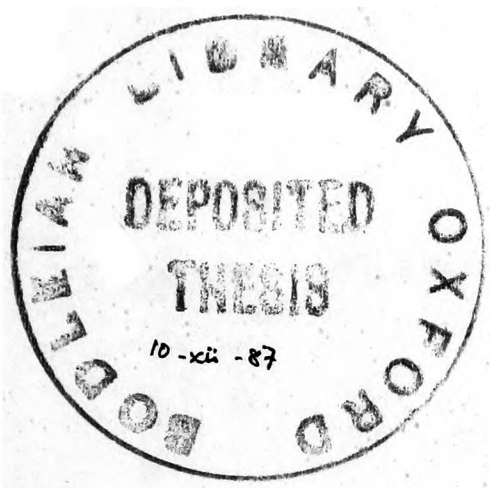
The Large-Scale Structure of the Universe: Some Theoretical Considerations



Colin Andrew M^{c}Gill

Department of Theoretical Physics
and
Magdalen College,
Oxford.

A thesis submitted to the University of Oxford in partial fulfilment of the requirements for the degree of Doctor of Philosophy.

September, 1987

To my parents,
with love.

Abstract

In this thesis, several theoretical concepts relating to the large-scale structure of the universe are presented. In particular, various aspects of the hierarchical scenario are investigated. The initial perturbation field and its early evolution are discussed in Chapter 3. Chapter 4 is concerned with two-point correlation functions for galaxies, clusters and super-clusters. In Chapter 5, some effects of using velocity as a distance measure are examined. In particular, it will be argued that caustics in redshift space are an almost inevitable feature of the hierarchical scenario. Chapter 6 concentrates on the possibity that quasar Ly-α absorption lines are redshift caustics.

ERRATA: PLEASE READ 'FURTHER' FOR 'FUTHER' THROUGHOUT.

Acknowledgements

First and foremost, I should like to thank my supervisor, Dr. James Binney for much help and guidance. Discussions with Dr. Andrew Hamilton are always enlightening. On a more personal level, numerous people in Theoretical Physics have helped make the last three years extremely pleasant. The secretarial and administrative staff managed to turn morning coffee into the highlight of the day. I would especially like to thank past and present members of the Astrophyics group. It wouldn't have been the same without you.

Outside of the department, my erstwhile housemates, Alan, Dave and Gavin, together with the Churchill boys, Mal and Andy, did their best to keep me from working. Those perennial visitors, Alastair, Chris and Phil, had similar success. Howard Street would have been very different without Jackie, Gina and Catherine. Meri—I *shall* return.

I am grateful to Princeton University, particularly the Graduate College and the Department of Astrophysical Sciences, for their hospitality during the Spring of 1986. Many of the ideas in Chapters 5 and 6 were formulated during that period. This work was supported by an SERC studentship.

Contents

Chapter 1. Introduction

Many people have studied or philosophised about the large-scale structure of the universe. Aristotle, for example, argued in favour of a finite, closed world [Koyré 1962]. His conception of 'large' extended no farther than Saturn and the classical constellations. Aristotle's vision of the universe, is shown in Figure 1.1 [taken from Apian 1593]. This anthropocentric world view was generally unchallenged until the advent of Copernicus and Galileo in the fifteenth century.

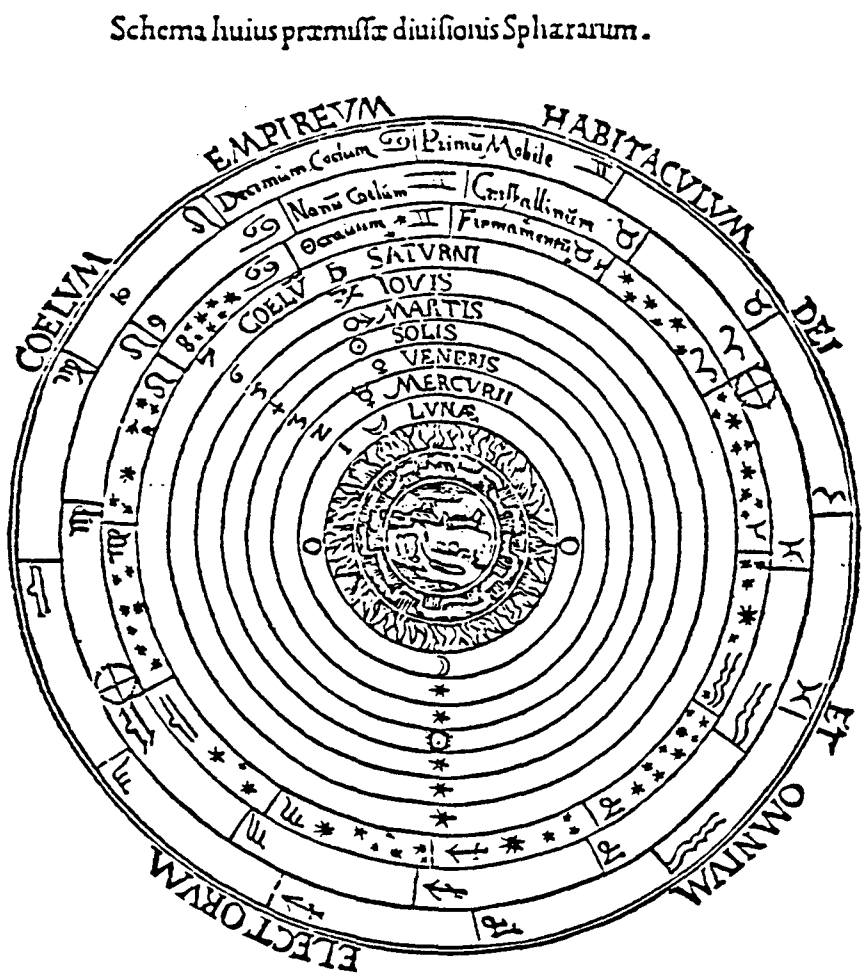


Figure 1.1. A typical pre-Copernican diagram of the universe. Taken from Apian [1593].

With the invention and use of increasingly powerful telescopes, Man's conception of scale has gradually changed. In the early years of this century 'the large-scale distribution of matter' was usually taken to mean the distribution of stars within the Milky Way. Today, there are indications of structure on scales

of up to $100h^{-1}$ Mpc[1] [Bahcall and Burgett 1986]. Although this distance is a significant fraction of the horizon scale, $cH_o \sim 3000h^{-1}$ Mpc, we would be naïve in the extreme *not* to expect structures of larger scale. It is likely that $100h^{-1}$ Mpc is nothing more than a statement of our observing prowess. Distance is subjective; our conception of 'large' says more about us than the universe we inhabit.

Bearing this in mind, this thesis explores aspects of one class of models attempting to account for the observed large-scale structure: the hierarchical scenario. The starting point is a presentation of the evidence for homogeneity on the very largest scales, followed by a discussion of departures from homogeneity on smaller scales.

1.1 Homogeneity

1.1.1 Evidence from Number Counts

Homogeneity has traditionally been examined by using number counts. If the distribution of a set of objects (stars, galaxies) is homogeneous, then the number of objects brighter than a limiting magnitude, m, would vary as [e.g., Rowan-Robinson 1981]

$$N(<m) \sim 10^{0.6m}. \tag{1.1.1}$$

This relationship was used by Kapteyn in his study of the distribution of stars in the Galaxy [Kapteyn 1922]. He found that the counts dropped off with magnitude far more rapidly than suggested by (1.1.1), indicating that the Galaxy has an edge. In addition, the counts varied with direction—the earth was not quite at the centre of the system. In point of fact, the picture Kapteyn drew was incorrect; the effects of extinction, caused by interstellar material in the galactic

[1] h is the present day Hubble constant measured in units of $100\,\mathrm{km\,s^{-1}\,Mpc^{-1}}$.

plane, dominated his counts. Shapley's assessment of the Galaxy, derived from globular cluster data, was much more accurate [Shapley 1918–1919]. Globular clusters occur out of the galactic plane where they are significantly less affected by galactic extinction.

Hubble was the first to apply (1.1.1) to the 'spiral nebulae'. He had just established that at least the brightest nebulae were stellar systems, similar to the Milky Way; it was natural to enquire after the homogeneity of their distribution. Preliminary indications [Hubble 1926] were encouraging enough for Hubble to begin a program of systematic galaxy counts. Ten years later he published his results; galaxy counts from several fields to 5 limiting magnitudes in the range $m \sim 18.5$–21. Hubble deduced that [1936]

$$N(<m) \sim 10^{0.4m}. \tag{1.1.2}$$

At the time, Hubble ascribed the deviation from the $10^{0.6m}$ law to a problem with relativistic models. With hindsight, 0.4 is about right once the K-correction is taken into account. This allows for the shift of the galaxy spectrum toward the red—photographic plates are less sensitive at these wavelengths.

In spite of this minor discrepancy, Hubble [1934] was led to conclude:

> "There are as yet no indications of a super-system of nebulae analogous to the system of stars."

Shapley, however saw things differently. If Hubble was taken with the homogeneity, Shapley [1933] was struck by the departures from it.

> "Irregularities [in the distribution of galaxies] are obviously too pronounced to be attributed to chance; they are rather a demonstration of evolutionary tendencies in a meta-galactic system."

He even suggested that deviations from homogeneity may be the cause of the observed discrepancy with equation (1.1.1). Other authors have found the evidence for homogeneity from galaxy counts similarly unconvincing [de Vaucouleurs 1960,1971; van den Bergh 1961].

1.1.2 Modern Constraints on Homogeneity

More recent indications of homogeneity are the low anisotropy limits on the diffuse X-ray background. A substantial fraction of X-rays come from active galaxies and galaxy clusters. The X-ray background is a measure of the distribution of these objects, integrated to the horizon. The present upper limit on the anisotropy implies that fluctuations in the projected density, $\delta f/f$, are less than 4% on an angular scale of 5° [Schwartz *et al.* 1976]. Another observation pointing to homogeneity is the almost random distribution of extra-galactic radio sources. The 5000 sources in the 4C catalogue are distributed in a manner virtually indistinguishable from random [Seldner and Peebles 1978].

Whilst these result are clear evidence for homogeneity, stronger constraints on uniformity come from observations of the micro-wave background. In the standard model [Weinberg 1972; Zel'dovich and Novikov 1983], the micro-wave background probes the epoch when matter and radiation were last strongly coupled. Inhomogeneities in the matter distribution should leave their imprint; anisotropies in the micro-wave background are a direct measure of inhomogeneity at high redshifts ($z \sim 1500$) [Section II, Inner Space/Outer Space 1984].

Current limits on the level of anisotropy are stringent. The micro-wave background is parameterised in terms of T, the temperature of a black body which will reproduce the observed intensity. Upper limits on $\delta T/T$ range from 10^{-3}–10^{-4}, depending on the angular scale [Wilkinson 1984]. Only a dipole aniso-

topy, indicating the sun's motion with respect to the micro-wave background, has been clearly observed, although Davies [1987] reports a positive detection of $\delta T/T \sim 5 \times 10^{-5}$ at an angular scale of 8°. Taken together with the X-ray data and the distribution of radio sources, there is a strong observational case for homogeneity.

1.1.3 Theoretical Arguments for Homogeneity

Until recently, homogeneity and isotropy had to be postulated as part of the initial conditions for general relativistic models of the universe (the homogeneity/isotropy problem). Indeed, the isotropy of the micro-wave background on scales larger than the horizon at the epoch of decoupling was somewhat of a puzzle (the horizon problem). With the advent of inflation [Guth 1981] the situation changed.

Guth's starting point was to acknowledge several problems with the standard model. Two of the most pressing have already been mentioned. Another was the flatness problem, why the average density of the universe appears to be so close to the critical density—in geometrical terms, why is the universe so flat? Guth realised that these problems would be resolved at a stroke if a period of exponential growth had occured [e.g., Linde 1984]. The entire universe would have come from one small causally connected region—goodbye to the horizon and homogeneity/isotropy problems! Magnifying a surface, no matter how crumpled will eventually remove creases from view. The same is true for the universe in the inflationary picture; after inflation the geometry must be close to flat—a solution to the flatness problem.

In a modified form, the inflationary scenario solves a host of problems related to coupling Grand Unified Theories to general relativity. Many of the same

theories naturally provide the conditions for inflation to occur; a phase transition from a state with a positive vacuum energy to a state with zero or negligible vacuum energy. The new inflationary model, at least in principle, is also capable of providing an *a priori* calculation of the power spectrum of initial perturbations (see §1.4) [Brandenberger 1985]. Inflation provides a firm theoretical backing for observational evidence of large-scale homogeneity.

1.2 Observed Inhomogeneities and their Characterisation

Whilst the case for large-scale homogeneity is strong, there are obvious inhomogeneities (see Shapely's comment above). Hubble himself realised that the distribution of galaxies was not random. He noted that the frequency distribution of the *log* of the number of galaxies in a field was Gaussian, clearly not a result of a random galaxy distribution [Hubble 1934].

Structures in the galaxy distribution are readily discernible. Regions on the sky with projected galaxy density greatly in excess of the average, i.e., galaxy clusters, are common [Abell 1958]. Redshift surveys, which yield distance information, reveal a rich variety of three-dimensional forms. There is at least one large region, volume $\sim 10^6 h^{-3}\,\mathrm{Mpc}^3$, of significantly below average galaxy density [Kirshner *et al.* 1983; 1987]. Long chains of galaxies have been identified [in the Perseus-Pisces supercluster; Giovanelli and Haynes 1986], shells of galaxies are clearly visible in redshift slices [de Lapparent *et al.* 1986]. There are indications that similar stuctures may also be present in the distribution of clusters. An absence of clusters in a volume $\sim 3 \times 10^6 h^{-3}\,\mathrm{Mpc}^3$ has been reported by Bahcall and Soneira [1982].

Inhomogeneities can also be made apparent through studies of recession velocities. Bulk motions in addition to universal expansion have been discovered

in a region $\sim 60h^{-1}$ Mpc in diameter around the galaxy. Both eliptical galaxies and clusters [Collins *et al.* 1986; Dressler *et al.* 1987] indicate that large-scale streaming velocities of around 500-1000 $\mathrm{km\,s^{-1}}$ occur in our neighbourhood. Coherent motion on such a large scale has caused difficulties for some standard scenarios (§1.5.3).

1.2.1 Attempts at Characterisation

If the inhomogeneities themselves are obvious, a quantitative method of characterising them is not. A popular measure of clustering has been the two-point correlation function and its variants. If the density at $\mathbf{r}$ is given by $\rho(\mathbf{r})$, then the dimensionless two-point correlation function, $\xi(\mathbf{r})$, is defined as:

$$\xi(\mathbf{r}) = \frac{\langle [\rho(\mathbf{r_1}) - \langle\rho\rangle][\rho(\mathbf{r_1}+\mathbf{r}) - \langle\rho\rangle]\rangle}{\langle\rho\rangle^2} \tag{1.2.1}$$

where $\langle \ldots \rangle$ represents a spatial average, in this case over $\mathbf{r_1}$. Isotropy implies that $\xi(\mathbf{r})$ is a function of $r = |\mathbf{r}|$ only. A similar statistic, the angular correlation function, $w(\theta)$, is useful if only the density projected on the sky is known. The relationship between $w(\theta)$ and $\xi(r)$ has been explored in some detail [Peebles 1980, §§49–54]. The lag, r_o, for which $\xi(r_o) = 1$, is called the *correlation length.*

Modern estimates of $w(\theta)$ and $\xi(r)$ for galaxies [Totsuji and Kihara 1969; Peebles 1980 and references therein] are:

$$w(\theta) \sim \theta^{1-\gamma} \qquad \xi(r) = \left(\frac{r}{r_o}\right)^{-\gamma}. \tag{1.2.2}$$

γ is found to be $\sim$ 1.8, $r_o \sim 5h^{-1}$ Mpc. Similar forms have been found for clusters [Bahcall and Soneira 1983] and super-clusters [Bahcall and Burgett 1986], although the correlation lengths are somewhat larger—$50h^{-1}$ Mpc and $100h^{-1}$ Mpc respectively.

Other statistics have also been used. Bogart and Wagoner [1973] used the nearest neighbour distance to establish the clustering of Abell clusters. Gott and Turner [1977] introduced a multiplicity function, an analogue of the luminosity function, to characterise the distribution of groups and clusters. Higher order correlation functions, defined in analogy to $\xi(r)$ have been used by Peebles and co-workers [Groth and Peebles 1975, Fry and Peebles 1978]. For example, the three-point correlation function, $\zeta(\mathbf{u}, \mathbf{v})$, is defined to be:

$$\zeta(\mathbf{u}, \mathbf{v}) = \frac{\langle [\rho(\mathbf{r}) - \langle\rho\rangle][\rho(\mathbf{r}+\mathbf{u}) - \langle\rho\rangle][\rho(\mathbf{r}+\mathbf{v}) - \langle\rho\rangle]\rangle}{\langle\rho\rangle^3} \tag{1.2.3}$$

The assumptions of homogeneity and isotropy imply that ζ is a function of only three variables, conveniently taken to be r_a, r_b and r_c, the sides of the triangle defined by $\mathbf{u}$ and $\mathbf{v}$. Isotropy additionally implies that $\zeta(r_a, r_b, r_c)$ is a symmetric function of these lengths. Observations indicate [Peebles 1980 §54] that the three-point correlation function for galaxies can be written as:

$$\zeta(r_a, r_b, r_c) = Q\left[\xi(r_a)\xi(r_b) + \xi(r_b)\xi(r_c) + \xi(r_c)\xi(r_a)\right]. \tag{1.2.4}$$

Estimates of Q vary from 0.8 to 1.3 [Peebles 1980 §58 and references therein].

Using the higher-order correlation functions is a systematic approach. A distribution is completely characterised by the full set of n-point correlation functions. White [1979] examined the expansion of various measures of clustering in terms of the hierarchy of correlation functions. He put particular emphasis on the small and intermediate-scale limits of the expansion (small is defined to be less than the mean inter-object distance, a few Mpc for bright galaxies). Using the observed forms of the two and three-point correlation functions (equations (1.2.2) and (1.2.4)), White found that many statistics contained little more information than the two-point function. His analysis led him to suggest an alternative statistic; the probability that a random volume, V, contains no galaxies, $P\{\Phi_o(V)\}$. White showed that $P\{\Phi_o(V)\}$ was strongly influenced

by the complete hierarchy of n-point correlation functions, a potentially useful complement to the two and three-point functions.

Recent methods for characterising the inhomogeneities include cluster analysis (percolation theory) [Zel'dovich *et al.* 1982] and a quantitative assement of the topology of the large scale structure [Hamilton, Gott and Weinberg 1986]. The correlation function function in redshift space (see §5.5) provides additional information [Bean *et al.* 1983; Davis and Peebles 1983].

There is *still* no satisfactory way to quantitatively characterise a distribution. A large part of the problem involves the human eye; it is notoriously good at finding patterns (perhaps too good—see Barrow and Bhavsar [1987]). On the basis of the statistics discussed above it may be impossible to distinguish between two distibutions which look quite different. The two-point correlation function in particular suffers from this defect (§4.5). Whilst there are inherent difficulties in comparing distributions by eye [Barrow and Bhavsar 1987], it remains most valuable in distinguishing between spatial distributions.

1.3 Theoretical Models of Large-Scale Structure

During the thirties, theorists rapidly accepted the evidence for homogeneity, largely as a result of theoretical prejudices; only the information on number counts was available to them at the time. Einstein, for example, wrote [1933]:

> "Hubble's research has further shown that the objects [galaxies] are distributed in a statistically uniform fashion, by which the schematic assumption of the theory of a uniform mean density receives experimental confirmation."

The homogeneous and isotropic world-model was thus established amongst theorists as a good zeroth-order approximation for the universe. The next step was to account for departures from homogeneity.

Homogeneous world-models have one characteristic in common—they are unstable; deviations from uniformity will grow. Lemaître suggested that initially small perturbations may evolve into the inhomogeneities observed today. He left the form of the initial pertubations an open question.

Although the idea was inherently plausible, it had a poor start. Lifshitz [1946] produced an analysis of linear perturbations in Friedman-Lemaître models and concluded that gravitational instability could not form stable systems. What he in fact showed was that a perturbation crossing the horizon whilst the universe was still radiation dominated would not grow. Novikov [1964] rescued the situation by pointing out that if a perturbation crossed the horizon after decoupling it *could* grow and evolve into a stable system. The form of the initial perturbation remained the only question.

Other theoretical models attempting to account for the large-scale structure include cosmic strings [Vilenkin 1985] and an early universe dominated by exploding objects [Carr *et al.* 1984; Ikeuchi and Ostriker 1986]. Little else will be said about other models in this thesis (although see §1.5).

Having chosen or guessed a form for the initial perturbations, the obvious question to ask is: how do they evolve? First-order perturbation theory predicts the rate of growth of the perturbations, but preserves their form (§2.1). Applications of this theory are obviously somewhat limited. A better analytic approach is to use the Zel'dovich approximation (§2.2) [1970]. Pertubations can be followed fairly reliably into the non-linear regime, but only until the first

structures have collapsed [Efstathiou 1987]. A general solution to the problem does not exist, leaving numerical calculations the only real alternative.

1.4 N-body Simulations

Numerical solutions to the problem have so far consisted of N-body simulations. The universe is represented as a set of points (N bodies) which generally interact only through gravity. The heart of the problem lies in accurately evaluating the force on a particle at one instant. Once this has been determined it is relatively straightforward to integrate the equations of motion [Efstathiou *et al.* 1985].

Only three approaches to this problem have yet been applied to cosmological problems. The most obvious way of calculating the force on a particle is to sum the force contributions from the remaining particles. Each contribution can be directly calculated from the inverse square law [Aarseth, Gott and Turner 1979; Aarseth 1984]. Forces determined in this way are highly accurate. This technique has the additional advantage of being able to evaluate the forces on one particle at a time. Forces on particles change at different rates; the more rapid the change, the more frequently a force should be evaluated. Each particle can have its own timestep in this method, making integration of the equations of motion efficient. The drawback to this approach is the amount of computer time required. The number of operations required to calculate the force on every particle varies as N^2. Present technology limits the use of this method to systems of less than 10^4 particles.

The second approach is to calculate the forces via a mesh and Poisson's equation. Masses are assigned to a discrete grid and Poisson's equation is solved using a Fast Fourier Transform [Klypin and Shandarin 1983; Centrella and Melott 1983]. The force on a particle is interpolated from the mesh (hence the name PM

for these codes—Particle-Mesh). The advantage of this approach is that forces are calculated efficiently; the number of operations needed to solve Poisson's equation scales as N log N. Its major drawback is the resolution of the scheme, limited by the separation of grid points. The code also has to evaluate the forces simultaneously. Present generation computers allow systems in excess of 10^7 particles to be studied by this method, although the limited force resolution substantially reduces its applicability. It is the most appropriate code to use if the initial conditions have a large coherence length [Efstathiou *et al.* 1985].

The third, and perhaps most promising approach, involves a combination of the first two techniques. Long-range forces are calculated by the efficient PM scheme, short-range forces using direct summation. In this way it is possible to calculate forces at a resolution smaller than the mesh size [Efstathiou and Eastwood 1981; Efstathiou *et al.* 1985; Hockney and Eastwood 1981]. Such schemes are denoted P^3M—Particle-Particle Particle-Mesh. Currently P^3M codes can be run with up to 10^5 particles. For most initial conditions this compromise between direct summation and PM codes represents the 'best buy'. The exception to this rule has already been mentioned.

1.4.1 New Numerical Methods

New approaches to this problem have recently been suggested; tree-codes [Appel 1985] and the orrery technique. As an example of a hierarchical tree-code, we shall consider dynamic gridding [Barnes and Hut 1986]. It is a form of direct summation in that the inverse square law is used to calculate the force between two particles. The force contribution from distant particles is relatively insensitive to their exact positions. Dynamic gridding exploits this idea; more distant particles can be grouped into volumes (gridding) and treated as though the total mass in a volume is situated at the centre of mass. Where dynamic

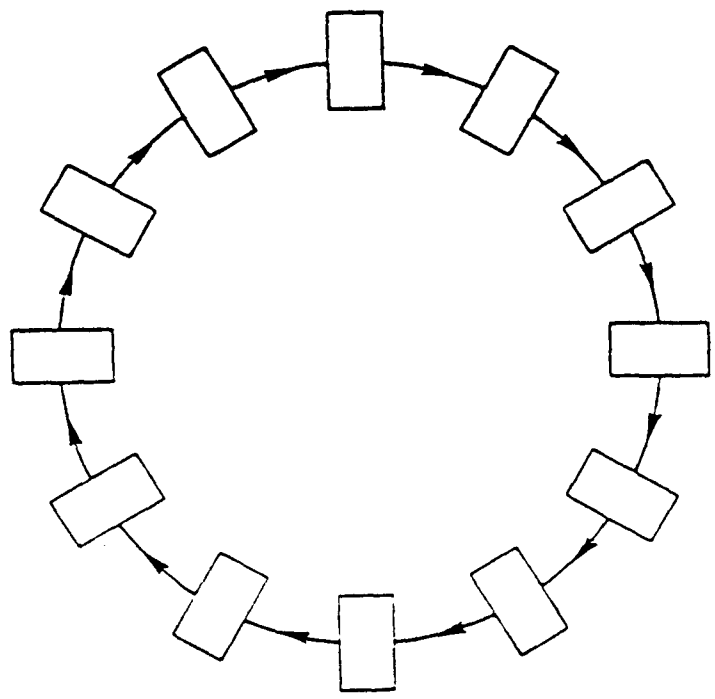


Figure 1.2. Illustrates the inter-connection of processors necessary to implement the 'orrery'. Information can only be passed in the direction of the arrows.

gridding differs from previous tree-codes is that a new grid is established at each time step. The grid responds to the particle distribution—hence *dynamic* gridding.

In the Barnes-Hut code, there are many levels of grouping. In general, the further from the selected particle, the larger the grouping volume allowed. The technique is efficient in that the number of operations required to find the force on a single particle varies as $\log N$. Forces can be evaluated for one particle at a time, with the advantages that entails (discussed above). The major disadvantage of this method is the complexity of the computer code necessary to implement it. It has yet to be applied in a cosmological setting.

The final numerical method to be discussed is the orrery technique. This method is appropriate for use on a multi-processor machine. Each particle in the simulation is assigned its own processor, which we can visualise as being connected in a ring—see Figure 1.2. The best way to understand this technique is to follow the cycle for one timestep.

At the start of the cycle each processor stores the position of its own particle and places this information in the ring. The next stage in the process is to move all the positions one step around the ring. Since it now knows another particle position, each processor can calculate the first contribution to the force acting on its particular particle. This stage is repeated until the positions have been stepped right around the ring. After each step, every processor has access to a new particle position; an additional force contribution can be calculated. At the end of the cycle each processor has calculated the total force acting on its particle and can update the velocity and position accordingly. The cycle is repeated as necessary.

This method is very efficient. The number of operations per timestep scales only as N. The force is calculated exactly and the coding is simple. Apart from having to calculate all the forces at once, the only drawback to this method is that N processors are needed. This approach was adopted by Sussman [1986] in designing a machine specifically to study the Solar System. It forms the basis of a current proposal to build a computer dedicated to N-body computations, but is a long way from being applied to cosmological problems.

1.4.2 Problems with the N-body approach

With numerical solutions, the degree to which the numerical model represents the physical/mathematical model is not always clear. This is particularly true for the N-body methods discussed above. In addition to any physical length scales in the problem, there are at least two scales which solely reflect the numerical technique. The first is the box scale, L_b, the size of the volume being represented in the model. Periodic boundary conditions are frequently imposed. The second is the mean inter-particle distance, L_i. There is generally one other length scale—the grid size, L_g, if a mesh is used or the 'smoothing length',

L_s, if a particle does not represent a point mass but rather an extended mass distribution. If L_l is the greatest of L_i,L_g,L_s, length scales outside the range $L_l < L < L_b$ cannot be followed by these techniques.

Because of restriction in computer resources, the range of length scales which can usefully be probed is fairly small. P^3M codes, for example, typically have 10^4–10^5 particles. The smallest scale which can be followed is thus a fraction $\left(10^5\right)^{-1/3} \approx 1/40^{\text{th}}$ of L_b. If scales near L_b or L_l are significantly affected by the inability to follow larger or smaller scales repectively, then the range of scales which can accurately be probed maybe fairly small [Bouchet *et al.* 1985].

Such numerical restrictions make it virtually impossible for an N-body technique to accurately model the evolution of certain perturbations. Once scales comparable to the box size collapse there is little point continuing the simulation. If the initial conditions are such that these scales collapse rapidly, subsequent evolution will be hard to follow.

Whilst these codes generate lots of data, in the form of many particle positions, it is not simple to compare them with observations. Results must be cast in a form which takes into account the way observations are made [White 1987]. The codes have yet to usefully incorporate physical processes other than gravitational interactions, although some steps have been taken in this direction [Carlberg 1987]. An approach combining P^3M and smooth particle hydrodynamics is presently being developed by Evrard [Private Communication].

It seems inevitable that N-body methods will play an increasing rôle in our understanding of the large-scale structure of the universe. However, it is important to develop insights into the nature of gravitational clustering; it is not enough to leave it all to the machine. Only with a deeper theoretical understanding will real progress occur.

1.5 Initial Conditions and other Parameters

The form of the initial perturbations has so far not been discussed. There are only two ways to specify initial conditions—guess, or attempt to calculate them from a knowledge of other physics. The former method is not as ludicrous as it may appear. By studying the evolution of a range of initial perturbations it should be possible to predict how any particular set of perturbations will evolve. The aim is to work out the form of the initial perturbations which evolved into the observed large-scale structure. This approach is adopted throughout the remainder of the thesis, particularly in Chapter 3. The latter method is discussed below.

What else can be altered in the models? The parameters which describe the overall evolution of the universe, H_o, Ω_o and Λ, can be varied—although all of these are observationally constrained in some way. One last parameter, related to biasing, is discussed next.

1.5.1 Biasing

Biasing was originally introduced by Kaiser [1984] to explain the discrepancy between the correlation length for galaxies and clusters (§1.2.1). Since clusters are correlated over a significantly larger scale than galaxies, Kaiser argued that both could not be fair tracers of the mass (but see §4.3). Kaiser showed that if only regions which initially had a density greater than some threshold value were included in the calculation, $\xi(r)$ would be enhanced. The higher the threshold, the greater the enhancement. If clusters only occur in the most overdense regions, the discrepancy in correlation lengths can be explained.

Since then, biasing has found many applications. It allows many theoretical calculations to be reconciled with the observed values. Application of the virial

theorem on a variety of scales [Huchra 1984] implies that $\Omega = 0.2$, incompatible with inflation which predicts $\Omega = 1$ (§1.1.3). Galaxy formation may be biased, in the sense that only reasonably overdense regions become galaxies. If this is the case, then there are many 'un-galaxies' which lack the required density to become fully-fledged L_* galaxies. $\Omega = 0.2$ would then only be a measure of luminous galaxies.

This idea has been taken over into N-body simulations. Adjusting the threshold density at which galaxies form enables simulations to establish a much better correspondence with observation.

1.5.2 Calculation of the form of the Initial Perturbations

Since N-body codes have yet to include any physics other than gravity, these methods are only able to model perturbation growth after decoupling. Fortunately, perturbations are likely to be small at this epoch (implied by the isotropy of the micro-wave background—§1.1.2), allowing first order perturbation theory to be used. Linear perturbation theory in a matter dominated universe is presented in §2.1. Under these conditions, every mode grows at the same rate. In general, this will not be the case.

Let us denote the Fourier components of the overdensity at a time t by $\widetilde{\delta}(\mathbf{k}, t)$ (see §2.1). Perturbation theory can be used to calculate the transfer function, $T(\mathbf{k};\, t_i, t_f)$, defined to be:

$$T(\mathbf{k};\, t_i, t_f) = \frac{\widetilde{\delta}(\mathbf{k}, t_f)}{\widetilde{\delta}(\mathbf{k}, t_i)}. \tag{1.5.1}$$

If a perturbation is known at a time, t_i, before decoupling, it can be followed through to a time, t_f, after decoupling through the use of equation (1.5.1). The form of the transfer function is determined by high-energy physics alone; for a

particular set of assumptions, a definite transfer function results [e.g., Bond and Efstathiou 1984] Once a suitable model has been adopted, these calculations are relatively accurate [Fall 1987].

It just remains to specify a form for the perturbation at t_i. A common speculation is that the overdensity is a Gaussian-random field—§2.3 [Adler 1981; Bardeen *et al.* 1986]. In this case, the perturbation is specified by its power spectrum (see §2.3),

$$P(\mathbf{k}, t_i) = \langle\langle|\widetilde{\delta}(\mathbf{k}, t_i)|^2\rangle\rangle. \tag{1.5.2}$$

$\langle\langle\ldots\rangle\rangle$ implies averaging over an ensemble of systems. As usual, isotropy implies that $P(\mathbf{k}, t_i)$ is a function of k only. The initial power spectrum is frequently taken to be of the form:

$$P(k, t_i) \sim k^n, \tag{1.5.3}$$

which may be appropriate since there are no conspicuous length scales in the problem. Various values of the spectral power index, n, have some theoretical or aesthetic appeal. If $n = +1$, the spectrum takes on the constant curvature form [Harrison 1970; Zel'dovich 1972]. $n = +1$ and $n = -3$ emerge from inflationary models; $n = -3$ from flicker noise in the density field, $n = +1$ from flicker noise in the corresponding gravitational potential field [Bardeen *et al.* 1986].

1.5.3 Standard Spectra

It would be impossible to conclude this section without some mention of currently popular spectra. They generally divide into two groups. The first comprises spectra with large coherence lengths so that the largest structures collapse first (pancake models). Smaller structures develop as the pancake gradually fragments. Spectra in this category are generated in hot dark matter models. Spectra in the second group cause smaller scales to collapse first. These scales

become gradually incorporated into larger structures (hierarchical clustering). Spectra in this group are predicted by warm/cold dark matter models. Perhaps the most succesful model to date falls into the later category: cold dark matter with biasing [Davis *et al.* 1985; Blumental *et al.* 1984].

Whilst these models have had their successes, all fail to explain some aspect of the observed large-scale structure. Even the cold dark matter model has difficulty matching some observations—in particular the large-scale streaming motions. Sometime in the future it will undoubtedly be possible to make an *a priori* prediction of the initial power spectrum. Until then, the methods adopted in Chapter 3 will have their uses.

1.6 Quasar Absorption Lines.

Most of the information on the large-scale distribution of matter is inferred from galaxy surveys. Whilst this is convenient, galaxies can only be observed easily to distances of a few hundred Mpc, equivalently, a redshift of ~ 0.1. Quasar absorption lines allow the distribution to probed up to a redshifts of ~ 4.

The first quasar was discovered by Schmidt [1963]. He was observing the apparently stellar object 3C 273, which exhibited emission lines that matched no known transitions. Schmidt realised that the problem disappeared if the object was put at a redshift of 0.158. After that, many quasi-stellar objects (QSOs) were discovered to be at large redshifts.

Many people noted that a redshift of 2 was large enough to move the H Lyman series into the optical band. Once a quasar was discovered at $z \gtrsim 2$, it would be possible to search for intervening hydrogen. Gunn and Peterson [1965] suggested looking for a discontinuity in the continuum level at the Ly-α emission line.

The discontinuity would be indicative of a continuous distribution of unionised hydrogen (H I). In the same issue of the *Astrophysical Journal*, Bahcall and Salpeter [1965] coined the phrase 'absorption trough' for features of this type. They also noted that if the hydrogen was clustered, sharp absorption lines would result. Making the assumption that the gas only occurs in galaxy clusters, they estimated that ~ 7 lines might be observed per unit redshift interval, Δz.

Observations soon established that there was *no* detectable Gunn-Peterson effect and that 80 or 90 lines per unit Δz was not uncommon. A small proportion could be assigned to redshift systems, but the majority of the lines were single, assumed to be Ly-α. Bahcall and Spitzer [1969] proposed that the redshift systems arose from absorption in galactic haloes. If each galaxy has a halo $\sim 100\,\mathrm{kpc}$ in extent, the predicted number agrees with that observed.

No such simple explanation for the single Ly-α lines was forthcoming. It was not even obvious whether the lines were caused by material intrinsic to the quasar (the ejection model) or by cosmologically distributed hydrogen (the intervening hypothesis). Only the latter made any prediction; the statistical properties of the lines should be *independent* of the quasar towards which they are observed. Sargent *et al.* [1980] tested this hypothesis in some detail. They concluded that the observations were in complete agreement with the intervening hypothesis.

The structures causing the Ly-α lines were expected to be kept at around $10^4\,\mathrm{K}$ by the background u.v. flux. Since an isothermal, gravitationally bound gas cloud is unstable [e.g., Rees 1986] the structures had to be confined in some way. Sargent *et al.* proposed that the Ly-α lines represented clouds of cool ($10^4\,\mathrm{K}$) gas embedded in a hotter and more tenuous intergalactic medium (IGM). If the IGM was totally ionised, it would not cause a Gunn-Peterson effect.

This model has been used to place constraints on the evolution of the IGM [Ostriker and Ikeuchi 1983; Ikeuchi and Ostriker 1986]. For the clouds to survive (not to evaporate into the IGM), the IGM must have a temperature $\lesssim 10^6$ K. This temperature is too low to account for the diffuse X-ray background by two orders of magnitude [Guilbert and Fabian 1986] *There is no reason to believe in an* IGM *in this form at all, except for our inability to create a viable mechanism to confine the clouds.* Only the exploding galaxy scenario can account for an IGM at 10^6 K; other theories must find a different explanation for the Ly-α lines.

One suggestion [Rees 1986], is that the lines are caused by absorption in 'mini-haloes' of dark matter. Regions which were not quite dense enough to form galaxies (see §1.5.1), have baryonic matter sitting in a potential well, generated by the dark matter. The u.v. flux prevents the hydrogen from falling further into the well. Although this model is attractive, it has very little predictive power—it would be next to impossible to model the evolution of these systems [Rees 1986].

Modern observations at higher resolution have confirmed the intervening hypothesis. A summary of the properties of the absorption lines is presented in §6.2.

1.6.1 Projection in Redshift Space

One convenient measure of the distance to an object is its recession velocity. Hubble's law converts the velocity to a distance. The problem with this approach is that there are two components to the velocity of any particular object—the velocity due to expansion and a peculiar velocity because of motion with respect to the co-moving frame (§2.1.1). This motion will affect the distance estimate.

Whilst the absolute distance of the object may not be overly affected, in relation to other objects nearby the additional displacement can be significant. A well known example is the 'Finger of God' [Jackson 1972]. Internal motions in a cluster make it appear highly elongated along the line of sight. Other peculiar velocities may also be important.

In order for a perturbation to grow, matter must flow towards it. This flow will cause the perturbation to be enhanced when viewed in redshift space (§5.2). As Kaiser [1987] has emphasised, projection effects may well account for some of the forms seen in galaxy redshift surveys. The latter part of the thesis explores the effect in greater detail. In particular, it examines the hypothesis that quasar absorption lines are structures apparent only in redshift space.

1.7 Thesis Outline

The remainder of the thesis is set out as follows. Chapter 2 presents a brief summary of the mathematical techniques used throughout the thesis. Some standard formulae are re-worked into a form more suitable for theoretical study. Chapter 3 applies some of these techniques to a study of various initial conditions and how they evolve. In Chapter 4, correlation functions are examined in some detail. From the considerations in these two chapters, it will be argued that there is a strong case for the initial perturbations to have a spectral power index of -3.

Chapter 5 presents the theory of the effects of projecting particle positions into redshift space. These concepts are applied in Chapter 6, where the idea that quasar absorption lines are caustics in redshift space is studied in depth. Conclusions arising from this work are presented in Chapter 7.

Chapter 2. Mathematical Background

This chapter presents a summary of the techniques used throughout the thesis. In particular, first order perturbation theory is covered in some depth. All the material in this chapter has been published elsewhere—most can be found in chapter two of *The Large-Scale Structure of the Universe* by Peebles [1980].

2.1 First Order Perturbation Theory

2.1.1 The Background Solution

As with any perturbation approach, it is important to establish the background solution. In this case, the unperturbed solution represents a homogeneous and isotropic world model. The cosmological constant, Λ, is assumed to be zero throughout.

It is appropriate to use co-moving co-ordinates. If the physical co-ordinate of a particle is $\mathbf{r}(t)$, then the equivalent co-moving co-ordinate, $\mathbf{x}(t)$, is given by

$$\mathbf{r}(t) = a(t)\mathbf{x}(t), \tag{2.1.1}$$

where $a(t)$ is the expansion scale factor at time t. The co-moving co-ordinate of a particle remains constant in the background solution. The explicit dependence of a on t will generally not be indicated in what follows. Taking the derivative of (2.1.1) with respect to time yields:

$$\dot{\mathbf{r}} = \dot{a}\mathbf{x} + a\dot{\mathbf{x}}. \tag{2.1.2}$$

The first term represents the expansion velocity, the second can be interpreted as a peculiar velocity, the motion of the particle with respect to the co-moving frame.

With the assumption of homogeneity and isotropy, Einstein's gravitational field equations in a matter dominated universe reduce to

$$\ddot{a} = -\tfrac{4}{3}\pi G a \rho_b, \tag{2.1.3}$$

where $\rho_b = \rho_b(t)$ is the background density at time t. ρ_b is given by

$$\rho_b = \rho_b(t_o)\left(\frac{a_o}{a}\right)^3, \tag{2.1.4}$$

where $a_o = a(t_o)$, the scale factor at some time t_o. $\rho_b(t_o)$ can be expressed as a fraction Ω_o of the critical density, $\rho_{crit}(t_o)$, the density just sufficent to close the universe:

$$\rho_b(t_o) = \Omega_o \rho_{crit}(t_o) = \Omega_o \frac{3H^2(t_o)}{8\pi G}. \tag{2.1.5}$$

$H(t)$ is defined to be the ratio $\dot{a}/a$. Solutions to these equations, together with a derivation of the expression for ρ_{crit}, can be found in standard texts [e.g., Weinberg 1972; Zel'dovich and Novikov 1983].

One futher cosmological result, needed in a later chapter, will be derived now. The redshift used to identify an epoch, $z(t)$, is given by

$$1 + z(t) = \frac{a_o}{a(t)}. \tag{2.1.6}$$

t_o represents the present for the remainder of this section. Taking the derivative with respect to time gives

$$\frac{dz}{dt} = -\frac{a_o}{a^2}\dot{a} = -(1+z)H(z). \tag{2.1.7}$$

$H(z)$ is well known to be given by

$$H(z) = H_o(1+z)(1+\Omega_o z)^{1/2}, \tag{2.1.8}$$

where $H_o \equiv H(t_o)$.

The result we wish to obtain is the following. Suppose there are two points on the same line of sight, separated by a redshift dz at a redshift z. What is their physical separation? To obtain this result, let us consider the propagation of photons in these models. Neglecting displacements perpendicular to the line of sight, the metric can be written in the form:

$$ds^2 = c^2dt^2 - a^2dx^2. \tag{2.1.9}$$

x is the co-moving co-ordinate along the line of sight. A photon emitted at some epoch will travel towards us down the line of sight, governed by $ds = 0$. This implies that

$$dx = -\frac{c}{a}dt = -\frac{c}{a_o}(1+z)dt. \tag{2.1.10}$$

The $-$ sign is approriate here since an increase in dt corresponds to a decrease in dx. Substituting for dt from (2.1.7),

$$dx = \frac{c}{a_o H(z)}dz. \tag{2.1.11}$$

Thus, when a photon emitted at an epoch identified by z has moved a co-moving distance dx towards us, it will be in an epoch characterised by $z - dz$. Hence dx is the co-moving separation corresponding to dz. The physical separation of the two points is given by

$$dr = a\,dx = \frac{c}{(1+z)H(z)}dz. \tag{2.1.12}$$

2.1.2 Equations of Motion

The perturbation in the density field is specified in terms of the dimensionless density contrast, $\delta(\mathbf{x}, t)$, defined to be

$$\delta(\mathbf{x}, t) = \frac{\rho(\mathbf{x}, t) - \rho_b}{\rho_b}. \tag{2.1.13}$$

This in turn is specified by its Fourier components:

$$\delta(\mathbf{x},t) = \sum_{\mathbf{k}\neq\mathbf{0}} \widetilde{\delta}(\mathbf{k},t)\mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}}. \tag{2.1.14}$$

The $\mathbf{k} = \mathbf{0}$ term is excluded since $\langle\delta(\mathbf{x},t)\rangle = 0$. Since the overdensity is real, the Fourier components obey the relation $\widetilde{\delta}(\mathbf{k},t) = \widetilde{\delta}^*(-\mathbf{k},t)$.

Poisson's equation in co-moving co-ordinates takes the form [Peebles 1980 eqn. 7.9]

$$\nabla^2\phi = 4\pi G\rho_b a^2\delta(\mathbf{x},t), \tag{2.1.15}$$

whilst the equations governing the motion of an ideal pressureless fluid are [Peebles 1980 eqn. 9.17]

$$\frac{\partial\mathbf{v}}{\partial t} + \frac{1}{a}(\mathbf{v}\cdot\nabla)\mathbf{v} + \frac{\dot{a}}{a}\mathbf{v} = -\frac{1}{a}\nabla\phi, \tag{2.1.16a}$$

$$\frac{\partial\delta}{\partial t} + \frac{1}{a}\nabla\cdot(1+\delta)\mathbf{v} = 0. \tag{2.1.16b}$$

In the last two equations, $\mathbf{v}$ is the peculiar velocity of the fluid, measured in physical units (equivalent to $a\dot{\mathbf{x}}$ in (2.1.2)).

These equations can be combined to give the linearised perturbation equation [Peebles 1980 eqn. 10.2]

$$\frac{\partial^2\delta}{\partial t^2} + 2\frac{\dot{a}}{a}\frac{\partial\delta}{\partial t} = 4\pi G\rho_b\delta, \tag{2.1.17}$$

where terms second order in δ or $\mathbf{v}$ have been neglected.

It will be convenient to re-express this equation in terms of the dimensionless time, μ, defined as

$$\mu \equiv H_o t \tag{2.1.18}.$$

Together with the expression for ρ_b, (2.1.17) becomes

$$\frac{\partial^2\delta}{\partial\mu^2} + 2\frac{a'}{a}\frac{\partial\delta}{\partial\mu} = \frac{3\Omega_o}{2}\left(\frac{a_o}{a}\right)^3\delta, \tag{2.1.19}$$

where $'$ denotes differentiation with respect to μ.

The derivatives in this equation are all with respect to μ. Solutions are of the form

$$\delta(\mathbf{x}, \mu) = \Delta(\mathbf{x}) D(\mu). \tag{2.1.20}$$

Substituting this expression into (2.1.19) yields an equation for $D(\mu)$, the linear growth factor for perturbations:

$$D'' + 2\frac{a'}{a}D' = \frac{3\Omega_o}{2}\left(\frac{a_o}{a}\right)^3 D. \tag{2.1.21}$$

Solutions to this expression can be found in Weinberg [1972]. Since this is a second order differential equation, there are two solutions. However, only the growing mode will cause structures—equation (2.1.20) represents the most general solution of interest to us. We can now express the Fourier components of $\delta(\mathbf{x}, \mu)$ as

$$\delta(\mathbf{k}, \mu) = \Delta_{\mathbf{k}} D(\mu), \tag{2.1.22}$$

where the $\Delta_{\mathbf{k}}$ are the Fourier components of the initial perturbation, $\Delta(\mathbf{x})$.

2.1.3 Velocity Field

An expression for the velocity field can also be found in the linear approximation. Let us define $\mathbf{u}$ to be the peculiar velocty of the fluid, measured in co-moving distance units per dimensionless time. This is related to $\mathbf{v}$ by

$$\mathbf{v} = H_o a \mathbf{u}. \tag{2.1.23}$$

Substituting this expression for $\mathbf{v}$ into equation (2.1.16a), re-expressed in terms of the dimensionless time gives:

$$\frac{\partial \mathbf{u}}{\partial \mu} + 2\frac{a'}{a}\mathbf{u} = -\frac{1}{a^2 H_o^2}\nabla\phi. \tag{2.1.24}$$

Again, second order terms have been neglected.

To continue, we need an expression for ϕ. Substituting (2.1.14) into (2.1.15) results in

$$\phi = -4\pi G\rho_b a^2 \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\widetilde{\delta}(\mathbf{k},\mu)}{|\mathbf{k}|^2} e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{2.1.25}$$

Using the linear approximation for $\widetilde{\delta}(\mathbf{k},\mu)$, and the expression for ρ_b, the right hand side of equation (2.1.24) becomes

$$-\frac{1}{a^2 H_o^2}\nabla\phi = \frac{3\Omega_o}{2}\left(\frac{a_o}{a}\right)^3 D(\mu) \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\mathbf{k}\Delta_\mathbf{k}}{|\mathbf{k}|^2} e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{2.1.26}$$

Defining

$$\Psi(\mathbf{x}) = \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\Delta_\mathbf{k}}{|\mathbf{k}|^2} e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{2.1.27}$$

our final equation for $\mathbf{u}$ is given by

$$\frac{\partial\mathbf{u}}{\partial\mu} + 2\frac{a'}{a}\mathbf{u} = \frac{3\Omega_o}{2}\left(\frac{a_o}{a}\right)^3 D(\mu)\nabla\Psi. \tag{2.1.28}$$

This has a growing solution

$$\mathbf{u} = D'(\mu)\nabla\Psi. \tag{2.1.29}$$

2.2 The Zel'dovich Approximation

The perturbation theory presented above has somewhat limited applications. A better approximation, due to Zel'dovich [1970] can easily be derived using (2.1.29). The approximation assumes that a particle placed initially at $\mathbf{q}$ (at $\mu = 0$), always has the velocity given by (2.1.29). Its position at any future time is given by integrating that equation:

$$\mathbf{x}(\mathbf{q},\mu) = \mathbf{q} + D(\mu)\nabla\Psi(\mathbf{q}). \tag{2.2.1}$$

$\mathbf{q}$ may be viewed as the Lagrangian co-ordinate of the particle, $\mathbf{x}$ as the corresponding Eulerian co-ordinate.

It is also possible to calculate the density field under this approximation. Particles are distributed uniformly in the Lagrangian co-ordinate system; calculating the Jacobian of the transformation to Eulerian co-ordinates provides the key to evaluating densities. If the volume element in the $\mathbf{q}$ co-ordinate system is $dV_{\mathbf{q}}$, that in the $\mathbf{x}$ system, $dV_{\mathbf{x}}$, is given by

$$dV_{\mathbf{x}} = dV_{\mathbf{q}} \left| \frac{\partial x_i}{\partial q_j} \right| . \tag{2.2.2}$$

$|\partial x_i/\partial q_j|$ is the Jacobian of the transformation (2.2.1), the determinant of the matrix $J_{ij} = \partial x_i/\partial q_j$. x_i is the i^{th} component of the vector $\mathbf{x}$, q_j the j^{th} component of $\mathbf{q}$.

The real space density associated with the point $\mathbf{q}$ at a dimensionless time, μ, can be written as

$$\rho(\mathbf{q}, \mu) = \rho_b \frac{dV_{\mathbf{q}}}{dV_{\mathbf{x}}} = \frac{\rho_b}{|\partial x_i/\partial q_j|}. \tag{2.2.3}$$

The Eulerian co-ordinate, $\mathbf{x}$, with this density can be determined from equation (2.2.1).

Let us consider this last expression in more detail. From (2.2.1),

$$J_{ij} = \frac{\partial x_i}{\partial q_j} = \delta_{ij} + D(\mu) \frac{\partial^2 \Psi(\mathbf{q})}{\partial q_i \partial q_j}. \tag{2.2.4}$$

The last term on the right-hand side, $\Sigma_{ij} = \partial^2 \Psi(\mathbf{q})/\partial q_i \partial q_j$, is the tidal tensor; δ_{ij} is the Kronecker delta. Since J_{ij} is symmetric, it can be reduced to diagonal form; the eigen-vectors are those of Σ_{ij} since δ_{ij} is an isotropic tensor.

Let the eigen-values of Σ_{ij} be $\lambda_i(\mathbf{q})$. Following Zel'dovich, the Jacobian can be written in the form

$$|J_{ij}| = \prod_{i=1}^{d} \{1 + D(\mu)\lambda_i(\mathbf{q})\}. \tag{2.2.5}$$

d is the dimensionality of the system under study. This equation shows that if one of the eigen-values of the tidal tensor is negative (compressive), at some time the Jacobian will have the value zero. The real space density, given by (2.2.3), will be infinite. This is the process of caustic formation. Caustics will, of course, affect those points with the most negative eigenvalues first. We will denote the growth factor which will first cause a point of infinite density D_o.

There is one last point to be noted before moving on. As one of the factors in equation (2.2.5) tends to zero, it will dominate the determination of the density. The collapse becomes pseudo one-dimensional close to caustic formation.

The approximation has been found to work surprisingly well up until the collapse of the first structures. It is used frequently in this thesis as a convenient means of extrapolating into the non-linear regime ($\delta \gtrsim 1$).

2.3 Initial Conditions

The preceeding sections in this chapter have dealt with analytic approximations for the evolution of perturbations in a Friedman Universe. This section considers the specification of the initial conditions—the problem is to generate a random density field. The presentation in this section closely follows that of Zel'dovich and Novikov [1983 chapter 13].

A reasonable definition of a random function is a function whose Fourier coefficients are random. If periodic boundary conditions are assumed, the problem has become one of specifying a set of discrete numbers rather than a continuous function. If the Fourier coefficient is $A_{\mathbf{k}}$, then let us define $B_{\mathbf{k}}$ and $C_{\mathbf{k}}$ to be the real and imaginary components of $A_{\mathbf{k}}$ respectively. Let us now consider an

ensemble of N similar systems. The number of systems, dN, which have $B_{\mathbf{k}}$ in the range $B_{\mathbf{k}}$ to $B_{\mathbf{k}} + dB_{\mathbf{k}}$ can be written

$$dN = NP(B_{\mathbf{k}})\, dB_{\mathbf{k}}. \tag{2.3.1}$$

A similar expression can be constructed for the number of systems, dN' with $C_{\mathbf{k}}$ in the range $C_{\mathbf{k}}$ to $C_{\mathbf{k}} + dC_{\mathbf{k}}$. If the distributions of $B_{\mathbf{k}}$ and $C_{\mathbf{k}}$ are independent, then the number of systems, dN'', which satisfy both pairs of inequalities is given by:

$$dN'' = NP(B_{\mathbf{k}})P'(C_{\mathbf{k}})\, dB_{\mathbf{k}}\, dC_{\mathbf{k}}. \tag{2.3.2}$$

Any volume at a late epoch will consist of many volumes which were causally unconnected (and presumably independent—although see the remarks on inflation, §1.1.3) at an earlier epoch. The central limit theorem then implies that $P(B_{\mathbf{k}})$ and $P'(C_{\mathbf{k}})$ are Gaussian in form;

$$P(B_{\mathbf{k}}) = \frac{1}{\sqrt{2\pi}\beta_{\mathbf{k}}} \exp\left(\frac{-{B_{\mathbf{k}}}^2}{2{\beta_{\mathbf{k}}}^2}\right), \tag{2.3.3}$$

with a similar expression for $P'(C_{\mathbf{k}})$, parameterised by $\gamma_{\mathbf{k}}$. Homogeneity and isotropy imply that $\beta_{\mathbf{k}} = \gamma_{\mathbf{k}} \equiv \alpha_{\mathbf{k}}/\sqrt{2}$, and that $\alpha_{\mathbf{k}} = \alpha(k)$ respectively. The $\sqrt{2}$ is introduced for convenience later.

It will be useful at this point to transform to polar co-ordinates. Let us define $a_{\mathbf{k}}$ and $\phi_{\mathbf{k}}$ to be the amplitude and phase of $A_{\mathbf{k}}$ respectively. Then using (2.3.3), the joint probability distribution of $a_{\mathbf{k}}$ and $\phi_{\mathbf{k}}$, $P(a_{\mathbf{k}}, \phi_{\mathbf{k}})$, is given by

$$P(a_{\mathbf{k}}, \phi_{\mathbf{k}})\, da_{\mathbf{k}}\, d\phi_{\mathbf{k}} = \frac{2a_{\mathbf{k}}}{{\alpha_{\mathbf{k}}}^2} \exp\left(-\frac{{a_{\mathbf{k}}}^2}{{\alpha_{\mathbf{k}}}^2}\right) da_{\mathbf{k}}\, \frac{d\phi_{\mathbf{k}}}{2\pi}. \tag{2.3.4}$$

This expression shows that the phases of the Fourier coefficients are uniformly randomly distributed in the range 0–2π. The amplitudes of the coefficients are drawn from a Rayleigh distribution, characterised by the function ${\alpha_{\mathbf{k}}}^2$.

With the above distribution, the random field created is an example of a Gaussian Random Field [Adler 1981; Bardeen *et al.* 1986]. Such fields have mathematical properties which make them amenable to analytic study. These fields will be studied on the computer in the following chapter, in a way which provides more insight than the formal analytic approach.

One useful result that will be needed later can be derived from (2.3.4): the ensemble average of $|A_{\mathbf{k}}|^2$.

$$\begin{aligned} \langle\langle |A_{\mathbf{k}}|^2 \rangle\rangle &= \int_0^\infty {a_{\mathbf{k}}}^2 \frac{2a_{\mathbf{k}}}{{\alpha_{\mathbf{k}}}^2} \exp\left(-\frac{{a_{\mathbf{k}}}^2}{{\alpha_{\mathbf{k}}}^2}\right) da_{\mathbf{k}} \\ &= {\alpha_{\mathbf{k}}}^2 \end{aligned} \tag{2.3.5}$$

(the reason for the $\sqrt{2}$ introduced earlier). $\alpha_{\mathbf{k}}$ is called the *power spectrum* of the perturbation and is frequently written as $P(k)$. *The statistical properties of the field are completely specified by its power spectrum.*

2.4 The Two-Point Correlation Function

The dimensionless two-point correlation function is frequently used to characterise a distribution (§1.2.1). Its relationship to the Fourier representation is developed below.

Using the definitions of $\xi(\mathbf{r})$ and $\delta(\mathbf{x}, t)$, the dimensionless two-point correlation function of $\rho(\mathbf{x}, t)$ is given by

$$\xi(\mathbf{r}, t) = \langle \delta(\mathbf{x}, t) \delta(\mathbf{x} + \mathbf{r}, t) \rangle. \tag{2.4.1}$$

Substituting the expression for $\delta(\mathbf{x}, t)$ in terms of its Fourier components, equation (2.1.14),

$$\xi(\mathbf{r}, t) = \sum_{\mathbf{k},\mathbf{k}' \neq \mathbf{0}} \langle \widetilde{\delta}(\mathbf{k}', t) \widetilde{\delta}(\mathbf{k}, t)\, \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \mathrm{e}^{i\mathbf{k}'\cdot(\mathbf{r}+\mathbf{x})} \rangle. \tag{2.4.2}$$

Only the $e^{i\mathbf{k}\cdot\mathbf{x}}$ and $e^{i\mathbf{k}'\cdot(\mathbf{r}+\mathbf{x})}$ terms depend on $\mathbf{x}$. Using the result

$$\langle e^{i\mathbf{k}\cdot\mathbf{x}} e^{i\mathbf{k}'\cdot\mathbf{x}} \rangle = \delta_{\mathbf{k},-\mathbf{k}'} = \begin{cases} 1, & \text{if } \mathbf{k} = -\mathbf{k}'; \\ 0, & \text{otherwise,} \end{cases} \tag{2.4.3}$$

the double sum reduces to

$$\xi(\mathbf{r}, t) = \sum_{\mathbf{k}\neq\mathbf{0}} |\widetilde{\delta}(\mathbf{k}, t)|^2 e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{2.4.4}$$

We have used $\widetilde{\delta}(\mathbf{k}, t) = \widetilde{\delta}^*(-\mathbf{k}, t)$ to obtain this result. This expression is an example of the Wiener-Khintchine theorem: the two-point correlation function is the Fourier transform of the power spectrum.

The above result highlights an important point; $\xi(\mathbf{r})$ only depends on the amplitudes of the Fourier coefficients, not their phases. One implication of this is discussed in Chapter 4.

This completes the summary of mathematical techniques.

Chapter 3. Study of Initial Conditions

The aim of this chapter is to establish a conceptual framework which will allow initial perturbations and their early evolution to be understood. Only initial conditions which are examples of Gaussian random fields (§2.2) will be examined. Particular attention will be paid to perturbations which have power spectra of the form[1]

$$P(k) \sim k^n. \tag{3.1}$$

n is the spectral power index. As mentioned in chapter 1, such spectra may be suitable since there is no conspicuous length scale. Also, any power spectrum can be represented locally by a power law. Studying power spectra which have the form k^n is a step towards understanding an arbitrary power spectrum.

3.1 Classification as a Function of Spectral Index

In this section, the expected overall evolution of the perturbations as a function of n is considered. Before this, however, let us briefly discuss the effects and necessity of a short-wavelength cut-off.

3.1.1 Effect of Cut-off

Instead of employing power spectra which are a pure power law, we will actually study spectra of the form

$$P(k) \sim k^n W(k/k_c). \tag{3.1.1}$$

$W(s)$ is a monotonically decreasing function which has the properties that

$$W(s) = \begin{cases} 1, & \text{if } s \ll 1; \\ 0, & \text{if } s \gg 1. \end{cases} \tag{3.1.2}$$

[1] The expression 'power law field' will be used as a convenient way of denoting perturbations with power spectra of this form.

Thus $W(k/k_c)$ acts as a high frequency filter. Let us also make the transformation $s = k/k_c$;

$$P(s) \sim s^n \, W(s). \tag{3.1.3}$$

We see that the function W effectively sets the scale of the structures.

We can envision the effect of removing the cut-off by allowing k_c to tend to infinity. Equation (3.1.3) shows that structure is merely identically reproduced at a smaller scale. It is meaningless to discuss a power law spectra without a cut-off; structures would occur on an infinitely small scale, effectively at a point.

The high frequency cut-off may be more than just a mathematical construct. In some scenarios, high frequency modes are damped because of actual physical processes. Silk damping [Silk 1967], for example, prevents the growth of small scale adiabatic density perturbations in baryonic universes. A filter of a different kind is applied when using point objects as probes of a continuous field. This is futher discussed in relation to correlation functions in the next chapter.

Whether the function W is viewed as a filter or as a smoothing function is a matter of taste. The convolution theorem for Fourier transforms shows that a convolution (smoothing operation) in real space is equivalent to multiplying by a filtering function in Fourier space and vice-versa. The convolution function and the filtering function, $W^{1/2}$, are a Fourier transform pair. Thus, choosing a form for W which has too sharp a cut-off will cause 'ringing' (side lobes) in the convolution function. Conversely, if too smooth a function is used for the convolution, too much small scale power will be allowed through.

Two filtering functions are used in this thesis:

$$W_G(k/k_c) = \exp\left(-2\frac{k^2}{k_c^2}\right) \tag{3.1.4}$$

and

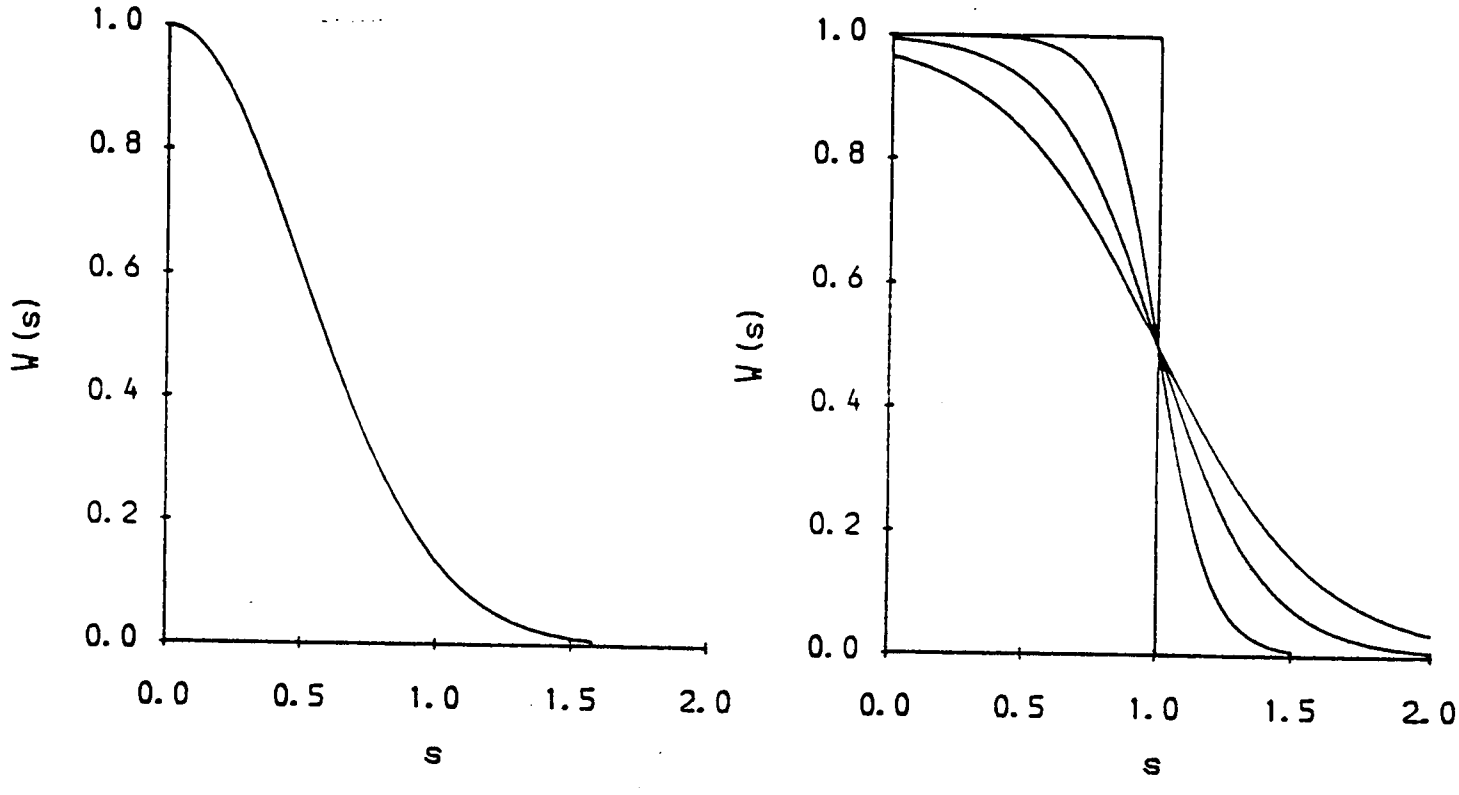


Figure 3.1. Graphs of the filtering functions. a) The Gaussian cut-off. b) The Fermi-Dirac cut-off for α values of 0, 0.1, 0.2 and 0.3. The sharper the cut-off, the smaller the value of α.

$$W_{FD}(k/k_c;\alpha) = \left\{\exp\left(\frac{k/k_c - 1}{\alpha}\right) + 1\right\}^{-1}. \qquad (3.1.5)$$

The Gaussian function optimally combines smoothing and filtering. The Fermi-Dirac cut-off allows the balance between filtering and smoothing to be changed by altering α. These functions are shown in Figures 3.1a and b.

Choosing a value for k_c effectively selects a length scale, the wavelength equivalent to k_c,

$$x_s = \frac{2\pi}{k_c}. \qquad (3.1.6)$$

Alternatively, x_s may be viewed as the smoothing scale. Lengths in the following graphs will be in terms of x_s.

3.1.2 Classification

Let us consider the mean square amplitude of the perturbations as a function of n and k_c. If there are d dimensions then

$$\Delta(\mathbf{x}) = \int d^d k\, \tilde{\Delta}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \tag{3.1.7}$$

where $\tilde{\Delta}(\mathbf{k})$ is the Fourier transform of $\Delta(\mathbf{x})$. Thus

$$\langle|\Delta(\mathbf{x})|^2\rangle \sim \int d^d k\, |\tilde{\Delta}(\mathbf{k})|^2. \tag{3.1.8}$$

Taking an ensemble average we can replace $|\tilde{\Delta}(\mathbf{k})|^2$ by $P(k)$. Using (3.1.1),

$$\langle|\Delta(\mathbf{x})|^2\rangle_{k_c} \sim \int d^d k\, k^n\, W(k/k_c). \tag{3.1.9}$$

Again making the substitution $s = k/k_c$,

$$\langle|\Delta(\mathbf{x})|^2\rangle_{k_c} \sim k_c^{n+d} \int d^d s\, s^n\, W(s). \tag{3.1.10}$$

Since the integral is not a function of k_c, we have the final result

$$\langle|\Delta(\mathbf{x})|^2\rangle_{k_c} \sim k_c^{n+d} \sim x_s^{-n-d}. \tag{3.1.11}$$

This last expression represents the mean square fluctuation of a power law field when smoothed over a length x_s.

Thus we see that if $n > -d$, the r.m.s. fluctuation smoothed over smaller scales will be greater than that smoothed over larger scales. Smaller scales will collapse first. This idea forms the basis of the hierarchical scenario [Peebles 1980 and references therein]. It is also possible to formulate similar results for the gravitational potential and the velocity field. Equations (2.1.27) and (2.1.29) show that we can go through exactly the same analysis for both, but replacing k^n by k^{n-4} and k^{n-2} respectively.

$$\langle|\Psi(\mathbf{x})|^2\rangle_{k_c} \sim k_c^{n+d-4} \tag{3.1.12}$$

and

Spectral Index n	Potential dominated by	Velocity dominated by	Density dominated by
$n > 4-d$	small scales.	small scales.	small scales.
$4-d > n > 2-d$	large scales.	small scales.	small scales.
$2-d > n > -d$	large scales.	large scales.	small scales.
$-d > n$	large scales.	large scales.	large scales.

Table 1. Summary of the classification scheme, discussed in the text.

$$\langle |\mathbf{u}(\mathbf{x})|^2 \rangle_{k_c} \sim k_c^{n+d-2}. \tag{3.1.13}$$

The last three equations form the basis for the classification scheme summarised in Table 1.

3.1.3 Comments

Each case will be considered separately.

(i) $n > 4-d$. The potential, velocity and density fields are all dominated by small-scale perturbations. The smallest scales will collapse first, but will have high binding energies. Larger scales have lower possible binding energies and cannot incorporate the smaller objects. Hierarchical clustering will not take place. In order for hierarchical clustering to occur the gravitational potential should be dominated by large-scale perturbations.

(ii) $4-d > n > 2-d$. Hierarchical clustering will take place in this case—the potential is dominated by large-scale perturbations. The velocity field, however, is dominated by small scale perturbations. Coherent motion will only occur on the scale which is just collapsing. Large-scale coherent motions will not occur for this range of n (unless n is close to $2-d$—see §3.5.4).

(iii) $2-d > n > -d$. Only the density field is dominated by the small-scale perturbations. Hierarchical clustering and large-scale coherent motions will

occur. Strictly speaking, a large-scale cut-off should also be included in the power spectrum. The velocity field is formally infinite otherwise.

(iv) $-d > n$. The largest scales (the entire universe) will collapse first. This is perhaps the least interesting case.

3.2 One-Dimensional Fields

A useful step towards understanding the full three-dimensional problem is to make a detailed study of perturbations in only one dimension. The fields are much easier to display and computations are simplified. The real advantage, however, is that a much wider range of scales can be studied than in three dimensions. Periodic boundary conditions are imposed.

It is a reasonable assumption that structures form at the maxima of the density field; the properties of the maxima are suitable targets for investigation. Our approach will follow Peacock and Heavens [1985], rather than Bardeen *et al.* [1986]. Zeroes in the perturbing density field will also prove to be of interest. This is another reason for studying perturbations in one-dimension—zeroes of the field define points, rather than surfaces. In future, 'zeroes' will generally mean the set of points defined by the zeroes of the field.

3.2.1 Equivalent Realisations

A realisation of a power spectrum is set up by choosing the Fourier components of the field. The amplitudes and phases are drawn from the probability distribution (2.3.4). The aim of this chapter is to examine the properties of different power spectra; realisations of different spectra must be set up which are as similar as possible—*equivalent realisations*. The realisations should not alter the statistical properties of the fields.

Let the amplitudes and phases of a particular realisation be $a_{\mathbf{k}}$ and $\phi_{\mathbf{k}}$ respectively, the power spectrum be $P(k)$. An equivalent realisation of a spectrum $\tilde{P}(k)$ is established by setting $\tilde{\phi}_{\mathbf{k}} = \phi_{\mathbf{k}}$ and $\tilde{a}_{\mathbf{k}}^2 = a_{\mathbf{k}}^2 \tilde{P}(k)/P(k)$. $\tilde{a}_{\mathbf{k}}$ and $\tilde{\phi}_{\mathbf{k}}$ are the amplitudes and phases of the equivalent realisation.

Figures 3.2a–f show equivalent realisations of power spectra with different spectral indices but the same Gaussian cut-off. The field amplitude of a particular realisation is scaled in terms of σ, the r.m.s. fluctuation of the realisation. Several points should be noted:

(i) The range of amplitudes is similar for each realisation.

(ii) The positions of extrema can be traced from one realisation to another. Only the phases in the different realisations remain constant. This suggests that it is the relative phases (rather than the amplitudes) of the Fourier coefficients which determine the position of structures in the field.

(iii) For the Figures 3.2a–c maxima are almost all positive, minima negative. An increasing number of maxima become negative, minima positive, in later figures.

(iv) As the spectral index decreases, larger connected regions above or below zero become more prominent. This is obviously connected to the previous point.

(v) Figure 3.2f, which has $-d > n$, is dominated by a perturbation on the largest possible scale—the size of the box.

These considerations set the tone for later investigations. The positions of maxima are relatively insensitive to the spectral index; the value of n has more effect on the height of the extrema. As power is shifted to larger scales, extended connected regions above or below zero become apparent.

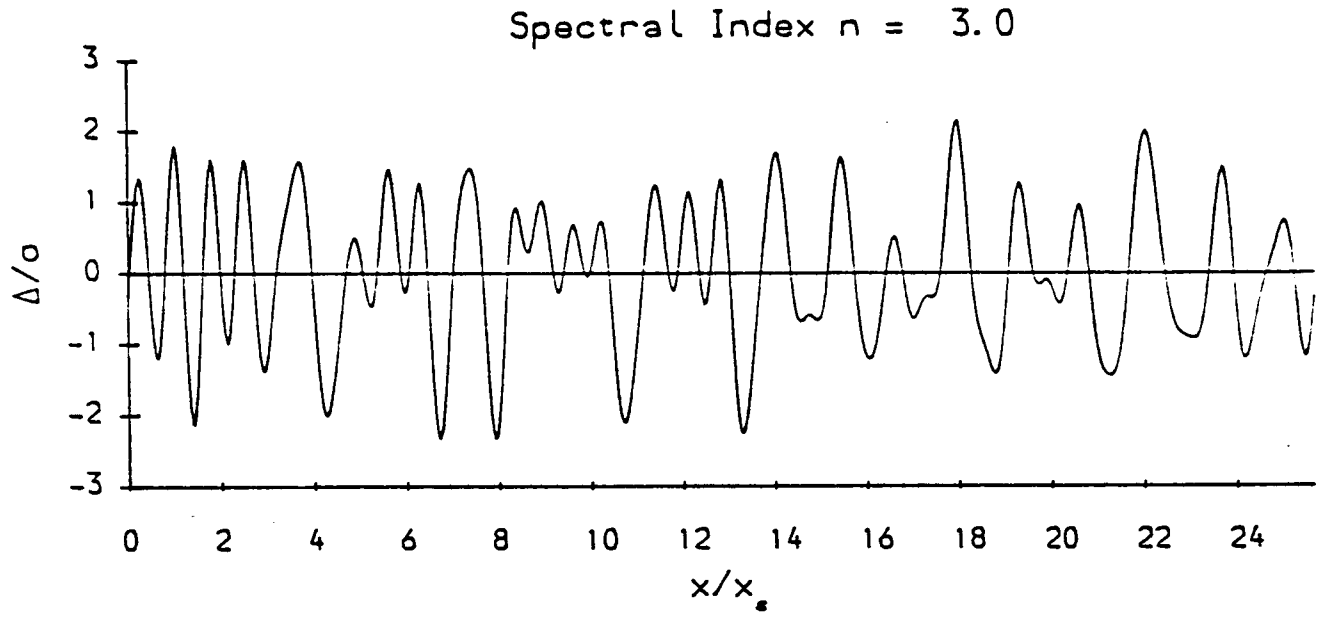
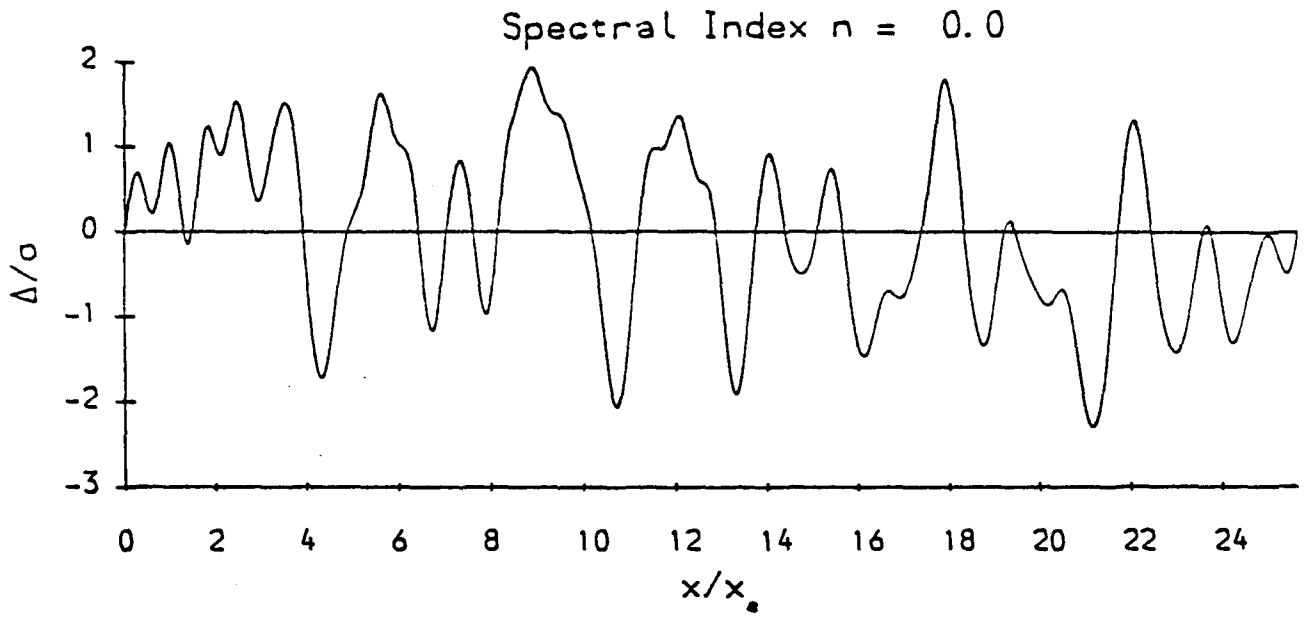
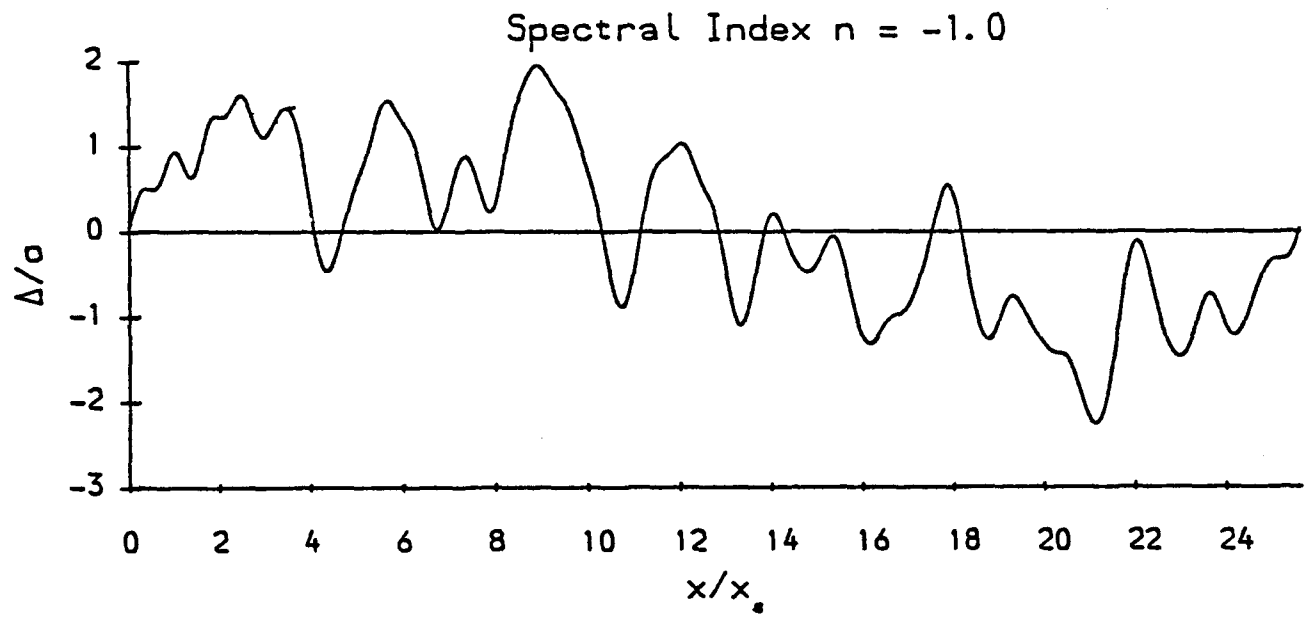


Figure 3.2a–c. Equivalent realisations of one-dimensional Gaussian random fields with varying spectral index.





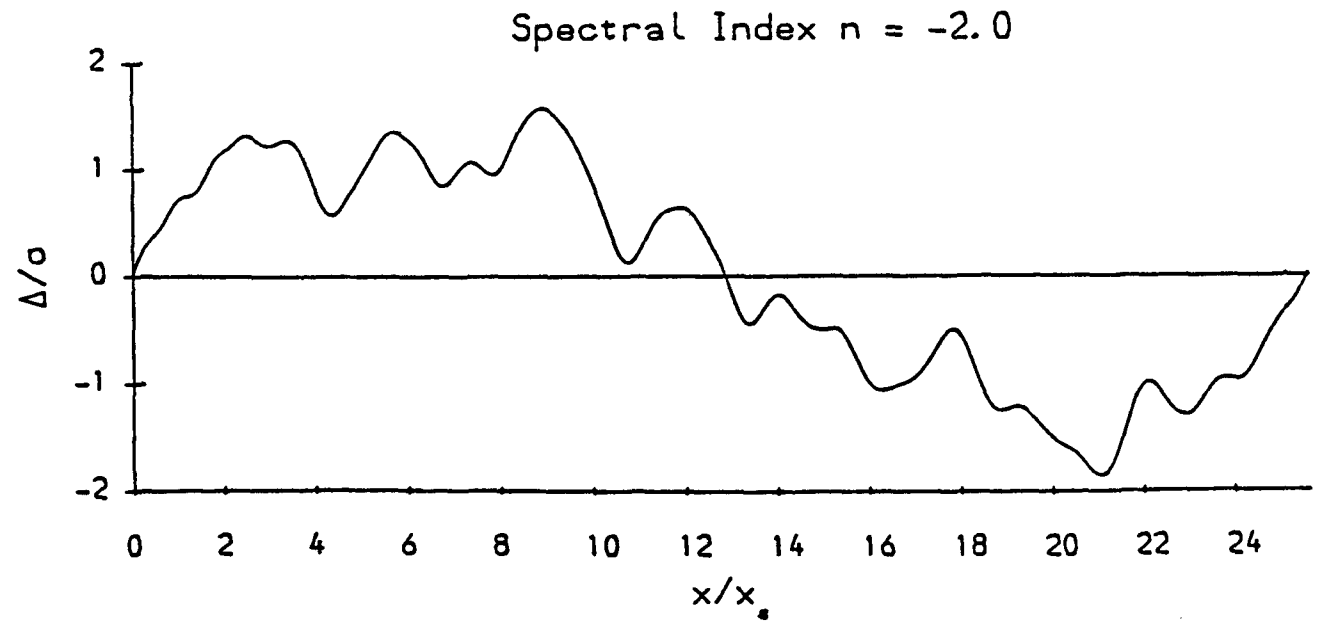


Figure 3.2d–f. Equivalent realisations of one-dimensional Gaussian random fields with varying spectral index.

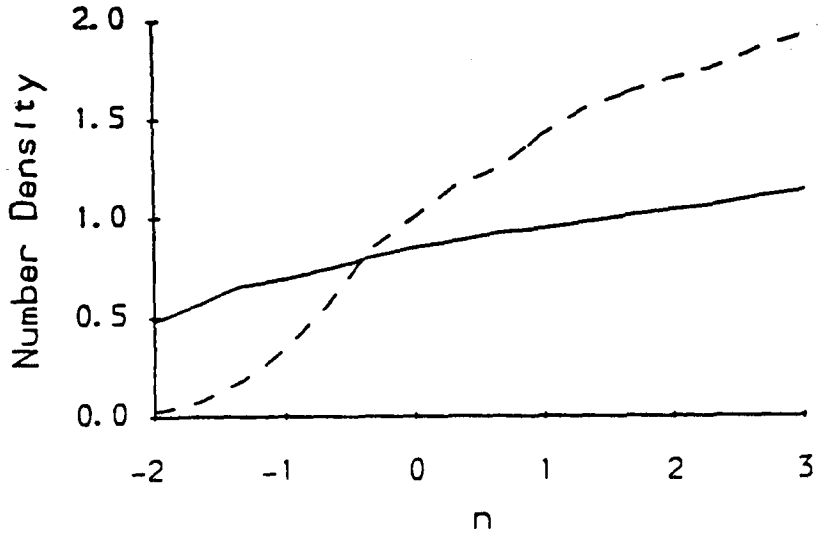


Figure 3.3. The mean number densities of zeroes and peaks plotted against the spectral index. Zeroes are represented by the dashed line.

3.3 Distribution of Peaks and Zeroes

This section is mainly concerned with the spatial distribution of zeroes and peaks, again for a one-dimensional Gaussian random field. The distribution of peak heights, however, is briefly discussed in §3.3.6. Except for the results presented in §3.3.7, a Gaussian window function is employed. When comparison is made between fields of different spectral power index, equivalent realisations are used. The positions of zeroes are determined using linear interpolation between grid points, peaks using quadratic interpolation. There are typically 6 grid points per length x_s. Unless otherwise stated, the realisations used are $\sim 330x_s$ in length.

3.3.1 Mean Number Densities

Let us define the mean number density of a set of points to be the average number of points per length x_s. The mean number densities of zeroes, $\langle n_z \rangle$, and peaks, $\langle n_p \rangle$, are plotted as functions of n in Figure 3.3.

The figure shows that $\langle n_p \rangle$ is only slightly dependent on the spectral index; $\langle n_p \rangle \sim 1$ for all n. The mean number density of zeroes, on the other hand,

depends significantly on n. For high values of n there are roughly two zeroes per peak—most maxima are positive, minima negative, implying two sign changes between peaks. As n decreases, so does $\langle n_z \rangle$.

3.3.2 Separation of Zeroes

Figures 3.4a–d show the frequency distribution of the separation of adjacent zeroes for various values of n. The distribution becomes increasingly extended towards large separations as n becomes less positive. Figure 3.5 shows the maximum separation of adjacent zeroes as a function of n. It is clear that the maximum separation increases rapidly as n decreases.

3.3.3 Separation of Peaks

Figures 3.6a–d show the frequency distribution of the separation of adjacent maxima for various n. Although the distribution does become extended as n decreases, the effect is less marked than in the separation of adjacent zeroes. The maximum separation of peaks never exceeds $4x_s$.

3.3.4 Peak-Peak Correlation Functions

The two-point correlation function for a set of discrete points can be defined in analogy to equation (1.2.1). Let the expected number of point pairs with separations in the range x to $x + \Delta x$ be $N_r(x)\Delta x$ for an equivalent random distibution. If the number for the actual distribution is $N_d(x)\Delta x$ then

$$\xi(x) = \frac{N_d(x) - N_r(x)}{N_r(x)} \tag{3.3.1}$$

(since we are dealing with a finite number of points, $\xi(x)$ strictly depends on the size of the bins, Δx, also).

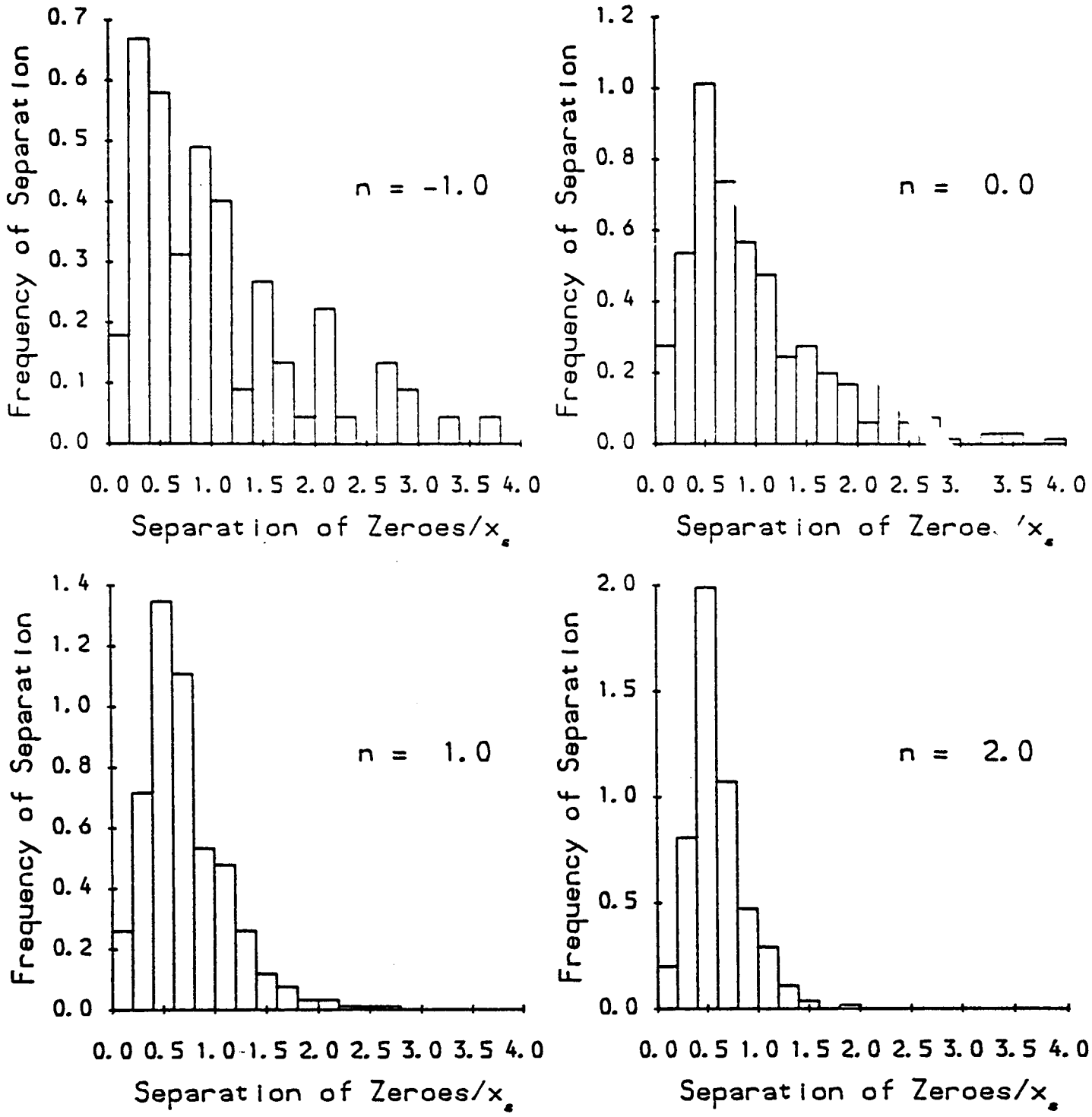


Figure 3.4. The frequency distributions of the separation of adjacent zeroes in a one-dimensional Gaussian random field for various n.

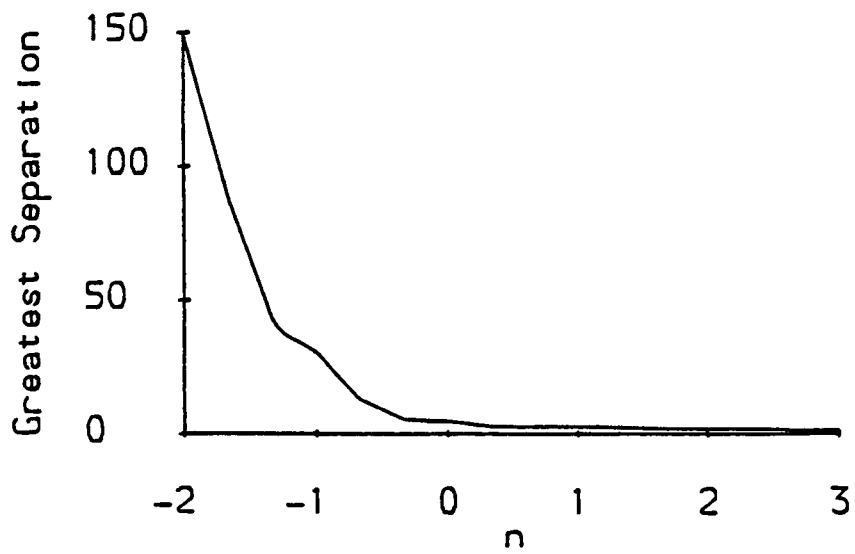


Figure 3.5. The maximum separation of adjacent zeroes, plotted as a function of n. Refer to text for details of the realisations used.

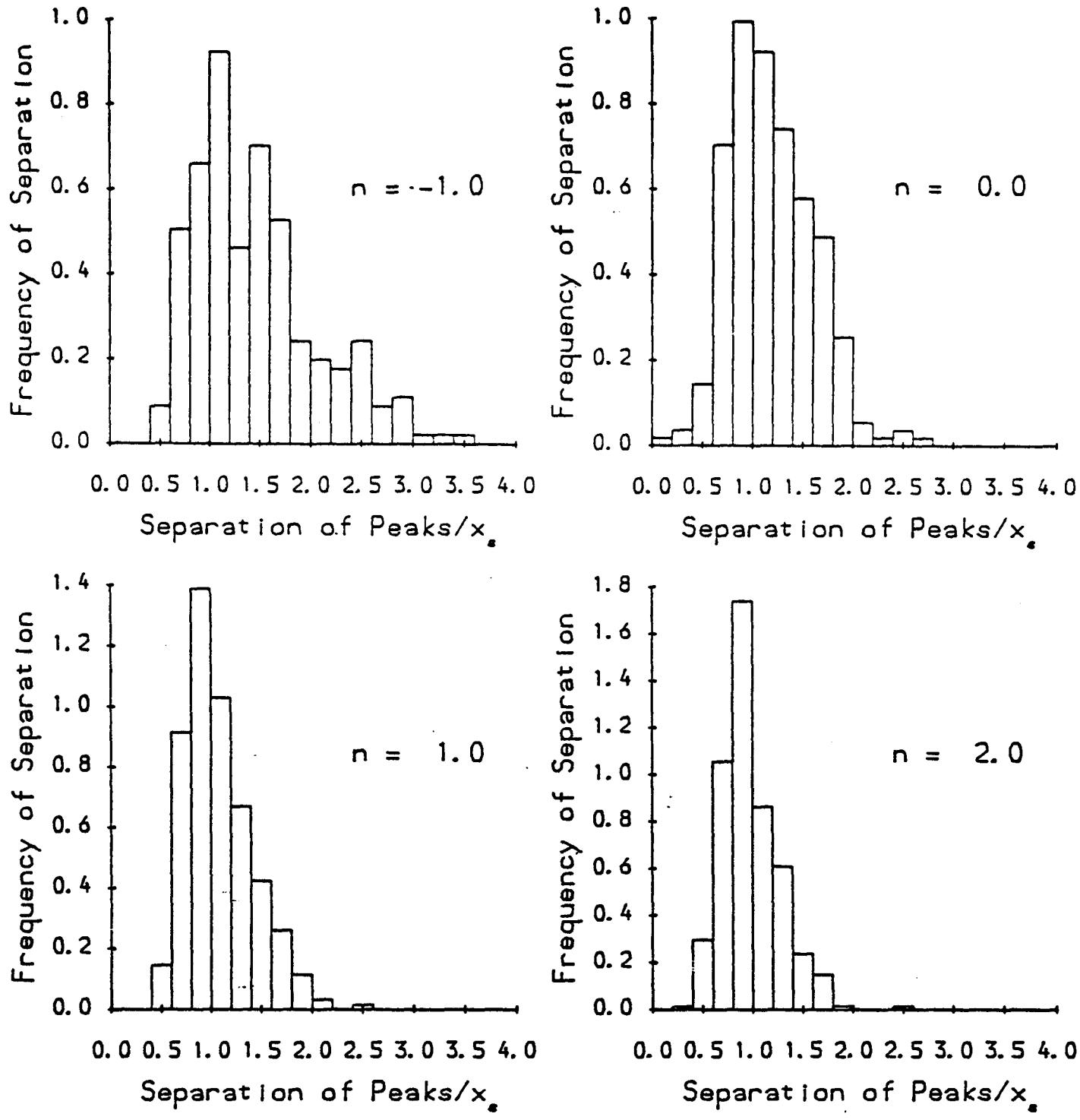


Figure 3.6. The frequency distributions of the separation of adjacent maxima for various spectral indices.

It is possible to make this expression more formal. Let us consider a set of m points at positions $x_1, x_2, \ldots, x_m$ distributed throughout a region of length X. Now define a function

$$Q(r, x; \Delta x) = \begin{cases} 1, & \text{if } x < r < x + \Delta x; \\ 0, & \text{otherwise.} \end{cases} \tag{3.3.2}$$

$N_d(x)\Delta x$ can be expressed in terms of Q as

$$N_d(x)\,\Delta x = \sum_{i>j} Q(|x_i - x_j|, x; \Delta x). \tag{3.3.3}$$

The correlation function can now be written

$$\xi(x) = \frac{\sum_{i>j} Q(|x_i - x_j|, x; \Delta x) - N_r(x)\Delta x}{N_r(x)\Delta x}. \qquad (3.3.4)$$

The expression for $\xi(x)$ has been written in this form in anticipation of the following section.

The peak-peak correlation function for a range of spectral indices is shown in Figures 3.7a–d. Realisations of length $\sim 5000x_s$ were used. The functions look remarkably similar; each resembles the radial density function for a glass. The amplitude of the oscillations about zero increases with n, reflecting the tighter distribution of adjacent peak separations discussed above.

The analogy with a radial density function is quite strong. It is difficult for peaks to be much closer than x_s; there is little power on smaller scales because of the cut-off. On the other hand, adjacent peaks cannot have separations much wider than x_s; large-scale power cannot efficiently suppress peaks on smaller scales.

3.3.5 Biasing

The idea behind biasing is that large-scale power will cause the heights of nearby peaks to be correlated. This concept is only normally applied to peaks above some high threshold; these peaks are taken to have more astrophysical significance. However, it should be equally applicable to all peaks. Previous authors have studied biasing by examining the distibution of peaks lying above the threshold [Peacock and Heavens 1985; Bardeen *et al.* 1986]. When imposing a threshold has any effect, the results are messy and hard to interpret [Peacock and Heavens 1985, Figure 6.]. This approach also discards information from other peaks.

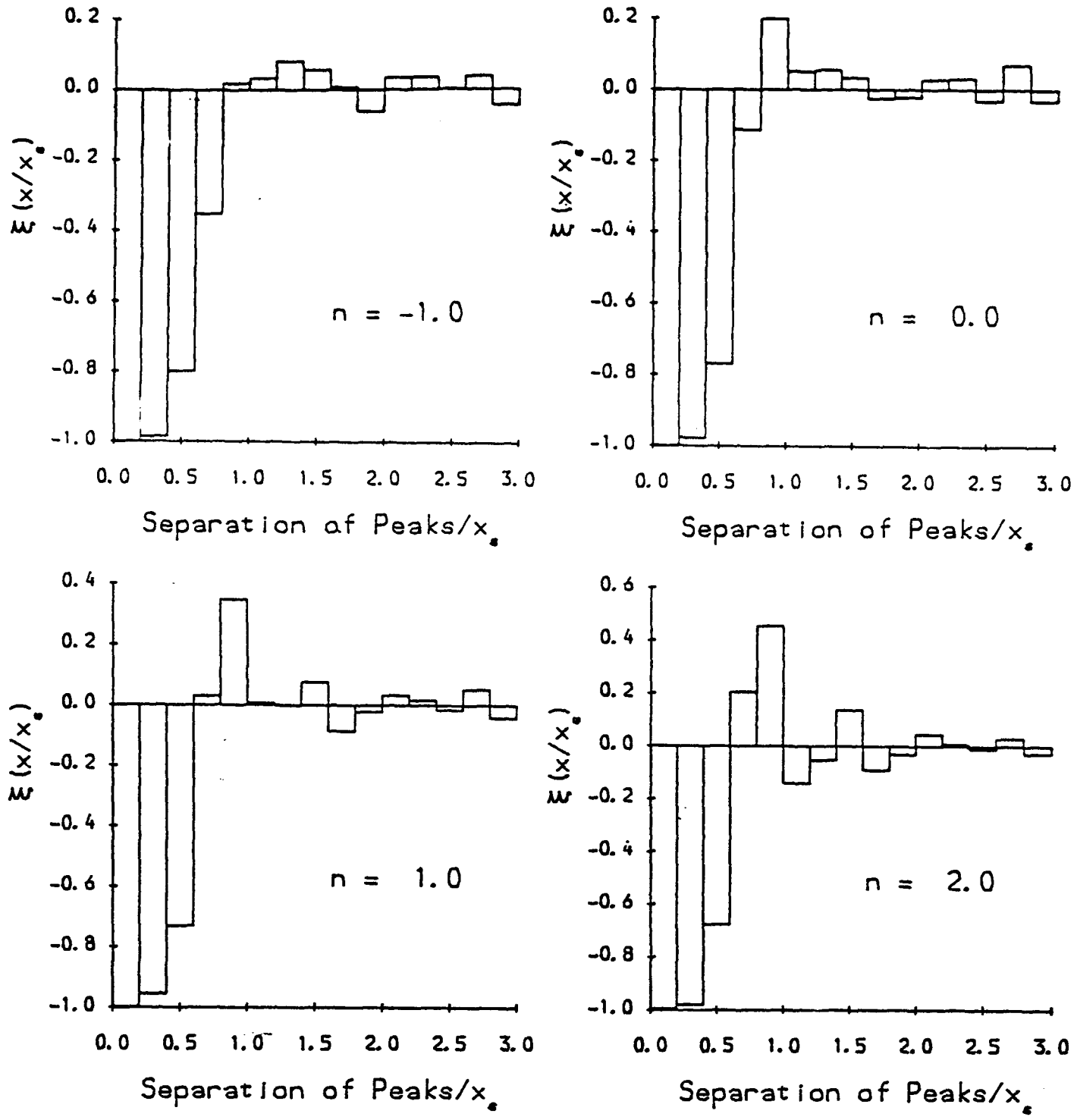


Figure 3.7. The peak-peak correlation functions for several values of n.

A different method is adopted here. Let the peaks have heights $\rho_1, \rho_2, \ldots, \rho_m$ at positions $x_1, x_2, \ldots, x_m$. Again assume that the peaks are distributed throughout a region of length X. Let

$$\langle \rho_p \rangle = \frac{1}{m} \sum_{i=1}^{m} \rho_i \tag{3.3.5}$$

and

$$\sigma_p^2 = \frac{1}{m} \sum_{i=1}^{m} (\rho_i - \langle \rho_p \rangle)^2 \tag{3.3.6}$$

be the mean peak height and the r.m.s. fluctuation of heights about the mean respectively. Consider now the weighted sum

$$S(x;\Delta x) = \sum_{i>j} Q(|x_i - x_j|, x; \Delta x) \frac{(\rho_i - \langle\rho_p\rangle)(\rho_j - \langle\rho_p\rangle)}{\sigma_p^2}, \tag{3.3.7}$$

i.e., the sum of $(\rho_i - \langle\rho_p\rangle)(\rho_j - \langle\rho_p\rangle)$ for all peak pairs with separations in the range x to $x + \Delta x$, divided by σ_p^2. If the heights of peaks with this separation are correlated, then the sum will be positive, if anti-correlated negative.

To make the sum as analgous to the definition of $\xi(x)$ as possible, it should futher be divided by the number of expected pairs with this separation. We will study the *weighted two-point correlation function*, $\Xi(x)$, defined by

$$\Xi(x) = \frac{\sum_{i>j} Q(|x_i - x_j|, x; \Delta x)(\rho_i - \langle\rho_p\rangle)(\rho_j - \langle\rho_p\rangle)}{\sigma_p^2 N_r(x) \Delta x}. \tag{3.3.8}$$

Note that the expected value of the function is zero if the peak distribution is random; there is no term subtracted from the numerator of the sum.

$\Xi(x)$ for a range of n is shown in Figures 3.8a–d. Again realisations of length $\sim 5000 x_s$ were used. Peak heights are correlated over an extended region only if n is close to -1. Figures 3.8c–d show that adjacent peaks are significantly anti-correlated for $n > 0$, i.e., neighbouring peak heights tend to be on opposite sides of the mean peak amplitude.

Let us futher define the *biasing length*, L_B, to be the separation at which $\Xi(x)$ first changes sign. A plot of L_B vs. n is shown in Figure 3.9. Realisations of length $\sim 1300 x_s$ were used in this calculation. Again this shows that biasing is only significant for n close to -1.

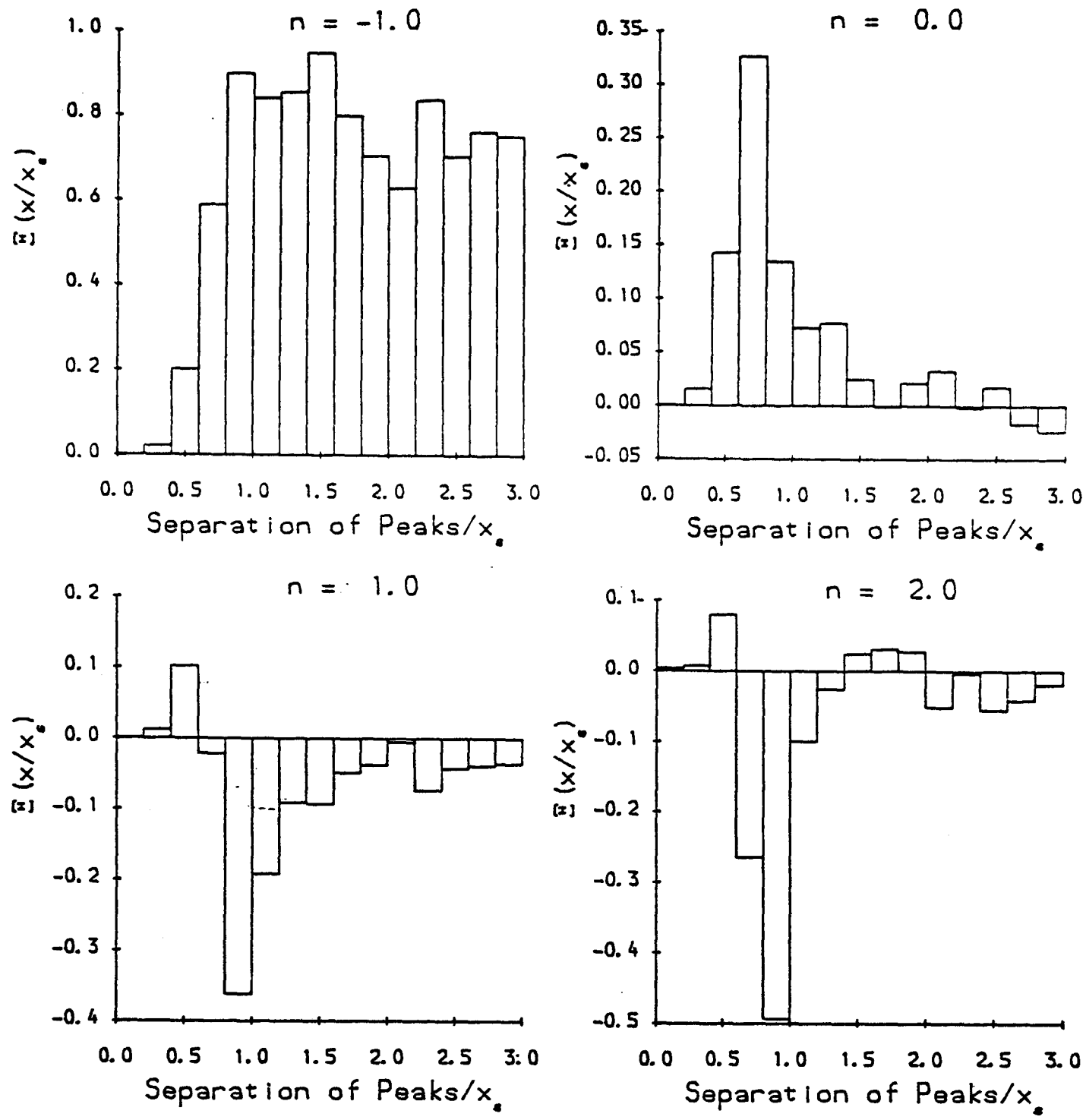


Figure 3.8. The *weighted* peak-peak correlation functions for several values of n.

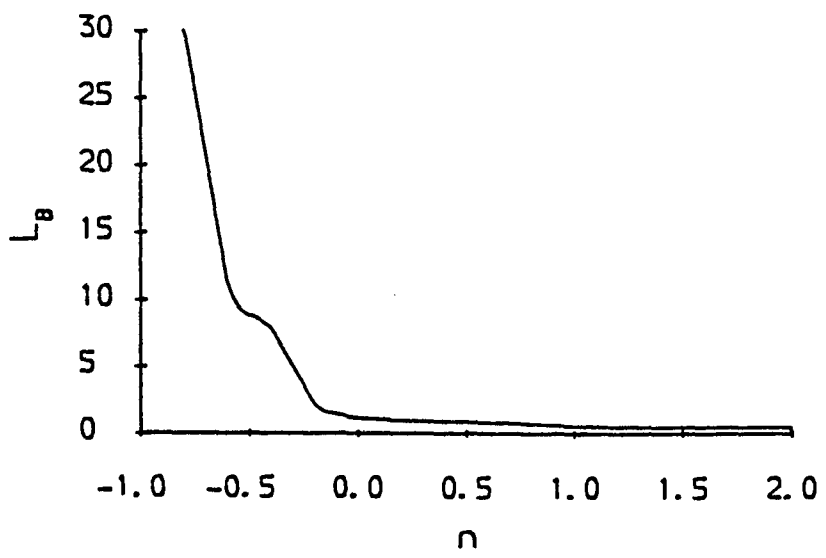


Figure 3.9. The biasing length, L_B, plotted against the spectral index.

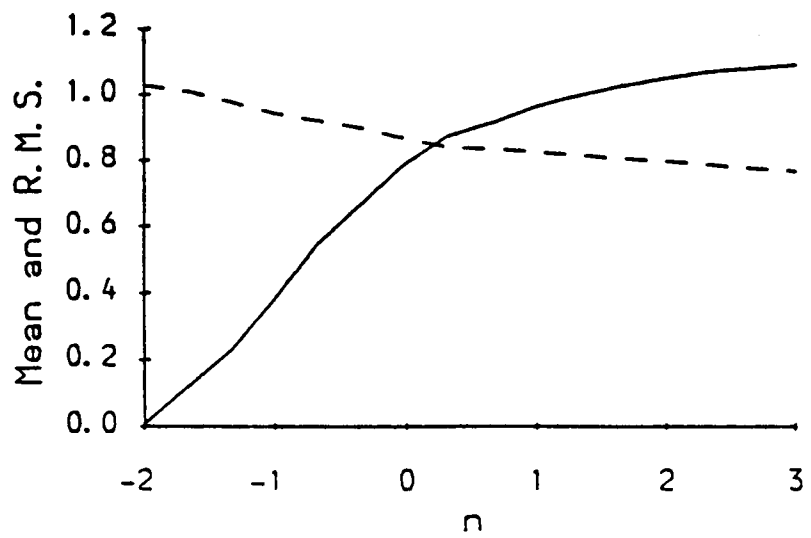


Figure 3.10. The solid curve shows the variation of the mean peak height with spectral index. The dashed curve plots the r.m.s. fluctuation of the peak heights about the mean.

3.3.6 Distribution of Peak Heights

It is relatively straightforward to calculate the distribution of peak heights directly from the power spectrum [Bardeen *et al.* 1986; Couchman 1987]. We will content ourselves with examining $\langle \rho_p \rangle$ and σ_p as functions of n, plotted in Figure 3.10. Both are plotted in units of σ, the r.m.s. fluctuation of the entire field. The figure shows that the distribution of peak heights depends significantly on n.

From results in previous sections, it is reasonable to conclude that the positions of peaks depends on the phases of the Fourier components near the high-frequency cut-off. The height of the peaks, however, must also depend on the large-scale modes. As n increases, more power is transfered to the small scale modes, the same modes which determine the positions of the maxima. Hence as n increases the distribution of peak heights becomes narrower. Fewer peaks occur at lower amplitudes, increasing the mean height. Conversely, as n decreases, the amplitudes of the peaks are more strongly influenced by power on large scales. Peaks no longer always occur at the higher amplitudes. The distribution of heights becomes wider and the mean decreases.

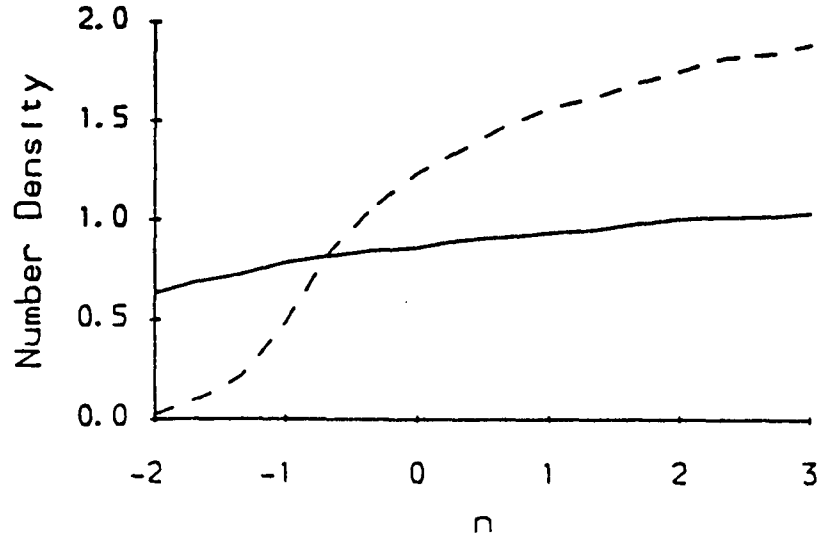


Figure 3.11. The mean number densities of zeroes and peaks for a Gaussian random field but with a Fermi-Dirac cut-off. Plotted as a function of n.

3.3.7 Effect of Cut-off Revisited

Earlier in this chapter, it was shown that a high-frequency cut-off to the power spectrum established the scale of structure in the fields. In this section, the effect of altering the *form* of the cut-off is examined. The Fermi-Dirac function with $\alpha = 0.1$ is employed throughout. Reference to Figures 3.1a and b will show that this is a somewhat sharper cut-off than the Gaussian function.

Figure 3.11 shows the mean number density of peaks and zeroes as a function of n. Comparison with Figure 3.3 reveals that $\langle n_p \rangle$ is slightly flatter with the sharper cut-off, whilst the number density of zeroes rises more sharply as n increases. Figures 3.12a–d show the frequency distribution of the separation of adjacent zeroes, Figures 3.13a–d the peak-peak correlation function, for a range of n. Whilst the distribution of zero separations is somewhat tighter, the peak-peak correlation reveals that there is a significantly tighter distribution of adjacent peak separations with the Fermi-Dirac cut-off.

With a sharper cut-off, the range of modes with amplitudes near the maximum is reduced. It is reasonable to interpret the above observations as evidence that the positions of maxima are determined by a restricted range of frequen-

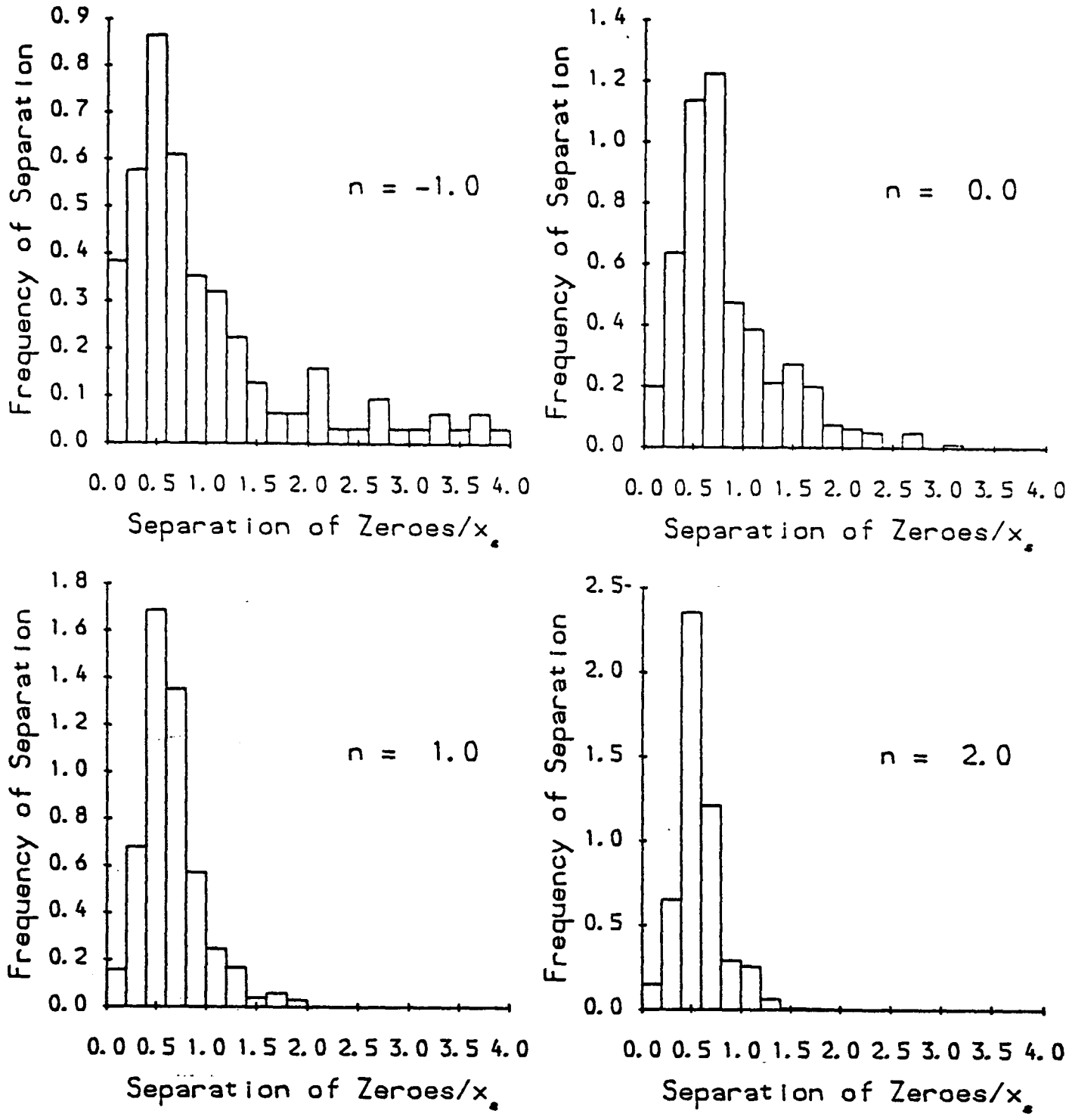


Figure 3.12. Frequency distributions of the separation of adjacent zeroes for various values of the spectral index. A Fermi-Dirac cut-off was used.

cies with the Fermi-Dirac function. Again, this leads to the conclusion that the positions of maxima are determined by the Fourier components of scales close to the cut-off.

3.3.8 Conclusions

The results presented above lead to the following conclusions:

(i) The spatial distribution of peaks is relatively insensitive to the spectral index.

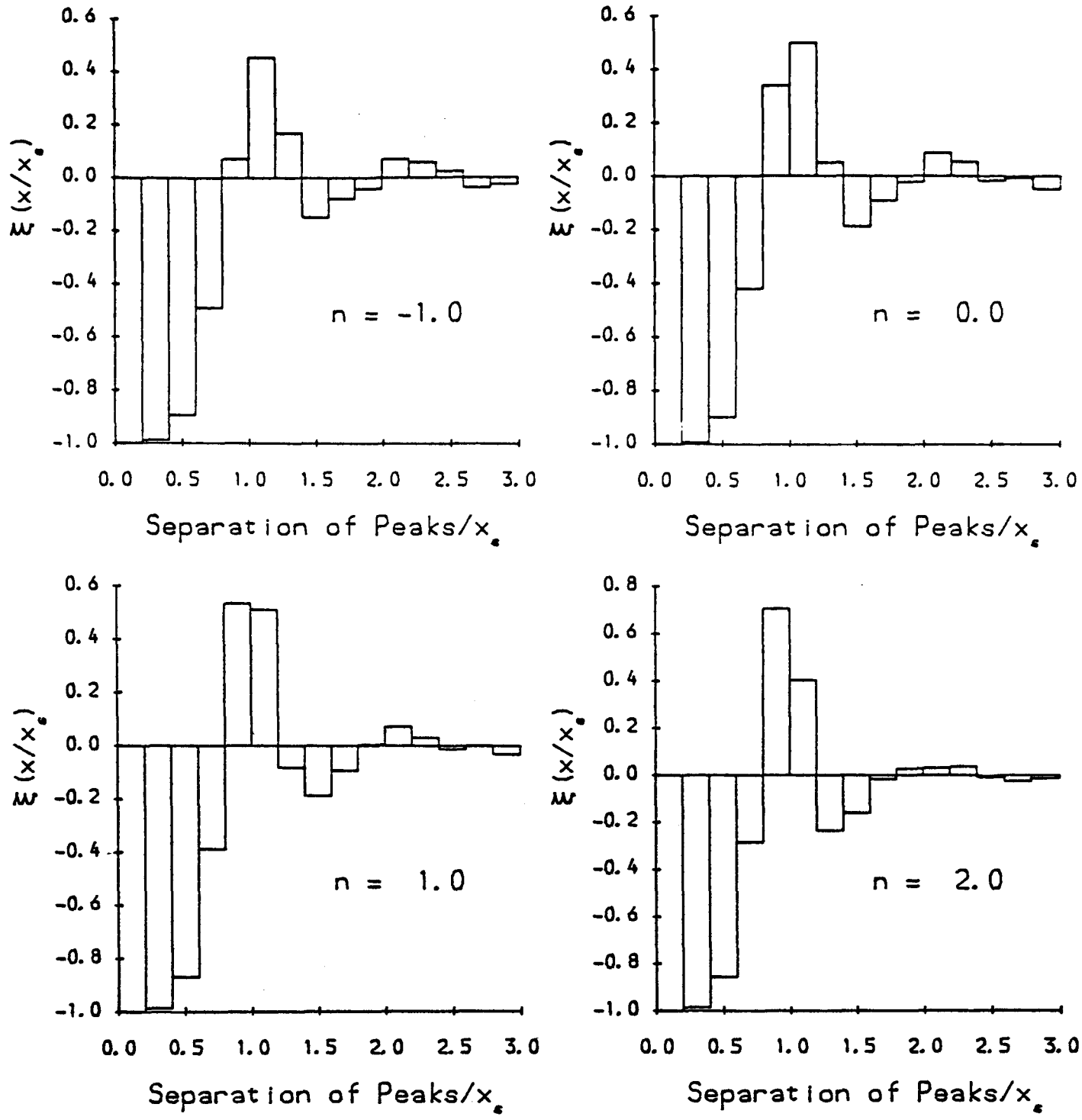


Figure 3.13. Peak-peak correlation functions with a Fermi-Dirac cut-off.

The positions of peaks are determined by the phases of the Fourier coefficients near to the high-frequency cut-off. Peaks tend to occur a distance x_s apart, the wavelength of the cut-off frequency. Relatively little information is contained in the peak-peak correlation function, which is analogous to the radial density function of an amorphous material. There is no evidence for correlation between the peaks, other than the tendency to occur at a fixed separation.

(ii) The amplitude distribution of peaks does depend on n—the height of a peak

is dependent on the large-scale modes. As more power is transferred to large scales (by decreasing n), a wider distribution of peak heights, with a lower mean amplitude, is produced.

(iii) Biasing is only important as n approaches -1. This is clearly related to the change over between large and small scale domination of the density field, discussed in §3.1.3. For a d-dimensional field, biasing will become more important as n approaches $-d$. The question remains, however, whether biasing will be important only for $n \sim -d$, or for $n < 0$. A similar study should be made of bias in three-dimensions to resolve the issue (although see §4.2).

(iv) The spatial distribution of zeroes does depend significantly on n—the distribution of zeroes is connected to the amplitude distribution of peaks. Extended regions above or below zero exist if n is close to -1. The last point is clearly related to the importance of biasing as n approaches -1.

Cosmological implications of these conclusions will be presented later in the chapter.

3.4 Non-Linear Evolution

Previous sections have been concerned with the form of the initial perturbations. This section will study their early non-linear evolution. The aim is to present a heuristic view, to provide an intuitive picture of initial gravitational collapse. Again only one-dimensional systems will be studied. Apart from the reasons discussed earlier in the chapter, collapse in three dimensions is generally dominated by motion in one direction; pseudo one-dimensional collapse [Zel'dovich 1970; §2.1]. The Zel'dovich approximation will be used throughout (§2.1). The approximation is ideally suited to the aims expressed above.

3.4.1 Spatial Evolution

We will consider the spatial evolution of the initial perturbation $(\sin x)/x$, as D/D_o increases. D is the growth factor for linear perturbations (§2.1.2), D_o the growth factor which will cause the first caustics to form (§2.1).

Figures 3.14a–d show the evolution. The amplitude, $\delta(x)$, is plotted in units of $\sigma(D)$, the r.m.s. fluctuation of the perturbation if purely linear evolution had occurred; changes from figure to figure are due to non-linear evolution alone. Scrutiny of Figures 3.14a–d shows that non-linear evolution results in a change of shape. There is a marked asymmetry between the growth of overdense and underdense regions. Overdense regions become narrow and strongly peaked with high densities. Underdense regions become rounder, wider and shallower. The growth of underdense regions is outward, rather than downward.

3.4.2 Fourier Evolution

To study the evolution of perturbations in Fourier space, we will use the initial perturbation $\exp\left(-x^2/\lambda^2\right)$. Figures 3.15a–d show the evolution of the amplitudes of the Fourier coefficients as D/D_o increases. The amplitudes are scaled such that the maximum amplitude for pure linear evolution, $a_{max}(D)$, would be 1. Changes in the figure sequence again represent non-linear evolution. The wavevectors, k, are plotted in units of $k_o = 2\pi/\lambda$.

The figures show that power is transferred to smaller scales (higher k) as the perturbation evolves. This reflects the increasingly peaked spatial evolution. The phases of the small scale components must be highly correlated to produce such sharp features (cf., the Dirac δ-function).

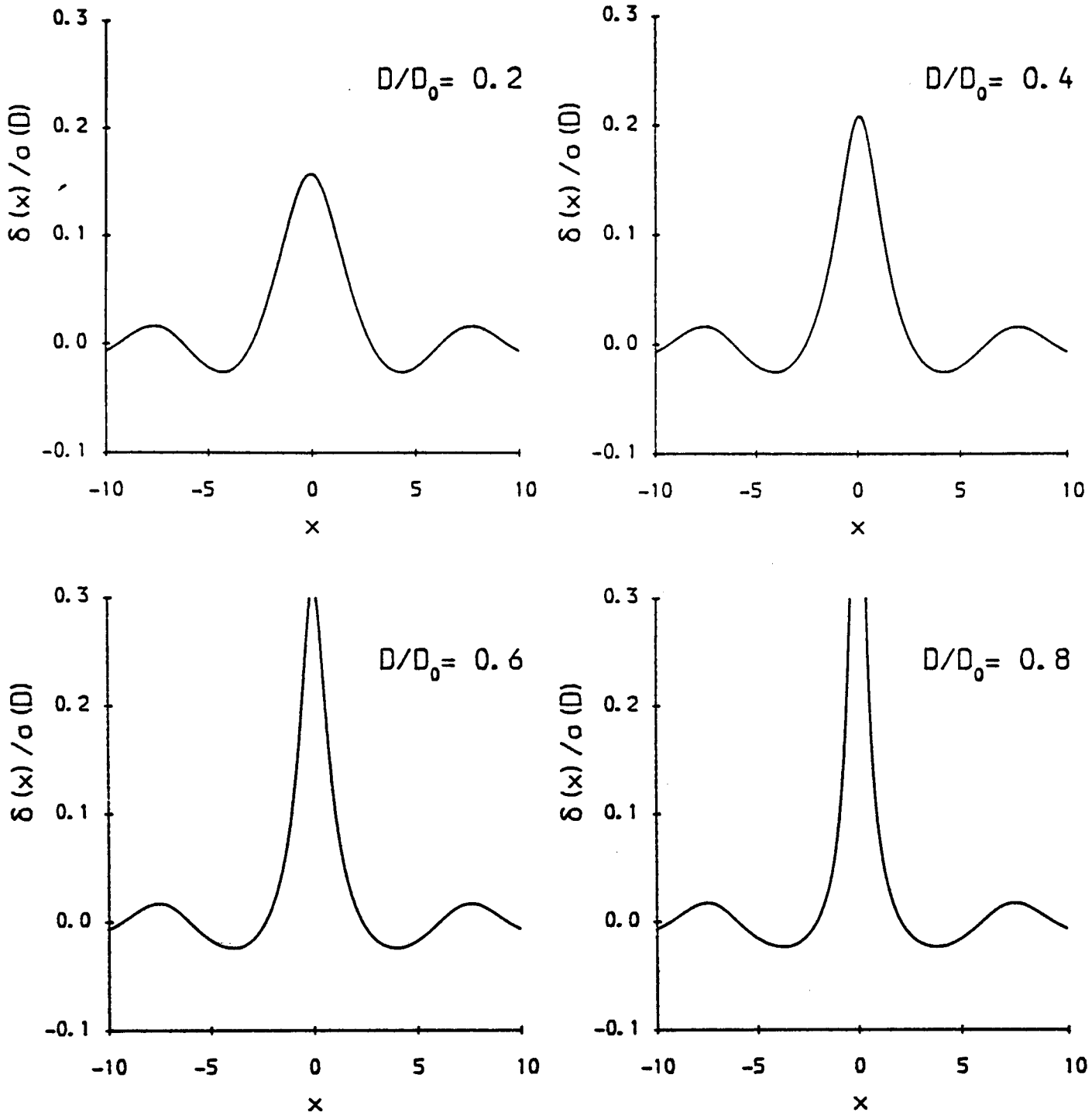


Figure 3.14. The spatial evolution of an initial perturbation $(\sin x)/x$. The linear evolution of the perturbation has been accounted for in the scaling of the y-axis. Differences between the figures represent non-linear evolution.

In reality, other effects will prevent the perturbation from forming a caustic; virialisation will stop the transfer of power to small scales. Once the perturbation has collapsed to a state in virial equilibrium, evolution will slow markedly. Power will not be transferred to scales much smaller than the virial radius.

3.4.3 Evolution of Power Law Fields

The gravitational forces between distant matter are relatively insensitive to the

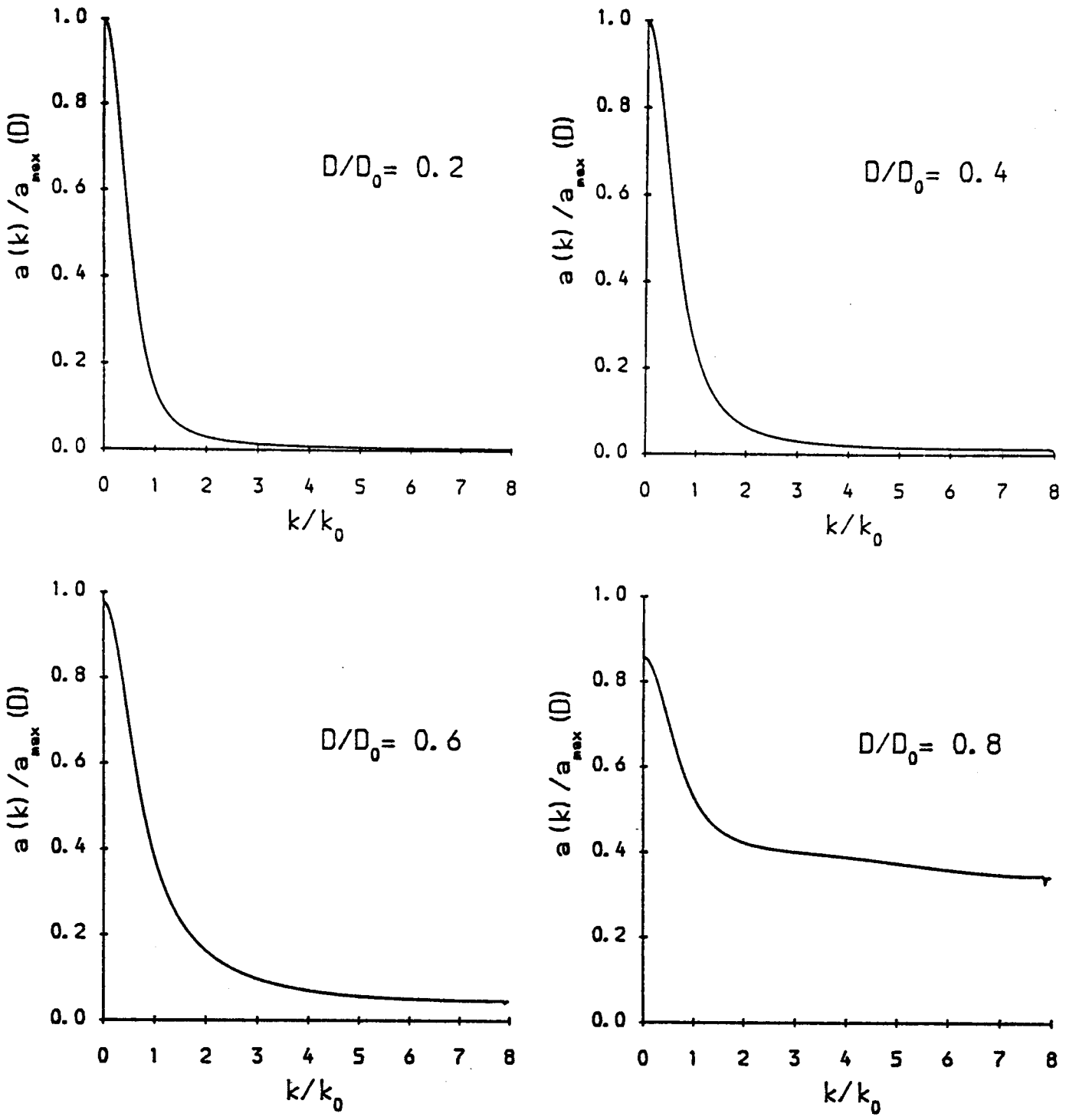


Figure 3.15. Fourier evolution of an initial perturbation $\exp(-x^2/\lambda^2)$. The linear evolution of the perturbation has been accounted for in the scaling of the y-axis. Differences between the figures represent non-linear evolution.

mass distribution on smaller scales. Perturbations of large scale will grow in blissful ignorance of the evolution on smaller scales [Peebles 1980 §28-C]. In a hierarchical scenario, large scale modes will evolve linearly when smaller scales have entered the non-linear regime ($\delta \gtrsim 1$).

Let r_o be the scale which is currently going non-linear. To be more precise, let us define r_o to be the smoothing scale at which $\langle|\delta(\mathbf{x}, \mu_o)|^2\rangle_{r_o}$ would have become one if only linear evolution had occurred. μ_o is the present time, expressed in

dimensionless form.

r_o is a characteristic scale; the power spectrum must evolve away from a power law. The processes discussed above will have transferred power on scales close to r_o to smaller scales. It will be argued in the next chapter that virialisation orchestrates the small-scale power to produce the observed galaxy-galaxy correlation function. Be that as it may, non-linearities will impose a characteristic scale for structure in the universe.

3.5 Cosmological Implications

Several implications of cosmological interest can be inferred from the work presented in this chapter. Since the majority of the calculations were performed for only one-dimensional systems, some care must be taken in applying the results to the three-dimensional problem. Whilst the general form of the conclusions presented in §3.3.8 may be expected to be the same in three dimensions, the details may differ. This is true particularly with regard to the phrase 'close to -1'. In d dimensions, this will become 'close to $-d$'. One question remains: how close is close? Biasing in one dimension, for example, becomes increasingly important as n decreases from 0 to -1. It is not obvious whether 'close' means $n < 0$ or $n \sim -d$. The covariance function in three dimensions (see §4.2) sheds some light, but work in three dimensions needs to be done to resolve the issue. It is important to keep this thought in mind for the remainder of the chapter.

3.5.1 Range of Collapsed Scales

From equation (3.1.11), the mean square fluctuation of a three dimensional perturbation field with spectral index n, when smoothed over the scale r, can

be written in the form

$$\langle|\delta(\mathbf{x},\mu_o)|^2\rangle_r = \left(\frac{r}{r_o}\right)^{-n-3} \tag{3.5.1}$$

r_o is defined above (§3.4.3). If $n > -3$, larger scales should still be in the linear regime.

We wish to use this result to examine the range of scales over which significant structure has developed. Let us first establish a standard by which we may judge whether or not structure will have formed on a scale r, larger than r_o. A reasonable criterion is that the peaks in the field, once smoothed over the scale r, should have δ in excess of 1. If the highest peaks we are willing to consider have amplitudes ν times the r.m.s. amplitude, structures will have formed on a scale r if

$$\nu^2\langle|\delta(\mathbf{x},\mu_o)|^2\rangle_r > 1. \tag{3.5.2}$$

This implies that structures will have formed on scales

$$r < r_o\nu^{2/(n+3)}. \tag{3.5.3}$$

This inequality brings out an important point. As discussed above, non-linearities will impose a characteristic scale, r_o. The inequality (3.5.3) shows that structures will only form on scales close to r_o, unless n is close to -3.

We can use the inequality to get some idea of the value of n. The spatial distribution of bright galaxies suggests that $r_o \sim 8h^{-1}$ Mpc [Peebles 1982]. If galaxy formation is biased, r_o may be even smaller. There are indications from the super-clusters correlation function (§1.2.1) that structures may be present at scales of up to $100h^{-1}$ Mpc. Let us take this as a lower limit for the scale of the largest structures and ν to be 3. Using (3.5.3) yields

$$n < -2.1. \tag{3.5.4}$$

Although this constraint is very rough, it does highlight the point that we inhabit a universe with structure on a wide range of scales. To explain this in a hierarchical scenario, a spectral index of close to -3 is needed.

3.5.2 Galaxy Clustering

In common with recent work [Peac ock and Heavens 1985; Bardeen *et al.* 1986], let us assume that galaxies formed at the peaks of some Gaussian random field. Since the initial positions of the peaks were uncorrelated (§3.3.8), the galaxies too should be spatially uncorrelated. With the above assumption, the galaxy-galaxy correlation function should have been negligible over all scales at early times. We will return to this point in relation to quasar absorption lines (§6.3.4). Correlation functions and biasing will be discussed in the following chapter.

Since the galaxies are presently strongly clustered (§1.2.1), non-linear evolution must be the cause if this model is correct (see also §4.4.3). We can examine the tendency of peaks to approach each other using the expression for the peculiar velocity field, equation (2.1.29). In one dimension this takes the form

$$u(q) = D'(\mu)\frac{d\Psi}{dq}. \tag{3.5.5}$$

The velocity difference between peaks at q_1 and q_2 is given by

$$u(q_1) - u(q_2) = D'(\mu)\int_{q_1}^{q_2}\frac{d^2\Psi}{dq^2}\,dq \;\propto\; \int_{q_1}^{q_2} -\Delta(q)\,dq. \tag{3.5.6}$$

Equation (2.1.27) was used to derive the last expression.

This result (unsurprisingly) shows that two peaks will approach each other if the integrated overdensity between them is positive. Moreover, the larger the integrated overdensity, the faster they will come together. Again unsurprisingly, this shows that peaks in a region with an overdensity above zero will become more clustered. *We may expect stronger clustering if n is close to -3.*

3.5.3 Galaxy Clusters and Voids

Clusters represent very overdense regions of the universe; of the order of one hundred galaxies in a volume more typically occupied by one. It is interesting to ask how they became so overdense, or alternatively, why their binding energies are so large.

To put this in perspective, consider a gravitating system with an initial nett energy of $-E$. Let the kinetic energy of the system be K, the potential energy U. After virialisation, K and U must obey

$$2K + U = 0. \tag{3.5.7}$$

This implies that $K = E$, $U = -2E$ if energy is conserved; since a cluster has a presently large gravitational potential, it must always have had a large nett binding energy. Admittedly this argument has assumed that the cluster is isolated. Two-body relaxation processes would have increased the binding energy. It is difficult to believe that such effects can be the sole cause of the observed large overdensities. As for the considerations in the previous section, this suggests that clusters formed from extended regions with significant overdensity—a yet futher indication that n is close to -3.

We have previously concluded tht if galaxies formed at peaks in a Gaussian random field, they will be distributed more or less uniformly throughout the universe. Let us take a typical cluster to have ~ 125 members and the mean separation of galaxies in the universe to be $\sim 5\,\mathrm{Mpc}$. With these values, material in a cluster must initially have come from a region $\sim 25\,\mathrm{Mpc}$ in diameter.[2]

A very naïve picture of a cluster is that it contains all the material from a region with an overdensity initially above zero. What about extended regions

[2] Just spreading out the material in a cluster at the average density in the universe would give much the same answer.

with a negative overdensity? By the symmetry of the Gaussian random field, they should have statistically the same sizes as the regions of positive overdensity. Voids with diameters of tens of Mpc arise naturally in this picture. They are the negative density counterparts of clusters.

3.5.4 Large-Scale Streaming Motions

Extended regions with a coherent velocity field are a natural consequence of initial power spectra with spectral indices of < -1. However, in order for the expression for the velocity field to converge, the effective spectral index,

$$n_{eff} = \frac{d\log P(k)}{d\log k}, \tag{3.5.8}$$

must become > -1 on large scales. In the C.D.M. spectrum for example, n_{eff} becomes > -1 on scales in excess of 10 Mpc [taken from Couchman 1987, Figure 6].

In one dimension, the velocity field is exactly analagous to the overdensity field, but with a power spectrum of k^{n-2} (§3.1.2). Extended regions with the same velocity sense will occur if n is close to $+1$. Extending this result to three dimensions, extended regions with a coherent velocity field will also occur if n is greater than -1, but close to it. The bulk motions discussed in §1.1 are not totally inexplicable.

3.5.5 Summary

The above implications are strongly suggestive if we are to believe in a hierarchical scenario with initial perturbations described by a Gaussian random field. Non-linearities should have imposed a characteristic scale; power on smaller scales would be controlled by non-linear effects, including virialisation. Since

no such scale is apparent, n_{eff} should be close to -3 over the range of scales for which structures have been detected. Other considerations are more circumstantial but indicate the same result. Whether other observations, the low-levels of micro-wave background anisotropy in particular, will rule out this model is a subject for future consideration.

Chapter 4. Correlation Functions

This chapter is devoted to the study of the two-point correlation function. As discussed in §1.2.1, this function has been used extensively to characterise the spatial distribution of galaxies, clusters and super-clusters. Several considerations related to their interpretation are presented below. We shall begin by distinguishing between the correlation function for a continuous field and that for a point distribution.

4.1 Covariance Functions

We have already defined the two-point correlation function for a continuous field and a spatial distribution of discrete objects. In common with previous work [e.g., Peebles 1980 Chapter 3], no distinction between these cases has yet been made. There are, however, important differences. The relationship between them is far from clear. For the remainder of this chapter, the correlation function for a continuous field will be called the *covariance function*. It will be denoted by the symbol Υ.

4.1.1 The Covariance Function for Power Law Fields

As in the previous chapter, we will take power spectra to be of the form

$$P(k) \sim k^n \, W(k/k_c), \tag{4.1.1}$$

where W is a window function. The same reasons for employing a cut-off apply. Although we will use the Gaussian filtering function throughout, it is simple to understand the effects of altering the form of the cut-off. We discussed in §3.1.1 the relationship betwen filtering in Fourier space and convolving in real space. The convolving function, the Fourier transform of $W^{1/2}$, may have features such

as side lobes. This will introduce a correlation at the position of the feature, purely as a result of the form chosen for W.

Using the Wiener-Khintchine theorem, the covariance function can be written in the form

$$\Upsilon(\mathbf{x}) \sim \int_{-\infty}^{\infty} k^n \, W(k/k_c) \mathrm{e}^{i\mathbf{k}\cdot\mathbf{x}} \, d^d k. \tag{4.1.2}$$

d is the dimensionality of the system. The fields we are studying are isotropic; no information is lost by averaging the covariance function over the direction of $\mathbf{x}$. In one dimension,

$$\Upsilon(x) \sim \int_{0}^{\infty} k^n \, W(k/k_c) \cos kx \, dk, \tag{4.1.3}$$

in three dimensions,

$$\Upsilon(x) \sim \int_{0}^{\infty} k^{n+2} \, W(k/k_c) \frac{\sin kx}{kx} dk. \tag{4.1.4}$$

The covariance function can be conveniently normalised by setting the r.m.s. fluctuation of the field to one. With this choice, $\Upsilon(0) = 1$.

It will be convenient to display the covariance functions for various values of the spectral power index, n, in the form of a contour plot. The lag distance, x, is plotted along the x-axis, n along the y-axis. In the two plots shown, the contour heights start at zero and work up towards one in six equal steps. The last contour in the series, $\Upsilon(x) = 1$, lies along the x-axis because of the normalisation discussed above.

4.1.2 One-Dimensional Fields

Figure 4.1 shows how the covariance function depends on the spectral index for a one-dimensional system. We can understand this figure by considering the Fourier makeup of the function $\Upsilon(x)$ for $n = 0$. This is one of a series of

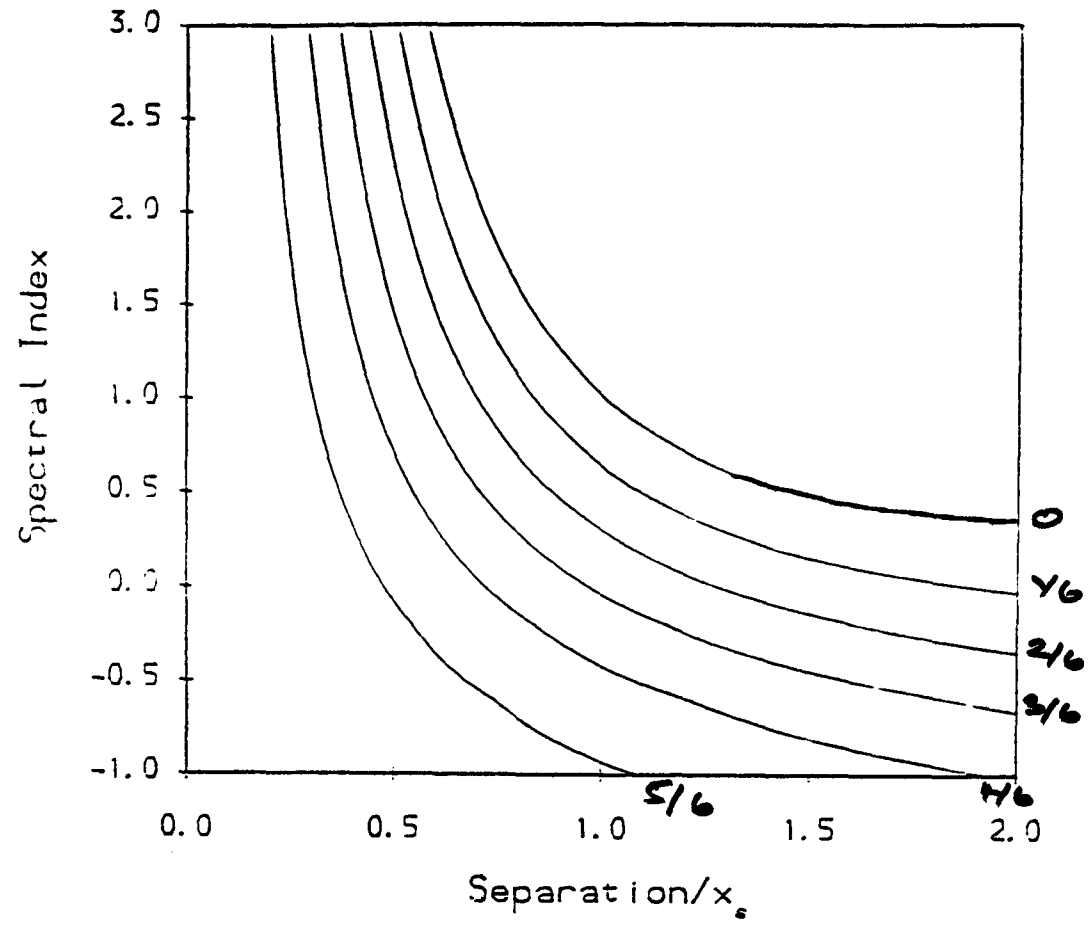


Figure 4.1. Contour plot showing the dependence of the covariance function on the spectral index for a one-dimensional system.

functions which tend towards a δ-function as k_c tends to infinity [e.g., Lighthill 1958]. For this function, all the Fourier components are in phase at the origin; constructive interference causes a maximum amplitude. The components continue to constructively interfere out to a lag $\sim x_s/2$, where x_s is the smoothing length, defined in §3.1.1. At this distance, the contributions from the highest frequency components have begun to change sign. Destructive interference causes the amplitude of the covariance function to decrease. It has little or no amplitude at higher separations. As for white light interference fringes, the function is 'washed out' away from the origin.

The structure of the covariance function for other values of n is similar. If there is more small scale power ($n > 0$), the function will be washed out for smaller lag separations. Conversely, as n decreases, correlation will occur for larger separations.

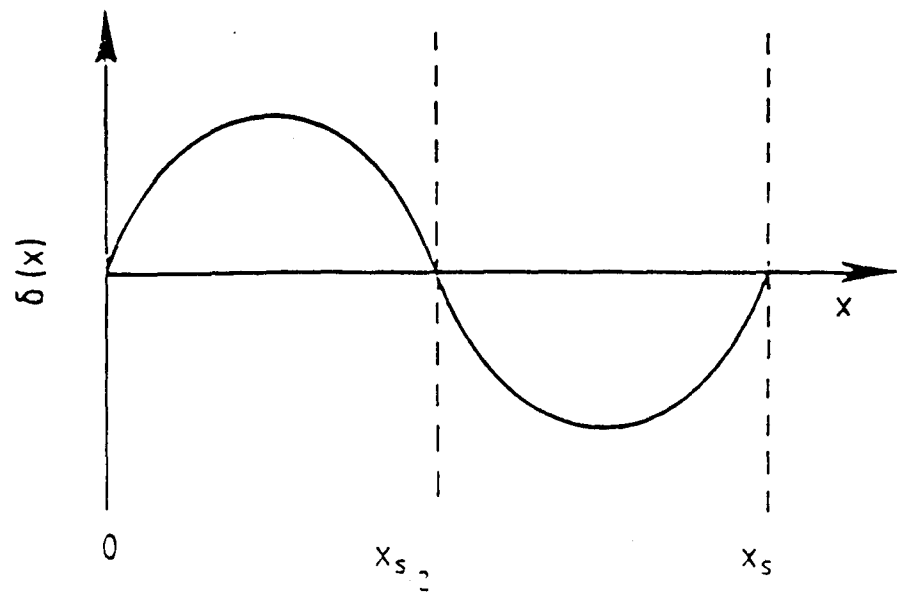


Figure 4.2. Schematic diagram showing that the field only has the same sense within an individual maximum or minimum. Correlation should not extend much futher than $x_s/2$.

It is also possible to understand the figure by considering the field in real space. In chapter 3, we saw that the majority of the peaks are positive, minima negative, unless n is close to -1 (see Figures 3.2a–f; Figure 3.10 implies the same). We also concluded that maxima, and hence minima by the symmetry of the Gaussian random field, tend to occur at a separation of x_s. Together, these results imply that the field has the same sense (positive or negative) only within an individual maximum or minimum—see Figure 4.2. The major contribution to the covariance function comes from correlation within individual extrema. Correlation only occurs to a lag $\sim x_s/2$. There is little correlation between separate extrema.

The situation changes as n approaches -1. Extended regions with the same sign begin to occur (see Figures 3.2a–f and 3.5). Separate extrema are correlated; the covariance function extends to larger separations. These considerations are reinforced by our study of biasing in the previous chapter.

4.1.3 Three-Dimensional Fields

Figure 4.3 is a contour plot of the covariance function against n for three-

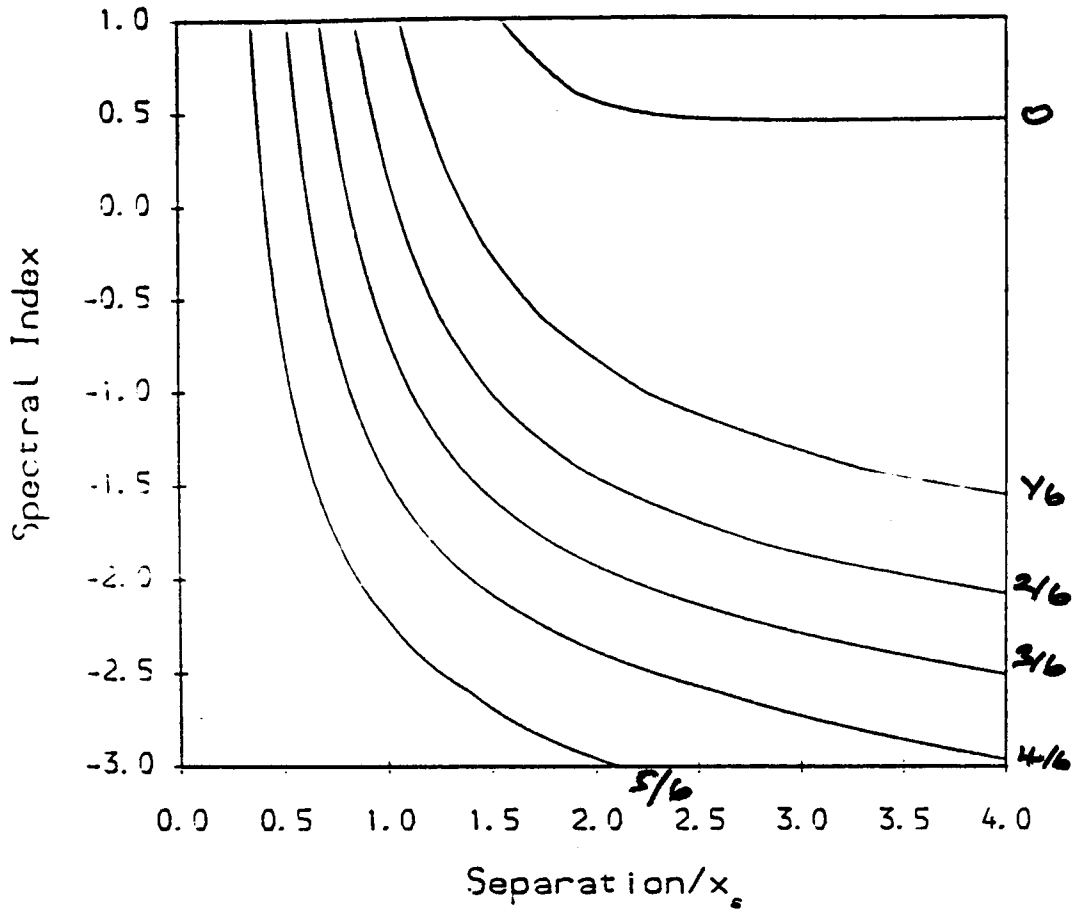


Figure 4.3. Contour plot showing the dependence of the covariance function on the spectral index for a three-dimensional system.

dimensional fields. We can understand this figure in the same way that we understood the one-dimensional case, although with two minor differences. Equation (4.1.4) shows that k^n becomes k^{n+2} in three dimensions. The other difference is related to using $(\sin kx)/kx$ as opposed to $\cos kx$. This in turn is the result of directional averaging in a space of different dimensionality. The former function is positive over a length twice that for $\cos kx$. Correlation extends roughly twice as far for the three-dimensional case.

It is reasonable to assume that correlation for three-dimensional fields arises in the way previously discussed for one-dimensional systems. Since correlation extends twice as far, extrema must be somewhat more extended. Their separation is greater in a three-dimensional field for the same cut-off. We may suppose that their separation is roughly x_s per dimension. The total separation is thus $(1^2 + 1^2 + 1^2)^{1/2} x_s \sim 2x_s$. The truth or otherwise of this supposition will be determined in future work on three-dimensional fields.

The considerations above suggest an interesting conclusion. Correlation only occurs over lag separations for which the Fourier components interfere constructively. In other words, *there is only correlation on scales where there is NO significant power.* Unless there is a marked feature in the power spectrum, this result should hold whatever its form.

4.2 Biasing (for the last time!)

In the last chapter, we were unable to decide whether or not biasing would be important for $n < 0$ or only $n \sim -d$. With the shift in the y-axis and the rescaling of the x-axis, Figure 4.3 is remarkably similar to Figure 4.1. The covariance function for three-dimensional fields has little amplitude at separations $\gtrsim x_s$, unless n is close to -3. This similarity strongly suggests that biasing is only important for $n \sim -d$.

Whatever the outcome of the three-dimensional study, we have yet to explain why biasing works. Kaiser's 1984 calculation, for example, discussed in §1.5.1, shows that imposing a threshold will cause the correlation function to be enhanced whatever the power spectrum. Let us extend the arguments of the previous section. Correlation in regions with field amplitudes above a threshold will almost entirely occur within the tips of individual peaks (Figure 4.4). These regions are manifestly highly correlated, explaining the amplification Kaiser found. Coles [1986] suggested that correlation within these areas would make a significant contibution to the enhancement. Our study can push this result a little futher: unless n is close to -3, they are the entire cause.

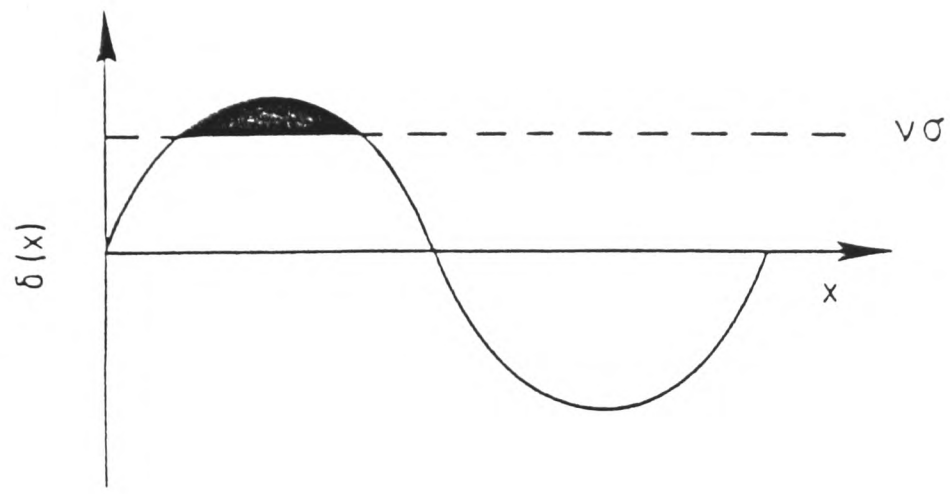


Figure 4.4. If a threshold is imposed, correlation will only occur within the tips of individual peaks. These regions are manifestly highly correlated.

4.3 Correlations in the Cosmological Density Field

It is frequently assumed that galaxies trace the cosmological matter distribution. On one hand this assumption is motivated by simplicity, on the other, ignorance. Other assumptions are somewhat *ad hoc*. They generally add a free parameter, which of course can only aid when matching theory to observation. An obvious example is the use of thresholding in N-body simulations. Galaxies are taken to form only in regions where the density is greater than some threshold value. The value of the threshold level is clearly an extra parameter.

The tracing assumption does have its own drawbacks. If galaxies trace the matter distribution, why should not clusters, or indeed super-clusters? Since the correlation lengths are different for each of these, Kaiser argued that at most one could be a fair tracer of the mass distribution; biasing was born.

4.3.1 Smoothing Volumes

Kaiser's argument is not the only possible interpretation of the differences in correlation lengths. Consider the definition of density at a point x. If $M(V)$

is the total mass within a volume V, centred on $\mathbf{x}$, then the density, $\rho(\mathbf{x})$, is defined by:

$$\rho(\mathbf{x}) = \frac{M(V)}{V}. \tag{4.3.1}$$

This simple definition shows that by itself, density is meaningless. V, the averaging or smoothing volume, also needs to be specified. In general, changing V will alter the density.

The same is true if we consider a density field. By altering the smoothing volume a different, albeit related, density field is created. The effects of varying the cut-off frequency for power law fields can be thought of in this light. Different cut-off frequencies correspond to different averaging volumes.

4.3.2 Galaxies etc. as Probes

Exactly how the cosmological density field is related to the distribution of galaxies or clusters is far from obvious. However, it is clear that by using clusters as probes of the density field, as opposed to galaxies, a much larger smoothing volume is implied. The density fields probed by galaxies, clusters and super-clusters are *not* the same. To distinguish between the correlation functions for the different objects, ξ will be appropriately subscipted. For example, the cluster-cluster correlation function will be denoted 'ξ_{cc}'.

To progress futher we need to make some assumption about the effects of using different sets of objects as probes of the density field. The simplest possible assumption is that the smoothing length, $L = V^{1/3}$, for a particular set of objects is proportional to their mean separation, $\bar{x}$.

In three dimensions, the covariance function is given by equation (4.1.4). Our assumption allows k_c to be written as

$$k_c = \frac{\alpha}{\bar{x}}, \tag{4.3.2}$$

where α is a constant of proportionality. Making our habitual transformation $s = k/k_c = k\bar{x}/\alpha$, (4.1.4) becomes

$$\Upsilon(x) \sim \left(\frac{\alpha}{\bar{x}}\right)^{n+3} \int_0^\infty s^{n+2}\, W(s) \frac{\sin(\alpha s\, x/\bar{x})}{\alpha s\, x/\bar{x}}\, ds. \tag{4.3.3}$$

Writing the integral as $\Gamma(x/\bar{x}, \alpha; n)$, the covariance function takes the form:

$$\Upsilon(x) \sim \bar{x}^{-n-3}\Gamma(x/\bar{x}, \alpha; n). \tag{4.3.4}$$

Once the length is written as the dimensionless combination $x/\bar{x}$, the covariance function takes the same form whatever $\bar{x}$. Only the amplitude of the function depends on the mean separation of the objects.

4.3.3 The Universal Correlation Function

Bahcall and Burgett [1986] recently examined the correlation functions for various sets of objects. They found that the functions do take the form suggested by (4.3.4). They also discovered that the amplitude of the function was independent of the probe used, providing that the lag was written in the above dimensionless combination. Bahcall and Burgett proposed a universal correlation function.

If it were really felt that the present power spectrum is of the form k^n, then the amplitude of the covariance function can only be independent of $\bar{x}$ if $n = -3$.[1] However, as we discussed in §3.4.3, even if the power spectrum was initially of that form it would not be so now. Non-linearities would in

[1] In fact, the dependence will become logarithmic.

general cause the power spectrum to evolve away from a power law form. This is particularly relevant since it will be argued later in the chapter that the observed correlations are mainly due to objects in non-linear regions.

All is not lost however. Consider the evolution of an initially power law field. For other than a spectral index of -3, scales will enter the non-linear regime at different times. The scale invariance would be broken. If $n = -3$. however, every scale must enter the non-linear regime at the same time. At any chosen time the evolved field must still be scale invariant. Let us assume that there are scales large enough for gravity to be the only significant force. In this situation, structure on one scale must be reproduced on any other. Moreover, the distribution of the structures must also be scale invariant. The correlation functions for these structures will naturally be 'universal'.

There is a strong case to be made for taking the initial density perturbation to be an $n = -3$ power law field. Together with the arguments from the previous chapter, there is a strong case indeed.

4.4 Observed Correlation Functions

Although the averaging argument makes the existence of a universal correlation function plausible, there is nothing to suggest its form. In the language of the previous section, this depends on the window function, W. Unfortunately, it is difficult or impossible to make an *a priori* calculation of this function. Such a calculation would depend on the details of structure formation and probably on their current distribution. Deeming [1976], for example, shows that the discrete Fourier transform of unequally spaced data depends significantly on the distribution of data points. It is inevitable that the form of W must depend on the current spacing of the structures.

An alternative approach is to account for the form of the galaxy-galaxy correlation function. If the arguments for an initially $n = -3$ power law field are to be believed, the other correlation functions will follow because of scale invariance. We shall begin by establishing that correlation functions are dominated by overdense regions.

4.4.1 Non-linear Domination

Let us consider the structure of the covariance function. The contribution towards correlation from a region is approximately proportional to δ^2. δ is the dimensionless density contrast, defined in Chapter 2 as $(\rho - \bar{\rho})/\bar{\rho}$. Here ρ is the mean density in the region, $\bar{\rho}$ the mean density of the entire field. This scaling shows that the importance of the contribution increases rapidly with overdensity. A small amount of material at a high density may have more effect than a much larger amount at a more average density. In a cluster, for example, the galaxy density is typically a few hundred times the average galaxy density. This implies that the small fraction of galaxies in clusters, maybe only 5 or 10%, can significantly alter the galaxy-galaxy correlation function. Conversely, areas of underdensity can only have a maximum $|\delta|$ of 1. Any contribution from galaxies in these regions is totally swamped by the contribution from cluster galaxies.

There is also more direct observational evidence that correlation functions are dominated by the contribution from non-linear regions. That is, regions with δ in excess of one. The galaxy-galaxy correlation function in redshift space is strongly elongated along the velocity separation axis [Davis and Peebles 1983; Bean *et al.* 1983]. This contrasts strongly with the linear calculation; compression should occur in this direction (see §5.5.5). We can conclude that the velocity correlation function for galaxies is dominated by the contributions from galaxies in non-linear regions. The spatial correlation function is merely an ap-

propriately averaged version of the velocity correlation function. It too must be dominated by the contribution from galaxies in overdense regions.[2]

The cluster-cluster velocity correlation function is similarly elongated along the velocity direction [Bahcall, Burgett and Soneira 1986]. We reject Bahcall *et al.*'s suggestion that this elongation is evidence for spatial elongation along the line of sight for two reasons. So that we would not live in a privileged observing position, Bahcall *et al.* proposed that our galaxy should lie in the edge of a yet undiscovered 'shell' of clusters. This idea requires that most clusters should lie in only *one* surface. Bahcall *et al.* specifically refer to elongation in three super-clusters. This implies that there are at least three separate cluster surfaces. Do we observe from the point of intersection of these surfaces? If so then we must be at the vertex of a shell—an undoubtedly privileged position.

The second point is related to the existence of the universal correlation function. Consider the process which arranged galaxies, clusters and super-clusters in this fashion. It is not hard to believe that it could also have produced similar (if not universal) velocity correlation functions. What is hard to imagine is that the observed elongation for galaxies is due to peculiar velocities, whilst the elongation for clusters reflects an elongated spatial distribution. We can reasonably conclude that the cluster-cluster correlation function is also dominated by non-linear regions.

4.4.2 Cluster Model

Correlation functions are dominated, or at least greatly influenced, by high density regions. With this realisation, we are in a position to estimate the

[2] For the rest of this chapter, there will be many statements such as 'the correlation function is dominated by galaxies in clusters'. This is merely a shorthand way of referring to the contribution towards the correlation function made by these galaxies. This short form is considerably less clumsy. Moreover, meanings are made clearer.

form of the galaxy-galaxy correlation function, ξ_{gg}. The function will reflect the correlation within an 'average' cluster. Clusters are the regions with the highest galaxy density.

Let us build a very simple model. We will assume that galaxies are only correlated within clusters, which are themselves distributed at random. This last assumption is manifestly not true but the added correlation is probably unimportant. Clusters have typical radii of a few Mpc. Since galaxies are only correlated over a similar distance, correlation between galaxies in separate clusters cannot make much of a contribution. They must be futher apart than a typical cluster radius.

Similar models were at the heart of the program proposed by Neyman and Scott [1952]. Their models were slightly more sophisticated in that they allowed the clusters to be correlated too. Unfortunately, this did not help a great deal. They showed that for any chosen cluster profile, a distribution of cluster centres can be found which makes the model fit the observations.

Neyman and Scott tried to adjust their model parameters to fit the observed clustering pattern, an *a posteriori* approach. The method here is slightly different. Using a standard model for the cluster density profile, we will make an *a priori* calculation of the correlation due to clusters. There are no free parameters in the model, only theoretical and observational uncertainties. Within these limits we can confirm or reject the model.

The correlation function we are trying to calculate is of the directionally averaged variety. There is little or no point in considering other than spherically symmetric cluster models. We will also consider every cluster to be identical. We will specify 'the' cluster profile in terms of a dimensionless profile function, K, a characteristic galaxy density, ρ_o, and a characteristic scale, r_o. The other

parameters we shall need are the average galaxy density for the universe as a whole, $\bar{\rho}$, and the mean cluster separation, R_o.

Taking the origin to be at a cluster centre, the density at a point $\mathbf{r}$ within the cluster can be written as

$$\rho(\mathbf{r}) = \rho_o K(\mathbf{r}/r_o). \tag{4.4.1}$$

K is the dimensionless cluster profile. The corresponding dimensionless density contrast is given by

$$\delta(\mathbf{r}) = \frac{\rho_o K(\mathbf{r}/r_o) - \bar{\rho}}{\bar{\rho}}. \tag{4.4.2}$$

Since the galaxy density in a cluster is much greater than the mean, the $\bar{\rho}$ term in the numerator can be ignored.

Each cluster occupies a mean volume of R_o^3. Since the correlation function is a spatial average, the predicted galaxy-galaxy correlation function[3] takes the form

$$\xi_{gg}(\mathbf{r}) = \frac{1}{R_o^3} \int \delta(\mathbf{r}_1)\delta(\mathbf{r}+\mathbf{r}_1)\, d^3 r_1. \tag{4.4.3}$$

The integral is over all space. Making the substitutions $\mathbf{u}_1 = \mathbf{r}_1/r_o$, $\mathbf{u} = \mathbf{r}/r_o$ and using equations (4.4.1) and (4.4.2), this expression becomes

$$\xi_{gg}(\mathbf{u}) = \left(\frac{r_o}{R_o}\right)^3 \left(\frac{\rho_o}{\bar{\rho}}\right)^2 \int K(\mathbf{u}_1)K(\mathbf{u}+\mathbf{u}_1)\, d^3 u_1. \tag{4.4.4}$$

For convenience, let us write the integral as $\Lambda(\mathbf{u})$.

We are now in a position to average over the direction of $\mathbf{u}$. $\xi_{gg}(\mathbf{u})$ and $\Lambda(\mathbf{u})$ become functions of $u = |\mathbf{u}|$ only. Since the cluster models are spherically

---

[3] Even though we are calculating a covariance function, we are attempting to predict the galaxy-galaxy correlation function. It is appropriate to retain the name 'correlation function'.

symmetric, $K(\mathbf{u})$ is already a function of u only. It is straightforward to show that

$$\Lambda(u) = 4\pi \int_0^\infty k^2 \, |\widetilde{K}(k)|^2 \frac{\sin ku}{ku} \, dk \tag{4.4.5}$$

where

$$\widetilde{K}(k) = \left(\frac{2}{\pi}\right)^{1/2} \int_0^\infty u^2 \, K(u) \frac{\sin ku}{ku} \, du \tag{4.4.6}$$

Transforming back to the real space co-ordinate, $r = r_o u$, our final expression for the predicted galaxy-galaxy correlation function takes the form

$$\xi_{gg}(r) = \left(\frac{r_o}{R_o}\right)^3 \left(\frac{\rho_o}{\bar{\rho}}\right)^2 \Lambda(r/r_o). \tag{4.4.7}$$

To apply this model, we must establish a suitable form for the function K and values for the various parameters. Observational studies of cluster profiles could in principle be used for this purpose. In practice, however, this approach is far from trivial. It is not obvious which cluster, if any, is an 'average' cluster. Moreover, there is inherent difficulty in evaluating ρ_o and $\bar{\rho}$, or more precisely, their ratio. The bright galaxies used in studies of the correlation function are known to have a separation of a few Mpc. The problem is that the number of 'bright' galaxies in a cluster is hard to estimate from published material. A detailed knowledge of the galaxy luminosity function is needed to accurately assess this ratio.

Fortunately, there is a more satisfactory approach based on a determination of the cluster-galaxy cross-correlation function, $\xi_{cg}(r)$. As Peebles notes [1980 §77-A], $\bar{\rho}\big(1 + \xi_{cg}(r)\big)$ should reproduce an 'average' cluster profile. Moreover, Peebles used the Lick catalogue of galaxies to determine the cross-correlation function, the same catalogue he used to determine $\xi_{gg}(r)$. There is no need to know the galaxy luminosity function.

Peebles found that the mean galaxy density at a distance r from a cluster centre, $\rho(r)$, could be represented by [1980 eqn. 77.1, slightly rewritten]

$$\rho(r) = \bar{\rho}\left(1 + 130(r/r_s)^{-2.5} + 70(r/r_s)^{-1.7}\right). \qquad (4.4.8)$$

$\bar{\rho}$ is again the mean galaxy density, $r_s = 1h^{-1}$ Mpc. In terms of our model we can identify $\rho_o = \bar{\rho}$, $r_o = r_s$. The last parameter we need is the mean cluster separation, R_o. We will take $R_o = 50h^{-1}$ Mpc throughout the following calculations.

Unfortunately, this expression becomes infinite as r tends to zero. Since this is obviously unphysical, we must alter the function for small r and trust that the correlation function will not be overly affected. One convenient modification is to set $\rho(r) = \rho(r_{min})$ for $r < r_{min}$. Thus, in the altered function, the density is constant within a distance r_{min} of the cluster centre.

Typical results are shown in Figures 4.5a and b. In Figure 4.5a a value of $r_{min} = 0.05h^{-1}$ Mpc was used, in Figure 4.5b, $r_{min} = 0.2h^{-1}$ Mpc. The dashed line in each figure represents the observed correlation function. Whilst the predicted correlation functions are remarkably similar to the observed function, there is an obvious dependence on the small-scale behaviour of $\rho(r)$. The adopted modification is clearly not sufficient.

A better way to treat the small-scale behaviour of $\rho(r)$ is to try to fit standard cluster profiles to the form suggested by $\xi_{cg}(r)$. We shall use two profiles for this purpose. The first generated by a King model, the second by the de Vaucouleurs $r^{1/4}$ model.

King models are specified by three parameters: the core density, ρ_c, the core radius, r_c, and the concentration index, $c = \log(r_t/r_c)$. r_t is the tidal radius of the model. Galaxies outside this radius are no longer bound to the cluster

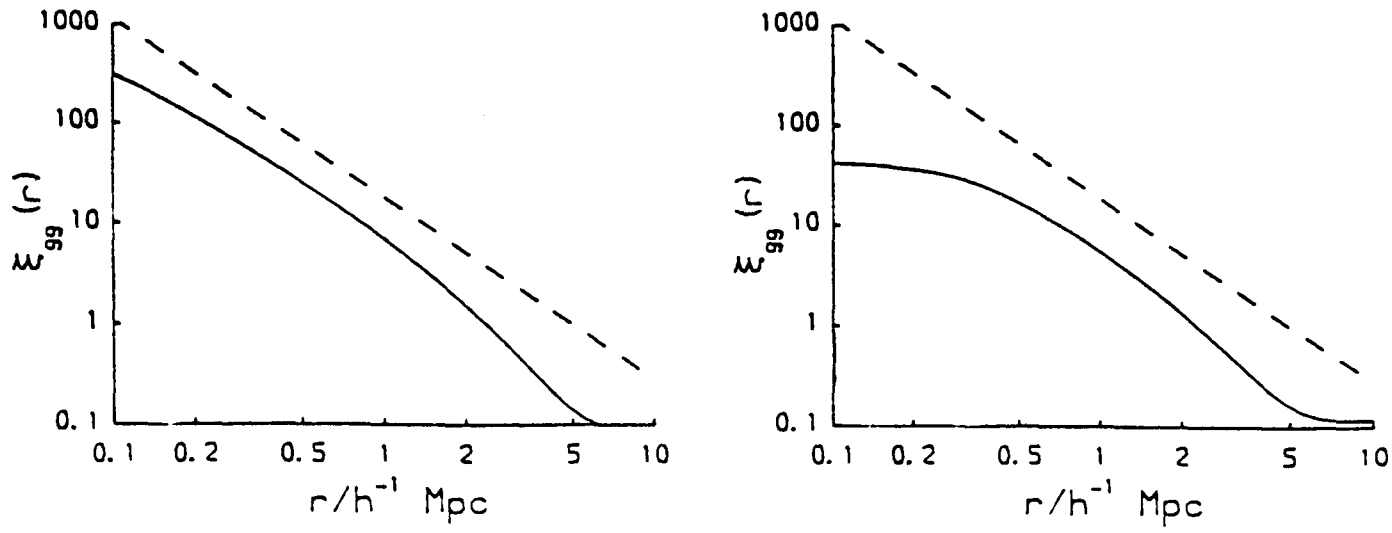


Figure 4.5. Shows the estimated galaxy-galaxy correlation functions under the assumptions discussed in the text. The dashed lines indicate the observed form. $(r/5h^{-1}\,\mathrm{Mpc})^{-1.8}$. a) $r_{\mathrm{min}} = 0.05h^{-1}$ Mpc. b) $r_{\mathrm{min}} = 0.2h^{-1}$ Mpc.

[Binney and Tremaine 1987 Chapter 4]. In terms of our model, we can identify $\rho_o = \rho_c$ and $r_o = r_c$.

The $r^{1/4}$ model is specified by only two parameters: the effective radius of the cluster, r_e, and some normalising parameter, conveniently taken to be the total number of galaxies in the cluster, N_c [Young 1976]. For the purposes of our model, we may identify $r_o = r_e$ and $\rho_o = N_c/r_e^3$.

Figure 4.6a shows the fit of a King model to $1 + \xi_{cg}(r)$. This corresponds to values of $r_c = 0.3h^{-1}$ Mpc, $\rho_c = 10000\bar{\rho}$ and $c = 2.3$. This value of the concentration index is perhaps slightly higher than may have been expected. However, the cluster-galaxy cross-correlation function indicates that the average cluster profile has a significant tail. The King model has to be relatively extended to account for this. Figure 4.6b shows a similar fit for the $r^{1/4}$ model. This corresponds to values of $r_e = 3h^{-1}$ Mpc and $\rho_o = 1000\bar{\rho}$.

Figures 4.7a and b show the predicted correlation functions, calculated from the profiles shown in Figures 4.6a and b. Again the dashed lines represent the observed correlation function. The figures show that the model of $\xi_{gg}(r)$ is

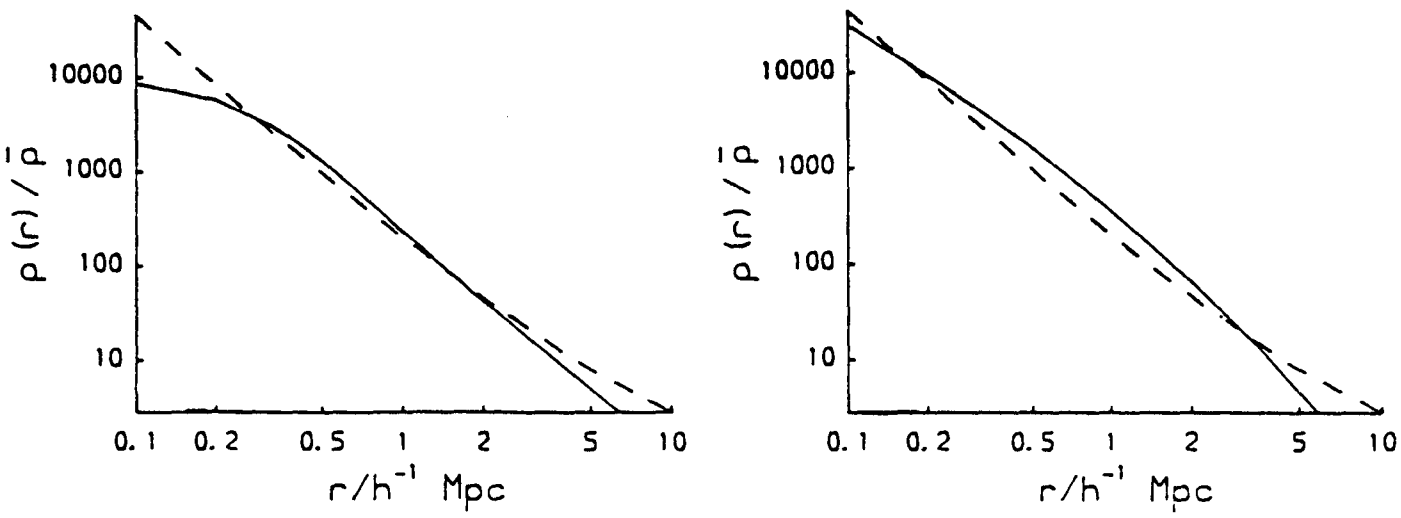


Figure 4.6. Fits of standard models to the 'average' cluster profile suggested by ξ_{cg}. a) King model. b) $r^{1/4}$ model. The parameters used to produce the fits are discussed in the text.

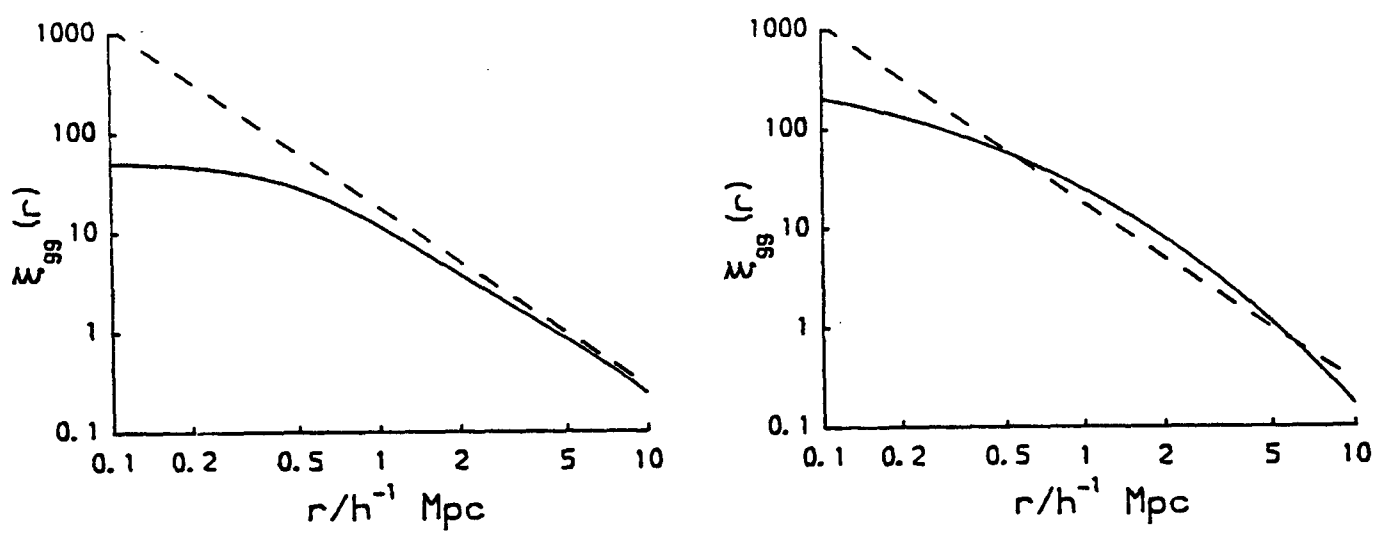


Figure 4.7. Galaxy-galaxy correlation functions predicted using the cluster profiles shown in Figure 4.6. a) Using a King model. b) Using the de Vaucouleurs $r^{1/4}$ model.

remarkably successful. Within observational and theoretical uncertainties the galaxy-galaxy correlation function can be accounted for.

4.4.3 Discussion

The work presented in this section has few conclusions which are wholely incontestable. There is little doubt, however, that the observed correlation functions are dominated by non-linear regions. This imme diately implies that there is no direct connection between the power spectrum of the initial conditions and the observed correlations. Non-linear evolution must be accounted for. The non-linear domination also rescues the concept that galaxies form at the peaks

of some initially Gaussian random field. Without this domination, as discussed in §3.5.2, galaxies in this scenario would be uncorrelated.

There is also little doubt that the correlation arising from within galaxy clusters strongly influences the form of the galaxy-galaxy correlation function. Figures 4.5 and 4.7 show that it is entirely possible that cluster galaxies completely determine $\xi_{gg}(r)$. Fortunately, the proposal can be tested relatively easily in N-body simulations. Simply removing the cluster galaxies from the calculation of ξ_{gg} will demonstrate the worth or otherwise of this suggestion.

We have already concluded that the form of the observed correlation functions is determined by non-linear evolution. If cluster galaxies dominate $\xi_{gg}(r)$ then relaxation effects will also be important. The information imprinted by the initial conditions will have been wiped out. West *et al.* [1987], for example, showed that cluster profiles are insensitive to the form of the initial perturbations. This would explain why different initial conditions develop similar correlation functions. $\xi_{gg}(r)$ would be a poor probe of the cosmological density field. It would merely reflect the structure of clusters.

Most of this discussion has been concerned with the galaxy-galaxy correlation functions. What of the cluster-cluster and super-cluster–super-cluster correlation functions, $\xi_{cc}(r)$ and $\xi_{ss}(r)$ respectively. The former at least is dominated by non-linear regions and it is reasonable to assume that the latter is too. Their form is presumably influenced by clusters of clusters and clusters of super-clusters, respectively.

If $\xi_{gg}(r)$ *is* dominated by galaxies in clusters, then there are strong implications for the other correlation functions. Their form is very similar to the galaxy-galaxy correlation function. All vary as $r^{-1.8}$. If $\xi_{gg}(r)$ takes this form because of virialisation, then $\xi_{cc}(r)$ and $\xi_{ss}(r)$ probably take this form for the

same reason. The clusters of clusters and clusters of super-clusters would also have to be virialised. Collapse on an extremely large scale ($> 100h^{-1}$ Mpc) would be implied. Again an argument for an $n = -3$ power law field appears. This of course, is no more than the smoothing volume argument discussed earlier.

These ideas are somewhat unconventional. However, they do follow naturally from the realisation that the galaxy-galaxy correlation function is strongly influenced by cluster galaxies. It is important to establish how strongly. If $\xi_{gg}(r)$ is entirely determined by the structure of clusters, then, as we have seen, there are many far reaching implications.

4.5 The Importance of Being in Phase

One of the important features of the Wiener-Khintchine theorem, discussed in Chapter 2, is that the correlation function depends only upon the amplitudes of the Fourier coefficients. Amplitude information is retained in full, whilst the phase information is totally discarded. We have already seen that the positions of peaks in a Gaussian random field are mainly dependent on the phases of the Fourier coefficients near the cut-off. This implies that phases contain valuable information. A similar implication can be drawn from studies of image transmission via Fourier transforms.

Engineers have made strenuous efforts to reduce the amount of information necessary to restore an image. Of the many methods employed, two are of interest to us. In the first, only amplitude information is transmitted. The restoring device supplies random phases for the Fourier components. In the second, only phase information is used. The restoring device draws amplitudes from a suitable distribution.

Perhaps unsurprisingly, the first method produces totally unrecognisable images. The second produces pictures which are clearly a reproduction of the transmitted image [M. Rose, private communication]. This suggests that there is more information about patterns in phases than in amplitudes. It also explains why the correlation function has difficulty distinguishing between two galaxy distributions which are visually quite different. The appearance is more determined by the phase information that the correlation function discards.

The most important implication is that a good complimentary statistic to the two-point correlation function should be based on phase information. Unfortunately, the author has yet to think of a suitable statistic. Until someone does, the visual appearance will remain highly valuable.

4.6 Conclusions

Several conclusions arise out of the work presented in this chapter.

(i) There is no correlation on scales which have significant power.

(ii) The heights of neighbouring peaks are unlikely to be correlated unless $n \sim -3$.

(iii) Galaxies, clusters and super-clusters probe different, if related, density fields. There is no reason for their correlation functions to be the same.

(iv) A universal correlation function arises naturally if the initial density perturbation was an $n = -3$ power law field. Together with the arguments presented in chapter 3, there is a strong case for considering initial conditions of this form.

(v) The observed correlation functions are strongly influenced by non-linear regions. $\xi_{gg}(r)$ probabbly reflects nothing more than the 'average' cluster profile. This proposal has important implications and should be tested in N-body simulations.

(vi) Phase information is more important than amplitude information for distinguishing between patterns.

Chapter 5. The Redshift Projection

The position of an object on the sky provides two of the three co-ordinates necessary to specify its location completely. Because of Hubble's law, a convenient third co-ordinate is its recession velocity. These define the the position of the object in *redshift* space. As we shall see, this is not quite its position in *real* space. The purpose of the chapter is to study some of the effects of using velocity as a distance measure.

5.1 The Mathematics of the Redshift Projection

Every cosmological object has six kinematic degrees of freedom: three in position and three in velocity. It is thus possible to produce a complete kinematical description of the universe in terms of a six-dimensional phase space. Let us denote this phase space Γ.

Unfortunately, it is not possible to measure all six kinematic quantities. There are restrictions on the data which may be obtained. It is improbable, for example, that the transverse motions of galaxies will ever be determined. It is also difficult to measure the distance to a galaxy. We are therefore left with only three easily observable kinematic quantities: the two angular co-ordinates of an object on the sky and its redshift. These can be used to define the position of the object in *redshift space*. This space will be denoted R. Observations are almost entirely restricted to redshift space.

Calculations, on the other hand, are free from this restriction. It is possible to know all six kinematic quantities in theoretical models of the universe. Many results are presented in real space, X. The two spaces X and R are the most

obvious three-dimensional projections of Γ. In what follows, we shall develop expressions for these projections and examine the connection between X and R.

5.1.1 Notation

One of the most important parameters which specify the projection of Γ on to R is the viewing direction. We will only consider the effects of the projection within a limited volume. If it is sufficiently distant, every line of sight through the volume will be approximately parallel. The viewing direction is then characterised by $\mathbf{n}$, the unit vector along a line of sight.

For many real space quantities there is a corresponding redshift space quantity. We will denote this by the accent $\check{}$. For example, if the real space quantity is y, then the corresponding redshift space quantity will be written $\check{y}$. This notation serves to link corresponding quantities whilst clearly distinguishing between them.

5.1.2 Calculation of the Redshift Space Co-ordinate

Consider a particle with position and velocity vectors $\mathbf{r}$ and $\dot{\mathbf{r}}$ respectively. Employing the notation set out above, its position vector in redshift space is $\check{\mathbf{r}}$. Let us define its position in phase space as $\boldsymbol{\omega} = (\mathbf{r}, \dot{\mathbf{r}})$. We may now formally define the projections from Γ on to X and R, P_X and P_R respectively:

$$P_X : \Gamma \to X, \qquad \boldsymbol{\omega} \mapsto \mathbf{r}; \tag{5.1.1}$$

$$P_R : \Gamma \to R, \qquad \boldsymbol{\omega} \mapsto \check{\mathbf{r}}. \tag{5.1.2}$$

It only remains to derive an expression for $\check{\mathbf{r}}$ in terms of $\mathbf{r}$ and $\dot{\mathbf{r}}$, the components of $\boldsymbol{\omega}$.

The corresponding position and velocity vectors in co-moving space, $\mathbf{x}$ and $\dot{\mathbf{x}}$, are related to $\mathbf{r}$ and $\dot{\mathbf{r}}$ through equations (2.1.1) and (2.1.2). For convenience, these two equations are reproduced below:

$$\mathbf{r} = a\mathbf{x}; \tag{2.1.1}$$

$$\dot{\mathbf{r}} = \dot{a}\mathbf{x} + a\dot{\mathbf{x}}. \tag{2.1.2}$$

The above expressions form the basis of Hubble's law. If $\dot{\mathbf{x}} = 0$, then $\dot{\mathbf{r}} = \dot{a}\mathbf{x}$. Substituting for $\mathbf{x}$ from (2.1.1) yields

$$\dot{\mathbf{r}} = \left(\frac{\dot{a}}{a}\right)\mathbf{r}. \tag{5.1.3}$$

For a homogeneous universe, $\mathbf{r}$ and $\dot{\mathbf{r}}$ may equally well represent the separation and relative velocities of any two objects. If neither has a peculiar velocity, then equation (5.1.3) shows that their relative velocity is proportional to their separation. The constant of proportionality is the Hubble constant, $\dot{a}/a$. Equation (5.1.3) is a mathematical statement of Hubble's law.

Consider the corresponding position vector of the particle in redshift space, $\check{\mathbf{r}}$. This vector can be constructed as the sum of two others. The first is the component of $\mathbf{r}$ perpendicular to the line of sight, $\mathbf{n}$. The redshift projection has no effect in this direction. The second vector is the component of $\dot{\mathbf{r}}$ in the direction of $\mathbf{n}$, scaled by the factor $a/\dot{a}$, representing Hubble's law. $\check{\mathbf{r}}$ can thus be expressed as

$$\check{\mathbf{r}} = [\mathbf{r} - (\mathbf{r}\cdot\mathbf{n})\mathbf{n}] + \frac{a}{\dot{a}}(\dot{\mathbf{r}}\cdot\mathbf{n})\mathbf{n}. \tag{5.1.4}$$

The two terms in the above expression are logically and physically separate; the scaling of the second term is in principle arbitrary. However, if the particle only moves with the general expansion of the universe, its velocity is given by equation (5.1.3). Substituting this expression into (5.1.4) yields $\check{\mathbf{r}} = \mathbf{r}$. If there are no peculiar velocities the transformation from real to redshift space is an

identity with the scaling $a/\dot{a}$. With this choice $\check{\mathbf{r}}$ is as analagous as possible to $\mathbf{r}$.

Whilst possible to study the effects of the redshift projection using equation (5.1.4), it is more convenient to use co-moving co-ordinates. The redshift space equivalent of $\mathbf{x}$ is $\check{\mathbf{x}}$. To make the latter vector as analagous as possible to $\mathbf{x}$, we shall define $\check{\mathbf{r}} = a\check{\mathbf{x}}$. Substituting the expressions for $\mathbf{r}$, $\dot{\mathbf{r}}$ and $\check{\mathbf{r}}$ in terms of $\mathbf{x}$, $\dot{\mathbf{x}}$ and $\check{\mathbf{x}}$ in equation (5.1.4) yields

$$\check{\mathbf{x}} = \mathbf{x} + \frac{a}{\dot{a}}(\dot{\mathbf{x}} \cdot \mathbf{n})\mathbf{n}. \tag{5.1.5}$$

One property of the redshift projection is imme diately obvious from equation (5.1.5). For a given viewing angle, the projection from Γ to R is uniquely determined. It is also obvious that there is *no* unique transformation from X to R or vice-versa. This transformation depends on the details of the peculiar velocity components. Attempts to analyse redshift space data in this way are doomed to failure unless other information is available.

Equation (5.1.5) shows more clearly than (5.1.4) that if peculiar velocities can be ignored, the transformation from real to redshift space is an identity. This is the fundamental connection between X and R. However, this happy state of affairs is upset by the second term in equation (5.1.5). It represents the additional displacement along the line of sight due to the peculiar velocity of the particle. Much of the remainder of the chapter will be concerned with the effects of this term.

5.2 The Evolution of Perturbations in Redshift Space

5.2.1 Particle Positions

The evolution of perturbations in real space has been discussed at some length in Chapters 2 and 4. In particular, the Zel'dovich approximation was presented as a reasonable way to explore evolution in the non-linear regime. The co-moving position of a particle, initially at $\mathbf{q}$, is given by equation (2.2.1). Again for convenience, this equation is reproduced below:

$$\mathbf{x}(\mathbf{q},t) = \mathbf{q} + D(t)\nabla\Psi(\mathbf{q}). \tag{2.2.1}$$

The dimensionless time co-ordinate of the original expression has been replaced by the physical time, t. Other variables were defined in Chapter 2.

The approximation also allows us to explore the evolution of perturbations in redshift space. Taking the time derivative of the above expression gives the co-moving velocity of the particle,

$$\dot{\mathbf{x}}(\mathbf{q},t) = \dot{D}(t)\nabla\Psi(\mathbf{q}). \tag{5.2.1}$$

Combining the last two equations with equation (5.1.5) yields the corresponding co-moving redshift space co-ordinate,

$$\check{\mathbf{x}}(\mathbf{q},t) = \mathbf{q} + D(t)\left\{\nabla\Psi(\mathbf{q}) + f\big(\nabla\Psi(\mathbf{q})\cdot\mathbf{n}\big)\mathbf{n}\right\}, \tag{5.2.2}$$

where

$$f = \frac{a}{\dot{a}}\frac{\dot{D}}{D}. \tag{5.2.3}$$

The function f depends only upon the cosmological density parameter, Ω. To a good approximation, $f(\Omega) \approx \Omega^{0.6}$ [Peebles 1980 §14].

In equation (5.2.2), the two terms in the braces represent the displacement of the particle from its initial position, $\mathbf{q}$. The first term is the actual displacement

in real space, the second the additional displacement because of the redshift projection.

Let us consider the equation in more detail. The term $(\nabla\Psi \cdot \mathbf{n})\mathbf{n}$ is the component of $\nabla\Psi$ in the direction of $\mathbf{n}$. The additional displacement it causes must therefore always augment the real space displacement. Moreover, if θ is the angle between $\nabla\Psi$ and $\mathbf{n}$, the ratio of the magnitude of the displacements is given by

$$\frac{|f\,(\nabla\Psi \cdot \mathbf{n})\,\mathbf{n}|}{|\nabla\Psi|} = f\cos\theta. \tag{5.2.4}$$

Since Ω is close to one, so is f. Equation (5.2.4) shows that the two displacements have similar magnitudes. The effects of the redshift projection *must* be taken into account. In addition, this equation shows that the effect will be most important for small θ, i.e., in those regions for which $\nabla\Psi$ is almost parallel to $\mathbf{n}$. Since displacements along the line of sight are most strongly enhanced, structures in redshift space will tend to be *perpendicular* to the line of sight. These effects will be demonstrated in two-dimensional simulations later in the chapter.

5.2.2 Densities

In Chapter 2, it was shown that the real space density field could be determined under the Zel'dovich approximation (§2.1). The density was calculated from the Jacobian of the transformation (2.2.1). Similarly, the redshift space density, $\check{\rho}$, can also be calculated from the Jacobian of the transformation (5.2.2), $|\check{J}_{ij}|$. Here, $|\check{J}_{ij}|$ is the determinant of the matrix $\check{J}_{ij} = \partial\check{x}_i/\partial q_j$. As in Chapter 2, $\check{\rho}$ can be expressed as $\rho_b/|\check{J}_{ij}|$, where ρ_b is the background density. Alterna-

tively, we may calculate the dimensionless density contrast in redshift space, $\check{\delta} = (\check{\rho} - \rho_b)/\rho_b$:

$$\check{\delta}(\mathbf{q}, t) = \frac{1}{|\check{J}_{ij}|} - 1. \tag{5.2.5}$$

In suffix notation, equation (5.2.2) can be written

$$\check{x}_i = q_i + D\{\Psi_{,i} + f\Psi_{,k}n_k n_i\}. \tag{5.2.6}$$

We have adopted Einstein's conventions: $_{,i}$ stands for $\partial/\partial q_i$ and repeated suffices imply summation unless otherwise stated. The Jacobian matrix is thus given by

$$\check{J}_{ij} = \delta_{ij} + D\{\Psi_{,ij} + f\Psi_{,jk}n_k n_i\}. \tag{5.2.7}$$

In Chapter 2, we were able to reduce the expression for $|J_{ij}|$ since J_{ij} is symmetric. Because $\check{J}_{ij}$ is in general asymmetric, the matrix cannot be cast in diagonal form. The exact value of $|\check{J}_{ij}|$ must be calculated from equation (5.2.7).

However, we may still calculate the density in the linear regime, using the first order approximation

$$\det|\delta_{ij} + B_{ij}| = 1 + B_{ii}. \tag{5.2.8}$$

Summation over i is implied on the right hand side of this equation. Applying this approximation to (5.2.7),

$$|\check{J}_{ij}| \approx 1 + D\{\Psi_{,ii} + f\Psi_{,ij}n_i n_j\}. \tag{5.2.9}$$

Using the expression (2.1.27)

$$\Psi(\mathbf{q}) = \sum_{\mathbf{k}\neq\mathbf{0}} \frac{\Delta_\mathbf{k}}{|\mathbf{k}|^2} e^{i\mathbf{k}\cdot\mathbf{q}}, \tag{2.1.27}$$

yields

$$\Psi_{,ii} = -\sum_{\mathbf{k}\neq\mathbf{0}} \Delta_\mathbf{k} e^{i\mathbf{k}\cdot\mathbf{q}} \tag{5.2.10}$$

and

$$\Psi_{,ij} n_i n_j = -\sum_{\mathbf{k}\neq\mathbf{0}} \Delta_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{q}} \frac{(\mathbf{k}\cdot\mathbf{n})^2}{|\mathbf{k}|^2} \,. \tag{5.2.11}$$

Substitututing these expressions into (5.2.5), we may write the linearised expression for $\check{\delta}$ as

$$\check{\delta}(\mathbf{q},t) = D(t) \sum_{\mathbf{k}\neq\mathbf{0}} \Delta_{\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{q}} \left(1 + f\frac{(\mathbf{k}\cdot\mathbf{n})^2}{|\mathbf{k}|^2}\right). \tag{5.2.12}$$

Strictly speaking, we have only calculated the value of $\check{\delta}$ associated with the Lagrangian co-ordinate, $\mathbf{q}$. However, it is staightforward to show that the first order expression in terms of the Eulerian co-ordinate, $\check{\mathbf{x}}$, is identical. $\mathbf{q}$ merely has to be replaced by $\check{\mathbf{x}}$. Defining the angle between $\mathbf{k}$ and $\mathbf{n}$ to be $\phi_{\mathbf{k}}$, we have the final result:

$$\check{\delta}(\check{\mathbf{x}},t;\mathbf{n}) = D(t) \sum_{\mathbf{k}\neq\mathbf{0}} \Delta_{\mathbf{k}} (1 + f\cos^2\phi_{\mathbf{k}}) e^{i\mathbf{k}\cdot\check{\mathbf{x}}}. \tag{5.2.13}$$

The dependence of $\check{\delta}$ upon $\mathbf{n}$ has been acknowledged explicitly in this equation. This expression was first derived somewhat differently by Kaiser [1987].

5.3 The Redshift Projection in Two Dimensions

There are several reasons for making an initial study of the effects of the redshift projection in only two dimensions. As in the study of initial conditions, this can only aid our understanding of the three-dimensional problem. Particle positions can be completely displayed in two dimensions, allowing a straightforward interpretation of the results. Moreover, in two dimensions it is possible to alter the viewing angle. There is no new physics introduced by extending the calculation to three dimensions. There is also the signifi cant advantage that a wider range of scales can be probed in than in the three-dimensional case.

5.3.1 *The Simulation*

The simulation involves moving the particles according to the Zel'dovich approximation. The positions in real space are determined from equation (2.2.1), those in redshift space from equation (5.2.2). Let us take the initial perturbation field, $\Delta(\mathbf{x})$, to be a Gaussian random field with a power spectrum

$$P(k) \sim k^n \, W_{FD}(k/k_c; \alpha). \tag{5.3.1}$$

The Fermi-Dirac cut-off was defined in §3.1.1. Morover, let us choose to normalise $\Delta(\mathbf{x})$ such that $\langle |\Delta(\mathbf{x})|^2 \rangle = 1$. With this choice of normalisation, D would be the r.m.s. value of the evolved field if only linear evolution had occured. Periodic boundary conditions are imposed.

For the simulations shown, $n = -1$, $k_c = 12\pi$ and $\alpha = 0.3$. The same realisation was used to produce each figure. Particles were initially placed on a square Cartesian grid. The cosmological density parameter, Ω, is taken to be 1 throughout.

Figures 5.1a–d show the evolved particle positions in real space for various values of D. It should be noted that these figures are independent of the value of Ω. Figures 5.2a–d show the equivalent particle positions in redshift space, projected down the line of sight shown. Figures 5.3a–d show the positions in redshift space for $D = 0.4$, projected down various lines of sight.

5.3.2 *Remarks*

Several points should be noted:

(i) Structures are significantly more pronounced in redshift space.

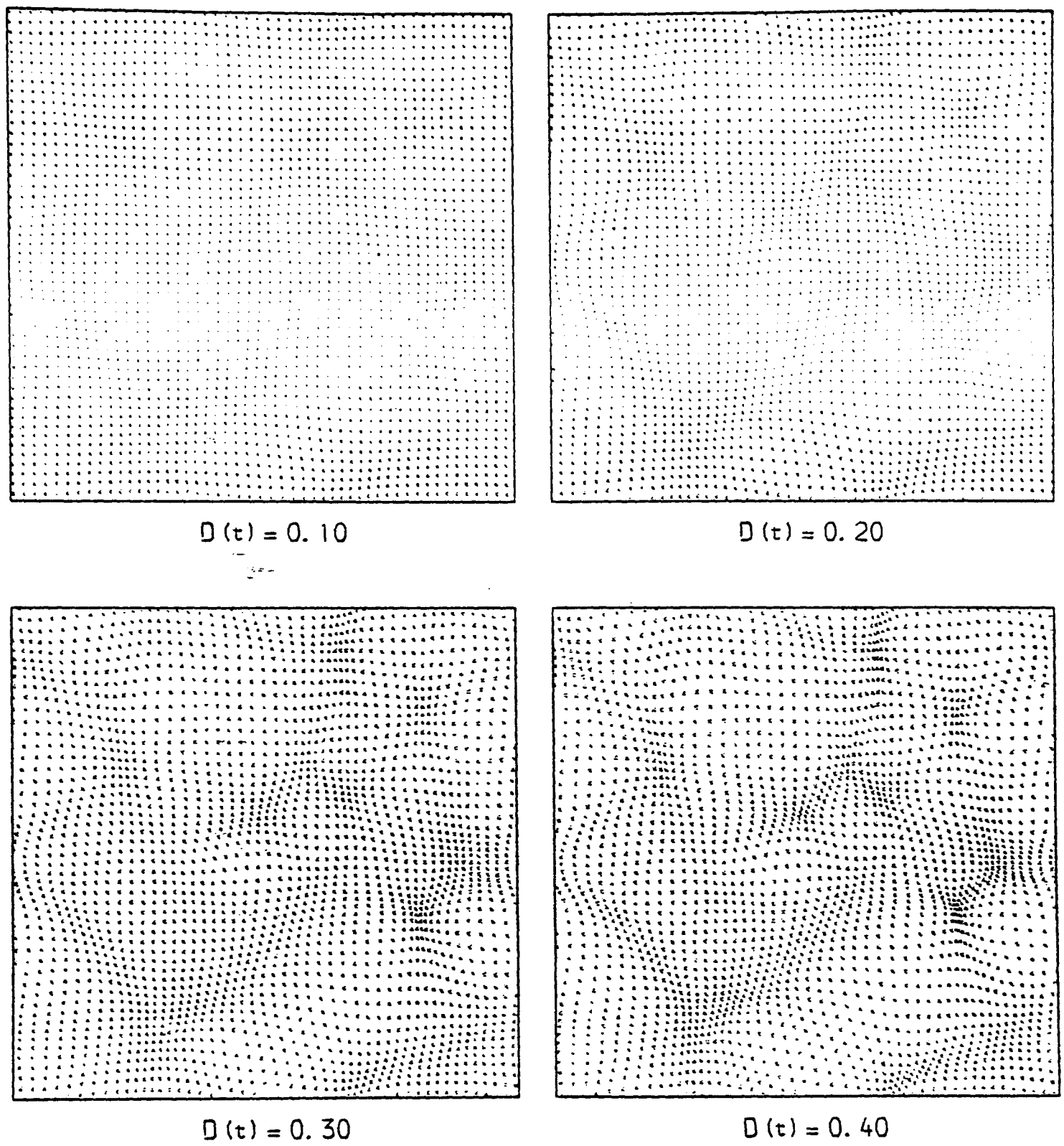


Figure 5.1. Evolved particle positions in real space for various values of the linear growth factor, D.

(ii) Displacement is enhanced along the line of sight, creating structures perpendicular to $\mathbf{n}$.

(iii) Choosing a different viewing direction selects different structures in redshift space. Many of the structures apparent in Figure 5.3a are significantly less

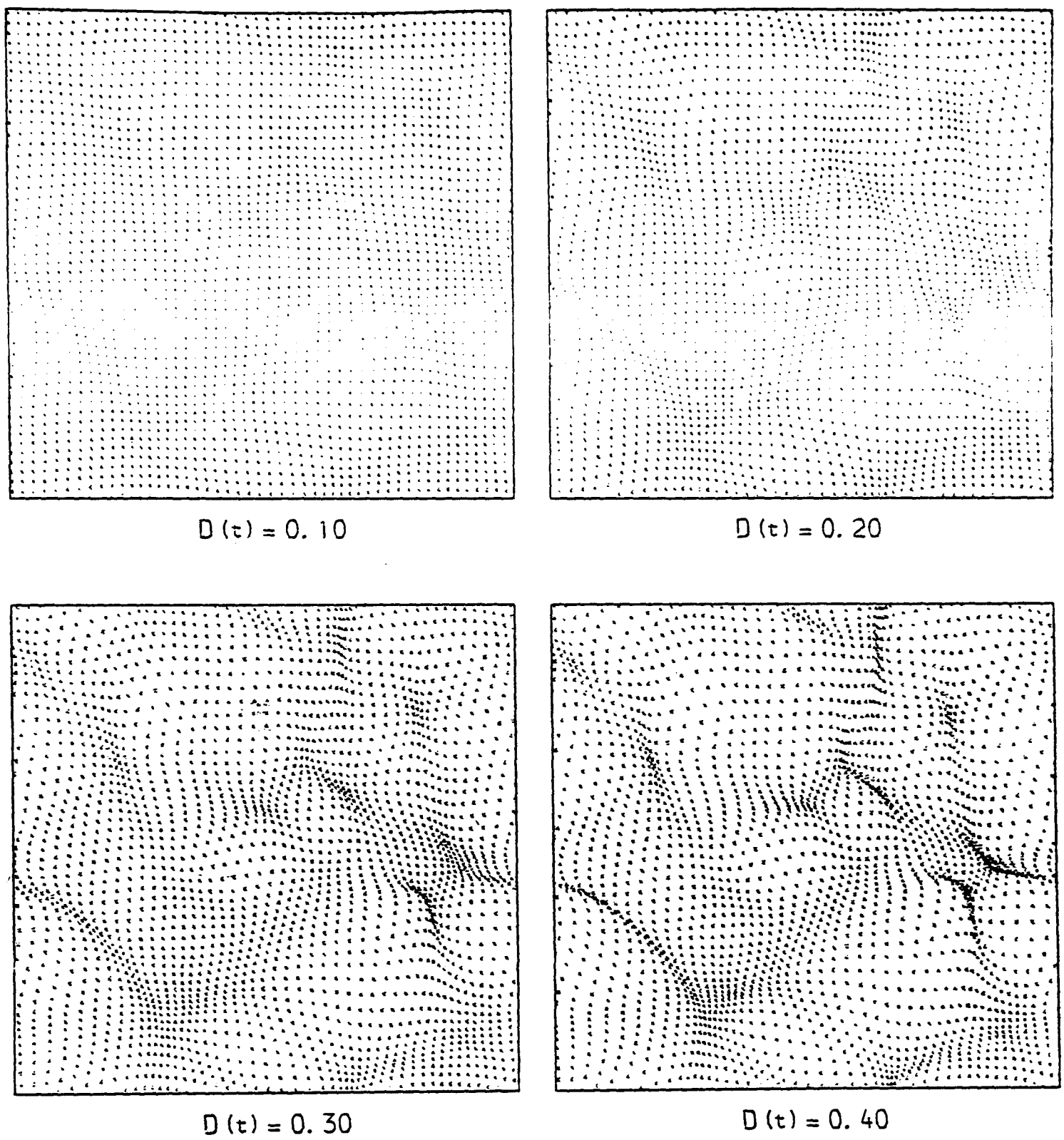


Figure 5.2. Evolved particle positions in redshift space. The line of sight used in the projection is indicated.

so in Figure 5.3d.

(iv) Caustics appear in redshift space whilst particles are at a relatively low density in real space. We will call these structures *redshift caustics*.

Most of these comments were anticipated in §5.2.1. Nothing is said there or

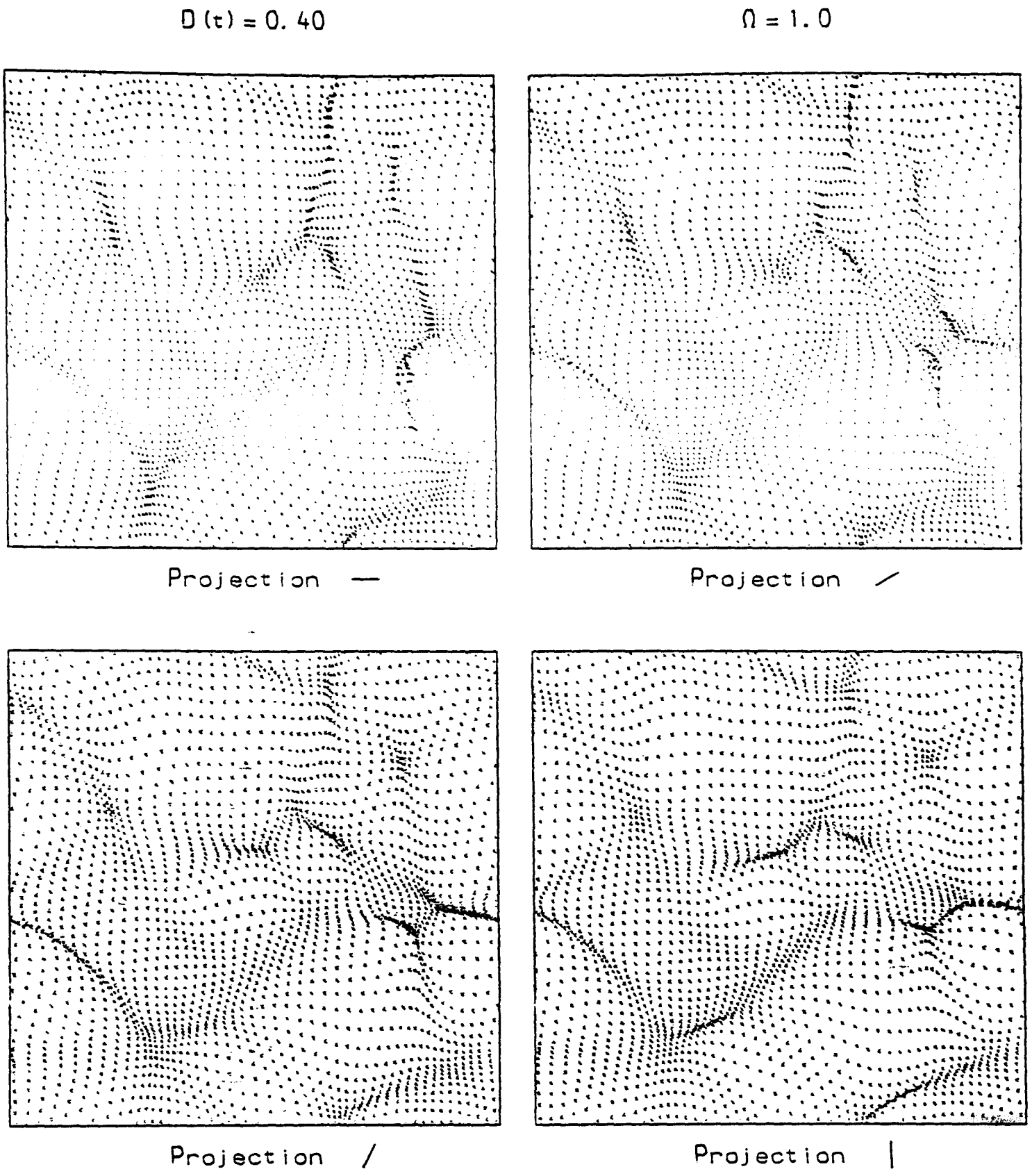


Figure 5.3. Evolved particle positions in redshift space. D is the same for each figure; only the viewing direction changes.

above which will not be equally applicable in three dimensions. There are no extra considerations involved for three-dimensional systems.

5.4 Redshift Caustics

In the previous section, it was noted that caustics form in redshift space. The

purpose of this section is to study their formation and properties in greater depth. We will also be concerned with the details of redshift caustics within the hierarchical scenario.

5.4.1 *Formation and Properties*

In Chapter 2, we remarked that caustics in real space form when the Jacobian of the transformation (2.2.1) is zero. Similarly, caustics in redshift space form when $|\check{J}_{ij}|$ is zero. The differences result from the anisotropy of redshift space.

Consider the process of caustic formation. The collapse is dominated by motion in a particular direction. Let us denote this direction by the vector **c**. In an isotropic universe, **c** may equally well point in any direction. In redshift space, however, collapse is enhanced along the line of sight. **c** will tend to be parallel to **n**. The resulting structures will thus tend to form perpendicular to the line of sight.

A second property of redshift caustics is related to the condition $|\check{J}_{ij}| = 0$. This is one constraint in a d-dimensional system. Caustics will in general be $(d-1)$-dimensional. In our three-dimensional world, redshift caustics will tend to form sheets rather than filaments.

One last property of redshift caustics can be seen by comparing Figures 5.1a–d and Figures 5.2a–d. Caustics form in redshift space whilst densities in real space are still moderate. Let us estimate the real space density for a point lying in a caustic in redshift space. We can use the following simple argument. All distances refer to separations in co-moving co-ordinates, whether in real or redshift space.

Consider two points in a one-dimensional system initially separated by a small distance, Δq. This distance is the same in both real and redshift space since there are no peculiar velocities at $t = 0$. Now allow the system to evolve such that the $\nabla\Psi$ term in equation (2.2.1) has decreased the real space separation by an amount Δx. Equation (5.2.2) tells us that the redshift space separation will have decreased by a futher amount, $f\Delta x$. The remaining separations between the points in real and redshift space, Δr and $\Delta\check{r}$, respectively, are given by

$$\Delta r = \Delta q - \Delta x \quad \text{and} \quad \Delta\check{r} = \Delta q - (1+f)\Delta x. \qquad \text{(5.4.1a and b)}$$

The values of δ and $\check{\delta}$, the dimensionless density contrasts in real and redshift space respectively, are given by

$$\delta = \Delta q/\Delta r - 1 \quad \text{and} \quad \check{\delta} = \Delta q/\Delta\check{r} - 1. \qquad \text{(5.4.2a and b)}$$

The condition for a redshift caustic to form is that $\check{\delta} = \infty$, or alternatively that $\Delta\check{r} = 0$. Equation (5.4.1b) implies that the displacement in real space necessary to form a caustic, Δx_c, is $\Delta q/(1+f)$. Using this value with equations (5.4.1a) and (5.4.2a) yields

$$\delta_c = \frac{1}{f}, \qquad (5.4.3)$$

where δ_c is the value of δ for a point in a redshift caustic.

Although this argument is for a one dimensional system, it should hold for any number of dimensions. We have alredy discussed the pseudo one-dimensional nature of caustic formation. However, it is possible to check this assertion using the simulations which have already been presented. For every point which satisfies $|\check{J}_{ij}| = 0$, the correponding value of $|J_{ij}|$ can be calculated. Let us denote the value of $|J_{ij}|$ for such a point $|J_{ij}|_c$. This is related to δ_c by the expression

$$|J_{ij}|_c = \frac{1}{1+\delta_c}. \qquad (5.4.4)$$

Using our estimate for δ_c, we may expect to find that $|J_{ij}|_c = f/(1+f)$. In an $\Omega = 1$ universe, this corresponds to a value of 0.5.

Figures 5.4a–d show the frequency distributions of $|J_{ij}|_c$ for various values of D. The line of sight used to determine the positions of points on redshift caustics is indicated. Figures 5.5a–d show similar distributions for a value of $D = 0.4$, but for different lines of sight. A value of $\Omega = 1$ is used throughout. These figures show that redshift caustics do tend to form from material whose real space dimensionless density contrast, δ, is close to $1/f$.

5.4.2 Redshift Caustics in the Hierarchical Scenario

The results of the previous section show that redshift caustics occur whilst real space densities are relatively low. Indeed, in an $\Omega = 1$ universe, they occur whilst $\delta \sim 1$. Moreover, these are the most dense regions in the evolved field. The value of D for any figure in this chapter has not exceeded 0.4. We recall that D would represent the r.m.s. fluctuation of the field if only linear evolution had occured.

What does this imply within the context of a hierarchical scenario? For the moment, let us ignore the effects of scales well into the non-linear regime. The arguments in the preceeding paragraph suggest that redshift caustics should occur on the scale just becoming non-linear. Redshift caustics should form on increasingly large scales.

Smaller scales will cause noise in the peculiar velocity field. This will have the effect of broadening the redshift caustics. Unfortunately, it is not simple to estimate the magnitude of the broadening. It is entirely possible that the caustics will be so smeared out as to be totally unrecognisable. However, if this

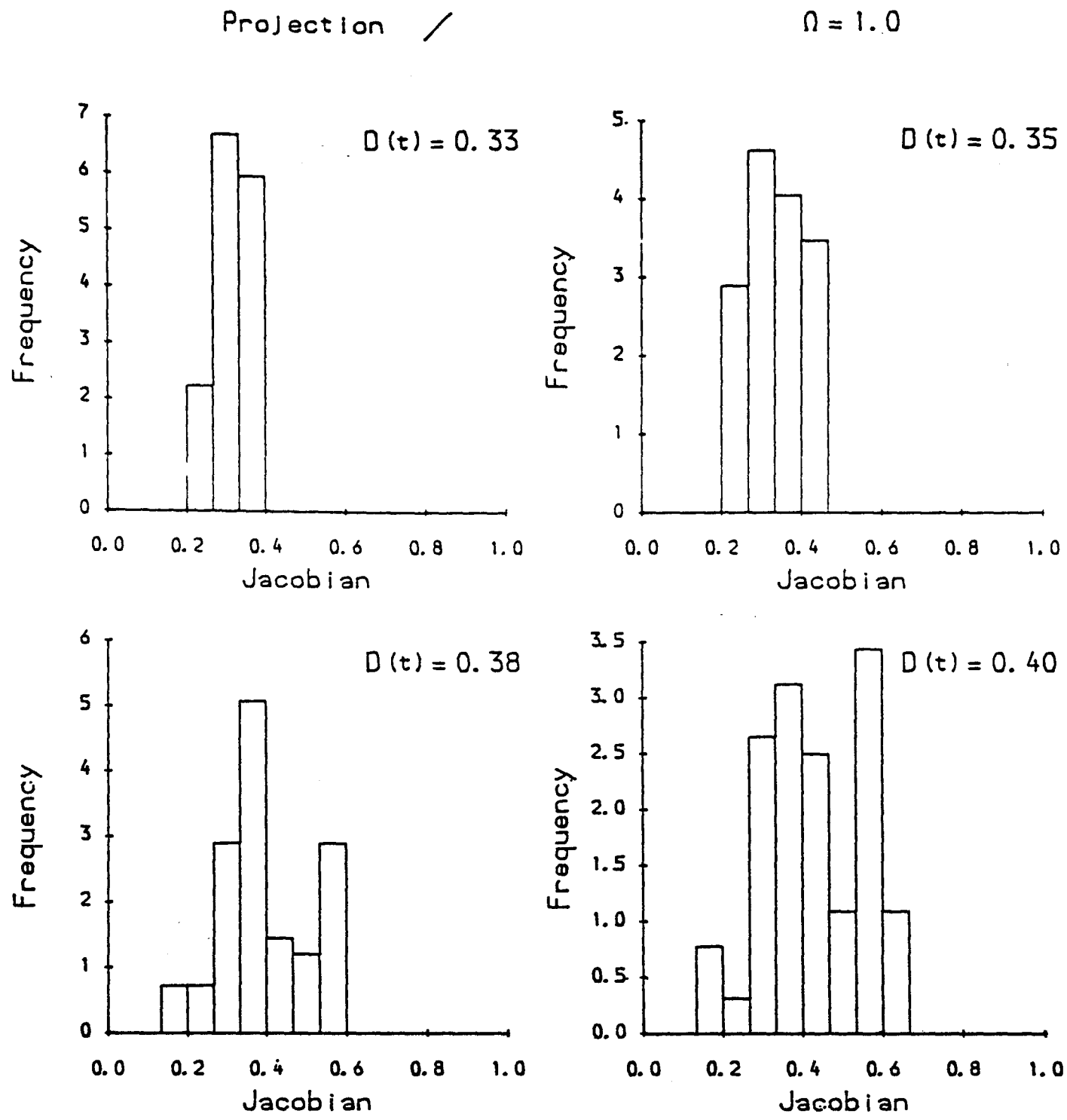


Figure 5.4. The frequency distributions of $|J_{ij}|_c$ for various values of D.

is the case, then it is unlikely that *any* sharp features will be seen in redshift space. Recent galaxy redshift surveys [e.g., de Lapparent *et al.* 1986; Giovanelli and Haynes 1986] have shown that such sharp features do exist. Given their existence, broadening must be unimportant if these structures are to be explained by a hierarchical model.

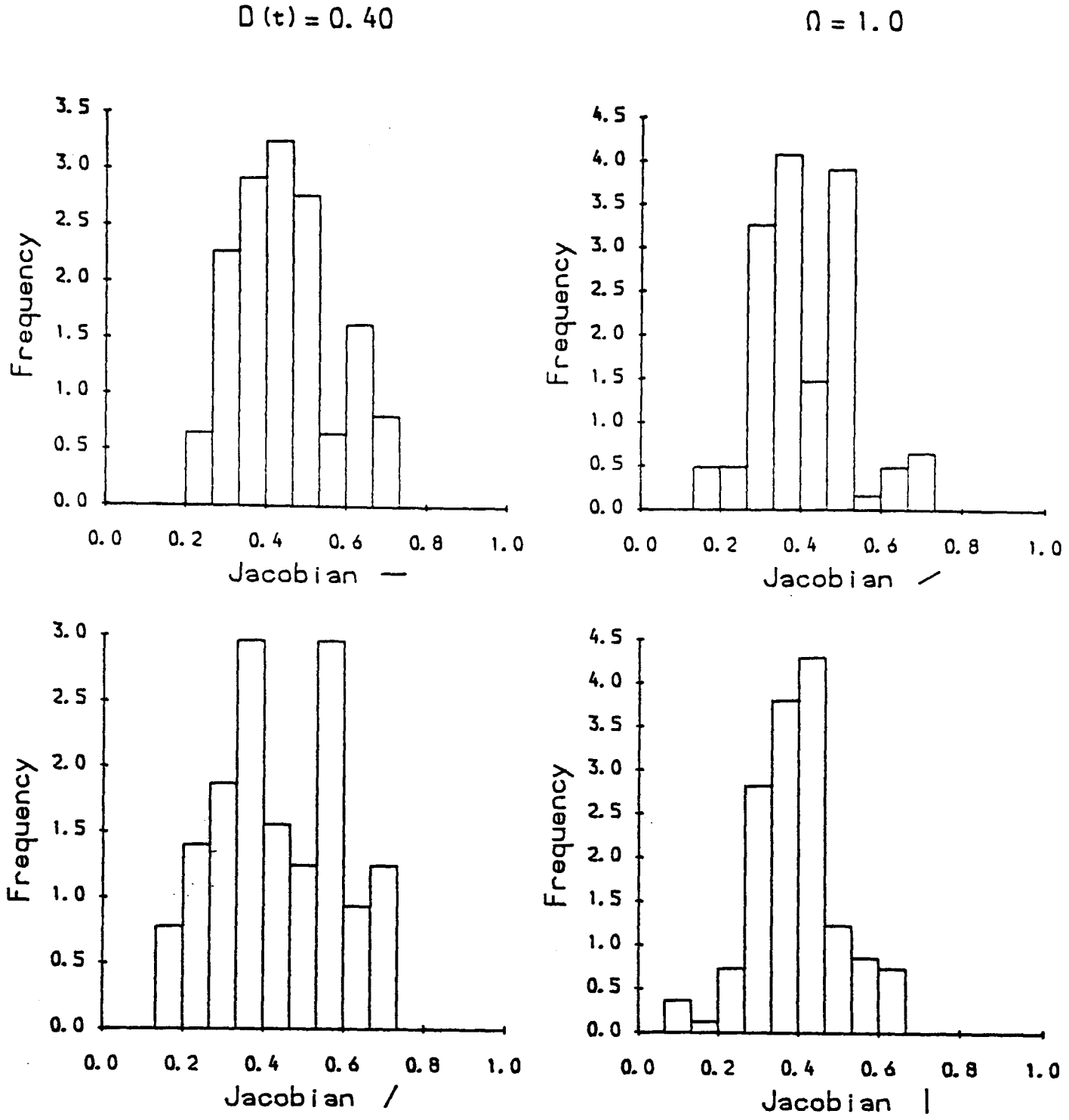


Figure 5.5. The frequency distributions of $|J_{ij}|_c$ for the same value of D but varying the line of sight.

5.4.3 Summary

We may summarise the above in the following points:

(i) Caustics in redshift space occur and tend to form structures perpendicular to the line of sight.

(ii) Redshift caustics should in general be two-dimensional.

(iii) Caustics occur in redshift space whilst the correponding real space density is low. For points on a redshift caustic, $\delta \sim 1/f$.

(iv) Redshift caustics must inevitably form in the hierarchical scenario on the scale entering the non-linear regime. The non-linear evolution of smaller scales will broaden the caustics. It is possible that the broadening may smear out the caustics completely. If so, then *no* sharp features in redshift space should be seen in a hierarchical model.

Cosmological implications of the above will be presented in the following chapter.

5.5 Velocity Correlation Functions

As we discussed in Chapters 1 and 4, the galaxy-galaxy correlation function has been extensively used to characterise the distribution of galaxies. With the advent of systematic galaxy redshift surveys, more positional information is available. It is now possible to determine correlation functions in redshift space, *velocity correlation functions* [e.g., Davis and Peebles 1983; Bean *et al.* 1983]. As we shall see, there are a variety of possible velocity correlation functions, depending on the way directional averages are taken. This section is devoted to calculating these functions in the linear regime of perturbation growth. We shall begin by establishing an expresion for the most general velocity correlation function.

5.5.1 The Fundamental Velocity Correlation Function

Let us calculate the velocity correlation function when particle positions are projected down a particular line of sight, $\mathbf{n}$. This is defined by

$$\check{\xi}(\check{\mathbf{x}};\mathbf{n}) = \langle \check{\delta}(\check{\mathbf{x}}';\mathbf{n})\check{\delta}(\check{\mathbf{x}}'+\check{\mathbf{x}};\mathbf{n})\rangle \qquad (5.5.1)$$

where $\langle \ldots \rangle$ represents an average over the co-ordinate $\check{\mathbf{x}}$. $\check{\delta}$ is the dimensionless density contrast in redshift space.

In the linear regime we may use the expression (5.2.13) for $\check{\delta}(\check{\mathbf{x}}; \mathbf{n})$. In what follows it will be convenient to assign $D(t) = 1$. The time dependence for all the correlation functions derived below may be recovered by multiplying the resulting expressions by $D^2(t)$. Substituting the linear approximation for $\check{\delta}$ into equation (5.5.1) yields

$$\check{\xi}(\check{\mathbf{x}}; \mathbf{n}) = \left\langle \sum_{\mathbf{k},\mathbf{k}' \neq \mathbf{0}} e^{i\mathbf{k}'\cdot\check{\mathbf{x}}'} e^{i\mathbf{k}\cdot\check{\mathbf{x}}'} e^{i\mathbf{k}\cdot\check{\mathbf{x}}} \Delta_{\mathbf{k}} \Delta_{\mathbf{k}'} (1 + f\cos^2\phi_{\mathbf{k}})(1 + f\cos^2\phi_{\mathbf{k}'}) \right\rangle . \tag{5.5.2}$$

Performing the average over $\check{\mathbf{x}}$ retains only those terms in the double sum for which $\mathbf{k} = -\mathbf{k}'$. Thus the above expression reduces to

$$\check{\xi}(\check{\mathbf{x}}; \mathbf{n}) = \sum_{\mathbf{k}\neq\mathbf{0}} |\Delta_{\mathbf{k}}|^2 (1 + f\cos^2\phi_{\mathbf{k}})^2 e^{i\mathbf{k}\cdot\check{\mathbf{x}}}. \tag{5.5.3}$$

This is the fundamental velocity correlation function. All other velocity correlation functions are derived from it by taking appropriate directional averages.

5.5.2 The Directionally Averaged Velocity Correlation Function

As for the spatial correlation function we studied in Chapter 4, the directionally averaged correlation function is calculated by averaging over the direction of $\check{\mathbf{x}}$. Since we are studying isotropic fields (in real space) we can also average over the direction of $\mathbf{n}$ without loss of information. Let us follow Kaiser's notation [1987] and denote this function $\xi_{\langle s \rangle}$.

It is relatively straightforward to average $\check{\xi}(\check{\mathbf{x}}; \mathbf{n})$ over the directions of $\mathbf{n}$ and $\check{\mathbf{x}}$. The averaging proces yields

$$\xi_{\langle s \rangle}(\check{x}) = (1 + \tfrac{2}{3}f + \tfrac{1}{5}f^2) \sum_{\mathbf{k}\neq\mathbf{0}} |\Delta_{\mathbf{k}}|^2 \frac{\sin k\check{x}}{k\check{x}}, \tag{5.5.4}$$

where $\check{x} = |\check{\mathbf{x}}|$.

This expression was first derived by Kaiser [1987]. The author claims no originality in its derivation; it is included here merely for completeness.

5.5.3 *The Line of Sight Correlation Function*

For some studies, the lag separation, $\check{\mathbf{x}}$, is restricted to lie along the line of sight. This is the case for 'pencil beam' galaxy surveys [e.g., Kirschner *et al.* 1983] and for studies of quasar absorption line systems [e.g., Sargent *et al.* 1980]. To compare the predictions of linear theory with these studies, we should calculate the line of sight correlation function, for which $\check{\mathbf{x}} = \check{x}\mathbf{n}$. Again since we are studying isotropic fields, we may average over the direction of $\mathbf{n}$. Once more following Kaiser, we will denote this correlation function ξ_l.

Let us set $\check{\mathbf{x}} = \check{x}\mathbf{n}$ in equation (5.5.3) and consider averaging over the direction of $\mathbf{n}$. Let us define

$$L_\nu(k\check{x}) = \langle e^{i\mathbf{k}\cdot\mathbf{n}\check{x}} \cos^{2\nu} \phi_{\mathbf{k}} \rangle, \tag{5.5.5}$$

where $k = |\mathbf{k}|$ and $\langle \ldots \rangle$ implies averaging over the direction of $\mathbf{n}$. We can thus write the line of sight correlation function in the form

$$\xi_l(\check{x}) = \sum_{\mathbf{k}\neq\mathbf{0}} |\Delta_{\mathbf{k}}|^2 \big(L_0(k\check{x}) + 2fL_1(k\check{x}) + f^2 L_2(k\check{x})\big). \tag{5.5.6}$$

Expressing the average as an integral, L_ν can be written as

$$L_\nu(k\check{x}) = \frac{1}{2}\int_0^\pi e^{ik\check{x}\cos\phi_{\mathbf{k}}} \cos^{2\nu}\phi_{\mathbf{k}} \sin\phi_{\mathbf{k}}\, d\phi_{\mathbf{k}}. \tag{5.5.7}$$

We recall that $\phi_{\mathbf{k}}$ was defined to be the angle between $\mathbf{k}$ and $\mathbf{n}$. It is straightforward, though tedious, to show that L_ν may be obtained from the recurrence relation

$$L_\nu(\chi) = \frac{\sin\chi}{\chi} + \frac{2\nu\cos\chi}{\chi^2} - \frac{2\nu(2\nu-1)}{\chi^2} L_{\nu-1}(\chi) \tag{5.5.8}$$

with

$$L_0(\chi) = \frac{\sin\chi}{\chi}. \tag{5.5.9}$$

It is possible to reduce equation (5.5.6) to the form obtained by Kaiser. However, to aid comparison with the next calculation it will be convenient to leave the expression as it is.

5.5.4 The Full Velocity Correlation Function

Perhaps the most useful of the velocity correlation functions is that defined by Peebles [1980 §76] as ξ_v. The anisotropy of redshift space is acknowledged by writing the lag separation $\check{\mathbf{x}}$ in two parts; $\check{\mathbf{x}}_\parallel$ along the line of sight, $\check{\mathbf{x}}_\perp$ perpendicular to it. Moreover, redshift space is symmetric around the line of sight. It is possible to average over the direction of $\check{\mathbf{x}}_\perp$, for a given $\check{\mathbf{x}}_\parallel$, without loss of information.

It is rather complicated to calculate the corresponding correlation function from equation (5.5.3). We must calculate the average value as the directions of $\mathbf{n}$ and $\check{\mathbf{x}}$ vary, but in such a way as to keep the angle between them constant. For a particular value of $\mathbf{k}$ we must calculate expressions of the form

$$I_\nu(k\check{x}_\perp, k\check{x}_\parallel) = \left\langle \left(e^{i\mathbf{k}\cdot\check{\mathbf{x}}}\right)\left(\frac{\mathbf{k}\cdot\mathbf{n}}{|\mathbf{k}|}\right)^{2\nu}\right\rangle, \tag{5.5.10}$$

where $\langle\ldots\rangle$ implies taking the appropriate average. $\check{x}_\perp = |\check{\mathbf{x}}_\perp|$ and $\check{x}_\parallel = |\check{\mathbf{x}}_\parallel|$. With this definition of I_ν we may write

$$\xi_v(\check{x}_\perp, \check{x}_\parallel) = \sum_{\mathbf{k}\neq\mathbf{0}} |\Delta_\mathbf{k}|^2 \left(I_0 + 2fI_1 + f^2 I_2\right). \tag{5.5.11}$$

Since the calculation of the I_ν is somwhat lengthy, the details are contained in the Appendix. We will merely state the result

$$I_\nu(k\check{x}_\perp, k\check{x}_\parallel) = \sum_{l=0}^{\nu}(4l+1)(-1)^l j_{2l}(k\check{x})\frac{2^{2l}(2\nu)!(\nu+l)!}{(2\nu+2l+1)!(\nu-l)!}P_{2l}(\cos\psi). \tag{A.2.16}$$

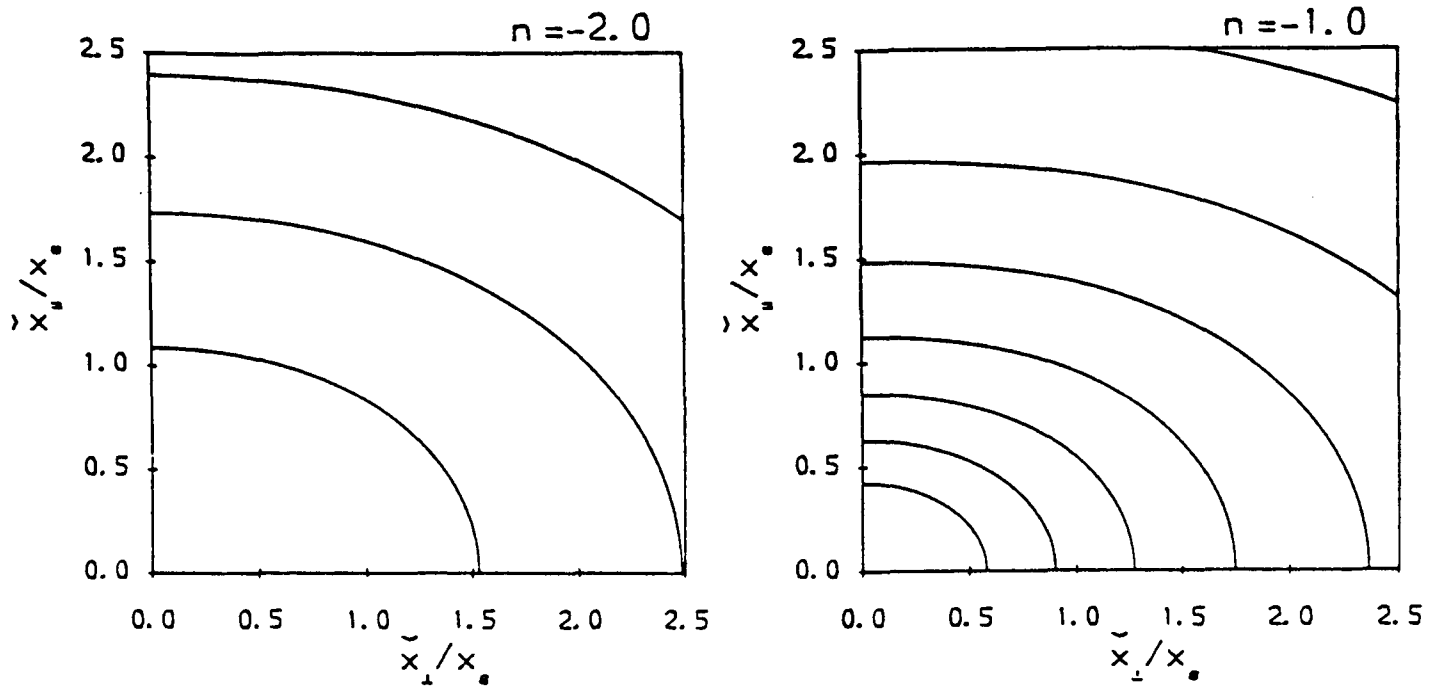


Figure 5.6. The full velocity correlation functions according to linear theory for various values of the spectral index. Note that all are compressed along the line of sight. THE CONTOUR NEAREST THE ORIGIN IS HEIGHEST - 10/6. CONTOUR HEIGHTS DECREASE OUTWARDS IN STEPS OF 1/6.

In this expression, $\cos\psi$ is the angle between $\check{\mathbf{x}}$ and $\mathbf{n}$. j_{2l} and P_{2l} are the spherical Bessel function and Legendre polynomial of order $2l$, respectively.

Taking an ensemble average, we may replace $|\Delta_{\mathbf{k}}|^2$ by the power spectrum, $P(k)$. Figures 5.6a–d show the resulting ξ_v for power spectra of the form

$$P(k) \sim k^n \, W_G(k/k_c). \tag{5.5.12}$$

The Gaussian cut-off was defined in §3.1.1. Power spectra were normalised so that $\langle|\Delta(\mathbf{x})|^2\rangle = 1$.

Each figure displays ξ_v in the form of a contour plot. Contour heights decrease outward in steps of 1/6. The contour nearest the origin is the highest, with a height of 10/6. $\check{x}_\perp$ and $\check{x}_\parallel$ are plotted in units of x_s, defined in §3.1.1 to be $k_c/2\pi$. A value of $f = 1$, corresponding to $\Omega = 1$ was used throughout.

5.5.5 *Discussion*

Before discussing these figures, let us consider the relationship between the three correlation functions derived above. It is obvious that $\xi_l(\check{x}) = \xi_v(0, \check{x})$. It should also be obvious that $\xi_{\langle s \rangle}(\check{x})$ is an averaged version of ξ_v. If one so wished, $\xi_{\langle s \rangle}(\check{x})$ could be derived by averaging ξ_v over the angle between $\mathbf{n}$ and $\check{\mathbf{x}}$. Together with the normalisation condition, equivalent to setting $\sum_{\mathbf{k}\neq\mathbf{0}} |\Delta_\mathbf{k}|^2 = 1$, this provides a convenient way to calculate the amplitude of all the functions at the origin.

The value of $\xi_v(0,0)$ is independent of the angle between $\mathbf{n}$ and $\check{\mathbf{x}}$. Averaging over this angle only produces the same value. By the arguments above, this must equal $\xi_{\langle s \rangle}(0)$ and $\xi_l(0)$. Of these functions, $\xi_{\langle s \rangle}(0)$ is the easiest to evaluate. Using equation (5.5.4) and the normalistion condition, $\xi_{\langle s \rangle}(0) = (1 + \frac{2}{3}f + \frac{1}{5}f^2)$. In an $\Omega = 1$ universe, this corresponds to a value of 28/15.

Apart from being of some intrinsic interest, this result reveals an error in one of Kaiser's figures [Kaiser 1987]. Whilst his Figures 2b–d show this result, Figure 2a does not. Moreover, the correlation function for a continuous function should be a maximum at the origin. Kaiser's Figure 2a is clearly wrong. However, since his equations are correct, this is probably only some slight plotting or computational error.

Let us return to Figures 5.6a–d. The most obvious feature in all the figures is that ξ_v is compressed along the line of sight. This is hardly surprising since

we have already seen that structures tend to form perpendicularly to the line of sight. However, this result strongly contrasts with the observed form of ξ_v. Both Davis and Peebles [1983] and Bean *et al.* [1983] find that ξ_v is *elongated* along the line of sight. This is presumably due to the random motion of galaxies within clusters. We may conclude that linear perturbation theory is inappropriate to analyse the distribution of galaxies in redshift space. Moreover, the spatial correlation function is a suitably deprojected form of ξ_v. Linear perturbation theory must also be inappropriate for analysing $\xi_{gg}(r)$.

Superficially, this conclusion seems to contradict an argument earlier in the chapter. How can random galaxy motions elongate ξ_v but not broaden caustics? Perhaps part of the answer can be found amongst the considerations of the last chapter. Correlation functions tend to be dominated by the densest regions. It is just these regions which will have collapsed first and formed Fingers of God. Redshift caustics, on the other hand, should be found in regions which are just becoming non-linear; there are no real random velocities in these regions. This argument is probably too simple. It only really considers collapse on one scale; it may not be applicable to the real universe.

Be that as it may, observations do show both elongation of ξ_v and sharp features in redshift space. The CfA 'slice' redshift survey [de Lapparent *et al.* 1986] reveals a Finger of God in close proximity to a structure which may be a redshift caustic—see Figure 6.1. Similar forms are seen in N-body calculations [White *et al.* 1987]. Whether the above argument applies or no, the universe *has* managed to elongate ξ_v without providing appreciable random velocities everywhere.

Whilst the major conclusion to be drawn from the above calculations is fairly negative, there is at least one positive aspect. It is often stated that the form of

correlation functions at large lag separations will be dominated by scales still in the linear regime. If this is so, then it can be tested by looking for compression, rather than elongation, of ξ_v along the line of sight.

5.6 Summary

In this chapter, we have examined some effects of using velocity as a distance measure. For any scenario in which structures grow from small perturbations, we have shown that these considerations are important.

In §5.3 the formation and properties of caustics in redshift space were studied. We concluded that redshift caustics tend to form as structures perpendicular to the line of sight. Moreover, they form whilst the density in real space is quite modest. We also discused the importance of redshift caustics in the hierarchical scenario.

The last section in this chapter considered the calculation of various velocity correlation functions in the linear regime. Our study led us to conclude that the predictions of linear perturbation theory are at odds with observation. Linear theory is an inappropriate tool for the study of correlation functions. This conclusion has already been used in Chapter 4 as evidence that non-linear regions dominate the real space correlation function.

In this chapter, we have mainly been concerned with the concepts of the redshift projection. In the following chapter these concepts will be applied in a cosmological setting. In particular, we shall examine the proposal that quasar absorption lines represent redshift caustics.

Chapter 6. Quasar Absorption Lines and Redshift Caustics

The formation and properties of redshift caustics were discussed in the preeceeding chapter. One conclusion was that caustics in redshift space are an almost inevitable feature of the hierarchical scenario. Only the extent of the broadening caused by the non-linear evolution of small scales remained in doubt. The purpose of this chapter is to examine the evidence that various structures in redshift space *are* caustics. Broadening will be assumed to be unimportant throughout. We will be particularly concerned with the structures causing the Ly-α forest in quasar spectra. We will begin, however, by discussing the structural forms found in galaxy redshift surveys.

6.1 Bubbles and Filaments

Perhaps the most obvious place to look for redshift caustics is in the recent galaxy redshift surveys. The structures in two of the surveys will be discussed here. The first to be considered will be the CfA 'slice' redshift survey [de Lapparent *et al.* 1986]. The CfA slice is reproduced in Figure 6.1.

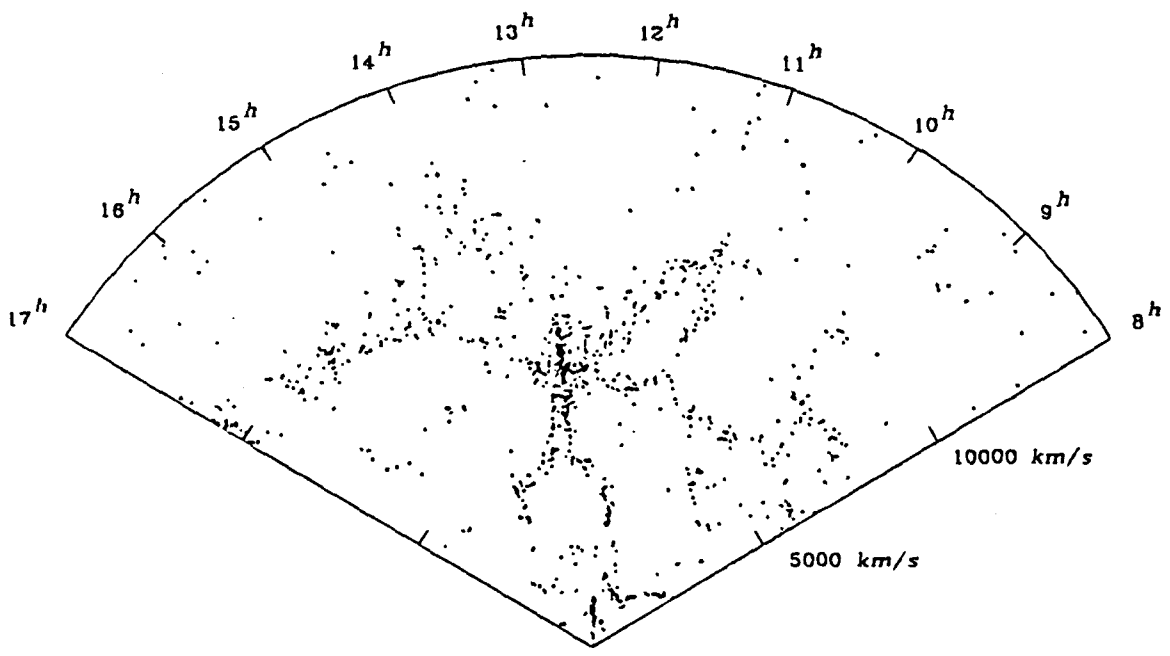


Figure 6.1. Map of the observed galaxy velocities vs. right ascension in the declination wedge $26.5° \leq \delta \leq 32°$. Taken from de Lapparent *et al.* [1986].

The figure shows sharply defined structures with large density contrasts. Moreover, there is a tendency for the structures to lie parallel or perpendicular to the line of sight. Those perpendicular to the viewing direction are good candidates to be redshift caustics. Those parallel are 'Fingers of God'.

Structures in this survey show many of the features expected of redshift caustics. Unfortunately, however, there are at least two structures which are neither parallel nor perpendicular to the line of sight—see Figure 6.1. With the present data it would be unwise to conclude that the structures are redshift caustics.

The second survey we will consider is the 21cm survey of the Pisces-Perseus super-cluster [Giovanelli and Haynes 1984,1986; Giovanelli *et al.* 1986]. The best candidate to be a redshift caustic is a filament in the major ridge of the super-cluster. This is particularly well shown in Giovanelli *et al.* Figure 4. This is a large, well defined filament of high galaxy density lying perpendicular to the line of sight. Again, whilst this structure shows most of the features of a redshift caustic, it is only one-dimensional. Although it is possible to have a one-dimensional caustic, it does require a certain fine tuning of initial conditions. If many more filaments are found it will not be possible to explain them as redshift caustics.

The evidence for redshift caustics from galaxy redshift surveys is inconclusive. The fact remains, however, that peculiar velocities will have a significant effect on the morphology of structures in redshift surveys. This point has already been emphasised by Kaiser [1987].

6.2 Quasar Absorption Lines: A Review

Later in the chapter, it will be argued that quasar Ly-α absorption lines represent redshift caustics. In preparation, this section reviews properties of the observed Ly-α lines in some depth. It should be borne in mind that these are the properties of the *detected* lines. There may be many lines which are currently undetectable. It is perhaps the distributions of line properties, rather than the properties of typical lines, which contain the most fundamental information.

6.2.1 Overview

In 1980, Sargent *et al.* (SYBT) presented a systematic statistical study of Ly-α absorption lines. The statistical properties of the lines were independent of the quasar spectrum in which they were observed. SYBT concluded that the lines represented a cosmologically distributed population of absorbing structures, 'clouds'.

There are typically 60 detected lines per unit redshift interval, although this number increases with redshift. The equivalent widths of the lines are in the range 0.1–1.0 Å [SYBT]. Whilst there is no reason to miss lines with larger equivalent widths, the lower limit may only represent the resolution at which the spectra were obtained [see, however, Carswell *et al.* 1984]. These widths correspond to neutral hydrogen column densities, $N_c(\mathrm{H\,I})$, of around 10^{14} atoms/cm^2.

With high resolution spectra, it is possible to do a little better. Theoretical line profiles may be fitted to many of the absorption features [Carswell *et al.* 1984]. The resulting $N_c(\mathrm{H\,I})$ are in the range 10^{13}–10^{15} atoms/cm^2. The fitting procedure also yields velocity dispersions. These are in the range 10–50 km s^{-1}, with an average value of 25 km s^{-1}.

Again, the lower limit to the column density range probably reflects the resolution. There is, however, a theoretical lower limit to the velocity dispersion of the lines. Absorbing structures are expected to be highly ionised by the background u.v. flux [Ikeuchi and Ostriker 1986; Rees 1986]. This ionisation will maintain the hydrogen at a temperature of 10^4 K. This correponds to a velocity dispersion of $10\,\mathrm{km\,s^{-1}}$. The lower limit to the range of velocity dispersions may have physical significance [but see Hogan 1987].

6.2.2 Distribution of Equivalent Widths

There does not seem to be any relationship between the redshift of a line and its equivalent width, W. Performing a rank correlation test between these two quantities reveals no significant correlation [SYBT]. This suggests that the distribution of equivalent widths is *not* dependent on redshift. The number of lines with equivalent widths in the range W to $W + dW$, $n(W)dW$, is found to be well fitted by the exponential form

$$n(W)\,dW \propto \exp\left(-W/W^*\right)\,dW, \tag{6.2.1}$$

where $W^* \approx 0.3\,\text{Å}$ [SYBT].

6.2.3 Distribution of Column Densities

The neutral hydrogen column densities, $N_c(\mathrm{H\,I})$, are obtained by fitting theoretical line profiles to absorption features. The distribution of column densities is *not* well fitted by the exponential form. This is somewhat surprising since most of the lines are optically thin. The number of clouds with column densities in the range N_c to $N_c + dN_c$, $n(N_c)dN_c$, is found to be well fitted by

$$n(N_c)\,dN_c \propto N_c^{-s}\,dN_c, \tag{6.2.2}$$

with $s \approx 2$ [Carswell *et al.* 1984; Atwood, Baldwin and Carswell 1985].

It is possible to reproduce this distribution by assuming that the clouds are optically thin, two-dimensional structures with constant neutral hydrogen densities. If they are randomly oriented in space, then the distribution of path lengths, l, varies as l^{-2}. The distribution of column densities would thus follow equation (6.2.2). This simple model also has the advantage that there is a natural minimum column density, set by the thickness of the structures. Barcons and Fabian [1987] independently studied this idea in some depth and reached the same conclusion.

6.2.4 Number Density of Lines

The number density of lines, dN/dz, is the number of detected lines per unit redshift interval. Observations indicate that dN/dz depends upon redshift; more lines are found at higher redshifts than at lower. The number density of lines is fitted by

$$\frac{dN}{dz} \propto (1+z)^{\gamma}, \tag{6.2.3}$$

with $\gamma \approx 2$ [Murdoch *et al.* 1986]. For comparison, if the lines were caused by an un-evolving population of absorbers, γ would be expected to be 0.5 in an $\Omega = 1$ universe. If the absorbing structures are un-evolving in co-moving co-ordinates, γ should be -1.5 if $\Omega = 1$.

6.2.5 Line of Sight Correlation Functions

To look for any clustering of the lines, SYBT determined the two-point correlation function along the line of sight (see §5.5.3). They employed a variety of co-ordinate systems to define line separations, but no matter which was used,

the correlation function was flat on all scales. They concluded that there was no evidence for clustering on any scale.

6.2.6 Variation of Parameters with Sample

The values of the parameters W^*, s and γ of equations (6.2.1)–(6.2.3) vary with the sample of lines used. γ increases as the redshift range of the sample increases, W^* decreases as the resolution of the spectra improve. A recent paper [Murdoch *et al.* 1986] suggests that absorption lines close to the Ly-α emission line are weak. Murdoch *et al.* proposed that hydrogen close to the quasar was more strongly ionised by the higher than average u.v. flux. Hence absorbers near to the quasar would give rise to lower equivalent widths and are less likely to be detected.

According to Murdoch *et al.*, if this is the case a lower proportion of the lines will be affected as the redshift range increases.[1] Hence proportionally more lines will be detected which will increase the value of γ. Similarly, as the resolution improves, more of the lower width lines can be detected. This will have the effect of decreasing W^*. Murdoch *et al.* strengthened this viewpoint by examining the trends amongst lines from individual spectra. There are statistically fewer lines than expected near to the Ly-α emission line and they do tend to have low equivalent widths. They called this tendency the 'inverse effect'. Futhermore, they showed that by excluding those lines closest to the Ly-α emission line, the parameters obtained from various samples come into closer agreement.

Whilst this may prove to be the reason that γ varies with redshift range, it cannot completely explain the variation of W^* with resolution. As resolution improves, many previously single lines become resolved into separate compo-

[1] Since dN/dZ increases with redshift, this is not as obvious to the author as it is to them.

nents. Moreover, new features should emerge from the continuum. This must reduce the value of W^*. For example, compare the values of W^* obtained from two samples taken from the same quasar, Q1101-264 [Carswell *et al.* 1984]. For Ly-α lines which are free from possible blending effects, $W^* = 0.07 \pm 0.02$ Å. Including all the lines gives the value $W^* = 0.16 \pm 0.02$ Å. Since blended lines are in some sense less well resolved, this does seem to imply that W^* changes with resolution—unless un-blended lines also occur closer to the quasar!

6.2.7 Cloud Sizes

A lower limit to the transverse size of a cloud is implied by observations on the quasar pair 2345+007A,B [Foltz *et al.* 1984]. Foltz *et al.* noted that many of the Ly-α absorption features were common to both images. Using this information, a lower size limit may be established.

Spectra strongly indicate that the quasar pair are images of the same quasar, evidence for a gravitational lens. With this assumption, Foltz *et al.* calculated the separation of the light paths at the redshifts of the common features. The maximum separation yields a lower limit for the transverse size of a cloud. This limit depends upon the redshift of the deflecting mass, but must be somewhere in the range 2–$15h^{-1}$ kpc.

Since two unmatched lines were also observed, some authors have claimed that this is a typical size rather than a lower limit. Let us examine this claim under the simplifying assumption that the light paths are co-planar and parallel. If s is the separation between the paths and r_t the mean transverse size of the absorbing structures, then the probability that an absorber will intersect exactly one path is given by $\beta s/r_t$. β is a dimensionless parameter, of order unity, which depends on the geometry of the structures. If the clouds are spherical,

for example, $\beta = 4/\pi$ for $r_t \gg s$. For the double quasar 2345 + 007, 13 matched and 2 unmatched lines were observed. The probability for this to occur, p, is given by the binomial distribution:

$$p = \left(\frac{\beta s}{r_t}\right)^2 \left(1 - \frac{\beta s}{r_t}\right)^{13} \frac{15!}{2!\,13!}. \tag{6.2.4}$$

Maximising the probability gives the maximum likelihood value

$$\frac{\beta s}{r_t} = \frac{2}{2+13} = 0.13. \tag{6.2.5}$$

This calculation indicates that the most likely size for the clouds is ~ 10 times the average path separation. Moreover, there is greater than a 10% chance that $\beta s/r_t$ is as low as 0.04. The characteristic transverse cloud size, r_t, is clearly poorly constrained by these observations.

Relaxing the assumption that the light paths are plane parallel will decrease the expected size; the two unmatched lines have the widest path separations. This relaxation may decrease the expected size significantly if the separation of the light paths varies substantially over the observed range. There are also at least two futher candidates to be considered as unmatched lines. The signal to noise ratio for the present spectra is too low to determine whether or not these lines are real. Better observations combined with a more thorough analysis may yield a statistically significant value for the characterisic cloud size.

6.2.8 Gunn-Peterson Test

Measurements on the main body of the quasar spectrum cannot determine whether or not there is neutral hydrogen between absorption lines. Variations in the continuum level are indistinguishable from variations in the density of a smoothly distributed H I component. The only constraint on such a component

comes from comparing the continuum levels blueward and redward of the Ly-α emission line. The redward continuum is unaffected by the intervening hydrogen. A discontinuity would reveal the presence of H I. This is the basis of the Gunn-Peterson test [1965].

Let us calculate the amount of absorption that the H I component will cause at an observed frequency, $\nu_{\rm obs}$. Consider neutral hydrogen atoms at a redshift z. In their rest frame, the absorption cross-section at a frequency ν can be written:

$$\sigma(\nu) = \sigma_{\rm o}\phi(\nu - \nu_\alpha), \tag{6.2.6}$$

where $\sigma_{\rm o}$ is the total absorption cross-section and ν_α is the (rest frame) frequency of the Ly-α transition. ϕ is a function describing the line profile and has the property that

$$\int_0^\infty \phi(\nu - \nu_\alpha)\,{\rm d}\nu = 1. \tag{6.2.7}$$

$\sigma_{\rm o}$ is given by $(\pi e^2/m_e c)\, f_\alpha$ where f_α is the oscillator strength of the transition. $\sigma_{\rm o}$ has the numerical value $0.011\,{\rm cm}^2\,{\rm Hz}$ [Allen 1973].

In the rest frame at a redshift z, the observed frequency $\nu_{\rm obs}$ has a value $\nu_{\rm obs}(1+z)$. The optical depth at $\nu_{\rm obs}$ can thus be expressed as:

$$\tau(\nu_{\rm obs}) = \int_0^{z_{\rm em}} \check{n}(z)\sigma_{\rm o}\phi\big((1+z)\nu_{\rm obs} - \nu_\alpha\big)\frac{dr}{dz}\,dz. \tag{6.2.8}$$

In this expression, $z_{\rm em}$ is the redshift of the Ly-α emission line of the quasar under study. $\check{n}(z)$ is the number density of neutral hydrogen atoms in redshift space at the redshift z. dr/dz is the physical length of the path per unit dz; $(dr/dz)dz$ is the physical path length, dr, for the redshift increment dz.

For the system we are considering, the line profile will be fairly sharp. The major contribution to the width is from the thermal broadening; ϕ will typically be $10\,{\rm km\,s^{-1}}$ wide. Since neither of the functions $\check{n}(z)$ nor dr/dz will alter

significantly over such a small velocity range, we may treat ϕ as a δ-function. Performing the integral yields

$$\tau(\nu_{\rm obs}) = \frac{\check{n}(z)\sigma_{\rm o}}{\nu_{\rm obs}}\frac{dr}{dz}, \tag{6.2.9}$$

where the functions $n(z)$ and dr/dz are evaluated at a redshift, z, satisfying $(1+z)\nu_{\rm obs} = \nu_\alpha$. Let us now rewrite the expression for τ as a function of z. Using the expression (2.1.12) for dr/dz and expressing $\nu_{\rm obs}$ in terms of $(1+z)$ and ν_α yields

$$\tau(z) = \frac{\check{n}(z)\sigma_{\rm o}c}{\nu_\alpha H(z)}. \tag{6.2.10}$$

This derivation closely follows that of the original calculation by Gunn and Peterson. It is repeated here to allow a slight extension later in the chapter.

Observations by Steidel and Sargent [1987b] indicate that there *is* a discontinuity between the continuum levels on either side of the Ly-α emission line. For the eight quasars in their sample, with a mean redshift of 2.64, the mean observed discontinuity corresponds to an optical depth, $\tau_{\rm o} = 0.27 \pm 0.02$. However, Steidel and Sargent attributed much of the discontinuity to the presence of unobserved Ly-α lines. At their 4 Å resolution, many of the sharper features are smoothed over, effectively lowering the continuum [Steidel and Sargent 1987a]. Using Murdoch *et al.*'s distributions of column densities etc., Steidel and Sargent calculated that there should be an optical depth, $\tau_c = 0.25 \pm 0.02$. They proposed an upper limit for the Gunn-Peterson optical depth of

$$\tau(2.64) < 0.02 \pm 0.03. \tag{6.2.11}$$

It is obvious from their calculations that it is not trivial to determine an accurate limit. Steidel and Sargent's observations are entirely consistent with a Gunn-Peterson optical depth of zero. On the other hand, their value is the

difference of two much larger values. The distributions of line properties are not known precisely. Since they are used in the calculation of τ_c, the systematic error in this calculation is probably much larger than the formal error Steidel and Sargent quote. The true limit for the Gunn-Peterson optical depth could be much larger. Certainly a value of 0.05, and probably as large as 0.1, is not excluded by the observations.

A model of the structures causing the Ly-α forest should endeavour to calculate the magnitude of this effect. If the calculated optical depth between lines is less than 0.1, it cannot (yet) be ruled out by the Gunn-Peterson test.

6.2.9 Comments

The view of the Ly-α absorption lines presented above is fairly standard. There are, however, certain features of the distributions which are puzzling, at least to the author. For example, there is apparently no correlation between the redshift of a line and its equivalent width—§6.2.2.

Most models of the Ly-α clouds expect the hydrogen to be almost fully ionised by the background u.v. flux [e.g., Ikeuchi and Ostriker 1986; Rees 1986]. The inverse effect alone is evidence that the structures are affected by this flux. Since the flux is probably produced by quasars, it is expected to vary with redshift. The column densities of the clouds should also change. The absence of correlation may have several implications. There might be a self-similar distribution of column densities or perhaps an unnatural balance between the column density distribution and the evolution of the u.v. background.

There may be an alternative explanation, however. It is interesting to note that the form of the distribution of equivalent widths does *not* change with resolution. Since many lines are resolved into several components as the resolution

improves, it is not obvious that the form should remain the same. Liu and Jones [1987] have shown that blending effects are important. It is at least possible that the distribution of equivalent widths is purely a result of line blending.

Whilst this possibility may be worth examining, there does seem to be a self-similar distribution of column densities (§6.2.3). The number density of detected lines then becomes a reflection of the background flux variations. In their paper, Murdoch *et al.* found that $\gamma = 2.17 \pm 0.36$. They compared the observed distribution of lines with the proposed form by perfoming a one-sided Kolmogorov-Smirnov test. The probability that deviations from the proposed number density occurred by chance was 0.95. This suggests a very good fit to the data. Futhermore, this implies that γ does not vary with redshift. How does the ionising flux change over the redshift range in their study, 1.5–3.78? Whilst it is not known with any certainty, it is unlikely to be as a power of $(1 + z)$. Tytler [1985] even suggests that the u.v. flux at the Lyman limit would peak at a redshift of around 2.5. The implications for the Ly-α lines are not obvious. If the number density of lines does not follow the u.v. flux, then presumably the evolution is dominated by other effects.

Column densities of detected clouds vary from 10^{13}–10^{15} atoms/cm^2. If the distribution of column densities really is self-similar over this range then it will prove quite a challenge to explain theoretically. It is, however, possible to examine the self-similarity observationally. The distribution of column densities, or alternatively equivalent widths, should be determined at two different redshifts. Differences in the distributions can be found through the application of a two-sided Kolmogorov-Smirnov test.

Recent observations have indicated that the distribution of column densities may not be self-similar [Carswell *et al.* 1987]. The same paper also questions

whether or not quasars can generate the u.v. background necessary to ionise the clouds at high redshifts ($z \gtrsim 3$). We may conclude that whilst the overall properties of the Ly-α forest are well established, the exact details are less certain. In part this is a reflection of the poor theoretical understanding of the structures causing the absorption lines. No theory yet presented is entirely compelling.

Perhaps the reason that both currently popular models for these structures (discussed in §1.5) are un-persuasive is that they lack predictive power. It is impossible to decide whether or not they can account for the known properties of the Ly-α lines. In the following section a radically different model of the 'clouds' is presented. The model proposes that the absorption lines represent redshift caustics. As we shall see, this model can account for many of the properties discussed above.

6.3 Quasar Absorption Lines as Redshift Caustics

In the previous chapter, it was argued that caustics in redshift space arise naturally in the hierarchical scenario. Seen in absorption, the caustics would appear as sharp lines. The proposal that these caustics are the structures causing the Ly-α forest is attractive. The remainder of this chapter is devoted to exploring the proposal in more detail. We shall assume that the initial density perturbation is a Gaussian random field with a power spectrum such that the evolution will be hierarchical—see Chapter 3. As a preliminary, let us discuss the likely ionisation state of the hydrogen.

6.3.1 *Equilibrium Ionisation Calculation*

Since the recombination timescale is much shorter than the Hubble time [Ikeuchi and Ostriker 1986], the hydrogen in the absorbing structures is likely to be in equilibrium with the ionising u.v. background. Let n_e, n_p and n_H be the densities of electrons, protons and (neutral) hydrogen respectively. The ionisation balance equation can be written:

$$n_e n_p \alpha_H(T) = n_H \int_{\nu_c}^{\infty} J(\nu)\sigma(\nu)\, d\nu \qquad (6.3.1a)$$

$$\equiv n_H J(\nu_c) G_H. \qquad (6.3.1b)$$

The left hand side of this equation represents collisional recombination. $\alpha_H(T)$ is the collisional recombination coefficient at a temperature, T. The right hand side repesents the ionisation by the background flux, $J(\nu)$. Collisional ionisation is *not* important. $\sigma(\nu)$ is the ionisation cross-section at a frequency, ν. G_H is defined by equation (6.3.1b); $J(\nu_c)$ is the flux at the Lyman limit frequency, ν_c. This is likely to be in the range $10^{-19} > J(\nu_c)/\,\mathrm{erg\,cm^{-2}} > 10^{-21}$ [Black 1981]. Assuming that $J(\nu) \propto \nu^{-1}$, $G_H = 2.54\,\mathrm{erg^{-1}\,cm^2\,Hz}$ [Black 1981].

If n_t is the total hydrogen density, $(n_p + n_H)$, and χ the neutral fraction, this equation can be rewritten as

$$\chi = \frac{(1-\chi)^2 n_t \alpha_H(T)}{J(\nu_c)G_H} \approx \frac{n_t \alpha_H(T)}{J(\nu_c)G_H}. \qquad (6.3.2)$$

The approximation follows since $\chi \ll 1$.

We now wish to determine an expression for n_t, the total hydrogen density in an absorbing structure. It should be emphasised that n_t is the *real space* hydrogen density. Let Ω_H be the fraction of the critical density contributed by

hydrogen.[2] The background number density of (ionised and neutral) hydrogen atoms at a redshift z, $n_b(z)$, is thus given by:

$$n_b(z) = \frac{\rho_{\rm crit}\Omega_H(1+z)^3}{m_H}, \tag{6.3.3}$$

where m_H is the mass of a hydrogen atom. $\rho_{\rm crit}$ is the density just sufficient to close the universe at the current epoch—see §2.1.1.

Let the mass density in the absorbing structures be ρ_a and the background mass density at a redshift z be $\rho_b(z)$; n_t can be expressed as $n_b(z)\rho_a/\rho_b$. With the above equation and the standard expression for $\rho_{\rm crit}$,

$$n_t = \frac{3H_{\rm o}^2\Omega_H(1+z)^3}{8\pi G m_H}\left(\frac{\rho_a}{\rho_b}\right). \tag{6.3.4}$$

Substituting this expression into equation (6.3.2):

$$\chi = \frac{3H_{\rm o}^2\Omega_H(1+z)^3\alpha_H(T)}{8\pi G m_H J(\nu_c) G_H}\left(\frac{\rho_a}{\rho_b}\right). \tag{6.3.5}$$

Using numerical values where appropriate, together with the analytic approximation for $\alpha_H(T)$ [Black 1981]

$$\alpha_H(T) = 4.4\times10^{-10}T^{-\beta}\,{\rm cm^3\,s^{-1}}, \tag{6.3.6}$$

where $\beta \approx 0.76$, gives our final result:

$$\chi = 1.7\times10^{-6}\Omega_H(1+z)^3h^2\left(\frac{\rho_a}{\rho_b}\right)\left(\frac{T}{10^4\,{\rm K}}\right)^{-\beta}\left(\frac{J(\nu_c)}{10^{-20}\,{\rm erg\,cm^{-2}}}\right)^{-1}. \tag{6.3.7}$$

In this chapter, we shall mainly be concerned with hydrogen in redshift caustics; ρ_a/ρ_b should not exceed a factor of a few. Thus $\chi \ll 1$. This justifies the assumption made above.

---

[2] Unless $\Omega_{\rm o} = 1$, the value of this fraction will change with redshift. All Ω values in this thesis refer to the $z = 0$ value unless otherwise stated.

6.3.2 *Column Densities*

If the lines represent redshift caustics, a substantial fraction of the hydrogen component must be in the caustics. From Figures 5.2a–d, we may estimate the fraction of hydrogen in lines, η, to be somewhere in the range 0.05–0.5. If Δr is the mean physical separation of the caustics, then the total hydrogen column density between caustics is $n_b(z)\Delta r$. Allowing for the ionisation of the hydrogen and the factor η gives the neutral hydrogen column density:

$$N_c(\mathrm{H\,I}) = \eta n_b(z)\Delta r\,\chi. \tag{6.3.8}$$

The number of lines per unit redshift interval is dN/dz. The mean redshift increment between lines, Δz, is thus $(dN/dz)^{-1}$. Using equation (2.1.12), the equivalent physical separation, Δr, is given by

$$\Delta r = \frac{c}{H(z)(1+z)}\left(\frac{dN}{dz}\right)^{-1}. \tag{6.3.9}$$

Together with the expressions for n_b and χ, equation (6.3.8) becomes:

$$N_c(\mathrm{H\,I}) = 5.6\times10^{14}\frac{\eta}{(1+\Omega_o z)^{1/2}}\left(\frac{\rho_c}{\rho_b}\right)\left(\frac{h}{0.5}\right)^3\left(\frac{1+z}{1+2.5}\right)^4\left(\frac{\Omega_H}{0.1}\right)^2 \times\left(\frac{dN/dz}{60}\right)^{-1}\left(\frac{T}{10^4\,\mathrm{K}}\right)^{-\beta}\left(\frac{J(\nu_c)}{10^{-20}\,\mathrm{erg\,cm^{-2}}}\right)^{-1}\mathrm{cm}^{-2}. \tag{6.3.10}$$

Here, ρ_c is the matter density in a caustic. From the work in Chapter 5, $\rho_c/\rho_b \approx (1+f)/f$; $f \approx \Omega^{0.6}$. The factor $\eta(\rho_c/\rho_b)$ must be close to one.

This calculation indicates that the column densities produced in the model are reasonable. Unfortunately, it is not simple to estimate the likely *distribution* of column densities. Although the caustics are thin and two-dimensional, they will not be oriented randomly in space. As we concluded in §5.3, redshift caustics tend to lie perpendicularly to the line of sight. The argument in §6.2.3 that

produced N_c^{-2} is not valid for redshift caustics. Calculation of the distribution of column densities is a topic for the future.

6.3.3 Line Number Density

Consider an $\Omega = 1$ universe where the initial power spectrum of the density field is of the form, $P(k) \propto k^n$. In the model there are initially no characteristic length or time scales. The universe must evolve self-similarly [Davis and Peebles 1977]. However, in Chapter 3 it was argued that non-linear evolution would break the length scale invariance. All lengths must vary as the scale entering the non-linear regime, the only characteristic length scale in the model.

The co-moving length scale becoming non-linear, x, varies with z as [e.g., Kaiser 1986]

$$x \sim (1+z)^{-2/(n+3)}. \tag{6.3.11}$$

The corresponding physical scale, r, is given by $x/(1+z)$:

$$r \sim (1+z)^{-1-2/(n+3)}. \tag{6.3.12}$$

This is related to the line number density, dN/dz, through equation (6.3.9), with $\Delta r = r$. Setting $\Omega_o = 1$ since the similarity solution requires it, $H(z) \sim (1+z)^{3/2}$. Thus

$$\frac{dN}{dz} \sim (1+z)^{2/(n+3)-3/2}. \tag{6.3.13}$$

The observations discussed in §6.2.4 show that dN/dz varies with z as $(1+z)^\gamma$, with $\gamma \approx 2$. Equating the coefficients of $(1+z)$,

$$n = -3 + \frac{4}{2\gamma+3} \approx -2.4 \quad \text{if} \quad \gamma = 2. \tag{6.3.14}$$

The caustics model can thus account for the observations if $n \approx -2.4$.

Using this result, we may write (6.3.12) in the form

$$r = r_o(1+z)^{-5/2-\gamma}. \tag{6.3.15}$$

If the caustics continue to evolve in the same way, r_o should be the current separation of the caustics. Let us assume that we know the value of dN/dz at some redshift, z. Equating the expression for r with equation (6.3.9)

$$r_o = \frac{c(1+z)^\gamma}{H_o}\left(\frac{dN}{dz}\right)^{-1}. \tag{6.3.16}$$

For lines with an equivalent width in excess of 0.32 Å, SYBT found that dN/dz was about 60 at a redshift of 2.4. With these values, r_o has a value of $\sim 600h^{-1}$ Mpc. This is about an order of magnitude larger than the size of the voids in the CfA slice redshift survey. There is a problem if we are to identify quasar absorption lines as the forerunners of the structures in this survey.

Equation (6.3.16) represents a perfectly definite prediction of the current separation of redshift caustics. There are, however, at least two considerations which may substantially alter the value of r_o obtained. (6.3.14) indicates that γ depends sensitively on the value of n, particularly if n is close to -3. If the spectral index is not constant, r_o will be altered. n only has to increase to a value of -1.7 for γ to become zero. This will lower the predicted value of r_o by the required factor.

Alternatively, the distribution of column densities varies as N_c^{-2} over the range 10^{13}–10^{15} atoms cm^{-2}. As discussed in §6.1, the lower limit probably reflects the resolution at which the spectra were obtained. If this form continues to hold down to 10^{12} atoms cm^{-2}, ten times as many lines will be detected by improving the resolution. This will also lower the prediction by the required amount.

We can conclude that the observed evolution of dN/dz can be accounted for in this model. It is also possible to extrapolate the evolution to the current epoch. However, the parameters necessary to determine r_o are too poorly constrained to be of much use. It should be noted that no allowance has been made for the evolution of the u.v. flux in these calculations.

6.3.4 Line of Sight Correlations

In Chapter 3, it was argued that there is no correlation between the spatial postion of peaks in a one-dimensional Gaussian random field. This result should apply whatever the dimensionality. In Chapter 5, Figures 5.2a–d show that redshift caustics occur close to the most dense regions of the evolved field, i.e., the peaks. If we take the initial density perturbation to be a Gaussian random field, there will naturally be no correlation between the *initial* positions of the caustics.

In Chapter 4, it was argued that the present galaxy-galaxy correlation function is dominated by highly non-linear regions. Redshift caustics, however, occur on scales which are just becoming non-linear. Non-linear evolution cannot cause the caustics to be clustered. In this model, the line of sight correlation function will be naturally flat. One of the strongest predictions of the model is that there should be no clustering of the Ly-α lines.

6.3.5 Transverse Sizes

Figures 5.2a–d show that the transverse size of a caustic is roughly the same as the inter-caustic separation. We may thus estimate the transverse physical size

of a caustic through equation (6.3.9), with $\Delta r = r_t$. Normalising this equation in terms of the SYBT values: $dN/dz \sim 60$ at $z = 2.5$,

$$r_t = \frac{4.1}{(1+z\Omega_o)^{1/2}} \left(\frac{1+z}{1+2.5}\right)^{-2} \left(\frac{dN/dz}{60}\right)^{-1} h^{-1}\,\mathrm{Mpc}. \qquad (6.3.17)$$

$4.1h^{-1}$ Mpc is clearly much larger than the minimum transverse size implied by the observations on the double quasar (§6.2.7).

As noted in §6.3.3, dN/dz may depend significantly on the resolution at which the spectra were obtained. Since only more lines can be detected as resolution improves, the above estimate of r_t is an upper limit. If lines with column densities as low as 10^{12} atoms cm^{-2} could be detected, r_t may decrease by a factor of ten. However, if the characteristic transverse size of clouds is determined observationally to be as low as 10 kpc, it will be difficult to reconcile with the above estimate. An accurate determination of r_t may succeed in ruling out the caustics model.

6.4 Synthetic Quasar Spectra

Perhaps the most interesting feature of the caustics model is that it is possible to synthesise quasar spectra. Any of the properties discussed in §6.1 can in principle be determined. This predictive power is unique amongst models of the structures causing the Ly-α lines. This section presents the formulae necessary to calculate the synthetic spectra. They are applied to two simple one-dimensional systems.

The optical depth corresponding to a unform distribution of H I was calculated in §6.2.8. In performing the integral in equation (6.2.8), the function ϕ was treated as a δ-function. This approximation was valid in the context of the Gunn-Peterson calculation since we were considering a uniformly distributed H I

component. The variation of $\check{n}(z)$ with z is then much slower than the variation of ϕ. When calculating a synthetic spectrum, however, this approximation is no longer sufficient. $\check{n}(z)$ may vary quite as rapidly as ϕ.

On the other hand, dr/dz will still vary much more slowly than ϕ. To a good approximation, dr/dz may be taken outside the integral and evaluated at a redshift satisfying $(1+z)\nu_{\rm obs} = \nu_\alpha$. Equation (6.2.8) may thus be expressed as

$$\tau(\nu_{\rm obs}) = \sigma_{\rm o} \frac{dr}{dz} \int_0^{z_{\rm em}} \check{n}(z') \phi\big((1+z')\nu_{\rm obs} - \nu_\alpha\big)\, dz'. \tag{6.4.1}$$

Let us transform to a new variable $\nu = \nu_{\rm obs}(1+z')$. Defining $\nu_{\rm em} = \nu_{\rm obs}(1+z_{\rm em})$,

$$\tau(\nu_{\rm obs}) = \frac{\sigma_{\rm o}}{\nu_{\rm obs}} \frac{dr}{dz} \int_{\nu_{\rm obs}}^{\nu_{\rm em}} \check{n}(\nu) \phi(\nu - \nu_\alpha)\, d\nu. \tag{6.4.2}$$

$\check{n}(\nu)$ is the number density of hydrogen atoms in *redshift* space. We may express this as

$$\check{n}(\nu) = \check{n}_t(\nu)\chi, \tag{6.4.3}$$

where $\check{n}_t(\nu)$ is the total hydrogen density in redshift space. $\check{n}_t(\nu)$ may be calculated in a similar fashion to n_t, but with ρ_a replaced by $\check{\rho}_\nu$. $\check{\rho}_\nu$ is the total hydrogen density in redshift space at the redshift corresponding to ν. Thus, using equation (6.3.4),

$$\check{n}(\nu) = \frac{3H_{\rm o}^2 \Omega_H (1+z)^3}{8\pi G m_H} \left(\frac{\check{\rho}_\nu}{\rho_b}\right) \chi. \tag{6.4.4}$$

Using the expression (6.3.7) for χ:

$$\begin{aligned} \check{n}(\nu) = 1.9 \times 10^{-11} \Omega_H^2 h^4 (1+z)^6 \left(\frac{\rho_\nu}{\rho_b}\right) \left(\frac{\check{\rho}_\nu}{\rho_b}\right) \\ \times \left(\frac{T}{10^4\ {\rm K}}\right)^{-\beta} \left(\frac{J(\nu_c)}{10^{-20}\ {\rm erg\ cm^{-2}}}\right)^{-1} {\rm cm}^{-3}. \end{aligned} \tag{6.4.5}$$

Here, ρ_a in the original expression has been replaced by ρ_ν, the real space matter density evaluated at the redshift corresponding to ν. Substituting this expression into (6.4.2) and using equation (2.1.12) for dr/dz,

$$\tau(\nu_{\rm obs}) = \frac{0.8\Omega_H^2 h^3 (1+z)^5}{(1+\Omega_{\rm o} z)^{1/2}} \left(\frac{T}{10^4\,{\rm K}}\right)^{-\beta} \left(\frac{J(\nu_c)}{10^{-20}\,{\rm erg\,cm^{-2}}}\right)^{-1} \times \int_{\nu_{\rm obs}}^{\nu_{\rm em}} \left(\frac{\rho_\nu}{\rho_b}\right) \left(\frac{\check{\rho}_\nu}{\rho_b}\right) \phi(\nu - \nu_\alpha)\, d\nu. \tag{6.4.6}$$

The integral depends on $\nu_{\rm obs}$ through the variable ν.

The above expression may be simplified somewhat by writing

$$\tau(\nu_{\rm obs}) = Q \int_{\nu_{\rm obs}}^{\nu_{\rm em}} \left(\frac{\rho_\nu}{\rho_b}\right) \left(\frac{\check{\rho}_\nu}{\rho_b}\right) \phi(\nu - \nu_\alpha)\, d\nu \tag{6.4.7}$$

where

$$Q = 0.3 \left(\frac{\Omega_H}{0.1}\right)^2 \left(\frac{h}{0.5}\right)^3 \left(\frac{1+z}{1+2.5}\right)^5 \times \left(\frac{T}{10^4\,{\rm K}}\right)^{-\beta} \left(\frac{J(\nu_c)}{10^{-20}\,{\rm erg\,cm^{-2}}}\right)^{-1} \left(\frac{1+\Omega_{\rm o} z}{1+1\times 2.5}\right)^{-1/2}. \tag{6.4.8}$$

The expression for the multiplying factor, Q, is quite complicated. Moreover, there is a strong dependence on some of the factors. Q is best reguarded as a poorly constrained parameter of the theory, probably lying in the range 0.03–3.0.

6.4.1 Synthetic Spectra for One-Dimensional Systems

There are two reasons for initially simulating spectra for only one-dimensional systems. Primarily, the computations are simplified considerably. However, there is a secondary consideration. Rees [private communication] has suggested that the lines in the simulated spectra will not have both large optical depths and small velocity dispersions. The lines produced by a one-dimensional system may not reproduce the observed velocity dispersions and equivalent widths. There would then be little point in extending the calculation to three dimensions.

It will prove convenient to plot spectra against a velocity, rather than a frequency. To define a suitable velocity co-ordinate, consider hydrogen atoms from the viewpoint of a rest frame at a redshift z. To first order in v/c, atoms moving with a velocity v along the line of sight will appear to absorb photons at a frequency $\nu_\alpha(1 - v/c)$. This corrsponds to an observed frequency at $z = 0$ of $\nu_\alpha(1 - v/c)(1 + z)^{-1}$. We are thus led to define a velocity co-ordinate, v, by

$$\nu_{\rm obs} = \frac{\nu_\alpha}{(1+z)}\,(1 - v/c)\,. \tag{6.4.9}$$

With this definition, it is straightforward to show that:

$$\tau(v) = Q \int_{-\infty}^{\infty} \left(\frac{\rho_{v'}}{\rho_b}\right)\left(\frac{\check{\rho}_{v'}}{\rho_b}\right) \Phi(v' - v)\, dv', \tag{6.4.10}$$

where $\Phi(v' - v) = (\nu_\alpha/c)\phi\big(\nu(v') - \nu(v)\big)$. The limits of the integral have been taken to be at $\pm\infty$ since there is little contribution to the integral unless v' is close to v.

We shall assume that the line profile function is determined by Doppler broadening only. With this assumption, Φ takes the form, [Mihalas 1978 §9-2]

$$\Phi(v) = \frac{1}{\pi^{1/2} v_{\rm o}} \exp\left(-\frac{v^2}{v_{\rm o}^2}\right), \tag{6.4.11}$$

where $v_{\rm o} = (2kT/m_H)^{1/2}$. T is the temperature of the hydrogen in the absorbing system.

Before we apply these equations, let us consider the velocity scale of the structures we are trying to explain. If the mean physical separation of the structures is Δr, the mean velocity separation, Δv is $H(z)\Delta r$. Using equation (6.3.9) we may write:

$$\Delta v = \frac{5000}{(1+z)} \left(\frac{dN/dz}{60}\right)^{-1} \ {\rm km\,s^{-1}}. \tag{6.4.12}$$

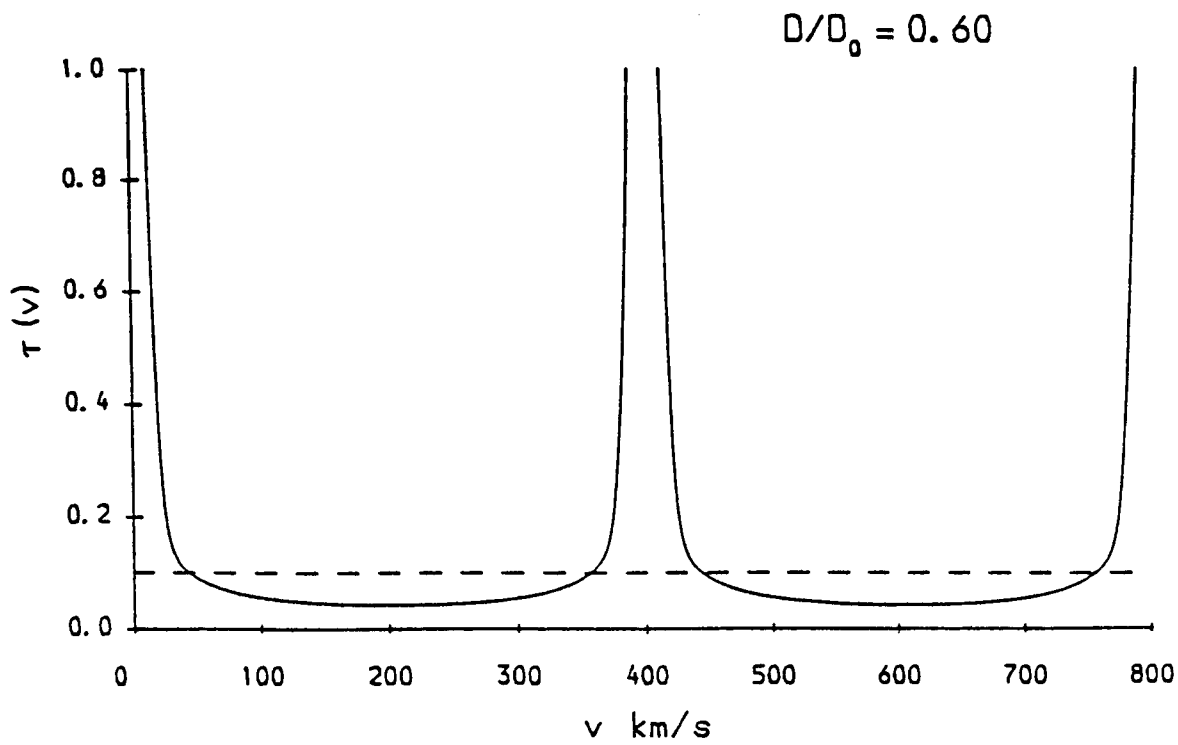


Figure 6.2. The optical depth as a function of v, calculated for an initial perturbation $\cos kx$. The dashed line indicates the constraint imposed by the Gunn-Peterson test. Futher details are provided in the text.

At a redshift of 2.5, $\Delta v \sim 1400\,\mathrm{km\,s^{-1}}$. However, this is the mean separartion in a three-dimensional system. We might expect, say, a third of all possible caustics to be suitably inclined to the line of sight. Since all structures in a one-dimensional system must be suitably inclined, we should take $\Delta v \sim 400\,\mathrm{km\,s^{-1}}$.

As a first application of these equations, let us consider a one-dimensional system with an initial density perturbation given by $\cos kx$. k should be chosen to correspond to a wavelength of $400\,\mathrm{km\,s^{-1}}$. Ω will be taken to be 1 throughout. Figure 6.2 shows the optical depth as a function of v, calcuated using (6.4.10). Q and T were chosen to be 0.15 and $10^4\,\mathrm{K}$ respectively. The factors $\rho_{v'}/\rho_b$ and $\check{\rho}_{v'}/\rho_b$ were calculated using the Zel'dovich approximation. The linear growth factor, D, was set to a fraction 0.6 of D_o. D_o is the growth factor which would cause the first *real* space caustics to form. With this value of D, the maximum value of $\delta\rho/\rho$ will not exceed 1.5.

The figure indicates that the lines *are* both narrow and have significant optical depth. The dashed line represents an optical depth of 0.1. Between the

lines, the optical depth is substantially less than this. The constraint imposed by the Gunn-Peterson test is satisfied.

To proceed futher, we need to know several observational details. For the highest resolution spectra obtained [Carswell, private communcation]:

(i) The full-width half-maximum resolution of the spectrometer is $20\,\mathrm{km\,s^{-1}}$.

(ii) The spectra are sampled in wavelength bins of width 0.07 Å. At a redshift of 2.5, the corresponding velocity width is $5\,\mathrm{km\,s^{-1}}$.

(iii) The number of photons per bin is typically fifty for the continuum.

For the moment, we shall ignore Poisson noise. Let us define the number of photons in the bin correponding to a velocity v to be $F(v)$. We may write:

$$F(v) = F_{\mathrm{o}}\langle \exp\left(-\tau(v)\right)\rangle, \qquad (6.4.13)$$

where F_{o} is the number of photons per bin in the continuum. $\langle \ldots \rangle$ implies smoothing over the resolution of the instrument. Once $F(v)$ has been calculated, it is simple to add Poisson noise.

Figure 6.3 shows the spectra calculated from the optical depth profile in Figure 6.2. F_{o} was taken to be fifty. The FWHM instumental resolution and the bin widths were taken to be $20\,\mathrm{km\,s^{-1}}$ and $5\,\mathrm{km\,s^{-1}}$ respectively. The figure shows remarkably thin lines—the FWHM are only $\sim 40\,\mathrm{km\,s^{-1}}$. The corresponding velocity disperson parameter is $\sim 24\,\mathrm{km\,s^{-1}}$. If the spectrum represents absorption at a redshift of 2.5, the equivalent widths of the lines are 0.45 Å.[3] These values compare vary favourably with those observed.

[3] For computational simplicity, these values were obtained from the line profiles *before* the Poisson noise was added. Since adding noise should only increase the error in the measurement, this should make no essential difference.

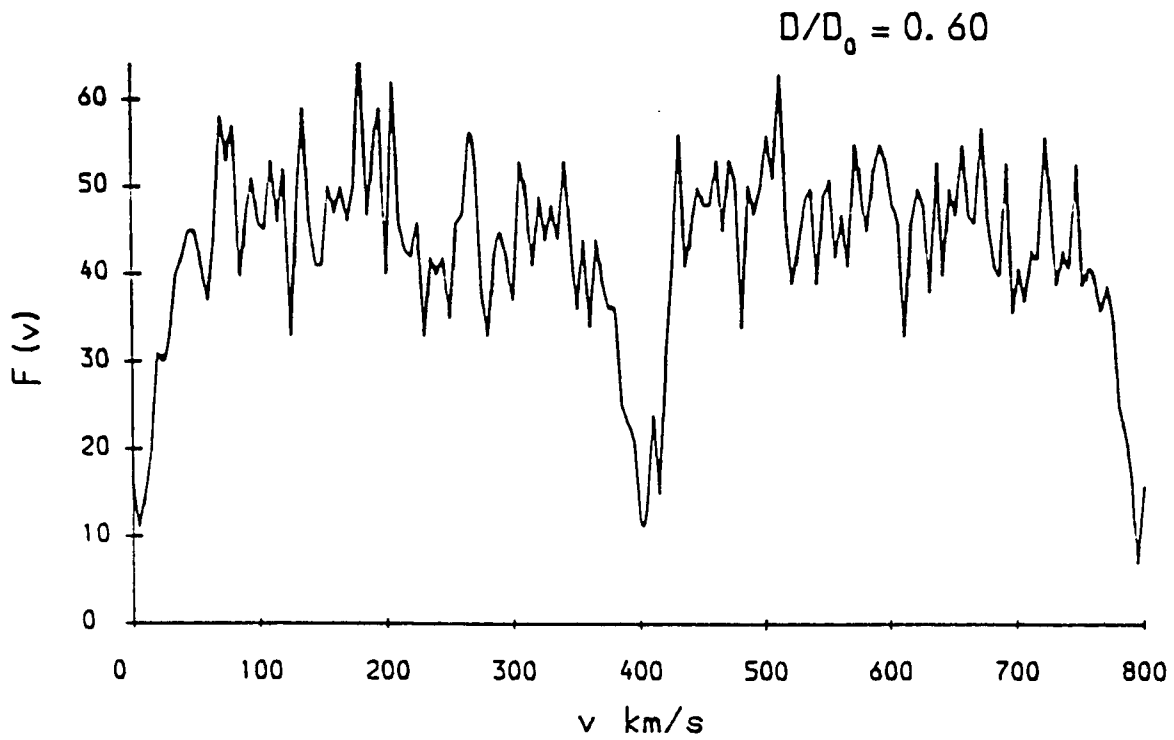


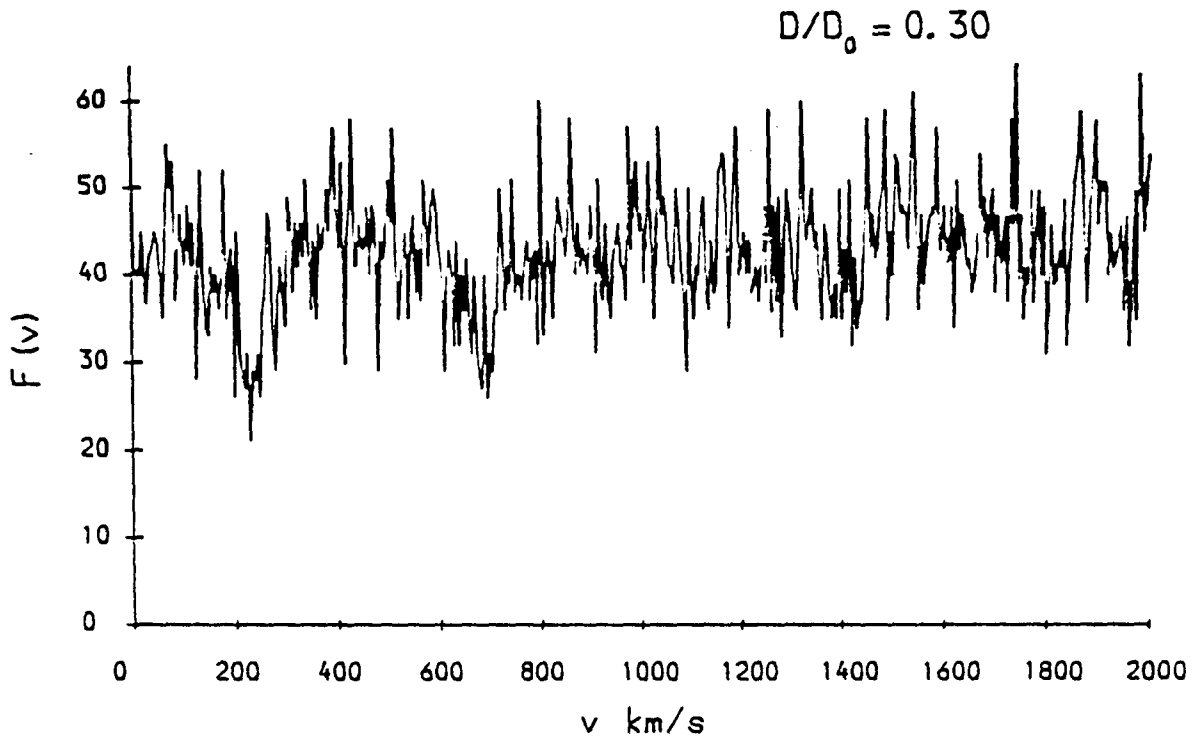
Figure 6.3. Simulated spectrum, calculated using the optical depth profile shown in Figure 6.2.

Our final example will involve a more complicated initial perturbation. We will take this to be a Gaussian random field with a power spectrum given by

$$P(k) \sim kW_{FD}(k/k_c; 0.1). \tag{6.4.14}$$

The Fermi-Dirac cut-off was defined in §3.1.1. The wavelength of k will be chosen to correspond to $400\,\mathrm{km\,s^{-1}}$. Figures 6.4a–d show the resulting spectra for various values of D/D_o. It should be emphasised that this sequence does not represent time evolution. The velocity scale is *not* co-moving. Q, T and the observational parameters were the same as for Figure 6.3.

Two absorption features are clearly visible in these spectra. They occur at velocity co-ordinates of 260 and $650\,\mathrm{km\,s^{-1}}$. In the last figure of the sequence, the velocity dispersion parameters are 27 and $24\,\mathrm{km\,s^{-1}}$ respectively. At a redshift of 2.5, the equivalent widths correspond to 0.53 Å and 0.42 Å. The calculated velocity dispersions and equivalent widths agree well with those observed—see §6.1.



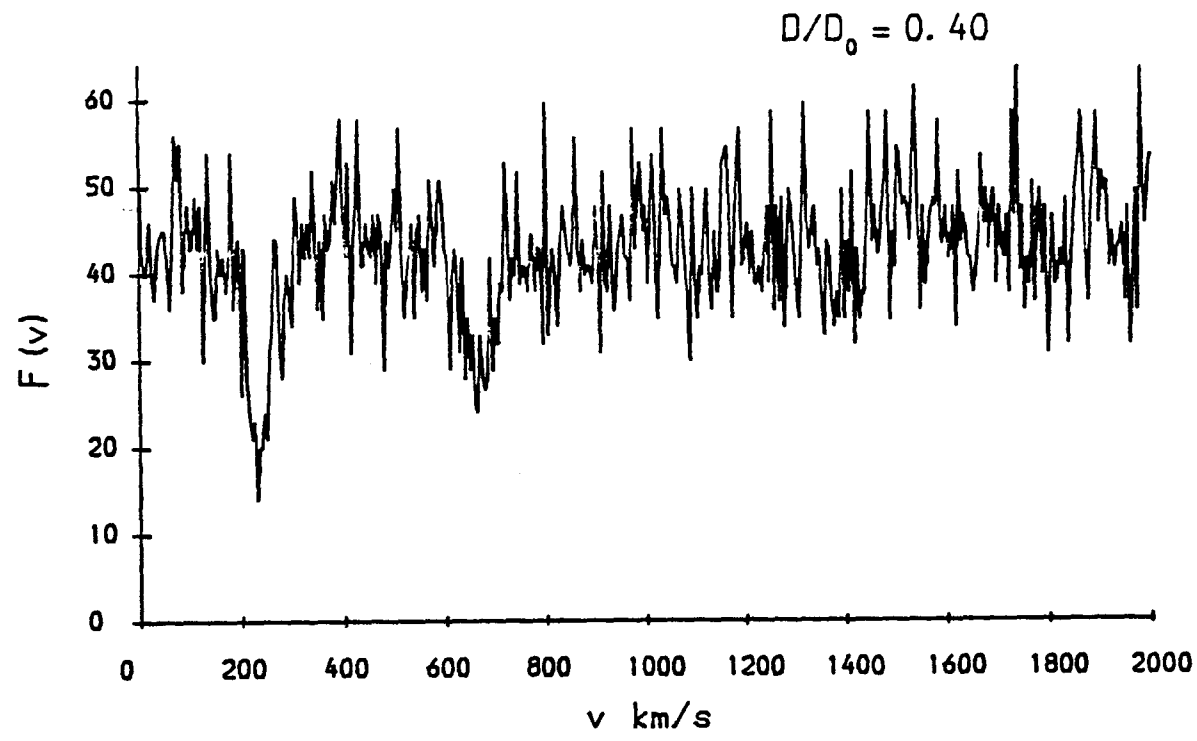


Figure 6.4a–b. Simulated spectra for the sytem discussed in the text.

6.5 Conclusions

The majority of this chapter has been concerned with the proposal that quasar Ly-α absorption lines represent redshift caustics. Within the larger context of the hierarchical scenario, this proposal is entirely natural. §§6.3—6.4 demonstrated that many of the line properties discussed in §6.2 can be accounted for by this model. In particular, together with the conclusions of Chapter 3, the model strongly suggests that the lines should not be clustered. The synthetic

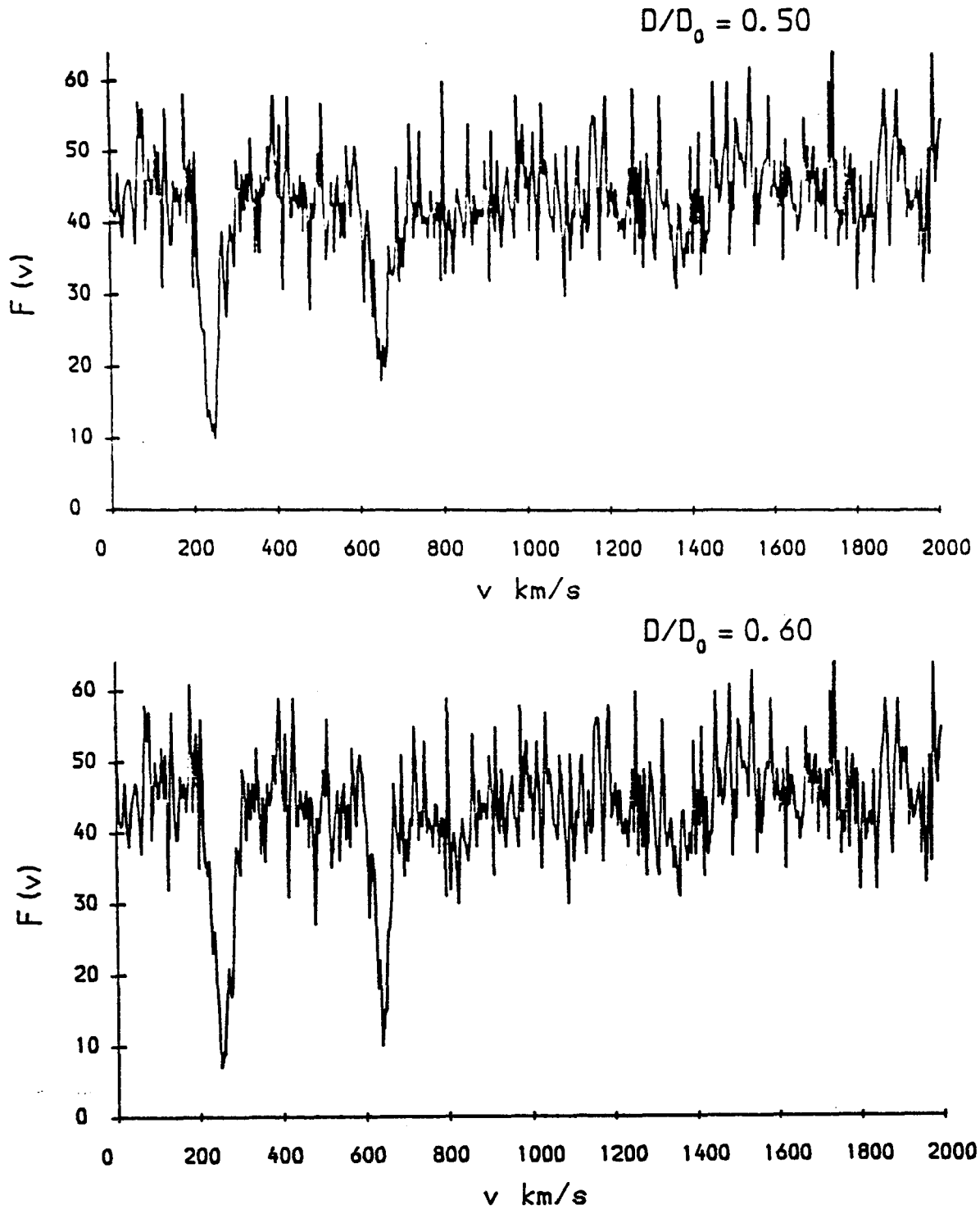


Figure 6.4c–d. Simulated spectra for the sytem discussed in the text.

spectra generated in the previous section indicate that the model produces lines of roughly the correct equivalent width and velocity dispersion. Moreover, the optical depth of the neutral hydrogen between lines satisfies the Gunn-Peterson test. It should be noted that *all* the model parameters were suggested by observational considerations of one sort or another. For the model to have performed as well as it has with no real free parameters is very encouraging.

There are some problems with the model however. The estimated trans-

verse size of the structures is probably too large by an order of magnitude. Simulating quasar spectra for a three-dimensional system will produce a more accurate estimate. Considering the way that the estimate was made in §6.3.5, the more accurate estimate will almost certainly be lower. It should even be possible to simulate spectra corresponding to the light paths of the quasar pair 2345+007A,B. A direct comparison with Foltz *et al.*'s [1984] observations can be made. There is also the possibility, discussed in §6.3.3, that more lines will be observed as the resolution of the spectra improves. This will also reduce the estimate of the transverse size. Whether these two considerations can combine to reduce the estimated size sufficiently remains to be seen.

For the model to be successful, the hydrogen lying between us and the quasar must be highly ionised. There are growing indications, both theoretical [Shapiro 1986] and observational [Carswell *et al.* 1987; §6.2.9] that quasars may not be able to produce the necessary u.v. flux. However, we know that the hydrogen must be ionised. The directly observed (neutral) hydrogen contributes less than one part in a million to the closure density. A suitable ionising source *must* exist if current cosmological thinking is not completely awry. The problem is that unless we know the source of the ionistion, the magnitude of the u.v. flux cannot be estimated. Without this knowledge, it is impossible to make detailed predictions, particularly of the evolution of line number density. The predictive power of this model will be severly resticted.

In conclusion, a strong case can be made for the proposal that Ly-α lines are redshift caustics. If the model is validated, it will open the way to an exciting and immeadiate probe of the distribution of matter at high redshifts.

Chapter 7. Conclusions and Future Work

Conclusions have been presented at the end of every chapter. What follows is more of a summary than a conclusion.

Chapter 3 was devoted to understanding initial perturbations that are examples of Gaussian random fields. Particular emphasis was placed on power law fields. As a step towards understanding the full three-dimensional problem, one-dimensional fields were studied in some depth, particularly the properties of peaks. It was demonstrated that the positions of peaks are determined by the phases of the Fourier components close to the high frequency cut-off. The peaks are thus spatially uncorrelated, except for a tendency to occur with a separation equal to the wavelength of the cut-off frequency, k_c. The heights of neighbouring peaks were shown to be uncorrelated unless the spectral index, n, is close to $-d$. Close probably means $-d+1 < n < -d$. It was also demonstrated that structures appear to be forming over a wide range of scales. This was interpreted as a constraint on the effective spectral power index, n_{eff}. n_{eff} must be close to -3.

Chapter 4 was mainly concerned with correlation functions. It was argued that ξ_{gg}, ξ_{cc} and ξ_{ss} do not necessarily probe the same density field. Using clusters as probes rather than galaxies implies a larger smoothing volume. It was suggested that the correlation length, r_o for a particular set of objects should increase with the mean separation of the objects, $\bar{r}$. Observations indicate that $r_o \propto \bar{r}$. It was futher argued that this arises naturally if the initial density perturbation was an $n = -3$ power law field.

The remainer of the chapter presented a model of the galaxy-galaxy correlation function. It was argued that galaxies in non-linear regions dominate ξ_{gg}.

This led to the proposal that the form of ξ_{gg} was determined by galaxies in clusters. It was demonstrated that this proposal is entirely consistent with observations. This then implies that violent relaxation is responsible for the form of the galaxy-galaxy correlation function. Little information about the initial density pertubations is contained in ξ_{gg}. Since ξ_{cc} and ξ_{ss} appear to take the same form as ξ_{gg}, this proposal has significant implications. Clusters and super-clusters may also occur in relaxed structures, implying collapse on an impressively large scale.

Chapter 5 examined some of the effects of using velocity as a distance measure. It was demonstrated that density contrasts are significanly enhanced in redshift space if structures have grown from small perturbations. In particular, caustics form in redshift space whilst the associated real space $\delta\rho/\rho \sim 1$. It was argued that redshift caustics should continually form in the hierarchical scenario on the scale entering the non-linear regime. It was noted that the effects of collapse on smaller scales would cause noise in the velocity field which would broaden the caustics. However, since sharp features are observed in redshift space, this broadening must be unimportant. If we are to explain these structures within the hierarchical scenario, broadening must also be unimportant in these models. In the concluding section to this chapter, expressions for several velocity correlation functions were derived using linear theory. It was demonstrated that the predictions of linear theory are totally incompatible with observations. Linear theory is totally inappropriate for the study of correlation functions.

In Chapter 6, the possiblity that various structures were redshift caustics was examined. The majority of the chapter was concerned with the proposal that quasar Ly-α absorption lines represent redshift caustics. Evidence for caustics in galaxy redshift surveys was inconclusive. It was demonstrated that many of

the observed properties of the Ly-α lines can be accounted for by this model. In particular, the lines should *not* be clustered. There is a stong case to be made for considering the structures causing Ly-α lines to be redshift caustics.

In a sense, little of the work presented in this thesis is complete. Whilst most of the concepts are equally applicable to systems of any dimensionality, few of the calculations have been in three dimensions. It is obviously necessary to extend some of the calculations. The author stongly believes, however, that few, if any, of the conclusions presented above will need to be changed.

There are at least two important proposals arising from this work. The first concerns the initial density perturbation field. To account for the observed large-scale distribution of matter it may be appropriate to consider the initial pertubation to be an $n = -3$ power law field. The second is that quasar absorption lines represent redshift caustics. These two possiblities merit consideration in greater depth. Both have important implications for our understanding of the large-scale structure of the universe.

'Stuff and Nonsense!', said Alice.

[Caroll 1868]

Appendix: Angular Correlation Functions

The purpose of the appendix is to determine an analytic expression for the functions I_ν which were defined in Chapter 5. As we shall see, the I_ν are two-point angular cross-correlation functions. As a first step in the calcualtion, we will determine a general expression for such functions.

A.1 Notation and Definitions

Many of the formulae in this appendix concern rotations. We shall specify a particular rotation in terms of Euler angles [e.g., Rose 1957 p.50]. Let us define α to be the rotation angle about the original z-axis, β the angle about the *new* y-axis and γ the angle about the *new, new* z-axis. It will be convenient to refer to the Euler angles through the triplet $\omega \equiv (\alpha, \beta, \gamma)$. The symbol

$$\int_\Omega X(\omega)\, d\omega$$

denotes the integral of $X(\omega)$ over the space of all rotations.

Continuing in the same vein, it will frequently be necessary to specify a vector in terms of its spherical polar co-ordinates. If the vector is $\mathbf{r}$ then we shall denote them: $(r, \theta_\mathbf{r}, \phi_\mathbf{r})$. As a convenient shorthand, the angular co-ordinates will be referred to through the doublet $\Omega_\mathbf{r} = (\theta_\mathbf{r}, \phi_\mathbf{r})$.

To conclude this section, we will consider the transformation of a vector, $\mathbf{r}$ under a rotation, ω. In general, $\mathbf{r}$ will be mapped to a new vector, $\mathbf{r}'$. We may write $\mathbf{r}'$ in terms of the rotation operator, $\mathbf{R}(\omega)$:

$$\mathbf{r}' = \mathbf{R}(\omega)\mathbf{r}. \tag{A.1.1}$$

A.1.1 *Properties of Sperical Harmonics*

Many of the following calculations depend on the properties of sperical harmonics, $Y_{lm}(\Omega)$. Those needed are set out below. Derivations of all of these properties can be found in *The Elementary Theory of Angular Momentum* by Rose [1957].

PROPERTY 1: *Value at the poles.*

The value of $Y_{lm}(0,0)$ is given by [Rose 1957 p.60]

$$Y_{lm}(0,0) = \left(\frac{2l+1}{4\pi}\right)^{1/2} \delta_{m,0}, \tag{A.1.2}$$

where $\delta_{i,j}$ is the Kronecker delta.

PROPERTY 2: *The Addition Theorem.*

It may be shown that [Rose 1957 p.60]

$$\sum_m Y_{lm}(\Omega_{\mathbf{q}}) Y^*_{lm}(\Omega_{\mathbf{r}}) = \left(\frac{2l+1}{4\pi}\right) P_l(\cos\psi). \tag{A.1.3}$$

Here, P_l is the Legendre polynomial of degree l. $\cos\psi$ is the angle between $\mathbf{q}$ and $\mathbf{r}$.

PROPERTY 3: *Behaviour under a rotation.*

Under a rotation, ω, the spherical harmonics transform according to [Rose 1957 p.52]

$$Y_{lm}(\Omega_{\mathbf{r}'}) = \sum_{m'} R^l_{mm'}(\omega) Y_{lm'}(\Omega_{\mathbf{r}}). \tag{A.1.4}$$

$R^l_{mm'}(\omega)$ are the rotation matrices for the rotation ω.

PROPERTY 4: *Orthogonality of the rotation matrices.*

It may be shown that [Rose 1957 p.75]

$$\int_\Omega R^l_{mm'}(\omega) R^{L*}_{MM'}(\omega)\, d\omega = \frac{8\pi^2}{2l+1} \delta_{l,L} \delta_{m,M} \delta_{m',M'}. \tag{A.1.5}$$

A.2 The Two-Point Angular Cross-Correlation Function

We wish to calculate the two point angular cross-correlation function, κ, of the functions $f(\mathbf{q})$ and $g(\mathbf{r})$. κ may be defined as

$$\kappa = \langle f(\mathbf{q}) g^*(\mathbf{r}) \rangle, \tag{A.2.1}$$

where $\langle \ldots \rangle$ implies averaging over all directions of the vectors $\mathbf{q}$ and $\mathbf{r}$ such that the angle between them, ψ, remains constant. We may evaluate this average in terms of an integral over rotations. Using the rotation operator, $\mathbf{R}(\boldsymbol{\omega})$,

$$\kappa = \frac{1}{8\pi^2} \int_\Omega f\big(\mathbf{R}(\boldsymbol{\omega})\mathbf{q}\big) g^*\big(\mathbf{R}(\boldsymbol{\omega})\mathbf{r}\big)\, d\boldsymbol{\omega}. \tag{A.2.2}$$

In this expression, $\mathbf{q}$ and $\mathbf{r}$ represent the initial positions of the vectors, i.e., when $\boldsymbol{\omega} = 0$. The normalising term is the value of

$$\int_\Omega d\boldsymbol{\omega} = \int_0^{2\pi} d\alpha \int_0^{\pi} \sin\beta\, d\beta \int_0^{2\pi} d\gamma \equiv 8\pi^2. \tag{A.2.3}$$

To proceed futher, let us expand f in terms of spherical harmonics. Defining $\mathbf{q}' = \mathbf{R}(\boldsymbol{\omega})\mathbf{q}$,

$$f(\mathbf{q}') = \sum_{l,m} F_{lm}(q) Y_{lm}(\Omega_{\mathbf{q}'}). \tag{A.2.4}$$

We have written F_{lm} as a function of q since $|\mathbf{q}| = |\mathbf{q}'|$. Under rotations, spherical harmonics transform according to property 3. Thus we may write

$$Y_{lm}(\Omega_{\mathbf{q}'}) = \sum_{m'} R^l_{mm'}(\boldsymbol{\omega}) Y_{lm'}(\Omega_{\mathbf{q}}). \tag{A.2.5}$$

Substituting this expression into (A.2.4),

$$f(\mathbf{q}') = \sum_{l,m,m'} F_{lm}(q) R^l_{mm'}(\boldsymbol{\omega}) Y_{lm'}(\Omega_{\mathbf{q}}). \tag{A.2.6}$$

A similar expression may be formed for $g^*(\mathbf{r}')$.

Substituting the expansions of f and g into equation (A.2.2),

$$\kappa = \frac{1}{8\pi^2} \sum_{l,m,m'} \sum_{L,M,M'} F_{lm}(q) G^*_{LM}(r) Y_{lm'}(\Omega_{\mathbf{q}}) Y^*_{LM'}(\Omega_{\mathbf{r}}) \times \int_\Omega R^l_{mm'}(\omega) R^{L^*}_{MM'}(\omega)\, d\omega. \tag{A.2.7}$$

Using property 4, this expression reduces to

$$\kappa = \sum_{l,m} \frac{F_{lm}(q) G^*_{lm}(r)}{2l+1} \sum_{m'} Y_{lm'}(\Omega_{\mathbf{q}}) Y^*_{lm'}(\Omega_{\mathbf{r}}). \tag{A.2.8}$$

We may now use the addition theorem for spherical harmonics, property 2. This leaves us with a final expression for κ:

$$\kappa(q, r, \cos\psi) = \frac{1}{4\pi} \sum_{l,m} F_{lm}(q) G^*_{lm}(r) P_l(\cos\psi), \tag{A.2.9}$$

where $\cos\psi$ is the angle between the two vectors.

A.2.1 The Calculation of the I_ν

It is evident from the definitions of κ and I_ν, equations (A.2.1) and (5.5.10), that the I_ν are angular cross-correlation functions. We may identify

$$f(\mathbf{q}) = e^{i\mathbf{k}\cdot\mathbf{q}} \tag{A.2.10}$$

and

$$g(\mathbf{r}) = \cos^{2\nu}\theta, \tag{A.2.11}$$

where θ is the angle between $\mathbf{r}$ and $\mathbf{k}$. To apply the above results, it only remains to determine the expansions of f and g in terms of spherical harmonics.

The expansion of $e^{i\mathbf{k}\cdot\mathbf{q}}$ is given by [e.g., Rose 1957 p.99]:

$$e^{i\mathbf{k}\cdot\mathbf{q}} = 4\pi \sum_{l,m} i^{-l} j_l(kq) Y_{lm}(\Omega_{\mathbf{q}}) Y^*_{lm}(\Omega_{\mathbf{k}}). \tag{A.2.12}$$

j_l is the spherical Bessel function of order l. Since we are averaging over directions, we may choose $\mathbf{k}$ to lie along the z-axis: $\Omega_{\mathbf{k}} = (0,0)$. Using property 1,

$$F_{lm}(q) = \{(2l+1)4\pi\}^{1/2}\, i^{-l}\, j_l(kq)\delta_{m,0}. \tag{A.2.13}$$

The coefficients G_{lm} in the expansion of $\cos^{2\nu}$ are zero unless l is even and no greater than 2ν. If these conditions are satisfied,

$$G_{lm} = \{(2l+1)\pi\}^{1/2}\, \frac{2^{l+1}2\nu!(\nu+l/2)!}{(2\nu+l+1)!(\nu-l/2)!}\delta_{m,0}. \tag{A.2.14}$$

To take account of the odd/even nature of G_{lm}, it will be convenient to define L such that $l = 2L$.

Substituing these expressions into (A.2.9),

$$\kappa(q,r,\cos\psi) = \sum_{L=0}^{\nu}(4L+1)(-1)^L\, j_{2L}(kq)\frac{2^{2L}(2\nu)!(\nu+L)!}{(2\nu+2L+1)!(\nu-L)!}P_{2L}(\cos\psi). \tag{A.2.15}$$

It should be noted that this expression is independent of r. In terms of the original variables (refer to Chapter 5), and replacing L by l, we may write

$$I_\nu(k\check{x}_\perp, k\check{x}_\parallel) = \sum_{l=0}^{\nu}(4l+1)(-1)^l j_{2l}(k\check{x})\frac{2^{2l}(2\nu)!(\nu+l)!}{(2\nu+2l+1)!(\nu-l)!}P_{2l}(\cos\psi). \tag{A.2.16}$$

References

Aarseth, S.J. (1984), in *Methods of Computational Physics*, ed. Brackbill J.U and Cohen B.I. (New York: Academic) p.1.

Aarseth, S.J., Gott, J.R. and Turner, E.L. (1979) *Astrophys. J.*, **236**, 43.

Abell, G.O. (1958) *Astrophys. J. Suppl.*, **3**, 211.

Adler, R.J. (1981) *The Geometry of Random Fields* (Chichester: J. Wiley).

Allen, C.W. (1973) *Astrophysical Quantities* (London: Althone Press).

Apian, P. (1593) *Cosmographia.*

Appel, A. (1985) *SIAM J. Sci. statist. Comput.*, **6**, 85.

Atwood, B., Baldwin, J.A. and Carswell, R.F. (1985) *Astrophys. J.*, **292**, 58.

Bahcall, N.A. and Burgett, W.S. (1986) *Astrophys. J. Lett.*, **300**, 35.

Bahcall, J.N. and Salpeter, E.E. (1965) *Astrophys. J.*, **142**, 1677.

Bahcall, N.A. and Soneira, R.M. (1982) *Astrophys. J.*, **262**, 419.

Bahcall, N.A. and Soneira, R.M. (1983) *Astrophys. J.*, **270**, 20.

Bahcall, J.N. and Spitzer, L. (1969) *Astrophys. J. Lett.*, **156**, 63.

Barcons, X. and Fabian, A.C. (1987) *Mon. Not. Roy. Astron. Soc.*, **224**, 675.

Bardeen, J.M., Bond, J.R., Kaiser, N. and Szalay, A.S. (1986) *Astrophys. J.*, **304**, 15.

Barnes, J, and Hut, P. (1986) *Nature*, **324**, 446.

Barrow, J.D. and Bhavsar, S.P. (1987) *Quarterly Journal Roy. Astron. Soc.*, **28**, 109.

Bean, A.J., Efstathiou, G., Ellis, R.S., Peterson, B.A. and Shanks, T. (1983) *Mon. Not. Roy. Astron. Soc.*, **205**, 605.

van den Berg, S. (1961) *Publ. Astron. Soc. Pacific*, **73**, 46.

Binney, J.J. and Tremaine, S. (1987) *Galactic Dynamics* (Princeton: Princeton University Press).

Black, J.H. (1981) *Mon. Not. Roy. Astron. Soc.*, **197**, 553.

Blumenthal, G.R., Faber, S.M., Primack, J.R. and Rees, M.J. (1984) *Nature*, **311**, 517.

Bogart, R.S. and Wagoner, R.V. (1973) *Astrophys. J.*, **181**, 609.

Bond, J.R. and Efstathiou, G. (1984) *Astrophys. J. Lett.*, **285**, 45.

Bouchet, F.R., Adam, J.-C., and Pellat, R. (1985) *Astron. Astrophys.*, **144**, 413.

Brandenberger, R.H. (1985) *Rev. Mod. Phys.*, **57**, 1.

Carlberg, R. (1987), in *Kinematics and Dynamics of Elliptical Galaxies*, ed. de Zeeuw, P.T. Proceedings of the IAU Symposium 127 (Dortrecht: Reidel).

Caroll, L. (1868) *Alice's Adventures in Wonderland.*

Carswell, R.F. Webb, J.K., Baldwin, J.A. and Atwood, B. (1987) *Astrophys. J.*, **319**, 709.

Carr, B.J., Bond, J.R. and Arnett, W.D. (1984) *Astrophys. J.*, **277**, 445.

Centrella J. and Melott, A.L. (1983) *Nature*, **305**, 196.

Coles, P. (1986) *Mon. Not. Roy. Astron. Soc.*, **222**, 9P.

Collins, A., Joseph, R.D. and Robertson,N.A. (1986) *Nature*, **320**, 506.

Couchman, H.M.P. (1987) *Mon. Not. Roy. Astron. Soc.*, **225**, 777.

Davies, R.D. (1987) Talk presented at the *Dark Matter in the Universe* meeting, organised by RAL.

Davis, M., Efstathiou, G., Frenk, C.S., and White, S.D.M. (1985) *Astrophys. J.*, **292**, 371.

Davis, M. and Peebles, P.J.E. (1977) *Astrophys. J. Suppl.*, **34**, 425.

Davis, M. and Peebles, P.J.E. (1983) *Astrophys. J.*, **267**, 465.

Deeming, T.J. (1975) *Astrophysics and Space Science*, **36**, 137.

Dresler, A., Faber, S.M., Burstein, D., Davies, R.L., Lynden-Bell, D., Terlevich R.J. and Ewgner, G. (1987) *Astrophys. J. Lett.*, **313**, 37.

Efstathiou, G. (1987), in *Kinematics and Dynamics of Elliptical Galaxies*, ed. de Zeeuw, P.T. Proceedings of the IAU Symposium 127 (Dortrecht: Reidel).

Efstathiou, G., Davis, M., Frenk, C.S. and White, S.D.M. (1985) *Astrophys. J. Suppl.*, **57**, 241.

Efstathiou, G. and Eastwood, J.W. (1981) *Mon. Not. Roy. Astron. Soc.*, **194**, 503.

Einstein, A. (1933) *Structure Cosmologique de l'Espace* (Paris: Hermann).

Fall, S.M. (1987), in *Kinematics and Dynamics of Elliptical Galaxies*, ed. de Zeeuw, P.T. Proceedings of the IAU Symposium 127 (Dortrecht: Reidel).

Foltz, C.B., Weymann, R.J., Röser, H.J. and Chaffee, F.H. (1984) *Astrophys. J. Lett.*, **281**, 1.

Fry, J.N. and Peebles, P.J.E (1978) *Astrophys. J.*, **221**, 19.

Giovanelli, R. and Haynes, M.P. (1985) *Astron. J.*, **90**, 2445.

Giovanelli, R. and Haynes, M.P. (1986) *Astrophys. J. Lett.*, **306**, 55.

Giovanelli, R., Haynes, M.P. and Chincarini, G.L. (1986) *Astrophys. J.*, **300**, 77.

Gott, J.R. and Turner, E.L. (1977) *Astrophys. J.*, **216**, 357.

Groth, E.J. and Peebles, P.J.E. (1975) *Astron. Astrophys.*, **41**, 143.

Guilbert, P.W. and Fabian, A.C. (1986) *Mon. Not. Roy. Astron. Soc.*, **220**, 439.

Gunn, J.E. and Peterson, B.A. (1965) *Astrophys. J.*, **142**, 1633.

Guth, A.H. (1981) *Phys. Rev. D*, **23**, 347.

Hamilton, A.J.S, Gott, J.R. and Weinberg, D. (1986) *Astrophys. J.*, **309**, 1.

Harrison, E.R. (1970) *Phys. Rev. D*, **1**, 2726.

Hogan, C. (1987) Pre-print.

Hockney, R.W. and Eastwood, J.W. (1981) *Computer Simulation Using Particles* (New York: McGraw Hill).

Hubble, E. (1926) *Astrophys. J.*, **64**, 321.

Hubble, E. (1934) *Astrophys. J.*, **79**, 8.

Hubble, E. (1936) *Astrophys. J.*, **84**, 517.

Huchra, J.P. (1984), in *Inner Space/Outer Space*, ed. Kolb, E.W. *et al.* (Chicago: Chicago University Press).

Ikeuchi, S. and Ostriker, J.P. (1986) *Astrophys. J.*, **391**, 522.

Jackson, J.C. (1972) *Mon. Not. Roy. Astron. Soc.*, **156**, 1P.

Kaiser, N. (1984) *Astrophys. J. Lett.*, **284**, 9.

Kaiser, N. (1986) *Mon. Not. Roy. Astron. Soc.*, **222**, 323.

Kaiser, N. (1987) *Mon. Not. Roy. Astron. Soc.*, **227**, 1.

Kapteyn, J.C. (1922) *Astrophys. J.*, **55**, 302.

Kirshner, R.P., Oemler, A., Schechter, P.L. and Schectman,S.A. (1983) *Astron. J.*, **88**, 1285.

Kirshner, R.P., Oemler, A., Schechter, P.L. and Schectman,S.A. (1987) *Astrophys. J.*, **314**, 493.

Klypin, A.A. and Shandarin, S.F. (1983) *Mon. Not. Roy. Astron. Soc.*, **204**, 891.

Koyré (1962) *From the Closed World to the Infinite Universe*(New York: Harper and Brothers).

de Lapparent, V., Geller, M.J. and Huchra, J.P. (1986) *Astrophys. J. Lett.*, **302**, 1.

Lifshitz, E.M. (1946) *ZhETF.*, **16**, 587.

Lighthill, M.J. (1958) *Fourier Analysis and Generalised Functions* (Cambridge: Cambridge University Press).

Linde, A.D. (1984) *Rep. Prog. Phys.*, **47**, 925.

Liu, X.D. and Jones, B.J.T. (1987) Pre-print

Mihalas, D. (1978) *Stellar Atmospheres* (Chicago: Chicago University Press).

Murdoch, H.S., Hunstead, R.W., Pettini, M. and Blades, J.C. (1986) *Astrophys. J.*, **309**, 19.

Neyman, J. and Scott, E.L. (1952) *Astrophys. J.*, **116**, 144.

Novikov, I.D. (1964) *ZhETF.*, **46**, 686.

Ostriker, J.P. and Ikeuchi, S. (1983) *Astrophys. J. Lett.*, **268**, 63.

Peacock, A.H. and Heavens, A.F. (1985) *Mon. Not. Roy. Astron. Soc.*, **217**, 805.

Peebles, P.J.E. (1980) *The Large-Scale Structure of the Universe* (Princeton, N.J.: Princeton University Press).

Peebles, P.J.E. (1982) *Astrophys. J. Lett.*, **263**, 1.

Rees, M.J. (1986) *Mon. Not. Roy. Astron. Soc.*, **218**, 25P.

Rose, M.E. (1957) *Elementary Theory of Angular Momentum* (New York: Wiley).

Rowan-Robinson, M. (1981) *Cosmology* (Oxford: Oxford University Press).

Sargent, W.L.W, Young, P.J,, Boksenberg, A. and Tytler, D. (1980) *Astrophys. J. Suppl.*, **42**, 41.

Schmidt, M. (1963) *Nature*, **197**, 1040.

Schwartz, D.A., Murray, S.S., and Gursky, H. (1976) *Astrophys. J.*, **204**, 315.

Seldner, M. and Peebles, P.J.E. (1978) *Astrophys. J.*, **215**, 703.

Shapley, H. (1918) *Astrophys. J.*, **48**, 154.

Shapley, H. (1918) *Publ. Astron. Soc. Pacific*, **30**, 42.

Shapley, H. (1919) *Astrophys. J.*, **49**, 249.

Shapley, H. (1919) *Astrophys. J.*, **49**, 311.

Shapley, H. (1919) *Astrophys. J.*, **50**, 107.

Shapley, H. (1933) *Harvard Bull.*, **890**, 1.

Shapiro, P.R. (1987) Pre-Print.

Silk, J. (1967) *Nature*, **215**, 1155.

Steidel, C.S. and Sargent, W.L.W. (1987a) *Astrophys. J.*, **313**, 171.

Steidel, C.S. and Sargent, W.L.W. (1987b) *Astrophys. J. Lett.*, **318**, 11.

Sussman, G. (1986) Talk at *Supercomputers and Supercomputing* conference in Princeton.

Totsuji, H. and Kihara, T. (1969) *Publ. Astron. Soc. Japan*, **21**, 221.

de Vaucouleurs, G. (1960) *Astrophys. J.*, **131**, 585.

de Vaucouleurs, G. (1971) *Publ. Astron. Soc. Pacific*, **83**, 113.

Vilenkin, A. (1985) *Phys. Reports*, **121**, 263.

West, M.J., Dekel, A. and Oemler, A. (1987) *Astrophys. J.*, **316**, 1.

Weinberg, S. (1972) *Gravitation and Cosmology* (New York: Wiley).

White, S.D.M. (1979) *Mon. Not. Roy. Astron. Soc.*, **186**, 145.

White, S.D.M, Frenk, C.S., Davis, M. and Efstathiou, G. (1987) *Astrophys. J.*, **313**, 506.

White, S.D.M (1987), in *Kinematics and Dynamics of Elliptical Galaxies*, ed. de Zeeuw, P.T. Proceedings of the IAU Symposium 127 (Dortrecht: Reidel).

Wilkinson D. (1984), in *Inner Space/Outer Space*, ed. Kolb, E.W. *et al.* (Chicago: Chicago University Press).

Zel'dovich, Ya.B. (1970) *Astron. Astrophys.*, **5**, 84.

Zel'dovich, Ya.B. (1972) *Mon. Not. Roy. Astron. Soc.*, **160**, 1P.

Zel'dovich, Ya.B. and Novikov, I.D. (1983) *The Structure and Evolution of the Universe* (Chicago: University of Chicago Press).

Zel'dovich, Ya.B., Einasto, J. and Shandarin, S.F. (1982) *Nature*, **300**, 407.