curve may take place if, on top of the galaxy, group, or cluster magnification, a microlensing event occurs due to individual stars or MACHOs in the galaxy halo or intracluster medium. The timescale for microlensing-induced variations is determined by two competing effects: the expansion velocity of the SN photosphere and the relative transverse motion between the microlens and the light bundle. The radius of a SN at maximum light is $r_{\rm SN} \approx 10^{15}$ cm, corresponding to an observed angular dimension of $\theta_{\rm SN} = r_{\rm SN}/D_s$. A compact object along the line-of-sight to the SN will give rise to a detectable time dependent amplification only when $\theta_E \gg \theta_{\rm SN} \gtrsim |\theta|$ (Schneider & Wagoner 1987), where θ is the angular impact parameter. Modeling the deflector with a Schwarzschild lens of mass M (i.e. neglecting the shear induced by the gravitational field of the host halo), the first condition gives

$$M \gg M_E = 1.83 \times 10^{-4} \, M_\odot \, h \left(\frac{cD_l}{H_0 D_s D_{ls}} \right) . \tag{5.24}$$

[4] In the notation of Fukugita & Turner, our approach is equivalent to a flux-dependent lensing optical depth, as fainter objects are typically at higher redshifts than bright ones.

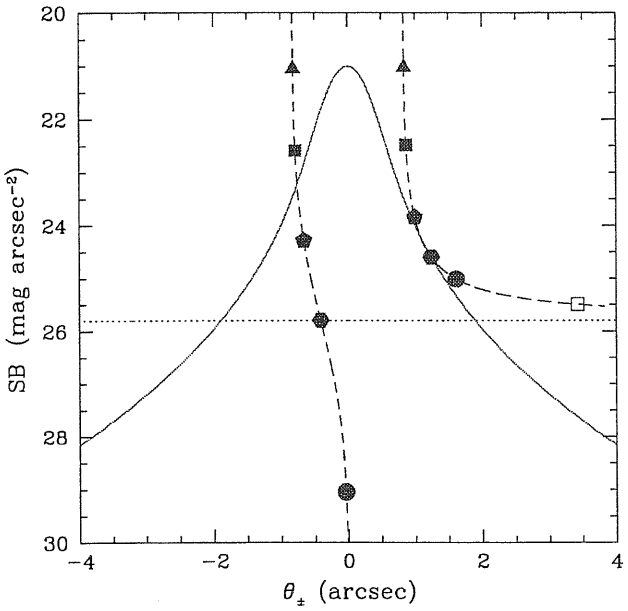


Figure 5.9: Observed I-band light profile for an elliptical galaxy at $z_l = 0.5$ (with typical luminosity and color, $M_B = -19.37 + 5\log h$, rest-frame $B - V = 0.95$), as a function of angular distance from the galaxy center (*solid line*). A de Vaucouleurs' law has been adopted with effective radius $R_e = 4$ kpc. The galaxy acts as a SIS lens with $\sigma_v = 225\ \mathrm{km\,s^{-1}}$ for a background Type Ia SN at $z_s = 1.5$ ($\theta_E = 0.83''$). *Dashed lines:* peak surface brightness of the dimmest ($\theta_- = \theta - \theta_E$, *left curve*) and brightest ($\theta_+ = \theta + \theta_E$, *right curve*) SN images for different source-lens alignments θ. The *filled points* along the curves identify multiple events corresponding to the same impact parameter. Only one image is produced at $\theta_+ > 2\theta_E$. The *empty square* corresponds to a single image with $\Delta m = 0.3$ mag. The light profiles of SN and galaxy have been convolved with a Gaussian point spread function of FWHM $0.6''$. *Dotted line:* SN brightness in the absence of lensing. A SCDM cosmology has been assumed.

Less massive MACHOs can efficiently magnify the SN only during its fast light-rising time. For $M \gg M_E$, the duration of the microlensing event roughly corresponds to the time needed by the photosphere to fill the Einstein ring of the lens (see also Gould 1994). For a typical photospheric expansion velocity of $\approx 10^9$ cm s^{-1}, one obtains

$$t_{\rm ph} \approx 27\ \mathrm{days}\ h^{-1/2} \left(\frac{H_0 D_s D_{ls}}{c D_l}\right)^{1/2} \left(\frac{M}{10^{-3} M_\odot}\right)^{1/2} (1+z_s)\ . \tag{5.25}$$

This is usually much shorter than the characteristic crossing time of the Einstein ring

$$t_E \approx 3.7\ \mathrm{yr}\ h^{-1/2} \left(\frac{H_0 D_l D_{ls}}{c D_s}\right)^{1/2} \left(\frac{M}{10^{-3} M_\odot}\right)^{1/2} \left(\frac{200\ \mathrm{km\,s^{-1}}}{v_\perp}\right) (1+z_l)\ , \tag{5.26}$$

where $v_\perp \approx \sqrt{2}\sigma_v$ is the relative transverse velocity between microlens and source. In a SCDM cosmology with $z_s = 2$ and $z_l = 0.5$, one finds $t_{\rm ph} \approx 49.2\ \mathrm{days}\,(M/10^{-3}M_\odot)^{1/2}$ and $t_E \approx 2.9\ \mathrm{yr}\,(M/10^{-3}M_\odot)^{1/2}(200\ \mathrm{km\,s^{-1}}/v)$. The microlensing timescale is then comparable with $t_{\rm SN} \approx 20(1+z_s)$ days (SN Ia) for very small microlens masses. Only MACHOS with $10^{-4} \lesssim M/M_\odot \lesssim$

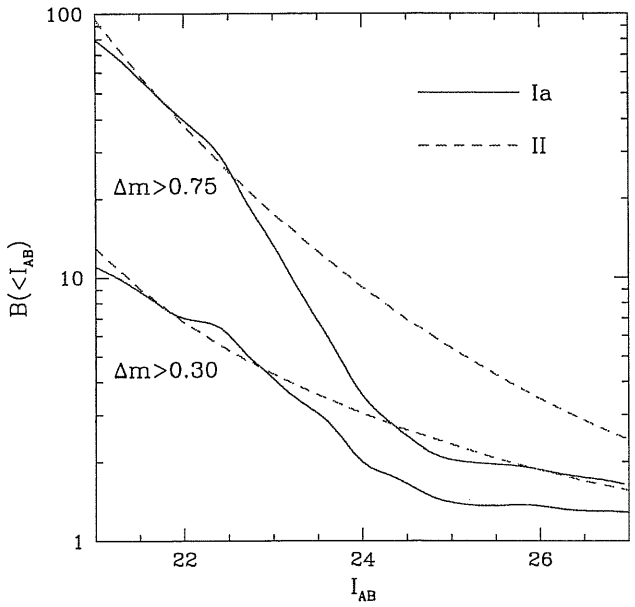


Figure 5.10: Magnification bias versus apparent limiting magnitude for Type Ia (*solid lines*) and Type II (*dashed lines*) SNe. The two set of curves have been computed for magnifications $\Delta m > 0.3$ and $\Delta m > 0.75$ mag. In the case of strong lensing, only the brightest image has been included in the counts.

10^{-3} can then produce detectable 'noise' that will degrade the light curves of the macroimages. More massive objects will boost the magnification of the SN image for the entire duration of the macroevent.

5.4 Summary and conclusions

All previous studies of the influence that a population of deflectors has on the observational properties of background objects have considered sources like galaxies or quasars that are quite inhomogeneous, i.e. have a wide range of intrinsic luminosities, ellipticities, etc. By contrast, Type Ia SNe represent a population of objects with well-known intrinsic properties, and could – in addition to their use as reliable cosmological distance indicators and valuable tracers of the star formation history in galaxies – potentially flag the occurrence of a gravitational lensing event even in the absence of multiple or highly distorted images. We have investigated the lensing effect of background SNe due to virialized dark matter halos in CDM cosmogonies, and computed lensing frequencies, rates of SN explosions, and distributions of arrival time differences and image separations. Within the assumption that dark halos approximate singular isothermal spheres, distributed in mass according to the Press-Schechter theory, about one every 20 Type Ia SNe at $z \approx 1$ will be magnified by $\Delta m \geq 0.1$. [5]

[5]An improved version of the model presented in this chapter is now in preparation. The main modifications we introduced concern the density profile and the mass function of the dark haloes. In particular, we replaced the SIS approximation with the Navarro-Frenk-White profile and the PS mass function with the fitting formula calibrated against N-body simulations by Sheth & Tormen (1998). These changes allow better agreement with the

Using recent estimates of the global history of star formation to compute the expected frequencies of SNe as a function of cosmic time, we have derived a detection rate of Type Ia's with magnification $\Delta m \geq 0.3$ mag of a few events $\mathrm{yr}^{-1}\ \mathrm{deg}^{-2}$ at maximum B-light and $I_{\mathrm{AB}} \leq 25$. Strong lensing events are 7 times less frequent. Only about one third of them give rise to observable multiple images: the SN will appear as multiple events with identical light curves, separated in time and differing only by the scaling of their amplitudes. Because of the flat K-correction and wide luminosity function, we find that Type II SNe will dominate the number counts at $I_{\mathrm{AB}} > 25$ and have the largest fraction of lensed objects. At bright magnitudes, the effect of magnification bias on the apparent frequency of strongly lensed SNe is huge, a factor of 40 in samples with $I_{\mathrm{AB}} \leq 22$. The enhancement beyond the number expected if each source represented a random line-of-sight drops to a factor of 3.5 at 24 mag (26 mag) in the case of Type Ia SNe (SN II). The apparent magnitude of the lens galaxy introduces a serious selection effect, as it reduces the detectability of small-separation multiple events in ground-based searches.

On the face of their rarity, it is not unconceivable that large consortia of new and existing SN search teams may conduct in the near future deep pencil beam surveys and be able to detect magnified Type Ia's up to $z \sim 2$ (Wang 1998). Current searches have been limited so far to $R \sim 24$ mag. A second-epoch deep observation of the *Hubble Deep Field* by Gilliland & Phillips (1998) has recently revealed two likely $z \sim 1$ SNe out to $I_{\mathrm{AB}} \sim 27$ mag. The combination of a new generation of large telescopes with 1 degree field of view and multi-fiber spectrometers, together with faster techniques such as the snapshot distance method (Riess et al. 1998), could soon yield more than ~ 500 new events per year. As suggested by Kolatt & Bartelmann (1998), one could also look for lensed SNe in the background of massive clusters. A nearby rich cluster at $z_l = 0.1$ with velocity dispersion $\sigma_v = 1000\ \mathrm{km\,s}^{-1}$ will have an Einstein radius for a source at $z_s \lesssim 2$ of about 25 arcsec. The expected rate of SNe Ia behind such a cluster with a mismatch of ≥ 0.3 mag is ~ 0.85 events per year with $I_{\mathrm{AB}} \leq 26$ (SCDM, $\tau = 1$ Gyr). Along the same lines, Kovner & Paczynski (1988) have proposed the possible identification of Type II SNe in the lensed giant arcs observed in the cores of clusters of galaxies. If a SN goes off within the caustic and close to the cusp of the cluster potential, it will be detected as three (or more) spatially-separated events with time delays as short as a few weeks or even days. The blue color of the arcs suggests significant star formation hence a high rate of SN explosions. Incidentally, if the cluster potential were understood well enough that a reasonable estimate of the time delay was available, one could exploit the occurrence of delayed mirror events to study such transient phenomena in much greater details that would be possible without prior warning.

Looking further into the future, it is interesting to consider the detectability of lensed Type Ia and Type II SNe with the *Next Generation Space Telescope* (NGST). A typical SN II at $z = 5$ would give rise to an observed flux of 8 nJy (SCDM) at 2.6 μm. At this wavelength, the imaging sensitivity of an 8m NGST is around 1.5 nJy (10^4 s exposure and 10σ detection threshold), while the moderate resolution ($\lambda/\Delta\lambda = 1000$) spectroscopic limit is about 50 times higher (10^5 s exposure per resolution element and 10σ detection threshold, Stockman et al. 1998). Depending on the history of star formation at high redshifts, the NGST could detect ~ 10 Type II SNe (at peak brightness) per $4' \times 4'$ field per year at $z > 4$ (MDP; cf. Dahlén & Fransson 1998; Miralda-Escudé & Rees 1997). (In principle, Type Ia's could also be imaged at extreme redshifts, $z > 10$, but the long delay times from stellar birth characteristic of these events make a detection rather improbable.) We find that, even for sources at these early epochs, the probability of strong lensing is small, $\mathcal{P} \approx 0.008$ for $z_s = 7$ and SCDM (filled beam). Uncertainties in the relation between angular diameter distance and redshift are significant at high-z; the corresponding value in the Ehlers-Schneider

observed distribution of angular separations for high-z quasars. In a future work we will also address the problem of substructure (e.g. galaxies) in cluster-sized haloes.

formalism is $\mathcal{P} \approx 0.02$. Marri & Ferrara (1998) have recently explored in detail the observational perspectives for Pop III SNe detection with the NGST when gravitational magnification is taken into account. They computed lensing frequencies and magnification maps using a ray-shooting numerical technique, and assuming point-like lenses. Our probabilities tend to be much lower than theirs, since isothermal spheres are inefficient magnifiers compared to point-like deflectors of the same mass.

Chapter 6

Biased galaxy formation

Astronomical observations are mainly concerned with the luminous parts of galaxies and clusters which represent only a small fraction of the total mass of the universe. On the other hand, the complexity of galaxy formation (for which we do not yet have a standard model) limits theoretical investigations to the analysis of the dark matter distribution. Therefore, to compare theory and observations and, say, give estimates for the cosmological parameters, it is first necessary to specify how mass and galaxies are related. The simplest hypothesis, widely adopted in the past, would be to assume that galaxies are faithful tracers of the underlying mass density field. One of the major advances of the comological research in the last 15 years was the realisation that this assumption needs not be true. In principle, it is well possible that galaxies may be distributed differently from the underlying mass density field, as a consequence, for example, of enviromental effects influencing the process of galaxy formation. This idea, which has come to be called biased galaxy formation, is also observationally supported: different classes of cosmic objects are in fact observed to cluster differently, implying that at least some of them do not directly trace the mass distribution. The aim of this chapter is to quickly review the history and the present status of the theory of biased galaxy formation.

6.1 Bias models

Biases are statistical differences arising when one compares different populations of objects. As an example, suppose that one uses a series of photographs of the Earth taken at night from a satellite to draw inference about the spatial distribution of emerged lands. Probably, the best thing to do would be to directly relate the positions of the bright spots in the picture (presumably generated by hurban agglomerates) with the continents, and the dark regions with the oceans. Results would be satisfying in densely populated and industrialized regions (e.g. Europe) but would give a biased description, say, of Australia, Antartica and Greenland which are mainly desert. Suppose that one instead uses the same picture to estimate the spatial distribution of the human population on Earth. The light density on the photographic plate would probably represent a good estimator and the differences between the two populations may well be much less than in the previous case. It is then clear that biases involve two factors: the selection of the sample population and the measurement to be performed.

In the cosmological case, the sample population (a set of galaxies) is selected by choosing a frequency band and a strategy to perform the observations. The measurement, instead, generally regards some statistical properties of the mass distribution on large length scales. Thus, the final question is: to which extent studies on the clustering and motions of galaxies furnish biased

information of the mass distribution on cosmological scales? Knowledge, at least approximate, of the biases would be fundamental:

- to reconstruct the mass power spectrum directly from the observed clustering statistics (e.g. Peacock & Dodds 1994; ; Peacock 1997; Mann, Peacock & Heavens 1998)
- to infer the value of Ω_0 either from the analysis of the peculiar velocity field (e.g. Dekel 1994; Dekel & Lahav 1998) or via redshift distortions (e.g. Kaiser 1987; Hamilton 1992, 1997; Fisher, Sharf & Lahav 1994; Fisher & Nusser 1996);
- to compare the predictions of cosmological models with the observed clustering properties of high-z objects (e.g. Matarrese et al. 1997; Moscardini et al. 1998; Bagla 1998a,b; Coles et al. 1998; Cress & Kamionkowski 1998; Arnouts et al. 1999; Magliocchetti & Maddox 1999; Magliocchetti et al. 1999; Mo, Mao & White 1999).

Unfortunately, at present we know little about the details of star and galaxy formation. Hence the general attitude is to adopt a phenomenological approach to the problem of galaxy bias. Simple "toy" models, often dealing with the distribution of dark matter haloes (hopefully hosting the luminous matter), are developed to understand what observational tests are sensitive to the biasing effect. Observational data are then used to quantify the galaxy bias by constraining the free parameters of the models.

Essentially a bias scheme is a rule to compute the galaxy number density field $\rho_g(\mathbf{x},t)$ [or its fluctuation with respect to the mean value $\delta_g(\mathbf{x},t)$] given a cosmological model and its mass distribution $\rho(\mathbf{x},t)$. [1] In most cases, $\rho_g(\mathbf{x},t)$ denotes a continuous field: galaxy positions (a point process) can be recovered, for example, through Poisson sampling. Denoting the mean number density of galaxies by n_g, this is obtained by associating to each infinitesimal volume d^3x centred on $\mathbf{x}$ the probability $dP_g(\mathbf{x}) = n_g\rho_g(\mathbf{x},t)\,d^3x/\langle\rho_g\rangle$ to contain a galaxy (Layzer 1956; Peebles 1980; Coles 1993). [2] Notice that sampling probabilities at different points are assumed to be independent (although the density field may be correlated). In this way, the galaxy correlations at every order come out identical to the correlations of ρ_g; in other words, Poisson sampling generates a discrete distribution that "traces" ρ_g. Note, however, that this procedure to account for galaxy discreteness is a model and may not hold in practice (see e.g. Williams, Heavens & Peacock 1991 for a different discretization procedure). Since comparison with observations often requires some smoothing of the distributions, the continuous character of ρ_g does not represent a problem in every practical application.

The simplest parametrization of a bias model is obtained by introducing one of the following functions

$$b_\sigma^2(R_f,t) \equiv \frac{\sigma_g^2(t,R_f)}{\sigma^2(t,R_f)}\,, \qquad b_\xi^2(r,t) \equiv \frac{\xi_g(r,t)}{\xi(r,t)}\,, \qquad b_P^2(k,t) \equiv \frac{P_g(k,t)}{P(k,t)}\,, \tag{6.1}$$

where the subscript g denotes quantities referring to the galaxy distribution. While b_σ is a measure of the bias in galaxy counts, b_ξ and b_P compare the two-point properties of mass and galaxies. Many other parametrizations for the bias can be given, each one leading to different values and scale dependences (e.g. Coles 1993). In general, on very large scales (where $\sigma(R_f,t) \ll 1$ or $\xi(r,t) \ll 1$), $b_\sigma \simeq b_\xi = \text{const}$ and one can speak about a single bias parameter, b.

[1] In this chapter, with the exception of §6.6.1, we use the word galaxy in a rather liberal way, meaning a selected class of astronomical objects. Note that some authors prefer to define δ_g as the field of ligth-density fluctuations, whose time evolution is less sensitive to galaxy mergers (e.g. Coles 1993; Dekel & Lahav 1998).

[2] In many bias models the overall normalization of ρ_g is not specified. This explains why n_g and $\langle\rho_g\rangle$ denote different quantities in the main text.

6.2 A bit of history

One of the first observational evidences for the existence of biasing came from comparing the two-point correlation functions of galaxies (ξ_g) and rich Abell clusters (ξ_{cl}). Although both correlation functions have the same power law slope with exponent -1.8, ξ_{cl} has an amplitude about 20 times larger than ξ_g. Furthermore the cluster correlation length increases when richer and richer objects are considered. Kaiser (1984) suggested that the origin of this enhancement is essentially statistical. His argument develops as follows. Abell clusters are the most massive objects which have collapsed and relaxed to a roughly spherical shape by the present time. They are also extremely rare, as their number density is orders of magnitude smaller than the number density of galaxies. In the gravitational instability scenario, at a fixed mass scale, the first objects to collapse are the densest. It is therefore reasonable to assume that clusters form where the linear mass density contrast δ_ℓ, when averaged over their mass scale (corresponding to $R_{cl} \sim 10\,h^{-1}$Mpc), lay above some relatively high threshold value δ_{cl}. Even though peculiar motions may have rearranged the clusters on small scales ($\lesssim 10\,h^{-1}$Mpc, the distance covered by an object with a peculiar velocity of $1000\,\mathrm{km\,s^{-1}}$ in a Hubble time), the large-scale correlation function should still approximate that of the "peaks" of the primordial density field. It can be shown that objects selected in this way give a biased estimate of the large-scale density correlations; in particular, for large separations and when δ_{cl} is much larger than the rms mass fluctuation on the cluster scale, one finds that there is a multiplicative amplification of the correlations of peaks relative to those of the underlying matter distribution, $\xi_{cl} \simeq b^2\xi$, with $b \gg 1$ (Kaiser 1984; see §6.3.1 for mathematical details). This is because the probability of finding a high value of δ_ℓ is enhanced in regions where there is a positive fluctuation on scale much larger than the smoothing radius.

The idea of the statistical biasing was readily accepted by the scientific community, for it could also help to reconcile the observational estimate $\Omega_0 \sim 0.2$ (obtained assuming that the cluster mass to light ratio is universal) with the theoretically favoured value $\Omega_0 = 1$ stemming from inflationary models (see e.g. Dekel & Rees 1987 for a review). [3]

It was soon recognized that the same argument of threshold biasing could also be applied to the formation of galaxies. The seminal work by Dressler (1980; see also Postman & Geller 1984) presented strong evidence that galaxy morphological type correlates with the local galaxy density. For instance, ellipticals are most frequently found in clusters of galaxies while spirals favour lower density regions. This supported the idea about the presence of effective environmental effects during the process of galaxy formation. Moreover, from large scale observations, it was becoming clear that vast regions depleted of galaxies (voids) are quite common; this picture could be more easily reconciled with the expected amplitude of density fluctuations at recombination (estimated from observational upper limits on CMB anisotropies) if a biasing mechanism segregated dark and luminous matter (Dekel & Rees 1987).

The realization of the first N-body simulations with a CDM spectrum (Davis et al. 1985) gave further motivation for galaxy biasing. Since the CDM spectrum curves slowly between the effective indices of $n = -3$ and $n = 1$, the mass correlation function clearly steepens with time (see §3.3). There is therefore a unique epoch when ξ will have the observed slope $\gamma = -1.8$. Davis et al. (1985) identified this epoch as the present, and then noted that it implied a rather low correlation length, $r_0 = 1.3\,h^{-2}$ Mpc, with respect to the observed value for galaxies $r_{0g} \sim 5\,h^{-1}$ Mpc. What seemed to be required was a galaxy correlation function which was an amplified version of that for the

[3] Note that biasing is naturally realized in the presence of a non-clustering component like the cosmological constant or a sea of ultra-hot weakly interacting particles formed, for example, from non-radiative decay of some heavy dark matter component.

mass by a factor $\sim 2.4^2$. [4] The analogy with the argument by Kaiser was readily done, and the CDM model adopted the idea that the sites of galaxy formation might in some way be associated with the peaks of a suitably smoothed version of the linear density field. Note that, even though they are described through the same mathematical model, cluster and galaxy biasing must be given different physical interpretations. Biasing of clusters is purely statistical and follows directly from their rareness, that they correspond to (3-4)σ objects: cluster-scale peaks with lower overdensities would not have collapsed by now. However, they will become non-linear in the future and the bias will be progressively reduced. This is not the case for galaxies. The linear overdensity on galaxy scales is $\sigma_\ell(R_g \simeq 1\,h^{-1}\mathrm{Mpc}) \sim 3$, hence even 1σ regions would have already collapsed. In this case, it is the local physics that inhibits galaxy formation in the lowest peaks.

A physical motivation was then needed to explain why baryons may have condensed more efficiently into galaxies in regions of above-average density. The bias could have been determined in each protogalaxy autonomously, for example by its binding energy or local density (e.g. Dekel & Silk 1986), or it may have been a result of feedback from other galaxies. For instance, Silk (1983), Rees (1985) and Dekel & Rees (1987) proposed some mechanisms that could explain how energy from early galaxies could prevent or foster the condensation of further galaxies. This influence may propagate by gas transport to limited distances, or by radiation and fast particles to larger distances.

Many observational results gave additional support to the idea of galaxy biasing. For example, the growing evidence that galaxies of different luminosities and morphological types (Loveday et al. 1992; Willmer, Da Costa & Pellegrini 1998; Tegmark & Bromley 1999) or selected at different wavelength (Saunders, Rowan-Robinson & Lawrence 1992; Willmer, Da Costa & Pellegrini 1998; see also the compilation by Peacock & Dodds 1994) cluster differently showed that the correlation statistics of all types of galaxy cannot separately parallel those of the mass distribution. It has been also realized that the motivations which led to the concept of galaxy biasing do not single out any specific form for it. The threshold model was then just considered as a particularly plausible and convenient biasing prescription, and new, equally viable, schemes have been proposed. Their main features are described in the next two sections. We stress once again that such statistical models are overly simplistic. Galaxies are not "painted" onto a background field; they form through non-linear gravitational and gas dynamic processes.

6.3 Lagrangian bias schemes

A bias model is called Lagrangian if it computes the galaxy density field at the Lagrangian position $\mathbf{q}$ by using only the information encoded in the linear mass density field $\rho_\ell(\mathbf{q})$, i.e.

$$\rho_g(\mathbf{q}, t) = \mathcal{F}[\rho_\ell(\mathbf{q})] \tag{6.2}$$

where $\mathcal{F}$ denotes an arbitrary functional that maps functions onto functions. In this case the bias parameters introduced in eq. (6.1) are obviously computed with respect to the linear mass variance, autocorrelation function and power spectrum.

Linear bias

The simplest model is obtained by assuming $\delta_g(\mathbf{q}, t; R_f) = b(R_f)D_+(t)\delta_\ell(\mathbf{q}; R_f)$ where, at fixed R_f, $b(R_f)$ is a constant, called the linear bias parameter. Since both δ_g and δ must assume values larger

[4] The opposite problem appears in HDM models: when $\gamma = -1.8$, the HDM correlation length has increased to $5(\Omega_0 h^2)^{-1}$ Mpc, which is large in comparison with r_{0g} unless $\Omega_0 h \geq 1$ (Centrella & Melott 1983; White, Frenk & Davis 1983). One would then need anti-bias, i.e. galaxies should somehow be less clustered than the dark matter.

than, or equal to -1, this model is acceptable only for large smoothing radii, when $\sigma_g = b^2\sigma \ll 1$. Note, however, that, even in this limit, linear biasing cannot be exact since one expects $\delta_g = -1$ where $\delta = -1$. Thus, in the most general case, non-linear biasing, where b varies with δ_ℓ, is inevitable.

Threshold bias

As previously discussed, Kaiser (1984) proposed the idea that "galaxy" formation might be associated with those regions in which the smoothed density field exceeds a given threshold value. From the mathematical point of view, the algorithm is embodied in the relation

$$\rho_g(\mathbf{q}, t) \propto \Theta[\delta_\ell(\mathbf{q}; R_f) - \nu(t)\sigma_\ell(R_f)] \tag{6.3}$$

where $\nu(t) \equiv \delta_g(t)/\sigma_\ell$, the value of the threshold in units of $\sigma_\ell(R_f)$, and R_f are both spatially invariant. In this case, it can be shown that, if the local PDF of δ_ℓ has a steeply decreasing tail for $\delta_\ell \gg \sigma_\ell$, the galaxy field displays enhanced clustering over the density contrast of the underlying matter distribution. For example, considering a Gaussian density field, for $\nu \gg 1$ and $\xi_\ell(\mathbf{q}_i - \mathbf{q}_j) \ll \sigma_\ell$, one obtains (Politzer & Wise 1984; Jensen & Szalay 1986)

$$\xi_{n,g}(\mathbf{q}_1, \ldots, \mathbf{q}_n) \simeq \exp\left[\sum_{i>j} \frac{\nu^2}{\sigma_\ell^2}\xi_\ell(\mathbf{q}_i - \mathbf{q}_j)\right] - 1 \,, \tag{6.4}$$

where $\xi_{n,g}$ denotes the galaxy n-point irreducible correlation function and the dependences on the smoothing radius and time are understood. The expression for the two-point function then is (Kaiser 1984; Politzer & Wise 1984; Jensen & Szalay 1986)

$$\xi_g(r) = \exp\left[\frac{\nu^2}{\sigma_\ell^2}\xi_\ell(r)\right] - 1 \simeq \left(\frac{\nu^2}{\sigma_\ell^2}\right)\xi_\ell(r) \,, \tag{6.5}$$

where the last equality holds for $\xi_g \ll 1$. Detailed numerical analyses showed that, far from this asymptotic regime, the biasing properties deriving from the Kaiser model are much more complex and strongly scale dependent (e.g. Coles & Davies 1993).

Local Lagrangian schemes

The argument by Kaiser can be generalized by replacing the sharp Heaviside threshold in eq. (6.3) with a different smooth function. For example, Kaiser & Davis (1985) considered an exponential bias $\rho_g \propto \exp(\nu\delta_\ell/\sigma_\ell)$ and showed that the asymptotic behaviour of the galaxy two-point correlation function is identical to the sharp threshold case. Alternatively, Fry (1986) proposed a power law biasing scheme $\rho_g \propto (1 + \delta_\ell/\sigma_\ell)^\alpha$ which enhances the non-Gaussian features of the galaxy distribution. We can then define the subclass of local Lagrangian bias models containing those schemes in which $\rho_g(\mathbf{q}, t)$ is fully determined by a single number: the value of the mass density field linearly extrapolated at the time t and evaluated at the same Lagrangian position $\mathbf{q}$,

$$\rho_g(\mathbf{q}, t) = f[\rho_\ell(\mathbf{q})]; \,, \tag{6.6}$$

with f an arbitrary function, the principal features of which are the suppression of galaxy counts in low-density regions and their enhancement in regions of high density. Note that, barring trivial transformations (e.g. linear bias), even local algorithms do not preserve the correlation statistics of the density field. It can be shown that, for Gaussian initial conditions, local Lagrangian bias

models always give $\xi_g \propto \xi_\ell^n$ with $n \geq 1$ to leading order as $\xi_\ell \to 0$. Then, as far as ξ_ℓ decreases monotonically with r (i.e. for separations smaller than its its zero-crossing length), b_ξ must be a non-increasing function of scale (Coles 1993). A similar theorem does not hold for b_P since a local biasing scheme in configuration space is extremely non-local in Fourier space. In general, when $\rho_g \propto f(\delta_\ell/\sigma_\ell)$ and the initial conditions are Gaussian, galaxy correlations are fully determined by the Hermite coefficients of the non-linear threshold functions $f(y) = \sum_{k=0}^{\infty}(J_k/k!)H_k(y)$, with $H_n(x) = e^{x^2/2}(-d/dx)^n e^{-x^2/2}$ (Szalay 1988). In particular, one gets

$$1 + \xi_g = \sum_{k=0}^{\infty} \simeq \frac{J_k^2}{k!}\xi^n \simeq 1 + J_1^2\xi \ , \tag{6.7}$$

where the last equality holds in the linear regime, i.e. for small correlations. In the same limit, the galaxy three-point correlation is

$$\xi_{3,g} = Q(\xi_{12}\xi_{23} + \xi_{13}\xi_{23} + \xi_{13}\xi_{12}) + Q^3\xi_{12}\xi_{23}\xi_{31} + \frac{1}{2}Q(T - Q^2)[\xi_{12}^2(\xi_{23} + \xi_{31}) + \text{cyclic}] \ , \tag{6.8}$$

with $Q = J_2/J_1^2$ and $T = J_3/J_1^3$. Sharp threshold and exponential bias models, having $Q = T = 1$, provide the simplest possible three-point disconnected correlation function $\langle\delta_g\delta_g\delta_g\rangle = (1+\xi_{12})(1+\xi_{23})(1+\xi_{31})$, called the Kirkwood superposition (e.g. Matarrese, Lucchin & Bonometto 1986; as shown by Coles and Davies 1993, this expansion is valid when the correlations are extremely small and practically there is no difference between Kirkwood and hierarchical scalings; for smaller triangles, a hierarchical $\xi_{3,g}$ gives a better, though not perfect, fit to the correlations obtained using a sharp threshold). One could then think about reconstructing the dominant shape of the threshold function by determining the Q and T coefficients from observed data (Szalay 1988). Because of the Gaussian assumption for δ, clusters of galaxies might be more suitable than galaxies (since they form from the collapse of larger perturbations). However,in the gravitational instability scenario, we expect non-Gaussian effects to become soon important (Fry 1986) and more refined analyses are required to determine the amplitude of the biasing effect from astronomical observations (see e.g. §6.6).

Bias of density maxima

The threshold model developed by Kaiser to explain the strong correlations of Abell clusters was originally formulated in terms of density maxima. The idea was to associate proto-objects with the maxima of the (suitably smoothed) linear density field, provided their height exceeded a given threshold value $\nu\sigma$. The tremendous complexity involved in the analytical derivation of the corresponding correlations, led Kaiser to simplify the problem by considering regions lying above the threshold. The statistical properties of density maxima in a three-dimensional Gaussian field have been studied by Bardeen et al. (1986; see also Peacock & Heavens 1985 and Bond & Efstathiou 1987), who followed the methods described in the seminal works by Rice (1945; reprinted in Wax 1954), Cartwright & Longuet-Higgins (1956) and Longuet-Higgins (1957). In general, in 3-dimensions, the high-peak ($\nu \gg 1$) probability density comes out as

$$N_{\max}(>\nu) \simeq \frac{1}{(2\pi)^2}\left(\frac{\gamma}{R_\delta}\right)^3 (\nu^2 - 1)\exp\left(-\nu^2/2\right) \ , \tag{6.9}$$

where the quantities R_δ and γ are simple functions of the spectral moments of δ and are defined in Appendix B.

Density maxima have some inbuilt clustering as a result of the statistics of the linear field: a peak form because of constructive interference of waves with wavelengths near R_g in its neighbourhood,

and the main role of long waves is to modulate the height of the peaks, which creates biasing. The effect is easy to work out, using the so-called peak-background split (Bardeen et al. 1986). Here one decomposes the density field into two parts $\delta_\ell = \delta_{\rm pk} + \delta_{\rm bg}$, where $\delta_{\rm pk}$ contains all the high-frequency modes and $\delta_{\rm bg}$ is generated by the Fourier terms with much longer wavelength. Since we are interested in a global threshold, the large scale modulation given by $\delta_{\rm bg}$ will drive some peaks of $\delta_{\rm pk}$ just below threshold above it and some above it below. This is more or less equivalent to identifying galaxies with the maxima in the peak field such that $\delta_{\rm pk} > \nu\sigma_\ell - \delta_{\rm bg}$. This varying effective threshold will now produce more peaks in the regions of high $\delta_{\rm bg}$, leading to amplification of the clustering pattern. For high peaks in 3 dimensions, using eq. (6.9) one gets $\delta N_{\rm max}/N_{\rm max} \simeq (\nu/\sigma_\ell)\delta_{\rm bg}$ in agreement with eq. (6.5). This result can be qualitatively understood as follows: on the rare high ν side, there are few maxima above threshold and many below, hence a little positive modulation will result in many peaks above threshold and a little negative modulation will result in a severe paucity. In proto-voids, which have a large scale underdensity, there will be dramatically fewer galaxy-scale peaks than in protocluster regions.

The exact solution, in the large separation limit, gives a similar expression $\xi_{\rm max}(r) \simeq \langle(\nu_{\rm eff}\rangle/\sigma_\ell)^2$ $\xi(r)$ (Bardeen et al. 1986). It is worth stressing that, for realistic values of ν (2-3, say), the effective threshold $\langle\nu_{\rm eff}\rangle < \nu$. When we consider all peaks, i.e let $\nu \to -\infty$, the effective threshold $\langle\nu_{\rm eff}\rangle \to 0$. This is because density maxima are determined by the small-scale structure in the power spectrum and not by any additional large scale power.

Natural bias

Non-linear gravitational evolution itself can generate a bias of the dense clumps formed by gravitational instability relative to the mass. In hierarchical models, such biasing occurs because formation of the most (least) massive haloes is enhanced (depressed) in regions of high initial density, which collapse earlier and accrete faster than their surroundings (Dekel & Rees 1987; White et al. 1987; Efstathiou et al. 1988). The importance of this effect can be calculated approximately using the Press-Schechter formalism (Efstathiou et al. 1988; Cole & Kaiser 1989; Kashlinsky 1991; Mo & White 1996; Mo, Jing & White 1997; Catelan et al. 1998; Porciani et al. 1998; Catelan, Matarrese & Porciani 1998). The basic idea is to use the peak-background split that we already introduced to discuss the biasing properties of the density maxima. In full analogy with the peak case, a non-vanishing background field is equivalent to a modulation of the threshold for structure formation (i.e. t_f must be replaced with $t_f - \delta_{\rm bg}$). In other words, $\delta_{\rm bg}$ accelerates or retards the halo collapse at a given location depending on its sign. The corresponding bias can then be quantified by the ratio between the induced halo overdensity and $\delta_{\rm bg}$ (Efstathiou et al. 1988; Cole & Kaiser 1989)

$$b_{\rm ps}(M, z_f) \equiv \left.\frac{\log n(M, z_f)}{\delta_{\rm bg}}\right|_{\delta_{\rm bg}=0} = -\frac{\partial \log n(M, z_f)}{\partial t_f} = \frac{t_f}{\Lambda(M)} - \frac{1}{t_f} \ , \ . \tag{6.10}$$

which reduces to the standard form ν/σ_ℓ in the limit of high thresholds when the term $1/t_f$ can be neglected. Thus, haloes with $M > M_*$ have $b_{\rm ps} > 0$, i.e. are biased with respect to the mass distribution in Lagrangian space. Notice that $b_{\rm ps}$ can be very large when $M \gg M_*$. On the contrary, objects with $M < M_*$ have $-1/t_f < b_{\rm ps} < 0$, i.e. are moderately antibiased.

Mo & White (1996) and Mo, Jing & White (1997) used the excursion set formalism to compute the Lagrangian bias factors for dark matter haloes at any order. The average number of haloes with mass M identified at redshift z_f that are contained in a large spherical region with Lagrangian radius R_L (corresponding to the mass $M_L \gg M$ and the linear variance Λ_L) and overdensity $\delta_L < t_f$

is

$$\mathcal{N}(M, z_f|R_L, \delta_L)\, dM \equiv \frac{M_L}{M}\mathcal{P}_1(\Lambda - \Lambda_L, \delta - \delta_L)\frac{d\Lambda}{dM}\, dM\ , \tag{6.11}$$

where $\mathcal{P}_1$ is the probability given in eq. (4.27) and top-hat smoothing is used to compute the function $\Lambda(M)$. The corresponding average overdensity of haloes in the "background sphere" with volume $V_L = 4\pi R_0^3/3$ is then given by

$$\delta_h^L(M, z_f|R_L, \delta_L) \equiv \frac{\mathcal{N}(M, z_f|M_L, \delta_L)}{n(M, z_f)V_L} - 1\ . \tag{6.12}$$

and, in the limit $|\delta_L| \ll t_f$, one can write

$$\delta_h^L(M, z_f|R_L, \delta_L) = \sum_{k=0}^{\infty} \frac{b_k^L}{k!}\delta_L^k\ . \tag{6.13}$$

Let us explicitly write the first five Lagrangian bias factors (we assume $\Lambda_L \ll \Lambda$)

$$\begin{aligned}
b_0^L(\Lambda) &= 0\ , \\
b_1^L(\Lambda) &= \frac{t_f}{\Lambda} - \frac{1}{t_f}\ , \\
b_2^L(\Lambda) &= \frac{t_f^2}{\Lambda^2} - \frac{3}{\Lambda}\ , \\
b_3^L(\Lambda) &= \frac{t_f^3}{\Lambda^3} - \frac{6t_f}{\Lambda^2} + \frac{3}{t_f\Lambda}\ , \\
b_4^L(\Lambda) &= \frac{t_f^4}{\Lambda^4} - \frac{10t_f^2}{\Lambda^3} + \frac{15}{\Lambda^2}\ .
\end{aligned} \tag{6.14}$$

The linear bias term, which for $M \neq M_*$ dominates the halo correlation at large separation, coincides with $b_{\rm ps}$. Notice that both $b_{\rm ps}$ and the b_k^L are mean quantities over the ensemble and cannot be used to compute halo correlations. A careful analysis of the correlation functions in the context of the excursion set theory have been performed by Porciani et al. (1998; see also Catelan et al. 1997). Even though we already presented their results in chapter 4, it is convenient here to re-discuss them in terms of bias coefficients. For instance, at large separations, the halo two-point function given in eq. (4.40) can be expanded in powers of the filtered mass auto-correlation function $\xi(r; \Lambda_{\rm m}) \ll 1$,

$$\xi^{obj}(r) \equiv \xi^{pts}(r) = \sum_{n=1}^{\infty} \frac{1}{n!}\, b_n^L(\Lambda_1)\, b_n^L(\Lambda_2)\, \xi^n(r; \Lambda_{\rm m})\ , \tag{6.15}$$

where the factors $b_n(\Lambda)$ coincide with the Lagrangian bias coefficients introduced in eq. (6.13). Notice that, in the limiting case $M = M_*$, $b_1 = 0$ and the leading term of ξ^{hh} is proportional to $\xi(r; \Lambda_{\rm m})^2$, implying much lower halo correlations compared with different mass ranges. In the next chapter we will discuss other approximated models for the two-point correlations of dark matter haloes that have been suggested by Catelan et al. (1997). Also these schemes give the same series of Lagrangian bias factors of eq. (6.15) provided the lag is a few times larger than the Lagrangian halo size.

6.4 Eulerian bias schemes

The Lagrangian models we discussed so far have been extremely useful to understand how a general biasing scheme can work. However, they cannot be adopted for most practical applications dealing with observational data. In fact, the maps resulting from these schemes are only accurate at extremely small resolution ($R_f \gg 10\,h^{-1}$Mpc) where perturbations are in the linear regime. For smaller smoothing lengths, a fully Eulerian description, including non-linear dynamical evolution, is then needed to quantify galaxy biasing.

We term Eulerian those biasing models that output the galaxy distribution in configuration space at a given instant t. These can be divided into two main classes:

- The first set is constituted by models in which the galaxy distribution is related to the Eulerian mass density field $\delta(\mathbf{x},t)$ irrespective of the initial conditions of the regions in which galaxies formed, i.e.

$$\rho_g(\mathbf{x},t) = \mathcal{F}_{\mathcal{E}}[\rho(\mathbf{x},t)]\,, \tag{6.16}$$

  with $\mathcal{F}_{\mathcal{E}}$ an arbitrary functional. This is the direct generalization of eq. (??) to the Eulerian world. All the Lagrangian biasing schemes we described above can be easily adapted to the Eulerian case. For instance, it is possible to associate galaxies to the highest peaks of the density distribution or use a local transformation of $\delta(\mathbf{x},t)$ to build $\delta_g(\mathbf{x},t)$. The main difficulty is to deal with the non-Gaussian PDF of the evolved density field. Often, the lognormal model by Coles & Jones (1991) is used for simple analytical computations. For example, in the lognormal case, power law biasing (i.e. $\rho_g = \rho^\alpha$) would give $1+\xi_g = (1+\xi)^{\alpha^2}$ (Mann, Peacock & Heavens 1998). However, N-body simulations do not confirm this simple scaling relation suggesting, instead, that $[1+k^3P_g(k) \simeq [1+\alpha^{1.8}k^3P(k)]^{0.7}$ (Mann, Peacock & Heavens 1998). Note that, analytic and numerical results do not agree even in the asymptotic large scale regime where they give $b_P = 1$ and $b_P \simeq 1.3$, respectively. This shows that the lognormal model is only a rough approximation and accurate numerical investigation is needed in the general case.

- Alternatively, one can select in Lagrangian space the patches of fluid that will form galaxies and then follow their Eulerian evolution. For instance, the present-day clustering of Lagrangian density maxima differs substantially from the original one because of gravitationally induced motions and merging processes. An interesting approach is to use the Zel'dovich approximation to map smoothed regions in Lagrangian space to their Eulerian positions (Bond & Couchman ; Bond & Myers 1996; Catelan et al. 1997; Catelan, Matarrese & Porciani 1998).

Both approaches represent extreme point of views. For example, models of the second type describe a sort of infinite memory process, since the sites for galaxy formation are known from the beginning and evolution simply changes their spatial distribution. On the other hand, in biasing algorithms belonging to the first class, there is no heritage from the past and galaxies are simply paint on a snapshot of the density field. However, even though the real galaxy formation is probably a process with intermediate characteristics with respect to the two classes of biasing schemes presented above, recent models based on a Lagrangian selection of the sites for object formation were shown to be extremely succesful in reproducing the clustering properties of dark matter haloes found in numerical simulations (Mo & White 1996; Mo, Jing & White 1997; Catelan et al. 1998; Catelan, Matarrese & Porciani 1998; Jing 1998; Sheth & Lemson 1998; Porciani, Catelan & Lacey 1999; Sheth & Tormen 1999; Jing 1999; Mo, Sheth & Tormen in preparation; Matarrese, Porciani & Catelan in preparation). Detailed discussions of these results are given in §6.4.4 and in the next chapter.

6.4.1 Local Eulerian biasing

One can assume that the (smoothed) galaxy density at a given point $\mathbf{x}$ and time t is a local function of the matter density evaluated at the same spatial location and at the same time. If $f(\delta)$ admits a Taylor expansion around $\delta = 0$, it is possible to define the Eulerian bias parameters b_k at any order k through

$$\delta_g(\mathbf{x}, t; R_f) = f_E[\delta(\mathbf{x}, t; R_f)] = \sum_{k=0}^{\infty} \frac{b_k}{k!}\delta^k \,, \tag{6.17}$$

where $b_0 = -\sum_{k=2}^{\infty} b_k \langle \delta^k \rangle / k!$ to have $\langle \delta_g \rangle = 0$, and the constraint $\delta_g(\delta = -1) = \sum_{k=0}^{\infty} f b_k k!(-1)^k = -1$ provides a second relation between the parameters. The numerical analysis by Mann, Peacock & Heavens (1998) showed that, commonly adopted local Eulerian bias models produce a scale dependent b_P which is a monotonically decreasing function of scale tending to an asymptotic value on scales where the density fluctuations are linear.

Fry & Gaztañaga (1993) showed that if the volume-averaged $j-$point mass correlation function $\bar{\xi}_j(R_f)$ have the hierarchical form $\bar{\xi}_j(R_f) = S_j \bar{\xi}_2^{j-1}(R_f)$, then

$$\bar{\xi}_{g,2} = b_1^2 \bar{\xi}_2 + \left(b_1 b_2 S_3 + b_1 b_3 + \frac{b_2^2}{2} \right) \bar{\xi}_2^2 + \mathcal{O}(\bar{\xi}_2^3) \,, \tag{6.18}$$

$$\begin{aligned} \bar{\xi}_{g,3} = (b_1^3 S_3 + 3 b_1^2 b_2)\bar{\xi}_2^2 + \Bigg(& \frac{3}{2} b_1^2 b_2 S_4 + \frac{9}{2} b_1^2 b_3 S_3 + 6 b_1 b_2^2 S_3 + \frac{3}{4} b_1^2 b_4 + \\ & + 6 b_1 b_2 b_3 + b_2^3 \Bigg) \bar{\xi}_2^3 + \mathcal{O}(\bar{\xi}_2^4) \,. \end{aligned} \tag{6.19}$$

To leading order in $\bar{\xi}_2$ one then gets (Fry & Gaztañaga 1993)

$$\bar{\xi}_{j,g} = S_{j,g} \bar{\xi}_{j,g}^{j-1}(R_f) \,, \tag{6.20}$$

i.e. the transformation given by equation (6.17) preserves the hierarchical structure in the limit $\bar{\xi}(R_f) \ll 1$. In particular, for $j = 2, 3$ and 4 one has

$$S_{g,2} = b_1^2 \,, \quad S_{g,3} = \left(S_3 + 3\frac{b_2}{b_1} \right) \frac{1}{b_1} \,, \quad S_{g,4} = \left[S_4 + 12 \frac{b_2}{b_1} S_3 + 4 \frac{b_3}{b_1} + 12 \left(\frac{b_2}{b_1} \right)^2 \right] \,. \tag{6.21}$$

Thus, any local (non-linear) biasing scheme gives a contribution to the connected moments of the galaxy distribution which is comparable to that deriving from non-linear gravitational evolution. This is because both biasing and gravitational instability skew the density field away from its initial Gaussian symmetry. It is then interesting to try to devise some methods to distinguish between the two effects when dealing with observational data. If the large scale structure formed via gravitational instability from Gaussian initial conditions, then measuring the hierarchical amplitudes $S_{j,g}$ and their redshift evolution can reveal a clear signature of biasing (Fry & Gaztañaga 1993; Fry 1994). This is because perturbation theory makes precise predictions for the unbiased S_j in the quasi-linear regime (see chapter 2). In a similar fashion, the shape dependence of the bispectrum and three-point correlation function in leading-order perturbation theory can be used for the same purposes (Fry & Gaztañaga 1993; Fry 1994; Frieman & Gaztañaga 1994). For example, the galaxy bispectrum to leading order in pertubation theory is

$$B_g^{(0)}(k_1, k_2, k_3, t) = [2 b_1^3 F_2^{(S)}(\mathbf{k}_1, \mathbf{k}_2) + b_1^2 b_2] P_g^{(0)}(k_1) P_g^{(0)}(k_2) + \text{cyclicterms} \,, \tag{6.22}$$

where, to the same accuracy, $P_g^0(k) = b_1^2 P^{(0)}(k)$. Since $F_2^{(S)}$ is almost independent on Ω_0 and λ, a set of triangles with different shapes can be used to obtain multiple, independent constraints on the combination of the constant b_1 and b_2 parameters (Matarrese, Verde & Heavens 1997). The ultimate goal is to combine measurements of b_1 with estimates of $\beta = \Omega_0^{0.6}/b_1$ either from linear redshift distortions studies or from cosmic flows analyses to get an estimate of the density parameter Ω_0. Similar results can be obtained from the three-point correlation function in real space. Other factors, like redshift distortions, boundary effects, shot-noise, and a possible scatter in the biasing relation (see §6.6) complicate the comparison of these predictions with observations. It is anyway expected that the Sloan digital sky survey (SDSS) and the Anglo-Australian 2 degree-field survey (2dF) will permit to constrain b_1 and b_2 to better than 10% accuracy (Matarrese, Verde & Heavens 1997; Verde et al. 1998).

Weighted and censoring biases

Let us consider, once again, the original threshold bias model by Kaiser and let $\nu \to -\infty$. In this limit, galaxy formation is seeded with constant probability throughout space and the net result is an unclustered (Poisson) distribution of objects. This is in agreement with the statistics for density maxima and derives from the fact that all the regions above threshold are assigned the *same* chance to form a galaxy. Catelan et al. (1994) modified the threshold model to build a new biasing scheme which, in the same limit $\nu \to -infty$, originates an unbiased distribution of galaxies. According to their prescription:

$$\rho_g(\mathbf{q}, t) \propto \rho(\mathbf{q}, t)\Theta[\delta(\mathbf{q}, t; R_f) - \nu\sigma(R_f)] \,, \tag{6.23}$$

i.e. the galaxy density is proportional to the local mass density in those regions where δ exceeds the threshold value $\nu\sigma(R_f)$. [5]

The idea of "weighted biasing" may readily be combined with any kind of local bias, to yield a biasing scheme where galaxy counts are particularly enhanced in high density regions and simultaneously suppressed in regions of low density:

$$\rho_g(\mathbf{q}, t) \propto f[\rho(\mathbf{q}, t)]\Theta[\delta(\mathbf{q}, t; R_f) - \nu\sigma(R_f)] \,. \tag{6.24}$$

Mann, Peacock & Heavens (1998) studied the behaviour of the two-point correlation function generated by this censoring process when $f(\rho) = \rho^\alpha$. Following Catelan et al. (1994), they assumed a lognormal PDF for δ to compute the large scale limit of the bias function $b_\infty \equiv b_\xi(r \to \infty)$

$$b_\infty = \alpha + \frac{\sqrt{2/\pi}\exp\left[-(\nu_G - \alpha\sigma_G)^2/2\right]}{\sigma_G\left[1 - \mathrm{erf}\left(\nu_G - \alpha\sigma_G\right)/\sqrt{2}\right]} \,, \tag{6.25}$$

where σ_G and ν_G are, respectively, the variance and the threshold for censoring in the Gaussian process, δ_G, that generates the Eulerian mass density field through the transformation $1 + \delta = \exp\left(\delta_G - \sigma_G^2/2\right)$. Notice that the case $\alpha = 0$ corresponds to the Kaiser model in Eulerian space and that, for sufficiently large ν_G, one always obtains $b_\infty \simeq \nu_G/\sigma_G$ irrespective of the value for α. A more direct way (that can be extended also to non-lognormal fields) to define the threshold is in terms of the fraction of volume, f_V, and mass, f_M, which are censored (i.e. given zero galaxy

[5] Notice that this model cannot be used in Lagrangian space if one assumes that the linear density perturbations are Gaussian. In fact, eq. (6.23) would associate a "negative probability" for galaxy formation to those regions lying in the unphysical tail of the density PDF at $\delta < -1$. This effects is negligible only when $\sigma_\ell \ll 1$ or $\nu \gg 1/\sigma_\ell$ (see Catelan et al. 1994 for a detailed analysis).

density). In the simple case discussed above one obtains (Catelan et al. 1994; Mann, Peacock & Heavens 1998)

$$f_V = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\nu_G}{\sqrt{2}}\right)\right] , \qquad f_M = \frac{1}{2}\left[1 + \mathrm{erf}\left(\frac{\nu_G - \sigma_G}{\sqrt{2}}\right)\right] . \tag{6.26}$$

Tests of the censoring procedure applied to N-body simulations showed that the scale dependence of b_P increases with f_M and that the lognormal approximation for b_∞ gives a fair estimate of the level of large scale bias, especially for $f_M \lesssim 0.5$ (Mann, Peacock & Heavens 1998).

6.4.2 Non-local schemes

The features imprinted on the two-point galaxy correlation function by a local biasing scheme can be summarized as follows (Coles 1993; Scherrer & Weinberg 1998; Mann, Peacock & Heavens 1998): i) on small and intermediate scales (where $\xi \gg 1$ and $\xi \gtrsim 1$, respectively), biasing, as a consequence of its non-linear behaviour, tends to alter the shape of ξ_g with respect to ξ; ii) on large scales (where $\xi \ll 1$), ξ_g always parallels the mass autocorrelation function and a constant bias parameter can be defined.

It is interesting to investigate if these conclusions are altered by dropping the hypothesis of locality of the biasing process. This can be verified as follows. In analogy with eq. (6.17), one can conceive a power-law expansion for non-local biasing such as

$$\delta_g(\mathbf{x},t) = b_0 + \int b_1(\mathbf{x}')\delta(\mathbf{x}-\mathbf{x}')d^3x' + \int b_2(\mathbf{x}',\mathbf{x}'')\delta(\mathbf{x}-\mathbf{x}',t)\delta(\mathbf{x}-\mathbf{x}'',t)\,d^3x'\,d^3x'' + \ldots . \tag{6.27}$$

To lowest order in δ, eq. (6.27) gives the two-point function (Fry & Gaztañaga 1993)

$$\xi_g(x_{12}) = \int b_1(\mathbf{x}_1')b_1(\mathbf{x}_2')\xi(|\mathbf{x}_{12} - \mathbf{x}_{12}'|)d^3x_1\,d^3x_2 , \tag{6.28}$$

which is distorted with respect to ξ on scales comparable with the characteristic range of action of the non-linear kernel $b_1(\mathbf{x})$. Non-local processes must then be invoked to justify any differences in shape between galaxy and matter power spectra on scales $\gtrsim 10\,h^{-1}$Mpc (e.g. to try to reconcile observed data on galaxy clustering with the standard CDM scenario). As a concluding remark, notice that, when both δ_g and δ are smoothed on scales much larger than the range of non-locality, eq. (6.17) should provide an adequate approximation of eq. (6.27).

Only a few non-linear biasing schemes have been discussed in the literature, mainly because it is not easy to think about many physical processes able to influence galaxy formation on cosmologically relevant scales.

Biasing from bright quasars

Following the suggestions by Silk (1985) and Dekel & Rees (1987), Babul & White (1991) explored the possibility that quasars may inhibit the formation of galaxies in their neighbourhood by heating and ionising pregalactic gas. From the mathematical point of view, biasing is obtained by imagining that each high-redshift quasar suppresses galaxy formation in a spherical region seen as a void in the present galaxy distribution. Assuming that these "spheres of avoidance" have radius R_v, and are Poisson distributed with mean number density n_v, the two-point correlation function of any uniform material component that fills the intervoid regions is (Babul 1991)

$$1 + \xi_\mathrm{v}(r) = \begin{cases} \exp\left\{\dfrac{f_\mathrm{v}}{2}\left[\left(\dfrac{r}{2R_v}\right)^3 - 3\left(\dfrac{r}{2R_\mathrm{v}}\right) + 2\right]\right\} & \text{if } r \le 2R_\mathrm{v}, \\ 1 & \text{otherwise,} \end{cases} \tag{6.29}$$

with $f_{\rm v} = 4\pi n_{\rm v} R_{\rm v}^3/3$ the nominal volume filling factor of the spherical void regions. Then, assuming that the spatial distribution of quasars is independent of that of the dark haloes (characterized by the two-point correlation ξ_h), one gets

$$1 + \xi_g(r) = [1 + \xi_{\rm v}(r)][1 + \xi_h(r)] \, . \tag{6.30}$$

Babul & White (1991) showed that, if $R_{\rm v} \simeq 15\, h^{-1}$Mpc and $f_{\rm v} \sim 0.5$, the extra correlation given by $\xi_{\rm v}$ would be sufficient to reconcile the standard CDM scenario with the APM determination of the correlation function. These figures for the parameters do not appear problematic. In fact, it seems indeed posible that a single quasar can ionise gas (and hence affect galaxy formation) out to a comoving distance of $\sim 10\, h^{-1}$Mpc during its lifetime (Rees 1988). The main problem with this scenario is that, in contradiction with observational evidences, it predicts the existence of failed galaxies and galaxy clusters in voids.

Co-operative bias

Bower et al. (1993) presented a model in which galaxy formation occurs at high peaks of the mass density field but is further enhanced by the presence of nearby galaxies. This modification is accomplished by assuming the threshold for galaxy formation to be modulated by large-scale density fluctuations rather than to be spatially invariant,

$$\rho_g(\mathbf{q}, t) \propto \Theta \left\{ \delta(\mathbf{q}, r; R_g) - [\nu\sigma(R_g) - \kappa\delta(\mathbf{q}, t; R_s)] \right\} \, . \tag{6.31}$$

Here, R_s and κ characterize the scale and the amplitude of co-operative effects; in particular, $R_s > R_g$ denotes the typical size of the regions influenced by the presence of galaxies in their neighbourhood, while κ is a constant which quantifies the modulation effect on the threshold for galaxy formation. If κ is positive, peaks in "protosuperclusters" regions have to cross a lower threshold in order to form a galaxy than peaks in "protovoids". Note that adopting too large (small) values for the modulation parameter (influence radius) would dramatically change the predicted abundance of galaxies with respect to the standard threshold model by Kaiser. This problem can be prevented by relating κ to R_s. For example Bower et al. (1993), once selected a value for R_s, determine the coefficient κ so as to have a 2.5% rms modulation of the formation threshold.

It can be shown that, for $R_s \gg R_g$, the co-operative model reduces to the Kaiser biasing scheme at small scales ($r < R_s$) but has extra large scale power when compared with the standard threshold model. Bower et al. (1993) showed how this extra power could be used to reproduce the galaxy autocorrelation function from the APM survey (Maddox et al. 1990) in the context of the standard CDM scenario. An optimal fit is obtained for $R_g = 0.5\, h^{-1}$ Mpc, $\nu = 2.8$, $R_s = 20\, h^{-1}$ Mpc, $\kappa = 2.29$ and, in any case, a value for R_s in the range between $10 - 30\, h^{-1}$ Mpc seems to be required. From a physical point of view, these numbers are indeed impressive. In fact, even though it is expected that co-operative effects arise through radiative and hydrodynamical processes during protogalactic evolution, it is not clear if they can be so efficient to coherently influence such large lenght scales.

6.4.3 From Lagrangian to Eulerian biasing

Both galaxies and dark particles flow through Eulerian space toward mass concentrations. Then, even assuming that the spots of galaxy formation can be easily identified in Lagrangian space, one has to consider large scale motions in order to compute the statistics of the present-day structure.

The simplest case of linear biasing has been addressed, adopting the peak-background split technique, by Bardeen et al. (1986) for the clustering of density maxima and Cole & Kaiser (1989)

for the Press-Schechter case. The key idea is that a positive (negative) value for $\delta_{\rm bg}$ induces a gravitational contraction (expansion) of the volume on the background scale that causes some enhancement (rarefaction) of the halo number density. In fact, in the single stream regime, mass conservation gives $d^3x = d^3q/[1+\delta_{\rm bg}[\mathbf{x}(\mathbf{q},t),t]$ and the number of objects per unit Eulerian volume is then $n_E[\mathbf{x}(\mathbf{q},t),t] = n(\mathbf{q},t)\{1+\delta_{\rm bg}[\mathbf{x}(\mathbf{q},t),t]\}$. Then, the linear Eulerian bias in terms of its Lagrangian counterpart is

$$b_1 \equiv \left.\frac{\partial \log n_E}{\partial \delta^E_{\rm bg}}\right|_{\delta^E_{bg}=0} = 1 + b_L \ , \tag{6.32}$$

with b_L the linear Lagrangian biasing factor. This method has been generalized by Mo & White (1996; see also Mo, Jing & White 1997) to higher order biasing. The key idea is to compute the overabundance of haloes in large spherical cells of given Eulerian radius R and overdensity δ. Assuming that each of these "background" spheres evolves according to the top-hat model (see §2.4.1), it is possible to relate their overdensity in Lagrangian space, δ_L, to its present-day counterpart. In particular, for $|\delta| \ll 1$ and in an Einstein-de Sitter model, one gets

$$\delta_L = \sum_{k=0}^{\infty} a_k \delta^k \tag{6.33}$$

where, for example, $a_0 = 0$, $a_1 = 1$, $a_2 = -17/21$, $a_3 = 341/567$, $a_4 = -55805/130977$, and the coefficients a_k depend only very weakly on the cosmological parameters (Bernardeau 1992). Since the haloes are identified at some specific redshift and no shell-crossing occurs during the spherical evolution of the background, the halo overdensity in the cells is

$$\delta^E_h(M, z_f|R,\delta) \equiv \frac{\mathcal{N}(M, z_f|R_L,\delta_L)}{n(M,z_f)V_E} - 1 \ , \tag{6.34}$$

where $V_E = 4\pi R^3/3$, $R_L = R(1+\delta)^{1/3}$ and the function $\mathcal{N}(M, z_f|R_L,\delta_L)$ is given in eq. (6.11). When $M_L \gg M$ and $|\delta_L| \ll t_{\rm f}$, one then has

$$\delta^E_h(M, z_f|R,\delta) = \sum_{k=0}^{\infty} \frac{b_k}{k!}\delta^k \ . \tag{6.35}$$

The first few bias parameters are related to their Lagrangian counterparts given in eq. (6.14) as follows:

$$\begin{aligned}
b_0 &= 0 \ , \\
b_1 &= 1 + b_1^L \ , \\
b_2 &= 2(1+a_2)b_1^L + b_2^L \ , \\
b_3 &= 6(a_2+a_3)b_1^L + 3(1+2a_2)b_2^L + b_3^L \ , \\
b_4 &= 24(a_3+a_4)b_1^L + 12[a_2^2 + 2(a_2+a_3)]b_2^L + 4(1+3a_2)b_3^L + b_4^L \ .
\end{aligned} \tag{6.36}$$

The same method can be adopted to compute the Eulerian statistics of Lagrangian density peaks (Jing, Mo & White 1997). In this case eq. (6.36) is still valid, provided one uses suitable Lagrangian bias parameters (given, e.g., in Mo, Jing & White 1997). In analogy with the Lagrangian version of the model, the halo overdensity is computed using quantities that are averaged over the ensemble and random fluctuations are not allowed for. For this reason, this algorithm is expected to work

only for quantities such as $S_{j,g}$ that are related to the averages of the $j-$point correlation functions over orientation and configurations. Indeed, in this case, the predictions of the model are in good agreement with numerical simulations (e.g. Mo, Jing & White 1997). More recent analyses tested In particular, it is now clear that b_1 gives an excellent description of clustering for $M > M_*$ while it tends to underestimate for $M < M_*$ (Jing 1998; Sheth & Lemson 1998; Porciani, Catelan & Lacey 1999; Jing 1999). We postpone the discussion of these results into the next chapter. A fully stochastic model for Eulerian biasing that keeps into account also the anisotropic gravitational evolution of the background is also presented there.

6.5 Stochastic biasing

It is difficult to believe that the sole property affecting the efficiency of galaxy formation is the value of the underlying mass density at a certain smoothing scale. Even considering only dark matter haloes, we can think to many other variables which are able to influence the Eulerian local abundance of virialized objects; for example, the dynamical status of the background region (e.g. the eigenvalues of its deformation tensor; for a quantitative analysis see the model developed in the next chapter) or the presence of random density fluctuations on small scales. The situation is even worse for galaxies, where complex hydrodynamical processes play a fundamental role. It is then likely that the action of all these hidden variables will show up as scatter in the δ_g vs. δ relation. In other words, since δ does not contain enough information to fully determine δ_g, any biasing scheme that attempts to link the galaxy counts to the mass density field must be stochastic and not deterministic.

Recently, Dekel & Lahav (1998; see also Scherrer & Weinberg 1998) proposed a new formalism to account for the stochastic character of galaxy biasing. The key idea is to consider the relation between galaxies and mass as a random process specified by the biasing conditional distribution function $\mathcal{P}(\delta_g|\delta)$ of measuring δ_g for a given δ. The average properties of the bias relation are described via the biasing function $b(\delta)$ given by the conditional mean

$$b(\delta)\delta \equiv \langle \delta_g|\delta\rangle = \int \mathcal{P}(\delta_g|\delta)\delta_g d\delta_g \ . \tag{6.37}$$

This is a natural generalization of the deterministic linear biasing relation; note, however, that $b(\delta)$ allows for any possible non-linearity in the biasing scheme. The fluctuating part of the biasing relation is, instead, embodied in the random field

$$\epsilon_g \equiv \delta_g - \langle \delta_g|\delta\rangle, \tag{6.38}$$

for which, by definition, $\langle \epsilon_g|\delta\rangle = 0$.

In order to deal with observational data, it is convenient to characterize $b(\delta)$ through its lowest order moments defined by

$$\hat{b} \equiv \frac{\langle b(\delta)\delta^2\rangle}{\sigma^2} \qquad \text{and} \qquad \tilde{b}^2 \equiv \frac{\langle b^2(\delta)\delta^2\rangle}{\sigma^2} \ . \tag{6.39}$$

The parameter $\hat{b}$ measures the mean biasing, while the ratio $\tilde{b}/\hat{b}$ quantifies the non-linearity in the biasing relation (e.g. $\tilde{b}/\hat{b} = 1$ for a linear scheme). Similarly, the stochasticity in the local δ_g vs. δ relation can be measured by the the biasing scatter function

$$\sigma_b^2(\delta) \equiv \frac{\langle \epsilon_g^2|\delta\rangle}{\sigma^2} \ . \tag{6.40}$$

The (constant) global biasing scatter in instead given by

$$\sigma_b^2 \equiv \frac{\langle \epsilon_g^2 \rangle}{\sigma^2} \,. \tag{6.41}$$

One can then investigate how non-linearity and stochasticity affect galaxy statistics. Considering galaxy counts, it is easy to show that (Dekel & Lahav 1998)

$$b_\sigma^2 = \tilde{b}^2 + \sigma_b^2 \,, \quad \text{and} \quad r \equiv \frac{\langle \delta_g \delta \rangle}{\sigma_g \sigma} = \frac{\hat{b}}{b_\sigma} \,. \tag{6.42}$$

For the two-point correlation function one instead gets

$$\xi_g(r) = \langle b(\delta_1)\delta_1 b(\delta_2)\delta_2 \rangle + \langle \epsilon_{g1}\epsilon_{g2} \rangle \,. \tag{6.43}$$

A preliminary study dealing with simulated dark matter haloes in a CDM cosmology shows only weak non-linearities, $\tilde{b}^2/\hat{b}^2 = 1.08$, but important stochasticity, $\sigma_b^2/\hat{b}^2 = 0.15$ (Dekel & Lahav 1998; Somerville et al. in preparation). These data suggest that the biasing relation can induce 30% fluctuations in the dynamical estimates of the parameter $\beta \propto \Omega_0^{0.6}$. It can be shown that redshift distortions, being less dependent on biasing stochasticity, come out to be a more reliable instrument to measure β (Dekel & Lahav 1998).

The mean bias relation extracted from the simulations is well approximated by the function

$$b(\delta)\delta = \begin{cases} (1+b_0)(1+\delta)^{b_{\rm neg}} - 1 & \text{if } \delta \leq 0 \,, \\ b_{\rm pos}\delta + b_0 & \text{otherwise} \,, \end{cases} \tag{6.44}$$

with $b_{\rm neg} \sim 2$ and $b_{\rm pos} \sim 1$ at $z = 0$. At high redshift both parameters become significantly larger. We refer the reader to the original paper by Dekel & Lahav (1998) for the analyses of higher order moments (e.g. skewness) and redshift distortions.

A stochastic, non-local and non-linear bias model that automatically accounts for the gravitational evolution of the halo population has been developed by Catelan et al. (1998) and successfully tested against numerical simulations by Catelan, Matarrese & Porciani (1998) and Matarrese, Catelan & Porciani (in preparation). Its detailed description will be presented in the next chapter.

6.6 Bias evolution

Deep multiband observations with the HST and from the ground have detected a population of high redshift galaxies through the Lyman-break technique (Steidel & Hamilton 1993). This photometric method allows the identification of candidate high-z objects through the shifting of their Lyman limit cutoff into a particular colour filter. The robustness of the method for U-band ($z \sim 3$), B-band ($z \sim 4$) and V-band ($z \sim 5$) dropout galaxies has been demonstrated by the spectroscopic confirmation of their redshifts with the Keck telescopes (e.g. Steidel et al. 1996). At $z \sim 3$, these galaxies were found to be extremely clustered, more or less as strongly as their present day counterparts (Steidel et al. 1998; Giavalisco et al. 1998; Adelberger et al. 1998). The observed clustering amplitude fits nicely within the standard gravitational instability scenario, provided the Lyman-break galaxies are strongly biased tracers of the mass density field (e.g. Mo & Fukugita 1996; Governato et al. 1998; Wechsler et al. 1998; Coles et al. 1998). Depending on the adopted cosmological model, their typical linear bias factor comes out roughly $b \sim 3 - 6$ (Giavalisco et al. 1998; Adelberger et al. 1998). The standard interpretation is that these galaxies form into the most massive virialized haloes at $z \sim 3$.

On the other hand, a number of studies (e.g. Klypin, Primack & Holtzmann 1996; Cole et al. 1997; Jenkins et al. 1998) compared accurate measurements of the cluster correlations from local surveys with estimates for ξ extracted from dissipationless numerical simulations of popular CDM models. These analyses indicate that in order to fit the observational data: i) at scales $\lesssim 8\,h^{-1}$Mpc, significant anti-bias ($b \sim 1/3$ at $r = 1\,h^{-1}$Mpc) is required for flat, low-Ω_0 models; ii) little anti-bias is needed on the same scales by open models with vanishing cosmological constant; iii) Flat models with primordial power index $n \sim 0.8$ and vanishing cosmological constant require instead a scale-dependent bias of order $1-2$ that becomes stronger on larger scales. Semi-analytical studies generally agree with these conclusions: the local population of bright galaxies seem to require a bias of order unity (e.g. Peacock & Dodds 1994; Peacock 1997).

If the observed structure formed indeed through hierarchical gravitational instability and the Lyman-break galaxies are representative of the high$-z$ population (i.e. their strong clustering is not caused by selection effects), then, regardless of the cosmological model, galaxy bias has decreased significantly from early epochs to the present. Understanding the evolution of bias then represents a crucial step for the development of cosmological research. This section contains an overview of the results obtained so far.

Gone with the flow

A simple model for the evolution of galaxy bias has been developed by Fry (1996; see also Dekel 1986; Dekel & Rees 1987; Nusser & Davis 1994). This is based on the assumption that galaxies form at the time t_* by an arbitrary local process but then move with the streaming velocity given by the gravitational potential. The underlying ideas are that:

1. bias in the galaxy distribution, parametrized as in eq. (6.17), arises at the epoch of galaxy formation;

2. the processes that determine where galaxy form cannot affect how they move; thus, after t_* the galaxy distribution evolves only under the influence of gravity.

Fry uses Eulerian perturbation theory to follow the evolution of mass density fluctuations. The galaxy distribution is, instead, described through its continuity equation,

$$\frac{\partial \delta_g}{\partial t} + \frac{1}{a}\nabla \cdot (1+\delta_g)\mathbf{v} = 0\,, \tag{6.45}$$

which is coupled to the Eulerian fluid equations for δ by the mass velocity field $\mathbf{v}$. Solving the system of equations at linear order, one obtains that, for $t > t_*$, the effective bias $b_{\rm eff} = \delta_g^{(1)}/\delta^{(1)}$ depends on time as

$$b_{\rm eff}(t) = \frac{d(t) + b_* - 1}{d(t)} \tag{6.46}$$

with $d(t) \equiv D_+(t)/D_+(t_*)$ and $b_* = b_1(t_*)$. Thus, if galaxy motions are determined by gravity, a large bias cannot be sustained and the galaxy distribution eventually relaxes to that of the mass. The reverse of this is that any present bias must have been even larger in the past. Similar conclusions hold for higher order statistics like the "skewness" $S_{3,g}$ and the bispectrum amplitude Q_g computed to leading order in perturbation theory (Fry 1996; see also Buchalter & Kamionkowski 1999).

Tegmark & Peebles (1998) have generalized this argument to the case of stochastic (but linear) biasing, showing that both $\langle \delta_g \delta \rangle / \sigma_g \sigma$ and σ_g/σ approach unity as time passes. They show that

this is true even in simple, continuous models of galaxy formation. Note however that in open cosmologies, where D_+ saturates, the asymptotic value of b as $t \to \infty$ is generally different from 1 (Catelan, Matarrese & Porciani 1998).

The simple model we presented in this section has been generalized to the fully non-linear regime by Catelan et al. (1998; see also chapter 7).

6.6.1 Galaxy biasing

Semi-analytic models

Since the connection between the galaxies and their host haloes is far from being trivial, observational data need to be interpreted within the framework of a complete scheme of galaxy formation and evolution. One possibility is given by the so-called semi-analytic models of galaxy formation (e.g. Kauffmann, White & Guiderdoni 1993; Cole et al. 1994; Somerville & Primack 1998; see also section 1 of Kauffmann et al. 1999a for a complete list of references). In these models, the merger history of dark matter haloes is followed either through N-body simulations or using a Monte Carlo approach based on the PS formalism. Within each halo, galaxy formation is then described using a set of simple recipes that model the processes of gas cooling, star formation, feedback from supernovae, dynamical friction and galaxy merging (which, for example, determines the morphological type). Eventually, spectrophotometric models are used to convert the star formation and dust production histories of each galaxy into a set of magnitudes relative to different wavebands. The final output can then be "observed" adopting the same selection criteria used for real galaxies. These models contain a number of free-parameters that are fixed by constraining the outputs to match selected properties of the local galaxy distribution. The fact that the models are able to reproduce a larger number of observationally established relations than the constraints, it is reassuring about the reliability of the method.

Recently, a series of studies (Kauffmann, Nusser & Steinmetz 1997; Kauffmann et al. 1999a; Kauffmann et al. 1999b; Benson et al. 1999) combined semi-analytic models with high-resolution N-body simulations to investigate the evolution of galaxy clustering in popular cosmological scenarios. Their main results are summarized below.

- Galaxies do not faithfully trace the mass density field. The corresponding bias is dependent on the criterion adopted to select the galaxy sample. In particular, differences in the clustering amplitude (and evolution) between galaxies of differing star formation rate, colour, morphological type and luminosity have been found. For instance, at a given z, red galaxies, bright galaxies and early-type galaxies are more clustered than the whole population (Kauffmann, Nusser & Steinmetz 1997; Kauffmann et al. 1999a, 1999b). In fact, these galaxies occur preferentially in the cores of groups and clusters. Conversely, star-forming galaxies are less clustered on small scales because they are predominantly central halo galaxies and thus spatially exclusive.

- Galaxy biasing is strongly time dependent (Kauffmann et al. 1999b). Once selected a galaxy sample (e.g. considering galaxies brighter than some given absolute magnitude) this can be seen by tracing the evolution of b_ξ at a fixed comoving separation. The general trend is a progressive decrement of b_ξ towards the unbiased case, the details depending on the cosmological model. In particular, when b_ξ becomes roughly 1, no further evolution is seen. Moreover, except at very high redshifts ($z \gtrsim 3$), the statistical bias b_ξ is scale independent (Kauffmann, Nusser & Steinmetz 1997; Kauffmann et al. 1999b), at least on scales $\gtrsim 2 - 3\,h^{-1}$Mpc.

- Alternatively, one can follow directly the evolution of ξ_g at fixed comoving separation. In this case, depending on the background cosmological scenario, the clustering amplitude either decreases continuously from high redshift to the present epoch or reaches a minimum when $1 < z < 2$ and then starts growing. In terms of b_ξ this can be understood as follows. At high$-z$, the bias is rapidly decreasing and ξ is growing at a modest pace. Then $\xi_g = b_\xi^2 \xi$ has a large, negative, time derivative. On the other hand, at more recent epochs, non-linear gravitational evolution speeds up the growth of ξ while b_ξ slows down its evolutionary rate. These trends tend to balance and ξ_g keeps roughly constant. Eventually, once $b_\xi \simeq 1$, the galaxy correlation function matches ξ and starts growing at a fast pace tracing the fully non-linear behaviour of the mass density field. This last phase starts before the present epoch only in some cosmological models. In particular, in the ΛCDM scenario, ξ_g grows with time for $z \lesssim 1.5$ and the resulting evolution for the galaxy correlation length agrees remarkably well with the data by Carlberg et al. (1998).

- On scales smaller than $2 - 3\,h^{-1}$Mpc galaxies are anti-biased with respect to the dark matter distribution (Benson et al. 1999). This scale dependent effect is probably determined by the complex interplay of various physical effects including the spatial exclusion of dark matter clumps, the internal structure of the most massive haloes, the PDF of the number of galaxies hosted in the haloes with similar structure and dynamical friction.

- The normalization of the linear power spectrum of density fluctuation plays a fundamental role in determining the present-day clustering properties of the galaxy population. This is because, higher normalizations correspond to larger values for M_* at $z = 0$ and, in this case, luminous galaxies are hosted in relatively low mass haloes which are unbiased tracers of the dark matter distribution. In these conditions, the dependence of galaxy clustering on luminosity can be completely cancelled (Kauffmann et al. 1999a).

- No velocity bias is present. [6]

Hydrodynamic simulations

A second approach to theoretically explore the evolution of galaxy clustering is to use hydrodynamic simulations incorporating a heuristic model for galaxy formation (e.g. Cen & Ostriker 1992, 1999; Katz, Hernquist & Weinberg 1992, 1998; Steinmetz 1996). In this case, to follow the baryonic, stellar and dark components at the same time, one combines numerical gasdynamical simulations with large scale N-body codes and uses physically motivated prescriptions to convert baryonic fluid patches into collisionless stellar particles. In words, the star formation criterion can be expressed as follows: if a cell's density is overdense enough, if the cooling time of the gas in it is shorter than its dynamical time, if the baryonic mass in it is greater than the Jeans mass, and if the flow around the cell is converging, it will have stars forming inside of it. The baryonic fluid component is then turned into stars at a rate proportional to $m_b/t_{\rm dyn}$, where m_b is the mass of the gas in the cell and $t_{\rm dyn}$ is the local dynamical time. These galaxy particles, whose mass in the most recent codes ranges from about 10^6 to 10^9 $M_\odot$, subsequently contribute to metal production and to the background ionising UV radiation. Notice that, to avoid resolution problems, in this kind of studies it is customary to define a smooth galaxy mass density field instead of grouping the stellar particles into galaxies.

[6] Carlberg & Couchman (1989) found that the dispersion in the relative velocities of pairs of simulated dark matter clumps was much less than that of pairs of randomly selected mass particles of the same separation. Differences in velocity statistics derived from galaxies of different types or compared with the matter are called velocity biases.

A number of authors recently addressed the problem of galaxy biasing using hydrodynamical simulations (Cen & Ostriker 1992, 1993, 1999; Blanton et al. 1998, 1999; Katz, Hernquist & Weinberg 1992; 1998). Their results can be summarized as follows.

- Galaxies are biased tracers of the mass density field, in particular underdense regions in mass correspond to empty voids in the galaxy distribution (Blanton et al. 1998). The bias b, defined as the ratio between the rms fluctuation in galaxy and total mass inside a sphere of given radius, decreases with increasing smoothing scale (Blanton et al. 1998). For example, in a ΛCDM model, the present-day bias for all galaxies at $30\,h^{-1}$Mpc is 1.2 while it becomes sligthly larger than 2.6 at $1\,h^{-1}$Mpc (Blanton et al. 1998; Cen & Ostriker 1999). The same behaviour characterizes b_ξ (Cen & Ostriker 1999). The predicted value at the fiducial smoothing length of $8\,h^{-1}$Mpc, $b = 1.41$ ($b_\xi = 1.35$), is consistent with observational data (Loveday et al. 1996).

- Adopting the formalism developed by Dekel & Lahav (1998), it results that the scale dependence of galaxy biasing is caused mainly by scatter in the δ_g vs. δ relation, while non-linearity is not very significant. The scatter, however, is not random and correlates with the large scale mass density field. A more detailed description of galaxy biasing as a two parameter relation $\delta_g(\delta, X)$, with X the gas temperature or the velocity dispersion for dark matter particles, sensibly reduces the scatter (Blanton et al. 1998). This works because both the gas temperature and the matter velocity dispersion are tightly connected to the value of the local gravitational potential that has a stronger dependence on the large scale modes of the density field than δ itself. A good fit that reproduces the scale dependence of b is

$$1 + \delta_g = 1.23\,(1+\delta)^{1.9}\left(1 + \frac{T}{40,000\,\mathrm{K}}\right)^{-0.66}, \qquad (6.47)$$

  but for very young galaxies the dependence on the gas temperature, T, is stronger.

- Galaxy bias is a monotonically increasing function of redshift and the predictions of hydrocodes are generally in excellent agreement with observations of Lyman-break galaxies (Katz et al. 1998; Cen & Ostriker 1999). The correlation function of dark matter haloes decreases from very high redshifts to $z \sim 1$ and then rises towards zero redshift (Cen & Ostriker 1999; see also Kauffmann et al. 1994 and Colin et al. 1998). However, in contrast with semi-analytic modelling (Kauffmann et al. 1999b), no dip is seen in the redshift evolution of galaxy correlations that, on the contrary, has a peak at $z \sim 1$ (Cen & Ostriker 1999). A plausible explanation is that, at $z \lesssim 1$, galaxy formation preferentially avoids the densest spots (e.g. clusters of galaxies) because they are too hot to permit cooling and gravitational instability to occur effectively. This possibility has been investigated by Blanton et al. (1999) who studied the clustering properties of star forming galaxies. Their results show that, for newly born galaxies, the linear correlation coefficient r, defined in eq. (6.42), changes from $r \sim 0.9$ at $z = 3$ (when $b(8\,h^{-1}\mathrm{Mpc}) \sim 4.5$ for these objects) to $r \sim 0.25$ at $z = 0$ (corresponding to $b(8\,h^{-1}\mathrm{Mpc}) \sim 1$). Notice that, even though $b \to 1$ as $z \to 0$, a decreasing r means that star-forming galaxies become poorer tracers of the mass density field.

- Accordingly, at any epoch older galaxies are more strongly clustered (and more correlated to the mass distribution) than younger galaxies. The metallicity of the first generation of stars in a galaxy is another source of segregation: galaxies with initial high metallicity are more strongly biased than the others. The effect is anyway weak compared to the age segregation (Cen & Ostriker 1999).

- At $z = 0$, the peculiar velocity field of young galaxies is colder than that for old galaxies (the difference between the average velocity amplitudes is typically $100\,\mathrm{Km}\,s^{-1}$ for $1\,h^{-1}\mathrm{Mpc}$ smoothing). This reflects the fact that older galaxies tend to reside in deeper potential holes than do younger ones. In general, the whole galaxy population tends to move more slowly than does the dark matter. This (weak) velocity anti-bias is probably of gasdynamical origin (Cen & Ostriker 1999).

Chapter 7

The bias field of dark matter haloes

7.1 Introduction

The theory proposed by PS to obtain the relative abundance of matter condensations in the Universe has strongly influenced all later studies on the statistical properties of dark matter haloes and led to a large variety of extensions, improvements and applications. An important aspect of the PS model is that, being entirely based on linear theory, suitably extrapolated to the collapse time of spherical perturbations, it is, by definition, local in Lagrangian space. While this Lagrangian aspect of the theory does not have immediate implications for the study of the mass function of dark matter haloes, it is, instead, of crucial importance for their spatial clustering properties. This point was recognized by Cole and Kaiser (1989) and, more recently, by Mo and White (1996, hereafter MW), who proposed a bias model for halo clustering in Eulerian space, by a suitable extension of the original PS algorithm for the mass function. With their formalism MW studied the clustering of dark matter haloes with different formation epochs (see also Mo, Jing & White 1996). The comparison of their theoretical predictions with the spatial distribution of haloes obtained by a *friends-of-friends* group finder and by a spherical overdensity criterion in numerical simulations proved extremely successful.

These very facts imply that there exists a local version of the PS algorithm providing a *mapping* between points of Lagrangian space and the haloes in embryo which will come into existence at the various epochs. For a given realization of the initial density field, the PS mapping is such that, at a fixed redshift z, each Lagrangian point $\mathbf{q}$ can be assigned to a matter clump of some mass M, identified by a suitable Lagrangian filter, which is collapsing at the epoch $z_f = z$. One can therefore exploit the existence of this mapping to assign a stochastic halo process, our *halo counting field* below, to each point $\mathbf{q}$. This will be the starting point of our analysis.

What the PS ansatz cannot account for is the fact that the fluid elements are moved apart by gravity, so that the halo which the PS mapping assigns to the fluid patch with Lagrangian coordinate $\mathbf{q}$ is not going to collapse in the same position, i.e. at $\mathbf{x} = \mathbf{q}$, but, rather in the Eulerian point $\mathbf{x}(\mathbf{q}, z) = \mathbf{q} + \mathbf{S}(\mathbf{q}, z)$, with $\mathbf{S}(\mathbf{q}, z)$ the displacement vector, corresponding to the Lagrangian one at the epoch $z = z_f(\mathbf{q}, M)$ of its collapse. This fact, while not affecting in any way the PS result for the mean mass function, as the average halo abundance cannot change by scrambling the objects, sensibly modifies their spatial clustering properties. Modelling the latter effect is one of the main purposes of the present work. In their derivation of the Eulerian halo bias MW took into some account this problem by allowing for the local compression, or expansion, of the volumes where the haloes are located, an effect which is of crucial importance for the derivation of the correct halo density contrast. Their derivation, however, is formally flawed by the fact that they only deal with mean halo number densities, so that they are forced to define the bias in terms of

them. For reasons to be shown below, however, this heuristic treatment can be put on sounder statistical grounds, by applying a suitable coarse-graining procedure.

Of course, the PS model has its own limitations. The comparison of its predictions for the mass function with the outputs of N-body simulations (e.g. Efstathiou et al. 1988; Gelb & Bertschinger 1994; Lacey & Cole 1994; Gross et al. 1998; Lemson & Kauffmann 1999), while surprisingly successful in its general trends, given the simplicity of the assumptions, showed a number of problems. Gelb and Bertschinger (1994), for instance, found that the simulated haloes are generally less massive than predicted, the reason being that merging does not erase substructure in large haloes as fast as required by the PS recipe. On the other hand, there is nowadays general agreement about the fact that the PS method overestimates the abundance of haloes in the low-mass tail by a factor of ~ 1.5 (Lacey & Cole 1994; Gross et al. 1998; Lemson & Kauffmann 1999).

There have been many attempts to improve the original PS model. If cosmic structures preferentially formed at the peaks of the initial density fluctuation field this would affect their mean mass function (Bardeen et al. 1986; Bond 1988; Colafrancesco, Lucchin & Matarrese 1989; Peacock & Heavens 1990; Manrique & Salvador–Solé 1995, 1996). Bond and Myers (1996) developed a *peak-patch* picture of cosmic structure formation, according to which virialized objects are identified with suitable peaks of the Lagrangian density field. The peak-patch collapse dynamics is then followed in terms of the homogeneous ellipsoid model, which allows for the influence of the external tidal field, while the Zel'dovich approximation (Zel'dovich 1970) is used for the external peak-patch dynamics. The effects of non-spherical collapse on the shape of the mass distribution were studied by Monaco (1995). Lee and Shandarin (1998) analytically derived the mass function of gravitationally bound objects in the frame of the Zel'dovich approximation.

We prefer here to follow the simple lines of the PS theory to set up the 'initial conditions' for our stochastic approach to the evolution of halo clustering. Nevertheless, one should keep in mind that our approach is flexible enough to accept many levels of improvement in the treatment of the Lagrangian initial conditions.

A relevant part of the following analysis will be devoted to the study of the evolution of halo clustering away from the linear regime. It turns out that the problem can be solved exactly in terms of the evolved mass density. An important result of this analysis is that the general Eulerian bias factor, defined as the local ratio between the halo density contrast and the mass fluctuation field, is both non-linear and non-local. The latter property follows directly from our selection criterion of candidate haloes out of the linear density field.

Our algorithm can also be seen as a specific example of a bias model which is local in Lagrangian space. This is expected to have relevant consequences on galaxy clustering. Because of this local Lagrangian character, our model strongly differs from the local Eulerian bias prescription applied by Fry and Gaztañaga (1993) to the analysis of the hierarchical correlation functions. A simple test of our theory can be obtained by analysing the behaviour of the bispectrum (or the skewness), whose shape (scale) dependence will be shown to be directly sensitive to the assumption of local bias in Lagrangian vs. Eulerian space.

Our results for the evolved halo distribution generally allow to study their statistical properties at the required level of non-linearity, and could be further implemented to generate mock halo catalogues starting from low-resolution numerical simulations of the dissipationless matter component. These results have important implications for the study of the redshift evolution of galaxy clustering, a problem made of compelling relevance by the growing body of observational data at high-redshift which are being produced by the new generation of large telescopes. A general study of this problem has been recently performed by Peacock (1997), Matarrese et al. (1997), Moscardini et al. (1998) and Arnouts et al. (1999); these authors pointed out that knowledge of the evolution of the effective bias for the various classes of objects is a key ingredient in the comparison of theoretical

scenarios of structure formation with observational data on clustering at high redshift. Kauffmann, Nusser and Steinmetz (1997) and Kauffmann et al. 1999a,b used both semi-analytical methods and N-body techniques to study the physical origin of bias in galaxies of different luminosity and morphology.

The plan of this chapter is as follows. In §7.2 we define our halo counting field, within the linear approximation, both in the Lagrangian and Eulerian context. The non-linear evolution of the halo clustering is analysed in §7.3, where we also compute the bispectrum and skewness of the evolved halo distribution. §7.4 contains a general discussion of our results and some conclusions.

7.2 Stochastic approach to halo counting and clustering

7.2.1 Basic tools and notation

Let us assume that the mass density contrast linearly extrapolated to the present time, $\epsilon(\mathbf{q})$, is a statistically homogeneous and isotropic Gaussian random field completely determined by its power-spectrum $P(k)$. [1] Here $\mathbf{q}$ represents a comoving Lagrangian coordinate. As discussed in §2.2.3, a smoothed version of the field $\epsilon(\mathbf{q})$ is obtained by convolving it with a rotationally invariant filter $W(q,R)$, containing a resolution scale R, with associated mass $M \sim \rho_b R^3$, ρ_b being the background mean density at $z=0$,

$$\epsilon_{_R}(\mathbf{q}) = \int d\mathbf{q}'\, W(|\mathbf{q}-\mathbf{q}'|,R)\,\epsilon(\mathbf{q}') \equiv \epsilon_{_M}(\mathbf{q})\ . \tag{7.1}$$

The smoothed field is also Gaussian with one–point distribution function $G_{\sigma_{_M}}(\epsilon_{_M}) = (2\pi\,\sigma_{_M}^2)^{-1/2}$ $\exp(-\epsilon_{_M}^2/2\sigma_{_M}^2)$, where $\sigma_{_M}^2$ denotes the variance of $\epsilon_{_M}$, $\sigma_{_M}^2 \equiv \langle\epsilon_{_M}^2\rangle = (2\pi^2)^{-1}\int_0^\infty dk\,k^2\,P(k)\widetilde{W}(kR)^2$. In the following, we will often be concerned with the joint distribution of the fields $\epsilon_{_{M_1}}(\mathbf{q})$ and $\epsilon_{_{M_2}}(\mathbf{q})$. The two–point correlation function of the linear density contrast smoothed on the scale R_1 and R_2 is

$$\xi_{12}(q) = \langle\,\epsilon_{_{M_1}}(\mathbf{q}_1)\,\epsilon_{_{M_2}}(\mathbf{q}_2)\,\rangle = \frac{1}{2\pi^2}\int_0^\infty dk\,k^2\,P(k)\,\widetilde{W}(kR_1)\,\widetilde{W}(kR_2)\,j_{\rm o}(kq)\ , \tag{7.2}$$

where $q=|\mathbf{q}_1-\mathbf{q}_2|$ and $j_{\rm o}(x)$ is the spherical Bessel function of order zero. We term σ_{12}^2 the value assumed by ξ_{12} in the limit $q\to 0$.

The properties of the filtered quantities clearly depend upon the choice of the window function. For instance, the relation between the mass enclosed by a top-hat filter is the standard $M(R) = 4\pi\rho_b R^3/3$, while, for a Gaussian window, the enclosed mass is $M(R) = (2\pi)^{3/2}\rho_b R^3$. These two masses coincide for $R_{\rm G} = 0.64\,R_{\rm TH}$ (Bardeen et al. 1986).

In the literature, the sharp top-hat filtering has been alternatively adopted in Fourier space, $\widetilde{W}_R(k) = \Theta(k_R - k)$, where $k_R = 1/R$ and $k=|\mathbf{k}|$. The most remarkable property of this filter is that each decrease of the smoothing radius adds up a new set of uncorrelated modes (Bardeen et al. 1986; Bond et al. 1991; Lacey & Cole 1993). This also implies that, for example, the correlation function in eq. (7.2) simplifies to $\xi_{12} = \xi_{11}$, whenever $k_{R_1} < k_{R_2}$; consequently, $\sigma_{12} = \sigma_{11} \equiv \sigma_1$. In practice, the information is always erased below the largest of the two smoothing lengths. This property will be particularly useful in the next sections. For this 'sharp k-space' filter, the main difficulty is how to associate a mass $M(R)$ to the cutoff wavenumber k_R. Lacey and Cole (1993)

[1] The field $\epsilon(\mathbf{q})$ coincides with the quantity denoted by $\delta_\ell(\mathbf{q})$ in chapter 2. Here, the new symbol is introduced to simplify the notation.

give the expression $M(R) = 6\pi^2\rho_b k_R^{-3}$, which coincides with the mass within a top-hat filter if one takes $k_R = 2.42/R_{TH}$.

In the next section we introduce the halo counting random fields that allow a fully stochastic description of the biased haloes distribution. To illustrate how our formalism works, we first show how to derive the PS mass function by performing a simple averaging of our stochastic counts.

7.2.2 Lagrangian mass function: Press-Schechter theory

A local version of the PS approach can be built up in terms of stochastic counting operators acting on the underlying Gaussian density field, as follows. The number of haloes per unit mass, contained in the unit comoving volume centered in $\mathbf{q}$, identified by the collapse-threshold $t_f(z_f) \equiv \delta_c/D(z)$ (where $D(z)$ denotes the linear growth factor of density perturbations normalized as $D(0) = 1$), is described as a density field of a point process by

$$\mathcal{N}_h^L(\mathbf{q}|M, t_f) = -2\,\frac{\rho_b}{M}\,\frac{\partial}{\partial M}\Theta\left[\epsilon_M(\mathbf{q}) - t_f\right] . \tag{7.3}$$

Note that the quantity $\mathcal{N}_h^L(\mathbf{q}|M, t_f)$ is non-zero only when the filtered density contrast in $\mathbf{q}$ upcrosses, or downcrosses, the threshold t_f, by varying the smoothing length R (or the corresponding mass M). The factor of 2, appearing in the expression of $\mathcal{N}_h^L(\mathbf{q}|M, t_f)$, is needed in order to obtain the right normalization of the mass function, in which case it has been shown to be intimately related to the solution of the cloud-in-cloud problem (Peacock & Heavens 1990; Bond et al. 1991; Cole 1991; see also chapter 4), at least for sharp k-space filtering. At this level, our description should be thought of as a sort of differential version of Kaiser's bias model (Kaiser 1984), that defines a population of objects with the right average halo abundance and their related clustering properties, rather than a detailed modelling of how structures form from the primordial density field. The results that we presented in chapter 4 (see also Porciani et al. 1998) show, however, that the present approach is fully consistent with a rigorous treatment of the cloud-in-cloud problem. We remember that, in chapter 4, halo correlations have been obtained from pairs of first-upcrossing 'times' for spatially correlated random walks above the collapse threshold t_f.

It can be seen from equation (7.3) that a population of haloes is uniquely specified by the two parameters M and t_f. In the standard PS formulation t_f is interpreted as a sort of time variable, related to the formation redshift z_f, which decreases with real time, as every halo continuously accretes matter. In this sense one can say that, for a continuous density field with infinite mass resolution, each halo disappears as soon as it forms to originate another halo of larger mass.

Alternatively, instead of considering t_f as a time variable, one can use it simply as a label attached to each halo. The haloes so labelled can be thought as keeping their identity during the subsequent evolution at any observation redshift z. This is not in contrast with the fact that in the real Universe dark matter haloes undergo merging at some finite rate (e.g. Lacey & Cole, 1993, 1994). Within such a picture, in fact, the physical processes of accretion and merging reduce to the trivial statement that haloes identified by a given threshold are necessarily included in catalogues of lower threshold, so that, in the limit of infinite mass resolution, only haloes with $z_f = z$ would actually survive. Nevertheless, keeping z_f distinct from z may have several advantages, among which the possibility of allowing for a more realistic description of galaxy and cluster formation inside haloes, for both the evolution of the luminosity function (Cavaliere, Colafrancesco & Menci 1993; Manrique & Salvador-Solé 1996) and of the galaxy bias (e.g. MW; Kauffmann et al. 1997). Let us stress, however, that we are not addressing here the issue of galaxy or cluster merging: our method is completely general in this respect and allows to span all possible models, from the instantaneous merging hypothesis ($z_f = z$) to the case of no merging at all (z_f fixed for changing $z \leq z_f$).

In what follows, therefore, we will assume that we can deal with the halo population specified by the formation threshold t_f at any redshift z. Only in this sense we will say that we 'ignore' the effects of merging in our description: merging can be exactly recovered at any step, and with any assumed mass resolution, as the relation between z_f and z. To implement this idea it is enough to scale appropriately the argument of the Heaviside function in eq.(7.3), which can be recast in the form

$$\mathcal{N}_h^L(\mathbf{q}, z|M, z_f) = -2\,\frac{\rho_b}{M}\,\frac{\partial}{\partial M}\Theta\left[\epsilon_{_M}(\mathbf{q}, z) - \delta_f(z, z_f)\right]\,, \tag{7.4}$$

where $\epsilon_{_M}(\mathbf{q}, z) \equiv D(z)\epsilon(\mathbf{q})$ and $\delta_f(z, z_f) \equiv \delta_c D(z)/D(z_f)$. It can be easily seen that the ensemble average of the counting field $\mathcal{N}_h^L(\mathbf{q}, z|M, z_f)$ corresponds to the PS mass function

$$\langle \mathcal{N}_h^L(\mathbf{q}, z|M, z_f)\rangle\, dM = n_{_{PS}}(z|M, z_f)\, dM\,, \tag{7.5}$$

where

$$n_{_{PS}}(z|M, z_f)\, dM \equiv \frac{1}{\sqrt{2\pi}}\,\frac{\rho_b}{M}\,\frac{\delta_f(z, z_f)}{\sigma_{_M}^3(z)}\,\exp\left[-\frac{\delta_f^2(z, z_f)}{2\sigma_{_M}^2(z)}\right]\left|\frac{d\sigma_{_M}^2(z)}{dM}\right| dM\,. \tag{7.6}$$

Note that we emphasized the z-dependence of the comoving mass function, though it is straightforward to verify that the value of $n_{_{PS}}(z|M, z_f)$ does not change with z. In fact, since we are ignoring the effects of merging, once a class of haloes has been identified, their mean comoving density keeps constant in time. Thus, as far as the mass function is concerned, the introduction of the observation redshift z is somewhat more formal than physical. However, this distinction will be far more significant in the next sections, where, in order to compute the halo-to-mass bias factor, we will relate the Lagrangian distribution of a population of haloes selected at z_f to the mass density fluctuation field linearly extrapolated to the redshift z. Models of galaxy formation which assume that galaxies form at a given redshift z_f with some initial bias factor and that their subsequent motion is purely caused by gravity (e. g. Dekel 1986; Dekel & Rees 1987; Nusser & Davis 1994; Fry 1996) can be easily accommodated into this scheme.

To conclude this section, let us consider the integral stochastic process

$$\int_M^\infty dM' M' \mathcal{N}_h^L(\mathbf{q}, z|M', z_f) = 2\,\rho_b\,\Theta\left[\epsilon_{_M}(\mathbf{q}, z) - \delta_f(z, z_f)\right]\,, \tag{7.7}$$

representing the fraction of mass, in the unit Lagrangian comoving volume centered in $\mathbf{q}$, which at redshift z_f has formed haloes more massive than M. This coincides with the original Kaiser bias model (Kaiser 1984) up to the multiplicative factor $2\rho_b$, which is irrelevant for calculating correlation functions.

7.2.3 Conditional Lagrangian mass function

The PS theory reviewed in the previous section describes the overall distribution of halo masses in a homogeneous universe of mean density ρ_b. However, of cosmological interest is also, for instance, the estimate of the halo distribution within rich or poor environments (which can be related to the galaxy number enhancement per unit mass in rich clusters or in voids), thus justifying the investigation of the distribution of halo masses conditioned to lie within a larger uncollapsed container of given density. The conditional mass function has been studied by several authors (e.g. Bond et al. 1991; Bower 1991; Lacey & Cole 1993).

We extend here the approach introduced in the previous section in order to derive the conditional mass function. Specifically, we calculate the comoving mass function, in the mass range M, $M+dM$,

for objects contained in a large region of dimension $R_\circ$, corresponding to a mass $M_\circ$, with local density contrast $\epsilon_\circ \equiv \epsilon_{M_0}$. We will require $\epsilon_\circ \ll \delta_f$ and $R_\circ \gg R$, to ensure that the container is not collapsed yet by the epoch z_f and that it encloses a non-negligible population of objects.

In order to mimic these environmental effects, we modify the halo counting field according to

$$\mathcal{N}_h^L(\mathbf{q}, z|M, z_f|M_\circ, \epsilon_\circ) = -\frac{2}{N_\circ}\frac{\rho_b}{M}\frac{\partial}{\partial M}\Theta\big[\epsilon_M(\mathbf{q}, z) - \delta_f(z, z_f)\big]\,\delta_D\big[\epsilon_{M_0}(\mathbf{q}, z) - \epsilon_\circ\big]\,, \tag{7.8}$$

where $\delta_D(\mathbf{q})$ denotes the Dirac delta function and $N_\circ \equiv \langle \delta_D[\epsilon_{M_0}(\mathbf{q}, z) - \epsilon_\circ]\rangle$ is the normalization constant. Here the scalar $\epsilon_\circ$ indicates the value of the random field $\epsilon_{M_0}(\mathbf{q}, z)$. Taking the ensemble average (and using the cross-variance σ_{ij} for a sharp k-space filter) one eventually obtains

$$\langle\, \mathcal{N}_h^L(\mathbf{q}, z|M, z_f|M_\circ, \epsilon_\circ)\,\rangle\, dM = n_{PS}(z|M, z_f|M_\circ, \epsilon_\circ)\, dM\,, \tag{7.9}$$

where

$$\begin{aligned} n_{PS}(z|M, z_f|M_\circ, \epsilon_\circ)\, dM = \frac{1}{\sqrt{2\pi}}\frac{\rho_b}{M}\frac{\delta_f(z, z_f) - \epsilon_\circ}{[\sigma_M^2(z) - \sigma_\circ^2(z)]^{3/2}} \times \\ \times \exp\left\{-\frac{[\delta_f(z, z_f) - \epsilon_\circ]^2}{2\,[\sigma_M^2(z) - \sigma_\circ^2(z)]}\right\}\left|\frac{d\sigma_M^2(z)}{dM}\right| dM\,. \end{aligned} \tag{7.10}$$

This straightforward calculation shows how to obtain results already known in the literature by simply starting from the random field in eq.(7.8): averaging that halo counting field leads to the expected conditional mass function. But not only: unlike previous treatments, once the halo counting field has been consistently defined, other statistics, like the two-point halo correlation function, can be calculated. We will carry out this program in the next section.

7.2.4 Lagrangian clustering: halo-to-mass bias from correlations

In this section we will compute the halo-halo correlation function which coincides with the correlation function of our random counting field. Specifically, we will calculate the Lagrangian halo correlation function from the Lagrangian counting field $\mathcal{N}_h^L(\mathbf{q}, z|M, z_f)$. By definition, the correlation function of this stochastic process is given by

$$\xi_{hh}^L(q, z|M_1, z_1; M_2, z_2) = \frac{\langle\, \mathcal{N}_h^L[\mathbf{q}_1, z|M_1, \delta_f(z, z_1)]\, \mathcal{N}_h^L[\mathbf{q}_2, z|M_2, \delta_f(z, z_2)]\,\rangle}{\langle\, \mathcal{N}_h^L[\mathbf{q}_1, z|M_1, \delta_f(z, z_1)]\,\rangle\, \langle\, \mathcal{N}_h^L[\mathbf{q}_2, z|M_2, \delta_f(z, z_2)]\,\rangle} - 1\,, \tag{7.11}$$

where $q = |\mathbf{q}_1 - \mathbf{q}_2|$. Performing the ensemble average over the Gaussian fields $\epsilon_{M_1}(\mathbf{q})$ and $\epsilon_{M_2}(\mathbf{q})$, we obtain

$$\begin{aligned} &\langle\, \mathcal{N}_h^L[\mathbf{q}_1, z|M_1, \delta_f(z, z_1)]\, \mathcal{N}_h^L[\mathbf{q}_2, z|M_2, \delta_f(z, z_2)]\,\rangle = \\ &\qquad = \frac{4\rho_b^2}{M_1 M_2}\frac{\partial}{\partial M_1}\frac{\partial}{\partial M_2}\int_{\delta_f(z,z_1)}^{\infty}\int_{\delta_f(z,z_2)}^{\infty} d\alpha_1\, d\alpha_2\, G_2(\alpha_1, \alpha_2)\,, \end{aligned} \tag{7.12}$$

where $G_2(\alpha_1, \alpha_2)$ denotes the bivariate Gaussian distribution

$$G_2(\alpha_1, \alpha_2) = \left[2\pi\sigma_1\sigma_2\sqrt{1 - \Omega^2}\right]^{-1} \exp\left[-\left(\frac{\alpha_1^2}{\sigma_1^2} + \frac{\alpha_2^2}{\sigma_2^2} - 2\,\Omega\frac{\alpha_1}{\sigma_1}\frac{\alpha_2}{\sigma_2}\right)\Big/2(1 - \Omega^2)\right], \tag{7.13}$$

with normalized correlation $\Omega(q) = \xi_{12}(q)/\sigma_{M_1}\sigma_{M_2}$ and $\sigma_i \equiv D(z)\sigma_{M_i}$.

The full exact expression for the halo-halo correlation function can be obtained after an incredibly long algebraic computation. We report here only the final expression. Defining $\delta_{fi} \equiv \delta_f(z, z_i)$, we have

$$
\begin{aligned}
1 + \xi^L_{hh}(q, z|M_1, z_1; M_2, z_2) = \\
&= \frac{1}{\sqrt{1-\omega^2}} \left\{ \frac{d\sigma_1}{dM_1}\frac{d\sigma_2}{dM_2} + \frac{\sigma_2^2}{\delta_{f2}(1-\omega^2)} \left(\frac{\delta_{f1}}{\sigma_1} - \omega \frac{\delta_{f2}}{\sigma_2} \right) \frac{d\sigma_1}{dM_1}\frac{\partial\omega}{\partial M_2} + \right. \\
&+ \frac{\sigma_1^2}{\delta_{f1}(1-\omega^2)} \left(\frac{\delta_{f2}}{\sigma_2} - \omega \frac{\delta_{f1}}{\sigma_1} \right) \frac{\partial\omega}{\partial M_1}\frac{d\sigma_2}{dM_2} + \frac{\sigma_1^2\sigma_2^2}{\delta_{f1}\delta_{f2}} \frac{\partial^2\omega}{\partial M_1 \partial M_2} + \\
&+ \frac{\sigma_1^2\sigma_2^2}{\delta_{f1}\delta_{f2}(1-\omega^2)^2} \left[\omega(1-\omega^2) + (1+\omega^2)\frac{\delta_{f1}}{\sigma_1}\frac{\delta_{f2}}{\sigma_2} - \omega\left(\frac{\delta_{f1}^2}{\sigma_1^2} + \frac{\delta_{f2}^2}{\sigma_2^2} \right) \right] \times \\
&\left. \times \frac{\partial\omega}{\partial M_1}\frac{\partial\omega}{\partial M_2} \right\} \exp\left[- \frac{\omega^2\left(\frac{\delta_{f1}^2}{\sigma_1^2} + \frac{\delta_{f2}^2}{\sigma_2^2} \right) - 2\omega \frac{\delta_{f1}}{\sigma_1}\frac{\delta_{f2}}{\sigma_2}}{2(1-\omega^2)} \right] \left(\frac{d\sigma_1}{dM_1}\frac{d\sigma_2}{dM_2} \right)^{-1} .
\end{aligned}
\tag{7.14}
$$

This expression can be easily shown to be independent of the observation redshift z. A remark is now appropriate. Our formalism describes the halo distribution as a discrete point process. However, actual haloes are extended in size. This is why, as also seen in numerical simulations, for separation smaller than the typical Lagrangian radius of the halo, the correlation function abruptly reaches the value -1: a sort of 'exclusion principle' for extended haloes. Thus, we expect that the correlation function in eq.(7.14) can be a reliable description of halo clustering only for $q \gtrsim \mathrm{Max}(R_1, R_2)$. Another point concerns the use of finite mass resolution as in N-body simulations. The proper analytical correlation to compare with in that case is the integral of $\xi^L_{hh}\, n_{PS}(M_1)\, n_{PS}(M_2)$ over the specified mass interval, appropriately normalized.

Since the action of the window functions on the correlations is negligible for lags q much larger than the smoothing lengths, $q \gg R_1$ and $q \gg R_2$, for the normalized correlation we obtain $\omega(q) \simeq \xi_m(q)/\sigma_{M_1}\sigma_{M_2}$ [where $\xi_m(q)$ is the linear mass autocorrelation function] and, eventually, for the halo correlation

$$
\begin{aligned}
&\xi^L_{hh}(q, z|M_1, z_1; M_2, z_2) = \\
&\quad = b_1^L(z|M_1, z_1)\, b_1^L(z|M_2, z_2)\, \xi_m(q, z) + \frac{1}{2}\, b_2^L(z|M_1, z_1)\, b_2^L(z|M_2, z_2)\, \xi_m^2(q, z) + \cdots .
\end{aligned}
\tag{7.15}
$$

Explicitly, the first two bias parameters read

$$
b_1^L(z|M, z_f) = \frac{\delta_f(z, z_f)}{\sigma_M^2(z)} - \frac{1}{\delta_f(z, z_f)} = \frac{D(z_f)}{D(z)} \left[\frac{\delta_c}{D(z_f)^2\sigma_M^2} - \frac{1}{\delta_c} \right] , \tag{7.16}
$$

$$
b_2^L(z|M, z_f) = \frac{1}{\sigma_M^2(z)} \left[\frac{\delta_f^2(z, z_f)}{\sigma_M^2(z)} - 3 \right] = \frac{1}{D(z)^2\sigma_M^2} \left[\frac{\delta_c^2}{D(z_f)^2\sigma_M^2} - 3 \right] . \tag{7.17}
$$

These expressions for the bias factors generalize, in a sense, those concerning the clustering properties of dark matter haloes in Lagrangian space obtained by MW and Mo et al. (1996), with the relevant difference that we have obtained the bias factor from the behaviour of the halo two-point correlation function. Also relevant is the fact that, unlike MW, we obtained our Lagrangian correlation function without introducing any background scale $R_\circ$, which allows to extend its validity down to spatial separation comparable to the halo size $R \ll R_\circ$. A calculation of the leading

behaviour of the correlation deriving from equations (7.11) and (7.12) has been already attempted by Kashlinsky (1987) who, however, missed the contributions originated by the differentiation of ω with respect to M_1 and M_2, thereby obtaining an incomplete expression for b_1^L.

The halo correlation function in Lagrangian space, ξ_{hh}^L from eq.(7.14), with $M_1 = M_2 = M$ and $z_1 = z_2 \equiv z_f$, is shown in Figure 7.1 for two scale-free power-spectra, $P(k) \propto k^n$, with spectral index $n = -2$ and $n = -1$, in the Einstein-de Sitter case. The two-point function is calculated for various halo masses in units of the characteristic mass, M_*, defined so that $\sigma_{M_*} \equiv t_f = \delta_c/D(z_f)$, with top-hat filtering; the spatial dependence is shown as a function of the scaling variable q/R, which eliminates any residual redshift dependence. Also shown is the mass autocorrelation function ξ_m and an estimate of the Lagrangian halo two-point function obtained as $(b_1^L)^2\xi_m$, for $M \neq M_*$, and $(b_2^L)^2\xi_m^2/2$, for $M = M_*$, as in this case $b_1^L = 0$. Note that such an estimate of ξ_{hh}^L always provides an accurate fit to its analytical expression for separation a few times larger than the halo size. The characteristic behaviour of the halo correlation function for $M = M_*$, where the linear bias vanishes, is actually a peculiarity of the Lagrangian case (see also chapter 6 and MW). As we will see below, the Eulerian halo correlation function does not show such a drastic change of slope in the same mass range.

7.2.5 Peak-background split

In the previous paragraphs we computed number densities and correlation functions of haloes in Lagrangian space. However, after their identification, these haloes in embryo move following the gravitational field, modifying their original spatial distribution. One issue to address is how, for instance, the conditional halo number density per unit mass changes as a consequence of gravitational evolution. Furthermore, of interest is to quantify the evolution of clustering in terms of the halo correlation functions, or in terms of the halo-to-mass bias. Both problems can be dealt with by defining Eulerian halo counting fields, in the same spirit as we did for the Lagrangian case.

Essentially, our approach to the clustering evolution can be based on a generalization of the so-called peak-background split, first proposed by Bardeen et al. (1986), which basically consists in splitting the mass perturbations in fine-grained (*peak*) and coarse-grained (*background*) components[2]. As discussed in chapter 6, the underlying idea is to ascribe the collapse of objects on small scales to the high-frequency modes of the density field, while the action of large-scale structures on these non-linear condensations is due to the remaining low-frequency modes. At the linear level the resulting effect of these long wavelengths is simply modeled as a shift of the local background density.

In the spirit of the peak-background split, we define the linear density field smoothed on a given scale ϵ_M as consisting of two complementary and superimposed components, namely $\epsilon_M = \epsilon_{bg} + \epsilon_{pk}$. Adopting as window function the sharp k-space filter, we define as 'background' component the density contrast smoothed on the scale $R_\circ = 1/k_\circ$

$$\epsilon_{bg}(\mathbf{q}, z) \equiv \int \frac{d\mathbf{k}}{(2\pi)^3}\, \tilde{\epsilon}(\mathbf{k}, z)\, \Theta(k_\circ - k)\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{q}}\,. \tag{7.18}$$

The 'peak' component is instead obtained by smoothing the mass density fluctuation with the band-pass filter $\Theta(k_M - k)\,\Theta(k - k_\circ)$, namely

$$\epsilon_{pk}(\mathbf{q}, z) \equiv \int \frac{d\mathbf{k}}{(2\pi)^3}\, \tilde{\epsilon}(\mathbf{k}, z)\, \Theta(k_M - k)\, \Theta(k - k_\circ)\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{q}}\,, \tag{7.19}$$

[2]We are here making a rather liberal use of the word 'peak', to mean the fine-grained component of the linear density field.

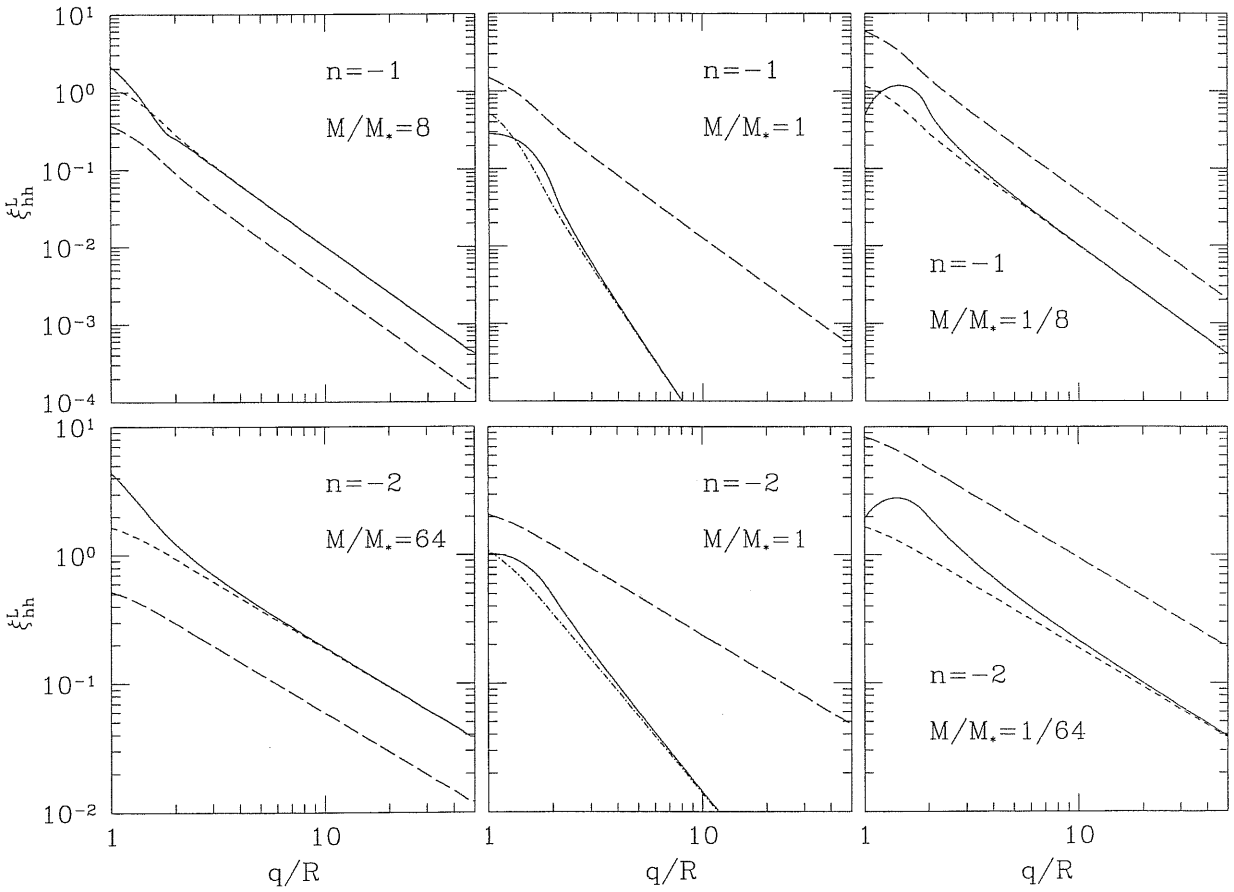


Figure 7.1: The exact Lagrangian halo correlation function in an Einstein-de Sitter universe (*solid lines*) is shown for scale-free models with spectral index $n = -1$ and $n = -2$, and for various masses. In each panel we set $M_1 = M_2 = M$ and $z_1 = z_2$. Results are plotted in terms of the scaling variables M/M_* and q/R (with R the top-hat radius corresponding to the halo mass M), which makes the resulting curves redshift independent. For comparison, the linear mass autocorrelation function smoothed on the halo scale is also shown with *long-dashed lines*. The *short-dashed lines* represent the linear bias approximation for the halo correlations: $(b_1^L)^2\xi_m$. In the central panels, where $b_1^L = 0$, the *dot-dashed lines* show, instead, the second-order approximation for ξ_{hh}^L. Each column contains panels that refer to the same mass variance ($\sigma_M^2/t_f^2 = 1/4,\ 1,\ 4$) and so to the same Lagrangian bias factors. Notice that, for separation a few times the halo size the first non-vanishing term of eq.(15) always gives an accurate approximation to the exact halo correlation. This implies that, for M_* objects, $\xi_{hh}^L \approx (b_2^L)^2\xi_m^2/2$.

where $k_M = 1/R_M$, with $M \propto \rho_b R_M^3$ and $M_\circ \propto \rho_b R_\circ^3$ the masses enclosed by the two filters. So, the peak component contains only modes with wavenumber in the interval $[k_\circ, k_M]$. Note that in the linear regime, with Gaussian initial conditions, the peak and background components are statistically independent, i.e.,

$$\langle \epsilon_{pk}(\mathbf{q}_1, z)\, \epsilon_{bg}(\mathbf{q}_2, z) \rangle = 0 \,, \tag{7.20}$$

for, by construction, the two fields do not share any common Fourier mode. To summarize: provided the collapsed object is described according to the spherical model, as in the PS theory, the peak field $\epsilon_{pk}(\mathbf{q}, z)$ can be thought as evolving in a local environment with effective mean density $\rho_b[1 + \epsilon_{bg}(\mathbf{q}, z)]$. This implies that the collapse condition can be written as $\epsilon_{pk}(\mathbf{q}, z) = \delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)$.

7.2.6 Eulerian halo counting field and bias

The previous analysis shows how the PS and the conditional Lagrangian mass functions can be obtained by averaging properly defined halo counting random fields. It is thus legitimate to explore the possibility of building up analogous counting processes in the Eulerian world. Our approach here will be based on the peak-background split technique described above.

Let us define the Eulerian counting field of haloes collapsed at redshift z_f and observed at z as

$$\begin{aligned} \mathcal{N}_h^E(\mathbf{q}, z|M, z_f) &\equiv \left[1 + \epsilon_{bg}(\mathbf{q}, z)\right] \mathcal{N}_h^L(\mathbf{q}, z|M, z_f) = \\ &= -\frac{2\rho_b}{M}\left[1 + \epsilon_{bg}(\mathbf{q}, z)\right] \frac{\partial}{\partial M}\Theta\left[\epsilon_{pk}(\mathbf{q}, z) - \left(\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)\right)\right] . \end{aligned} \tag{7.21}$$

The watchful reader might wonder about our use of the Lagrangian variable $\mathbf{q}$ within the Eulerian framework, however, in linear theory, $\mathbf{x} = \mathbf{q}$. Once again, the redshift z must be thought of as the redshift the sampled objects have at the epoch of observation. Noteworthy, eq.(7.21) is fully consistent with the analysis in Cole and Kaiser (1989). Most importantly, our treatment allows for a local description. Let us stress here that the factor $(1 + \epsilon_{bg})$, connecting the Eulerian to the Lagrangian counting field, simply comes from mass conservation in Eulerian space [see also §6.4.3, §7.3.1 and, in particular, eq.(7.38)]; this point has been discussed in more detail by Kofman et al. (1994).

Now consider the integral stochastic process

$$\int_M^\infty dM' M' \mathcal{N}_h^E(\mathbf{q}, z|M', z_f) = 2\,\rho_b\,[1 + \epsilon_{bg}(\mathbf{q}, z)]\,\Theta[\epsilon_M(\mathbf{q}, z) - \delta_f(z, z_f)] \,; \tag{7.22}$$

this represents the fraction of mass, in the unit Eulerian comoving volume centered in $\mathbf{q}$, which at redshift z_f will form haloes more massive than M. For $M_\circ \to M$, $\epsilon_{bg} \to \epsilon_M$ and the above relation coincides (up to the usual fudge factor of 2, having no effect on correlations) with the *weighted bias* model of Catelan et al. (1994). Thus, the weighted bias is just the Eulerian version, within linear theory, of Kaiser (1984) bias model.

Of course, further specifications could be added to our Eulerian counting field. For instance, we might ask that the background scale has not yet collapsed by the epoch z_f; in such a case we should multiply the above stochastic process by the factor $\Theta[\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)]$. Extra details of this kind would however make negligible changes to our final results, provided $\sigma_M^2 \gg \sigma_\circ^2$ (see, e.g., §8.2).

Like in the Lagrangian case, to calculate the mean halo number density per unit mass, one needs to ensemble average $\mathcal{N}_h^E(\mathbf{q}, z|M, z_f)$. Let us analyse this operation in more detail. Because of the way the Eulerian counting process has been defined, it is clear that $\mathcal{N}_h^E(\mathbf{q}, z|M, z_f)$ depends on two

random fields, specifically ϵ_{bg} and ϵ_{pk}. So, the ensemble average $\langle \mathcal{N}_h^E \rangle$ can be interpreted as a double average over these fields, i.e. $\langle \mathcal{N}_h^E \rangle \equiv \langle \langle \mathcal{N}_h^E \rangle_{\epsilon_{pk}} \rangle_{\epsilon_{bg}}$. The statistics of the field $\mathcal{N}_h^E$ can be described in terms of n-th order correlation functions, $\langle \langle \mathcal{N}_h^E(\mathbf{q}_1) \cdots \mathcal{N}_h^E(\mathbf{q}_n) \rangle_{\epsilon_{pk}} \rangle_{\epsilon_{bg}}$. The exact calculation of these quantities is rather difficult. However, because of the short-scale coherence of the peak field, implied by the 'infrared' cutoff at $k_\circ$, its covariance $\langle \epsilon_{pk}(\mathbf{q}_i)\, \epsilon_{pk}(\mathbf{q}_i + \mathbf{r}) \rangle_{\epsilon_{pk}}$ vanishes whenever $r \gg R_\circ$, so that we can simplify the general halo correlations above as $\langle \langle \mathcal{N}_h^E(\mathbf{q}_1) \cdots \mathcal{N}_h^E(\mathbf{q}_n) \rangle_{\epsilon_{pk}} \rangle_{\epsilon_{bg}} \approx \langle \langle \mathcal{N}_h^E(\mathbf{q}_1) \rangle_{\epsilon_{pk}} \cdots \langle \mathcal{N}^E(\mathbf{q}_n) \rangle_{\epsilon_{pk}} \rangle_{\epsilon_{bg}}$, provided we consider sets of points $\mathbf{q}_i$, $i = 1, \ldots, n$, with relative separation $r_{ij} \equiv |\mathbf{q}_i - \mathbf{q}_j| \gg R_\circ$. Therefore, with the purpose of calculating the mean Eulerian halo number density per unit mass and Eulerian halo correlations, we can make the replacement $\mathcal{N}_h^E \to \langle \mathcal{N}_h^E \rangle_{\epsilon_{pk}} \equiv N_h^E$, with only negligible loss of accuracy. According to the definition of $\mathcal{N}_h^E$ in eq.(7.21), the latter ensemble average gives

$$N_h^E(\mathbf{q}, z|M, z_f) = \frac{1}{\sqrt{2\pi}} \frac{\rho_b}{M} \left[1 + \epsilon_{bg}(\mathbf{q}, z)\right] \frac{\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)}{[\sigma_M^2(z) - \sigma_\circ^2(z)]^{3/2}} \times \\ \times \exp\left\{ -\frac{[\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)]^2}{2\,[\sigma_M^2(z) - \sigma_\circ^2(z)]} \right\} \left| \frac{d\sigma_M^2(z)}{dM} \right| , \tag{7.23}$$

which, having averaged over the fine-grained mass fluctuations, represents a sort of coarse-grained halo counting field. Notice that the fine-grained ensemble average has replaced the original step-function operator of eq.(7.21) by a smoother function, which can then be consistently expanded in series of the background field, as shown below.

Let us stress that the expression in eq.(7.23) is just the Eulerian analog of eq.(10) in MW, but the field ϵ_{bg} is here a true random field, and so is the process N_h^E. The knowledge of N_h^E allows us to define the Eulerian halo number density fluctuation as

$$\delta_h^E(\mathbf{q}, z|M, z_f) \equiv \frac{N_h^E(\mathbf{q}, z|M, z_f) - \langle N_h^E(\mathbf{q}, z|M, z_f) \rangle_{\epsilon_{bg}}}{\langle N_h^E(\mathbf{q}, z|M, z_f) \rangle_{\epsilon_{bg}}} \equiv \mathrm{b}^E(\mathbf{q}, z|M, z_f)\, \epsilon_{bg}(\mathbf{q}, z) , \tag{7.24}$$

where we introduced the Eulerian 'bias field' $\mathrm{b}^E(\mathbf{q}, z|M, z_f)$. The second equality in the above equation does not mean that the Eulerian fluctuation field δ_h^E is proportional to the background density field ϵ_{bg}. In fact, b^E in general depends upon ϵ_{bg} itself. Its functional dependence can be understood by expanding $N_h^E(\mathbf{q}, z|M, z_f)$ in powers of ϵ_{bg} to obtain

$$\delta_h^E(\mathbf{q},z|M, z_f) = b_1^E(z|M, z_f)\, \epsilon_{bg}(\mathbf{q}, z) + \frac{1}{2}\, b_2^E(z|M, z_f)\, \epsilon_{bg}^2(\mathbf{q}, z) + \cdots = \\ = \left[1 + b_1^L(z|M, z_f)\right] \epsilon_{bg}(\mathbf{q}, z) + \frac{1}{2} \left[b_2^L(z|M, z_f) + 2\, b_1^L(z|M, z_f)\right] \epsilon_{bg}^2(\mathbf{q}, z) + \ldots , \tag{7.25}$$

where, for $\sigma_M^2 \gg \sigma_\circ^2$, the first and second-order Lagrangian bias factors b_1^L and b_2^L are those of eq.(7.16) and eq.(7.17), respectively. Accounting for the transformation from the Lagrangian to the Eulerian distribution (e. g. Kofmann et al. 1992), one has $\langle N_h^E(\mathbf{q}, z|M, z_f) \rangle_{\epsilon_{bg}} = n_{PS}(z|M, z_f)$. It can be useful to give explicit expressions for the first two Eulerian bias parameters of linear theory

$$b_1^E(z|M, z_f) = 1 + \frac{D(z_f)}{D(z)} \left[\frac{\delta_c}{D(z_f)^2 \sigma_M^2} - \frac{1}{\delta_c} \right] , \tag{7.26}$$

$$b_2^E(z|M, z_f) = \frac{1}{D(z)^2 \sigma_M^2} \left[\frac{\delta_c^2}{D(z_f)^2 \sigma_M^2} - 3 \right] + \frac{2D(z_f)}{D(z)} \left[\frac{\delta_c}{D(z_f)^2 \sigma_M^2} - \frac{1}{\delta_c} \right] . \tag{7.27}$$

The set of linear theory Eulerian bias factors $b_\ell^E(z)$ can be obtained from the Lagrangian ones according to the general rule

$$b_\ell^E = \ell\, b_{\ell-1}^L + b_\ell^L \ , \tag{7.28}$$

with $b_{\ell=0}^L \equiv 1$.

The same method can be applied to the Lagrangian expression, in the sense that we can obtain, similarly,

$$N_h^L(\mathbf{q}, z|M, z_f) = \frac{1}{\sqrt{2\pi}} \frac{\rho_b}{M} \frac{\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)}{[\sigma_M^2(z) - \sigma_\circ^2(z)]^{3/2}} \exp\left\{ -\frac{[\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)]^2}{2\,[\sigma_M^2(z) - \sigma_\circ^2(z)]} \right\} \left| \frac{d\sigma_M^2(z)}{dM} \right| . \tag{7.29}$$

One has, exactly, $\langle N_h^L(\mathbf{q}, z|M, z_f)\rangle_{\epsilon_{bg}} = \langle \mathcal{N}_h^L(\mathbf{q}, z|M, z_f)\rangle = n_{PS}(z|M, z_f)$. By expanding the coarse-grained Lagrangian counting field $N_h^L(\mathbf{q}, z|M, z_f)$ we can define Lagrangian bias factors at any order. For $\sigma_M^2 \gg \sigma_\circ^2$ these turn out to be identical to those deriving from the expansion of the halo correlation in Lagrangian space, eq.(7.14). This suggests, however, that these bias factors can be used to describe halo clustering on distances $r > R$, without any further restriction introduced by the background scale $R_\circ$.

The very fact that, for practical purposes, we can replace the exact operator $\mathcal{N}_h^E$ by the locally averaged one N_h^E demonstrates that the MW treatment can be made self-consistent, provided their small-scale density field is replaced by the peak field, and the value of the threshold is modified accordingly. Most importantly, our local averaging procedure implies that, up to the scale $R_\circ$, we are indeed correctly accounting for the cloud-in-cloud problem. This is because at each point $\mathbf{q}$, characterized by a random value of the background field $\epsilon_{bg}(\mathbf{q})$, the coarse-grained stochastic process $N_h^E(\mathbf{q}, z|M, z_f)$ (and its Lagrangian equivalent) actually represents the local mean mass function, for which the cloud-in-cloud problem is exactly solved in terms of first passage 'times' across the local barrier $\delta_f(z, z_f) - \epsilon_{bg}(\mathbf{q}, z)$, with initial condition $\epsilon_{pk}(\mathbf{q}, z) = 0$ at $R = R_\circ$. Therefore, to the aim of calculating correlations on lags $r \gg R_\circ$, we can safely state that our coarse-grained halo counting fields are unaffected by the cloud-in-cloud problem.

The shift by 1 of the linear bias factor, here implied by the transformation from the Lagrangian to the Eulerian world, was also noticed in the weighted bias approach by Catelan et al. [1994; their eq.(21)], where an underlying Lognormal distribution was assumed to avoid negative-mass events.

The above expression for $b_1^E(z|M, z_f)$ coincides with the formula by MW [their eq.(20)], who, however, only presented it for $z = 0$. As noticed by MW, an important feature of this linear bias is that it predicts that large-mass objects (actually those characterized by $\sigma_M < t_f$) are biased with respect to the mass ($b_1^E > 1$), while small-mass ones ($\sigma_M > t_f$) are anti-biased ($b_1^E < 1$). Haloes with mass close to the characteristic one, M_*, have non-vanishing linear bias, unlike the Lagrangian case. As we will see in §7.3.1, this one-to-one classification of biased and anti-biased objects according to their mass is no longer valid in the non-linear regime, as the shear field at the Lagrangian location of the halo also contributes to the determination of its Eulerian bias factor.

The effect of merging can be easily accommodated into this scheme. In the real universe, haloes undergo merging at some finite rate, which can be suitably modeled (e.g. Lacey & Cole 1993). As mentioned above, in the simple PS theory such a rate is actually infinite, for infinite mass resolution, implying that only haloes 'just formed' can survive, so that $z_f = z$. So, if one gives up singling out the individuality of haloes selected at different threshold, i.e. with different formation redshifts $z_f \geq z$, one immediately obtains (e.g. Matarrese et al. 1997)

$$b_1^E(z|M) = 1 + \left[\frac{\delta_c}{D(z)^2 \sigma_M^2} - \frac{1}{\delta_c} \right] , \tag{7.30}$$

which implies a quadratic redshift dependence in the Einstein-de Sitter universe,

$$b_1^E(z|M) = 1 + \left[\frac{\delta_c(1+z)^2}{\sigma_M^2} - \frac{1}{\delta_c}\right] . \tag{7.31}$$

The latter form coincides with the result by Cole and Kaiser (1989) [their eq.(6)], who however define the bias factor of haloes at redshift z with respect to the mass fluctuation at the present time, which then scales the latter expression by a factor $(1+z)^{-1}$.

On the other hand, for fixed z_f and varying z, i.e. for objects which survived till the epoch z after their birth at z_f, the Eulerian bias of eq.(7.26) gets a completely different evolution, namely

$$b_1^E(z|M) = 1 + \frac{D(z_f)}{D(z)}\left[b_1^E(z_f|M) - 1\right] , \tag{7.32}$$

which implies a linear redshift dependence in the Einstein-de Sitter case,

$$b_1^E(z|M) = 1 + \frac{1+z}{1+z_f}\left[b_1^E(z_f|M) - 1\right] . \tag{7.33}$$

The latter form coincides with that obtained by Dekel (1986), Dekel and Rees (1987), Nusser and Davis (1994) and Fry (1996). This relation can be relevant for galaxies which were conserved in number after their formation, i.e. that maintained their individuality even after their hosting haloes merged.

It is trivial, at this point, to obtain the Eulerian halo-halo correlation function within our approximations. For lags $r \gg R_\circ$, one has

$$\begin{aligned}\xi_{hh}^E(r,z|M_1,z_1;M_2,z_2) = \\ =b_1^E(z|M_1,z_1)\, b_1^E(z|M_2,z_2)\, \xi_m(r,z) + \frac{1}{2}\, b_2^E(z|M_1,z_1)\, b_2^E(z|M_2,z_2)\, \xi_m^2(r,z) + \cdots .\end{aligned} \tag{7.34}$$

The main limitation of this formula, however, is that it only provides a link between the Eulerian halo correlation function and that of the mass within linear theory. What one would really need, instead, is a similar relation in the fully non-linear regime. This problem will be solved in the next section.

7.3 Halo counting and non-linear dynamics: Eulerian description

One can derive a general expression for the Eulerian halo-to-mass bias by integrating the continuity equations for the mass and for the halo number density, assuming that haloes move according to the velocity field determined by the matter. The Lagrangian analysis carried out in the previous section is crucial to the present purposes, since it allows for the natural initial conditions necessary to integrate the Eulerian equations. As we will show below, the Eulerian halo-to-mass bias obtained in such a way holds for any cosmology and in any dynamical regime. This turns out to be a remarkable generalization of the biasing proposed by Cole and Kaiser (1989) and MW.

7.3.1 Eulerian bias from dynamical fluid equations

Let us consider the mass density fluctuation field $\delta(\mathbf{x},\tau(z)) = \delta(\mathbf{x},z)$ which obeys the mass conservation equation

$$\frac{d\delta}{d\tau} = -(1+\delta)\nabla\cdot\mathbf{v} , \tag{7.35}$$

where τ is the conformal time of the background cosmology and the differential operator $d/d\tau \equiv \partial/\partial\tau + \mathbf{v}\cdot\nabla$ is the convective derivative. The peculiar velocity field $\mathbf{v} \equiv d\mathbf{x}/d\tau$ satisfies the Euler equation $d\mathbf{v}/d\tau + (a'/a)\mathbf{v} = -\nabla\phi_g$, where a is the expansion factor and a prime denotes differentiation with respect to τ. For later convenience, let us also define the scaled peculiar velocity $\mathbf{u} \equiv d\mathbf{x}/dD = \mathbf{v}/D'$. The peculiar gravitational potential ϕ_g is determined by the matter distribution via the cosmological Poisson equation $\nabla^2\phi_g = 4\pi G a^2 \rho_b(\tau)\delta$, where $\rho_b(\tau)$ is the background mean density at time τ. If we assume that our halo population of mass M and formation redshift z_f is conserved in time, and evolves exclusively under the influence of gravity, its number density fluctuation $\delta_h(\mathbf{x},z) = \delta_h(\mathbf{x},z|M,z_f)$ has to satisfy the continuity equation (e.g. Fry 1996)

$$\frac{d\delta_h}{d\tau} = -(1+\delta_h)\nabla\cdot\mathbf{v}\ , \tag{7.36}$$

from which, eliminating the expansion scalar $\nabla\cdot\mathbf{v}$, we obtain

$$\frac{d\ln(1+\delta_h)}{d\tau} = \frac{d\ln(1+\delta)}{d\tau}\ . \tag{7.37}$$

This equation can be integrated exactly in terms of Lagrangian quantities, and the solution reads

$$1+\delta_h(\mathbf{x},z) = \left[1+\delta_h(\mathbf{q})\right]\left[1+\delta(\mathbf{x},z)\right] \tag{7.38}$$

(see also the discussion in Peacock & Dodds 1994), where $\mathbf{q}$ is the Lagrangian position corresponding to the Eulerian one via $\mathbf{x}(\mathbf{q},z) = \mathbf{q}+\mathbf{S}(\mathbf{q},z)$, with $\mathbf{S}(\mathbf{q},z)$ the displacement vector. In eq.(7.38), by $\delta_h(\mathbf{q}) = \delta_h(\mathbf{q}|M,z_f)$ we mean the Lagrangian halo density fluctuation, whereas, for simplicity, we assumed that $\lim_{z\to\infty}\delta[\mathbf{x}(\mathbf{q},z),z] \equiv \delta(\mathbf{q}) = 0$, i.e. that the mass was initially uniformly distributed (this amounts to taking purely growing-mode initial perturbations). Defining the Eulerian halo bias field through

$$\delta_h(\mathbf{x},z) \equiv \mathrm{b}^E(\mathbf{x},z)\,\delta(\mathbf{x},z)\ , \tag{7.39}$$

we end up with the *exact* relation

$$\mathrm{b}^E(\mathbf{x},z) = 1 + \frac{1+\delta(\mathbf{x},z)}{\delta(\mathbf{x},z)}\,\delta_h(\mathbf{q})\ . \tag{7.40}$$

The key problem now is how to calculate the field $\delta_h(\mathbf{q})$. We cannot simply take the Lagrangian halo distribution as $\delta_h(\mathbf{q}) = \mathrm{b}^L(\mathbf{q})\delta(\mathbf{q})$, because $\delta(\mathbf{q}) = 0$; thus, we are forced to adopt some limiting procedure. To specify the Lagrangian halo distribution, we can take advantage of the results of §7.2. By definition, the Lagrangian distribution of nascent haloes of mass M and formation epoch z_f is given by

$$\delta_h(\mathbf{q}|M,z_f) \equiv \lim_{z\to\infty}\mathrm{b}^E(\mathbf{q},z|M,z_f)\,\epsilon_{bg}(\mathbf{q},z) \equiv \mathrm{b}_\circ^L(\mathbf{q}|M,z_f)\,\epsilon_\circ(\mathbf{q})\ , \tag{7.41}$$

where $\mathrm{b}_\circ^L(\mathbf{q}|M,z_f) \equiv \mathrm{b}^L(\mathbf{q},z{=}0|M,z_f)$ is the Lagrangian halo bias field. Once again, let us stress that the second equality in the latter equation does not mean at all that $\delta_h(\mathbf{q})$ is proportional to $\epsilon_\circ(\mathbf{q})$. In fact, b^L is in general a functional of the background density field. To understand the above equation, one has to remember that, at sufficiently early times, the expression for the Eulerian bias field obtained in linear theory becomes exact (as linear theory gets more and more accurate), and $\delta(\mathbf{x},z) \to \epsilon_{bg}(\mathbf{x},z) = D(z)\,\epsilon_\circ(\mathbf{q})$, as $z\to\infty$. Because of our normalization of D, here $\epsilon_\circ(\mathbf{q})$ is the mass density fluctuation linearly extrapolated to the present time and filtered on the background scale $R_\circ$. The background smoothing scale $R_\circ$ actually has a twofold role in

our analysis. In the linear theory approach of §7.2 it was introduced and required to be much larger than the halo size, in order to get a self-consistent definition of halo counting fields, with the desirable feature of being free of the cloud-in-cloud problem. In the present non-linear analysis, however, the background mass scale must be chosen large enough to ensure that the halo velocity field coincides with the one of the matter.

The Lagrangian density contrast of haloes identified by a PS-type algorithm can be obtained from eq.(7.29) as $\delta_h(\mathbf{q}|M,z_f) = N_h^L(\mathbf{q},z|M,z_f)/n_{PS}(z|M,z_f) - 1$, which leads to

$$1+\delta_h(\mathbf{q}|M,z_f) = \left[1-\frac{D(z_f)\epsilon_{\rm o}(\mathbf{q})}{\delta_c}\right]\left(1-\frac{\sigma_{\rm o}^2}{\sigma_M^2}\right)^{-3/2}\times \\ \times\exp\left[-\frac{\epsilon_{\rm o}(\mathbf{q})^2-2\epsilon_{\rm o}(\mathbf{q})\delta_c/D(z_f)+\delta_c^2\sigma_{\rm o}^2/D(z_f)^2\sigma_M^2}{2(\sigma_M^2-\sigma_{\rm o}^2)}\right]. \tag{7.42}$$

For $\sigma_M^2 \gg \sigma_{\rm o}^2$ this expression simplifies to

$$\delta_h(\mathbf{q}|M,z_f) = \left[1-\frac{D(z_f)\epsilon_{\rm o}(\mathbf{q})}{\delta_c}\right]\exp\left[-\frac{\epsilon_{\rm o}(\mathbf{q})^2-2\epsilon_{\rm o}(\mathbf{q})\delta_c/D(z_f)}{2\sigma_M^2}\right]-1 = \\ =\sum_{\ell=1}^{\infty}\frac{b_{{\rm o}\ell}^L(M,z_f)}{\ell!}\,\epsilon_{\rm o}(\mathbf{q})^\ell\,. \tag{7.43}$$

The first four Lagrangian bias factors evaluated at $z=0$ read

$$b_{\rm o1}^L(M,z_f) = D(z_f)\left[\frac{\delta_c}{D(z_f)^2\sigma_M^2}-\frac{1}{\delta_c}\right], \tag{7.44}$$

$$b_{\rm o2}^L(M,z_f) = \frac{1}{\sigma_M^2}\left[\frac{\delta_c^2}{D(z_f)^2\sigma_M^2}-3\right], \tag{7.45}$$

$$b_{\rm o3}^L(M,z_f) = \frac{D(z_f)}{\sigma_M^2}\left[\frac{\delta_c^3}{D(z_f)^4\sigma_M^4}-\frac{6\,\delta_c}{D(z_f)^2\sigma_M^2}+\frac{3}{\delta_c}\right], \tag{7.46}$$

$$b_{\rm o4}^L(M,z_f) = \frac{1}{\sigma_M^4}\left[\frac{\delta_c^4}{D(z_f)^4\sigma_M^4}-\frac{10\,\delta_c^2}{D(z_f)^2\sigma_M^2}+15\right]. \tag{7.47}$$

Note that, in full generality, $b_{{\rm o}\ell}^L(M,z_f) = D(z)^\ell\, b_\ell^L(z|M,z_f)$. Adding the further requirement that the local fluctuation on the background scale $R_{\rm o}$ has not collapsed yet by the time of halo formation would make our object number density semi-positive definite both at the Lagrangian and Eulerian level, i.e. $\delta_h \geq -1$, at any time, only provided $\epsilon_{\rm o} \leq t_f$.

The general expression for the Lagrangian halo density contrast of eq.(7.42) is plotted in Figure 7.2 as a function of the background density field, for different halo masses. In the high-mass case positive mass fluctuations typically correspond to positive values of the Lagrangian halo density contrast (and viceversa), while the trend is the opposite at low masses. The transition, once again, corresponds to halo masses around M_*, in which case positive values of δ_h only occur in regions with background density close to the mean. Also shown are two approximations to the Lagrangian halo density contrast obtained by expanding eq.(7.42) up to first and second order in the background

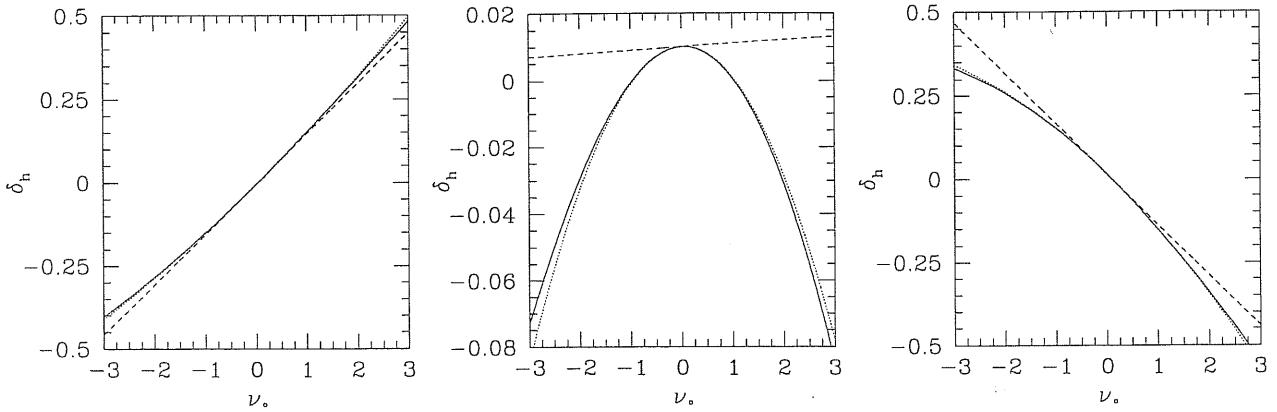


Figure 7.2: The exact expression for the Lagrangian halo density contrast of eq.(42) (*solid lines*) is plotted as a function of $\nu_o \equiv \epsilon_o/\sigma_o$. The three panels refer to values of the halo masses such that $\sigma_M^2/t_f^2 = 1/4$ (*left*), $\sigma_M^2/t_f^2 = 1$ (*centre*) and $\sigma_M^2/t_f^2 = 4$ (*right*), with $t_f = 1.69/D(z_f)$. The background mass scale is chosen so that $\sigma_o^2 = 0.01\,\sigma_M^2$. Also plotted are two estimates of δ_h obtained by expanding the r.h.s. of eq. (7.42) up to first (*dashed lines*) and second order (*dotted lines*) in ϵ_o. Note that, because of our choice of variables, all the curves are independent both of z_f and Ω_o.

field. Except for halo masses near M_*, where a quadratic bias is clearly needed, a linear Lagrangian bias generally provides an accurate fit to $\delta_h(\mathbf{q})$ within the bulk of the ϵ_o distribution.

The Eulerian bias field finally reads

$$\mathrm{b}^E(\mathbf{x}, z|M, z_f) = 1 + \frac{1+\delta(\mathbf{x},z)}{\delta(\mathbf{x},z)}\,\mathrm{b}_o^L(\mathbf{q}|M, z_f)\,\epsilon_o(\mathbf{q}) \ . \tag{7.48}$$

It can be seen that, in the linear regime, where $\delta(\mathbf{x},z) \approx D(z)\,\epsilon_o(\mathbf{q}) \ll 1$, the expression for b^E in MW [i.e. our eq.(7.26)] is recovered, provided $\mathrm{b}_o^L(\mathbf{q}|M,z_f)$ is replaced by its first-order approximation $b_{o1}^L(M,z_f)$. It is however important to realize that the exact expression in eq.(7.48) implies that the Eulerian bias field of dark matter haloes $\mathrm{b}^E(\mathbf{x},z|M,z_f)$ is both non-linear, in that it depends on $\delta(\mathbf{x})$, and non-local, as it depends on the Lagrangian position $\mathbf{q}$ through $\mathrm{b}_o^L(\mathbf{q}|M,z_f)\,\epsilon_o(\mathbf{q})$, simply because of inertia.

Our exact expression for the Eulerian halo bias, eq.(7.48), generally involves quantitative corrections to the MW approximate bias formula. In some cases, however, the MW relation may even fail to predict the correct qualitative behaviour of the halo-to-mass bias. This is the case, in fact, of those initially underdense fluid elements in Lagrangian space, $\epsilon_o(\mathbf{q}) < 0$, which, after an initial expansion phase, turn around to undergo a phase of local compression, so that the corresponding Eulerian fluid element eventually becomes overdense, $\delta(\mathbf{x}(\mathbf{q},z),z) > 0$, and collapses. This is a well-known non-linear effect caused by the shear component of the velocity field, i.e. by the tidal force of the surrounding matter. For Gaussian initial conditions, the occurrence of such an event can be estimated by the Zel'dovich approximation as affecting 42% of the overall Lagrangian volume (Doroshkevich 1970; Shandarin & Zel'dovich 1984); Hui and Bertschinger (1996), using a different approximation, estimated this effect as affecting at least 39% of the total Lagrangian volume. In all such cases the MW formula would incorrectly predict bias instead of anti-bias, for halo masses $M > M_*$, and anti-bias instead of bias, for $M < M_*$. The problem may be generally less severe than the above heuristic argument would suggest, as, at a fixed epoch z, only a smaller fraction of such Lagrangian patches have already turned around from their initial expansion; this is even more true for large mass haloes, which probe the underlying mass distribution in a more linear regime,

where the MW formula gets closer to the exact one. As a tentative conclusion let us say that one should be careful in applying the linear MW bias [i.e. our eq.(7.26)] at the Eulerian level especially in connection with halo masses much smaller than M_*.

The most important application of eq.(7.48) is that it allows to generate Eulerian maps of the local comoving halo number density per unit mass, $n_{PS}(M,z_f)[1+\delta_h(\mathbf{x},z|M,z_f)]$, given the non-linearly evolved mass density contrast $\delta(\mathbf{x},z)$ (with Lagrangian resolution $R_\circ$) and the corresponding Lagrangian mass and halo density fluctuation fields, $\epsilon_\circ(\mathbf{q})$ and $\delta_h(\mathbf{q}|M,z_f)$, respectively.

In order to account for halo merging, at this level, one just has to assume a suitable link between the formation and observation epochs, which, in the simple PS theory amounts to the replacement $z_f \to z$, in the above expressions for $\mathrm{b}_\circ^L$.

Recalling that mass conservation can be recast in terms of the Jacobian determinant $J \equiv \|\partial\mathbf{x}/\partial\mathbf{q}\|$ of the mapping $\mathbf{q}\to\mathbf{x}$, as $1+\delta[\mathbf{x}(\mathbf{q},z),z] = J(\mathbf{q},z)^{-1}$, one finds the exact relation

$$\mathrm{b}^E\left(\mathbf{x}(\mathbf{q},z),z|M,z_f\right) = 1+[1-J(\mathbf{q},z)]^{-1}\;\mathrm{b}_\circ^L(\mathbf{q}|M,z_f)\,\epsilon_\circ(\mathbf{q})\;. \tag{7.49}$$

It can be useful to illustrate the meaning of this expression by considering various approximations to the evolution of the mass density in the non-linear regime, i.e. to the particle trajectories $\mathbf{x}(\mathbf{q},z)$. Such approximation schemes should be thought of, not as self-consistent perturbative approaches to the actual dynamics, but as 'clever tricks' able to catch some aspects of the true dynamics, at least in the mildly non-linear regime. A detailed and systematic comparison of the performance of several approximations for different choices of the initial conditions has been made by Sathyaprakash et al. (1995; see also chapter 2).

Zel'dovich approximation

In the Zel'dovich approximation (ZEL; Zel'dovich 1970) the displacement vector is $\mathbf{S} = -D\,\nabla_\mathbf{q}\varphi_\circ(\mathbf{q})$, where $\varphi(\mathbf{q})$ is the linear peculiar gravitational potential, suitably rescaled so that $\nabla^2_\mathbf{q}\varphi_\circ(\mathbf{q}) = \epsilon_\circ(\mathbf{q})$. Indicating by $\lambda_\alpha(\mathbf{q})$ ($\alpha=1,2,3$) the eigenvalues of the deformation tensor $\partial^2\varphi_\circ(\mathbf{q})/\partial q_\alpha\,\partial q_\beta$, we obtain for the Eulerian bias field

$$\begin{aligned}\mathrm{b}^E_{\mathrm{ZEL}}(\mathbf{x}(\mathbf{q},z),z|M,z_f) &= 1+\frac{\epsilon_\circ(\mathbf{q})\,\mathrm{b}_\circ^L(\mathbf{q}|M,z_f)}{1-\prod_{\alpha=1}^3\left[1-D(z)\,\lambda_\alpha(\mathbf{q})\right]} = \\ &= 1+\frac{\mathrm{b}_\circ^L(\mathbf{q}|M,z_f)}{D(z)}\left[1-D(z)\frac{\mu_2(\mathbf{q})}{\mu_1(\mathbf{q})}+D(z)^2\frac{\mu_3(\mathbf{q})}{\mu_1(\mathbf{q})}\right]^{-1}\;.\end{aligned} \tag{7.50}$$

Here $\mu_1(\mathbf{q}) \equiv \lambda_1+\lambda_2+\lambda_3 = \epsilon_\circ(\mathbf{q})$, $\mu_2 \equiv \lambda_1\lambda_2+\lambda_1\lambda_3+\lambda_2\lambda_3$ and $\mu_3 \equiv \lambda_1\lambda_2\lambda_3$ are the three invariants of the deformation tensor. If one makes the further approximation of replacing the Lagrangian bias by its first-order estimate of eq.(7.44), it can be checked that the expression of $\mathrm{b}^E_{\mathrm{ZEL}}$ coincides with the MW result, both at sufficiently early times ($D \ll 1$) and in the case of one-dimensional perturbations, for which $\mu_2 = 0 = \mu_3$ and the Zel'dovich approximation represents the exact solution to the non-linear dynamics.

It is important to stress that we are not forced to take the above result as a perturbative expression. An accurate approximation to the Eulerian bias field would in fact consist in evolving the mass according to the truncated (on the scale $M_\circ$) Zel'dovich approximation (Kofman 1991; Kofman et al. 1992; Coles, Melott & Shandarin 1993) and using the full expression for the Lagrangian bias. Being a random field, the Eulerian halo bias is completely characterized by a probability density functional, thus for a given mass M and formation redshift z_f there exists a whole distribution of possible values of b^E, related to the particular environment where the object forms as well as to the initial conditions leading to that site. Starting from the ZEL expression in eq.(7.50) one could

explicitly obtain the probability distribution function $p(\mathrm{b}^E_{\rm ZEL})\, d\,\mathrm{b}^E_{\rm ZEL}$, by integrating over the joint distribution of the invariants μ_α [an expression for the latter is given in Kofman et al. (1994)]. These specific applications of our results will be discussed in chapter 8.

Equation (7.50) has the merit of clearly displaying the intrinsic non-locality of the Eulerian bias. Only in some simplified cases there exists a local mapping between b^E and δ, so that an expansion of the halo density contrast in a hierarchy of Eulerian bias factors, b^E_1, b^E_2, etc., makes sense. One example is provided by the linear theory approach of §7.2.6; further examples are given below.

Frozen-flow approximation

According to the *frozen-flow* approximation (FFA; Matarrese et al. 1992) the Eulerian density field can be written as

$$1+\delta(\mathbf{x}(\mathbf{q},z),z) = \exp\int_0^{D(z)} d\widetilde{D}\,\epsilon_{\rm o}[\mathbf{x}(\mathbf{q},\widetilde{D})]\,, \tag{7.51}$$

where the integral is calculated along the trajectory of the fluid element. Note that, since in the frozen-flow approximation shell-crossing never occurs, the mapping $\mathbf{q}\to\mathbf{x}$ can be inverted at any time. The solution (7.51) might be replaced in eq.(7.48) to obtain a non-local expression for the FFA bias parameter. However, we can make a further step by noting that, for Lagrangian points $\mathbf{q}_*$ corresponding to local extrema of the initial gravitational potential, $\nabla_{\mathbf{q}}\varphi_{\rm o}(\mathbf{q}_*)=\mathbf{0}$, FFA predicts $\mathbf{x}_*=\mathbf{x}(\mathbf{q}_*,z)=\mathbf{q}_*$, and

$$1+\delta(\mathbf{x}_*,z) = \exp\left[D(z)\,\epsilon_{\rm o}(\mathbf{x}_*)\right]\,, \tag{7.52}$$

One can speculate that such points represent the preferential sites for the formation of massive haloes, which could be associated to clusters of galaxies, and use this approximate expression to obtain

$$\mathrm{b}^E_{\rm FFA}(\mathbf{x}_*,z|M,z_f) \approx 1+\frac{1+\delta(\mathbf{x}_*,z)}{\delta(\mathbf{x}_*,z)}\,\ln\left[1+\delta(\mathbf{x}_*,z)\right]\frac{\mathrm{b}^L_{\rm o}(\mathbf{x}_*|M,z_f)}{D(z)}\,. \tag{7.53}$$

Expanding this expression in powers of δ, to first-order we recover the MW expression, eq.(7.26), while to second-order, we obtain

$$b^E_{\rm 2FFA}(z|M,z_f) = \frac{1}{D(z)^2\sigma_M^2}\left[\frac{\delta_c^2}{D(z_f)^2\sigma_M^2}-3\right]+\frac{D(z_f)}{D(z)}\left[\frac{\delta_c}{D(z_f)^2\sigma_M^2}-\frac{1}{\delta_c}\right]\,, \tag{7.54}$$

which differs from the linear theory prediction of eq.(7.27). Analogous results could be obtained using the *frozen-potential* approximation (Brainerd, Scherrer & Villumsen 1993; Bagla & Padmanabhan 1994), with the main difference that the δ evolution would be slowed down compared to FFA. Quite interesting is that the *lognormal model* by Coles and Jones (1991) assumes that the quantity $1+\delta(\mathbf{x},z)$ can be always approximated by the exponential of the linear density field at the same Eulerian position, so that the expressions above for the Eulerian bias factor in FFA would apply to all Eulerian points $\mathbf{x}$. Of course, the validity of these approximate expressions for the bias should be checked against the results of N-body simulations.

Another way to get a local mapping between the evolved halo density field and the underlying matter perturbations is to approximate the non-linear evolution of the mass by the spherical top-hat model. This method has been followed by MW and Mo et al. (1996).

7.3.2 Perturbative evaluation of the Eulerian halo density contrast

In the previous section we demonstrated that the Eulerian bias is a non-linear and non-local function of the density fluctuation field. The 'non-locality', in particular, comes from the fact that the halo number density fluctuation in $\mathbf{x}$ is determined by the initial halo number fluctuation at the Lagrangian position $\mathbf{q}$, which, in turn, is related to the linear mass fluctuation in the same point, through a hierarchy of Lagrangian bias parameters. Here we want to derive an approximate expression for $\delta_h(\mathbf{x}(\mathbf{q},D),D)$, by applying the second-order Eulerian perturbation theory. Whenever it will be necessary to go from the Lagrangian position $\mathbf{q}$ to the Eulerian one, the Zel'dovich approximation will show sufficient.

Within the linear regime, the Eulerian solution of the continuity equation is simply, $\delta^{(1)}(\mathbf{x},D) = D\,\epsilon_\circ(\mathbf{x})$. The mildly non-linear regime may be approximately described by the second-order solution (Bouchet et al. 1992; Bernardeau 1994; Catelan et al. 1995; see also chapter 2)

$$\delta^{(2)}(\mathbf{x},D) = \frac{1}{2}\left[1-\frac{E}{D^2}\right]\delta^{(1)}(\mathbf{x},D)^2 - D\,\mathbf{u}^{(1)}(\mathbf{x})\cdot\nabla\delta^{(1)}(\mathbf{x},D) + \\ + \frac{1}{2}D^2\left[1+\frac{E}{D^2}\right]\partial_\alpha u_\beta^{(1)}(\mathbf{x})\,\partial_\alpha u_\beta^{(1)}(\mathbf{x})\;, \tag{7.55}$$

in such a way that $\delta = \delta^{(1)} + \delta^{(2)}$ and higher-order corrections are neglected. Here $\mathbf{u}^{(1)}(\mathbf{x}) = -\nabla\varphi_\circ(\mathbf{x})$ is the (scaled) linear peculiar velocity, and $\varphi_\circ$ the (scaled) peculiar gravitational potential, linearly extrapolated to the present time. The second-order growth factor $E = E(D)$ is quite a complicated function of $D(\Omega)$ (see the Appendix in CLMP, for its explicit expression), but, in the vicinity of $\Omega = 1$ (actually in the range $0.05 \leq \Omega \leq 3$), it can be approximated by the expression $E \approx -\frac{3}{7}\Omega^{-2/63}D^2 + \mathcal{O}[(\Omega-1)^2]$ (see Bouchet et al. 1992). Therefore, the previous second-order solution is well approximated by the expression which holds in the Einstein de Sitter universe, namely (Fry 1984)

$$\delta^{(2)}(\mathbf{x},D) = \frac{5}{7}\,\delta^{(1)}(\mathbf{x},D)^2 - D\,\mathbf{u}^{(1)}(\mathbf{x})\cdot\nabla\delta^{(1)}(\mathbf{x},D) + \frac{2}{7}\,D^2\partial_\alpha u_\beta^{(1)}(\mathbf{x})\,\partial_\alpha u_\beta^{(1)}(\mathbf{x})\;. \tag{7.56}$$

We want now to compute the corresponding second-order perturbative correction, $\delta_h^{(2)}(\mathbf{x},D)$, to the linear halo density fluctuation field, $\delta_h^{(1)}(\mathbf{x},D)$. From equation (7.38) we obtain

$$\delta_h \approx \delta_x^{(1)} + \delta_x^{(2)} + b_1^L\,\delta_q^{(1)} + b_1^L\,\delta_x^{(1)}\,\delta_q^{(1)} + \frac{1}{2}\,b_2^L\,\delta_q^{(1)2}\;, \tag{7.57}$$

where, to maintain compact the notation we wrote e.g. $\delta_x^{(j)} \equiv \delta^{(j)}(\mathbf{x},D)$. The Lagrangian bias factors $b_1^L = b_1^L(z|M,z_f)$ and $b_2^L = b_2^L(z|M,z_f)$ are those given in eqs.(7.16) and (7.17). Notice that the perturbative expansion of δ_x holds at sufficiently early times and/or large scales, while the validity of the expansion of $\delta_h(\mathbf{q})$ in powers of $\epsilon_\circ(\mathbf{q})$ is based on assuming a suitably large smoothing radius $R_\circ$ on the background field $\epsilon(\mathbf{q})$.

The key point is that the first-order density field at the Lagrangian position $\mathbf{q}$ originates a non-local term, when written at the corresponding Eulerian position $\mathbf{x}$. Using the Zel'dovich approximation $\mathbf{x} = \mathbf{q} + D\,\mathbf{u}^{(1)}$, one obtains $\delta_q^{(1)} = \delta_x^{(1)} - D\,\mathbf{u}^{(1)}\cdot\nabla\delta_x^{(1)}$. Finally, defining $\delta_h = \delta_h^{(1)} + \delta_h^{(2)}$, one gets $\delta_h^{(1)} = (1+b_1^L)\,\delta^{(1)}$ and

$$\delta_h^{(2)} = \left[\frac{1}{2}\left(1-\frac{E}{D^2}\right) + b_1^L + \frac{1}{2}b_2^L\right]\delta^{(1)2} - D\,\left(1+b_1^L\right)\mathbf{u}^{(1)}\cdot\nabla\delta^{(1)} + \\ + \frac{1}{2}\,D^2\left(1+\frac{E}{D^2}\right)\partial_\alpha u_\beta^{(1)}\,\partial_\alpha u_\beta^{(1)}\;. \tag{7.58}$$

Thus, the non-locality has the effect of modifying the inertia term $\mathbf{u}^{(1)}\cdot\nabla\delta^{(1)}$, which gets multiplied by the factor $(1+b_1^L)$. The dynamical properties of the random field δ_h may be equivalently analysed in terms of its Fourier transform $\tilde{\delta}_h(\mathbf{k},t)$ where $\mathbf{k}$ is the comoving wavevector. Thus, the second-order solution (7.58) may be written as

$$\tilde{\delta}_h^{(2)}(\mathbf{k},D) = D^2 \int \frac{d\mathbf{k}_1 d\mathbf{k}_2}{(2\pi)^3}\,\delta_D(\mathbf{k}_1+\mathbf{k}_2-\mathbf{k})\,\mathcal{H}_S^{(2)}(\mathbf{k}_1,\mathbf{k}_2;b_1^L,b_2^L;\Omega)\,\tilde{\delta}_1(\mathbf{k}_1)\,\tilde{\delta}_1(\mathbf{k}_2)\;, \tag{7.59}$$

where the symmetrized kernel $\mathcal{H}_S^{(2)}$ reads

$$\mathcal{H}_S^{(2)}(\mathbf{k}_1,\mathbf{k}_2;b_1^L,b_2^L,\Omega) \equiv \left[\frac{1}{2}\Big(1-\frac{E}{D^2}\Big)+b_1^L+\frac{1}{2}b_2^L\right] + \frac{1+b_1^L}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2} + \\ + \frac{1}{2}\Big(1+\frac{E}{D^2}\Big)\left(\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2}\right)^2 . \tag{7.60}$$

The corresponding kernel for the Einstein-de Sitter case reads

$$\mathcal{H}_S^{(2)}(\mathbf{k}_1,\mathbf{k}_2;b_1^L,b_2^L,\Omega=1) \equiv \left[\frac{5}{7}+b_1^L+\frac{1}{2}b_2^L\right] + \frac{1+b_1^L}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2} + \\ + \frac{2}{7}\left(\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2}\right)^2 . \tag{7.61}$$

7.3.3 Halo bispectrum and skewness

A possible application of these results is the evaluation of the bispectrum and corresponding skewness of the halo distribution. A related calculation has been performed by Fry (1996), who assumed the bias to be local in Eulerian space at z_f. It should be clear that our model is quite different to the local Eulerian bias prescription applied to the analysis of the skewness by Fry and Gaztañaga (1993). Moreover, the latter treatment, unlike ours, lacks of any prediction for the value of the different bias parameters. We recall that the value of the gravitationally induced skewness of the mass is

$$S = \frac{\langle\delta^3\rangle}{\langle\delta^2\rangle^2} = 4-2\frac{E}{D^2}\;, \tag{7.62}$$

for unfiltered fields, and

$$S(\mathcal{R}) = \frac{\langle\delta^3\rangle}{\langle\delta^2\rangle^2} = 4-2\frac{E}{D^2}-\gamma(\mathcal{R})\;, \tag{7.63}$$

for a spherical top-hat filter, where $\gamma \equiv -d\ln\sigma(\mathcal{R})^2/d\ln\mathcal{R}$ (Bernardeau 1994). The smoothing radius $\mathcal{R}$ should not be confused with R, defining the halo mass: one is obviously interested in computing the skewness on a smoothing scale much larger than the typical size of the single objects. In the Einstein-de Sitter universe, and for a scale-free power-spectrum, with spectral index n, the latter reduces to $S(\mathcal{R}) = 34/7-(n+3)$, for $-3 \le n < 1$.

The derivation of the halo skewness $\langle\delta_h^3\rangle \approx 3\langle\delta_h^{(1)2}\delta_h^{(2)}\rangle$ is simple. Assuming that the Eulerian halo density field is smoothed by a top-hat filter, the halo skewness parameter S_h is, for a generic value of Ω,

$$S_h(\mathcal{R};z,\Omega) = 3\frac{\langle\delta_h^{(1)2}\,\delta_h^{(2)}\rangle}{\langle\delta_h^{(1)2}\rangle^2} = \\ = \frac{4-2\frac{E}{D^2}+6\,b_1^L(z|M,z_f)+3\,b_2^L(z|M,z_f)-\left[1+b_1^L(z|M,z_f)\right]\gamma(\mathcal{R})}{\left[1+b_1^L(z|M,z_f)\right]^2}\;. \tag{7.64}$$

The asymptotic value of $S_h(\mathcal{R};z,\Omega)$, for a fixed formation redshift z_f, is $S-\gamma(\mathcal{R})$ as $z\to -1$, in the open and flat cases, while, for $\Omega>1$, this value is attained at the time of maximum expansion, corresponding to $z=-1/\Omega_\circ$. This limit gives the value of the underlying mass skewness: in the absence of merging the haloes would evolve towards an unbiased distribution in the far future. It is of interest to write the halo skewness in the Einstein-de Sitter universe and for a scale-free linear power-spectrum,

$$S_h(n;z,\Omega=1)=\frac{\frac{34}{7}+6\,b_1^L(z|M,z_f)+3\,b_2^L(z|M,z_f)-(n+3)\left[1+b_1^L(z|M,z_f)\right]}{\left[1+b_1^L(z|M,z_f)\right]^2}\,. \tag{7.65}$$

As for the mass skewness, the dependence on the smoothing scale $\mathcal{R}$ now simply translates into a dependence on the spectral index n. Again, the standard value for the mass skewness $34/7-(n+3)$ is always recovered at the end of the expansion phase. The skewness parameter is shown in Figure 7.3 for different values of $\Omega_\circ$ and for a scale-free model with $n=-2$. For objects observed at the present time, $z=0$, we vary the collapse epoch z_f, which may simulate different models of galaxy formation inside dark haloes. By varying together $z=z_f$, we instead show the skewness evolution in the instantaneous merging model. We also consider the case of varying only z: this gives the evolution of the skewness in a model in which the objects did not suffer any merging after their formation at z_f. Finally, we show the evolution of the skewness parameter of filtered mass fluctuations; note that the Einstein-de Sitter case displays no redshift dependence, simply because of self-similarity; for sensible values of $\Omega_\circ\neq 1$ also the mass skewness of non-flat Friedmann models experiences very little evolution. The redshift dependence of S_h is therefore mostly due to that of the Lagrangian bias factors. Quite interesting, in this respect, is the fact that the halo skewness plotted in the two top panels of Figure 7.3 displays a turning point in its redshift dependence: this typically occurs when $M\approx M_*(z_f)$.

Of particular interest is also the expression for the halo bispectrum B_h defined by the relation

$$\langle\tilde\delta_h(\mathbf{k}_2,D)\,\tilde\delta_h(\mathbf{k}_2,D)\,\tilde\delta_h(\mathbf{k}_3,D)\rangle\equiv(2\pi)^3\delta_{\scriptscriptstyle D}(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\,B_h(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3;D)\,. \tag{7.66}$$

The leading term shows the characteristic hierarchical pattern

$$\begin{aligned}B_h(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3;D)=&\,D^4\left[1+b_1^L(z|M,z_f)\right]^2\times\\&\times\left[2\,\mathcal{H}_S^{(2)}(\mathbf{k}_1,\mathbf{k}_2;b_1^L,b_2^L,\Omega)\,P(k_1)\,P(k_2)\;+\;\text{cyclic terms}\right]\,,\end{aligned} \tag{7.67}$$

where $P(k)$ is the primordial density power-spectrum defined by $\langle\tilde\delta_1(k_1)\,\tilde\delta_1(k_2)\rangle=(2\pi)^3\delta_{\scriptscriptstyle D}(\mathbf{k}_1+\mathbf{k}_2)\,P(k_1)$, and the two cyclic terms are obtained by the substitutions $\mathbf{k}_1\to\mathbf{k}_2$, $\mathbf{k}_1\to\mathbf{k}_3$ and $\mathbf{k}_2\to\mathbf{k}_3$. Typically, as for the hierarchical mass bispectrum, the halo bispectrum is largely scale dependent, while its dependence on the k-shape is rather weak. One way to eliminate the scale dependence and look at the residual shape dependence is to analyse the 'effective' bispectrum amplitude Q (Fry 1984; see also chapter 2)

$$Q\equiv\frac{B_h(\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3;D)}{P_h(k_1,D)\,P_h(k_2,D)+P_h(k_1,D)\,P_h(k_2,D)+P_h(k_2,D)\,P_h(k_3,D)}\,. \tag{7.68}$$

The halo power-spectrum is biased with respect to the mass one, $P_h(k,D)=D^2\left[1+b_1^L(z)\right]^2P(k)$. For a power-law spectrum, the amplitude Q generally depends on the spectral index n, owing to the wavenumber modulation introduced by the kernel $\mathcal{H}_S^{(2)}(\mathbf{k}_1,\mathbf{k}_2)$ (cf. Figure 7.4). For equilateral triangle configurations, Q gets an n-independent value, namely

$$Q_{eq}(\Omega;z)=\frac{\frac{1}{4}\left[1-3\,\frac{E}{D^2}\right]+2\,b_1^L(z|M,z_f)+b_2^L(z|M,z_f)}{\left[1+b_1^L(z|M,z_f)\right]^2}\,, \tag{7.69}$$

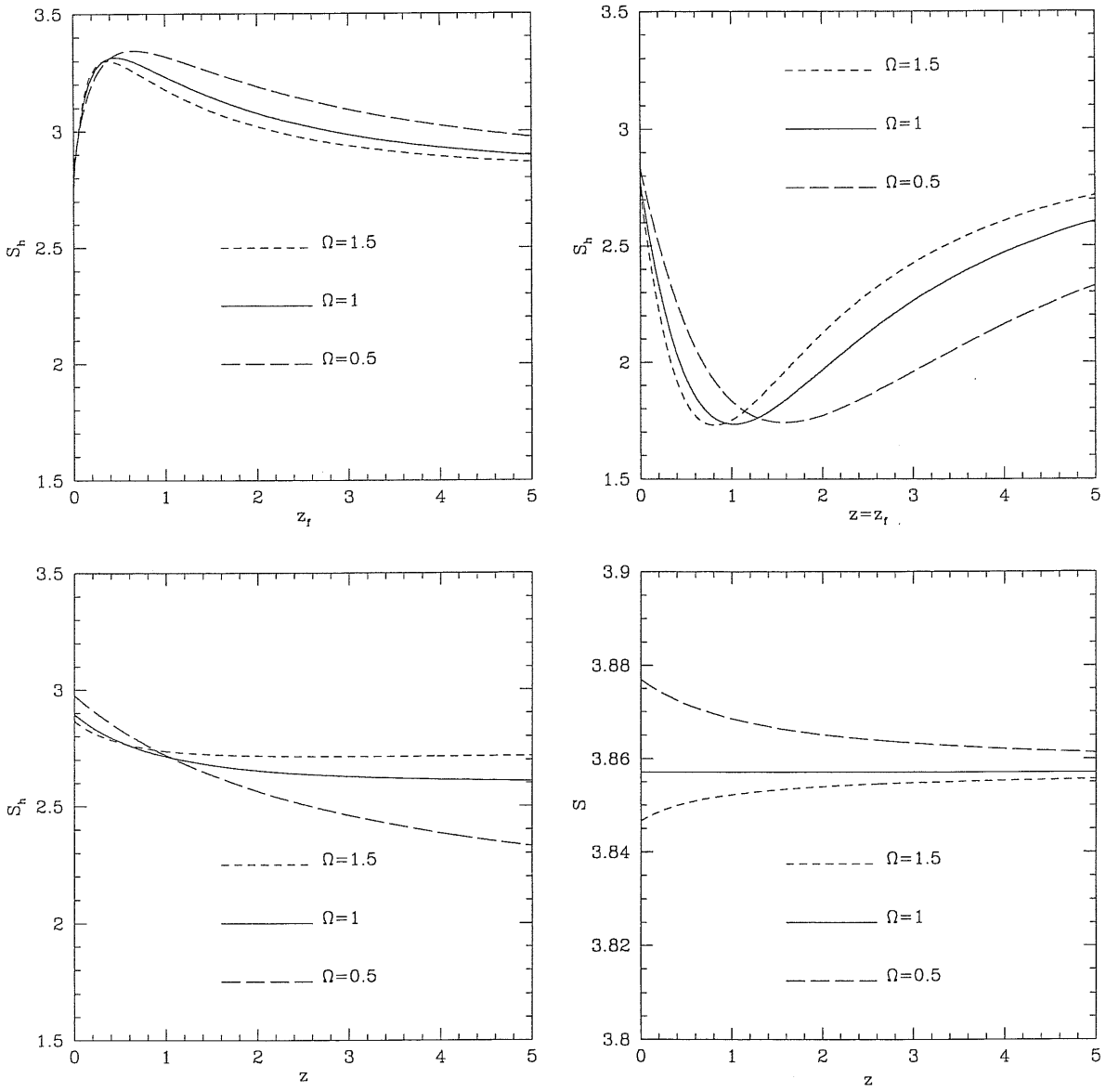


Figure 7.3: The filtered skewness parameter is plotted, for $\Omega_o = 0.5,\ 1,\ 1.5$, for a scale-free model with $n = -2$. The halo masses are selected with the same linear mass variance $\sigma_M^2 = 10$, corresponding to the same present-day bias parameters. We take everywhere $\delta_c = 1.69$. The *top left* panel refers to objects observed at $z = 0$, with varying formation redshift z_f. The *top right* panel shows the effect of varying simultaneously $z = z_f$. In the *bottom left* panel we fix $z_f = 5$ and look at different observation redshifts $z \leq z_f$. The *bottom right* panel, finally, shows the evolution of the skewness parameter of filtered mass fluctuations.

reducing to

$$Q_{eq}(\Omega=1;z)=\frac{\frac{4}{7}+2\,b_1^L(z|M,z_f)+b_2^L(z|M,z_f)}{\left[1+b_1^L(z|M,z_f)\right]^2}\,, \tag{7.70}$$

in the Einstein-de Sitter universe.

7.3.4 Local Lagrangian bias

So far, our model has been treated as being fully predictive. Once the cosmological scenario and the structure formation model have been fixed, our algorithm contains no fitting parameters. This is because we used a local version of the PS theory to generate the Lagrangian halo density contrast. One could, however, take a more general point of view and assume that the Lagrangian halo density contrast $\delta_h(\mathbf{q})$ is specified in terms of the linear background density field $\epsilon_{bg}(\mathbf{q},z)=D(z)\epsilon_o(\mathbf{q})$ by a set of unknown bias parameters $\hat{b}_\ell^L(z)$, as follows,

$$\delta_h(\mathbf{q})=\sum_{\ell=1}^{\infty}\frac{\hat{b}_{o\ell}^L}{\ell!}\,\epsilon_o(\mathbf{q})^\ell=\sum_{\ell=1}^{\infty}\frac{\hat{b}_\ell^L(z)}{\ell!}\,\epsilon_{bg}(\mathbf{q},z)^\ell\,. \tag{7.71}$$

Defining now $b_1\equiv b_1(z)=1+\hat{b}_1^L(z)$ and $b_2\equiv b_2(z)=2\hat{b}_1^L(z)+\hat{b}_2^L(z)$, according to eq.(7.28), and replacing these expansions in our previous treatment, we recover the general expression (7.59) for the second-order halo density contrast, with the more general kernel

$$\mathcal{H}_S^{(2)}(\mathbf{k}_1,\mathbf{k}_2;b_1,b_2,\Omega)=\frac{1}{2}\left[\left(1-\frac{E}{D^2}\right)+b_2\right]+\frac{b_1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2}+ \\ +\frac{1}{2}\left(1+\frac{E}{D^2}\right)\left(\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2}\right)^2\,. \tag{7.72}$$

Comparing this relation with the analogous one obtained with a local Eulerian bias expansion (e.g. Fry, Melott & Shandarin 1995; Matarrese, Verde & Heavens 1997), we see that the bispectrum for a set of objects selected by a local Lagrangian bias differs from the results of the local Eulerian bias by the extra inertia term

$$\frac{b_1-1}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right)\frac{\mathbf{k}_1\cdot\mathbf{k}_2}{k_1\,k_2}\,, \tag{7.73}$$

which implies a different shape dependence.

The halo bispectrum amplitude $Q(\theta)$, at $z=z_f=0$, for configurations with sides $k_1=1$, $k_2=1/2$, separated by an angle θ, is shown in Figure 7.4, for scale-free models with $n=-2$ and $n=-1$, with $\Omega=1$. Two different cases are considered: our local Lagrangian bias model, with linear Eulerian parameters $b_1=2$ and $b_2=1$ and the local Eulerian bias model of Fry and Gaztañaga, with the same Eulerian bias parameters.

Similar reasoning would apply to the skewness, for which the local Lagrangian vs. Eulerian bias hypothesis implies a change of the scale dependence, through the extra term

$$-\frac{b_1-1}{(b_1)^2}\,\gamma(\mathcal{R})\,. \tag{7.74}$$

With adequate modeling of galaxy formation inside dark matter haloes (e.g. Kauffmann et al. 1997, and references therein) the results of this section can be used to predict the clustering properties

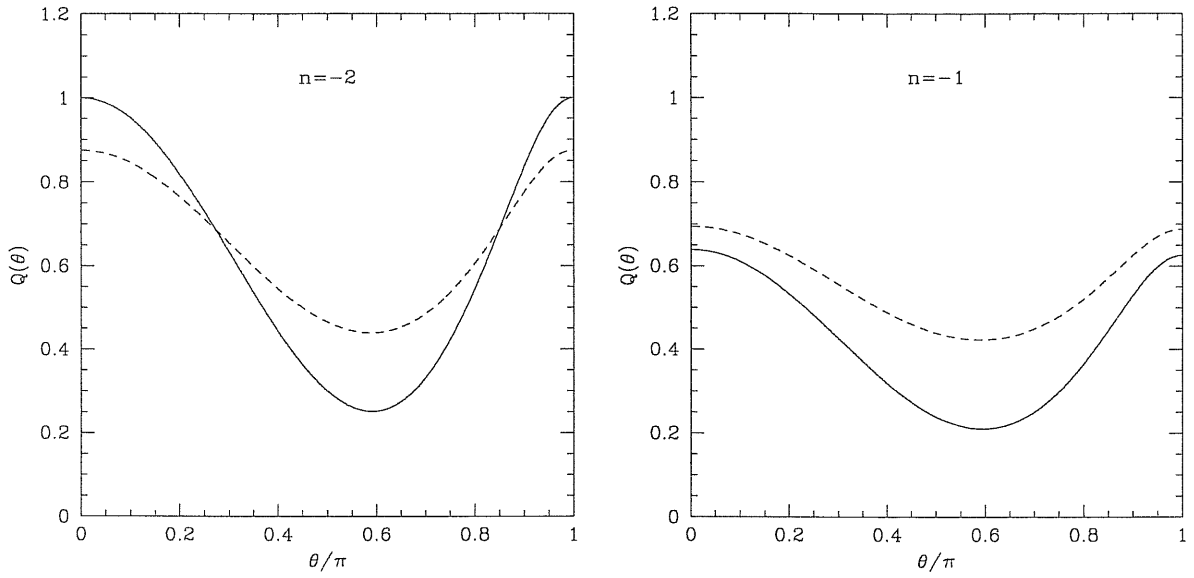


Figure 7.4: The halo bispectrum amplitude $Q(\theta)$ for configurations with sides $k_1 = 1$, $k_2 = 1/2$, separated by an angle θ is plotted vs. θ for scale-free models with $n = -2$ and $n = -1$ at $z = z_f = 0$ and in a flat Universe. Two cases are shown for each panel: the local Lagrangian bias model, with linear Eulerian parameters $b_1 = 2$ and $b_2 = 1$ (*solid line*) and the local Eulerian bias model, with the same bias parameters (*dashed line*).

of galaxies at different redshifts. In particular, the specific shape dependence of the bispectrum (and related scale dependence of the skewness), implied by our local Lagrangian bias prescription, would reflect into a detectable signature in the statistical properties of the galaxy distribution. Our model, therefore, provides a valid alternative to local Eulerian bias schemes (e.g. Cen & Ostriker 1992; Coles 1993; Fry & Gaztañaga 1993; Catelan et al. 1994; Mann et al. 1998).

7.4 Conclusions

In this chapter we studied the non-linear evolution of the clustering of dark matter haloes, using a stochastic approach to single out the halo formation sites directly in Lagrangian space. Our model is based on a local version of the Press-Schechter theory, which becomes free of the cloud-in-cloud problem after a suitable coarse-graining procedure is applied. The non-linear evolution of the halo distribution is then followed exactly by relating it to the dynamics of the Lagrangian patch of fluid which the nascent halo belongs to.

This formalism allowed us to obtain the bias random field relating the local halo density contrast to the underlying mass distribution. The expression for the halo bias field, reported in eqs.(7.48) and (7.49), represents the most relevant result of this thesis. Because of the locality in Lagrangian space inherent in our approach, such a bias field turns out to be non-local in Eulerian coordinates, which has relevant implications for the clustering properties of luminous objects like galaxies and galaxy clusters that formed inside dark matter haloes.

Our method contains two Lagrangian smoothing scales. The scale R, selecting the halo mass, and the background scale $R_\circ \gg R$ allowing us to define the Lagrangian halo counting field as the local PS mass function in a patch with comoving background density $\varrho_b\,[1 + \epsilon_{bg}(\mathbf{q}, z)]$, ϵ_{bg} being the linear mass fluctuation smoothed on the scale $R_\circ$. Given the role of the latter, it would appear

that our description of halo clustering makes sense only on scales larger than R_o. On the other hand, the derivation of the Lagrangian correlation function in §7.2.4, which does not make use of the background field, suggests that we can actually extrapolate our Lagrangian results down to separation comparable to the halo size. This result is further confirmed by an analysis in terms of space-correlated Langevin equations (Porciani et al. 1997). The numerical results of MW and Mo et al. (1996) support the idea that such an extrapolation would apply even in the non-linearly evolved case. In our treatment of the non-linear regime, the background scale R_o appears with a complementary role. It is the minimum scale ensuring that the nascent haloes are indeed comoving with the Lagrangian fluid patch which they belong to. This would reasonably require that the Lagrangian fluid elements evolve with negligible orbit crossing (e.g. Kofman et al. 1994).

Once again, let us stress that our approach makes no assumptions about the merger rates of the considered objects. The clear distinction between observation and formation redshift, z and z_f, in our approach implies that the instantaneous merging hypothesis, implicit in the standard PS model, as well as any other realistic approximation can be easily accommodated into our scheme as just the way to relate z_f and z.

Our method for evolving the spatial distribution of the haloes is indeed much more general than the specific application we have considered so far. Given any Lagrangian population of objects specified by some set of physical properties $\mathcal{M}$ (like mass and formation threshold in our halo model), with conserved mean comoving number density $\bar{n}_{obj}(\mathcal{M})$ and local Lagrangian density contrast $\delta_{obj}(\mathbf{q}|\mathcal{M})$, our results imply that, at any redshift z, their comoving local density in Eulerian space is given by

$$n_{obj}(\mathbf{x}, z|\mathcal{M}) = \bar{n}_{obj}(\mathcal{M}) \int d\mathbf{q}\, [1 + \delta_{obj}(\mathbf{q}|\mathcal{M})]\; \delta_D[\mathbf{x} - \mathbf{x}(\mathbf{q}, z)] \;, \tag{7.75}$$

where $\mathbf{x}(\mathbf{q}, z) = \mathbf{q} + \mathbf{S}(\mathbf{q}, z)$, and $\mathbf{S}(\mathbf{q}, z)$ is the displacement vector of the $\mathbf{q}$-th Lagrangian element. Smoothing the initial gravitational potential on some scale R_o is again required, so that the objects assigned to the $\mathbf{q}$-th patch can be sensibly assumed to be comoving with it. This method could be used, for instance, to follow the clustering of the Lagrangian density maxima in the non-linearly evolved mass density field. This suggests that, starting from low-resolution numerical simulations, one can generate mock catalogues of the given class of objects, with local density correctly specified up to some resolution scale. One can understand the last relation as a local version of the Chapman-Kolmogorov equation of stochastic processes (e.g. van Kampen 1992), stating that the local Eulerian object distribution is the convolution of the Lagrangian object density with the 'conditional particle density', $\delta_D[\mathbf{x} - \mathbf{x}(\mathbf{q}, z)]$, i.e. the probability that a particle is found in $\mathbf{x}$ at redshift z given that it was in $\mathbf{q}$ as $z \to \infty$. The only underlying hypothesis being, once again, that these objects move exclusively by the action of gravity. It may be worth to notice that the latter equation is actually more general than eq. (7.38), as it also holds in the presence of multi-streaming.

Our non-linear stochastic approach can be already considered successful in that, besides recovering the PS mass function, it provides a self-consistent derivation of the Eulerian halo bias, which, to a first approximation, reduces to the MW formula. We, however, also predict both quantitative and qualititive corrections to the MW results, that clearly need to be checked against the outputs of numerical simulations. A definite prediction of our analysis is, for instance, the form of the skewness and of the bispectrum of the spatial halo distribution, which significantly deviates from that deduced with any local Eulerian bias prescription.

Chapter 8

Comparison with N-body simulations

In this chapter we compare the predictions of two different biasing schemes for the distribution of dark matter haloes with the output of high-resolution N-body simulations. In particular, the MW algorithm and the model that we presented in chapter 7 are examined. Our analysis is entirely based on two recently published papers (Catelan, Matarrese & Porciani 1998; Porciani, Catelan & Lacey 1999), and we report their full text here. Both of them are provided with a small introduction which explains their collocation in the present-day scientific discussion.

8.1 How dark matter haloes cluster in Lagrangian space

The MW model for the Eulerian biasing of dark matter haloes has been thoroughly tested against N-body simulations by MW, Mo, Jing & White (1997), Sheth & Lemson (1998), Jing (1998), and Sheth & Tormen (1999) using different statistics. These studies clearly show that the linear bias parameter given by MW is accurate in accounting for the clustering properties of haloes with $M > M_*$, while it always underestimates the value of b^2 found in numerical simulations for $M < M_*$ (e.g. Jing 1998). For example, this is clearly seen in fig. (8.1) which has been extracted from Jing (1998, his figure 2). Since the MW method is composed of two well-distinguished parts (see chapter 6), it is important to clarify if the failure of the model at the low-mass tail is caused by the halo-selection algorithm in Lagrangian space (i.e. the PS theory) or by the simple scheme adopted to transform Lagrangian overdensities into Eulerian ones (i.e. the spherical model). This issue is deeply explored in Porciani, Catelan & Lacey (1999) where we use scale-free spectra with index -1 and -2 to address the problem. The text of this paper is reported in the following pages. Our results clearly show that the discrepancy between analytical predictions and numerical data can be totally ascribed to the lack of accuracy of the PS method. A subsequent study by Jing (1999), based on simulations with very high-resolution, confirmed our results extending their validity to a broader range of spectral indices. Sheth & Tormen (1999) showed that modifying the PS result, so as to reproduce the low-mass tail of the mass function obtained in the simulations, automatically accounts for the correct bias parameter. This is obtained by assuming that the PS threshold is a function of M which can be determined using a best fitting method to match the numerical mass function. The corresponding linear bias parameter is then computed using the MW formula. An improved model for the mass function based on sounder theoretical arguments than the naive PS approach is then required. This problem surely deserves more investigation in the near future (e.g. Mo, Sheth & Tormen in preparation).

THE ASTROPHYSICAL JOURNAL, 513:L99–L102, 1999 March 10

HOW DARK MATTER HALOS CLUSTER IN LAGRANGIAN SPACE

CRISTIANO PORCIANI,[1,2] PAOLO CATELAN,[3] AND CEDRIC LACEY[3]
Received 1998 December 1; accepted 1999 January 15; published 1999 February 3

ABSTRACT

We investigate the clustering of dark matter halos in Lagrangian space in terms of their two-point correlation function. Analyzing a set of collisionless scale-free 128^3 particle N-body simulations with spectral indices $n = -2, -1$, we measure the first two Lagrangian bias parameters b_1 and b_2 relating halo and mass correlations. We find that the Mo & White leading-order formula for b_1 describes the clustering of halos with mass $M \gtrsim M_*$ (where M_* indicates the characteristic nonlinear mass) quite accurately. Smaller halos turn out to be less clustered in Lagrangian space than predicted by Mo & White. Our findings are consistent with the recent results of Jing for the clustering of halo populations in Eulerian space, demonstrating that the discrepancies between the N-body and analytical Mo & White prediction for the bias exist already in Lagrangian space. This shows that a more refined theoretical algorithm for selecting halos in the initial conditions needs to be developed. Finally, we present a very accurate fitting formula for the linear halo bias factor b_1 in Lagrangian space.

Subject headings: cosmology: theory — galaxies: statistics — large-scale structure of universe

1. INTRODUCTION

Virialized dark matter halos in the universe are not distributed in the same way as the underlying dark matter. This is true whether one takes the final positions of the halos in Eulerian space or their initial positions in Lagrangian space (Mo & White 1996, hereafter MW; Catelan et al. 1998, hereafter CLMP; Catelan, Matarrese, & Porciani 1998; Jing 1998; Sheth & Lemson 1998). The numerical tour de force by Jing (1998), who thoroughly investigated with unmatched accuracy the clustering of dark matter halos in Eulerian space, demonstrated explicitly that (1) the halo-to-mass bias is independent of the halo separation (at least in the scale-free case $n = -2$ and in the linear regime) and (2) the MW Eulerian linear bias correctly describes the clustering of halos of masses $M \gtrsim M_*$, but systematically underpredicts it for any value of the spectral index n if $M \lesssim M_*$, where M_* is the typical nonlinear mass. However, it is impossible to understand solely on the basis of the Eulerian investigation whether the discrepancies between the numerical results and the analytical MW predictions are due to (1) a failure of the algorithm for identifying the halo positions in Lagrangian space, (2) the effects of nonlinear shear dynamics (not accounted for in the original MW approach) on the mapping of halo positions from Lagrangian to Eulerian space, or (3) a combination of the two.

In this Letter, we employ scale-free collisionless N-body simulations [i.e., with density parameter $\Omega = 1$ and initial power spectra $P(k) \propto k^n$] to investigate the clustering of dark matter halos in Lagrangian space in terms of their two-point correlation function. We compare the halo correlation function to the correlation of the underlying dark matter for $n = -2$, -1 spanning more than 4 (3) orders of magnitudes in halo masses for $n = -2$ ($n = -1$).

We find that the theoretical "underclustering" reported by Jing for masses $M \lesssim M_*$ is already present in the initial conditions as well (but as "overclustering," since the first-order bias is negative for small masses), and it cannot be due exclusively to the subsequent nonlinear effects of the shear dynamics acting on small scales and unaccounted for in the original spherical collapse model of MW. We then argue that the standard Press-Schechter (1974) approach ("extended" or not), on which MW and CLMP based their speculations, is inadequate for identifying the locations of halos in the initial conditions, and a major effort should be devoted to finding an improved algorithm. Section 2 reports the details of the present investigation. Finally, an accurate fitting formula for the bias b_1 in Lagrangian space is given, which should be considered as the Lagrangian version of Jing's fitting formula for the Eulerian case. Section 3 contains our conclusions.

2. HALO CLUSTERING FROM N-BODY SIMULATIONS

2.1. *Simulations and Halos*

The simulations used here are similar to those of Lacey & Cole (1994), who used them to test halo merging histories. They were performed using the AP^{3}M code of Couchman (1991) with 128^3 particles. The force softening used was $L/1280$, with L the size of the periodic box. Initial positions and velocities were generated by displacing particles from a uniform 128^3 grid according to the Zeldovich approximation, assuming an initial scale-free power spectrum and Gaussian statistics. We considered four realizations for two different spectral indices $n = -2, -1$. For each simulation, we recorded, for many epochs, positions and velocities of all particles. The output times (38 for $n = -2$ and 35 for $n = -1$) were chosen so that M_* increased by a factor $\sqrt{2}$ between subsequent output times. In the last output, M_* corresponded to 33,748 particles for $n = -2$ and to 28,616 particles for $n = -1$.

For each output time, we selected dark matter halos in the simulations employing the "friends-of-friends" group finder with a linking length equal to 20% of the mean interparticle distance (e.g., Davis et al. 1985). We checked that results obtained using the spherical overdensity group finder (Lacey & Cole 1994) are essentially identical. We excluded halos containing fewer than 20 particles or more than 20,000 particles. We moved all of the particles belonging to a given halo back to their initial (Lagrangian) positions, then computed the position of their center of mass. We used the latter as the "halo position" in Lagrangian space. In such a way, for each output

[1] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218.
[2] Scuola Internazionale Superiore di Studi Avanzati, via Beirut 2-4, 34014 Trieste, Italy.
[3] Theoretical Astrophysics Center, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark.

time, we constructed a catalog of halos indicating their mass and position in Lagrangian space.

2.2. *Halo Correlation Function in Lagrangian Space*

We computed the mean correlation function $\bar{\xi}_h$ between halos in a given mass interval, where a bar denotes mass averaged quantities. Self-similar scaling allowed us to combine data from different output times in order to reduce the Poisson fluctuations due to the finite number of halos within the box. We considered every output time containing more than 100 halos in the same mass interval, and we binned the distributions of halo separations in units of r/R_*, where $R_*^3 \propto M_*$. Finally, we computed the halo correlation function using the estimator

$$\langle \bar{\xi}_h \rangle = \frac{\sum_i N_i[r/R_*(z_i)]}{\sum_i N_i^{\rm Poi}[r/R_*(z_i)]} - 1, \qquad (1)$$

where the index i runs over different output times z_i of the same simulation, $N_i(r/R_*)$ is the number of halo pairs in the ith output, and $N_i^{\rm Poi}(r/R_*)$ is the corresponding quantity for a Poisson process with the same number density. The average symbol (angle brackets) is introduced since we considered information coming from the different temporal outputs. In this way, for each mass interval, we collected four realizations of the Lagrangian correlation function of dark matter halos.

This procedure allows us to achieve two goals: (1) to reduce statistical fluctuations by increasing the number of halo pairs and (2) to extend the mass interval and the range of halo separations that may be sampled. In fact, the box size L and the minimum halo mass in the simulation are fixed, while R_* and M_* increase with time. This means that the correlation function for halos with $M \gg M_*$ is measured mainly from the early output times, while that for halos less massive than M_* comes mostly from later output times. In order to simulate bootstrap resampling, we assigned as the standard error the Poisson error bar multiplied by a factor of $\sqrt{3}$ (Mo, Jing, & Börner 1992).

2.3. *Lagrangian Bias Parameters*

Now we want to test whether the Lagrangian halo correlation ξ_h is related that of the mass (ξ_m) through a relation (see eq. [15] in CLMP and eq. [38] in Porciani et al. 1998)

$$\xi_h\left(\frac{r}{R_*}\right) = b_1^2 \xi_m\left(\frac{r}{R_*}\right) + \frac{b_2^2}{2}\,\xi_m^2\left(\frac{r}{R_*}\right) + \dots\,, \qquad (2)$$

where the symbol b_i denotes the ith Lagrangian bias factor, and ξ_m is calculated according to linear theory. For instance, the extended Press-Schechter approach (Bond et al. 1991) leads to the MW and CLMP expressions for b_1 and b_2, namely $b_1 = \delta_c/\sigma_M^2 - 1/\delta_c$ and $b_2 = (\delta_c^2/\sigma_M^2 - 3)/\sigma_M^2$, where σ_M^2 indicates the mass variance on scale M (Cole & Kaiser 1989; MW; Mo, Jing, & White 1997). Note that, by definition, $\sigma_M = \delta_c$ for $M = M_*$. To test this biasing model against simulations, we have to consider a finite range of halo masses. Equation (2) then implies $\bar{\xi}_h = \bar{b}_1^2\xi_m + \frac{1}{2}\bar{b}_2^2\xi_m^2 + \dots$, where $\bar{b}_k$ is the mean of b_k in the mass interval, weighted by the mass function. In order to obtain statistically reliable values for $\bar{b}_1$ and $\bar{b}_2$, we compute the averages $\langle \xi_m \rangle$ and $\langle \xi_m^2 \rangle$ *directly* from the initial conditions of the simulation, following the same averaging procedure we used for the halo correlations (in this case, it corresponds to averaging over output times using the volume of the r/R_* bin as a weight). Note that, after averaging, $\langle \xi_m^2 \rangle$ generally differs from $\langle \xi_m \rangle^2$, so that equation (2) implies $\langle \bar{\xi}_h \rangle = \bar{b}_1^2\langle \xi_m \rangle + \frac{1}{2}\bar{b}_2^2\langle \xi_m^2 \rangle + \dots$. This averaging procedure is important to account for the finiteness of the box: in this way, the lack of any Fourier components of the density would be equally experienced by both mass and halo distributions. Finally, to check the reliability of equation (2), we apply the χ^2 test, assuming that every $\langle \bar{\xi}_h \rangle$ is normally distributed around a mean value.

Halos are objects of finite size, and we expect exclusion

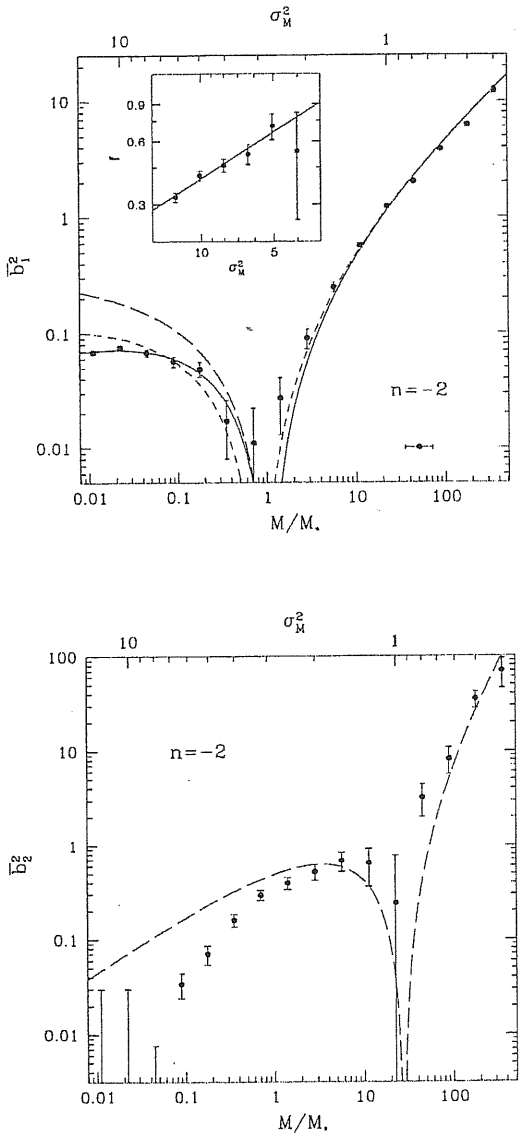


FIG. 1.—*Dots:* Best-fit values of $\bar{b}_1^2$ (*top*) and $\bar{b}_2^2$ (*bottom*) as a function of variance σ_M^2 and halo mass on the same scale for $n = -2$. Error bars denote 1 σ uncertainties. *Long-dashed line:* MW and CLMP Lagrangian bias with $\delta_c = 1.686$. *Short-dashed line:* Jing's fit to the Eulerian bias minus 1. *Continuous line:* Our fitting formula for b_1 given in the main text. Both the theoretical predictions and the fitting formula are averaged over the Press-Schechter mass function. The horizontal error bar in the bottom right-hand corner of the top panel indicates the mass intervals considered. The inset in the top panel shows the best-fit correction $f(\sigma_M)$ to the MW formula for b_1, when $M < M_*$: $\log f = (0.34 \pm 0.11) - (0.73 \pm 0.12)\log \sigma_M^2$.

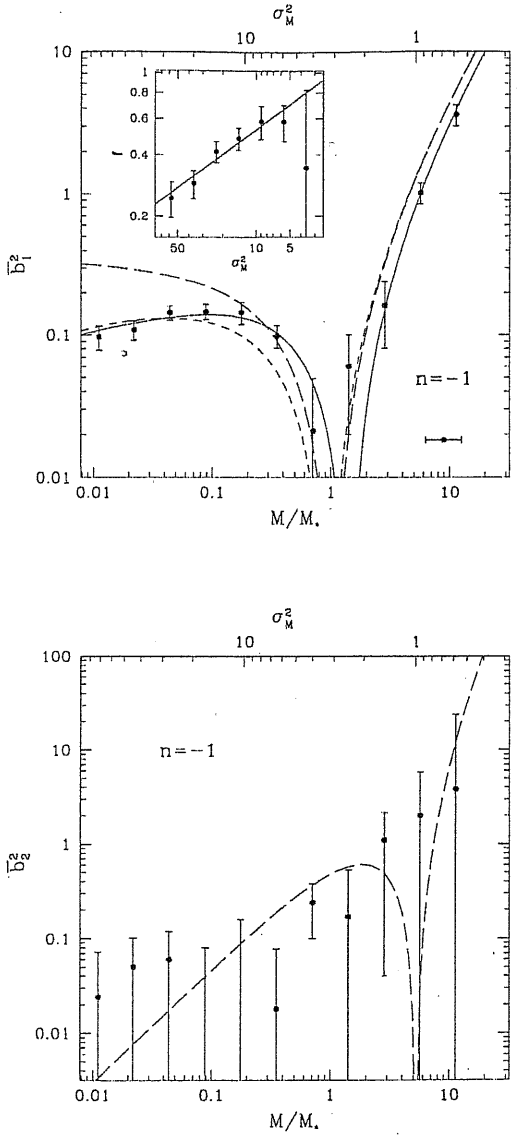


FIG. 2.—As in Fig. 1, but for $n = -1$. In this case, the best fit for f is $\log f = (0.13 \pm 0.13) - (0.41 \pm 0.10) \log \sigma_M^2$.

effects to dominate ξ_h at separations of the order of the halo Lagrangian size R_L. In a sample composed of identical spherical halos, spatial exclusiveness implies $\xi_h = -1$ for $r < 2R_L$ and a strong compensating positive correlation at $r \gtrsim 2R_L$. The magnitude of this positive correlation is related to the volume fraction occupied by the spheres, since $\int_0^\infty dr\, r^2\, \xi_h(r) = 0$. Inspired by this, we decided to exclude from the χ^2 test all data corresponding to separations $r \lesssim 2R_{max}$, where $R_{max} = R_*(M_{max}/M_*)^{1/3}$ is the characteristic Lagrangian radius of the largest halo in the mass interval considered. Note that, out of the exclusion region, the additional condition $|\xi_m| < 1$ is not needed to neglect higher order terms in equation (2). This is because, for $i > 1$ and $\sigma_M \to \infty$, $b_i \to 0$ as $\sigma_M^{-\alpha}$ with $\alpha \geq 2$. Anyway, the exact value of the minimum separation considered has been determined on a case-by-case basis, checking for the stability of our results with increasing the number of data points at small separations. We avoided averaging between different realizations of the same power spectrum: the entire data set of up to four values of $\langle \xi_h \rangle$ for each r/R_* interval is considered in the χ^2 test.

Minimization of χ^2 with nonholonomic constraints $\bar{b}_1^2 \geq 0$ and $\bar{b}_2^2 \geq 0$ is used to obtain estimates of the bias parameters. The resulting χ^2_{min} shows that the a priori bias hypothesis—the bias relation in equation (2)—is acceptable to a 95% confidence level over a wide interval of halo masses. However, for $n = -2$, samples with $M > 16M_*$ are inconsistent with equation (2) at the 99% confidence level. This is caused by the presence of noticeable differences (especially at very large separations) between the halo correlation functions extracted from different realizations. Indeed, very massive halos are exponentially rare and more affected by statistical fluctuations. On the other hand, we could also have underestimated the uncertainty in computing $\langle \xi_h \rangle$.

Figures 1 and 2 show the best-fit results for $\bar{b}_1^2$ and $\bar{b}_2^2$. The error bars are obtained by projecting along the b-axes the contours of constant χ^2 at $\chi^2_{min} + 1$. Concerning $\bar{b}_1^2$, the comparison against theoretical predictions shows a quantitatively good agreement for $M > M_*$. However, the square bias parameter of lower mass halos is significantly smaller than predicted by Mo & White (meaning that the bias $\bar{b}_1$ is *less negative* than the MW value). Findings of this kind have been published by Jing (1998), who investigated the Eulerian halo clustering in the quasi-linear regime, giving an accurate fitting formula for the Eulerian analog of b_1. Our result shows incontrovertably that the effects discovered by Jing are already present in Lagrangian space.

There is no rigorous way to apply Jing's bias-fitting formula in Lagrangian space, since the nonlocal dynamics of the mass density field enters the transformation. However, as a first approach, and only to check the order of magnitude of the effect, we can follow MW in assuming a spherical evolution of the coarse-grained mass density field to map from Lagrangian to Eulerian space. In this case, the first Eulerian bias term is given by $1 + b_1$. Jing's formula, $b_1 \equiv b_{Jing} - 1$, turns out to be a reasonably good description of our data, showing that the underclustering of small halos in Eulerian space detected by Jing is correspondingly already present in their Lagrangian clustering; the intervening dynamics is approximately well accounted for by the standard mapping from Lagrangian to Eulerian space in the laminar regime. We elaborate below our own fitting formula for the linear Lagrangian bias, which turns out to be accurate to 10% over the entire mass range investigated (with the exception of the zero crossing region for b_1).

Our results for b_2 are of course less conclusive than for b_1, mainly because of the larger uncertainties. However, they confirm the information extracted considering b_1. Quantitatively, for $n = -2$, the CLMP expression for b_2 gives a good estimate for halos with $M > M_*$, while it overpredicts the numerical outcome for b_2^2 when $M < M_*$. For $M \sim M_*$, our results are a factor of ~1.4 smaller than predicted by the CLMP formula. This is the range in which our method is best suited to compute b_2, since b_1 vanishes and the halo correlation function is proportional to ξ_m^2. No firm conclusion can be drawn for $n = -1$, since practically the entire data set only allows one to set an upper limit on $\bar{b}_2^2$. It is not unlikely that, for the less negative spectral indices, bigger simulations are better suited to quantify the halo clustering up to second-order biasing, above all in Lagrangian space. It will be worthwhile to address this point in a future work.

2.4. *Accurate Fitting Formula for b_1*

Our results for b_1 can be accurately parametrized by introducing a mass-dependent, multiplicative correction to the MW formula, namely

$$b_1^2(\sigma_M) = f(\sigma_M)(\delta_c/\sigma_M^2 - 1/\delta_c)^2. \quad (3)$$

For $n = -2$, the original MW formula (i.e., $f \equiv 1$) with $\delta_c = 1.686$, as suggested by the spherical collapse model, describes the data for $M > M_*$ to within 10% accuracy. Smaller masses instead require $f = 2.19/\sigma_M^{1.46}$. For $n = -1$ and $M > M_*$, our numerical data are very well described by the MW formula for b_1 with a lower collapse threshold $\delta_c \simeq 1.52$. In this case, for $M < M_*$ we obtain $f = 1.35/\sigma_M^{0.82}$. These fitting formulae are extremely accurate in describing our data set. Their simple power-law behavior encourages further theoretical investigation. For instance, with data for more than two spectral indices n, one may attempt to fit the dependence of b_1 on n.

3. DISCUSSION AND CONCLUSIONS

Employing 128^3-body scale-free simulations, we analyzed the clustering of dark matter halos in Lagrangian space. The main results of this investigation can be summarized as follows.

1. Assuming a correlation model as in equation (2), the first two Lagrangian bias factors b_1 and b_2 are strongly mass dependent over the 4 orders of magnitude in mass investigated.
2. The clustering of halos with mass above the nonlinear mass, $M \gtrsim M_*$, is fairly well described by the MW formula for the linear Lagrangian bias, both for $n = -1$ and $n = -2$.
3. Halos with nonlinear masses $M \lesssim M_*$ are less clustered (have a smaller correlation amplitude) than what the leading order Lagrangian bias of MW would predict (see also Sheth & Lemson 1998).

When these results are combined with the ones recently obtained by Jing (1998) about the clustering of halos in Eulerian space, we can disentangle the question of whether the discrepancies between Jing's numerical results and the MW theoretical predictions are mainly due to the effects of the nonlinear shear dynamics, effective on smaller scales, or to a possible failure of the halo selection algorithm in the initial conditions—a question actually left unsolved by Jing. Clearly, since, as we showed, the *same* effects discovered by Jing are essentially already present in Lagrangian space, our investigation suggests that it is time to improve on the classical Press-Schechter algorithm for identifying halos in Lagrangian space: since it assumes spherically symmetric collapse, it is not surprising that it fails in correctly counting the small halo masses, where departures from spherical collapse can be cosmologically relevant. The failure of the MW formula for the Lagrangian and Eulerian bias is presumably related to the departure of the halo mass function from the Press-Schechter form at low masses (e.g., Lacey & Cole 1994).

Finally, we derived a fitting formula for the linear Lagrangian bias b_1 that is relevant for accurately predicting the clustering of dark matter halos, above all in the low-mass tail. Our fitting formula can be considered as the Lagrangian equivalent of equation (3) in Jing (1998). These results are highly relevant for predicting the clustering of low-luminosity galaxies, most of which lie in lower-mass halos (e.g., Baugh et al. 1998). Modeling of galaxy clustering found in present and future galaxy redshift surveys will provide an important application of the present results.

C. P. acknowledges the hospitality of the Theoretical Astrophysics Center during the spring of 1997, when this work started. P. C. warmly thanks Sergio Gelato for insightful discussions. We all thank Sabino Matarrese for sharing his ideas on the matter, dark and not. C. P. thanks support from NASA ATP-NAG5-4236 grant and from Italian Ministero dell'Universita è della Ricerca Scientifica e Tecnologica. P. C. and C. L. have been supported by the Danish National Research Foundation at TAC.

REFERENCES

Baugh, C. M., et al. 1998, preprint (astro-ph/9811222)
Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Catelan, P., Lucchin, F., Matarrese, S., & Porciani, C. 1998, MNRAS, 297, 692
Catelan, P., Matarrese, S., & Porciani, C. 1998, ApJ, 502, L1
Cole, S., & Kaiser, N. 1989, MNRAS, 237, 1127
Couchman, H. M. P. 1991, ApJ, 368, L23
Davis, M., Efsthatiou, G., Frenk, C., & White, S. D. M. 1985, ApJ, 292, 371
Jing, Y. P. 1998, ApJ, 503, L9
Lacey, C., & Cole, S. 1994, MNRAS, 271, 676
Mo, H. J., Jing, Y. P., & Börner, G. 1992, ApJ, 392, 452
Mo, H. J., Jing, Y. P., & White, S. D. M. 1997, MNRAS, 284, 189
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347
Porciani, C., Matarrese, S., Lucchin, F., & Catelan, P. 1998, MNRAS, 298, 1097
Press, W. H., & Schechter P. 1974, ApJ, 187, 425
Sheth, R. K., & Lemson, G. 1998, preprint (astro-ph/9808138)

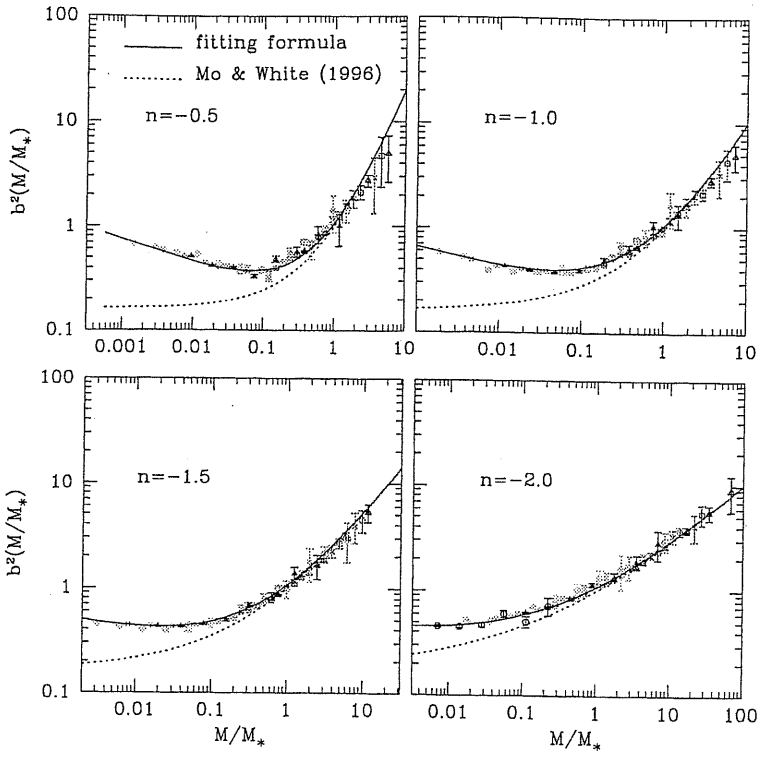


Figure 8.1: The square of the bias parameter as a function of the halo mass. Symbols with errorbars denote results from the numerical simulations by Jing (1998) at different output times. The dotted line is the prediction of the MW model. The continuous line is a fitting formula given by Jing (1998). This figure is extracted from Jing (1998).

8.2 On the spatial distribution of dark matter haloes

The biasing scheme we described in chapter 7 represents rather a big paradigm than a real model. In fact, in analogy with the MW method, it acts in two steps: first it selects the spots for halo formation in Lagrangian space, and then it follows the gravitational evolution of the large-scale mass density field. Obviously, both these processes need to be accurately modelled in order to make good predictions. In some sense, eq. (7.48) gives a recipe and the success of its outcome depends on the quality of the fundamental ingredients we use to carry it out. For example, low resolution N-body simulations can be certainly used to obtain a precise description of large-scale motions. It is, however, interesting to investigate if simple dynamical approximations can also be sufficiently accurate. In this case, an easy and fast numerical implementation of the biasing scheme and the possibility to obtain, at least for a restricted set of statistics, analytical solutions would represent the main advantages. In the paper that we report in the following pages, we use the truncated Zel'dovich approximation together with the simple PS theory to generate synthetic maps of the present-day halo population. The results are compared with the outcome of a high-resolution numerical simulation generated starting from the same initial conditions. A series of tests shows that our method gives an accurate description of the halo population. However, while the agreement is excellent on average, the lack of accuracy of the PS method introduces some scatter in a point-by-point comparison. This again reinforces the need for an improved halo selection method already invoked in §8.1. A different test involving the halo correlation functions is now in progress (Matarrese, Catelan & Porciani in preparation).

THE ASTROPHYSICAL JOURNAL, 502:L1–L4, 1998 July 20

ON THE SPATIAL DISTRIBUTION OF DARK MATTER HALOS

PAOLO CATELAN,[1] SABINO MATARRESE,[2] AND CRISTIANO PORCIANI[3,4]
Received 1998 April 24; accepted 1998 May 26; published 1998 July 1

ABSTRACT

We study the spatial distribution of dark matter halos in the universe in terms of their number density contrast, related to the underlying dark matter fluctuation via a nonlocal and nonlinear bias random field. The description of the matter dynamics is simplified by adopting the "truncated" Zeldovich approximation to obtain both analytical results and simulated maps. The halo number density field in our maps and its probability distribution reproduce with excellent accuracy those of halos in a high-resolution N-body simulation with the same initial conditions. Our nonlinear and nonlocal bias prescription matches the N-body halo distribution better than any Eulerian linear and local bias.

Subject headings: galaxies: statistics — large-scale structure of universe

1. INTRODUCTION

The simplest description for biasing assumes that the fluctuations in the number density of luminous objects, $\delta_{\rm lum}$, and in the mass, $\delta_{\rm mass}$, are proportional, $\delta_{\rm lum} = b\delta_{\rm mass}$, where b is the so called linear *bias factor*. Recently, Catelan et al. (1998, hereafter CLMP), following the seminal papers by Cole & Kaiser (1988) and Mo & White (1996, hereafter MW), showed how the relation between dark halos, recipient of the luminous matter, and the underlying mass is to be cast in terms of a bias random *field* b, which depends in a nonlocal way on the mass density field. Halo biasing is a process that evolves in time and depends on the scales and the collapse times of the selected objects, but it is additionally determined by the gravitational conditions of the environment. Most important, unlike previous models, CLMP treated halo biasing as a nonlocal process.

In this Letter, we apply the CLMP bias model to analyze the spatial halo distribution at several scales. Mass particles move according to the Zeldovich (1970) approximation. We generally find excellent agreement between our theoretical predictions and the distribution of halos extracted from an N-body simulation with the same initial conditions. In § 2 we present our bias model, in § 3 we test it against simulations; § 4 contains our conclusions.

2. THE DISTRIBUTION OF HALO FLUCTUATIONS

2.1. *The Model*

Let us consider a population of halos, selected in Lagrangian space through their mass M and formation redshift z_f. At any comoving position $\boldsymbol{x}$ and observation redshift $z \le z_f$, one can generate Eulerian maps of the halo number density fluctuation $\delta_h(\boldsymbol{x}, z|M, z_f)$, given the mass density contrast $\delta(\boldsymbol{x}, z)$ with Lagrangian resolution R_0 and the corresponding Lagrangian halo density fluctuation field, $\delta_h^{\rm L}(\boldsymbol{q}|M, z_f)$, through the relation

$$\delta_h(\boldsymbol{x}, z|M, z_f) = |1 + \delta_h^{\rm L}(\boldsymbol{q}|M, z_f)|\,|1 + \delta(\boldsymbol{x}, z)| - 1, \quad (1)$$

valid in the single-stream regime (CLMP). The nonlocality comes from the fact that the halo number density in $\boldsymbol{x}$ is determined by its initial value at the Lagrangian comoving position $\boldsymbol{q}$. Using a local version of the Press & Schechter (1974, hereafter PS) approach, we obtain

$$1 + \delta_h^{\rm L}(\boldsymbol{q}|M, z_f) = \sqrt{2\pi}\,|t_f - \epsilon_0(\boldsymbol{q})|\,\Theta\,[t_f - \epsilon_0(\boldsymbol{q})| \times \left\{ \sqrt{\frac{\pi}{2}}\frac{t_f\Sigma}{\sigma_M}\left[1 + {\rm erf}\left(\frac{t_f\Sigma}{\sqrt{2}\sigma_0\sigma_M}\right)\right] \exp\left(-\frac{t_f^2\Sigma^2}{2\sigma_0^2\sigma_M^2}\right)\right\}^{-1} + \sigma_0 \times \frac{\sigma_M^2}{\Sigma^2}\exp\left[-\frac{\epsilon_0(\boldsymbol{q})^2 - 2\epsilon_0(\boldsymbol{q})t_f + t_f^2\sigma_0^2/\sigma_M^2}{2\Sigma^2}\right]. \quad (2)$$

This result is obtained as follows: adopting the peak-background split, in the patch of fluid at $\boldsymbol{q}$ with Lagrangian size R_0, one writes the mean number density of halos in terms of the PS formula, with a local collapse threshold $t_f - \epsilon_0$ modulated by the background field ϵ_0, and finally removes the overall mean halo number density (eq. [3] below). Here $t_f \equiv \delta_c/D(z_f)$, with δ_c the critical threshold for collapse of a spherical perturbation and $D(z)$ the linear growth factor of density fluctuations normalized to unity at $z = 0$ [in the Einstein–de Sitter universe, $\delta_c \simeq 1.686$ and $D(z) = (1 + z)^{-1}$]; ϵ_0 is the linear mass fluctuation extrapolated to $z = 0$ and smoothed on R_0, with σ_0^2 its variance. Finally, $\sigma_M^2 = (2\pi^2)^{-1} \times \int_0^\infty dk\, k^2 P(k) W(kR)^2$, with $W(kR)$ the filter function and $P(k)$ the primordial power spectrum, is the variance on scale M of the linear density field ϵ_M and $\Sigma^2 \equiv \sigma_M^2 - \sigma_0^2$. Equation (2) actually generalizes equation (42) of CLMP in that collapse on the background scale R_0 is accounted for by the step function $\Theta[t_f - \epsilon_0]$: halos of mass $M \propto R^3$ cannot be present in a collapsed region of Lagrangian size $R_0 > R$. As stressed by CLMP (see also Porciani et al. 1998), this approach defines catalogs of halos unaffected by the cloud-in-cloud problem (see, e.g., Bond et al. 1991) up to the scale R_0. By expanding equation (2) to first order in ϵ_0 one obtains, for $\sigma_0 \ll \sigma_M$, $\delta_h^{\rm L}(\boldsymbol{q}|M, z_f) \simeq b_1^{\rm L}(M, z_f)\epsilon_0(\boldsymbol{q})$, where $b_1^{\rm L}(M, z_f) \equiv [t_f/\sigma_M^2 -$

[1] Theoretical Astrophysics Center, Juliane Maries Vej 30, 2100 Copenhagen Ø, Denmark.
[2] Dipartimento di Fisica "G. Galilei," Università di Padova, via Marzolo 8, 35131 Padova, Italy.
[3] Scuola Internazionale Superiore di Studi Avanzati, via Beirut 4, 34014 Trieste, Italy.
[4] Space Telescope Science Institute, 3700 San Martin Drive, Baltimore, MD 21218.

TABLE 1
HALO PARAMETERS

Class	N_{min}	N_{max}	Number of Halos	b_{MW}
S	25	35	1021	0.90
L	150	300	466	1.41

$1/t_f$] is the linear Lagrangian bias factor (MW). Note that in our approach the background scale R_0 appears as a fitting parameter that can be used to optimize the performance of the model. As we will see in § 3.2, the optimal value of σ_0 generally depends on the chosen Eulerian resolution scale and on the masses of the considered halo population.

When comparing with halos in numerical simulations, we will consider finite mass intervals, so we will have to replace δ_h^L in equation (2) with its weighted average, where the weight is given by the comoving conditional mass function

$$n_h(M, z_f) = \frac{\rho_h \exp(-t_f^2/2\sigma_M^2)}{2\pi M \sigma_M^2 \Sigma} \left|\frac{d\sigma_M^2}{dM}\right| \times \left\{ \sqrt{\frac{\pi}{2}} \frac{t_f \Sigma}{\sigma_M} \left[1 + \mathrm{erf}\left(\frac{t_f \Sigma}{\sqrt{2}\sigma_0 \sigma_M}\right)\right] + \sigma_0 \exp\left(-\frac{t_f^2 \Sigma^2}{2\sigma_0^2 \sigma_M^2}\right)\right\}, \quad (3)$$

with ρ_h the mean density. In addition, as discussed by CLMP, equation (1) can be generalized to multiple streaming as a Chapman-Kolmogorov–type relation

$$1 + \delta_h(x, z|M, z_f) = \int dq\, [1 + \delta_h^L(q|M, z_f)] \times \delta_D[x - x(q, z)], \quad (4)$$

with δ_D the Dirac function: each fluid element of Lagrangian size R_0 carries a "halo density charge" $n_h(M, z_f)[1 + \delta_h^L(q|M, z_f)]$ along its trajectory.

At every point $x = q + S$, with S the displacement of the qth Lagrangian element smoothed on R_0, we assign a halo density δ_h on the selected mass scale M according to equations (2) and (4). We extensively use this formulation of our bias scheme in § 3, where we test the model locally against a high-resolution N-body simulation.

2.2. *The Distribution of Halos*

In this section we compute the probability distribution $p(\delta_h)$ deriving from our bias model. We consider a mildly nonlinear density field in the laminar regime, though for comparisons with simulations we will also adopt the multistream generalization in equation (4).

Equation (1) can be recast, using mass conservation, in terms of the Jacobian determinant $J \equiv ||\partial x/\partial q||$ of the mapping from Lagrangian to Eulerian space, $q \rightarrow x(q, z) = q + S(q, z)$, namely $1 + \delta(x, z) = J(q, z)^{-1}$. In the Zeldovich approximation, $S(q, z) = -D(z)\nabla_q \varphi_0(q)$, where φ_0 is the linear peculiar gravitational potential, such that $\nabla_q^2 \varphi_0(q) = \epsilon_0(q)$. The probability distribution of the eigenvalues $\lambda_\alpha(q)$ ($\alpha = 1, 2, 3$) of the deformation tensor $\partial^2 \varphi_0(q)/\partial q_\alpha \partial q_\beta$ (Doroshkevich 1970) can be used to compute the one-point statistical properties of δ_h at any redshift $z \leq z_f$ in Eulerian space. Let us introduce the variables $L_0 \equiv [5(\mu_1^2 - 3\mu_2)]^{1/2}$ and $\mathcal{P}_0 \equiv (\lambda_1 - 2\lambda_2 + \lambda_3)/2(\mu_1^2 - 3\mu_2)^{1/2}$ (Reisenegger & Miralda-Escudé 1995), with $\mu_\alpha(q)$ ($\alpha = 1, 2, 3$) the invariants of the deformation tensor. Unlike the original eigenvalues, these variables are independent,

$$p_L(\epsilon_0, L_0, \mathcal{P}_0) = \left(\frac{e^{-\epsilon_0^2/2\sigma_0^2}}{\sqrt{2\pi}\sigma_0}\right)\left(\frac{2L_0^4 e^{-L_0^2/2\sigma_0^2}}{3\sqrt{2\pi}\sigma_0^5}\right)\left(\frac{3}{2} - 6\mathcal{P}_0^2\right), \quad (5)$$

where $0 \leq L_0 \leq \infty$ and $-1/2 \leq \mathcal{P}_0 \leq 1/2$. The Jacobian now reads $J = 1 - D\{675\epsilon_0 + 45D(L_0^2 - 5\epsilon_0^2) + D^2[2\sqrt{5}L_0^3\mathcal{P}_0(3 - 4\mathcal{P}_0^2) - 15L_0^2\epsilon_0 + 25\epsilon_0^3]\}/675$; it enters in equation (1) through mass conservation.

The probability $p(\delta_h)$ can then be computed by Monte Carlo generating realizations from the distribution in equation (5). Since we are interested in the Eulerian probability and equation (5) gives a Lagrangian distribution, we compute $p(\delta_h) = \int d\delta\, p_L(\delta_h, \delta)/(1 + \delta)$, where $p_L(\delta_h, \delta)$ is the joint Lagrangian probability for the Eulerian halo and mass overdensity fields (cf. eq. [14] in Kofman et al. 1994). In practice, $p(\delta_h)$ is obtained by (i) generating realizations for $\epsilon_0, L_0, \mathcal{P}_0$; (ii) computing J, and δ_h through equation (1) or equation (4); (iii) weighting the contribution to the probability of δ_h by the factor J.

3. TESTING THE MODEL

3.1. *Comparing the Probability Distribution Function*

We test our predictions for $p(\delta_h)$ against a high-resolution N-body simulation from the data bank of cosmological simulations provided by the Hydra Consortium and produced using the Hydra N-body code (Couchman, Thomas, & Pierce 1995). The simulation (RUN 501) evolves 128^3 particles on a 128^3 cubic mesh with periodic boundaries. The box size is 100 h^{-1} Mpc and the particle mass $1.32 \times 10^{11}\ h^{-1}\ M_\odot$. The initial conditions are Gaussian with a Cold Dark Matter spectrum with shape parameter $\Gamma = 0.25$, density parameter $\Omega = 1$ and zero cosmological constant. The simulation output corresponds to $\sigma_8 = 0.64$, where σ_8 is the rms linear mass fluctuation in spheres of 8 h^{-1} Mpc. At this epoch, the characteristic virializing halo mass, M_*, defined by $\sigma_{M_*} = \delta_c$, is $0.66 \times 10^{13}\ h^{-1}\ M_\odot$, i.e., 50.13 particles. To compare our predictions to the N-body outcome, we need a halo catalog from the simulation. We adopt the spherical overdensity (SO) halo finder (Lacey & Cole 1994) to identify spherical regions with mean overdensity $\kappa = 178$, leading to 5025 halos with more than 20 particles. We then consider two classes of objects: class S contains halos with $0.5 \lesssim M/M_* \lesssim 0.7$; class L has $3 \lesssim M/M_* \lesssim 6$ (see Table 1). In Figure 1 we plot $p(\delta_h)$ obtained with our bias scheme against the N-body outcomes. The model prediction for $z = z_f = 0$ is computed as in § 2. σ_0 has been tuned to optimize the agreement with the numerical outputs. The simulation probability distribution has been extracted after smoothing the halo distribution by a Gaussian filter $W = \exp(-k^2/2k_f^2)$ with resolution k_f. The prediction of a linear Eulerian bias model is also plotted in Figure 1: the mass distribution in the simulation has been smoothed with the same filter used for the halo overdensity and the resulting δ is multiplied by the Eulerian bias factor

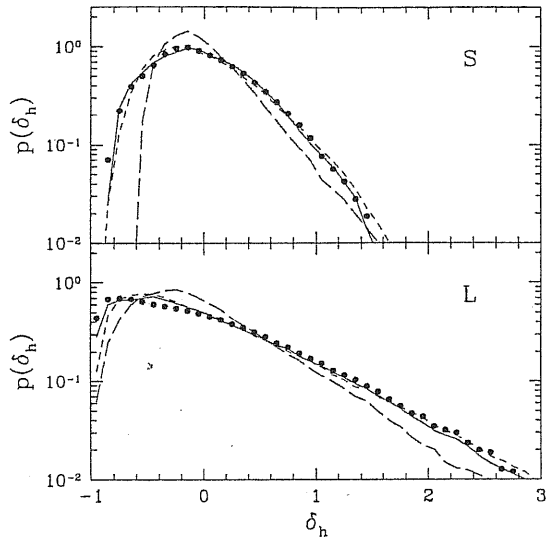


FIG. 1.—One-point distribution function of δ_h. Short-dashed lines refer to the data obtained by performing 10^5 random generations of the variables ϵ_0, L_0, and $\mathcal{P}_0$ ($k_0 = 0.180\ h$ Mpc^{-1}, $\sigma_0^2 = 0.217$); filled dots show the N-body output. The long-dashed lines represent the linear bias prediction, while the solid lines are the outcome of the numerical version of our model described in § 3.2. *Top:* Class S ($k_f = 0.314\ h$ Mpc^{-1}). *Bottom:* Class L ($k_f = 0.360\ h$ Mpc^{-1}).

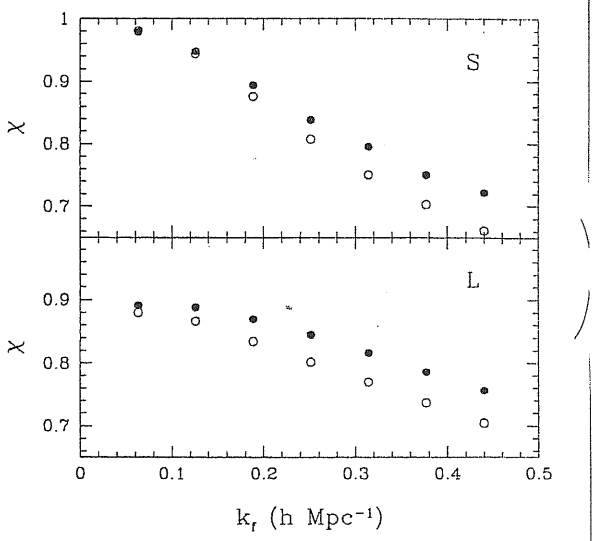


FIG. 2.—Cross-correlation coefficients vs. filtering scale. Filled and empty dots represent, respectively, our model and any Eulerian linear bias scheme.

$b_{\rm MW} = 1 + [\delta_c/\sigma_M^2 - 1/\delta_c]$ (MW), reported in Table 1. Our model accurately reproduces the tail of the distribution for positive δ_h, while for $\delta_h < 0$ it favors moderate underdensities ($\delta_h \sim -0.5$) with respect to the N-body simulation. The linear bias prescription instead produces much a less skewed distribution with a higher peak and severely underestimates the probability of very underdense regions. Our model can be further improved by adopting the multistream version introduced in § 3.2, whose predictions, also plotted in Figure 1, are in excellent agreement with the N-body outputs.

3.2. *Cross-Correlations*

We also performed a much more severe point-by-point test, implementing a fully numerical version of our bias scheme as follows. We consider a computational box as large as the N-body one, but sampled with lower resolution: 64^3 particles on a 64^3 grid (using 128^3 particles on a 128^3 mesh gave identical results). Each particle is moved to its final position according to the "truncated" Zeldovich approximation (Coles et al. 1993; Melott, Pellmann, & Shandarin 1994); i.e., prior computing the displacement, we remove initial power in high-frequency modes by a Gaussian filter with resolution k_0. We require the amplitudes and phases of the linear density field to be identical to the simulation ones (at least for the Fourier modes present in both grids). Each particle is then associated to the linear field $\epsilon_0(\boldsymbol{q}) = -\nabla_q \cdot \boldsymbol{S}(\boldsymbol{q})$. In such a way, every particle is endowed with its own halo density charge $n_h[1 + \delta_h^L(\boldsymbol{q})]$, computed as described in § 2.1. Note that, while for σ_0^2 we used Gaussian smoothing, σ_M is calculated with top-hat filtering. The halo density charge is carried by the particles and eventually assigned to the 64^3 grid through the triangular-shaped cloud scheme. The corresponding halo overdensity field in the multistream regime $\delta_h^{\rm mod}$ is then computed and smoothed with a Gaussian filter with $k_f \le k_0$. The result has to be compared with the halo field $\delta_h^{\rm sim}$ extracted from the N-body simulation, smoothed on the same scale. To quantify the agreement between the two, we compute their cross-correlation coefficient $\chi \equiv \langle \delta_h^{\rm mod} \delta_h^{\rm sim} \rangle / (\sigma_h^{\rm mod} \sigma_h^{\rm sim})$, where $\sigma_h^i \equiv \langle \delta_h^{i\,2} \rangle^{1/2}$, and the average is performed over the grid points. A value $|\chi| = 1$ means that the two fields are proportional, while $\chi = 0$ for uncorrelated fields. We tune the truncation k_0 to optimize our bias scheme. Since χ turns out to depend very weakly on k_0 (in a wide range around its optimal value), we can choose it so that $\sigma_h^{\rm mod} = \sigma_h^{\rm sim}$ while keeping the maximum allowed value of χ. For each resolution this is obtained with $\sigma_0^2 = 1.65\sigma_f^2 + C$, with $C = 1.45$ for class S and $C = 0.54$ for class L. Note that in the MW model the value of $\sigma_h^{\rm mod}/\sigma_h^{\rm sim}$ is not adjustable and generally differs from 1. For the resolutions considered in Figure 3, $\sigma_h^{\rm mod}/\sigma_h^{\rm sim}$ is 0.81 for class S and 0.90 for class L. Moreover, considering smaller smoothing lengths, this ratio can differ from 1 even by 30%. This implies that the MW method is unable to accurately predict the average properties of the bias distribution. Figure 2 reports χ-values for the optimized model vs. the resolution k_f.

To evaluate the performance of *any* Eulerian linear bias model, we also show the cross-correlation coefficient when the smoothed mass density field from the simulation is used instead of the output mass of our model. This test does not depend on the value of the linear bias. For each smoothing length our model reproduces the halo density field of the simulation much more accurately than the linear biasing scheme. The performance of the two models is similar only for very large smoothing lengths.

In Figure 3 we show the scatter obtained by plotting $\delta_h^{\rm mod}$ versus $\delta_h^{\rm sim}$ for our model and for the linear bias scheme. In this case, we adopt the MW bias factor in Table 1. Even though on average our bias scheme gives better predictions (especially for class S and for underdense regions), some scatter in the relation $\delta_h^{\rm mod}$ versus $\delta_h^{\rm sim}$ persists. To test whether this is a result of our simplified dynamics, we generate new halo maps taking the particle displacements directly from the N-body simulation.

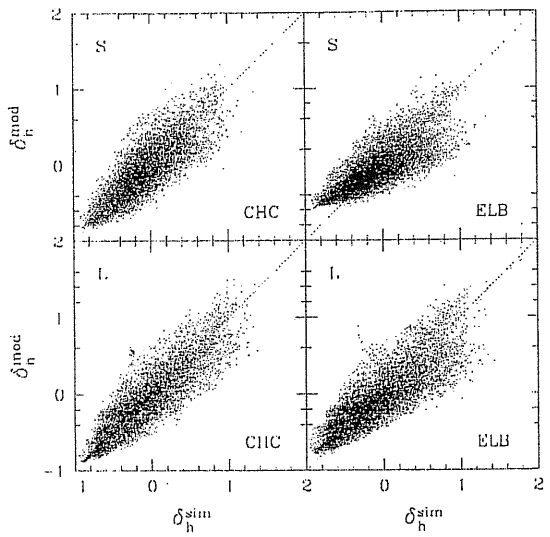


FIG. 3.—Model predictions for δ_h vs. N-body outcome. *Left panels:* Our bias scheme (conserved halo density charge, or CHC). *Right panels:* Eulerian linear bias (ELB) model in the MW version. *Top panels:* Class S (with $k_f = 0.260\ h\ \mathrm{Mpc}^{-1}$). *Bottom panels:* Class L (with $k_f = 0.301\ h\ \mathrm{Mpc}^{-1}$). A one-in-eight random selection is shown.

The results are in excellent agreement with those obtained with the Zeldovich approximation, indicating that the local PS approach, being unable to accurately model the Lagrangian halo counting (Fig. 8 in White 1996), is actually responsible for the scatter. We will address this point in a future work.

4. DISCUSSION AND CONCLUSIONS

We devised a simple and fast semianalytical technique that allows to study the spatial distribution of dark matter halos in terms of their local number density contrast. Our method, which is based on a Lagrangian halo identification algorithm plus the Zeldovich approximation for the matter dynamics, was successfully tested against the distribution of halos extracted from a high-resolution N-body simulation. Possible improvements should go in the direction of refining the halo selection criterion in Lagrangian space, e.g., using the ellipsoidal collapse model or the peak theory as in (Bond & Myers 1996; see also Monaco 1998).

As stressed by CLMP, our model can be applied to study the evolution of galaxy biasing, once the relation between the galaxies and the hosting dark matter halos is specified (see, e.g., Matarrese et al. 1997). In particular, defining the bias field such that $\delta_h \equiv \mathfrak{b}\delta$, from equation (1) one obtains $\mathfrak{b}[\boldsymbol{x}(\boldsymbol{q}, z), z] = 1 + \delta_h^L(\boldsymbol{q}|z_f)/[1 - J(\boldsymbol{q}, z)]$. Tegmark & Peebles (1998) have recently stressed the importance of the asymptotic trend of the bias factor. We can analyze this issue in the present context by considering a galaxy population conserved in number after an initial merging phase (i.e., for varying z at fixed z_f). In the Einstein–de Sitter case, we recover the "debiasing" predicted by linear theory: $\mathfrak{b} \to 1$ as $z \to -1$. On the other hand, if $\Omega < 1$, $\mathfrak{b}$ tends to a space-dependent value that generally differs from 1, linear theory would predict $b \to 1 + (b_0 - 1)/D_{-1}(\Omega_0)$, as $z \to -1$, with $b_0 \equiv b(z = 0)$ and $D_{-1} \equiv D(z = -1)$.

Our method can be used to analyze the coarse-grained statistical properties of galaxies and clusters at various redshifts, e.g., applying semianalytical techniques to relate the dark matter halo distribution to that of luminous objects like galaxies (see, e.g., Kauffmann, Nusser, & Steinmetz 1997). It can be further implemented to generate very large mock maps of these objects in our past light cone, a problem made of compelling relevance by the ongoing wide-field redshift galaxy surveys like the 2 Degree Field Survey (2dF) and the Sloan Digital Sky Survey (SDSS). Maps of the X-ray cluster distribution may also be produced by the present method.

We thank S. Cole and C. Lacey for providing the SO halo finder and the Hydra Consortium for N-body simulations (http://coho.astro.uwo.ca/pub/data.html). P. C. has been supported by the Danish NRF at TAC; S. M. and C. P. by the Italian MURST. C. P. thanks support from NASA ATP-NAG5-4236 grant.

REFERENCES

Bond, J. R., Cole, S., Efstathiou, G., & Kaiser, N. 1991, ApJ, 379, 440
Bond, J. R., & Myers, S. T. 1996, ApJS, 103, 1
Catelan, P., Lucchin, F., Matarrese, S., & Porciani, C. 1998, MNRAS, 297, 692 (CLMP)
Cole, S., & Kaiser, N. 1989, MNRAS, 237, 1127
Coles, P., Melott, A. L., & Shandarin, S. F. 1993, MNRAS, 260, 765
Couchman, H. M. P., Thomas, P. A., & Pearce, F. R. 1995, ApJ, 452, 797
Doroshkevich, A. G. 1970, Astrophysics, 6, 320
Kauffmann, G., Nusser, A., & Steinmetz, M. 1997, MNRAS, 286, 795
Kofman, L., Bertschinger, E., Gelb, J. M., Nusser, A., & Dekel, A. 1994, ApJ, 420, 44
Lacey, C., & Cole, S. 1994, MNRAS, 271, 676
Matarrese, S., Coles, P., Lucchin, F., & Moscardini, L. 1997, MNRAS, 286, 115
Melott, A. L., Pellman, T. F., & Shandarin, F. S. 1994, MNRAS, 269, 628
Mo, H. J., & White, S. D. M. 1996, MNRAS, 282, 347 (MW)
Monaco, P. 1998, Fundam. Cosmic Phys., 19, 157
Porciani, C., Matarrese, S., Lucchin, F., & Catelan, P. 1998, MNRAS, in press (astro-ph/9801290)
Press, W. H., & Schechter, P. 1974, ApJ, 187, 425
Reisenegger, A., & Miralda-Escudé, J. 1995, ApJ, 449, 476
Tegmark, M., & Peebles, P. J. E. 1998, ApJL, 500, L79
White, S. D. M. 1996, in Cosmology and Large-Scale Structure, ed. R. Schaeffer, J. Silk, M. Spiro, & J. Zinn-Justin (Elsevier), 349
Zeldovich, Ya.B. 1970, A&A, 5, 84

Appendix A

Random fields

In this appendix we introduce some basic concepts of the theory of random fields. After a general discussion based on functional methods, we focus our attention on Gaussian fields obtaining many relations that are widely used in the main text.

Historically speaking, the statistical analysis of random fields was first developed to discuss electrical noise in communication devices (Rice 1945; reprinted in Wax 1954). Nowadays it represents a well mature branch of mathematical statistics which finds a broad spectrum of applications in many problems of physics, engineering and economics. A complete discussion of the subject can be found in the books by Adler (1981) and Vanmarcke (1983).

A.1 Random variables

In nature there exist some systems for which it is not possible to predict the evolution (for instance because of the incomplete knowledge of the initial conditions and (or) some other unknown factors). From the mathematical point of view, these systems are described through random variables. A random (or stochastic) variable, $\mathcal{S}$, is determined by a set of numbers (its possible values) and a probability measure defined over this set. The idea is to consider a large (formally infinite) collection of identical systems (called, respectively, the ensemble and its realizations), and build the cumulative probability distribution function $I_{\mathcal{S}}(\alpha)$ by computing the fraction of systems in the ensemble for which $\mathcal{S}$ assumes values smaller than α. Denoting the operation of averaging over the ensemble by $\langle\cdot\rangle$ and using the symbol $\Theta(x)$ to denote the Heaviside step function, one has $I_{\mathcal{S}}(\alpha) \equiv \langle\Theta(\alpha - \mathcal{S})\rangle$. The probability density function, $\mathcal{P}_{\mathcal{S}}(\alpha)$, such that $\mathcal{P}_{\mathcal{S}}(\alpha)d\alpha$ gives the probability over the ensemble to find $\alpha < \mathcal{S} \leq \alpha + d\alpha$, is then obtained by differentiation: $\mathcal{P}_{\mathcal{S}}(\alpha) \equiv \langle\delta_D(\alpha - \mathcal{S})\rangle = dI_{\mathcal{S}}/d\alpha$ (with $\delta_D(x)$ the Dirac delta distribution). The statistical properties of $\mathcal{S}$ are completely determined by the probability density function, for any expectation value can be obtained from $\mathcal{P}_{\mathcal{S}}(\alpha)$ by integration

$$\langle f(\mathcal{S})\rangle = \int f(\alpha)\mathcal{P}_{\mathcal{S}}(\alpha)d\alpha \,, \tag{A.1}$$

with f an arbitrary function. Using $f(x) = 1$, it is possible to fix the normalization

$$\int \mathcal{P}_{\mathcal{S}}(\alpha)d\alpha = 1 \,. \tag{A.2}$$

Let us consider two stochastic variables, $\mathcal{S}_1$ and $\mathcal{S}_2$, we define their joint probability function as $\mathcal{P}_{\mathcal{S}_1\mathcal{S}_2}(\alpha_1, \alpha_2) = \langle\delta_D(\alpha_1 - \mathcal{S}_1)\delta_D(\alpha_2 - \mathcal{S}_2)\rangle$. We can also introduce a conditional probability density,

$\mathcal{P}_{\mathcal{S}_1|\mathcal{S}_2}(\alpha_1|\alpha_2)$, which is computed by considering only those realizations of the ensemble where $\mathcal{S}_2$ takes the fixed value α_2. In general,

$$\mathcal{P}_{\mathcal{S}_1|\mathcal{S}_2}(\alpha_1|\alpha_2) = \frac{\mathcal{P}_{\mathcal{S}_1\mathcal{S}_2}(\alpha_1,\alpha_2)}{\int \mathcal{P}_{\mathcal{S}_1\mathcal{S}_2}(\alpha_1,\alpha_2)d\alpha_1} . \tag{A.3}$$

We say that $\mathcal{S}_1$ and $\mathcal{S}_2$ are statistically independent if $\mathcal{P}_{\mathcal{S}_1|\mathcal{S}_2}(\alpha_1|\alpha_2) = \mathcal{P}_{\mathcal{S}_1}(\alpha_1)$, that is to say if their joint probability density function $\mathcal{P}_{\mathcal{S}_1\mathcal{S}_2}(\alpha_1,\alpha_2) = \mathcal{P}_{\mathcal{S}_1}(\alpha_1)\mathcal{P}_{\mathcal{S}_2}(\alpha_2)$.

The characteristic function of the stochastic variable $\mathcal{S}$ is the average

$$C_{\mathcal{S}}(u) = \langle e^{iu\mathcal{S}} \rangle = \int \mathcal{P}_{\mathcal{S}}(\alpha) e^{iu\alpha} d\alpha , \tag{A.4}$$

with u a real (non-stochastic) variable. In practice, $C_{\mathcal{S}}(u)$ is the Fourier transform of the probability density which can then be recovered through the inverse operation

$$\mathcal{P}_{\mathcal{S}}(\alpha) = \frac{1}{2\pi} \int C_{\mathcal{S}}(u) e^{-iu\alpha} du . \tag{A.5}$$

Because $\mathcal{P}_{\mathcal{S}}(\alpha) \geq 0$, $C_{\mathcal{S}}(u)$ must be a positive defined function, i.e., for every $n \geq 1$

$$\sum_{k=1}^{n} \sum_{j=1}^{n} C_{\mathcal{S}}(u_k - u_j) a_k^* a_j^* \geq 0 , \tag{A.6}$$

with $u_1, \ldots, u_n$ an arbitrary set of real numbers and $a_1, \ldots, a_n$ an arbitrary set of complex numbers (a^* denotes the complex conjugate of a). It is possible to show that every positive defined function with the property $C_{\mathcal{S}}(0) = 1$ is a characteristic function of some random variable (Bochner-Khinchine theorem; see e.g. Gnedenko 1989).

The n-th moment μ_n of $\mathcal{S}$, when it exists, can be obtained from the characteristic function by differentiation

$$\mu_n = \langle \mathcal{S}^n \rangle = \frac{1}{i^n} \frac{d^n C_{\mathcal{S}}(u)}{du^n}\bigg|_{u=0} . \tag{A.7}$$

Hence, the McLaurin expansion of the characteristic function is given by the moments according to

$$C_{\mathcal{S}}(u) = 1 + \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} \mu_n . \tag{A.8}$$

The cumulants or semi-invariants, k_n, are instead defined through the relation

$$k_n = \frac{1}{i^n} \frac{d^n \ln C_{\mathcal{S}}(u)}{du^n}\bigg|_{u=0} . \tag{A.9}$$

which gives

$$\ln C_{\mathcal{S}}(u) = \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} k_n . \tag{A.10}$$

It follows that the first n cumulants can be expressed by the first n moments and vice versa. These relations up to $n = 4$ read explicitly

$$\begin{aligned} k_1 &= \mu_1 , \qquad k_2 = \mu_2 - \mu_1^2 , \qquad k_3 = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3 , \\ & \qquad\qquad\qquad\qquad k_4 = \mu_4 - 3\mu_2^2 - 4\mu_1\mu_3 + 12\mu_1^2\mu_2 - 6\mu_1^4 . \end{aligned} \tag{A.11}$$

Note that the characteristic function of the sum of two independent random variables, $\mathcal{S}_1$ and $\mathcal{S}_2$, is given by the product $C_{\mathcal{S}_1}(u)C_{\mathcal{S}_2}(u)$. Consequently, the n-th order cumulants of $\mathcal{S}_1 + \mathcal{S}_2$ are simply obtained by summing up the k_n terms of the two independent random variables. It is clear from equations (A.10) and (A.5) that the whole set of cumulants (or moments) completely determines $\mathcal{P}_{\mathcal{S}}(\alpha)$. In particular, we have:

- for $k_2 = k_3 = \cdots = 0$: $C_{\mathcal{S}}(u) = \exp(iuk_1)$, $\mathcal{P}_{\mathcal{S}}(\alpha) = \delta_D(\alpha - k_1)$;
- for $k_3 = k_4 = \cdots = 0$: $C_{\mathcal{S}}(u) = \exp(iuk_1 - u^2k_2/2)$, $\mathcal{P}_{\mathcal{S}}(\alpha) = (2\pi k_2)^{-1/2}\exp[-(x-k_1)^2/2k_2]$.

These examples, respectively the Dirac delta and the Gaussian distribution, represent the only probability density functions with a finite number of non-vanishing cumulants (H anggi & Talkner, 1980).

A.2 Random fields

Let us consider a space of functions (or, more generally, of distributions), $\mathcal{F}$, characterized by some precise requirements (e.g. functions having a Lebesgue-integrable square modulus over a specified domain) and endowed with a scalar product and a norm. An n-dimensional scalar random field, $\mathcal{R}(\mathbf{x})$, is a set of random variables, one for each point in a n-dimensional real space, $\mathbf{x}$, defined by a probability functional, $\mathcal{P}[\hat{\mathcal{R}}(\mathbf{x})]$, which specifies the probability for the occurrence of a particular realization of the field (i.e. of the function $\hat{\mathcal{R}}(\mathbf{x}) \in \mathcal{F}$) over the ensemble. Such a probability distribution is normalized by requiring that

$$\int D[\mathcal{R}(\mathbf{x})]\mathcal{P}[\mathcal{R}(\mathbf{x})] = 1 \,, \tag{A.12}$$

where $\int D[\mathcal{R}]$ denotes functional integration over $\mathcal{F}$ with measure $D[\mathcal{R}]$.

Given a generic function $J(\mathbf{x})$ with finite norm, we introduce the partition functional

$$\mathcal{Z}[J(\mathbf{x})] = \int D[\mathcal{R}(\mathbf{x})]\mathcal{P}[\mathcal{R}(\mathbf{x})] \exp\left[i\int d^3x\, \mathcal{R}(\mathbf{x})J(\mathbf{x})\right] \,. \tag{A.13}$$

The disconnected correlation functions of order n are then defined as the functional derivatives of $\mathcal{Z}$

$$\xi_n^{\mathrm{dis}}(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \langle \mathcal{R}(\mathbf{x}_1)\ldots\mathcal{R}(\mathbf{x}_n)\rangle = \int D[\mathcal{R}]\mathcal{P}[\mathcal{R}]\,\mathcal{R}(\mathbf{x}_1)\ldots\mathcal{R}(\mathbf{x}_n) = \\ = \frac{1}{i^n}\frac{\delta^{(n)}\mathcal{Z}}{\delta J(\mathbf{x}_1)\ldots\delta J(\mathbf{x}_n)}\bigg|_{J=0} \,. \tag{A.14}$$

In other words, $\mathcal{Z}[J(\mathbf{x})]$ is the generating functional of the disconnected correlation functions, because the ξ_n^{dis} are the coefficients of its McLaurin expansion

$$\mathcal{Z}[J] = 1 + \sum_{n=1}^{\infty}\frac{i^n}{n!}\int d^3x_1\cdots\int d^3x_n\,\xi_n^{\mathrm{dis}}(\mathbf{x}_1,\ldots,\mathbf{x}_n)J(\mathbf{x}_1)\ldots J(\mathbf{x}_n) \,. \tag{A.15}$$

It is apparent that ξ_n^{dis} is unchanged by interchange of any subset of its arguments. In a similar way, the connected correlation functions, $\xi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n)$, are generated by the functional

$$\mathcal{W}[J(\mathbf{x})] = \ln\mathcal{Z}[J(\mathbf{x})] \,, \tag{A.16}$$

$$\mathcal{W}[J] = \sum_{n=1}^{\infty} \frac{i^n}{n!} \int d^3x_1 \cdots \int d^3x_n \, \xi_n(\mathbf{x}_1, \ldots, \mathbf{x}_n) J(\mathbf{x}_1) \ldots J(\mathbf{x}_n) \; . \tag{A.17}$$

The ξ_n have the nice property that the n-th connected correlation function of the sum of two independent random fields is simply given by the sum of the ξ_n of the original fields. More importantly, any contribution to $\xi_n^{\rm dis}$ which can be separated as $\langle \mathcal{R}(\mathbf{x}_1) \ldots \mathcal{R}(\mathbf{x}_m) \rangle \langle \mathcal{R}(\mathbf{x}_{m+1}) \ldots \mathcal{R}(\mathbf{x}_n) \rangle$ has been removed from ξ_n. Thus $\xi_n \to 0$ as any subset of $(\mathbf{x}_1, \ldots, \mathbf{x}_n)$ is removed to infinite separation. Eq. (A.16) can be used to relate the first few ξ_n and $\xi_n^{\rm disc}$; for example, assuming $\langle \mathcal{R}(\mathbf{x}) \rangle = 0$ and $n = 4$, one has

$$\begin{aligned} \xi_4(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) = \xi_4^{\rm dis}(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) - \xi_2^{\rm dis}(\mathbf{x}_1, \mathbf{x}_2) \xi_2^{\rm dis}(\mathbf{x}_3, \mathbf{x}_4) \\ - \xi_2^{\rm dis}(\mathbf{x}_1, \mathbf{x}_3) \xi_2^{\rm dis}(\mathbf{x}_2, \mathbf{x}_4) - \xi_2^{\rm dis}(\mathbf{x}_1, \mathbf{x}_4) \xi_2^{\rm dis}(\mathbf{x}_2, \mathbf{x}_3) \; . \end{aligned} \tag{A.18}$$

The functional series (A.15) and (A.17) show that the partition functional (and, consequently, the statistics of the system) is uniquely determined by the whole hierarchy of the correlation functions.

It is worth noticing that differentiation and integration of random fields may not always be possible especially when they possess power on all scales. Two approaches are then possible: define these operations as appropriate limits in the mean square (e.g. Vanmarcke 1983) or compute them after smoothing the field as described in §2.3.

A.3 Homogeneity and isotropy

Random fields are used in cosmology to describe fluctuations in the mass-energy distribution with respect to a smooth background. As a consequence of the cosmological principle, one is interested in homogeneous and isotropic stochastic fields with zero mean. A random field is homogeneous if all its connected correlation functions are invariant under simultaneous translations of their arguments by the same vector. For instance, this implies that $\langle \mathcal{R}(\mathbf{x}) \rangle$ is spatially independent and $\xi_2(\mathbf{x}_1, \mathbf{x}_2)$ is a function only of the relative separation $\mathbf{x}_{12} = \mathbf{x}_2 - \mathbf{x}_1$. Isotropy instead means invariance of the correlation functions under a global rotation of their arguments. In this case ξ_2 depends only on $x_{12} = |\mathbf{x}_{12}|$. Switching to Fourier space, spatial homogeneity and isotropy imply that the power spectrum, $P(\mathbf{k})$, defined through the relation $\langle \tilde{\mathcal{R}}(\mathbf{k}) \tilde{\mathcal{R}}^*(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}') P(\mathbf{k})$, is a function of $k = |\mathbf{k}|$ (here $\tilde{\mathcal{R}}$ denotes the Fourier transform of $\mathcal{R}$). In particular, if the random field assumes real values, one has $\langle \tilde{\mathcal{R}}(\mathbf{k}) \tilde{\mathcal{R}}(\mathbf{k}') \rangle = (2\pi)^3 \delta_D(\mathbf{k} + \mathbf{k}') P(k)$. In this case, the power spectrum and the two-point correlation function are simply related as follows (Wiener-Khinchine theorem)

$$\xi_2(x) = \frac{1}{2\pi^2} \int k^2 P(k) \frac{\sin(kx)}{kx} \, dk \; . \tag{A.19}$$

A.4 Gaussian random fields

A Gaussian scalar random field, $G(\mathbf{x})$, with vanishing mean value is defined by the probability functional

$$\mathcal{P}[G] = (\det K)^{1/2} \exp\left\{ -\frac{1}{2} \int G(\mathbf{x}) K(\mathbf{x}, \mathbf{y}) G(\mathbf{y}) \, d^3x \, d^3y \right\} \; , \tag{A.20}$$

where $K(\mathbf{x}, \mathbf{y})$ is an invertible operator, symmetric with respect to the variables $\mathbf{x}$ and $\mathbf{y}$. In this case, the partition functional and the generating functional for the connected correlations read

$$\mathcal{Z}[J] = \exp \mathcal{W}[J] = \exp\left\{ -\frac{1}{2} \int J(\mathbf{x}) K^{-1}(\mathbf{x}, \mathbf{y}) J(\mathbf{y}) \, d^3x \, d^3y \right\} \; , \tag{A.21}$$

with K^{-1} the functional inverse of the operator K, i.e.

$$\int K(\mathbf{x},\mathbf{y})\, K^{-1}(\mathbf{y},\mathbf{z})\, d^3y = \delta_D(\mathbf{x}-\mathbf{z})\ . \tag{A.22}$$

Eq. (A.17) then gives

$$\xi_2(\mathbf{x},\mathbf{y}) = K^{-1}(\mathbf{x},\mathbf{y}), \qquad \xi_n(\mathbf{x}_1,\ldots,\mathbf{x}_n) = 0 \ \text{ for } n>2\ . \tag{A.23}$$

Thus, a Gaussian random field with vanishing mean value is completely determined by its two-point statistics. For instance, for $\xi_n^{\rm dis}$ one finds: $\xi_n^{\rm dis}=0$ (n odd) and $\xi_n^{\rm dis}=(n-1)!!\,\xi_2^{n/2}$ (n even), where $(n-1)!!=(n-1)(n-3)\ldots 1$. Here, for brevity, labels on the connected functions are omitted: these are recovered by grouping the set of points at which $\xi_n^{\rm dis}$ is evaluated into all the possible pairs and by associating a function ξ_2 to each pair. For homogeneous and isotropic stochastic fields, one also has $P(k)=1/\tilde{K}(k)$.

The joint probability function of the values of $G(\mathbf{x})$ in the points $\mathbf{x}_1,\ldots,\mathbf{x}_n$ is a multivariate Gaussian distribution

$$P_n(\alpha_1,\ldots,\alpha_n)d\alpha_1\ldots d\alpha_n = [(2\pi)^n \det M]^{-1/2} \exp\left\{-\frac{1}{2}\alpha_i M^{-1}_{ij}\alpha_j\right\} d\alpha_1\ldots d\alpha_n\ , \tag{A.24}$$

with $M_{ij}=\langle\alpha_i\alpha_j\rangle=\xi_2(x_{ij})$ the covariance matrix. Note that this equation is simply the discrete version of eq. (A.20). Moreover, the field and its derivatives are jointly Gaussian. In this case, the elements of the covariance matrix depend only on moments over the power spectrum

$$\sigma_m^2 \equiv \frac{1}{(2\pi)^n}\int k^{2m}P(k)\, d^nk \tag{A.25}$$

where m and n denote the order of the spectral moment and the dimensionality of the random field, respectively. It is often convenient to define some new parameters by combining spectral moments of different orders,

$$R_G \equiv \sqrt{n}\,\frac{\sigma_1}{\sigma_2}\ , \qquad \gamma \equiv \frac{\sigma_1^2}{\sigma_0\,\sigma_2}\ . \tag{A.26}$$

The variable R_G is a measure of the coherence scale in the field, while γ is a measure of the width of the power spectrum: $\gamma=1$ corresponds to an infinitesimal shell in k-space.

Rigorously, in order for a field to be homogeneous and Gaussian, its Fourier modes, $\tilde{G}(\mathbf{k}) = |\tilde{G}(\mathbf{k})|e^{i\theta(\mathbf{k})}$, must be statistically independent, have random phases and have moduli which are Rayleigh distributed, i.e., for a given $\mathbf{k}$, their probability distribution is as follows:

$$\mathcal{P}[|\tilde{G}(\mathbf{k})|,\theta(\mathbf{k})] = \exp\left[-\frac{|\tilde{G}(\mathbf{k})|^2}{2P(k)}\right]\frac{|\tilde{G}(\mathbf{k})|}{P(k)}\, d|\tilde{G}(\mathbf{k})|\,\frac{d\theta(\mathbf{k})}{2\pi}\ . \tag{A.27}$$

The real and imaginary parts of $\tilde{G}(\mathbf{k})$ are then mutually independent and Gaussian distributed.

A.5 Hierarchical scaling

We say that a random field displays a scaling property if its connected correlation functions of any order can be expressed in terms of ξ_2. As an example, we introduce here the definition of hierarchical scaling that have used in the main text.

A field is termed hierarchical if

$$\xi_j(\mathbf{x}_1,\ldots,\mathbf{x}_n) = \sum_\alpha Q_{j,\alpha} \sum_{\{ab\}} \prod^{j-1} \xi_2(r_{ab}) \,, \tag{A.28}$$

where the formula must be interpreted as follows. Associated to each term in eq. (A.28) there is a tree graph (connected with no cycles) of j vertices and $j-1$ edges. The vertices correspond to the points $\mathbf{x}_1,\ldots,\mathbf{x}_j$, and the lines which join different vertices correspond to factors $\xi_2(x_{ab})$. The sum over α denotes topologically distinct graphs; the sum over $\{ab\}$ is over relabelings within α. This implies that the averaged correlation functions over a sphere of radius R obey

$$\bar{\xi}_j(R) = S_j\,[\bar{\xi}_2(R)]^{j-1} \,, \tag{A.29}$$

where the S_j are constants. Sometimes, eq. (A.29) is used to define a hierarchical distribution. Notice that the eq. (A.29) and the relation $\bar{\xi}_2(R) \propto R^{-\gamma}$ (with $\gamma = \text{const}$) characterize a fractal distribution.

A.6 Ergodicity

In cosmology, when we assume that fluctuations in the energy density (or in the gravitational potential) at some primordial epoch (e.g. after the end of the inflationary phase) form a 3-dimensional random field, we mean that they represent one particular realization extracted from an ideal ensemble of functions. Theoretical investigation is then only able to compute statistical quantities averaged over this ensemble. On the other hand, since astronomical observations refer to a single realization of the random field, in the observable universe we can only take spatial averages over large volumes. Therefore, theoretical results can be tested against observational data only if the primordial fluctuations are ergodic, i.e. if spatial averages over large volumes in a single realization are equal to expectations over the ensemble. A Gaussian random field can be shown to be ergodic if and only if its power spectrum is continuous (e.g. Adler 1981). In general, the validity of the ergodic hypothesis depends on the ratio between the length scale over which we perform the spatial average and the scale at which spatial correlations become negligibly small. When the observationally surveyed volume contains many statistically independent subsamples (i.e. it is a fair sample of the universe), ergodicity is expected to hold.

Appendix B

Kinetic theory

In this appendix, we briefly review the standard methods of kinetic theory both for long-range and short-range binary interactions. We first introduce the correlation function representation of the distribution functions. The exact dynamics of a collection of particles is then discussed in terms of the BBGKY hierarchy of equations. The case of N identical bodies interacting only through Newtonian gravity in an expanding universe is discussed in detail. Finally, particular attention is devoted to derive common fluid approximations from the kinetic equations.

B.1 Probability distributions and correlation functions

Let us consider a system composed by N classical particles with canonic coordinates $\mathbf{q}_i, \mathbf{p}_i$ $(i = 1, \ldots, N)$. [1] The actual configuration of the system at any given time is represented by a single point in the $6N$-dimensional Γ-phase space $\mathbf{q}_1, \ldots, \mathbf{q}_N, \mathbf{p}_1, \ldots, \mathbf{p}_N \equiv (1, \ldots, N) \equiv (q, p)$. Now imagine a large number of independent replicas (*ensemble*) of the same system. Suppose there are $\mathcal{N}$ systems in the ensemble, then at any given time the state of the ensemble is fully described by $\mathcal{N}$ points in Γ-space. Let us introduce the N-body joint probability density in Γ-space, which is defined as follows: $f_N(q, p, t)dqdp$ represents the fraction of system points in the phase volume $dqdp$ about the point (q, p) at time t. In other words, $f_N(q, p, t)dqdp$ is the probability (over the ensemble) that particle 1 is in the volume $d\mathbf{q}_1 d\mathbf{q}_1$ about the point $\mathbf{q}_1\mathbf{p}_1$ and particle 2 is in the volume $d\mathbf{q}_2 d\mathbf{p}_2$ about the point $\mathbf{q}_2\mathbf{p}_2$ and so on, at time t. f_N is normalized so that $\int f_N(q, p, t)dqdp = 1$ where the integration is over the entire phase space accessible to the system.

Now we direct our attention to a subsystem comprised of $s < N$ particles. The probability of finding this subsystem in the phase volume $d1, d2, \ldots, ds$ about the state $(1, 2, \ldots, s)$ is $f_s(1, 2, \ldots, s)d1d3 \ldots ds$ with $f_s(1, 2, \ldots, s) = \int f_N(1, \ldots, N)d(s + 1) \ldots dN$. It is convenient to introduce the s-tuple (or reduced s-point) distribution function, $F_s(1, \ldots, s)$, defined in the 6-dimensional single particle phase space $\mathbf{x}, \mathbf{p}$. The quantity $F_s(1, \ldots, s)d1 \ldots ds$ represents the probable number of s-tuples of particles such that one of the particles is in the phase volume $d1$ about the point 1, another in $d2$ about 2, and so on, at a given time. Both the joint probability s-particle distribution function f_s and the s-tuple distribution function F_s refer to the state of the same specific group of s particles. However F_s is independent of the specific manner in which the particles occupy this state. Thus the function $f_2(1, 2)$ refers to the configuration where particle 1 occupies the state 1, and particle 2 the state 2. The function $F_2(1, 2)$ is a measure of the number of

[1] We only analyse the classical case, where particles have a well-defined position and velocity. In general, this is a good approximation if the average de Broglie length of a particle is much smaller than the mean interparticle separation. For a fluid with characteristic particle number density n and temperature T, this requires $\hbar n^{1/3}/(2mk_{\rm b}T)^{1/2} \ll 1$ with m the mass of a single fluid unit.

pairs of particles that occupy these states. The relation between F_s and f_s is [2] $F_s = f_s N!/(N-s)!$ (e.g. Huang 1987; Liboff 1998) and the F_s normalization is given by $\int F_s d1 \dots ds = N!/(N-s)!$. Macroscopic fluid-dynamic variables are most easily defined in terms of the F_s distributions. For instance, if the particles have the same mass m and their velocities are simply given by $\mathbf{v} = \mathbf{p}/m$, one has

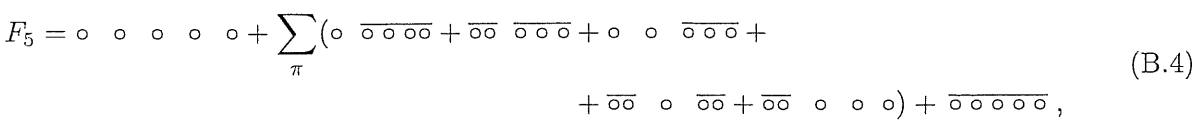
$$\text{number density of particles} = n(\mathbf{q},t) = \int F_1(\mathbf{q},\mathbf{p},t) d^3p \,, \tag{B.1}$$

$$\text{mean macroscopic (fluid) velocity} = \mathbf{u}(\mathbf{q},t) = \frac{\int \mathbf{p}\, F_1(\mathbf{q},\mathbf{p},t) d^3p}{m\, n(\mathbf{q},t)} \,. \tag{B.2}$$

Let us define the (reduced) 2-particle correlation function, $C_2(1,2)$, according to the relation $F_2(1,2) = F_1(1)F_1(2) + C_2(1,2)$. Note that, when $C_2 = 0$, particle 1 and 2 are statistically independent. The relation between F_2 and C_2 can be generalized by transforming the whole set of distribution functions $(F_1, F_2, \dots, F_N)$ into $(F_1, C_2, \dots, C_N)$ through the relations

$$\begin{aligned} F_2(1,2) &= F_1(1)F_1(2) + C_2(1,2) \\ F_3(1,2,3) &= F_1(1)F_1(2)F_1(3) + \sum_{\pi\{1,2,3\}} F_1(1)C_2(2,3) + C_3(1,2,3) \\ &\dots \end{aligned} \tag{B.3}$$

where $\pi\{1,2,3\}$ denotes permutations of 1,2,3. It is possible to represent these equations with the aid of diagrams. The transformation equation for F_5, say, has the representation:



(B.4)

where, for example, $\overline{\circ\circ}$ represents C_2, $\overline{\circ\circ\circ}$ stands for C_3, $\circ\ \circ$ means F_1F_1 and π is for $\pi\{1,2,3,4,5\}$. By integrating the reduced correlation functions over momenta, one obtains the corresponding spatial correlation functions $\xi_n(\mathbf{q}_1, \dots, \mathbf{q}_n, t)$:

$$\int C_2(1,2)\, d^3p_1\, d^3p_2 = n(\mathbf{q}_1,t) n(\mathbf{q}_2,t) \xi_2(\mathbf{q}_1,\mathbf{q}_2,t) \tag{B.5}$$

$$\int C_3(1,2,3)\, d^3p_1\, d^3p_2\, d^3p_3 = n(\mathbf{q}_1,t) n(\mathbf{q}_2,t) n(\mathbf{q}_3,t) \xi_3(\mathbf{q}_1,\mathbf{q}_2,\mathbf{q}_3,t) \tag{B.6}$$

$$\dots \tag{B.7}$$

In cosmology, particle dynamics is most conveniently described by adopting comoving coordinates and their canonical conjugate momenta as shown in §2.1.2. In this case, $\int F_1 d^3p = n a^3 =$ const with n the proper mean number density of particles and a the expansion factor of the FRW metric. Keeping also into account the underlying homogeneity and isotropy, one has (e.g. Peebles 1980)

$$\int C_2(1,2)\, d^3p_1\, d^3p_2 = n^2 a^6 \xi_2(x_{12},t) \tag{B.8}$$

$$\int C_3(1,2,3)\, d^3p_1\, d^3p_2\, d^3p_3 = n^3 a^9 \xi_3(x_1, x_2.x_3, t) \tag{B.9}$$

$$\dots \,. \tag{B.10}$$

[2] It is assumed here that f_s is a symmetric function of its arguments. If this is not the case, indicating by $\tilde{f}_s$ the given asymmetric distribution, it is possible to construct a symmetric function by defining $f_s(1,\dots,s) = \sum_\pi \tilde{f}_s(1,\dots,s)/s!$, where the sum is over permutations of the numbers $\{1,\dots,s\}$.

B.2 Exact dynamics: BBGKY equations

The Liouville theorem states that f_N is constant along system trajectories

$$\frac{df_N}{dt} = \frac{\partial f}{\partial t} + \sum_{\ell=1}^{N} \left(\frac{\partial \mathcal{H}}{\partial \mathbf{p}_\ell} \cdot \frac{\partial f_N}{\partial \mathbf{q}_\ell} - \frac{\partial \mathcal{H}}{\partial \mathbf{q}_\ell} \cdot \frac{\partial f_N}{\partial \mathbf{q}_\ell} \right) . \tag{B.11}$$

A differential equation describing the evolution of f_s is obtained from eq. (B.11) by integrating over variables associated with $N - s$ particles. If the Hamiltonian of the system is $\mathcal{H} = \sum_{i=1}^{N} (p_i^2/2ma^2) + \sum \sum_{i<j}^{N} \phi_{ij}$ with ϕ_{ij} the potential energy associated to the interaction between the i-th and j-th particles, one obtains

$$\frac{df_s}{dt} = -\sum_{i=1}^{s} \frac{\partial}{\partial \mathbf{p}_i} \cdot \int d(s+1) \ldots dN \sum_{j=s+1}^{N} \mathbf{G}_{ij} f_N \tag{B.12}$$

with $\mathbf{G}_{ij} = -\partial \phi_{ij} / \partial \mathbf{x}_j$. It is evident that $df_s/dt \neq 0$ because of interactions of the s-subgroup of particles with the remaining units in the system. To further reduce eq. (B.12), it is necessary to assume that the constituent particles in the system are identical and that f_s is symmetric under interchange of particle numbers. In this case, one obtains the so-called BBGKY equations (from Bogoliubov, Born, Green, Kaiser, Yvonne), i.e., in terms of the distribution functions F_s,

$$\frac{dF_s}{dt} + \sum_{i=1}^{s} \frac{\partial}{\partial \mathbf{p}_i} \int \mathbf{G}_{i,s+1} \, F_{s+1} \, d(s+1) = 0 . \tag{B.13}$$

This coupled sequence of N differential equations shows a typical hierarchical behaviour: the kinetic equation for F_s contains a term in F_{s+1}. Formally, they describe the exact dynamics of a collection of identical, binary interacting particles. In practice, however, some approximations are required to close the chain of equations.

We turn now to the explicit form of the first 2 equations of the hierarchy describing gravitational instability in a FRW background model (Fall & Severne 1976; Davis & Peebles 1977; Peebles 1980). We adopt comoving coordinates and the single-particle Hamiltonian (2.8). In this case,

$$\mathbf{G}_{ij} = \frac{Gm^2}{a} \frac{\mathbf{x}_{ij}}{x_{ij}^3} , \tag{B.14}$$

with $\mathbf{x}_{ij} = \mathbf{x}_i - \mathbf{x}_j$, and one obtains

$$\frac{\partial F_1}{\partial t} + \frac{Gm^2}{a} \frac{\partial}{\partial \mathbf{p}_1} \cdot \int C_2(1,2) \frac{\mathbf{x}_{21}}{x_{21}^3} \, d^3x_2 \, d^3p_2 = 0 , \tag{B.15}$$

$$\begin{aligned}
&\frac{\partial C_2(1,2)}{\partial t} + \frac{1}{ma^2} \left[\mathbf{p}_1 \cdot \frac{\partial}{\partial \mathbf{x}_1} + \mathbf{p}_2 \cdot \frac{\partial}{\partial \mathbf{x}_2} \right] C_2(1,2) + \\
&\quad + \frac{Gm^2}{a} \frac{\mathbf{x}_{21}}{x_{21}^3} \left(\frac{\partial}{\partial \mathbf{p}_1} - \frac{\partial}{\partial \mathbf{p}_2} \right) [F_1(1) F_1(2) + C_2(1,2)] + \\
&\quad + \frac{Gm^2}{a} \left\{ \frac{\partial F_1(1)}{\partial \mathbf{p}_1} \cdot \int \frac{\mathbf{x}_{31}}{|\mathbf{x}_{31}|^3} C_2(2,3) \, d^3x_3 \, d^3p_3 + \right. \\
&\quad + \frac{\partial}{\partial \mathbf{p}_1} \cdot \int \frac{\mathbf{x}_{31}}{|\mathbf{x}_{31}|^3} C_3(1,2,3) \, d^3x_3 \, d^3p_3 + \frac{\partial F_1(2)}{\partial \mathbf{p}_2} \cdot \int \frac{\mathbf{x}_{32}}{|\mathbf{x}_{32}|^3} C_2(1,3) \, d^3x_3 \, d^3p_3 + \\
&\quad \left. + \frac{\partial}{\partial \mathbf{p}_2} \cdot \int \frac{\mathbf{x}_{32}}{|\mathbf{x}_{32}|^3} C_3(1,2,3) \, d^3x_3 \, d^3p_3 \right\} = 0 .
\end{aligned} \tag{B.16}$$

The momentum moments of the BBGKY equations are often used to discuss clustering growth. By integrating the BBGKY-s equation over momenta (zero-order moments) one obtains an equation that expresses the conservation of the s-tuples of particles. In particular, eq. (B.15) gives the mass conservation equation $\partial(na^3)/\partial t = 0$, while (B.16) gives the pair conservation equation:

$$\frac{\partial \xi_2}{\partial t} + \frac{1}{x^2 a}\frac{\partial}{\partial x}\left[x^2(1+\xi_2)v\right] = 0\,. \tag{B.17}$$

Here, $v\mathbf{x}/x$ is the ensemble average of the peculiar velocity difference of a particle pair at separation $\mathbf{x}$, i.e.

$$v\frac{\mathbf{x}}{|\mathbf{x}|} = \frac{\int F_2(1,2)(\mathbf{p}_2 - \mathbf{p}_1)\, d^3p_1\, d^3p_2}{m\, a \int F_2(1,2)\, d^3p_1\, d^3p_2}\,, \tag{B.18}$$

with $\mathbf{x} = \mathbf{x}_{12}$. First and second order moments of the BBGKY-s equation give, respectively, relations between momenta and energies of a set of s particles. For example, the first momentum moment of (B.16) shows how the relative velocity dispersion and gravity determine the evolution of ξ_2 (see e.g. Peebles 1980, eq. 72.1).

B.3 Approximated kinetic equations

The BBGKY hierarchy describes the exact dynamics of a collections of binary interacting particles. Its intrinsic complexity, however, prevents its extensive use in practical applications. For this reason, approximated but mathematically simpler descriptions of the dynamics of a large set of particles are of great interest. In this section we will quickly review the construction and the validity regimes of two alternative kinetic equations: the Vlasov and the Boltzmann equations. The concept of perfect fluid will be finally introduced. We will consider only the simplest Newtonian case, however, the generalization of the method to the case of an expanding background is straightforward. The resulting dynamics is adopted in the main text to describe the phenomenon of gravitational instability.

B.3.1 Collisionless systems

Let us consider a fluid with characteristic particle number density n and suppose that we are able to characterize the system through the following parameters:

- the particle mass m;
- the mean thermal speed c_s linked to the temperature T through $mc_\mathrm{s}^2 \simeq k_\mathrm{b}T$ with k_b the Boltzmann constant;
- the typical potential energy of the fluid particles, ϕ, and the characteristic scale length r_0 of binary interactions between particles.

If $\phi/k_bT = \mathcal{O}(\epsilon) \ll 1$, correlations between particles in the fluid are expected to be small. Mathematically, this condition may be described by attaching an ϵ-factor with each coupling bar in the diagram equation (B.4). Considering long-range interactions ($n_0 r_0^3 = \mathcal{O}(1/\epsilon) \gg 1$) and keeping

only lowest-order terms in ϵ, it can be shown that the whole BBGKY hierarchy reduces to the Vlasov equation (e.g. Liboff 1998)

$$\left(\frac{\partial}{\partial t}+\frac{\mathbf{p}}{m}\cdot\frac{\partial}{\partial \mathbf{q}}+\mathbf{G}\cdot\frac{\partial}{\partial \mathbf{p}}\right)F_1(\mathbf{q},\mathbf{p},t)=0\,, \tag{B.19}$$

$$\mathbf{G}(\mathbf{q},t)=\int n(\mathbf{q}',t)\,\mathbf{G}(\mathbf{q},\mathbf{q}')\,d^3q'\,. \tag{B.20}$$

Note that

1. $F_1(\mathbf{x},\mathbf{v},t)$ develops under a force field that is the instantaneous average of all two-particles forces in the fluid. This "effective field" has a macroscopic character as it represents a collective effect of a large number of particles. Since this cumulative force changes slowly, we may expect particle trajectories to likewise change smoothly;
2. the Vlasov equation is non-linear since the force field is a functional of the distribution F_1;
3. the evolution described by the Vlasov equation cannot lead to statistical equilibrium and cannot increase the entropy of the system.

The dynamical state of a collection of N bodies interacting only through gravity can be accurately described through the Vlasov equation only for a limited lapse of time. In fact individual two-body encounters will progressively modify particle trajectories with respect to the effective field action. Denoting the characteristic scale-length of the system by L and the typical particle velocity by v, one obtains that the relaxation time is (Binney & Tremaine 1987)

$$t_{\rm rel}\simeq\frac{1}{8}\frac{N}{\ln N}\frac{L}{v}\,. \tag{B.21}$$

B.3.2 Collisional systems

Let us write the first BBGKY equation in a slightly different way,

$$\frac{\partial F_1}{\partial t}+\frac{\mathbf{p}}{m}\frac{\partial F_1}{\partial \mathbf{q}}=-\hat{O}F_2(1,2)\,, \tag{B.22}$$

where we have set all external and long-range forces to zero for simplicity, and $\hat{O}$ denotes the integro-differential operator introduced in the previous sections. A kinetic equation (i.e. a closed equation of motion for the system) results if we are able to effect the transformation $\hat{O}F_2(1,2)=\hat{J}[F_1]$, where the functional operator $\hat{J}$ maps functions onto functions. In this case the system is completely described by $F_1(\mathbf{x},\mathbf{p},t)$. The lowest-order closure of the hierarchy of kinetic equations is obtained by assuming that particles are uncorrelated, $F_2(1,2)=F_1(1)F_1(2)$, when eq. (B.15) reduces to Vlasov equation. In the simplest case, the effect of short-range interactions and of the graininess of the medium can be mathematically described by considering a non-vanishing two-particle correlation function $C_2(1,2)$. The corresponding term generated by $\hat{O}$ is called collision operator. The Boltzmann equation is obtained by choosing a particular analytical representation for the collision term (see e.g. Huang 1987).

B.4 Fluid dynamics

For general applications, the microscopic description given by F_1 is too much detailed, and a lot of information contained in the kinetic equations has to be discarded to deal with macroscopic

quantities. This reduction is achieved by taking velocity moments of the kinetic equations. The zeroth-order moment describes mass (or particle) conservation. If processes able to create, destroy or combinate particles are not present, the zeroth-order moment gives the continuity equation

$$\frac{\partial \rho}{\partial t} + \nabla (\rho \mathbf{u}) = 0 \,, \tag{B.23}$$

which describes mass (or particle) conservation. For interactions that do not depend on particle velocity, the first-order moment gives the momentum conservation equation

$$\rho \left[\frac{\partial}{\partial t} + (\mathbf{u} \cdot \nabla) \right] \mathbf{u} = -\nabla \mathbf{P} + n\mathbf{F} \,. \tag{B.24}$$

Here,

- $\mathbf{P} \equiv \rho \langle \mathbf{ww} \rangle = m \int \mathbf{ww}\, F\, d^3p$ denotes the stress (or pressure) tensor with $\mathbf{w} = (\mathbf{p}/m) - \mathbf{u}$ the random (thermal) velocity.
- $\mathbf{F}$ is an externally supported force field that permeates the fluid (or the mean field when long-range interactions between fluid elements are present).

The second-order moment gives the energy conservation equation

$$\rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) E = -\nabla \cdot \mathbf{Q} - \mathbf{P} \dddot{:} \mathbf{Y} \tag{B.25}$$

Here, $\mathbf{Q} = \rho \langle w^2 \mathbf{w}/2 \rangle$ is the heat flow vector, $E_K = \rho \langle w^2/2 \rangle$ is the kinetic energy density, and the tensor $\mathbf{Y}$ has components $(m/2)(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$. The term $\mathbf{P} \dddot{:} \mathbf{Y}$ represents the scalar $P_{ij} Y_{ij}$, where summation over repeated indices is understood.

B.4.1 The perfect fluid

We will briefly review here the Chapman-Enskog method to solve the Boltzmann transport equation by recursive approximations. This technique has been developed to obtain particular solutions that depends on time only implicitly through the local macroscopic variables $\rho, \mathbf{v}$ and T. It can be shown (see e.g. Huang 1987) that any solution of the Boltzmann equation will show only this implicit time dependence after a time of the order of a collision time ($\sim 10^{-11}$ s in ordinary gases). The Chapman-Enskog expansion deals with the large collision frequency regime and relies on the following assumptions:

1. we can introduce a bookeeping parameter $1/\epsilon$ ($\epsilon \ll 1$) to quantify the magnitude of the collision operator;
2. $F = F^{(0)} + \epsilon F^{(1)} + \epsilon^2 F^{(2)} + \ldots$;
3. n, $\mathbf{v}$ and T are all $\mathcal{O}(1)$ quantities in ϵ and stem from $F^{(0)}$;
4. F depends on time only implicitly through the hydrodinamic variables n, $\mathbf{v}$ and T.

The first order solution is (e.g. Huang 1987; Liboff 1998)

$$\begin{aligned} F^{(0)}(\mathbf{q},t) &= \frac{n(\mathbf{q},t)}{[2\pi R T(\mathbf{q},t)]^{3/2}} \exp\left(-\frac{w^2}{2RT(\mathbf{q},t)} \right) \qquad R \equiv k_\mathrm{b}/m \\ \mathbf{Q}^{(0)}(\mathbf{q},t) &= 0 \\ \mathbf{P}^{(0)}(\mathbf{q},t) &= p(\mathbf{q},t)\, \mathbf{I} = n(\mathbf{q},t)\; k_\mathrm{b} T(\mathbf{q},t)\, \mathbf{I} \end{aligned} \tag{B.26}$$

i.e. the one-particle distribution is a local Maxwellian, there is no heat-flow and the pressure tensor is diagonal and isotropic ($\mathbf{I}$ here denotes the (3×3) identity matrix). Substituting these expressions into the conservation equations gives the their first-order approximation

$$\begin{aligned} \frac{\partial \rho}{\partial t} + \nabla(\rho \mathbf{u}) &= 0 \\ \rho \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \mathbf{u} + \nabla p &= n\mathbf{F} \\ \left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \right) \frac{p}{\rho^{5/3}} &= 0 \end{aligned} \tag{B.27}$$

where the last equation express the adiabatic quality of the flow (the adiabatic index $5/3$ comes out because our fluid is composed by particles with only 3 degrees of freedom). A fluid is called perfect if its evolution is described by the set of equations (B.4.1) above. The value of the adiabatic index can obviously change from case to case.

The first-order solution in the Chapman-Enskog expansion is a good approximation to a real fluid if its local density, temperature and velocity are constant over characteristic scale-lengths much larger than the mean free-path of its constituent particles. Each successive iterate in the ϵ-expansion yields a more detailed set of hydrodynamic equations, better suited to describe spatial fluctuations in the fluid on small scales. The second-order solution gives the Navier-Stokes equations, the third the Burnett equations.

Bibliography

Abramowitz M., Stegun I.A., 1968, *Handbook of Mathematical Functions*, Dover, New York

Adelberger K.L., Steidel C.C., Giavalisco M., Dickinson M., Pettini M., Kellogg M., 1998, ApJ, 505, 18

Adler R.J., 1981, *The Geometry of Random Fields*, John Wiley & Sons, New York

Arnold V.I., 1986, *Catastrophe Theory*, Springer, Berlin

Arnold V.I., Shandarin S.F., Zel'dovich Ya.B., 1982, Geophys. Astrophys. Fluid Dyn., 20, 111

Arnouts S., Cristiani S., Moscardini L., Matarrese S., Lucchin F., Fontana A., Giallongo E., 1999, submitted to MNRAS, astro-ph/9902290

Babul A., 1991, MNRAS, 248, 177

Babul A., Lee M. H., 1991, MNRAS, 250, 407

Babul A., White S.D.M., 1991, MNRAS, 253, 31P

Bagla J.S., 1998a, MNRAS, 297, 251

Bagla J.S., 1998b, MNRAS, 299, 417

Bagla J.S., Padmanabhan T., 1994, MNRAS, 266, 227

Bagla J.S., Padmanabhan T., 1997, MNRAS, 286, 1023

Bahcall J., Casertano S., 1985, ApJ, 293, L7

Baleisis A., Lahav O., Loan A.J., Wall J.V., 1998, MNRAS, 297, 545

Bardeen J.M, 1980, Phys. Rev. D, 22, 1882

Bardeen J.M., Bond J.R., Kaiser N., Szalay A.S., 1986, ApJ, 304, 15

Baugh C.M., Efstathiou G., 1993, MNRAS, 265, 145

Baugh C.M., Efstathiou G., 1994, MNRAS, 270, 183

Baugh C.M., Gaztañaga E., 1996, MNRAS, 280, L37 (BG)

Baugh C.M., Gaztañaga E., Efstathiou G., 1995, MNRAS, 274, 1049

Baugh C.M. et al. , 1998, astro-ph/9811222

Benson A.J., Cole S., Frenk C.S., Baugh C.M., Lacey C.G., 1999, submitted to MNRAS, astro-ph/9903343

Bernardeau F., 1992, ApJ, 392, 1

Bernardeau F., 1994a, ApJ, 433, 1

Bernardeau F., 1994b, A&A, 291, 697

Bernardeau F., Kofman L., 1995, ApJ, 443,479

Bernstein G., Fischer P., 1999, astro-ph/9903274

Bertschinger E., 1985, ApJS, 58, 39

Bertschinger E., Jain B., 1994, ApJ, 431, 486

Bertschinger E., Hamilton A.J.S., 1994, ApJ, 435, 1

Bharadwaj S., 1996a, ApJ, 460, 28

Bharadwaj S., 1996b, ApJ, 472, 1

Bharadwaj S., Gupta A.K., Seshadri T.R., 1999, submitted to A&A, astro-ph/9903252

Biggs A.D., Browne I.W.A., Helbig P., Koopmans L.V.E., Wilkinson P.N., Perley R.A., 1999, MNRAS, 304, 349

Binney J.J., Tremaine S., 1987, *Galactic Dynamics*, Princeton University Press, Princeton, NJ

Birkinshaw M., 1999, Phys. Rept., 310, 97

Blandford R.D., Narayan R., 1992, ARA&A, 30, 311

Blandford R.D., Saust A. B., Brainerd T. G., Villumsen J. V., 1991, MNRAS, 251, 600

Blanton M., Cen R., Ostriker J.P., Strauss M.A., 1998, submitted to ApJ, astro-ph/9807029

Blanton M., Cen R., Ostriker J.P., Strauss M.A., Tegmark M., 1999, submitted to ApJ, astro-ph/9903165

Blumenthal G.R., Faber S.M., Primack J.R., Rees M.J., 1984, Nature, 311, 517

Bond J.R., 1988, in *Large-Scale Motions in the Universe*, eds. V. Rubin and G. Coyne, Princeton University Press, Princeton, New Jersey

Bond J.R., Couchman H.M.P., 1988, in *Proc. Second Canadian Conference on General Relativity & Relativistic Astrophysics*, eds. Coley A., Dyer C.C. & Tupper B.O.J., World Scientific, Singapore

Bond J.R., Efstathiou G., 1987, MNRAS, 226, 655

Bond J.R., Cole, S., Efstathiou, G., Kaiser, N., 1991, ApJ, 379, 440 (BCEK)

Bond J.R., Myers, S.T., 1996, ApJS, 103, 1

Bond J.R., Efstathiou G., Tegmark M., 1997, MNRAS, 291, 33

Bond J.R., Pogosyan D., Souradeep T., 1998, Class. Quant. Grav., 15, 2573

Bouchet F., Juszkiewicz R., Colombi S., Pellat R., 1992, ApJ, 394, L5

Bower R.G., 1991, MNRAS, 248, 332

Bower R.G., Coles P., Frenk C.S., White S.D.M., 1993, ApJ, 405, 403

Brainerd T.G., Scherrer R.J., Villumsen J.V., 1993, ApJ, 418, 570

Buchalter A., Kamionkowsky M., 1999, ApJ in press, astro-ph/9903462

Buchert T., 1989, A&A, 223, 9

Buchert T., 1993, A&A, 267, L51

Bunn E.F., White M., 1997, ApJ, 480, 6

Burles S., Tytler D., 1998a, ApJ, 499, 699

Burles S., Tytler D., 1998b, ApJ, 507, 732

Burles S., Nollett K.M., Truran J.N., Turner M.S., 1999, submitted to Phys. Rev. Lett., astro-ph/9901157

Burnstein D., Heiles C., 1984, ApJS, 54, 33

Bond J.R., Pogosyan D., Souradeep T., 1998, Class. Quant. Grav., 1998, 15, 2671

Cappellaro E., Turatto M., Tsvetkov D.Yu., Bartunov O.S., Pollas C., Evans R., Hamuy M., 1997, A&A, 322, 431

Carignan C, Freeman K., 1985, ApJ, 294, 494

Carlberg R.G., Couchman H.M.P., 1989, ApJ, 340, 47

Carlberg R. G., Yee H. K. C., Ellingson E., 1997, ApJ, 478, 462

Carlberg R.G., Cowie L.L., Songaila A., Hu E.M., 1997, ApJ, 484, 538

Carlberg R.G. et al. , 1998, astro-ph/9805131

Carroll S.M., Press W.H., Turner E.L., 1992, ARA&A, 30, 499

Cartwright D.E., Longuet-Higgins M.S., 1956, Proc. R. Soc. A., 237, 212

Catelan P., 1995, MNRAS, 276, 115

Catelan P., Moscardini L., 1994, ApJ, 426, 14

Catelan P., Theuns T., 1996a, 282, 436

Catelan P., Theuns T., 1996b, 282, 455

Catelan P., Coles P., Matarrese S., Moscardini L., 1994, 268, 966

Catelan P., Lucchin F., Matarrese S., Moscardini L., 1995, MNRAS, 276, 39

Catelan P., Lucchin F., Matarrese S., Porciani C., 1997, MNRAS, 297, 692 (CLMP)

Catelan P., Matarrese S., Porciani C., 1998, ApJ, 502, L1

Cavaliere A., Danese L., De Zotti G., 1979, A&A, 75, 322

Cavaliere A., Colafrancesco S., Menci N., 1993, ApJ, 415, 50

Cen R. & Ostriker J.P., 1992, ApJ, 399, L113

Cen R. & Ostriker J.P., 1993, ApJ, 417, 415

Cen R. & Ostriker J.P., 1999, ApJ in press, astro-ph/9811248

Centrella J., Melott A., 1983, Nature, 305, 196

Chaboyer B., 1998, to appear in Phys. Rep., astro-ph/9808200

Chaboyer B., Demarque P., Kernan P.J., Krauss L.M., 1998, ApJ, 494, 96

Chandrasekhar S., 1943, Rev. Mod. Phys., 15, 1

Chiba M., Yoshii Y., 1999, ApJ, 510, 42

Colafrancesco S., Lucchin F., Matarrese S., 1989, ApJ, 345, 3

Cole S., 1991, ApJ, 367, 45

Cole S., Kaiser N., 1988, MNRAS, 233, 637

Cole S., Kaiser N., 1989, MNRAS, 237, 1127

Cole S., Lacey C.G., 1996, MNRAS, 281, 716

Cole S., Aragón-Salamanca A., Frenk C.S., Navarro J.F., Zepf S.E., 1994, MNRAS, 271, 781

Cole S., Weinberg D.H., Frenk C.S., Ratra B., 1997, MNRAS, 289, 37

Coles P., 1990, MNRAS, 243, 171

Coles P., 1993, MNRAS, 262, 1065

Coles P, Jones B., 1991, MNRAS, 248, 1

Coles P., Davies A., 1993, MNRAS, 264, 261

Coles P., Lucchin F., 1995, *Cosmology: The Origin and Evolution of Cosmic Structure*, John Wiley & Sons, Chichester, New York

Coles P., Melott A.L., Shandarin S.F., 1993, MNRAS, 260, 765

Coles P., Lucchin F., Matarrese S., Moscardini L., 1998, MNRAS, 300, 183

Colin P., Klypin A., Kravtsov A., Khokhlov A., submitted to ApJ, astro-ph/9809202

Colombi S., Bouchet F.R., Hernquist L., 1996, ApJ, 465, 14

Connolly A.J., Szalay A.S., Brunner R.J., 1998, ApJ, 499, L125

Cornish N.J., Spergel D.N., Starkman G.D, 1998a, Phys. Rev. D., 57, 5982

Cornish N.J., Spergel D.N., Starkman G.D, 1998a, Class. Quant. Grav., 15, 2657

Couchman H.M.P., 1991, ApJ, 368, L23

Couchman H.M.P., Thomas P.A., Pearce F.R., 1995, ApJ, 452, 797

Courbin F., Magain P., Keeton C.R., Kochanek C.S., Vanderriest C., Jaunsen A.O., Hjorth J., 1997, A&A, 324, L1

Cowie L.L., Songaila A., Hu E.M., Cohen J.G., 1996, AJ, 112, 839

Cress C.M., Kamionkowski M., 1998, MNRAS, 297, 486

Dahlén T., Fransson C., 1998, in *The Next Generation Space Telescope: Science Drivers and Technological Challenges*, 34th Liege International Astrophysics Colloquium, ESA Publications

Davis M., Peebles P.J.E., 1977, ApJS, 34, 425

Davis M., Efstathiou G., Frenk C.S., White S.D.M., 1985, ApJ, 292, 371

Dekel A., 1986, ComAp, 11, 235

Dekel A., 1994, ARA&A, 32, 371

Dekel A., Silk J., 1986, ApJ, 303, 39

Dekel A., Rees M.J., 1987, Nature, 326, 455

Dekel A., Rees M.J., 1994, ApJ, 422, 1

Dekel A., Lahav O., 1998, astro-ph/9806193

Doroshkevich A.G., 1967, Astrofizika, 3, 175 (Astrophysics, 3, 84)

Doroshkevich A.G., 1970, Astrofizika, 6, 581 (Astrophysics, 6, 320)

Dressler A., 1980, ApJ, 236, 351

Dunlop J., Peacock J., Spinrad H., Dey A., Jimenez R., Stern D., Windhorst R., 1996, Nature, 381, 581

Dwek E. et al. , 1998, ApJ, 508, 106

Dyer C.C., Roeder R. C., 1973, ApJ, 180, L31

Efstathiou G., Frenk C.S., White S.D.M., Davis M., 1988, MNRAS, 235, 715

Ehlers J., Schneider P., 1986, A&A, 168, 57

Eisenstein D.J., 1997, submitted to ApJ, astro-ph/9709054

Eisenstein D.J., Hu W., 1998,ApJ, 496, 605

Eisenstein D.J., Hu W., Silk J., Szalay A.S., 1998, ApJ, 494, L1

Eke V.R., Cole S., Frenk C.S., 1996, MNRAS, 282, 263

Ellis R.S., Colless M., Broadhurst T., Heyl J., Glazebrook K., 1996, MNRAS, 280, 235

Evans R., van den Bergh S., McClure R.D., 1989, ApJ, 345, 752

Fabian A.C., Thomas P.A., White R.E. III, Fall S.M., 1986, MNRAS, 221, 1049

Falco E.E., Gorenstein M.V., Shapiro I.I., 1991, ApJ, 372, 364

Falco E.E., Kochanek C.S., Muñoz J.A., 1998, ApJ, 494, 47

Fall S.M., Severne G., 1976, MNRAS, 174, 241

Fassnacht C.D. et al. , 1999, in preparation

Filippenko A. V., 1997, ARA&A, 35, 309

Fillmore J.A., Goldreich P., 1984, ApJ, 281, 1

Fisher K.B., Sharf C.A., Lahav O., 1994, MNRAS, 266, 219

Fisher K.B., Nusser A., 1996, MNRAS, 279, L1

Fixsen D.J., Cheng E.S.M Gales J.M., Mather J.C., Shafer R.A., Wright E.L., 1996, ApJ, 473, 576

Forman W., Jones C., Tucker W., 1985, ApJ, 293, 102

Frieman, J.A., 1996, Comments Astrophys., 18, 323

Frieman J.A., Gaztañaga E., 1994, ApJ, 425, 392

Fry J.N., 1984, ApJ, 279, 499

Fry J.N., 1986, ApJ, 308, L71

Fry J.N., 1994, Phys. Rev. Lett., 73, 215

Fry J.N., Melott A., Shandarin S., 1992, ApJ, 393, 431

Fry J.N., Gaztañaga E., 1993, ApJ, 413, 447

Fry J.N., Melott A.L., Shandarin S.F., 1994, ApJ, 412, 504

Fry J.N., Melott A.L., Shandarin S.F., 1995, MNRAS, 274, 745

Fukugita M., Turner E. L., 1991, MNRAS, 253, 99

Fukugita M., Futamase T., Kasai M., Turner, E.L. 1992, ApJ, 393, 3

Fukugita M., Hogan C.J., Peebles P.J.E., 1998, ApJ, 503, 518

Fukushige T., Makino J. 1994, ApJ, 436, L111

Garnavich, P.M. et al. , 1998, ApJ, 493, L53

Gawiser E., Silk J., 1998, Science, 280, 1405

Gelb J.M., Bertschinger E., 1994, ApJ, 436, 467

Ghigna S., Moore B., Governato F., Lake G., Quinn T., Stadel J., MNRAS, 300, 146

Giavalisco M., Steidel C.C., Adelberger K.L., Dickinson M.E., Pettini M., Kellogg M., 1998, ApJ, 503, 543

Gilliland R. L., Phillips, M.M., 1998, IAUC 6810

Giovanelli R., 1998, in *Wide Field Surveys in Cosmology*, eds. Colombi & Y. Mellier, p. 109, Editions Frontières, Gif-sur-Yvette, France

Giovanelli R., Haynes M.P., Da Costa L.N., Freudling W., Salzer J.J., Wegner G. 1997, ApJ, 477, L1

Gnedenko B.V., 1989, *The theory of probability and the elements of statistics*, Chelsea, New York, NY

Goroff M.H., Grinstein B., Rey S.J., Wise M.B., 1986, ApJ, 311, 6

Gott J.R., 1975, ApJ, 201, 296

Gouda N., Nagashima M., 1997, MNRAS, 287, 515

Gould A., 1994, ApJ, 421, L71

Governato F., Baugh C.M., Frenk C.S., Cole S., Lacey C.G., Quinn T., Stadel J., 1998, Nature, 392, 359

Grainge K., Jones M.E., Pooley G., Saunders R., Edge A., Kneissl R., 1999, submitted to MNRAS, astro-ph/9904165

Greiner A., Strittmatter W., Honerkamp J., 1988, J. Stat. Phys, 51, 95

Grinstein B., Wise M.B., 1987, ApJ, 320,448

Grogin N.A., Narayan R., 1996a, ApJ, 464, 92

Grogin N.A., Narayan R., 1996b, ApJ, 473, 570

Gross M.A.K., Somerville R.S., Primack J.R., Holtzman J., Klypin A., 1998, MNRAS, 301, 81

Grossman S.A., Narayan R., 1988, ApJ, 324, L37

Gunn J.E., 1977, ApJ, 218, 592

Gurbatov S.N., Saichev A.I., Shandarin S.F., 1989, MNRAS, 236, 385

Hamilton A.J.S., 1992, ApJ, 385, L5

Hamilton A.J.S., 1997, MNRAS, 289, 285

Hamilton A.J.S., Kumar P., Lu E., Mathews A., 1991, ApJ, 374, L1

Hamuy M., Phillips M.M., Suntzeff N.B., Schommer R.A., Maza J., Aviles R., 1996, AJ, 112, 2391

Hamuy M., Phillips M.M., Wells L.A., Maza J. 1993, PASP, 105, 787

H anggi P, Talkner P., 1980, J. Stat. Phys., 22, 65

Heath D.J., 1977, MNRAS, 179, 351

Heath D.J., 1982, ApJ, 259, 9

Heavens A.F., Matarrese S., Verde L., 1998, MNRAS, 301, 797

Hindmarsh M, Kibble T.B.W., 1995, rep. Prog. Phys., 58, 477

Hoffmann Y., Shaham J., 1985, ApJ, 297, 16

Holz D.E., 1998, ApJ, 506, L1

Holz D.E., Wald R., 1998, submitted to Phys. Rev. D, astro-ph/9708036

Hu W., White M., 1997, ApJ, 479, 568

Hu W., Eisenstein D.J., 1999, submitted to Phys. Rev. D, astro-ph/9809368

Huang, K., 1987, *Statistical mechanics*, J. Wiley, New York, NY

Hui L., Bertschinger E., 1996, ApJ, 471, 1

Impey C.D., et al. , 1998, ApJ, 509, 551

Jain B., 1996, MNRAS, 287, 687

Jain B., Bertschinger E., 1994, 431, 495

Jain B., Mo H., White S.D.M., 1995, MNRAS, 276, L25 (JMW)

Jaroszyński M., 1992, MNRAS, 255, 655

Jenkins A. et al. , 1998, ApJ, 499, 20

Jensen L.G., Szalay A., 1986, ApJ, 305, L5

Jimenez R., 1998, to appear in IOP proc. series. "DARK98, Heidelberg, MPI", astro-ph/9810311

Jing Y.P., 1998, ApJL, 503, 9

Jing Y.P., 1998, ApJL, 515, L45

Jungmann G. et al. , 1996, Phys. Rev. D., 54, 1332

Juszkiewicz R., Bouchet F.R., Colombi S., 1993, ApJ, 412, L9

Kaiser N., 1984, ApJ, 284, L9

Kaiser N., 1987, MNRAS, 227, 1

Kaiser N., 1992, ApJ, 388, 272

Kaiser N., Davis M., 1985, ApJ, 297, 365

Kamionkowski M., Buchalter A., 1998, ApJ in press, astro-ph/9807211

Kantowski R., Vaughan. T., Branch D., 1995, ApJ, 447, 35

Kashlinsky A., 1987, ApJ, 317, 19

Kashlinsky A., 1991, ApJ, 376, L5

Kasai M., Futamase T., Takahara F., 1990, Phys. Lett. A, 147, 97

Katz N., Hernquist L., Weinberg D.H., 1992, ApJ, 399, L109

Katz N., Hernquist L., Weinberg D.H., 1998, submitted to ApJ, astro-ph/9806257

Kauffmann G., White S.D.M., 1993, MNRAS, 261, 921

Kauffmann G., White S.D.M., Guiderdoni B., 1993, MNRAS, 264, 201

Kauffmann G., Nusser A., Steinmetz M., 1997, MNRAS, 286, 795

Kauffmann G., Colberg J.M., Diaferio A., White S.D.M., 1999a, MNRAS, 303, 188

Kauffmann G., Colberg J.M., Diaferio A., White S.D.M., 1999b, submitted to MNRAS, astro-ph/9809168

Keeton C.R., Kochanek C.S., 1997, ApJ, 487, 42

Keeton C.R., Kochanek C.S., Falco E.E., 1998, ApJ, 509, 561

Kendall M., Stuart A., 1977, *The Advanced Theory of Statistics*, Vol. 1, 4th edn., Charles Griffin Company Limited, London

Kennicutt R.C. Jr., Freedman W.L., Mould J.R., 1995, AJ, 110, 1476

Kim A., Goobar A., Perlmutter S., 1996, PASP, 108, 190

Kirshner R.P. et al. , 1993, ApJ, 415, 589

Knuth D.E., 1981, *The Art of Computer Programming. Vol. 2: Seminumerical Algorithms*, Addison-Wesley, Reading

Kochanek C.S., 1993, ApJ, 419, 12

Kochanek C.S., 1994, ApJ, 436, 56

Kochanek C.S., 1995a, ApJ, 445, 559

Kochanek C.S., 1995b, ApJ, 453, 545

Kochanek C.S., 1996, ApJ, 466, 638

Kochanek C. S., Blandford R. D., Lawrence C. R., Narayan R., 1989, MNRAS, 238, 43

Kofman L., 1991, in *Primordial Nucleosynthesis and Evolution of the Early Universe*, eds. Sato K., Kluwer Academic Press, Dordrecht, the Netherlands

Kofman L., Pogosyan D., Shandarin S.F., Melott A.L., 1992, ApJ, 393, 437

Kofman L., Bertschinger E., Gelb J.M., Nusser A., Dekel A., 1994, ApJ, 420, 44

Kolatt T.S., Bartelmann M., 1998, MNRAS, 296, 763

Kolb E.W., Turner M.S., 1990, *The Early Universe*, Addison-Wesley, Reading, MA

Koopman L.V.E., Fassnacht C.D., 1999, in preparation

Kovner I., Paczynski B., 1988, ApJ, 335, L9

Krauss L.M., White M., 1992, Phys. Rev Lett., 69, 869

Krauss L., Turner M.S., 1995, Gen. Rel. Grav., 27, 1137

Kundič T. et al. , 1997a, ApJ, 482, 75

Kundič T., Cohen J.G., Blandford R.D., Lubin L.M., 1997b, AJ, 114, 507

Lacey C., Cole S., 1993, MNRAS, 262, 627

Lacey C., Cole S., 1994, MNRAS, 271, 676

Lachièze-Rey M., Luminet J.P., 1995, Phys. Rept., 254, 135

Lahav O., Rees M.J., Lilje P.B., Primack J.R., 1991, MNRAS, 251, 128

Lahav O., Piran T., Treyer M., 1997, MNRAS, 284, 499

Layzer D., 1956, AJ, 61, 383

Lee J., Shandarin S.F., 1998, ApJ, 500, 14

Le Fèvre O., Hudon D., Lilly S.J., Crampton D., Hammer F., Tresse L., 1996, ApJ, 461, 534

Leibundgut B. et al. , 1996, ApJ, 466, L21

Lemson G., Kauffmann G., 1999, MNRAS, 302, 111

Lewis A.D., Ellingson E., Morris S.L., Carlberg R.G., 1999, to appear in ApJ, astro-ph/9901062

Liboff R.L., 1998, *Kinetic theory: classical, quantum and relativistic descriptions*, J. Wiley, New York, NY

Liddle A.R., 1999, to appear in *Proceedings of ICTP summer school in high-energy physics, 1998*, astro-ph/9901124

Liddle A.R., Lyth D.H., 1993, Phys. Rep., 231, 1

Lidman C., Courbin F., Meylan G., Broadhurst T., Frye B., Welch W.J.W., 1999, ApJL in press, astro-ph/9902317

Lilly S.J., Tresse L., Hammer F., Crampton D., Le Fèvre O., 1995, ApJ, 455, 108

Lin C.C., Mestel L., Shu F.H., 1965, apJ, 142, 1431

Linde A.D., 1983, Phys. Lett. B., 129, 177

Linde A.D., 1990, *Particle Physics and Cosmology*, Harwood Academic, Chur, Switzerland

Linder E.V., 1998, ApJ, 497, 28

Linder E. V., Schneider P., Wagoner R. V., 1988, ApJ, 324, 786

Lineweaver C.H., 1998, ApJ, 505, L69

Lineweaver C.H., Tenorio L., Smoot G.F., Keegstra P, Banday A.J., Lubin P., 1996, ApJ, 470, 38

Lokas E.L., Juszkiewicz R., Weinberg D.H., Bouchet F.R., 1995, MNRAS, 274, 730

Longuet-Higgins M.S., 1957, Phil. Trans. R. Soc. Lond. A., 249, 321

Loveday J., Peterson B.A., Efstathiou G., Maddox S.J., 1992, 390, L338 CHECK altrimenti 400 L43 LEPM

Loveday J., Maddox S.J., Efstathiou G., Peterson B.A., 1995, ApJ, 442, 457

Loveday J., Efstathiou G., Maddox S.J., Peterson B.A., 1996, ApJ, 468, 1

Lovell J.E.J. et al. , 1998, ApJ, 508, L51

Lubin L.M., Cen R., Bahcall N.A., Ostriker J.P., 1996, ApJ, 460, 10

Lucchin F., Matarrese S., Mollerach S., 1992, ApJ, 401, L49

Lucchin F., Matarrese S., Melott A.L., Moscardini L., 1994, ApJ, 422, 430

Lynden-Bell D., 1964, ApJ, 139, 1195

Lynds R., Petrosian V., 1989, ApJ, 336, 1

Madau P., Della Valle M., Panagia N., 1998, MNRAS, 297, L17 (MDP)

Madau P., Pozzetti L., Dickinson M.E., 1998, ApJ, 498, 106

Maddox, S.J., Efstathiou G., Sutherland W.J., Loveday J., 1990, MNRAS, 242, 43P

Madore et al. , 1999, ApJ in press, astro-ph/9812157

Magliocchetti M., Maddox S.J., 1999, MNRAS in press, astro-ph/9811320

Magliocchetti M., Bagla J.S., Maddox S.J., Lahav O., 1999, submitted to MNRAS, astro-ph/9902260

Makino N., Sasaki M., Suto Y., 1992, Phys. Rev. D, 46, 585

Mann R.G., Heavens A.F., Peacock J.A., 1993, MNRAS, 263, 798 (MHP)

Mann R.G., Peacock J.A., Heavens A.F., 1998, MNRAS, 293, 209

Manrique A., Salvador–Solé E., 1995, ApJ, 453, 6

Manrique A., Salvador–Solé E., 1996, ApJ, 467, 504

Maoz D., Rix, H.W., 1993, ApJ, 416, 425

Marri S., Ferrara A., 1998, ApJ, 509, 43

Martel H., 1991, ApJ, 377, 7

Marzke R.O., da Costa L.N., Pellegrini P.S., Willmer C.N.A., Geller M.J., 1998, ApJ, 503, 617

Matarrese S., Lucchin F., Bonometto S.A., 1986, ApJ, 310, L21

Matarrese S., Lucchin F., Moscardini L., Saez D., 1992, MNRAS, 259, 437

Matarrese S., Coles P., Lucchin F., Moscardini L., 1997, MNRAS, 286, 115 (MCLM)

Matarrese S., Verde L., Heavens A., 1997, MNRAS, 290, 651

Matsushita K., Makashima K., Ikebe Y., Rokutanda E., Yamasaki N., Ohashi T., 1998, ApJ, 499, L13

Meiksin A., Madau P., 1993, ApJ, 412, 34

Meiksin A., White M., Peacock J.A., 1999, MNRAS in press, astro-ph/9812214

Melott A.L., Pellman T.F., Shandarin S.F., 1994, MNRAS, 269, 628

Metcalf R.B., 1998, astro-ph/9803319

Miralda-Escudé J., Rees, M.J., 1997, ApJ, 478, L57

Misner C.W., Thorne K., Wheeler J.A., 1973, *Gravitation*, Freeman, San Francisco

Mo H.J., Jing Y.P., Börner G., 1992, ApJ, 392, 452

Mo H.J., Fukugita M., 1996, ApJ, 467, L9

Mo H.J., White S.D.M., 1996, MNRAS, 282, 347 (MW)

Mo H.J., Jing Y.P., White S.D.M., 1996, MNRAS, 282, 1096

Mo H.J., Jing Y.P., White S.D.M., 1997, MNRAS, 284, 189

Mo H.J., Mao S., White S.D.M., 1999, MNRAS, 304, 175

Mo H.J., Sheth R.K., Tormen G., 1999, in preparation

Mohr J., Mathiensen B., Evrard G., 1999, ApJ in press, astro-ph/9901281

Monaco P., 1995, ApJ, 447, 23

Monaco P., 1998, Fund Cosm Phys, 19, 157

More J.G., Heavens A.F., Peacock J.A., 1986, MNRAS, 220, 189

Moscardini L., Coles P., Lucchin F., Matarrese S., 1998, MNRAS, 299, 95

Moutarde F., Alimi J.-M., Bouchet F.R., Pellat R., Ramani A., 1991, ApJ, 382, 377

Munshi D., Starobinsky A.A., 1994, ApJ, 428, 433

Myers S.T., Baker J.E., Readhead A.C.S., Lietch E.M., Herbig T., 1997, ApJ, 485, 1

Narayan R., White, S.D.M. 1988, MNRAS, 231, 97P

Navarro J., Frenk C.S., White S.D.M., 1995, MNRAS, 275, 720

Navarro J., Frenk C.S., White S.D.M., 1996, ApJ, 462, 563

Navarro J., Frenk C.S., White S.D.M., 1997, ApJ, 490, 493

Nityananda R., Padmanabhan T., 1994, MNRAS, 271, 976

Nusser A. & Davis M., 1994, ApJ, 421, L1

Nusser A., Dekel A., Bertschinger E., Blumenthal G.R., 1991, ApJ, 379, 6

Ostriker J.P., 1983, ARA&A, 31, 689

Ostriker J.P., Steinhardt P.J., 1995, Nature, 377, 600

Oswalt T.D., Smith J.A., Wood M.A., Hintzen P., 1996, Nature, 382, 692

Padmanabhan T., 1993, *Structure Formation in the Universe*, Cambridge University Press, Cambridge

Padmanabhan T., 1996, MNRAS, 278, 29P

Padmanabhan T., Subramanian K., 1993, ApJ, 410, 482

Padmanabhan T., Cen R., Ostriker J.P., Summers F.J., 1996, ApJ, 466, 604

Pain R. et al. , 1997, ApJ, 473, 356

Patat R., Barbon R., Cappellaro E., Turatto M., 1994, A&A, 282, 731

Pauls J.L., Melott A.L., 1995, MNRAS, 274, 99

Peacock J.A., 1997, MNRAS, 284, 885

Peacock J.A., 1999, *Cosmological Physics*, Cambridge University Press, Cambridge, England

Peacock J.A., Heavens A.F., 1985, MNRAS, 217, 805

Peacock J.A., Heavens A.F., 1990, MNRAS, 243, 133

Peacock J.A., Dodds S.J., 1994, MNRAS, 267, 1020

Peacock J.A., Dodds S.J., 1996, MNRAS, 280, L19

Peebles P.J.E., 1965, ApJ, 142, 1317

Peebles P.J.E., 1972, Comm. Ap. and Space Phys., 4, 53

Peebles P.J.E., 1974, A&A, 32, 391

Peebles P.J.E., 1980, *The Large-scale Structure of the Universe*, Princeton University Press, Princeton, New Jersey

Peebles P.J.E., 1984, ApJ, 284, 439

Penzias A.A., Wilson R.W., 1965, ApJ, 142, 419

Perlmutter S. et al. , 1998, Nature, 391, 51

Perlmutter S. et al. (The Supernova Cosmology Project), 1999, ApJ in press, astro-ph/9812133

Persic M., Salucci P., 1992, MNRAS, 258, 14P

Persic M., Salucci P., 1995, ApJS, 99, 501

Politzer D., Wise M.B., 1984, ApJ, 285, L1

Porciani, C., 1997, MNRAS, 290, 639

Porciani, C., Ferrini, F., Lucchin, F. & Matarrese, S., 1996, MNRAS, 281, 311

Porciani, C., Matarrese, S., Lucchin, F. & Catelan, P., 1998, MNRAS, 298, 1097

Porciani, C., & Madau, P., 1998, submitted to ApJ, astro-ph/9810403

Porciani C., Catelan P., Lacey C.G., 1999, ApJ, 513, L99

Postman M., Geller M.J., 1984, ApJ, 281, 95

Press W.H., Schechter P., 1974, ApJ, 187, 425 (PS)

Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W., 1992, Numerical recipes, Cambridge University Press, Cambridge

Protogeros Z.A.M., Scherrer R.J., 1997, MNRAS, 284, 425

Protogeros Z.A.M., Melott A.L., Scherrer R.J., 1997, MNRAS, 290, 367

Rees M.J., 1985, MNRAS, 213, 75P

Rees M.J., 1988, in *Large Scale Structures of the Universe*, eds. Audouze J., Pellettan M.C., Szalay A.S., Kluwert, Dordrecht, the Netherlands

Refsdal S., 1964, MNRAS, 128, 307

Reisenegger A., Miralda-Escudé J., 1995, ApJ, 449, 476

Rice S.O., 1945, Bell Systems Tech. J., 24, 46

Riess A.G., Press W.H., Kirshner R.P., 1996, ApJ, 473, 88

Riess A.G. et al. , 1997, AJ, 114, 722

Riess A. G., Nugent P., Filippenko A. V., Kirshner R. P., Perlmutter S., 1998, ApJ, 504, 935

Riess A.G. et al. (The High-z Supernova Search Team), 1998, AJ, 116, 1009

Rindler W., 1977, *Essential Relativity*, Springer-Verlag, New York

Rodrigues D.D.C., Thomas P.A., 1996, MNRAS, 282, 631

Roukema B.F., Peterson B.A., Quinn P.J., Rocca-Volmerange B., 1997, MNRAS, 292, 835

Roukema B.F., Blanloeil V., 1998, Class. Quant. Grav., 15, 264

Ruamsuwan L., Fry J.N., 1992, ApJ, 396, 416

Ruiz-Lapuente P., Canal R., Burkert A., 1997, in *Thermonuclear Supernovae*, ed. Ruiz-Lapuente P., Canal R. Isern J., p. 205, Kluwert, Dordrecht, The Netherlands

Rubin V.C., Ford W.K., Thonnard N., 1980, ApJ, 238, 471

Saha A., Sandage A., Labhardt L., Tammann G.A., Macchetto F.D., Panagia N., 1997, ApJ, 486, 1

Sahni V., Coles P., 1995, Phys. Rep., 262, 1

Salvador-Solé E., Solanes J.M., Manrique A., 1997, preprint astro-ph/9712080

Satyaprakash B.S., Sahni V., Munshi D., Pogosyan D., Melott A.L., 1995, 275, 463

Saunders W., Rowan-Robinson M., Lawrence A., 1992, MNRAS, 258, 134

Saunders, R. et al. , 1999, submitted to MNRAS, astro-ph/9904168

Scannapieco E., Levin J., Silk J., 1999, MNRAS, 303, 797

Schechter P.L, et al. , 1997, ApJ, 475, L85

Scherrer R.J., Weinberg D.H., 1998, ApJ, 504, 607

Schmidt B.P. et al. (The High-z Supernova Search Team), 1998, ApJ, 507, 46

Schneider P., Wagoner R.V., 1987, ApJ, 314, 154

Schneider P., Ehlers J., Falco, E.E., 1992, *Gravitational Lenses*, Springer-Verlag, Berlin, Germany

Schneider P., Bartelmann M., 1995, MNRAS, 273, 475

Schramm D.N., Turner M.S., 1998, Rev. Modern Phys., 70, 303

Scoccimarro R., 1997, ApJ, 487, 1

Scoccimarro R., Frieman J.A., 1996a, ApJS, 105, 37

Scoccimarro R., Frieman J.A., 1996b, ApJ, 473, 620

Scoccimarro R., Colombi S., Fry J.N., Frieman J.A., Hivon E., Melott A., 1998, 496, 586

Shandarin S.F., 1980, Astrofizika, 16, 769 (Astrophysics, 16, 439)

Shandarin S.F., Zel'dovich Ya.B., 1984, Phys. Rev. Lett., 52, 1488

Shandarin S.F., Zel'dovich Ya.B., 1989, Rev. Mod. Phys., 61, 185

Shepherd C.W., Carlberg R.G., Yee H.K.C., Ellingson E., 1997, ApJ, 479, 82

Sheth R.K., Lemson G., 1998, submitted to MNRAS, astro-ph/9808138

Sheth R.K., Tormen G., 1999, submitted to MNRAS, astro-ph/9901122

Sigad Y., Eldar A., Dekel A., Strauss M., Yahil A., 1998, 1998, ApJ, 495, 516

Silk J., 1983, Nature, 301, 574

Silk J., 1985, ApJ, 297, 1

Smoot G.F., 1999, astro-ph/9902027

Sneden C., McWilliam A., Preston G.W., Cowan J.J., Burris D.J., Armosky B.J., 1996, ApJ, 467, 819

Somerville R.S., Kolatt T.S., 1997, preprint astro-ph/9711080

Somerville R.S., Primack J.R., 1998, submitted to MNRAS

Soucail G., Mellier Y., 1994, in *Gravitational Lenses in the Universe*, eds. Surdej J., Fraipont-Caro D., Gosset E., Refsdal S., Remy M., p. 205, Univ. Liege, Liege, Belgium

Steidel C.C., Hamilton D., 1993, AJ, 105, 2017

Steidel C.C., Giavalisco M., Pettini M., Dickinson M., Adelberger K.L., 1996, ApJ. 462, L17

Steidel C.C., Adelberger K.L., Dickinson M., Giavalisco M., Pettini M., Kellogg M., 1998, ApJ, 492, 428

Steinmetz M., 1996, MNRAS, 278, 1005 Stockman H.S., Stiavelli M., Im M., Mather J.C., 1998, in *Science with the Next Generation Space Telescope*, eds. Smith E. & Koratkar A., p. 24, ASP, San Francisco, California

Strauss M.A., Willick J.A., 1995, Phys. Rept., 261, 271

Strittmatter W., 1987, preprint, University of Freiburg, THEP 87/12

Suntzeff N.B. et al. 1999, AJ in press, astro-ph/9811205

Sunyaev R.A., Zel'dovich Ya.B., 1972, Comm. Astrophys. Sp. Phys., 4, 173

Suto Y., Sasaki M., 1991, Phys. Rev. Lett., 66, 264

Sylos-Labini F., Montuori M., Pietronero L., 1998, Phys. Rept., 293, 61

Szalay A.S., 1988, ApJ, 333, 21

Tammann G.A., Schröder, A. 1990, A&A, 236, 149

Tammann G.A., Löffler W., Schröder A., 1994, ApJS, 92, 487

Taylor A.N., 1993, in *Cosmic Velocity Fields*, Bouchet F., Lachiéze-Rey M., p. 585, Editions Frontières, Gif-sur-Yvette, France

Tegmark M., 1999, ApJ, 514, L69

Tegmark M., Peebles P.J.E., 1998, ApJ, 500, L79

Tegmark M., Bromley, 1999, ApJ letters in press, astro-ph/9809324

Tomita K., 1998, Prog. Theor. Phys., 100, 1

Tormen G., Bouchet F.R., White S.D.M., 1997, MNRAS, 286, 865

Tormen G., Diaferio A., Syer D., 1998, MNRAS, 299, 728

Tozzi P., Governato F., Cavaliere A., 1996, in ASP. Conf. Series n.117, *Dark and Visible Matter in Galaxies and Cosmological Implications*, eds. Persic M. & Salucci P., Astron. Soc. Pac., San Francisco

Tremaine S., Richstone D.O., Byun Y.-I., Dressler A., Faber S.M., Grillmail C., Kormendy J., Lauer T.R., 1994, AJ, 107, 634

Treyer M., Scharf C., Lahav O., Jahoda K., Boldt E., Piran T., 1998, ApJ, 509, 531

Tully B., 1999, to appear in *Post-Hipparcos Cosmic Candles*, eds. Caputo F. & Heck A., Kluwer Academic Publishers, Dordrecht, The Netherlands, astro-ph/9809394

Turner M.S., 1993, FERMILAB-Conf-92/313-A, astro-ph/9304012

Turner E.L., Ostriker J.P., Gott, J.R. III, 1984, ApJ, 284, 1

Tyson J. A., Fisher P., 1995, ApJ, 446, L55

Tyson J. A., Kochanski G. P., Dell' Antonio I. P., 1998, ApJ, 498, L107

Uzan J.P., Lehoucq R., Luminet J.P., 1999, submitted to A&A, astro-ph/9903155

Van den Bergh S., McClure R.D., 1994, ApJ, 425, 205

van Dokkum P.G., Franx M., Kelson D.D., Illingworth G.D., 1998, ApJ, 504, L17

van Kampen N.G., 1992, *Stochastic Processes in Physics and Chemistry*, North-Holland, Amsterdam, The Netherlands

Vanmarcke E.H., 1983, *Random Fields: Analysis and Synthesis*, MIT Press, cambridge, Massachusetts

Verde L., Heavens A.F., Matarrese S., Moscardini L., 1998, MNRAS, 300, 747

Vietri M., Ostriker J.P., 1983, ApJ, 269, 487

Villumsen J.V., 1996, MNRAS, 281, 369

Villumsen J.V., Freudling W., da Costa L.N., 1997, ApJ, 481, 578

Wald, R.M., 1984, *General Relativity*, Univ. of Chicago Press, Chicago

Wallington S., Narayan R., 1993, ApJ, 403, 517

Wambsganss J., Cen R., Ostriker J.P., 1997, ApJ, 475, L81

Wang Y., 1998, submitted to MNRAS, astro-ph/9806185

Wang L., Caldwell R.R., Ostriker J.P., Steinhardt P.J., 1999, astro-ph/9901388

Watanabe K., Tomita K., 1990, ApJ, 355, 1

Wax N. (ed.), 1954, *Selected Papers on Noise and Stochastic Processes*, Dover, New York

Webster A., 1976, MNRAS, 175, 61

Wechsler R.H., Gross M.A.K., Primack J.R., Blumenthal G.R., Dekel A., 1998, ApJ, 506, 19

Weeks J.R., 1998, Class. Quant. Grav., 15, 2599

Weinberg S., 1972, *Gravitation and Cosmology*, J. Wiley & Sons, New York

Wheeler J.C., Swartz D.A., 1993, Space Sci. Rev., 66, 425

White S.D.M., 1996, in *Cosmology and Large-scale Structure*, eds. Schaeffer R., Silk J., Spiro M. & Zinn-Justin J., p. 349, Elsevier, Dordrecht, the Netherlands

White S.D.M., Rees M., 1978, MNRAS, 183, 341

White S.D.M., Silk J.I., 1979, ApJ, 231, 1

White S.D.M., Davis M., Efstathiou G., Frenk C.S., 1987, Nature, 330, 451

White S.D.M., Zaritsky D., 1992, ApJ, 394, 1

White S.D.M., Efstathiou G., Frenk C.S., 1993, MNRAS, 262, 1023

White S.D.M., Navarro J.F., Evrard A.E., Frenk C.S., 1993, Nature, 366, 429

White S.D.M., Fabian A.C., 1995, MNRAS, 273, 72

Williams B.G., Heavens A.F., Peacock J.A., 1991, MNRAS, 252, 43P

Williams B.G., Heavens A.F., Peacock J.A., Shandarin S.F., 1991, MNRAS, 250, 458

Willick J.A., Strauss M.A., 1998, ApJ, 507, 64

Willmer C.N.A., da Costa L.N., Pellegrini P.S., 1998, AJ, 115, 869

Woltjer L., 1997, A&A, 328, L29

Wu K.K.S., Lahav O., Rees M., 1998, submitted to Nature, astro-ph/9804062

Yamamoto K., Suto Y., 1999, ApJ in press, astro-ph/9812486

Yano T., Nagashima M., Gouda N., 1996, ApJ, 466, 1

Yano T., Gouda N., 1998, ApJ, 495, 533

Zaldarriaga M., Spergel D., Seljak U., 1997, ApJ, 488, 1

Zaritsky D., Smith R., Frenk C., White S.D.M., 1997a, ApJ, 478, 39

Zaritsky D., Smith R., Frenk C., White S.D.M., 1997b, ApJ, 478, 53

Zehavi I., Riess A.G., Kirshner R.P., Dekel A., 1998, ApJ, 503, 483

Zel'dovich Ya.B., 1964, Soviet Astron. AJ, 8, 13

Zel'dovich Ya.B., 1965, Adv. Astron. Ap., 3, 241

Zel'dovich Ya.B., 1970, A&A, 5, 84

Zlatev I., Wang L., Steinhardt P.J., 1999, Phys. Rev. Lett., 82, 896