Fermi National Accelerator Laboratory

FERMILAB-Pub-85/148-A
October 1985

THE STATISTICS OF PEAKS OF GAUSSIAN RANDOM FIELDS

J.M. BARDEEN[1,6], J.R. BOND[2,6], N. KAISER[3,4,6] AND A.S. SZALAY[5,6]

1. Physics Department, University of Washington

2. Physics Department, Stanford University

3. Astronomy Department, U.C. Berkeley

4. Institute of Astronomy, Cambridge University

5. Astrophysics Group, Fermilab

6. Institute of Theoretical Physics, U.C. Santa Barbara

ABSTRACT

Cosmological density fluctuations are often assumed to be Gaussian random fields. The local maxima of such fields are obvious sites for the formation of nonlinear structures. The statistical properties of the peaks can be used to predict the abundances and clustering properties of objects of various types. In this paper, we derive: (1) the number density of peaks of various heights $\nu\sigma_0$ above the rms σ_0; (2) the factor by which the peak density is enhanced in large scale overdense regions; (3) the n-point peak-peak correlation function in the limit that the peaks are well separated, with special emphasis on the two and three point correlations; and (4) the density profiles centered on peaks. To illustrate the predictive power of this semi-analytic approach, we apply our formulas to structure formation in the adiabatic and isocurvature $\Omega = 1$ cold dark matter (CDM) models. We assume bright galaxies form only at those peaks in the density field (smoothed on a galactic scale) that are above some global threshold height $\nu_t \sim 3$ fixed by normalizing to the galaxy number density. We find, for example, that the shapes of the peak-peak two- and three-point correlation functions for the adiabatic CDM model agree well with observations before any dynamical evolution, just due to the propensity of the peaks to be clustered in the initial conditions. Only moderate dynamical evolution is required to bring the amplitude of the correlations up to the observed level. The corresponding redshift of galaxy formation z_g in the isocurvature model is too recent ($z_g \sim 0$) for this model to be viable. Even for the adiabatic models $z_g \sim 3 - 4$ is predicted. We show that the mass-per-peak ratio in clusters, and thus presumably the cluster mass-to-light ratio, is substantially lower than in the ambient medium, alleviating the Ω problem. We also confirm that the smoothed density profiles of collapsing structures of height $\sim \nu_t$ are inherently triaxial.

Operated by Universities Research Association Inc. under contract with the United States Department of Energy

1. INTRODUCTION

Recent theories of the formation of cosmological structures focus attention on the linear and early nonlinear epochs appropriate to the collapse of regions of different length scales. The structure and clustering pattern of the objects forming reflect the initial conditions. These are embodied in a probability ensemble for linear density perturbation configurations $F(\vec{r},t) \equiv (\rho(\vec{r},t) - \langle\rho\rangle)/\langle\rho\rangle$. The fluctuation density $F(\vec{r},t)$ thus defines a three-dimensional random field. In this paper we derive some statistical properties of the local maxima of such fields, assuming they are Gaussian-distributed. Our results can form the core of an analytical framework within which to address the problem of structure formation from small amplitude initial density fluctuations.

The methods we use here complement the n-body and hydrodynamical techniques which are commonly applied to this problem. In principle, n-body methods allow one to follow the nonlinear evolution of any random density field by evolving enough realizations from the probability ensemble so that a combination of averaging over spatial volumes and over ensemble members converges. In practice, limitations arise from discreteness and from present computing capabilities: calculations can cover only limited spatial and temporal dynamic ranges; and the number of realizations from the ensemble that can be evolved is relatively small. (See Efstathiou et al. 1984 for a recent discussion.) For example, the development of rare condensations such as rich clusters are rather difficult to examine by n-body techniques (Barnes et al. 1984). The analytic methods described here are already ensemble-averaged and allow one to investigate easily the gross features of a broad class of initial conditions; and they are particularly suited for the study of rare events. (See Kaiser 1984a for an application to rich clusters.) At present, cosmological hydrodynamical studies require localized pre-collapse structures for their initial conditions. Probabilities of various initial shapes can best be obtained by the statistical methods of the sort we develop here.

In our approach to the problem of non-linear evolution of structure, we focus on the local maxima of the initial density perturbations. We assume that condensations of matter form around sufficiently high local density peaks. In order that the density field possess a well defined set of local maxima it must be smooth and differentiable; its harmonic content must be limited at high wavenumbers. It is often assumed that at very early times the spectrum of fluctuations had a power law form over a wide range of length scales. Deviations from any power law would have arisen naturally on some small scale when the fluctuations were generated, leading to differentiability. However, this length may be tiny, far below scales associated with cosmic structure. If fluctuation power is significant for arbitrarily short wavelengths, then noisy structure will exist on all scales. Universes dominated by cold dark matter have such fluctuation spectra (Peebles 1982, Bond and Szalay 1983). Every cloud would consist of ever smaller subclouds down to the small cutoff scale. During linear evolution, low-pass filtering on cosmologically interesting scales may occur due to a variety of physical processes. For example, adiabatic density fluctuations in baryon-dominated universes suffer Silk damping, and, in neutrino-dominated universes, they suffer collisionless damping as well. The first structures to form are of the physical filtering scale, with smaller scale structure generated by the 'pancaking' process (Zeldovich 1970).

If no physical filtering process exists, as in the cold dark matter picture, we can still treat the formation of objects of some characteristic size by applying an artificial low-pass filter of this scale to the density fluctuation spectrum. This has the effect of smoothing out all of the high frequency spikes in the density field. Since the filter is not physical, care must be taken to avoid the over-interpretation of smoothed density field statistics. This applies especially to the problem of shapes near peaks (§7).

Our model constitutes a non-local and non-linear operation which, when applied to an initial continuous density field $F(r)$, yields a 'population of objects' described by the density field of a point process, a sum of delta functions:

$$n_{pk}(\vec{r}) = \sum_p \delta^{(3)}(\vec{r} - \vec{r}_p), \tag{1.1}$$

where the $\vec{r}_p$ are the positions of the maxima satisfying certain physically-motivated conditions, such as that their height be above some threshold. The non-locality is embodied in the linear filtering, or smoothing, operation (if required) and the non-linearity is introduced by the subsequent application of a threshold criterion. The model has two adjustable parameters: the filtering length and the threshold parameter. We feel that this composite operation encompasses the essential features of the clustering action of gravity. It can also provide a model for a threshold mechanism which may operate, for instance, if galaxy formation were suppressed except at the high peaks of the density field.

We shall restrict ourselves to isotropic homogeneous Gaussian random fields with zero mean as descriptions of the initial cosmological density fluctuations. Such a field is completely specified, in a statistical sense, by a single function: the power spectrum $P(k)$; or, equivalently, its Fourier transform, the autocorrelation function, $\xi(r)$. The Gaussian nature is retained throughout the linear regime of evolution. In principle then, a complete statistical description of the local maxima can be extracted from the power spectrum.

That a Gaussian random field may provide a good description of the properties of density fluctuations could arise in a number of ways. The central limit theorem implies that a Gaussian distribution arises whenever one has a variable (or, more generally, a vector) which is a linear superposition of a large number of independent random variables (or vectors) which are all drawn from the same distribution. In particular, if the field $F(r)$ is written as a spatial Fourier decomposition, and its Fourier coefficients F_k are statistically independent, each having the same form of distribution, then the joint probability of the density evaluated at a finite number of points will be Gaussian under very weak conditions. Special cases of this include the random phase approximation, in which it is assumed that the phases of F_k are uniformly distributed from 0 to 2π. The specific form of the distribution of the moduli $|F_k|$ does not matter. We note that small amplitude curvature perturbations generated by quantum fluctuations in an inflationary phase of the very early universe would yield a Gaussian random density field. Other examples of density fluctuations which are approximately Gaussian are provided by shot noise distributions. A 'shot noise' density field consists of 'shots' (specific local density profiles such as Gaussians) centered on sites chosen from some random point process (e.g. Poisson statistics). On scales much larger than that of the shot, such distributions tend toward Gaussians as a consequence of the central limit theorem (Rice 1944). For example, gas density perturbations generated 'spontaneously' by inhomogeneous sources of radiation burning at redshifts of a few hundred (Hogan 1983) should be nearly Gaussian, though deviations at the high mass end are possible (Hogan and Kaiser 1983). High density loops of cosmic strings (Vilenkin 1984), the 'shots', would constitute a highly non-Gaussian density field locally. However, these loops would also have generated adiabatic perturbations in the matter as they entered the horizon and this component may be approximately Gaussian.

The statistical analysis of one-dimensional Gaussian random fields was pioneered by Rice (1944,1945) to analyze electrical noise in communication devices. He obtained expressions for the frequency of upward zero crossings, of maxima, and of extrema. The clustering properties were largely ignored. Longuet-Higgins (1957) extended Rice's analysis to two dimensions; his prototype random surface was an ocean surface rippled by gentle winds. He obtained the average number of maxima per unit area and the distribution of their heights, but was unable to obtain an expression for their spatial distribution except for a nearly monochromatic field. Less progress has been made for the case of higher dimensional fields due largely to the mathematical complexities. Adler (1981) has written a rigorous mathematical text on those few theorems that have been proved.

Doroshkevich (1970) was the first to apply these methods extensively to the study of the formation of cosmic structure. The many results given in this classic paper have mostly been applied to the pancake model, primarily by the Russian school. Adler and Doroshkevich give the number of peaks per unit volume in the limit where the amplitude of the peaks is extremely high (§4). They also obtain a closed analytic result for the average Euler characteristic expected per unit volume for contour surfaces of a given density; this useful result is twice the density of maxima in the high

peak limit (§4). Doroshkevich also evaluated the average shapes of contours around maxima (§7). Unfortunately, Doroshkevich only presents his results for a rather specific choice of power spectrum (an approximation to that which arises from initially adiabatic fluctuations in a baryon dominated universe), and focuses primarily on high threshold limits, which, as we shall show, are inaccurate in the regime of cosmological interest. He also does not address the important issue of the spatial distribution of maxima; in particular, determination of their correlation functions. Doroshkevich and Shandarin (1978a,1978b) have calculated some of the statistical properties of the maxima of the largest eigenvalue of the shear tensor. These maxima define the 'domain walls' where pancakes form rather than the isolated points which form the maxima of the density field. These authors also restrict their analysis to baryon-dominated 'pancake' models. Recently, a number of workers have attempted to determine the properties of bound objects which form in universes dominated by dark matter in which the perturbations obey Gaussian statistics (Peebles 1983, Hoffman and Shaham 1984, Shaeffer and Silk 1984); when the calculations involve the constraint that a point is an extremum of the density, our results differ from theirs. In §3, we discuss the procedure for constructing conditional probabilities subjected to the extremum constraint. These are required for the correct calculation of the spatial distribution of extrema and of the density profiles around them.

Kaiser (1984a) discussed the way in which the clustering of maxima reflects the correlations of the underlying density field. He obtained the 2-point correlation function of those regions with density fluctuations above some threshold ν (in units of the rms fluctuation level, σ_0); i.e. the autocorrelation of a Heaviside function of the density field: $Y(r) = \theta(F(r) - \nu\sigma_0)$. On large scales, $\langle Y(r)Y(0)\rangle$ is proportional to $\langle F(r)F(0)\rangle$ with an amplitude which increases with increasing threshold. This mechanism might account for the anomalously large correlation strength of rich clusters. Politzer and Wise (1984) extended this result and also gave an expression for the n-point correlation function of the highly non-Gaussian field $Y(r)$ valid in the limit of very high threshold. The large scale clustering properties of $Y(r)$ can also be obtained from an effective density field which is simply the exponential of the underlying density field $F(r)$ with its short-wavelength components filtered out; again this result is strictly valid only in the limit of high threshold. From this form $(exp(\nu F(r)/\sigma_0))$ for the effective density field it is easy to see that on large scales one has a simple linear enhancement of the correlation strength while on smaller scales one has the possibility of a non-linear enhancement of the number density of high peaks in initially overdense regions. It was recognized (Bardeen 1984, Kaiser 1984b) that if, for some reason, galaxy formation had been suppressed except at the high density peaks this process would give rise to strong segregation of light from mass and provide a natural way to reconcile low cluster mass-to-light ratios obtained from virial analysis with an $\Omega = 1$ universe. This statistical mechanism had also been invoked, though in a slightly different context, by Rees (1984) as a possible way to obtain mass/light segregation on large scales.

The outline of the paper is as follows. In §2, we define some general properties of random fields. In §3, we discuss the point process equation (1.1) and give the general formula for the average number density of peaks. We also discuss the problem of the proper conditional probability constraints appropriate to maxima using a one-dimensional illustration. In §4, we calculate the average density of maxima of a general three-dimensional Gaussian field as a function of the heights of the maxima, and present a useful analytic approximation to our formula. Our results are valid for arbitrary heights. We also obtain the high peak limit of our result and discuss the utility of it and of the average Euler characteristic density. We then compute the average density of 'upcrossing' points on density contour surfaces and show that it gives a good approximation to the peak density. In §5, we determine the number density of peaks subject to the constraint that the large scale density field be fixed. We use this to discuss the segregation of high peaks from the underlying mass distribution. In §6, we present the machinery to calculate n-point peak-peak correlation functions. We explicitly calculate the two- and the three-point function. In §7, we determine the shapes of the profiles about maxima. Finally, in §8, we summarize our main results and outline how these statistical results can be applied to cosmological problems, We have relegated the details of many of the derivations to seven appendices.

In discussing cosmological applications of our formal results, we assume a universe dominated

by cold dark matter (CDM) with density parameter $\Omega = 1$ and zero cosmological constant. Unless specifically stated otherwise, all spatial separations and length scales are described in comoving coordinates in the cosmological background, scaled so they correspond to physical distances at the present cosmological epoch. To reflect the observational determination of cosmological distance from redshift, all distances are given in units of $h^{-1}Mpc$, where h is related to the Hubble constant H_0 by

$$h \equiv H_0/\ (100\ km\ s^{-1}\ Mpc). \tag{1.2}$$

2. GAUSSIAN RANDOM FIELDS

An n-dimensional *random field* $F(\vec{r})$ is a set of random variables, one for each point $\vec{r}$ in n-dimensional real space, defined by the set of finite-dimensional joint probability distribution functions,

$$P(F(\vec{r}_1), F(\vec{r}_2), ..., F(\vec{r}_m))\ dF(\vec{r}_1)dF(\vec{r}_2)...dF(\vec{r}_m), \tag{2.1}$$

that the function F have values in the range $F(\vec{r}_j)$ to $F(\vec{r}_j) + dF(\vec{r}_j)$ for each of the $j = 1, ..., m$ with m an arbitrary integer and $\vec{r}_1, \vec{r}_2, ..., \vec{r}_m$ arbitrary points. By taking appropriate limits, the joint distribution for values of F and its derivatives, $\nabla F, \nabla\nabla F, ...$ follow. Differentiation and integration of random fields may not always be possible; they are usually defined as appropriate limits in the mean square. For example, without high frequency filtering, the density perturbation field for cold dark matter starting from an initial Zeldovich spectrum is arbitrarily spiky on small spatial scales, and this random field is not differentiable.

A *Gaussian random field* is one for which the various m-point probability distributions (equation 2.1) are multivariate Gaussians. A joint Gaussian probability distribution for random variables y_i is

$$\begin{aligned} P(y_1, ...y_n)dy_1...dy_n &= \frac{e^{-Q}}{((2\pi)^n det(M))^{1/2}} dy_1...dy_n, \\ Q &\equiv \sum \Delta y_i (M^{-1})_{ij} \Delta y_j/2. \end{aligned} \tag{2.2}$$

Only the means of the random variables $\langle y_i \rangle$ and their covariance matrix

$$M_{ij} \equiv \langle \Delta y_i \Delta y_j \rangle, \quad \Delta y_i \equiv y_i - \langle y_i \rangle \tag{2.3}$$

are required to specify the distribution. If $F(\vec{r})$ is Gaussian, joint distributions involving arbitrarily many values of the field and its derivatives, integrals and generally any linear functions of F are Gaussian. For a scalar Gaussian random field with zero mean, knowledge of only the 2-point correlation function of F, $\xi(r_1, r_2) = \langle F(r_1)F(r_2) \rangle$ and its derivatives and integrals is sufficient to calculate any statistical property.

A random field is *strictly homogeneous* if all finite dimensional distributions are invariant under simultaneous translations of the points $\vec{r}_i$ by the same vector. For Gaussian fields, this implies $\langle F(\vec{r}) \rangle$ is spatially-independent and ξ is a function of the difference $\vec{r}_1 - \vec{r}_2$ only; its Fourier transform is the power spectrum, $P(k)$, given by the variance

$$P(k) = \langle |F_k|^2 \rangle. \tag{2.4}$$

The spatial Fourier modes, $F_k = |F_k|\ e^{i\theta_k}$, are defined by the expansion

$$F(r) = 2 \sum_{k\ \epsilon\ uhs} |F_k| \cos(\vec{k} \cdot \vec{r} + \theta_k), \tag{2.5}$$

where *uhs* denotes the upper half of k-space, the limitation arising from the reality of F, so $\theta_{-k} = -\theta_k$.

A strictly homogeneous Gaussian random field is *isotropic* if ξ is rotation-invariant. The power spectrum is then only a function of $|\vec{k}|$.

Rigorously, in order for a field to be strictly homogeneous and Gaussian, its various spatial Fourier modes, F_k, must be mutually independent, have random phases and have moduli which are Rayleigh-distributed:

$$P(|F_k|,\theta_k)d|F_k|\ d\theta_k = exp(-\frac{|F_k|^2}{2P_k})\ \frac{|F_k|\ d|F_k|}{P(k)}\frac{d\theta_k}{2\pi}. \tag{2.6}$$

The real and imaginary parts of F_k are then mutually independent and Gaussian-distributed. However, as a consequence of the central limit theorem, a large variety of distributions will tend to a Gaussian. For example, random phases are essentially all that is required of equation (2.6) for the joint distribution of the density evaluated at a number of points to be a multivariate Gaussian. The quantum generation of metric fluctuations predicted to occur in inflationary models would give rise to a distribution of the form of equation (2.6).

An important property of Gaussian random processes is that they are *ergodic*. Our universe is unique, at least in the large. Thus, averages taken in our universe must be spatial ones, over a large volume. These averages will be equal to expectations over an ensemble of universes to which ours belongs if an ergodic theorem holds. It can be proved that a Gaussian random density fluctuation field is ergodic if and only if $P(k)$ is continuous (Adler 1981).

3. THE EXTREMUM CONSTRAINT

3.1 THE NUMBER DENSITY OF EXTREMA

The point process equation (1.1) describes the number density of points p selected to be maxima of the random field $F(\vec{r})$. We could further restrict the class of points we select if we consider, for example, only those maxima above a certain threshold height. Or, we might be interested in the less restrictive class of extremal points.

We can express the point process entirely in terms of the field and its derivatives. In the neighborhood of a maximum point $\vec{r}_p$ we can expand the field $F(\vec{r})$ and its gradient $\vec{\eta}(\vec{r}) \equiv \nabla F(\vec{r})$ in a Taylor series:

$$\begin{aligned} F(\vec{r}) &\approx F(\vec{r}_p) + \frac{1}{2}\sum_{ij}\varsigma_{ij}(r-r_p)_i(r-r_p)_j, \\ \eta_i(\vec{r}) &\approx \sum_j \varsigma_{ij}(r-r_p)_j. \end{aligned} \tag{3.1}$$

We have used the extremum constraint $\eta_i(\vec{r}_p) = 0$. For the extremum to be a maximum, the second derivative tensor of the field, $\varsigma_{ij} \equiv \nabla_i\nabla_j F(\vec{r})$, must be *negative definite* at $\vec{r}_p$. (We are following Rice's 1944 notation of η for the first derivative and ς for the second derivative.) Provided the ς-matrix is nonsingular at $\vec{r}_p$, we have

$$\vec{r} - \vec{r}_p \approx \varsigma^{-1}(\vec{r}_p)\vec{\eta}(\vec{r}),\ \text{hence}\ \delta^{(3)}(\vec{r}-\vec{r}_p) = |det\ \varsigma(r_p)|\ \delta^{(3)}(\vec{\eta}(\vec{r})).$$

This is true for each maximum; but the δ-function picks out *all* of the (*extremal*) points which are zeros of $\vec{\eta}(\vec{r})$. The expression for the number density of extrema in terms of field derivatives is then

$$n_{ext}(\vec{r}) = |det\ \varsigma(r)|\ \delta^{(3)}(\vec{\eta}(\vec{r})); \tag{3.2}$$

the expression for $n_{pk}(\vec{r})$ is identical except for the added restriction of negativity on the 3 eigenvalues of ς_{ij}. Further, if we select only those maxima whose heights are in the range F_0 to $F_0 + dF$, a $\delta(F - F_0)dF$ multiplies equation (3.2).

In principle, we could calculate an infinite hierarchy of correlation functions of equation (3.2). This is necessary for a complete description of the point process. In practice, only the ensemble average of 3.2 is easily obtainable:

$$\begin{aligned}\langle n_{ext}(\vec{r})\rangle &\equiv \langle |det\ \zeta(r)|\ \delta^{(3)}(\vec{\eta}(\vec{r}))\rangle \\ &= \int |det\ \zeta|\ p(F,\eta=0,\zeta)dFd^6\zeta.\end{aligned} \tag{3.3}$$

The homogeneity of the field guarantees that this will be independent of $\vec{r}$. To evaluate equation (3.3), we require the joint probability distribution $p(F,\eta,\zeta)dFd^3\eta d^6\zeta$ for the field at $\vec{r}=0$ being in the range F to $F+dF$, the field gradient η_i being in the range η_i to $\eta_i+d\eta_i$, and the second derivative matrix elements ζ_{ij} being in the range ζ_{ij} to $\zeta_{ij}+d\zeta_{ij}$. This distribution is evaluated as the usual multivariate Gaussian (equation 2.2). Since ζ_{ij} is symmetric, there are only 6 independent components. The computation of this integral is given in Appendix 1 and the results are discussed in §4.

The derivation of equation (3.3) does not require the introduction of the point process, but can be obtained directly from the probability distribution. We now sketch this development. In order to have a zero of $\vec{\eta}$ somewhere in the infinitesimal volume d^3r about $\vec{r}=0$, we must have $\eta_i \approx \zeta_{ij}r_j$, hence $d^3\eta$ can be replaced by $|det\ \zeta|\ d^3r$. The probability that there is a zero in d^3r is then

$$dFd^3r \int d^6\zeta\ |det\ \zeta|\ p(F,\eta=0,\zeta) = \langle n_{ext}(F)\rangle dFd^3r,$$

so the extremum number density is given by equation (3.3). The integral is over all values of ζ for an extremum.

Equations (3.2) and (3.3) can be generalized to the determination of the number density of points for which *any* (different) random fields, $y_1(\vec{r}), y_2(\vec{r}), y_3(\vec{r})$, take on specific values, y_{10}, y_{20}, y_{30}. Since $y_i(\vec{r}) \approx y_{i0} + \sum_j y_{i,j}(r_j - r_{0j})$, the mean number density is, for example,

$$\langle n(y_{10},y_{20},y_{30})\rangle = \int P(y_1=y_{10}, y_2=y_{20}, y_3=y_{30}, \{y_{i,j}\})\ |\ det\ (y_{i,j})\ |\ \prod dy_{i,j}. \tag{3.4}$$

In the application to extrema that we are most interested in, $y_i=\eta_i$ and $y_{i0}=0$. It is straightforward to obtain the density of points with a specific nonzero value of the gradient from this. A related problem would be to determine the density of points for which the velocity has a specific value. However, the y_i do not have to be the components of a vector for this density formula to be applicable (see §4.5).

3.2 The Extremum Constraint in Conditional Probabilities

The problem of density of maxima is especially easy in one-dimension and was solved by Rice (1944, 1945). We review the reasoning behind this calculation here because it illustrates an important issue: that care must be taken when constructing probabilities of various events subject to the constraint that an extremum exists at some given point ($r=0$ say). The conditional probability that event A occurs given that event E is true is given by the Bayes formula $P(A|E)=P(A,E)/P(E)$. For example, we might take the event A to be that the height of the density field at the extremum be in some prescribed range; or that the height of the field a distance r away from the extremum be in some prescribed range. In this subsection, we discuss the extremum rather than the maximum constraint; almost all sufficiently high extrema will be maxima.

At first sight, one might expect that the event that there is an extremum at zero, E, should be taken to be $\eta = dF/dr = 0$. However, probabilities for point events worded in this way are always zero. What we really mean by the constraint is that there is an extremum in an interval of length ϵ about $r=0$; we must form conditional probabilities for events E_ϵ of this form and only then can

we let the length of the interval go to zero. If r_e denotes the extremal point, then, since a Taylor expansion of $\eta(r_e)$ implies $r_e \approx -\eta(0)/\zeta(0)$, the condition that $|r_e| < \epsilon/2$ reduces to the condition $|\eta/\zeta| < \epsilon/2$. If we let $p(\eta,\zeta)d\eta d\zeta$ and $p(A,\eta,\zeta)d\eta d\zeta$ denote the obvious joint probabilties, then

$$\begin{aligned} P(E_\epsilon) &= \int d\zeta \int_{-\zeta\epsilon/2}^{\zeta\epsilon/2} d\eta\ p(\eta,\zeta) \\ &\to \epsilon \int |\zeta| d\zeta\ p(\eta=0,\zeta) \quad \text{and} \\ P(A,E_\epsilon) &\to \epsilon \int |\zeta| d\zeta\ p(A,\eta=0,\zeta) \quad \text{as } \epsilon \to 0. \end{aligned} \tag{3.5}$$

The conditional probability is the ratio of these two integrals. The presence of $|\zeta|$ in both integrands is an important feature. The integration is over both positive and negative ζ for extrema. The constraint that $r = 0$ be a maximum leads to identical equations except that the integration is only over negative ζ.

Another possible set of constraint events E'_ϵ which give an extremum at $r = 0$ in the limit $\epsilon \to 0$ is given by the condition: $|\eta| < \epsilon/2$. This class of events was impicitly assumed by Peebles (1984) and Hoffman and Shaham (1984) in their determination of various conditional probabilities of cosmological interest. The obvious advantage that the distribution of second derivatives is not required is offset by the fact the probabilities so obtained are biased in favor of extrema with extremely broad profiles. These are just the extrema which are of least interest cosmologically. The events E'_ϵ define intervals of length $\Delta r = \epsilon/\zeta$ within which the extrema occur. For fixed ϵ, it is clear that ζ being as small as possible probes the largest regions for zeros of dF/dr. The events E_ϵ exactly compensate for this so the lengths are all of equal size.

The average density of extrema as determined from equation (3.3),

$$\langle n_{ext} \rangle = \lim_{\epsilon \to 0} P(E_\epsilon)/\epsilon, \tag{3.6}$$

is the probability of finding one per unit length only if the constraint events E_ϵ are chosen. If the extremal points r_e are chosen to be only those for which the event A is true, then the associated mean density is

$$\langle n_{ext}(A) \rangle = \lim_{\epsilon \to 0} P(A,E_\epsilon)/\epsilon.$$

The conditional probability of event A occuring given that $r = 0$ is an extremum is thus the ratio of the numbers of special extremal points for which A is true to all extremal points:

$$P(A|E) = \langle n_{ext}(A) \rangle / \langle n_{ext} \rangle. \tag{3.7}$$

We are often interested in further constraint conditions, for example that the height of the extremum is some prescribed value. Denoting this condition by the event B, we have

$$P(A|B,E) = \langle n_{ext}(A,B) \rangle / \langle n_{ext}(B) \rangle. \tag{3.8}$$

This result, when extended to three-dimensions, is used extensively in the rest of the paper. The extension is straightforward as in §3.1

Various erroneous statements have appeared in the cosmological literature due to the choice of E'_ϵ as the constraint events. Peebles (1984) has claimed that the distribution of heights of an extremum is the same as that for arbitrary points, namely a Gaussian. Peebles and Hoffman and Shaham (1984) have claimed that the expectation of the shape of the profile around an extremum, or maximum, is approximately equal to the correlation function of F. As we now show, the distribution

of the heights of maxima is generally broader than that of field points (§4) and the correct average profile is narrower than the correlation function (§7).

4. THE PEAK NUMBER DENSITY OF 3-DIMENSIONAL FIELDS

4.1 THE NUMBER DENSITY FITTING FORMULA

In this section, we sketch the derivation of the differential number density $\mathcal{N}_{pk}(\nu)$, where $\nu \equiv F/\sigma_0$ with σ_0 the *rms* fluctuation level of F. (From now on, we let $\mathcal{N}_{pk}(\nu)d\nu$ denote average differential number densities, and $n_{pk}(\nu)$ denote number densities integrated over ν.) We also present an accurate fitting formula to our result. The derivation is given in detail in Appendix 1. The steps are as follows:

(1) To evaluate equation (3.3) with the added requirement that the points are maxima rather than just being extrema, we first obtain the Gaussian joint probability distribution $p(F,\eta,\zeta)dFd^3\eta d^6\zeta$. Since ζ_{ij} is symmetric, there are only 6 independent components. The integral is over all values of ζ for an extremum.

(2) To obtain the density of maxima, ζ must be negative definite. In that case, it is useful to rotate to its principal axes. A set of Euler angles defines the rotation matrix, R, which diagonalizes ζ: $\mathrm{diag}(\lambda_1,\lambda_2,\lambda_3) = -R\zeta R^\dagger$, $R^\dagger$ denoting the transpose. The eigenvalues λ_i are positive at a maximum and are ordered by

$$\lambda_1 \geq \lambda_2 \geq \lambda_3. \tag{4.1}$$

The condition that the extremum be a maximum is therefore simply $\lambda_3 \geq 0$. The density of maxima is

$$\mathcal{N}_{pk}(\nu) = \langle\, \delta(F/\sigma_0 - \nu)\, |\lambda_1\lambda_2\lambda_3|\, \theta(\lambda_3)\, \delta(\vec{\eta})\, \rangle, \tag{4.2}$$

where the average is taken using the probability distribution $p(F,\eta,\zeta)$. The density of minima is related to that of maxima by $\mathcal{N}_{min}(\nu) = \mathcal{N}_{pk}(-\nu)$.

(3) Partial integrations of this expression are of some interest, since they tell us about the distribution of shapes near the peak, where $F(r) \approx F(0) - \sum \lambda_i r_i^2/2$. This is discussed in §7. We have obtained an analytic expression (equation A1.14) for $\mathcal{N}_{pk}(\nu,x)$, where $x = (\lambda_1+\lambda_2+\lambda_3)/\sigma_2$ (σ_2 is a parameter defined in §4.2). The integral over x must be done numerically.

(4) We express the (comoving) differential peak density in terms of a function $G(\gamma,w)$, where $w = \gamma\nu$, and γ is a spectral quantity defined in §4.2:

$$\mathcal{N}_{pk}(\nu)d\nu = \frac{1}{(2\pi)^2 R_*^3} e^{-\nu^2/2}\, G(\gamma,\gamma\nu). \tag{4.3}$$

The (comoving) length R_* is also defined in §4.2. The explicit form of G is given by equation (A1.19).

(5) The following fitting formula designed to agree with the asymptotic large ν behavior is accurate to better than 1% over the range $0.3 < \gamma < 0.7$ and $-1 < w < \infty$ of cosmological interest, with the accuracy increasing to better than one part in a thousand for $w > 1$:

$$G(\gamma,w) = \frac{(w^3 - 3\gamma^2 w + [B(\gamma)w^2 + C_1(\gamma)]exp(-A(\gamma)w^2)}{(1 + C_2(\gamma)exp(-C_3(\gamma)w))}. \tag{4.4}$$

The coefficients A and B are taken from the asymptotic expansion and the C_i are fits:

$$\begin{aligned} A &= \frac{5/2}{(9-5\gamma^2)}, \\ B &= \frac{432}{(10\pi)^{1/2}(9-5\gamma^2)^{5/2}}, \\ C_1 &= 1.84 + 1.13(1-\gamma^2)^{5.72}, \\ C_2 &= 8.91 + 1.27\ exp(6.51\gamma^2), \\ C_3 &= 2.58\ exp(1.05\gamma^2). \end{aligned} \tag{4.5}$$

4.2 Spectral Parameters

The peak density equation involves the parameters γ and R_* which are related to various moments of the power spectrum $P(k) = \langle\, |F(k)|^2 \,\rangle$:

$$\gamma \equiv \frac{\langle k^2 \rangle}{\langle k^4 \rangle^{1/2}} = \frac{\sigma_1^2}{\sigma_2 \sigma_0}, \qquad R_* \equiv \sqrt{3}\frac{\sigma_1}{\sigma_2}. \tag{4.6a}$$

Here, the mean square density fluctuation at time t is

$$\sigma_0^2(t) \equiv \int \frac{k^2 dk}{2\pi^2} P(k,t) = \xi(0,t) \equiv \langle F(r=0,t)^2 \rangle. \tag{4.6b}$$

It is convenient to extrapolate F by linear perturbation theory to the present ($t = t_0$) to define $\sigma_0(t_0)$. Thus, in an Einstein-deSitter universe, the redshift at which the *rms* density fluctuations went nonlinear is given by $1 + z_{nl} \equiv \sigma_0(t_0)$. The z_{nl}-expression is more complicated for $\Omega < 1$ universes due to the deviation of the perturbation growth law from $(1+z)^{-1}$. The parameters σ_1 and σ_2 are members of a set of spectral moments weighted by powers of k^2:

$$\sigma_j^2(t) \equiv \int \frac{k^2 dk}{2\pi^2} P(k,t) k^{2j}. \tag{4.6c}$$

Of course, since the density field grows in a self-similar way in the linear regime, the ratios σ_j/σ_0 are time independent. Thus γ and R_* and hence the comoving density of peaks do not depend upon the time at which the density is measured. They only involve the moments

$$\langle k^2 \rangle \equiv \sigma_1^2/\sigma_0^2 = 3(\gamma/R_*)^2, \qquad \langle k^4 \rangle \equiv \sigma_2^2/\sigma_0^2 = 9\gamma^2/R_*^4, \tag{4.6d}$$

which are $-3\xi''(0)/\xi(0)$ and $5\xi^{(iv)}(0)/\xi(0)$, respectively, when expressed in terms of derivatives of the density correlation function at zero separation. Notice that the peak density only depends upon the spectrum through γ except for the overall multiplicative volume term R_*^3. The value of γ reflects the range over which $k^3 P(k)$ is large, since $\langle (\Delta k^2)^2 \rangle^{1/2} / \langle k^2 \rangle = (\gamma^{-2} - 1)^{1/2}$ measures the relative spectral width. If P is a delta-function then $\gamma = 1$, whereas, if $k^3 P$ is constant over a wide range of k, then γ is much less than unity.

Gaussian smoothing of the random field $F(\vec{r})$ on the comoving scale R_f leads to the new field

$$F(\vec{r}; R_f) \equiv \int exp(-|\vec{r} - \vec{r}\,'|^2/2R_f^2)\; F(\vec{r}\,') \; \frac{d^3 r\,'}{(2\pi R_f^2)^{3/2}} \tag{4.7}$$

whose Fourier components and power spectrum are Gaussian-filtered on this scale:

$$F(k; R_f) = exp(-R_f^2 k^2/2) F(k), \quad P(k; R_f) = e^{-R_f^2 k^2} P(k). \tag{4.8}$$

For the special case of a power law spectrum which has a Gaussian filtering,

$$P(k; R_f) \sim k^n \; exp(-(kR_f)^2), \tag{4.9}$$

the *rms* fluctuations scale with the filtering length as

$$\sigma_0(R_f) \propto R_f^{-(n+3)/2} \tag{4.10a}$$

and the other spectral quantities are given by

$$\sigma_1^2(R_f)/\sigma_0^2(R_f) = \frac{(n+3)}{2} R_f^{-2}, \quad \sigma_2^2(R_f)/\sigma_0^2(R_f) = \frac{(n+5)(n+3)}{4} R_f^{-4},$$
$$\gamma^2 = (n+3)/(n+5), \quad R_* = (\frac{6}{n+5})^{1/2} R_f. \tag{4.10b}$$

The number density is therefore $\propto R_f^{-3}$.

Another popular filtering choice is top hat smoothing:

$$F_{TH}(\vec{r}; R_{TH}) \equiv \int \theta(1 - |\vec{r} - \vec{r}\,'|/R_{TH})\ F(\vec{r}\,')\ \frac{d^3r\,'}{(4\pi R_{TH}^3/3)}\ ,$$

$$F_{TH}(k; R_{TH}) = W(kR_{TH})F(k), \quad P_{TH}(k; R_{TH}) = W^2(kR_{TH})P(k), \quad W(x) \equiv 3(\sin x - x\cos x)/x^3.$$

Here, θ is the Heaviside unit function. Sharp k-space filtering on the scale k_c leads to the smoothing function $(k_c^3/(6\pi^2))W(k_c r)$. In both cases, the sharpness results in oscillations in the correlation functions in the conjugate space which have significant amplitudes out to many times the cutoff scale. Gaussian filtering avoids this undesirable feature by balancing the smoothing and filtering requirements in an optimal manner. The identification of a given R_{TH} with an equivalent R_f depends upon which features of the fluctuation we are most interested in. For example, the mass enclosed by the smoothing function applied to the uniform background is the same for $R_f = 0.64R_{TH}$.

In Fig. 4.1, we plot the parameters γ and R_* for two currently popular spectra, one corresponding to the adiabatic cold dark matter (CDM) model (Peebles 1982, Blumenthal and Primack 1984, Bond and Efstathiou 1984, Bardeen 1984), the other corresponding to the isocurvature axion CDM model (Bardeen 1984, Efstathiou and Bond 1985). The transfer functions corresponding to these cases, and also to hot and warm dark matter, are given in Appendix 7. The initial conditions for the density fluctuations were taken to be the Zeldovich spectrum ($n = 1$) for the adiabatic model and flicker noise ($n = -3$) for the isocurvature axion model; both forms are predicted to arise in inflationary models. (For the Zeldovich spectrum, it is the gravitational potential which has the scale-invariant flicker noise spectrum.) The fluctuations evolve as the universe expands, leading to spectra with effective power law indices ($n_{eff}(k) \equiv d\ ln\ P(k)\ /d\ ln\ k$) ranging from -3 on small scales up to 1 on large scales. On galactic scales ($R_f \sim 0.5\ h^{-1}Mpc$), the index is ~ -2 for the adiabatic spectrum and ~ -2.5 for the isocurvature one. The moments of the spectrum clearly favor those wavelengths near the filtering scale, so the value of γ will reflect the value of n_{eff} at this scale. Indeed, n_{eff} should be defined by equation (4.10b) for a given R_f in order to use it in the peak density formula. In Figure 4.1, we also plot $log\ \sigma_0(R_f)$. Thus, Fig. 4.1 can be used to determine the relative epochs at which different scales go nonlinear.

4.3 Differential and Cumulative Density Results

To illustrate the results we have plotted the differential number density for various values of γ in Fig.4.2. Typically, there is a broad peak for maxima with heights about equal to the *rms*, $\nu \sim 1$. Note however that for small $\gamma \sim 0.3$, $n_{eff} \sim -2.8$, the peak is $\nu \sim 0$, reflecting the fact that on all scales, there is significant power, so there are many peaks even below the *rms*. For large $\gamma \sim 0.9$, $n_{eff} \sim +5.5$, all the power is at short wavelengths, and the peaks are therefore relatively isolated and typically quite high ($\nu \sim 2$). In all cases, the falloff is steep beyond $\nu \sim 2$. The cumulative number density of peaks higher than height ν,

$$n_{pk}(\nu) \equiv \int_\nu^\infty \mathcal{N}_{pk}(\nu) d\nu, \tag{4.11a}$$

is plotted in Fig.4.3. The asymptotic cumulative number giving the comoving density of peaks of arbitrary height is a useful quantity which can be evaluated analytically:

$$n_{pk}(-\infty) = \frac{29 - 6\sqrt{6}}{5^{3/2}2(2\pi)^2 R_*^3} = 0.016 R_*^{-3}. \tag{4.11b}$$

The fraction of peaks above a threshold can therefore be read off Figure 3. It can also be computed using equation (4.21) below.

A physical selection criterion for peaks which would form a given class of objects is unlikely to be so sharp as to make $n_{pk}(\nu_t)$ the relevant density, where ν_t is the threshold height. Instead, we might introduce a selection function $t(\nu/\nu_t)$, which gives the probability that a peak of height ν forms one of the objects. The number density would then be

$$n_{pk}[t] = \int_0^\infty t(\nu/\nu_t)\mathcal{N}_{pk}(\nu)d\nu \tag{4.12}$$

For the sharp threshold, the selection function is a Heaviside: $t(\nu/\nu_t) = \theta(\nu - \nu_t)$. Consider instead the class of functions

$$t(\nu/\nu_t, q) = \frac{(\nu/\nu_t)^q}{1 + (\nu/\nu_t)^q} \tag{4.13}$$

which also select sub-ν_t peaks, though with low probability. (We will usually let $n_{pk}(\nu_t)$ denote the density 4.12 if the selection function is of form 4.13.) The $q \to \infty$ limit reproduces the sharp threshold case. If q is too small, low peaks dominate even if the probability of their selection is small, due to their large abundance. These effects are illustrated in Fig. 4.4.

4.4 The High Peak Limit and the Euler Characteristic Density

In the limit of large $\gamma\nu$, the function G gives the high peak limit

$$\begin{aligned} \mathcal{N}_{pk}(\nu)d\nu &\to \frac{(\langle k^2\rangle/3)^{3/2}}{(2\pi)^2}(\nu^3 - 3\nu)e^{-\nu^2/2}d\nu \quad \text{as} \quad \nu \to \infty, \\ n_{pk}(\nu) &\to \frac{(\langle k^2\rangle/3)^{3/2}}{(2\pi)^2}(\nu^2 - 1)e^{-\nu^2/2} \quad \text{as} \quad \nu \to \infty. \end{aligned} \tag{4.14}$$

This result agrees with the expression given by Doroshkevich (1970) and Adler (1981) for the high peak limit. It is clear from Fig. 4.2 that this approximation is not accurate for moderate values of ν.

As discussed by Adler (1981), a useful approximation to the number density of extrema (or maxima) lying above a high threshold ν_t should be given by (one half of) $n_\chi(\nu_t)$, the density of the Euler characteristic. The Euler characteristic for a scalar field F is defined to be

$$\chi \equiv \textit{number of maxima} + \textit{number of minima} - \textit{number of saddle points}$$

of the *density contour surface* $F(\vec{r}) = \nu_t\sigma_0$. To define maxima and minima, we require that a direction be specified. Locally, the 2-dimensional surfaces can be defined by $x_3(x_1, x_2)$, where the x_3 coordinate has been singled out as the dependent variable. χ can then be calculated in this chosen coordinate frame. Of course, the result is independent of the particular choice of the coordinate axes and dependent variable, due to the isotropy of the random field. Following geometrical considerations similar to those in §3, the mean value of χ per unit volume is

$$n_\chi(\nu) = \langle\, \delta(F - \nu_t\sigma_0)\, |\eta_3|\, (\zeta_{11}\zeta_{22} - \zeta_{12}^2)\,\rangle \tag{4.15}$$

(Doroskevich 1970, Adler 1981), where $\eta_1 = \eta_2 = 0$ in the averaging procedure. This integral is straightforward to calculate (Appendix 1) once the Gaussian probability distribution

$$p(F, \vec{\eta}, \zeta_{11}, \zeta_{12}, \zeta_{22})dF d^3\vec{\eta} d\zeta_{11} d\zeta_{12} d\zeta_{22}$$

has been evaluated:

$$n_\chi(\nu_t) = 2\frac{(\langle k^2\rangle/3)^{3/2}}{(2\pi)^2}({\nu_t}^2 - 1)e^{-{\nu_t}^2/2}. \tag{4.16}$$

Except for the multiplicative factor out front, this agrees with Doroshkevich's (1970) expression.

For large values of ν_t (in units of the rms) these contour surfaces become predominantly simply connected 'bags' surrounding those extrema whose heights exceed the level ν_t and which are almost surely maxima. Thus, in the high ν_0 limit, the Euler characteristic counts one minimum plus one maximum for each bag, i.e. for each field extremum, so n_χ must approach *twice* the asymptotic limit of $n_{pk}(\nu_t)$, equation (4.14). However, neither provides an accurate estimation of peak density over the regime of cosmological interest. Indeed, for 4.16 to reproduce 4.11a with 10% accuracy, $\gamma\nu > 2.5$ is required.

4.5 The Number Density of Upcrossing Points on Contour Surfaces

A more accurate formula in the regime of cosmological interest for the integrated peak density can be obtained analytically by evaluating the density of a special class of points which are generalizations of the upcrossing points defined by Rice (1944). One of Rice's most useful analytic results for one-dimensional Gaussian fields was the number of upcrossings of a given threshold level that occur per unit 'time'; that is the density of points satisfying $F = F_t$, $dF/dt > 0$. Adler treats the Euler characteristic density as the appropriate generalization of this concept to multiple dimensional fields. We believe there is a natural class of points for isotropic random fields which better generalize the Rice points. These upcrossing points $\vec{r}_u$ lie on the contour surfaces of threshold F_t. As in the Euler characteristic case, we again choose a specific direction and label it the x_3-axis. These points belong to the set

$$\{\vec{r}_u | F(\vec{r}_u) = F_t, \frac{\partial F(\vec{r}_u)}{\partial x_3} > 0, \frac{\partial F(\vec{r}_u)}{\partial x_1} = 0, \frac{\partial F(\vec{r}_u)}{\partial x_2} = 0\}. \tag{4.17}$$

For given ν_t a set of disconnected bags each of which is generally multiply connected are defined. As we travel upward from negative to positive x_3, the upcrossing points are those points on the bag tangent to the $x_1 - x_2$ plane from which the high ν interior of the bag is entered when x_3 is further increased.

The statistics of the point process

$$n_{up}(\vec{r}) = \sum_u \delta(\vec{r} - \vec{r}_u) \tag{4.18}$$

can be evaluated using the methods of §3 (equation 3.4), where our 3 variables are now $y_1 = \eta_1$, $y_2 = \eta_2$, and $y_3 = F$, not η_3. The determinant of $(y_{i,j})$ is easy to evaluate. The point process is transformed into

$$n_{up}(\vec{r}) = \delta(F(\vec{r}) - F_t)\ \delta(\eta_1(\vec{r}))\delta(\eta_2(\vec{r}))|\eta_3|\ |\varsigma_{11}\varsigma_{22} - \varsigma_{12}^2|\theta(\eta_3). \tag{4.19}$$

The mean value is very similar to the Euler characteristic expression except that the absolute value of the ς-determinant is required and downcrossing points are excluded:

$$\begin{aligned}\langle n_{up}(\nu_t)\rangle &= \langle\ \delta(F - \nu_t\sigma_0)\ |\eta_3|\ \theta(\eta_3)\ |\varsigma_{11}\varsigma_{22} - \varsigma_{12}^2|\ \rangle \\ &= \frac{(\langle k^2\rangle/3)^{3/2}}{(2\pi)^2}(\nu_t{}^2 - 1 + \frac{4\sqrt{3}}{5\gamma^2(1-5\gamma^2/9)^{1/2}}\ e^{-5\gamma^2\nu_t^2/18})\ e^{-\nu_t{}^2/2}.\end{aligned} \tag{4.20}$$

The evaluation is given in Appendix 1. The extra exponential term ensures that n_{up} is always positive. Note that the result is symmetrical about $\nu_t = 0$ reflecting the fact that the contour interiors for negative ν_t are 'inside-out' versions of those for positive ν_t. This is in contrast to the peak density which is, of course, asymmetric. Nonetheless, (4.20) reproduces (4.11) with better than 10% accuracy for $\nu > 2.5\gamma$ and better than 1% by 3γ. Even for $\nu_t \approx 0$ (4.20) is never very far off; by then, the total peak number, equation (4.11b), is a good approximation.

The fraction of all peaks larger than height ν_t is therefore approximately the ratio of equations (4.20) and (4.11b)

$$f(>\nu_t) \approx 1.56\gamma^3(\nu_t{}^2 - 1 + \frac{4\sqrt{3}}{5\gamma^2(1-5\gamma^2/9)^{1/2}}\ e^{-5\gamma^2\nu_t^2/18})\ e^{-\nu_t{}^2/2}, \quad \nu_t > 2.5\gamma. \tag{4.21}$$

For very high thresholds, the number of upcrossing points provides a good estimate of the number of disconnected contour surfaces. An estimate of the average volume enclosed within one of these high contours can then be obtained by taking the ratio of the volume filled by those points above threshold within a volume V, $VP(>\nu_t) = V\ erfc(\nu_t/\sqrt{2})/2$, to the number of upcrossing points in the volume, $n_{up}(\nu_t)V$:

$$vol(\nu_t) \equiv P(>\nu_t)/n_{up}(\nu_t) \rightarrow (2\pi)^{3/2}(R_*/\gamma)^3\ \nu_t^{-3}\ \ as\ \ \nu_t \rightarrow \infty.$$

Therefore, the linear scale of the contour 'bags' decreases as ν_t^{-1}.

4.6 The Average Velocity Distribution Function for Peaks

Peaks move with a peculiar velocity $\vec{v}$. The distribution of velocity is given by the distribution function

$$f_{pk}(\vec{r},\vec{v}_0) = \delta(\vec{v}-\vec{v}_0)\ n_{pk}(\vec{r}), \tag{4.22}$$

where $n_{pk}(\vec{r})$ is the peak density 'operator' (equations 1.1 and 3.3). The statistical average of this expression is easy to obtain since, at any point, $\vec{v}$ is only correlated with $\vec{\eta}$, being statistically independent of the integration variables ν, ς. The average distribution function

$$\langle f_{pk}(\vec{r},\vec{v},\nu)\rangle d\nu d^3r d^3v = \frac{exp[-\frac{3\vec{v}^2}{2\sigma_{-1}^2(1-\gamma_v^2)}]}{[2\pi\sigma_{-1}^2/3]^{3/2}(1-\gamma_v^2)^{3/2}} d^3v\ \mathcal{N}_{pk}(\nu)d\nu d^3r \tag{4.23}$$

implies a conditional probability for the velocity of peaks

$$P(\vec{v}|\nu,\ peak) = \langle f_{pk}(\vec{r},\vec{v},\nu)\rangle/\mathcal{N}_{pk}(\nu) \tag{4.24}$$

which differs from the peculiar velocity distribution of ambient field points

$$P(\vec{v},\ field\ point) = \frac{exp[-\frac{3}{2}v^2/\sigma_{-1}^2]}{[2\pi\sigma_{-1}^2/3]^{3/2}} \tag{4.25}$$

through the presence of

$$\gamma_v \equiv \frac{\sigma_0^2}{\sigma_{-1}\sigma_1}. \tag{4.26}$$

Here, σ_{-1}, defined by equation (4.6c) for $j=-1$, is the 3-dimensional velocity dispersion of field points. Peaks have lower peculiar velocities than field points. For Gaussian filtering of power law spectra, $\sigma_{-1} \sim R_f^{-(n+1)/2}$ and $\gamma^2 = (n+1)/(n+3)$, $n > -1$, reflecting the fact that velocities are correlated over much larger distances than densities.

Such quantities as the distribution of line-of-sight velocity differences as a function of projected separation are of great interest for comparison with observations. However, they are quite difficult to calculate and are not considered in this paper.

5. SPATIAL MODULATION OF THE PEAK NUMBER DENSITY

In this section, we first discuss some of the issues and difficulties in relating observed cosmic structures to the peaks determined from the filtered linear perturbation spectrum (§5.1). A particularly simple selection criterion is adopted, which identifies a given class of cosmic objects (e.g., luminous galaxies) with peaks in the smoothed density perturbation field F_s exceeding a global threshold level F_t. The filtering scale R_s appropriate to the class of objects depends on the physical mechanisms which set the threshold and on the amount of subsequent infall to be expected, but should correspond to a mass not very much less than the characteristic mass of the objects. The motivation for this kind of selection criterion applied to galaxies and rich clusters is discussed in §5.2.

The number density of peaks above such a threshold depends strongly on whether the large scale environment in which the peak resides is overdense or underdense. To illustrate this, we show in §5.3 how the number density of peaks above threshold changes when the density perturbation field is biased by an externally imposed background level f_b.

The overdensity in a cluster or supercluster is not an externally imposed density excess. Rather, it comes from smoothing on a large scale R_b the same random field whose peaks when smoothed on the scale R_s are associated with galaxies. Depending on just how the field is smoothed on the larger scale, there are varying degrees of correlation between the small scale random field F_s and the large scale random field F_b. The number density of peaks above threshold at places where the field F_b has a given value is derived in Appendix 5 and discussed in §5.4. The ambiguity in the definition of F_b makes application of this formalism to estimating mass-to-light ratios in clusters of galaxies imprecise. We argue that a top hat smoothing for F_b probably best corresponds to the way clusters are defined observationally, and present some numerical results which suggest that an enhancement by ~ 5 in the number of galaxies per unit mass in rich clusters compared with the universe as a whole is plausible. Therefore, a global density parameter $\Omega = 1$ may not be inconsistent with virial mass estimates for rich clusters even if there is no separation of baryons from dark matter on large scales.

5.1 Selection Criteria for Cosmic Structures

As we have discussed in the introduction, we wish to identify the sites of formation of condensations of mass $\sim M$ with the local maxima of the initial density field when smoothed over a filtering scale R_f where $M \propto R_f^3$. The structures that have formed by time t will be among those that have gone nonlinear $F(R_f,t) = \nu(R_f)\sigma_0(R_f,t) > 1$. By itself, this is a threshold criterion. Other criteria would presumably have to be satisfied as well.

Once σ_0 is fixed at some scale (e.g. that of bright galaxies) at some time (e.g. the present), it is known at all other smoothing radii from the power spectrum, and at all other times by the linear fluctuation evolution law. That is, once the spectrum and cosmological parameters are set, the only other free parameter determines the overall spectrum normalization. One may take this to be, for example, the 'redshift of galaxy formation'. For a given smoothing scale, the nonlinear threshold function $\sigma_0^{-1}(R_f,t)$ decreases with time, so peaks of progressively lower height $\nu > \sigma_0^{-1}$ will have begun to collapse. In hierarchical scenarios, σ_0^{-1} is a monotonically increasing function of R_f. For the adiabatic and isocurvature cold dark matter spectra, Figure 4.1 shows $log(\sigma_0(R_f))$ up to a constant which would be fixed by the normalization criterion. At a given time, only rare high ν large scale objects would have gone nonlinear though even *rms* fluctuations may have collapsed on smaller scales. Even on the smaller scales, it may be that only rare high ν peaks form observationally interesting structures with the *rms* peaks ($\nu \approx 1$) forming under-luminous objects.

The sort of objects that these nonlinear structures form depends upon the details of the astrophysics of collapse. Ideally, there would be a unique selection function in $\nu - R_f$ or $F - M$ space for each type of cosmic object. This simplified view is at the heart of the extensive applications of the influential spherical top hat model of Gunn and Gott (1972) to hierarchical scenarios by Gott and Rees (1975), Rees and Ostriker (1977), White and Rees (1978), Peebles (1980), Faber (1982), Silk (1983) and Blumenthal *etal.* (1984), among others. We discuss this model in more detail in §5.2. However, such applications neglect one of the central difficulties in working with hierarchical random fields, the *cloud-in-cloud problem* that within clouds centered on peaks determined after large scale smoothing are smaller scale clouds which are themselves made up of subclouds. Though such substructure is smoothed away, the details of it may be crucial in determining the nature of the final object that forms. For example, stars may form in the subclumps and become supernovae early enough to blow away the gas before the larger scale fluctuation collapses; or, mergers of subclumps within the smoothed cloud may be deciding influences. The environment may also play a large role in determining the final object which forms. For example, if angular momentum is crucial, then the position and height of neighboring clouds may be necessary information to determine the degree of

tidal-torqueing (though the local tidal field smoothed on large scales might suffice). The asymmetry of the peak will certainly play a role if angular momentum is important, and will be significant for such issues as degree of ellipticity of the final object.

In this paper we focus mostly on the density and distribution of points required to be peaks of the smoothed density field with a certain height or range of heights. The statistics of these points is that of the point process $n_{pk}(\vec{r}) = \sum_p \delta(\vec{r} - \vec{r}_p)$ defined in §1. In principle, we can select the points $\vec{r}_p$ according to any prescription necessary to define the local initial conditions required to make a given class of cosmic objects. This may prove tractable if the added criteria are not too complex. For example, background field constraints are discussed in §5.3, and shape constraints are discussed in §7. An angular momentum constraint might also be amenable to a statistical treatment. However, ellipticals forming as a result of the merging of two spirals will surely not represent an analytically tractable class. Nonetheless, even in this case, analysis of numerical N-body studies within the framework of constrained point process densities should prove useful.

We regard the filtering operation as essential in dealing with fluctuation spectra with power covering a wide range of scales. Otherwise, the statistics is dominated by the smallest scale phenomena. However, the best choice of smoothing for a given astronomical object is debatable. One of the major difficulties with filtering the density field is that, in some cases, peaks on larger filtering scales may just be smoothed versions of peaks on smaller scales. This problem is especially acute for steep power spectra with high n, for most of the clouds are intrinsically of small scale. Fortunately, the cold dark matter spectra that we use throughout this paper to illustrate the use of our statistical results are not overly plagued by this difficulty due to their being relatively flat on small scales. The structure and average density of the peaks is predominantly determined by the nature of the spectrum near the filtering scale provided it is monotonically decreasing (§4.2). Nonetheless, the optimal choice of filter function is debatable and conclusions drawn which are sensitive to the specific choice are suspect.

The mass function for a given class of objects $n(M)dM$ would be highly desirable to obtain. Unfortunately, due to the cloud-in-cloud problem, we have not come up with an adequate definition. The obvious choice, analagous to that given by Press and Schecter (1974), is $n(M)dM = \frac{1}{3}(dn_{pk}/dlnR_f)\ dM/M$, with the total derivative including partial derivatives with respect to $\nu_t(R_f)$, $\gamma(R_f)$, and $R_*(R_f)$. However, this choice ignores the strong correlation between small mass clouds and large mass clouds that we derive in §5.4 and does not include the loss of small scale objects due to incorporation in larger ones, the essence of the hierarchical process.

In spite of the difficulties associated with the identification of proto-objects in the smoothed linear density field, and, if appropriate, in $\nu - R_f$ space, we feel our method offers a powerful approach within which to test in detail the viability of simple hypotheses for structure formation.

5.2 Global Thresholds

Here and in §6, we explore the properties of that subset of peaks which have initial amplitude above some threshold value ν_t, typically $\sim 2 - 3$. We specify that the threshold should be global, by which we mean that it should be constant throughout the whole of space.

Abell's rich clusters are an example of a set of objects which may reasonably be associated with high peaks above a global threshold. It is fairly clear that in these objects, which contain only a small fraction of all galaxies, we are seeing the high mass end of the distribution of those perturbations which had sufficient amplitude to have collapsed by now. The selection criterion is therefore that the linear extrapolation of the height of the peak in the density perturbation field, smoothed on an appropriately large mass scale, should exceed a value f_c, corresponding to complete collapse at the present time t_0: $\nu_t > f_c \sigma_0^{-1}(R_f, t_0)$. Such a threshold is certainly global.

For spherical top hat models, the perturbation turns around at $F = 1.06$ and is completely

collapsed when F is

$$f_c = \frac{3}{5}(\frac{3\pi}{2})^{2/3} = 1.69. \tag{5.1}$$

The collapsing structure virializes at half the radius of maximum expansion. This would give a density contrast at the time of 'complete collapse' of about 170. For an isolated structure, the density contrast would grow as $(1+z)^{-3}$ thereafter. These results hold for a flat universe, $\Omega = 1$. In a low density universe, collapsing structures virialize at a density contrast which is, very roughly, Ω^{-1} times larger than in the flat universe.

The simple spherical top hat model is clearly an unrealistic idealization. We will discuss the expected shapes of peaks in §7. Anisotropy will be amplified as the collapse proceeds and, at least for collapse without substructure, high densities will first be reached in a planar configuration, as discussed by Zeldovich (1970). However, in order to form a compact configuration, such as a rich cluster of galaxies or an individual galaxy, it is necessary to collapse in all three dimensions. We feel that equation (5.1) is a reasonable estimate of the linear density contrast required for such a collapse to have occurred. In this regard, it is encouraging to note that within an Abell radius of the center of a rich cluster, a scale which Abell clearly felt characteristic of the extent of these objects, the mean density is around 200 times the critical density, in good accord with the simple theory.

The same collapse criterion could also be applied to galaxies. *Bright galaxies* may tend to form preferentially around high peaks, with the collapsed structures of lower ν forming underluminous galaxies, or structures not recognized as galaxies at all. This hypothesis has been termed ***biased galaxy formation***. The suggestion is that bright galaxy formation is a 'self-limiting' process, so that once a small fraction of material has turned into galaxies by a time t_g (redshift z_g), the conditions in the external medium are modified sufficiently to suppress the formation of bright galaxies elsewhere. With f_c the linear extrapolation of the height of the peak to the time of complete collapse and R_s the galactic filtering scale, the threshold now becomes

$$\nu_t = f_c \sigma_0^{-1}(R_s, t_g) = f_c(1+z_g)\sigma_0^{-1}(R_s, t_0). \tag{5.2}$$

One can imagine many processes by which newly forming galaxies can influence their environment, including heating by radiation, shocks or energetic particles (e.g., Rees 1985 and Silk 1985). However, not all of these may give rise to segregation of the kind we are discussing here. If, for example, the feedback mechanism is of limited spatial range, then it is inappropriate to model the effect by a global threshold. Unfortunately, it is difficult to develop a convincing case for any particular threshold mechanism, given our poor understanding of the details of star formation and galaxy formation.

The feedback mechanism cannot be expected to take full effect instantaneously, so the physical threshold will not be perfectly sharp. The threshold function $t(\nu/\nu_t)$ introduced in equation (4.13) is our attempt to model this in an ad hoc way. The threshold level ν_t is defined so that the probability of a peak becoming a 'galaxy' is 1/2 when $\nu = \nu_t$. Since the differential number density of peaks increases rapidly with increasing ν for $\nu > 2$, unless the threshold function is fairly sharp more 'galaxies' may form from peaks with $\nu < \nu_t$ than from peaks with $\nu > \nu_t$.

An important constraint on the choice of ν_t (or on the consistency of the model if ν_t is determined from other considerations) is that the number density of peaks meeting the threshold criterion be at least roughly equal to the number density of the class of objects one is trying to represent. The ν_t thus chosen will depend on the filtering scale R_s used to pick out the peaks and on the sharpness of the threshold, through equation (4.12).

The precise relation between the galaxy filtering scale R_s and the characteristic mass of luminous galaxies in the present universe is not obvious. The mass enclosed by a Gaussian smoothing function is

$$M_f = (2\pi)^{3/2}\rho R_s^3 = 4.37 \times 10^{12}\ R_s^3\ h^{-1}\ M_\odot, \tag{5.3a}$$

and for top hat smoothing is

$$M_{TH} = (4\pi/3)\rho R_{TH}^3 = 1.16 \times 10^{12}\ R_{TH}^3\ h^{-1}\ M_\odot, \tag{5.3b}$$

with R_s and R_{TH} in units of $h^{-1}Mpc$. While R_s slightly less than 1 $h^{-1}Mpc$ would give a mass (including the dark halo) characteristic of present luminous galaxies, the threshold is presumably set considerably before the present, and the mass collapsing at the earlier time might be considerably smaller. Average density profiles around peaks (§7) suggest that a significant amount of mass will be associated with a high peak in excess of that given by equation (5.3a). On the other hand, if R_s is too small, the correspondence between peaks above threshold on the scale R_s and present galaxies will not be at all one-to-one, due to merging of neighboring peaks and deviations from the average profile. Our threshold hypothesis is at best a crude representation of what a complicated process of galaxy formation implies for overall average properties of galaxies.

Some attempts have been made to associate the morphological type of the galaxy with the height of the peak in the density perturbation field. Blumenthal *etal.* (1984) have argued that high ν peaks are relatively isolated, are therefore less subject to tidal-torqueing, and are more likely to become elliptical galaxies. However, tidal-torqueing is a local process, so the height of the peak relative to the global σ_0 is not relevant. The argument would suggest that there should be fewer low spin objects in the high density environment of a protocluster, contrary to observation. An alternative explanation for the distinction between ellipticals and spirals is that elliptical galaxies form from mergers of smaller scale peaks, while spirals form from relatively isolated peaks which are able to accumulate mass by infall over a substantial range of radius. This would also associate ellipticals with higher ν peaks smoothed on the mass scale of the final galaxy, but the difference in apparent angular momentum would be due to different degrees of radial infall of baryonic matter.

5.3 Modulation of Number Density by a Background Field

The field $F_s(\vec{r})$ defines a population of peaks, and a global threshold ν_t selects a subset of these peaks with a certain global number density $n_{pk}(\nu_t)$. However, one is often interested in the local number density of a class of objects, such as the density of galaxies in a cluster or the density of clusters of galaxies in a supercluster, and the relation of this local number density to large scale variations in the density perturbation amplitude. Here we consider adding an external, statistically independent background field f_b to F_s, with the understanding that f_b is a slowly varying function of position compared to the mean separation of the peaks in F_s. The number density of peaks of the combined field above the spatially uniform threshold characterized by ν_t will depend on the local value of f_b and therefore will fluctuate on large scales. The number density will be enhanced where f_b is positive and suppressed where f_b is negative. This construction will prove useful for calculating correlations of galaxies on large scales in §6, but, as we shall see in §5.4, it is not appropriate, except for providing qualitative insight, if the background is the large scale density perturbation associated with a particular cluster of galaxies.

Let $E(f_b)$ be the ratio of the local number density of peaks in the presence of f_b to the global number density with $f_b = 0$. To calculate the local number density, note that the combined field exceeds the global absolute threshold f_t when F_s exceeds a local threshold $f_t - f_b$. In terms of ν, the local threshold is then $\nu_t - f_b/\sigma_{0s}$. In the limit f_b is approximately uniform on the scale R_s, such quantities as γ and R_* are unaltered, and the enhancement factor is just

$$E(f_b) = n_{pk}(\nu_t - f_b/\sigma_{0s})/n_{pk}(\nu_t). \tag{5.4}$$

For a perfectly sharp threshold in the limit $\gamma\nu_t \gg 1$, equation (4.14) is a good approximation to $n_{pk}(\nu_t)$. Substituting this expression into equation (5.4) and keeping only the leading term, we find

$$E(f_b) \approx exp(\nu_t f_b/\sigma_{0s}), \quad for\ f_b \ll \sigma_{0s},\ and\ \nu_t \gg 1. \tag{5.5}$$

The enhancement can be highly nonlinear in f_b even though f_b may be small compared with one at the time galaxies form.

Unfortunately, equation (5.5), while qualitatively correct, is not quantitatively accurate for the range of ν_t relevant to galaxies. We find that a Taylor series expansion of $ln\ (E)$, keeping only the terms which are linear and quadratic in f_b,

$$E(f_b) \approx exp\left[\alpha(\nu_t,\gamma)(f_b/\sigma_{0s}) - \frac{1}{2}\beta(\nu_t,\gamma)(f_b/\sigma_{0s})^2\right], \tag{5.6}$$

is a good approximation as long as $\nu_t > 2$. The coefficients α and β also depend upon the sharpness of the threshold, as characterized by the index q (equation 4.13), and are obtained by differentiating equation (4.12) inside the integral and integrating by parts:

$$\begin{aligned}\alpha &= \int \frac{dt(\nu)}{d\nu}\mathcal{N}_{pk}(\nu;\gamma)/n_{pk}(\nu_t),\\ \beta &= \alpha^2 - \int \frac{d^2t(\nu)}{d\nu^2}\mathcal{N}_{pk}(\nu;\gamma)/n_{pk}(\nu_t).\end{aligned} \tag{5.7}$$

Table 5.1 shows how these parameters vary with q and ν_t. A perfectly sharp threshold ($q = \infty$) has the largest value of α for a given ν_t, and therefore the most rapid increase of peak density with f_b, but this maximum α is still less than the value $\alpha = \nu_t$ as estimated from the high threshold limit enhancement factor, equation (5.5). Equation (5.6) is accurate to within a few per cent out to $f_b/\sigma_{0s} \approx 1$ for $\nu_t \approx 2$ and out to $f_b/\sigma_{0s} \approx 2$ for $\nu_t \approx 3.5$.

5.4 The Segregation of 'Protogalaxies' from the Mass

The enhancement E is a strongly increasing function of f_b, particularly if the 'galaxies' contain a small fraction of the mass, so ν_t is high. Notice also that the quadratic term in equation (5.6) reveals an asymmetry. For fixed $|f_b|$ the suppression factor in a 'protovoid' will exceed the enhancement factor in a 'protocluster'. The enhancement factor saturates when the local effective threshold is small and almost all the peaks are counted as 'galaxies', but there is no upper limit to the suppression factor in a 'void'.

In the global threshold model 'galaxies' are born with large scale clustering which gives a distorted picture of the mass density contrast. One can think of the galactic peaks as having been 'painted on' to their Lagrangian sites at early times. The Lagrangian points flow through Eulerian space toward mass concentrations, carrying the painted points along with them. The 'painted on' density contrast has the same sign as the true mass density fluctuations and so the 'galaxy' clustering will always exceed that of the underlying density field which is determined solely by gravitational forces.

One might be able to observe the effects of this statistical enhancement directly, before there has been considerable dynamical evolution on the scale of the clusters. This would be the case for galaxies observed at high redshift: a protocluster would appear as a strong galaxy density enhancement even though the underlying density contrast was small at that time. Similarly, a supercluster at the present time may appear as a strong enhancement of rich clusters. In other applications the mass density field will be nonlinear. We can still obtain the number of 'galaxies' per unit mass, and the overall number density can then be determined if we know the motion of the mass.

The high ν_t result of equation (5.5) gives a very rough idea of the parameters required to obtain a strong non-linear enhancement of the galaxy density within a protocluster. Model the cluster as an externally imposed 'top hat' field with amplitude just sufficient to collapse by the present. The condition for non-linear enhancement, $\nu_t(f_b/\sigma_{0s}) > 1$, is satisfied provided

$$\nu_t^2 > 1 + z_g, \tag{5.8}$$

where z_g is the redshift at which peaks associated with galaxies collapse. Note that this condition is independent of the mass of the protocluster and the present *rms* amplitude of density fluctuations on the cluster scale. Since only collapsed systems are amenable to virial analyses, there is a lower limit to measured galaxy-per-unit-mass ratios and presumably a corresponding upper limit to measured mass-to-light ratios. Such an upper limit does seem to be present in the galaxy cluster data. While the upper limit has been interpreted as evidence for $\Omega < 1$, we see that it has a natural interpretation in an $\Omega = 1$ universe.

There are a couple of reasons to think that the enhancement factor in this externally imposed 'top hat' protocluster model, as calculated from equation (5.6) rather than from the crude approximation of equation (5.5), is an underestimate of the enhancement in a real protocluster. A more realistic model would not have a uniform density excess in the protocluster volume, and the average of an exponential is greater than the exponential of the corresponding average. Perhaps more seriously, a protocluster is not an externally imposed background. It is defined by a large scale smoothing of the same density perturbation field F which is smoothed on a smaller scale to get the field F_s.

Let $F_b(\vec{r})$ denote this large scale smoothed field. At least roughly, a region which has collapsed by the present is a region where the average overdensity of the extrapolated linear density perturbation field is greater than a value f_c. For a spherical collapse $f_c = 1.69$, and the appropriate average is over a sphere centered on the protocluster containing a mass equal to the final collapsed mass. This suggests defining F_b by a spherical top hat smoothing of F. Of course, actual collapses are usually not spherical, but the spherical model should be a reasonably good guide as to what sort of overdensity is required to produce an Abell cluster, as discussed in §5.2. (See §4.2 for the power spectrum associated with top hat filtering.)

In Appendix 5, we derive $\mathcal{N}_{pk}(\nu_s, \nu_b)d\nu_s d\nu_b$, the joint probability per unit volume that there is a peak in the F_s field with height $\nu_s \equiv F_s/\sigma_{0s}$ in the range $d\nu_s$ and that the background field has height $\nu_b \equiv F_b/\sigma_{0b}$ in the range $d\nu_b$ at the peak (equation A5.5). The differential number density of peaks at a point where ν_b has a specified value is the conditional density (equation A5.11) $\mathcal{N}_{pk}(\nu_s|\nu_b) = \mathcal{N}_{pk}(\nu_s, \nu_b)(2\pi)^{1/2} exp(\nu_b^2/2)$, and the local number density of peaks satisfying a global threshold criterion is

$$n_{pk}(\nu_t|\nu_b) = \int_0^\infty d\nu_s \; t(\nu_s/\nu_t) \; \mathcal{N}_{pk}(\nu_s|\nu_b). \tag{5.9}$$

The enhancement factor is

$$E(\nu_b) \equiv n_{pk}(\nu_t|\nu_b)/n_{pk}(\nu_t). \tag{5.10}$$

Numerical evaluation of equation (5.9) is straightforward since the form of the integral is similar to the standard number density integral.

The enhancement in $\mathcal{N}_{pk}(\nu_s|\nu_b)$ is much greater for the higher peaks than for the lower peaks. If elliptical galaxies are associated with high ν peaks, this could be part of the reason ellipticals are much more common in rich clusters than in the field.

Numerical results for the enhancement factor are sensitive to the smoothing function used to define F_b, even when the background scale R_b is very large compared to the scale R_s of the peaks. The parameter $\epsilon \equiv \langle \nu_s \nu_b \rangle$ is a measure of this, since in the high threshold limit $E(\nu_b) = exp(\nu_t \epsilon \nu_b)$. For a sharp filter in k-space defining F_b, $\epsilon = \sigma_{0b}/\sigma_{0s}$, leading to results similar to those in §5.3, but the corresponding smoothing function is an oscillating spherical Bessel function. For a Gaussian filter, $\epsilon \approx 2^{(n+3)/2} \; \sigma_{0b}/\sigma_{0s}$ where n is the effective spectral power law index on the background scale R_b. Top hat smoothing in real space gives intermediate results. We feel that top hat smoothing makes the most sense physically for treating rich clusters, though this is certainly debatable.

Another problem is that a top hat centered on the cluster gives an expected number density for a point near the center of the protocluster, which may be an overestimate of the average number density in the protocluster. Also, such a top hat may not be representative of all points with the same

value of F_b. In order to resolve some of these uncertainties it may be necessary to resort to numerical realizations of these random fields in which the relevant quantities can be directly measured.

Some numerical results for the enhancement factor E are presented in Table 5.2. The examples are from Table 6.1a (see also Fig.6.1a). The perturbation spectrum is the adiabatic one appropriate to cold dark matter (Appendix 7). In all cases, the cluster-scale top hat has a mass $5 \times 10^{14}\ h^{-1}M_\odot$, corresponding to the top hat filtering scale $7.6\ h^{-1}Mpc$, and the constrained value $F_b = 1.69$, corresponding to spherical collapse at the present. For a given sharpness parameter q, the threshold level ν_t is fixed by requiring that the number density of selected peaks equal the number density of galaxies in the CfA redshift survey (Davis and Huchra 1982), $0.01\ h^3Mpc^{-3}$. The range of values of the enhancement factor comfortably brackets the range desired to make cluster M/L values consistent with a global $\Omega = 1$. We also give in each case the value of ν_b determined by normalizing the overall density perturbation amplitude to make the present galaxy-galaxy correlation $5\ h^{-1}Mpc$, using the results of §6. The quantity z_t is the redshift of the collapse of a peak of height ν_t, while z_g is the average collapse redshift of a peak which becomes a 'galaxy'. The determination of these quantities is discussed more fully in §6.6.

The values of the enhancement factor would be reduced by about $30\% - 40\%$ if we were to ignore the correlation between the background field and the high frequency field which generates the peaks (i.e., take $\epsilon = \sigma_{0b}/\sigma_{0s}$).

A factor of 5 enhancement relative to the overall galaxy number density $0.01\ h^3Mpc^{-3}$ within the mass of $5 \times 10^{14}\ h^{-1}M_\odot$ corresponds to roughly 90 'luminous galaxies' ignoring mergers of individual peaks. Of course, what is perceived to be the cluster would be larger than the completely collapsed core. The mass of $5 \times 10^{14}\ h^{-1}M_\odot$ is just a little less than what is contained in an Abell radius if the mass overdensity is 170.

The values of ν_b in Table 5.2 are uncomfortably large, given that rich clusters are not all that rare. According to Gaussian statistics, less than 0.2% of the mass should have $\nu_b > 3$. Things are actually somewhat better than this, since a single point of the smoothed density field with $F_b > 1.69$ can mean a whole cluster will have collapsed. This large value of ν_b might be taken as an indication there is relatively more power on large scales in the real universe than predicted by the simple adiabatic CDM spectrum or as a suggestion that the simple uniform threshold biasing scenario needs modification. Of course, a larger value of R_{TH} for clusters makes the problem worse.

We now discuss some further consequences of the 'biasing' hypothesis. The mass-to-light ratio for any large system is, according to equation 5.6 or 5.10, determined only by the net initial density contrast, and therefore by the final density of the system. Hence the mass-to-light ratios should be *anticorrelated* with the mass density of the system. Applying this test is complicated by the fact that mass and density are both derived from virial analysis. In the absence of any intrinsic correlation we would expect any scatter in the observed quantities to introduce a correlation which is of the opposite sign to that we have predicted. It may therefore be very difficult to disentangle these two effects from the data which, at least in the case of very rich clusters, seem to be consistent with zero correlations between mass and density (Dressler 1978).

Another consequence of the hypothesis is that one should observe an increase of M/L in the outer parts of clusters, since F_b goes down. A problem here is that the radial dependence of M/L depends on the assumed anisotropy of the orbits. Kent and Gunn (1978) claim that the data for Coma are consistent with constant M/L and constant anisotropy. A trend of M/L in the direction predicted here would be consistent with the observations if the outer orbits tend to be more radial, as would be expected in the type of scenario we are considering. Also, we would only expect to see the trend at large radii, since in the inner cluster the galaxies and mass should be fairly well mixed. However, at large radii ($> 2\ h^{-1}Mpc$ say) the usual assumptions made in deriving mass-to-light ratios are not likely to be accurately obeyed.

So far we have concentrated on galaxies as 'rare events'. As noted above, it is more likely

that rich clusters can be associated with the high peaks of the initial density field and it is more straightforward to estimate ν_t and σ_{0s} than for the galaxies. The filtering scale R_s is now taken to be that associated with rich clusters ($\sim (7-11)\ h^{-1}Mpc$ for top hat filtering). In this case $\nu_t/\sigma_{0s} \sim 3-4$, whereas for galaxies this quantity is, according to our estimate above, close to unity. Thus, in large low-density systems such as superclusters, the density enhancement of clusters ($\sim exp((\nu_t/\sigma_{0s})F_b)$) can be much larger than that for galaxies. Similarly, one can expect that the suppression of structure formation in a 'void' would be much more complete for clusters and that regions which are devoid of clusters would be much larger than those which are devoid of galaxies.

We have shown here that if galaxies are associated with the high peaks of the initial density field then they will be strongly segregated from the mass in all objects which are sufficiently dense to have collapsed by the present. The effect is predicted to be much stronger for rich clusters and these may be segregated from the mass on much larger scales in systems which are still in the linear regime. This segregation means that galaxies may give a strongly distorted picture of the degree of matter clustering on cluster scales. We now turn to the clustering statistics for high peaks which show that galaxies and other objects may give a strongly distorted picture of the distribution of mass on *all* scales.

6. CORRELATIONS

6.1 OVERVIEW

We now turn to a statistical analysis of the effects of a threshold for galaxy formation on the clustering of galaxies. In §5 we saw how a modest overdensity on some large mass scale can lead to a strong enhancement in the local density of galaxies. The same mechanism has an effect on the statistical measures of galaxy clustering, such as the two-point correlation function. A statistically enhanced clustering of small groups of galaxies relative to the mass distribution may explain the relatively low amplitude of two-point velocity correlations found in redshift surveys (Davis and Peebles 1982) compared to what is expected in an $\Omega = 1$ cold dark matter universe if the galaxies trace the mass. Substantial fluctuations in how galaxies are distributed relative to the matter also imply a galaxy-galaxy correlation function which would have a significant amplitude even at early times when the density perturbation amplitude is relatively small on cluster scales, a prediction which should be testable in the near future.

Complete information about the statistical distribution of galaxies in space is contained in the hierarchy of n-point correlation functions (Peebles 1980). Reliable observational data are available for the two- and three-point functions. This data is an important constraint on any theory of large scale structure in the universe. Correlations between clusters of galaxies probe the structure of the universe out to even larger scales. The major goal of this section is to calculate two- and three-point correlation functions of galaxies starting from the hypothesis that galaxies are identified with high peaks of the linear density perturbation field. The qualitative picture is clear from §5: peaks above the threshold are much more likely to occur in regions where on a large scale the level of the density perturbation field is greater than zero.

Our primary focus will be on the initial correlations in the peaks above threshold, before any dynamical corrections due to the amplitude of the large scale perturbations becoming larger than the amplitude of the statistical fluctuations in the density of the peaks. To linear order in the amplitude of the density perturbations it is reasonably straightforward to set up a calculation of dynamical corrections, but all we will do here is to consider the simplest case, when the zeroth-order (statistical) correlations are small.

In principle, one can calculate the zeroth-order (in dynamics) correlations exactly using the machinery outlined in §2 and 3. The n-point correlation function requires constructing a joint probability distribution in $10n$ variables; at each point the value of the smoothed density field F_s, the three first derivatives η_i and the six second derivatives ς_{ij}. While the covariance matrix for these variables

at a single point has a reasonably simple form, there are a large number of non-zero correlations between the variables at different points. Even for the two-point function, the task of integrating over all of these variables is not pleasant to contemplate. Fortunately, approximations can be made to greatly simplify the calculations, and these approximations are reasonably accurate once the correlation lengths are more than a few times the smoothing radius R_s used to define the peaks.

In the context of a direct calculation of the peak correlations one key approximation is to *neglect derivatives* of the density-density correlation function of F_s, $\xi(r_{ij})$, which appear in the correlation matrix. If the density perturbation power spectrum is roughly a power law of index n, then the normalized mass density correlation function

$$\psi(r_{ij}) \equiv \xi(r_{ij})/\xi(0) \equiv \langle F_s(r_i) F_s(r_j) \rangle / \sigma_0^2 \tag{6.1}$$

falls off as $r^{-(n+3)}$, while $d^k\psi/dr^k$ falls off as $r^{-(n+3+k)}$. For values of the index $n < -2$ the neglect of the gradients of ψ may be justified when ψ itself is still not very much less than 1.

The only cross-correlations between points which survive in this approximation are

$$\langle \nu(r_i)\nu(r_j) \rangle = \psi(r_{ij}) = \psi_{ij}, \quad i \neq j. \tag{6.2}$$

The $8n \times 8n$ part of the correlation matrix involving the first derivatives and the variables describing the anisotropic part of the second derivative matrix is now diagonal, and the integrals over these variables are identical to those in the average peak number density formula.

A further approximation, justified if the ψ_{ij} are all at least moderately small compared to one, allows the integrals over the $x(i) \propto -\nabla^2 F_s(r_i)$ variables to be done analytically. The n-dimensional integral over the $\nu(i)$ must still be evaluated numerically, but this is quite feasible for $n = 2$ at least.

We also discuss an alternative approach to the n-point correlation functions in which the density perturbation field F_s, smoothed on the scale R_s of interest for the peaks, is considered as the sum of a "background" field F_b, filtered on a scale R_b substantially larger than R_s, and a "peak" field F_p. The local density of peaks in F_p is calculated as a function of the local value of F_b using the procedures of §5. Correlations of this peak density field $n_{pk}(\vec{r})$ depend on the probability distribution of F_b, which in turn can be expressed in terms of the normalized two-point correlation function of F_b with itself,

$$\psi_b(r_{ij}) \equiv \langle F_b(r_i) F_b(r_j) \rangle / \sigma_{0b}^2. \tag{6.3}$$

Clearly, the local peak density field contains no information about the correlations of peaks at separations less than R_b. Even for separations large compared with R_b the relationship of the correlation function of n_{pk} to the correlation functions of the peaks in F_s is rather obscure, since, strictly speaking, the local density of peaks depends upon the precise smoothing function used to define F_b. However, the choice of a filter for F_b is purely a mathematical device and can have no effect on the actual correlations of peaks of F_s.

Our procedure is as follows. Pick a convenient (e.g. Gaussian) filter to define F_b. Pretend F_p is an *independent* random field, so that the power spectrum of F_p is the difference of the power spectrum of F_s and that of F_b, ignoring the correlations between F_p and F_b. Calculate the local density of peaks n_{pk} as a function of F_b, using the number density formula of §4 with a local threshold in F_p biased by F_b to keep the threshold in F_s at the desired global value. Integrate over the $F_b(\vec{r}_i)$ weighted by the joint probability distribution for the $F_b(\vec{r}_i)$ to find the n-point correlation function of n_{pk}.

The advantage of the peak-background split is that n_{pk} can be rather accurately approximated by equation (5.6),

$$n_{pk}(F_b) = n_0 \; exp[\alpha F_b/\sigma_{0b} - \frac{1}{2}\beta(F_b/\sigma_{0b})^2], \tag{6.4}$$

with α and β chosen to fit the first and second derivatives of $n_{pk}(F_b)$ at $F_b = 0$. The form of equation (6.4) allows the integral for the n-point correlation function of n_{pk} to be done *analytically*.

As discussed in more detail below, we find good agreement between the direct calculation of the two-point peak correlations and the calculation using the peak-background split provided that:

(1) R_b/R_s is large enough (> 3) so that the *local* statistical properties of F_p are nearly the same as those of F_s. The power spectrum moments σ_1 and σ_2 are then nearly equal for F_p and F_s, though σ_0 may differ substantially, since σ_0 has important contributions from fairly long wavelengths if n is close to -3.

(2) r/R_b is large enough (> 4) so that the two-point density correlation functions $\xi_b(r)$ and $\xi_s(r)$ are nearly the same.

The peak-background approach gives reliable results for the peak correlation function only for $r/R_s > 12$; so, the direct calculation of the correlations will be preferred when it is feasible. We will use the peak-background approach to obtain results for 3-point correlations and to discuss the effects of dynamical evolution of the density perturbation field on the correlations. Much of the present amplitude of the galaxy-galaxy correlation function must be due to a non-zero amplitude of the density fluctuations, rather than the zeroth-order statistical correlations by themselves.

The following discussion will focus on how to apply our formal results for the n-point correlations, and in particular for the two-point and three-point correlations. Technical details of the derivation are given in Appendix 6.

Please note that all our approximations in this Section are predicated on the density perturbation spectrum having a substantial amount of power on large scales. The power spectrum should have an index $n < -1$. In particular, if $n = 0$ the mass correlation amplitude ψ falls off exponentially (when Gaussian-filtered), and the neglect of derivatives of ψ compared with ψ is not valid even at large separations. If $n \geq 0$, the correlations are dominated by fluctuations on scales considerably smaller than the separation of the points, so the rationale for the peak-background method is totally destroyed.

6.2 Direct Calculation of Correlations

Consider the n-point correlation function of peaks satisfying some sort of threshold criterion. The n-point correlation function is defined such that

$$1 + \xi_{pk}^{(n)}(\vec{r}_1, ..., \vec{r}_n) = \langle n_{pk}(\vec{r}_1)...n_{pk}(\vec{r}_n)\rangle / \langle n_{pk}\rangle^n \tag{6.5}$$

is the joint probability that a peak is in a volume dV_i about each $\vec{r}_i$, divided by the n^{th} power of the global average peak density times the product of the dV_i. The joint probability distribution of n peaks depends upon 10 variables $F, \vec{\eta}, \zeta_A$ at each point, with the same weighting factors at each point ($det\ (\zeta)$) as in the peak number density calculation of §4. Proper choice of variables (see Appendix 1) simplifies the correlations at a single point, so that only the correlation between $\nu = F/\sigma_0$ and $x = -\nabla^2 F/\sigma_2$ is non-zero. However, cross correlations between most of the variables at *different* points are nonzero.

Our approach is based on work in progress by Bardeen, Bond, Jensen and Szalay (1985). Cross-correlations involving any derivatives of F at either point are proportional to derivatives of the density correlation function $\xi(r_{ij})$. The derivatives of ξ fall off substantially more rapidly with increasing separation r_{ij} than ξ itself, particularly when ξ falls off slowly, as it does on galactic scales in cold dark matter scenarios. By the time r_{ij} is more than 4 times the smoothing length R_s which defines the peaks, the cross-correlations proportional to derivatives of ξ are small enough that it is a reasonable approximation to set them to zero and just retain cross-correlations between $F(i)$ and $F(j)$. This simplifies the correlation matrix enough to get an explicit expression for the joint probability distribution in the $\nu(i)$ and $x(i)$, after integrating over all other variables.

To describe the result, we use a matrix notation. Let $\tilde{\nu}$ and x be n-component column vectors, with components

$$\tilde{\nu}(i) = (\nu(i) - \gamma x(i))/(1 - \gamma^2) \tag{6.6}$$

and $x(i)$, respectively. Define an $n \times n$ matrix Ψ with off-diagonal elements the normalized density cross-correlations ψ_{ij} given by equation (6.2) and with zero diagonal elements. Also, we define a related matrix

$$\mathbf{C} \equiv \Psi \ (I + \Psi/(1-\gamma^2))^{-1}, \tag{6.7}$$

with I the unit matrix. The expectation value of the product of local peak densities which appears in equation (6.5) can then be written

$$\langle n_{pk}(\vec{r}_1)...n_{pk}(\vec{r}_n)\rangle = (4\pi^2 R_*^3)^{-n} \ \left(det \ (I + \Psi/(1-\gamma^2))\ \right)^{-1/2} \\ \times \ \left(\prod_i \int d\nu(i) \int_0^\infty dx(i) \ t(\nu(i)/\nu_i) \ g(x(i),\gamma,\gamma\nu(i)) \ e^{-\nu(i)^2/2}\right) \ exp(\frac{1}{2}\tilde{\nu}^\dagger \mathbf{C} \tilde{\nu}). \tag{6.8}$$

Here, g is the integrand of the integral over x (equation A1.19) in the function $G(\gamma,\gamma\nu)$ which appears in the number density formula (equation 4.3):

$$g(x,\gamma,\gamma\nu) \equiv f(x) \ \frac{exp\ \left(-\frac{(x-\gamma\nu)^2}{2(1-\gamma^2)}\right)}{(2\pi(1-\gamma^2))^{1/2}}. \tag{6.9}$$

In general, since the matrix C is intrinsically non-diagonal and since the vector $\tilde{\nu}$ depends on both the $\nu(i)$ and the $x(i)$, the integrals in equation (6.8) do not factor into independent terms. The $2n$-dimensional integral makes the evaluation of even the 2-point correlations difficult unless further approximations are made.

One regime in which the integral does simplify is the limit in which all the ψ_{ij} satisfy the condition

$$\nu_t^2 \psi_{ij} \ll 1, \tag{6.10}$$

for then

$$exp(\frac{1}{2}\tilde{\nu}^\dagger \mathbf{C} \tilde{\nu}) \approx 1 + \frac{1}{2}\tilde{\nu}^\dagger \mathbf{C} \tilde{\nu} \tag{6.11}$$

so the integrals over x and ν can now be done separately at each point. Let $\langle\tilde{\nu}\rangle$ be the average of $\tilde{\nu}$ over all the peaks selected by the threshold criterion. The n-point function becomes

$$\xi_{pk}^{(n)}(\vec{r}_1,...,\vec{r}_n) = \sum_{i<j} \langle\tilde{\nu}\rangle^2 \ \psi(r_{ij}) = \sum_{i<j} \xi_{pk}^{(2)}(r_{ij}). \tag{6.12}$$

This is the limit discussed by Kaiser (1984a) in regard to the statistical enhancement of clustering of Abell clusters. Compared to the two-point *mass* density correlation function $\xi_\rho(r_{ij})$, the two-point density correlation function is enhanced by a factor $\langle\tilde{\nu}\rangle^2/\sigma_0^2$.

To facilitate numerical evaluation of $\langle\tilde{\nu}\rangle$ we have found an approximate formula for the average value of x at fixed ν:

$$\langle x\rangle = \gamma\nu + \theta(\gamma,\gamma\nu), \tag{6.13}$$

$$\theta = \frac{\left(3(1-\gamma^2) + (1.216 - 0.9\gamma^4) \ exp\ (-\frac{1}{2}\gamma(\gamma\nu/2)^2)\right)}{\left((3(1-\gamma^2) + 0.45 + (\gamma\nu/2)^2)^{1/2} + \gamma\nu/2\right)}. \tag{6.14}$$

The value of θ is accurate to better than 1% over the range of values of γ and $\gamma\nu$ relevant for galaxies, $0.4 < \gamma < 0.7$ and $1 < \gamma\nu < 3$. Further, it has the correct asymptotic behavior as $\gamma\nu \to \infty$. In the high ν limit, $\theta \to 0$ and $\langle x\rangle \to \gamma\nu$. Using equation (6.13), $\langle\tilde{\nu}\rangle$ at fixed ν becomes

$$\langle\tilde{\nu}\rangle = \nu - \gamma\theta/(1-\gamma^2). \tag{6.15}$$

Typically, we have $\gamma \approx 0.6$, $\nu \approx 3$, so $\theta \approx 1$ and $\langle\tilde{\nu}\rangle$ is significantly less than ν.

We have tried more than one approach to simplifying equation (6.8) in the *nonlinear regime*, all based on approximating the integrand to allow analytic evaluation of the integrals over $x(i)$ in terms of accurate interpolation formulas, leaving only the integrals over the $\nu(i)$ to be done numerically. See Appendix 6 for a discussion of some of the alternatives. What we present here is an approach which seems to offer the best combination of accuracy and ease of use.

Consider approximating the function g (equation 6.9) by a Gaussian in x,

$$g(x,\gamma,\gamma\nu) \approx g(x_m,\gamma,\gamma\nu)\ exp(-\frac{1}{2}\beta_g(x_m)(x-x_m)^2). \tag{6.16}$$

We evaluated the function g numerically for each of several pairs of values of γ and $\gamma\nu$ at three evenly spaced values of x centered close to the maximum of g, with g at the smaller and larger values about 2/3 of the maximum. The Gaussian fit to these values determines x_m and β_g at each γ, $\gamma\nu$. We found interpolation formulas for $x_m(\gamma,\gamma\nu)$ and $\beta_g(\gamma,\gamma\nu)$ which give a good fit over the range of values of γ relevant to reasonable density perturbation spectra $(0.4-0.7)$ and over the range of $\gamma\nu$ where the integrand in equation (6.8) is large. Some care was needed because $g(x)$ is not particularly close to a Gaussian for x substantially different from x_m. The result of all this is

$$x_m = \gamma\nu + \frac{(3(1-\gamma^2)+(1.1-0.9\gamma^4)\ exp\ (-\gamma(1-\gamma^2)(\gamma\nu/2)^2))}{((3(1-\gamma^2)+0.45+(\gamma\nu/2)^2)^{1/2}+\gamma\nu/2)}, \tag{6.17}$$

$$\beta_g(x_m) = \frac{1}{1-\gamma^2} + \frac{3(x_m^4+3)}{(x_m^2-3)^2(x_m^2+164\ exp(-\frac{5}{8}x_m^2))}. \tag{6.18}$$

The value of $g(x_m)$ can be calculated directly through equation (6.9) with $f(x_m)$ approximated by

$$f(x_m) \approx x_m(x_m^2-3)+(2.41x_m^2+1.73)\ exp(-\frac{5}{8}x_m^2). \tag{6.19}$$

The interpolation formulas are based on large x asymptotic expansions since x_m is grreater than 2 even for $\gamma\nu < 1$; over this range, they are accurate to better than 1%.

Note that though the interpolation formula for x_m is similar to that for $\langle x\rangle$, it is not identical. The difference reflects the non-Gaussian nature of the true $g(x)$. A modification which improves the accuracy of the final results for the correlations significantly is to replace x_m in equations (6.18) and (6.19) by $\langle x\rangle$, so in equation (6.8)

$$g(x) \approx g(\langle x\rangle)\ exp(-\frac{1}{2}\beta_g(\langle x\rangle)(x-\langle x\rangle)^2). \tag{6.20}$$

In particular, this guarantees accuracy when the correlation amplitude is small and equation (6.12) applies.

Since the integrals over the $x(i)$ in equation (6.8) are now integrals over the exponential of a quadratic form, they can be evaluated analytically using standard techniques (see Appendix 6). Let $\hat{\nu}$ be the column vector with components

$$\hat{\nu}(i) = (\nu(i)-\gamma\langle x\rangle)/(1-\gamma^2), \tag{6.21}$$

and let β be the $n\times n$ matrix with the i^{th} diagonal element equal to $\beta_g(\langle x(i)\rangle)$ and all off-diagonal elements zero. Equation (6.8) reduces to

$$\langle\prod_i^n n_{pk}(\vec{r}_i)\rangle = (4\pi^2R_*^3)^{-n}\left(\prod_i\int d\nu(i)\ t(\nu(i)/\nu_t)\ g(\langle x(i)\rangle,\gamma,\gamma\nu(i))\ exp(-\nu(i)^2/2)\right)$$
$$\times\ exp\left(\frac{1}{2}\hat{\nu}^\dagger\ \mathbf{C}(\beta-\mathbf{C}\gamma^2/(1-\gamma^2)^2)^{-1}\ \beta\hat{\nu}\right)\ (2\pi)^{n/2}(det\ (\beta-\mathbf{C}\gamma^2/(1-\gamma^2)^2)\ )^{-1/2}. \tag{6.22}$$

Numerical evaluation is very manageable for $n = 2$ and feasible for at least $n = 3$.

For consistency, the global average number density of peaks $\langle n_{pk} \rangle$ should be calculated using the same approximation for $g(x)$ that was used to arrive at equation (6.22),

$$\langle n_{pk} \rangle = (4\pi^2 R_*^3)^{-1} \int d\nu \left(\frac{2\pi}{\beta_g(\langle x \rangle)} \right)^{1/2} t(\nu/\nu_t)\, g(\langle x \rangle)\, exp(-\nu^2/2), \tag{6.23}$$

so numerical errors in the absolute densities have a minimal effect on $\xi_{pk}^{(n)}$ as found from equation (6.5).

6.3 Peaks on a Fluctuating Background

A heuristic approach to understanding why the peak correlations are enhanced relative to the density perturbations is to consider the local density of peaks as a function of the mass overdensity averaged on a scale somewhat larger than than the mean separation of peaks. The local peak density increases rapidly with an increase in the *background* mass overdensity since the effective threshold is reduced. In this approach the peak correlations are an amplified reflection of the background density correlations. The local peak density is treated as a continuum process rather than the point process of §4 and 5.

To make this approach explicit and quantitative, write the full (smoothed) density field F_s as the sum of a *peak* field F_p and a *background* field F_b. The latter is defined by smoothing the full density field F on a scale R_b (we drop the subscript f in this section) larger than the scale R_s used to define the peaks in F_s. The smoothing can be thought of as a convolution integral as in §4,

$$F_b(\vec{r}) = \int C_b(|\vec{r} - \vec{r}\,'|) F(\vec{r}\,')\, d^3\vec{r}\,', \tag{6.24}$$

or as a low pass filter $C_b(k)$ acting on the Fourier transform of F just as for F_s itself. In some sense the peak field F_p should describe the local properties of the peaks in F_s, whose long range clustering properties are defined by the information in F_b. In this spirit, if $n_{pk}(F_b)$ is the local density of peaks in F_p, we want to compute the n-point correlations from an expression of the form

$$\langle n_{pk}(\vec{r}_1) ... n_{pk}(\vec{r}_n) \rangle = \int \prod_i n_{pk}(\nu_b(i))\, d\nu_b(i)\; P(\nu_b(1), ..., \nu_b(n)), \tag{6.25}$$

with

$$\nu_b = F_b/\sigma_b, \quad \sigma_b^2 = \langle F_b^2 \rangle. \tag{6.26}$$

Since F_b is a Gaussian field in its own right, its probability distribution for the values $\nu_b(i)$ in a range $d\nu_b(i)$ about $\nu_b(i)$ at n-points is, in matrix notation,

$$P(\nu_b(i), ..., \nu_b(i)) = (2\pi)^{-n/2} [det\ (I + \Psi)]^{-1/2}\ exp[-\frac{1}{2} \nu_b^\dagger (I + \Psi)^{-1} \nu_b]. \tag{6.27}$$

The matrix Ψ is the same normalized two-point density correlation matrix for the field as was introduced in §6.2 for F_s, with off-diagonal elements

$$\psi_{ij} = \xi_{\rho b}(r_{ij})/\sigma_b^2 = \langle \nu_b(i) \nu_b(j) \rangle. \tag{6.28}$$

In §5 we discussed how to calculate directly the probability per unit volume of a peak in F_s subject to the constraint that the field smoothed on a larger scale has a particular value F_b. It might seem reasonable to use this as an estimate of the local peak density $n_{pk}(\nu_b)$ in equation (6.25). However, the results for the correlation functions would depend on the choice of the smoothing function

C_b and would not agree with the correlations calculated directly, even at separations large compared with R_b. As we saw in §5, the problem is that the peak field defined as

$$F_p = F_s - F_b \tag{6.29}$$

is correlated with F_b, $\langle F_p F_b \rangle \neq 0$. The background is not statistically neutral (except for altering the effective local threshold), as it should be to use equation (6.25). Only if the background is defined by a perfectly sharp low pass filter in k-space, $C_b(k) = \theta(k_b - k)$, is $\langle F_p F_b \rangle = 0$, but then the background two-point function $\xi_{\rho b}$ oscillates strongly with separation r until r is very much larger than $R_b \sim k_b^{-1}$.

Our way around these problems is to define F_p as a Gaussian random field statistically independent of and uncorrelated with F_b. The field F_p is characterized by its power spectrum, which is equal to the difference of the power spectrum $P_s(k)$ of F_s and the power spectrum $P_b(k)$ of F_b. Furthermore, we choose a Gaussian smoothing for F_b, with

$$C_b(r) = (2\pi)^{-3/2} R_b^{-3}\ exp(-\frac{1}{2} r^2/R_b^2), \tag{6.30}$$

to eliminate the ringing of the mass correlation function $\xi_{\rho b}$. If $P(k)$ is the power spectrum of the unsmoothed density perturbation field F,

$$P_b(k) = exp(-k^2 R_b^2)\ P(k), \tag{6.31}$$

$$P_p(k) = [exp(-k^2 R_s^2) - exp(-k^2 R_b^2)]\ P(k). \tag{6.32}$$

With this definition of F_p, the field $F_s' = F_p + F_b$ has the same power spectrum as F_s and therefore has all of the statistical properties of F_s. Nevertheless, if F_b is calculated from F_s by an explicit convolution integral, F_p cannot be, and members of the ensemble $\{F_s'\}$ cannot be identified with members of the ensemble $\{F_s\}$.

The global threshold is applied to F_s', but the local density of peaks $n_{pk}(\nu_b)$ is given by the number density formula of equations (4.3)-(4.13) applied to F_p. Let

$$\begin{aligned} &\sigma_{0s}^2 = \langle F_s^2 \rangle = \langle (F_s')^2 \rangle, \quad \sigma_{0p}^2 = \langle F_p^2 \rangle, \\ &\sigma_{1p}^2 = \langle (\nabla F_s)^2 \rangle = \sigma_{1s}^2 - \sigma_{1b}^2, \quad \sigma_{2p}^2 = \langle (\nabla^2 F_s)^2 \rangle = \sigma_{2s}^2 - \sigma_{2b}^2, \\ &\gamma_p = \sigma_{1p}^2/(\sigma_{0p}\sigma_{2p}), \quad R_{*p} = \sqrt{3}\sigma_{1p}/\sigma_{2p}. \end{aligned} \tag{6.33}$$

The differential number density $\mathcal{N}_{pk}(\nu_p; \gamma_p, R_{*p})$, with

$$\nu_p \equiv F_p/\sigma_{0p}, \tag{6.34}$$

for peaks in F_p has the same form as equation (4.3). As a function of ν_b the local density of peaks above threshold is

$$n_{pk}(\nu_b) = \int_0^\infty d\nu_p\ t[(\nu_p \sigma_{0p}/\sigma_{0s} + \nu_b \sigma_{0b}/\sigma_{0s})/\nu_t]\ \mathcal{N}_{pk}(\nu_p). \tag{6.35}$$

What makes the peak-background calculation worthwhile in competition with the direct calculation of the correlations discussed in §6.2 is that equation (6.35) can be approximated accurately in a way which allows the integral in equation (6.25) to be done analytically. The form is (c.f. equation 5.6)

$$n_{pk}(\nu_b) = n_0\ exp(\alpha\nu_b - \frac{1}{2}\beta\nu_b^2). \tag{6.36}$$

The parameters n_0 and α are found from the Taylor series expansion of equation (6.35) at $\nu_b = 0$. Note that at $\nu_b = 0$ there is an effective threshold parameter

$$\nu_t' = \nu_t \sigma_{0s}/\sigma_{0p}. \tag{6.37}$$

The integral for n_0 is

$$n_0 = \int_0^\infty d\nu \; t(\nu/\nu_t') \; \mathcal{N}_{pk}(\nu; \gamma_p, R_{*p}), \tag{6.38}$$

and can easily be evaluated numerically using the interpolation formula, equation (4.4), for the differential number density. If $t^{(1)}$ is the first derivative of of $t(\nu/\nu_t')$ with respect to its argument, then

$$\alpha = (n_0 \nu_t')^{-1} (\sigma_{0b}/(\sigma_{0p}) \int_0^\infty d\nu \; t^{(1)} \; \mathcal{N}_{pk}(\nu), \tag{6.39}$$

While β could also be evaluated from the Taylor series, it is in practice better to determine β by forcing equation (6.36) to agree with a numerical evaluation of $n_{pk}(\nu_b)$ from equation (6.35) at $\nu_b = 2$, say.

The first use of equation (6.36) is to calculate the global average of the local number density. The spatial average is equal to the integral over the Gaussian probability distribution,

$$\begin{aligned} \langle n_{pk} \rangle &= n_0 (2\pi)^{-1/2} \int_{-\infty}^\infty d\nu_b \; exp[\alpha \nu_b - \frac{1}{2}(1+\beta)\nu_b^2] \\ &= n_0 (1+\beta)^{-1/2} exp[\frac{1}{2}\alpha^2/(1+\beta)]. \end{aligned} \tag{6.40}$$

An important consistency check on the peak-background split is to compare this $\langle n_{pk} \rangle$ with the global number density from F_s. We find good agreement to within a few percent once $R_b > 3R_s$ even though n_0 may be several times less than $\langle n_{pk} \rangle$. Also, $R_b > 3R_s$ ensures that the local properties of the peaks, as measured by the moments σ_1 and σ_2, are nearly the same for F_p and F_s.

Analytic evaluation of the integral in equation (6.25), using techniques similar to those of §6.2 and Appendix 6, gives

$$\langle \prod_i^n n_{pk}(\vec{r}_i) \rangle = n_0^n \left(det \; (I+\Psi) \; det \; (\beta I + (I+\Psi)^{-1}) \right)^{-1/2} exp\left(\frac{1}{2}\alpha^2 \sum_{ij} \left([\beta I + (I+\Psi)^{-1}]^{-1} \right)_{ij} \right). \tag{6.41}$$

Use equations (6.40) and (6.41) together to find

$$1 + \xi_{pb}^{(n)}(\vec{r}_1, ..., \vec{r}_n) \equiv \langle n_{pk}(\vec{r}_1) ... n_{pk}(\vec{r}_n) \rangle / \langle n_{pk} \rangle^n. \tag{6.42}$$

If $\beta \ll 1$, which holds in the limit $R_b \gg R_s$, so $\sigma_{0b}/\sigma_{0p} \ll 1$,

$$\xi_{pb} \approx exp[\alpha^2 \sum_{i<j} \psi_{ij}] - 1. \tag{6.43}$$

Equation (6.43) has the same form as the expression for the n-point correlation function derived by Politzer and Wise (1984) in the high threshold limit. However, for realistic thresholds, by the time equation (6.43) is valid the argument of the exponential is small compared with one and the n-point correlation function is just a sum over two-point correlations.

Of course, equations (6.40)-(6.42) can only be expected to give an accurate approximation to the real ξ_{pk} if the minimum separation r_{ij} of the points is large enough compared with R_b that the background mass correlation $\xi_{\rho b}$ is close to $\xi_{\rho s}$. A numerical comparison of the two-point ξ_{pb} and ξ_{pk} is given in §6.4. For parameters appropriate to galaxy correlations a ratio r_{ij}/R_b greater than 4 seems

to be sufficient for an accuracy of a few per cent, which implies a ratio r_{ij}/R_s greater than about 12. Asymptotically at large r_{ij}, where $\psi_{ij} \ll 1$, comparison of equations (6.40) - (6.42) and (6.12) shows that the two estimates of ξ_{pk} are the same if

$$\langle \tilde{\nu} \rangle \approx \alpha(\sigma_{0s}/\sigma_{0b})\ (1+\beta)^{-1}. \tag{6.44}$$

6.4 Two-point Statistical Correlations

To illustrate the application of the formalisms discussed in §6.2 and 6.3 we present numerical results for two-point correlations here and for three-point correlations in §6.5. Our calculations are designed for comparison with the data on the two-point galaxy-galaxy correlations from the CfA redshift survey as analyzed by Davis and Peebles (1983). The theoretical background is the cold dark matter class of cosmological models together with the origin of density perturbations during an inflationary epoch in the very early universe (Starobinski 1982, Guth and Pi 1982, Hawking 1982, Bardeen etal. 1983). We expand upon the discussion of these models in §4.2.

The cold dark matter cosmological scenario is the one most likely to be compatible with an assumption that galaxies form at peaks of the primordial density perturbation field. If the density perturbations are generated from quantum fluctuations in the Higgs scalar field which gives rise to an inflationary epoch, the primordial fluctuations should have a scale invariant Zeldovich spectrum, $P(k) \propto k^{-3}$, when expressed in terms of a hypersurface-independent measure of the metric fluctuations. The evolution of these perturbations through the epoch when the universe becomes matter dominated results in a well-defined power spectrum of density perturbations just prior to galaxy formation, the 'adiabatic' power spectrum, so named because initially there is no perturbation in the ratio of cold dark matter density to entropy density. This CDM adiabatic spectrum was first obtained by Peebles (1982). More refined calculations have been carried out by Blumenthal and Primack (1984), Bardeen (unpublished) and Bond and Efstathiou (1984), and an accurate fit is given in Appendix 7.

Another conceivable, though less strongly motivated, type of density perturbation spectrum to come out of inflation with cold dark matter is associated with quantum fluctuations in an axion field (Steinhardt and Turner 1983, Linde 1984). When these fluctuations become perturbations in the cold dark matter energy density at the chiral phase transition ($T \sim 200\ MeV$) the total energy density necessarily remains unperturbed, since the wavelengths are much larger than the horizon at that time. The evolved density perturbation spectrum just before galaxy formation is called the isocurvature spectrum since there is no initial perturbation in the spatial curvature - unlike the adiabatic case. Bardeen (unpublished) and Efstathiou and Bond (1985) have computed the form of this spectrum. Accurate fits are given in Appendix 7.

The isocurvature and adiabatic spectra become identical at long wavelengths. On shorter wavelengths, the isocurvature is generally flatter than the adiabatic, with more power on cluster scales, and with effective index $n \approx -2.5$ on galactic scales compared with $n \approx -2$ for the adiabatic. We shall see that these shape differences have a relatively important impact on the correlations we compute for the two models.

The precise scaling of these density perturbation spectra in relation to present observations depends on the value of the Hubble constant, since this determines the scale which comes inside the horizon just as the universe becomes matter dominated. If we assume that the present value of the cosmological constant model is zero and the cosmological density parameter $\Omega = 1$, constraints on the age of the universe from globular cluster evolution strongly suggest that the Hubble parameter h is no larger than 0.4 or perhaps at most 0.5.

The remaining parameters for our model of galaxy formation are the smoothing radius R_s for the peaks and the parameter q in the threshold function $t(\nu/\nu_t)$ which governs the sharpness of the threshold. Our best guess is that R_s should correspond to a mass somewhat smaller than the total mass (including the dark halo) of a typical luminous galaxy, since some of this mass would no doubt

have accreted onto a central core after the epoch at which the threshold was set. The mass within the Gaussian smoothing window is given by equation (5.3a).

The choice of q is also uncertain. However, in the context of a threshold hypothesis, a reasonable requirement on q is that the distribution in ν of peaks selected as 'galaxies' be not too asymmetric about its maximum and the maximum be reasonanbly close to ν_t. This requires a large q when ν_t is large and a smaller q when ν_t is not so large.

With the above as background, we first examine the asymptotic amplitude of the correlations in the limit $\xi_{pk} \ll 1$ when equation (6.12) is a good approximation to ξ_{pk}. The quantity $\langle\tilde{\nu}\rangle$ can be considered an effective threshold level for the correlations, since $\xi_{pk} = \langle\tilde{\nu}\rangle^2 \xi_\rho/\sigma_0^2$, and $\langle\tilde{\nu}\rangle$ approaches ν_t for a perfectly sharp very high threshold. The integral

$$\langle\tilde{\nu}\rangle = \int_0^\infty d\nu \; (\nu - \gamma\theta/(1-\gamma^2)) \; t(\nu/\nu_t) \; \mathcal{N}_{pk}(\nu) \tag{6.45}$$

is easily evaluated, and some illustrative results are shown in Table 6.1.

The masses associated with the values of R_s in Table 6.1 range from $M_s = 1.3 \times 10^{10} \; h^{-1} \; M_\odot$ for $R_s = 0.143$ to $M_s = 3.9 \times 10^{11} \; h^{-1} \; M_\odot$ for $R_s = 0.445$. The smaller pair of values of R_s and the larger pair of values of R_s each represent a single smoothing radius when the perturbation spectrum is rescaled to make it independent of h. For the two smaller values of R_s, $\gamma = 0.555$ for the adiabatic spectrum and $\gamma = 0.437$ for the isocurvature spectrum. For the two larger values of R_s, $\gamma = 0.599$ for the adiabatic spectrum and $\gamma = 0.500$ for the isocurvature spectrum.

The value of ν_t for each h, R_s, and q is chosen so that the global average number density of of peaks above threshold is $\langle n_{pk}\rangle = 0.01 \; h^3 Mpc^{-3}$, the number density of the galaxies counted in the CfA redshift survey (Davis and Huchra 1982). Note that as the threshold becomes less sharp (q decreases), there are more peaks with $\nu < \nu_t$ counted as 'galaxies' and ν_t must increase to keep the total number density constant. On the other hand, $\langle\tilde{\nu}\rangle$ decreases because the average height of the peaks decreases. The distribution of peaks in ν is reasonably symmetric about ν_t for $q = 16$ when $\nu_t \approx 3.5$ and for $q = 8$ when $\nu_t \approx 2.5 - 3.0$ (see Figure 4.4). By the time $\nu_t < 2$ the differential number density is no longer rapidly varying and $q = 4$ is probably as sharp a threshold as can be expected.

The last column in Table 6.1 gives values of the asymptotic expression for the zeroth-order two-point correlation function evaluated at $r = 5 \; h^{-1} Mpc$, where the observed correlation amplitude is one. The actual peak correlation amplitude is significantly larger unless $\langle\tilde{\nu}\rangle^2 \xi_\rho/\sigma_0^2 \ll 1$. These values give some idea of the trend of the overall amplitude of the peak correlations with h, R_s, and q. As R_s decreases, $\langle\tilde{\nu}\rangle$ increases because ν_t must increase to keep the global number density constant. σ_0 is also increasing at just about the same rate. In fact, the trend with R_s is opposite for $h = 0.4$ and $h = 0.5$. The correlation amplitude can be rather sensitive to h since at fixed observed separation (in units of $h^{-1} Mpc$), ξ_ρ/σ_0^2 decreases rapidly with increasing h. This is largely balanced by an increase in $\langle\tilde{\nu}\rangle$ at the larger R_s, but not at the smaller R_s.

In contrast, σ_0 rises more slowly as R_s decreases for the isocurvature spectrum. Also, once $\nu_t < 2.5$, $\langle\tilde{\nu}\rangle$ changes more rapidly with R_s. Both effects combine to make the correlation amplitude rather sensitive to the value of R_s for the isocurvature spectrum. Of course, for a given choice of $\langle n_{pk}\rangle$ there is a maximum possible value for R_s at which $\nu_t \to -\infty$.

For each type of spectrum, we pick two cases shown in Table 6.1 to show how the actual ξ_{pk} varies with $r = |\vec{r}_2 - \vec{r}_1|$. First consider the adiabatic spectrum. One case has a small value of R_s ($0.178 \; h^{-1} Mpc$ with $h = 0.4$) and a reasonable threshold sharpness ($q = 16$) for the fairly high value of ν_t. The second case has a larger $R_s = 0.356 \; h^{-1} Mpc$ with $h = 0.5$ and a correspondingly softer threshold ($q = 8$) for the smaller value of ν_t.

The direct calculation of the correlations from equation (6.22) is straightforward. The matrix C has components $C_{11} = C_{22} = -\psi_{12}^2/(1-\gamma^2)$, $C_{12} = C_{21} = \psi_{12}$. The double integral over $\nu(1)$

and $\nu(2)$ is evaluated numerically. The results for the adiabatic spectrum are plotted as the solid curves in Figure 6.1a. The approximations made in arriving at equation (6.22) are estimated to have errors of a few per cent at $r = 4.5R_s$, which rapidly diminish to less than 1% as r increases. To show how the statistical enhancement modifies the slope of the peak correlation function from that of the mass correlation function we plot $\langle\tilde{\nu}\rangle^2\xi_\rho/\sigma_0^2$ for $R_s = 0.178\ h^{-1}Mpc$. Note how this asymptotic approximation approaches ξ_{pk} once $\xi_{pk} < 1$.

In Figure 6.1a, it is remarkable how closely ξ_{pk} follows an $r^{-1.8}$ power law, like that of the observed galaxy correlation function, even though the mass correlation function ξ_ρ is not particularly close to a single power law. That this is so over the whole range of R_s likely to be relevant for galaxies is an accident of the particular shape of the adiabatic density perturbation spectrum and of the observed number density of luminous galaxies. If galaxies were rarer objects so that, for a given R_s, ν_t were larger, then the nonlinear statistical enhancement of the correlations when $\xi_{pk} > 1$ would be greater, and ξ_{pk} would rise more steeply at small r - as was predicted by Politzer and Wise (1984) from an approximation to ξ_{pk} valid only in the very high ν_t limit. If luminous galaxies were more common, ν_t and $\langle\tilde{\nu}\rangle$ would be smaller and ξ_{pk} would follow ξ_ρ in shape more closely at small r as well as at large r.

The amplitude of ξ_{pk} is 3 to 5 times smaller than the observed correlations, indicated by the $(5/r)^{1.8}$ line. However, remember that the ξ_{pk} plotted here is the correlation function of the peaks in the limit that $\xi_\rho \ll \xi_{pk}$, before any significant dynamical evolution of the correlation function. At the present time, ξ_ρ cannot be very much less than one on scales of several Mpc, since a significant fraction of the universe has collapsed on these scales to form bound clusters of galaxies. Also, σ_0 must be large enough so that galaxies formed at a reasonably early time.

To follow the dynamical evolution of ξ_{pk} when it and ξ_ρ exceed one is a job for numerical simulations. Davis *etal.* (1985) give some results of a pioneering attempt to follow the clustering of galaxies identified with high ν peaks in a cold dark matter dominated universe. The two-point correlation function ξ_g of their 'galaxies' increase in amplitude *without steepening* on comoving scales of $2\ Mpc < r < 10\ Mpc$ (see Figure 17 of their paper), suggesting that dynamical evolution of a zeroth-order ξ_{pk} like that shown in Figure 6.1 may be consistent with observation. The numerical simulation does show steepening on smaller scales once dynamical clustering becomes highly nonlinear; this is not seen in the observed correlations. However, the accuracy of these numerical simulations on such small scales is open to question. Much more work is needed. We will explore the dynamical evolution of the perturbations when both ξ_ρ and ξ_{pk} are less than one in §6.6.

One important further point in Figure 6.1a is the comparison of the $\xi_{pk}(r)$ calculated directly with $\xi_{pb}(r)$, which is the zeroth-order peak correlation function calculated using the peak-background split of §6.3. Equations (6.41) and (6.40) combine to give

$$1 + \xi_{pb} = [1 - (\beta\psi_{12}/(1+\beta))^2]^{-1/2}\ exp[(\alpha/(1+\beta))^2\psi_{12}/(1+\beta\psi_{12}/(1+\beta))], \qquad (6.46)$$

where ψ_{12} is the normalized two-point correlation function for the background field F_b as defined by equation (6.28).

For the case plotted, the background smoothing radius is $R_b = 0.593\ h^{-1}Mpc$, 3.33 times R_s. The values

$$\alpha = 2.169, \quad \beta = 0.426 \qquad (6.47)$$

(equation 6.36) and $n_0 = 0.00224\ h^3Mpc^{-3}$ follow. The effective threshold (equation 6.37) is at $\nu_t' = 4.249$, and the parameter γ_p for the peak field is 0.6534. The global number density from equation (6.40) is $\langle n_{pk}\rangle = 0.00976\ h^3Mpc^{-3}$ compared with 0.01009 calculated directly. The peak-background estimate of $\langle\tilde{\nu}\rangle$ is 2.64 compared with 2.62 in Table 6.1. The agreement of ξ_{pb} with ξ_{pk} is excellent once $r > 2\ Mpc \approx 3.4R_b$. This gives us confidence in the peak-background method, which we will use exclusively in §6.5 and §6.6.

Now consider the results for the isocurvature spectrum, as shown in Figure 6.1b. The two examples from Table 6.1 are $h = 0.4$, $R_s = 0.178\ h^{-1}Mpc$, $q = 8$ and $h = 0.4$, $R_s = 0.445\ h^{-1}Mpc$, $q = 4$. Taking $q = 8$ rather than 16 in the first example means that the average value of ν is somewhat smaller than ν_t. It appears that ξ_{pk} agrees very well with the observed galaxy-galaxy correlations at $r > (4-5)\ h^{-1}Mpc$, but remember that ξ_{pk} is the *unevolved* peak correlation function. If σ_0 is large enough for galaxies to form at a reasonable redshift, the peak-peak correlations at the present time will have grown to be considerably larger than the observed galaxy correlations (see §6.6).

There are still problems if we try to minimize ξ_{pk} at 5 $h^{-1}Mpc$, as in the second example. While the level of ξ_{pk} now does allow for some dynamical evolution, the level of σ_0 is still constrained to be too small to allow galaxies to form at a reasonable time. Also, $\langle\tilde{\nu}\rangle$ is so small that the statistical segregation of galaxies from mass in clusters and superclusters (§5) is too weak to be of much help with the missing mass problem. We conclude that a pure isocurvature density perturbation spectrum is incompatible with at least our simple version of a threshold hypothesis for galaxy formation.

From these preliminary explorations it seems promising that the threshold hypothesis together with the adiabatic density perturbation spectrum can give a good account of the observed correlations of galaxies (see also §6.5). The same formalism can be applied to applied to correlations of other rare objects such as rich clusters of galaxies (Kaiser 1984a). However, the adiabatic spectrum seems in some difficulty here, since, according to Bahcall and Soneira (1983), ξ_{cl} for the Abell cluster sample remains positive out to 100 $h^{-1}Mpc$, while ξ_ρ goes negative at $r \approx 20\ (\Omega h^2)^{-1}\ Mpc$ for the adiabatic spectrum. While our approximation schemes break down as ξ_ρ changes sign, it seems highly unlikely that the peak correlations will behave very differently from the density correlations.

6.5 THREE-POINT CORRELATIONS

The three-point correlation function for galaxies has been infered from angular correlation measurements (Groth and Peebles 1977) and has been directly estimated from redshift surveys (Efstathiou and Jedrezejewski 1985). The usual way of decomposing the three-point function is to write it as a sum of the two-point correlations ξ_{12}, ξ_{23}, ξ_{13} at each of the three separations r_{12}, r_{23}, r_{13}, plus a connected part denoted by $\varsigma_{123} = \varsigma(\vec{r}_1, \vec{r}_2, \vec{r}_3)$:

$$\xi^{(3)}(\vec{r}_1, \vec{r}_2, \vec{r}_3) = \xi_{12} + \xi_{23} + \xi_{13} + \varsigma_{123}. \tag{6.48}$$

The observational results are consistent with ς_{123} being a sum over products of two-point functions with a coefficient Q which is at least approximately scale independent,

$$\varsigma_{123} = Q(\xi_{12}\xi_{23} + \xi_{23}\xi_{31} + \xi_{31}\xi_{12}). \tag{6.49}$$

The value of Q is estimated to be close to one, with an uncertainty of about 0.2 either way.

The three-point function is also an important diagnostic for numerical simulations of clustering of galaxies. Davis etal (1985) find in their cold dark matter simulations that Q is usually in the range $1.5 - 2$, at least for smaller separations, when they force the three-point correlations into the form of equations (6.48) and (6.49). When they introduce biasing like that assumed in this paper they obtain a value of Q somewhat closer to one.

Politzer and Wise (1984) found that the connected part of the three-point function for *points* above a high threshold has the form

$$\varsigma_{123} = \xi_{12}\xi_{23} + \xi_{23}\xi_{31} + \xi_{31}\xi_{12} + \xi_{12}\xi_{23}\xi_{31}. \tag{6.50}$$

The last term dominates when $\xi_{ij} \gg 1$.

A direct calculation of the three-point function using equation (6.22) requires evaluating a triple integral over the $\nu(i)$. This is certainly feasible, but the analytic calculation using the peak-background split is much simpler and is all we present in this paper.

Reduction of the matrices in equation (6.41) is beginning to get a bit complicated for three-point correlations, but it is still straightforward. Let

$$\begin{aligned} D &= \left(det\ (I+\Psi)\ det\ (\beta I + (I+\Psi)^{-1})\ \right)^{-1/2}\ (1+\beta)^{-3} \\ &= 1 - (\psi_{12}^2+\psi_{23}^2+\psi_{31}^2)\ (\beta/(1+\beta))^2 + 2\psi_{12}\psi_{23}\psi_{31}\ (\beta/(1+\beta))^3. \end{aligned} \tag{6.51}$$

Then

$$\begin{aligned} 1+\xi_{pb}(\vec{r}_1,\vec{r}_2,\vec{r}_3) \equiv 1+\xi_{pb}^{(3)} &= D^{-1/2}\ e^X, \\ X = \frac{\alpha^2}{D(1+\beta)^2}\Big(\psi_{12}+\psi_{23}+\psi_{31} &- \frac{\beta}{1+\beta}(\psi_{12}^2+\psi_{23}^2+\psi_{31}^2 \\ &+ \psi_{12}\psi_{23}+\psi_{23}\psi_{31}+\psi_{31}\psi_{12}) + \frac{\beta^2}{(1+\beta)^2}3\psi_{12}\psi_{23}\psi_{31}\Big). \end{aligned} \tag{6.52}$$

Some simplification is possible if the three points $\vec{r}_1, \vec{r}_2, \vec{r}_3$ are at the corners of an equilateral triangle, so $r_{12} = r_{23} = r_{31} = r$ and $\psi_{12} = \psi_{23} = \psi_{31} = \psi$. We find

$$1+\xi_{pb}^{(3)} = [1-\beta\psi/(1+\beta)]^{-1}\ [1+2\beta\psi/(1+\beta)]^{-1/2}\ exp\Big(\frac{3(\alpha/(1+\beta))^2\psi}{1+2\beta\psi/(1+\beta)}\Big), \tag{6.53}$$

which is rather similar to equation (6.47) for $\xi_{pb}^{(2)}$.

Consider the limit when both $\beta\psi_{ij}$ and $(\alpha/(1+\beta))^2\ \psi_{ij}$ are small compared with one. Expanding the exponential and D in equation (6.52) to second order in ψ and then reassembling terms to replace the expansion in ψ by an expansion in $\xi_{pb}^{(2)}$ (see equation 6.47) gives for the connected part of the three-point function

$$\varsigma_{123} \approx (1-\beta(1+\beta)/\alpha^2)\ (\xi_{12}\xi_{23}+\xi_{23}\xi_{31}+\xi_{31}\xi_{12}). \tag{6.54}$$

As expected, this has the form of equation (6.49). The asymptotic value of Q at large separations is then

$$Q \approx 1-\beta(1+\beta)/\alpha^2. \tag{6.55}$$

While both α and β depend on the ad hoc choice of the background smoothing radius R_b, the scaling of α and β with R_b is such that Q is independent of R_b, at least when R_b/R_s is large, as it should be.

If $\beta\psi_{ij} \ll 1$, but $(\alpha/(1+\beta))^2\ \psi_{ij}$ is *not* small, which requires $(\alpha/(1+\beta))^2 \gg 1$,

$$\begin{aligned} \xi^{(3)} &\approx exp\Big((\alpha/(1+\beta))^2\ (\psi_{12}+\psi_{23}+\psi_{31})\Big) \\ &\approx (1+\xi_{12})(1+\xi_{23})(1+\xi_{31}) - 1. \end{aligned} \tag{6.56}$$

This is equivalent to the Politzer-Wise form, equation (6.50), and is incompatible with the form of equation (6.49) once the correlations are large.

What is the situation for peaks selected to model the distribution of galaxies? In order that the peak-background split apply when the correlation amplitudes are at least beginning to get large, R_s should be relatively small. Consider the case shown in Figure 6.1a (adiabatic spectrum), with α and β given by equation (6.47). Then $\alpha^2/(\beta(1+\beta)) = 7.74$ and the asymptotic $Q = 0.871$. At smaller r_{ij} define an effective Q, Q_e, which is calculated by taking $\xi^{(3)}$ from equation (6.52), subtracting the sum of the two-point correlations $\xi_{12}, \xi_{23}, \xi_{31}$ from equation (6.46), and then dividing by the sum of the pair-wise products of the ξ_{ij}. When $r_{12} = r_{23} = r_{31} = 2\ h^{-1}Mpc$, the smallest separation at which the peak-background split seems to be reliable, we find $\xi^{(3)} = 12.33$, $\xi^{(2)} = 1.638$, $\varsigma_{123} = 7.42$, and $Q_e = 0.922$. The Politzer-Wise form would give $\xi^{(3)} = 17.35$, almost 50% too large. Repeating

this calculation for a variety of different r_{ij}, including all kinds of triangle shapes, always gives values of Q_e in the range $0.87 - 0.92$. The moral is that as the correlation amplitude becomes large, what is a fairly small percentage error in the exponent in equation (6.52) can be a substantial error in the correlation.

The fact that Q_e for the peaks is approximately scale independent and has a value agreeing remarkably well with the observed Q for galaxies is encouraging, but should not be taken too seriously. The problem again is that what we have calculated for the peaks is zeroth-order in dynamics, while the galaxy three-point function is only measured on scales ($< 2\ h^{-1}Mpc$) on which there has been significant dynamical evolution. Since the evidence from the numerical simulations is that dynamical evolution drives Q toward one, at least one can say that our threshold hypothesis seems to be consistent with observations of three-point correlations of galaxies.

6.6 Dynamic Evolution of the Correlations

The precise calculation of the dynamic evolution of the correlations is rather complicated since the correlations are defined in terms of physical distances in our present universe whereas the probability distribution for the random field F_s and quantities associated with it are known in terms of the linear perturbation field as a function of Lagrangian (comoving) coordinates. At the outset one can restrict oneself to considering dynamics as a linear perturbation away from the homogeneous isotropic background, leaving the nonlinear problem to numerical simulations. However, the statistical effects are highly nonlinear in general, and the statistical and dynamical contributions do not couple in a simple way.

We leave the general problem to future work, and consider here only the much simpler limit when the separations of the points are very large compared with R_s and the statistical correlation amplitudes as well as the dynamical perturbations are small. In this limit, the n-point correlation function is just a sum over two-point functions, so we need only consider the two-point correlations. There is also plenty of room to apply the peak-background method, which is conceptually and mathematically relatively simple.

In the peak-background split of §6.3, there is a comoving density of peaks, n_{pk}, which depends only on the background overdensity $\nu_b(\vec{r},t)$ evaluated at $t = 0$. Here $\vec{r}$ is the comoving position and t is the time. While to lowest order in the dynamics ν_b is independent of time, the fractional perturbation in the physical density at time $t \neq 0$ should be written as

$$F_b(\vec{r},t) = \nu_b(\vec{r},t)\sigma_b(t), \tag{6.57}$$

since the correlations are second order in the density perturbation amplitude.

The number of peaks per unit physical volume is then

$$N_{pk} = n_{pk}(\nu_b(\vec{r},0))\ (1 + \nu_b(\vec{r},t)\sigma_b(t)). \tag{6.58}$$

The two-point correlation function can be written as

$$\xi_{pb}(x_{12},t) = \langle N_{pk}(\vec{x}_1,t)N_{pk}(\vec{x}_2,t)\rangle/\langle N_{pk}^2\rangle^2, \tag{6.59}$$

with $x_{12} = |\vec{x}_1 - \vec{x}_2|$ of fixed physical (Eulerian) separation in the present universe. The averages are uniform in physical volume. To evaluate them, we note that when $R_b \gg R_s$, equation (6.36) for n_{pk} reduces to

$$n_{pk} \approx n_0(1 + \alpha\nu_b(\vec{r},0)) \tag{6.60}$$

since α scales as σ_b/σ_s and β scales as $(\sigma_b/\sigma_s)^2$. (To be completely careful about second order terms, we should have kept terms of order β and α^2 here, but these obviously cancel between the numerator and denominator of equation (6.59).) Since

$$d^3x = d^3r/(1 + \nu_b(\vec{r},t)\sigma_b), \tag{6.61}$$

it is easy to see that $\langle N_{pk}\rangle_x = \langle n_{pk}\rangle_r = \langle n_{pk}\rangle$ as required for consistency. (The subscripts x and r denote volume averages in Eulerian or Lagrangian space.) Also, by definition,

$$\int d^3x\ \nu_b(\vec{r},t) = 0, \quad \int d^3r\ \nu_b(\vec{r},0) = 0. \tag{6.62}$$

Write the integral in the numerator of equation (6.59), keeping only terms up to second order in α or σ_b, as

$$\begin{aligned}
&\int d^3x_1\ n_0^2\ [1+\nu_b(\vec{r}_1,t)\sigma_b][1+\nu_b(\vec{r}_2,t)\sigma_b] \\
&+\int d^3r_1\ n_0^2\ \alpha\nu_b(\vec{r}_1,0)[1+\nu_b(\vec{r}_2,t)\sigma_b] \\
&+\int d^3r_2\ n_0^2\ \alpha\nu_b(\vec{r}_2,0)[1+\nu_b(\vec{r}_1,t)\sigma_b] \\
&+\int d^3r_1\ n_0^2\ \alpha^2\nu_b(\vec{r}_1,0)\nu_b(\vec{r}_2,0).
\end{aligned}$$

All explicit first order terms vanish, and in explicit second order terms we can identify $\nu_b(\vec{r},t)$ and $\nu_b(\vec{r},0)$. As in §6.3, let $\psi_{12} = \langle\nu_b(\vec{r}_1)\nu_b(\vec{r}_2)\rangle$. Then,

$$\begin{aligned}
\xi_{pb}(t) &= (\alpha+\sigma_b)^2\psi_{12} \\
&= \left((\xi_{pb}(0))^{1/2} + (\xi_\rho(t))^{1/2}\right)^2.
\end{aligned} \tag{6.63}$$

The second version of equation (6.63) may well be a reasonable approximation even when the statistical correlations are not really small enough for the above derivation to be strictly valid. We have compared the time evolution predicted by equation (6.63) wih the 'biased galaxy' numerical simulation of Davis etal. (1985) and find rather good agreement.

For a given zeroth-order correlation function, equation (6.63) can be used to estimate the present amplitude of density perturbations necessary for the evolved peak correlations to match those of galaxies, with $\xi_{pb} \approx 1$ at $r = r_0 = 5\ h^{-1}Mpc$. For instance, the adiabatic spectrum predicts a zeroth-order ξ_{pb} of about $0.2 - 0.3$ at r_0, so ξ_ρ at r_0 should also be about 0.25. In particular, the model in Table 6.1 and Figure 6.1a with $h = 0.5$, $R_s = 0.356\ h^{-1}Mpc$ and $q = 8$ has $\xi_{pb}(r_0,0) \approx 0.18$ which requires $\xi_\rho(r_0,t_0) \approx 0.33$. The corresponding $\sigma_{0s} \approx 2.4$. This value for σ_{0s} in turn allows an estimate of the cosmological time of collapse of a typical peak. For a spherical peak at the maximum of the distribution in ν, the collapse redshift is $z_g \equiv \sigma_{0s}\langle\nu\rangle/1.69 - 1 \approx 2.9$. The collapse redshift would be somewhat larger if, as is likely, the peaks are somewhat aspherical. This value of σ_{0s} was also used in Table 5.2 to estimate ν_b and $z_t \equiv \sigma_{0s}\nu_t/1.69 - 1$ for this model. A similar procedure was followed for the other models listed in Table 5.2.

To see the trouble with the isocurvature spectrum explicitly, consider the model in Table 6.1b with the smallest statistical correlation at $r = r_0$, the model with $h = 0.4$, $R_s = 0.445\ h^{-1}Mpc$ and $q = 4$. This has (see Figure 6.1b) $\xi_{pb}(r_0,0) \approx 0.31$, and $\xi_{pb}(r_0,t_0) = 1$ gives $\xi_\rho(r_0,t_0) \approx 0.20$. The corresponding $\sigma_{0s} \approx 0.9$, which together with the threshold at $\nu_t = 1.87$ implies a threshold collapse redshift $z_t \approx 0.0$. The typical peak counted as a 'galaxy' has $\langle\nu\rangle \approx 2.1$, so $z_g \approx 0.1$. Even in this extreme model, which would have relatively little biasing in the galaxy/mass ratio, 'galaxies' would form unacceptably late with a significant fraction not collapsed at present. Without any biasing at all, the collapse of a typical peak with $\nu \sim 1$ is at a redshift $z \approx 0.2$ (with the normalization $\xi_\rho(r_0,t_0) = 1$).

For scales $r > r_0$, equation (6.63) implies that the peak correlation function is amplified over the mass density correlation function, $\xi_{pk}(r,t) \approx b^2\xi_\rho(r,t)$, by the square of the spatially-uniform biasing factor $b(t) = \langle\tilde{\nu}\rangle/\sigma_{0s}(t) + 1$. A common way to normalize the overall amplitude of the linear perturbation spectrum is to relate the quantity $J_3(r) = \int_0^r \xi(r)r^2dr$ on scales $r \gg r_0$ to that

determined from galaxy redshift surveys, J_{3g}. The assumption that galaxies are unbiased tracers of the mass distribution corresponds to $J_{3\rho} = J_{3g}$ on sufficiently large scales, where $J_{3\rho}$ is determined using the mass autocorrelation function. Since the dominant contribution to $J_{3g}(r)$ comes from scales $> r_0$, with biasing $J_{3\rho}(r) = b^{-2}J_{3g}(r)$ is a good approximation. The power spectrum normalization with biasing is simply lowered by the factor b^{-2} over that determined assuming 'light traces mass'. Predictions of the microwave background anisotropies $\Delta T/T$ are lowered by b^{-1} over the values obtained using the 'light traces mass' assumption (e.g., Bond and Efstathiou 1984). For the adiabatic CDM model with $h = 0.5$, $R_s = 0.356\ h^{-1}Mpc$ and $q = 8$, $b = 1.7$. For the isocurvature CDM model with $h = 0.4$, $R_s = 0.445\ h^{-1}Mpc$ and $q = 4$, $b = 2.2$. The advantage of such a normalization procedure in the present context is that we can be relatively confident that equation (6.63) holds on scales $r > r_0$ even if it fails on much smaller scales. However, the differences between J_3 normalization and the normalization at r_0 that we have adopted here for simplicity are not large since the correlation functions we predict, especially in the adiabatic model, are quite similar to the observed galaxy correlation function.

7. DENSITY PROFILES AROUND PEAKS

If prominent cosmic structures do indeed arise from condensations of gas and dark matter around primordial density peaks, then the initial conditions for nonlinear collapse will be the density and velocity profiles in the neighborhood of the peak as determined in the linear regime. Here, we focus on the density only. Similar methods would be used to determine the velocity. In §7.1, we obtain the distribution of the asymmetry parameters of the peak, which determines the lowest order Taylor expansion of the density profile. In §7.2, we calculate the average shape and its dispersion as we go farther away from the maximum. In §7.3, we apply our results to clusters in neutrino-dominated universes.

7.1 The Triaxial Ellipsoid Approximation

In the immediate neighborhood of a peak, the density profile is given by the Taylor expansion

$$F(r) = F(0) - \sum \lambda_i r_i^2/2. \tag{7.1}$$

The axes are oriented along the principal axes. Since all λ_i are positive, the contour surface of constant density f, defined by $F(r) = f$, defines a triaxial ellipsoidal surface with semi-axes

$$a_i = \left(\frac{2(F(0) - f)}{\lambda_i}\right)^{1/2}, \tag{7.2}$$

at least provided $F(0) - f$ is small. Since λ_1 is by definition the largest eigenvalue, collapse will first occur along the 1-axis resulting in pancake formation as described by Zeldovich (1970). Depending upon the distribution of the other λ_i, collapse may or may not follow quickly along the other two axes.

We characterize the asymmetry by the parameters

$$\begin{aligned} e &= \frac{\lambda_1 - \lambda_3}{2\sum \lambda_i}, \\ p &= \frac{\lambda_1 - 2\lambda_2 + \lambda_3}{2\sum \lambda_i}. \end{aligned} \tag{7.3}$$

Thus, e (≥ 0) is a measure of the ellipticity of the distribution in the 1-3 plane, and p determines the degree of oblateness ($0 \leq p \leq e$) or prolateness ($0 \geq p \geq -e$) of the triaxial ellipsoid. Oblate spheroids have $p = e$ and prolate spheroids have $p = -e$. As we shall see, spheroidal distributions are highly improbable.

In terms of the asymmetry parameters and the variable $x = -\nabla^2 F/\sigma_2$ introduced in §6, the small r expansion of the profile is

$$\begin{aligned} F(r) &\approx \nu\sigma_0 - x\sigma_2\frac{r^2}{2}(1 + A(e,p)) \; as \; r \to 0, \\ A(e,p) &= 3e(1 - \sin^2\theta(1 + \sin^2\phi)) + p(1 - 3\sin^2\theta\cos^2\phi). \end{aligned} \tag{7.4}$$

We have adopted spherical coordinates with the 1-axis chosen as the azimuthal one and $x_3 = r\sin\theta\sin\phi$. Note that the angle-average of the asymmetry measure $A(e,p)$ vanishes. The distribution of shapes in the immediate neighborhood of a peak of given height ν can be expressed in terms of the distributions for x, e, and p.

The conditional probability for x given that the peak has a height ν is independent of the asymmetry parameters e and p:

$$P_x(x|\nu)dx = \mathcal{N}_{pk}(\nu,x)/\mathcal{N}_{pk}(\nu) = \frac{g(x,\gamma,x_*)}{G(\gamma,x_*)} = \frac{exp(-\frac{(x-x_*)^2}{2(1-\gamma^2)})}{(2\pi(1-\gamma^2))^{1/2}}\frac{f(x)\;dx}{G(\gamma,\gamma\nu)}. \tag{7.5}$$

Here g and G are defined by equations (6.9) and (4.3), and $f(x)$ is given by equation (A1.15); it is monotonic, $\sim x^8$ for small x and $\sim x^3$ for large x. This probability distribution function (7.5) is plotted in Fig. 7.1 for a specific value of γ appropriate to galactic scale peaks in the adiabatic CDM picture, and for a variety of ν. This function is sharply peaked at its maximum, which occurs at x_m, fit by equation (6.17).

7.1.1 The Distribution of Ellipticity and Prolateness

Only rarely occurring peaks are spherical. To determine the degree of asphericity expected, we need the conditional probability for the asymmetry parameters e and p subject to the constraint that the peak has a given height ν and x. This turns out to be independent of ν:

$$P_{ep}(e,p|\nu,x)dedp = P_{ep}(e,p|x)dedp == \frac{3^2 5^{5/2}}{(2\pi)^{1/2}}\frac{x^8}{f(x)}e^{-\frac{5}{2}x^2(3e^2+p^2)}W(e,p)dedp \tag{7.6}$$

This result and details of the subsequent discussion are given in Appendix 3. The function $W(e,p)$ is a polynomial of order five in both e and p which is constrained to vanish outside of the $|p| \le e, e \ge 0$ domain. Indeed, the allowed domain is the interior of a triangle bounded by the points in the (e,p)-plane $(0,0)$, $(\frac{1}{4},-\frac{1}{4})$ and $(\frac{1}{2},\frac{1}{2})$. See Appendix 3 for details of the form of W. We plot various contour lines of the probability distribution (7.6) in Fig.7.2 for a variety of $\gamma\nu$. We have taken $x = x_*$ as the constraint, which according to equations (6.13) and (6.19), is the mean value and most probable value for large $\gamma\nu$.

Notice that the most likely value of p quickly goes to zero. High ν peaks are neither oblate nor prolate, but they are definitely triaxially asymmetric, since $\lambda_2 \approx (\lambda_1 + \lambda_3)/2$. Indeed in the large x limit, e and p are small and we can approximate P by a Gaussian

$$P_{ep}(e,p) \approx P_{ep}(e_m,p_m)exp(-\frac{(e-e_m)^2}{2\sigma_e^2} - \frac{(p-p_m)^2}{2\sigma_p^2}),$$

where the most probable values and their dispersions are (for large x)

$$\begin{aligned} e_m &= \frac{1}{\sqrt{5}x(1+6/(5x^2))^{1/2}} \\ \sigma_e &= e_m/\sqrt{6} \\ p_m &= \frac{6}{5x^4(1+6/(5x^2))^2} \\ \sigma_p &= e_m/\sqrt{3}. \end{aligned} \tag{7.7}$$

The cross term $\propto (e-e_m)(p-p_m)$ has a coefficient which rapidly goes to zero for large x and can be ignored. Thus, though p_m goes to zero much more rapidly than e_m, the constant P_{ep} contours should be approximately elliptical with axial ratios 1.4 to 1, and with principal axes oriented along the e and p axes. This is evident from Fig. 7.2. *High peaks tend to be more spherically symmetric than low ones.* This is evident from 7.7: sphericity is approached as $(\sqrt{5}\gamma\nu)^{-1}$; the dispersion about sphericity approaches zero at about the same rate.

However, for CDM, we typically have $x_* \approx (0.4-0.7)\nu$, so significant deviation from sphericity is expected even for $\nu \sim 3$. For example, for peaks on the scale $R_s = 0.5\ h^{-1}Mpc$ of height $\nu = 2.7$ in the adiabatic cold dark matter model with $\Omega = 1$ and $h = 0.5$, we have $\gamma = 0.62$, hence the most probable value of $x \approx 1.5x_* \approx 2.5$; therefore the most probable $e_m \approx 0.17$ and $p_m \approx 5 \times 10^{-3}$ by equation (7.7). These values are in good accord with Fig.7.2. The eigenvalues would then be related by $\lambda_1 = 1.3\lambda_2 = 1.7\lambda_3$. The axial ratios (equation 7.2) immediately follow: the long (3) semi-axis is only 1.3 times the short one. However, since the short axis goes nonlinear first (collapsing when the expansion factor is 1.3 smaller than when the 2-axis collapses in the Zeldovich (1970) approximation), this asymmetry amplifies in the nonlinear regime (Lin, Mestel and Shu 1965, Zeldovich 1970). The generic collapsed structure will be pancake-like.

Note that artificial spherical smoothing tends to sphericalize pancakes and filaments, so the asymmetry parameters obtained from this distribution will generally be underestimates - unless the filtering is physical.

7.2 The Average Shape And Its Dispersion

Though we have determined the shape in the immediate neighborhood of the peak from §7.1, we still need to obtain the higher order terms in the Taylor expansion in r. We also wish to determine how far out we can go before the density at r becomes uncorrelated with that of the peak. The former requires the average of $F(r)$, the latter requires its dispersion.

A peak is characterized by the parameters $\mathcal{C} = \{\nu, \lambda_1, \lambda_2, \lambda_3, \alpha, \beta, \gamma\}$. The latter 3 parameters are the Euler angles which define the orientation of the principle axes of ζ_{ij}. We also let $\mathcal{C}$ include the information that $\vec{r} = 0$ is a peak. The shape about the peak would be fully determined if we could compute a hierarchy of conditional probability distributions $P[F(\vec{r}_1), ..., F(\vec{r}_N)|\mathcal{C}]$ as $N \to \infty$. Here, we will just compute $P[F(\vec{r}_1)|\mathcal{C}]dF(\vec{r})$, the probability that at a displacement $\vec{r}$ from the peak (taken to be at $\vec{r} = 0$), the density field has value F. As we show in Appendix 4, this is a Gaussian distribution characterized by the the mean value of F at $\vec{r}$ subject to the constraint $\mathcal{C}$,

$$\bar{F}(\vec{r})/\sigma_0 = \langle F(\vec{r})|\mathcal{C}\rangle/\sigma_0$$
$$= \frac{\nu}{(1-\gamma^2)}(\psi + \nabla^2\psi/3) - \frac{x/\gamma}{(1-\gamma^2)}(\gamma^2\psi + \nabla^2\psi/3) + \frac{5}{2}(x/\gamma)(\psi'/r - \nabla^2\psi/3)A(e,p), \tag{7.8}$$

and the variance ($\Delta F = F - \bar{F}$)

$$\langle(\Delta F(\vec{r}))^2|\mathcal{C}\rangle/\sigma_0^2 = 1 - \frac{\psi^2}{1-\gamma^2} - \frac{1}{\gamma^2(1-\gamma^2)}(2\gamma^2\psi + \nabla^2\psi/3)\nabla^2\psi/3 - \frac{5}{\gamma^2}(\psi'/r - \nabla^2\psi/3)^2 - \frac{1}{\gamma^2}(\psi')^2. \tag{7.9}$$

We have introduced the notation $\psi(r) \equiv \xi(\vec{r})/\xi(0)$ for the normalized density-density correlation function as in §6, r is measured in units of R_*, A is given by equation (7.4) and $\psi' \equiv d\psi/dr$. The parameters x, e and p are to be chosen from the distributions 7.5 and 7.6. The dispersion is independent of these parameters, and is spherically symmetric. At large distances, the mean also becomes spherically symmetric, though this reflects the lack of correlation out there rather than indicating sphericity of the actual structures. At short distances, a Taylor expansion of 7.8 in r does indeed reduce to 7.4 in lowest order.

A less useful quantity arises when we average the shape over all possible orientations of the principal axes. The result is of course spherically symmetric and depends only on the average curvature

of the peak and its height:

$$\bar{F}(r)/\sigma_0 = \langle F(r)|\nu, x, e, p\rangle/\sigma_0 \equiv \langle F(r)|\nu, x\rangle/\sigma_0$$
$$= \frac{\nu}{(1-\gamma^2)}(\psi + \nabla^2\psi/3) - \frac{x/\gamma}{(1-\gamma^2)}(\gamma^2\psi + \nabla^2\psi/3). \tag{7.10}$$

This orientation-averaged mean is independent of e and p and agrees with the mean of equation (7.8) if there is no asymmetry, $e = p = 0$. Also, the dispersion is still given by equation (7.9).

The mean shape averaged over all possible curvatures as well as orientations, $\langle F(\vec{r})|\nu\rangle$, is given by the same expression except that x is replaced by its conditional average $\langle x|\nu\rangle$, which is given by equation (6.13). We plot $\bar{F}$ together with the curves $\bar{F} \pm \Delta F$ in Fig. 7.3 for parameters appropriate to adiabatic and isocurvature CDM models. (Instead of $\langle x\rangle$, we used $x = x_*$ in this figure. The difference is small in this case.) Equations (7.10) and (7.9) for peaks should be compared with the average density structure and its dispersion around an arbitrary point with the same height ν as the peaks (Rice 1944, Dekel 1981):

$$\langle F(\vec{r})|\nu,\ no\ peak)\rangle/\sigma_0 = \nu\psi(\vec{r})$$
$$\langle(\Delta F(\vec{r}))^2|\nu,\ no\ peak)\rangle/\sigma_0^2 = (1-\psi^2(\vec{r})). \tag{7.11}$$

This correlation function profile is also plotted in Fig. 7.3. It falls off more slowly than 7.10. Generally, the added constraint that the point is a peak serves to make the profile steeper. Further, the dispersion is less than for the ambient field point. The asymptotic limit of the variance with the peak constraint,

$$\langle(\Delta F(\vec{r}))^2|\nu,\ peak)\rangle \sim \sigma_0^2(1-(1-\gamma^2)^{-1}\psi^2(\vec{r})), \tag{7.12}$$

approaches the uncorrelated limit σ_0^2 slower than the ambient point does by the $(1-\gamma^2)^{-1}$ factor. Typically, one can go several filtering radii before the dispersion becomes so large that knowledge of the central peak conditions gives us no information about the shape. The average spacing of peaks with $\nu > 1$ is $\sim 7R_f$ and with $\nu > 2.68$ is $\sim 15R_f$ in both cases. The shapes have long since lost their coherence before this.

The typical radius of a peak could be estimated as the distance at which the dispersion in the profile becomes unacceptably large. An alternative estimate is simply $d/4$, where $d = 4.0R_*$ is the average spacing between peaks of arbitrary height, as determined from equation (4.11b) for the integrated number density.

The average shapes are, of course, more asymmetric than in the example of Fig. 7.3. As an illustration, we plot (in Fig. 7.4) the density profile along the 1 and 3 axes for the same parameters as above, but for $p = 0$ and the extreme choice of ellipticity $e \sim 0.2$ which, according to the distributions of §7.1.1, is the most likely value. The mean sphericalized profile of Fig.7.3 and an intermediate ellipticity profile are drawn for comparison.

Though these statistically-averaged shapes are indicative of the profiles we may expect, we caution again that they will depend upon the filtering prescription. We believe that they do have physical meaning, however, since the profile is coherent out to so many filtering radii.

7.3 Application to Hot and Warm Dark Matter Models

The shapes are certainly meaningful if the filtering arises by a physical mechanism. Such is the case with collisionless damping in universes dominated by hot (massive neutrino) and warm dark matter with adiabatic fluctuations. Pancakes, where the largest eigenvalue of the shear tensor of the velocity field, σ_1, has a local maximum are apparently the first points where nonlinearity occurs (Zeldovich 1970). However, the matter flows away from these regions, accumulating ultimately at the points where the deepest potential wells exist - high maxima of the density field $F \propto \sum_i \sigma_i$, the trace of the shear tensor. We can analyze the structure and spacing of the typical peaks in these models.

According to Appendix 7, the transfer function for the massive neutrino fluctuation spectrum is an exponential filtering term times the cold dark matter transfer function. The exponential has both a linear and quadratic piece, but the linear term is small. Therefore, it is reasonable to approximate the massive neutrino spectrum by a Gaussian-filtered cold dark matter spectrum with filtering scale $R_{f\nu} \approx 2.6(\Omega_\nu h^2)^{-1}\ Mpc$. We now discuss the application of the results of §7 to neutrino-dominated models with $\Omega = 1$ and $h = 0.5$.

Using Figure 4.1, we obtain $\gamma \approx 0.73$, with the characteristic radius of peaks being $R_* \approx 1.28 R_{f\nu} \approx 13\ Mpc$. The most likely height of the peaks is ~ 1.5, 35% have $\nu > 2$ and 6% have $\nu > 3$. The mean separation of peaks of arbitrary height is 52 Mpc, those with $\nu > 2$ are separated by 76 Mpc, and those with $\nu > 3$ by 130 Mpc.

The average asymmetry parameters are $e \sim 0.24$, $p \sim 0.03$. The structures are thus slightly oblate (with a large dispersion) and relatively flattened so that, with nonlinear amplification upon collapse, pancake-like structures are the typical ones expected.

The flows are definitely not spherically symmetric into the peak, especially as later accretion from pancaked-regions occurs. In spite of the asymmetry, one can use a spherical top hat model to obtain a crude estimate of the properties of the 'virialized' state of such a system upon collapse. As the universe becomes nonlinear, much of its mass accretes onto the high peaks resulting in extremely deep potential wells (White, Frenk and Davis 1984). For definiteness, let us suppose that the universe goes nonlinear at redshift one ($\sigma_0(R_{f\nu}, t_0) = 2$). Then $\nu = 2$ peaks would not completely collapse until $z = 1.4$, assuming the collapse parameter of §5.2 is $f_c = 1.69$. Collapse along the shortest axis may have occurred as early as $z = 3$. The top hat collapsed system would have mass $M \sim 10^{15}\ M_\odot$, virialized radius $R_V \sim 25\ Mpc$, three-dimensional velocity dispersion $\bar{v} \sim 2000\ km\ s^{-1}$, and temperature $\sim 2 \times 10^8 K$. It is not clear that the gas will be sufficiently abundant in these deep 'clusters' to give very large X-ray luminosities, since it can suffer earlier shocking in pancakes which can separate the flow of gas from that of neutrinos. Nonetheless, these numbers do illustrate the difficulties in allowing neutrino-dominated universes to go nonlinear even at these redshifts. The properties of, and the potential problems with, these huge neutrino clusters were investigated by White, Frenk and Davis (1984). Our analysis gives similar cluster spacings and properties to their n-body results.

For warm dark matter, a best Gaussian filtering choice does not give a good fit to the transfer function. The best Gaussian filtering scale would be $R_f \sim 0.4(\Omega_X h^2)^{-1}\ Mpc$ (for $g_{Xdec} = 100$ -see Appendix 7 for notation). Thus, for $h = 0.5$ this is similar to the $R_s = 0.5\ h^{-1} Mpc$ example used in §7.1 and 7.2.

8. DISCUSSION

There are two aspects to this paper. One is to present a set of new mathematical results on the theory of Gaussian random fields. The other is to suggest how such calculations can be used in cosmology to treat rather detailed questions of structure formation.

8.1 MATHEMATICAL RESULTS

The main new mathematical results on the statistics of a random field F_s smoothed to filter short wavelengths are as follows:

(1) Our calculations of the average number densities of peaks, both differential in height $\mathcal{N}_{pk}(\nu)d\nu$ (equation 4.3), and integral in threshold height $n_{pk}(\nu_t)$ (equations 4.11a, 4.12 and 4.21), and of upcrossing points on contour surfaces of given threshold height $n_{up}(\nu_t)$ (equation 4.20).

(2) The calculation of the number density of peaks constrained to having a background field smoothed on a larger scale of given height F_b at the peak points, $n_{pk}(\nu_t|F_b)$ (equation A5.11). The factor $E(F_b)$ expresses this enhancement in the presence of an F_b over the average number in a dimensionless way. The related constrained probabilty $P(F_b|\nu_t,\ peak)$ (equation A5.8) of background field amplitude at points which are peaks above a threshold was also given.

(3) The shapes valid in the immediate neighborhood of a peak are to be determined from the distribution $P(x|\nu)$ (equation 7.5) and the distribution of ellipticity and prolateness $P(e,p|x,\nu)$ (equation 7.6). The shapes farther from the peak point are to be determined from the Gaussian probabilty distribution of values of $F(r)$ given that the center is a peak of given height, value of x and asymmetry parameters and orientation, with a mean profile $\langle F(r)|\nu,x,e,p\rangle$ (equation 7.8) and a dispersion about this mean given by equation (7.9). Spherically symmetric (orientation-averaged) results were also given (equation 7.10).

(4) Correlation functions of peaks were determined only with approximations whose validity depends upon the specific sort of power spectrum (§6.2). $\langle\prod_i n_{pk}(\nu_t,\vec{r}_i)\rangle$ was obtained in the limit that gradients of the normalized two point function ψ of the field could be ignored (generally in equation 6.8, approximated by 6.22, and in the linear limit of small ψ in equation 6.12). The results were shown to deviate substantially at moderate threshold from correlation functions computed using the interiors (Kaiser 1984a,b and Politzer and Wise 1984) or surfaces (equation A6.23) of contour regions. The contour region correlations overly weight the large contour regions. Another useful and conceptually simple method for the calculation of correlation functions uses the peak-background split, $\langle\prod_i n_{pk}(\nu_t|F_b(\vec{r}_i))\rangle$ (equations 6.25 and 6.41). The result was shown to agree with the direct calculation provided the background R_b is sufficiently large compared with the smoothing scale R_s. The first order 'dynamical' correlations $\langle\prod_i n_{pk}(\nu_t|F_b(r_i))\ (1+F_b(r_i,t))\rangle$ were also obtained in the limit that $F_b(t)$ is in the linear regime (e.g., equation 6.63 for the two point function).

8.2 Applications to Models of Structure Formation

8.2.1 *Framework*: We used the adiabatic and isocurvature cold dark matter spectra based on initial scale-invariant spectra to demonstrate the use of these statistical formulas in a cosmological setting. We emphasize that these methods are applicable to other power spectra which may be considered, though care is required in using some of the approximations (especially regarding the correlation functions) with steep (high n) spectra. Fits to the power spectra for the two cold dark matter models, and for the hot (massive neutrino) and warm adiabatic dark matter models are given in Appendix 7. The methods we use can also be applied to the old isothermal and adiabatic structure formation models appropriate to baryon-dominated universes. Further, even if it is found that most of the structure in the universe arose from explosive (Ostriker and Cowie 1981, Ikeuchi 1981) or radiative (Hogan 1983, Hogan and Kaiser 1984) events, rare peaks in an initial fluctuation field would still be required to initiate the activity, and these can be analyzed using our methods. The only requirement for applicability of this framework is that the field be Gaussian. As a consequence of the central limit theorem, this is the most likely case, so the implications of such an assumption deserve to be fully explored.

Our calculations give unambiguous answers about the texture of the linear density fluctuation field. However, relating structure present in the linear regime to final nonlinear luminous objects is a subject open to much debate. Our prescription here was to select a special class of points, the peaks of the fluctuation field smoothed by a filtering process applied to the power spectrum. We selected only those with heights in some specified range, typically taken to be above a global threshold, though restricted ranges of height may be more appropriate for some applications. For a given class of cosmic objects, the optimal choice of filter and of the selection function (equation 4.13) is questionable. Selection functions might not just be based on height, but on shape and internal or external environment of the peak. How to choose functions which select different Hubble types of galaxies is unclear. Filtering smooths small clouds into larger ones of lower heights leading to the possibility of overcounting of the objects which are truly associated with a given scale. Dynamical merging of peaks as structure evolves modifies the observed number of objects of a given scale. These difficulties complicate the comparison of the theory of the linear texture with the observations.

8.2.2 *Profiles*: Average smoothed profiles associated with the peaks were considered in §7. We found a triaxial ellipsoid approximation to be a valid description in the immediate neighborhood of the peak.

The average degree of prolateness or oblateness of the ellipsoids is found to go to zero faster than ellipticity as the height of the peak is increased, although there is typically a wide dispersion in peak asymmetries (equation 7.6). Predominantly spherical or even spheroidal collapses are extremely unlikely except for very high peaks. This is predicted in spite of the tendency of the smoothing operation to reduce asymmetry. The average shapes valid farther from the peak (equation 7.8) reproduce the triaxial ellipsoid in close but become spherically symmetric far from the peak, reflecting not the true structure but only that the outer points are uncorrelated with the central peak - expressed by the growth of the dispersion (7.9) with distance to the asymptotic *rms* value σ_0. Indeed, the mean spacing between peaks of arbitrary height, $4R_*$, implies that beyond $\sim 2R_*$, the variance in the shapes reflects the presence of other peaks. Nonetheless, since high peaks are relatively coherent in structure up to a few filtering radii, these shape calculations do give a strong indication of the smoothed generic collapse structure.

Unless the filter is physical (as in pancake models), these profiles cannot be used for hydrodynamic or n-body studies of collapse since substructure must be included. Subclumping within collapsing clouds will be important in determining their final configuration. This cloud-in-cloud problem is difficult due to the high degree of correlation of the structure on smaller scales with the smoothed structure. The probability $P(F_b|\nu_s,\ peak)$ can be used to get an idea of the sort of environment in which a given peak finds itself. For R_b near to R_s, the dispersion is small, and the average background height, $\nu_b \sim \left(R_s/R_b\right)^{(n+3)/2}\ \nu_s$ for Gaussian filtering, falls off slowly, showing that the typical small-scale peak is not isolated in the background. Though we have no adequate solution to the treatment of these subclouds, a crude estimate of the number of peaks in a given top hat region, $(4\pi/3)E(\nu_b)n_{pk}(\nu_t)R_{TH}^3$, was given in §5.4 for 'bright' galaxies within an Abell radius of the cluster center. The value obtained, $\sim 10^2$, is similar to that observed.

8.2.3 *Average Number Densities of Cosmic Structures*: Once a selection function is adopted, we can determine the average number of objects per comoving volume and relate it to observed values to fix, for example, the threshold. In §6.4 we found the threshold required for galaxy formation by equating the density of peaks smoothed on a scale $R_s \sim (0.2 - 0.4)\ h^{-1}Mpc$ to the density of bright galaxies $0.01\ h^3Mpc^{-3}$. Thresholds in the range $\nu_t \sim 2 - 4$ were obtained, depending on threshold function assumptions. That the density of bright galaxies is so low is an argument in itself that not all peaks collapsing on galactic scale can become luminous galaxies. Of course, this does not necessarily require a global threshold.

The same operation can be applied to Abell clusters with density $6 \times 10^{-6}\ h^3Mpc^{-3}$ for richness class ≥ 1. We present the steps in the calculation to illustrate the results of §4. We adopt a Gaussian smoothing radius $R_s \approx 5\ h^{-1}Mpc$ based on the mass $5 \times 10^{14}\ h^{-1}\ M_\odot$. If we assume an adiabitic CDM model with $\Omega = 1$, $h = 0.5$, we can use Figure 4.1 to get $\gamma \approx 0.73$ and $R_* \approx 1.27R_s$ at $10\ Mpc$. For simplicity, we adopt a sharp threshold. Using §4.3, we require the threshold satisfy $n_{pk}(\nu_t)/n_{pk}(-\infty) \approx 0.01$, i.e. $n_{pk}(\nu_t)R_*^3 \approx 1.5 \times 10^{-3}$. Interpolating between the $\gamma = 0.7$ and 0.9 lines in Fig. 4.3, we obtain $\nu_t \approx 2.8$. (Another method valid in this regime would be to invert the upcrossing density (4.20). A selection function with a soft threshold would require a numerical integration over the differential peak number density (4.13).) For clusters of this threshold height to have collapsed by now (in a spherical top hat model) would require $\sigma_0(5\ h^{-1}Mpc) = 1.69/\nu_t \approx 0.6$. This, in turn, implies that on galactic scales $\sigma_0(0.356\ h^{-1}Mpc) \approx 4$. Normalization to the galaxy clustering data in §6.6 gave the value 3 instead. Lowering the collapse factor f_c for rich clusters could bring these numbers closer. This problem is discussed more fully in §8.2.7 below. Similar procedures could be adopted for determining thresholds for each richness class of Abell clusters in turn (Kaiser 1984a).

Bond, Szalay and Silk (1985) have suggested that intergalactic Lyman α clouds are associated with primordial peaks filtered on $\sim 10^2\ kpc$ scales with a hydrodynamically-determined selection function arising as a consequence of re-ionization of the universe. The predicted density of selected peaks as a function of redshift can be compared with the observed density to test the theory. Appropriate

selection functions for dwarf galaxies (Dekel and Silk 1985) may also lead to a confrontation of cold dark matter theory with observation.

We have not been able to find an appropriate mass function $n(M)dM$ for the objects due to the cloud-in-cloud problem.

8.2.4 *Peak Enhancements in Overdense Regions*: Clustering is a consequence of two effects. One is *statistical* in origin: our peaks were already clustered when selected since peaks are preferentially found in overdense large scale regions; this effect depends only on the background height ν_b, not on the amount that the background field has grown, i.e. on $\sigma_0(t)$. The other is *gravitational* in origin: the peaks move with the mass as it flows into overdense regions out of underdense ones; this does depend on the amplitude $\sigma_0(t)$. In the CDM models, we find that each is about equally important in determining the current degree of correlation of galaxies.

Assuming that peaks flow with the mass, the enhancement factor $E(\nu_b)$ of peaks in overdense regions (5.10) derived from $n_{pk}(\nu_t|\nu_b)$ can be applied to the computation of peak-to-mass ratios, and, if a luminosity is associated with each object, of M/L ratios. We applied this to the enhancement of galaxy number in rich clusters in §5.4 and found $E \sim 3-10$ for peaks associated with the thresholds appropriate to 'bright' galaxies. These values are sensitive to the choice of cluster filter and collapse criterion. We normalized the time of galaxy formation with the galaxy correlation function. We also predict that M/L should be higher in the outer regions of clusters and be smaller in denser clusters. The same methods can be used to estimate the paucity of peaks in voids. Also, if we take supercluster environments as background constraints, we find even stronger enhancements in the cluster number per mass.

8.2.5 *Statistical Correlation Functions*: We presented a very detailed discussion of correlation function calculations in the two CDM models since this represents one of the most powerful applications of Gaussian statistics to cosmology. We first determined the correlation functions due to the statistical clustering effects alone, then included dynamical effects in the linear approximation. Two different calculational methods were applied to statistical clustering. The most accurate used equation (6.21), the n-point function determined assuming the peaks are sufficiently far apart that derivatives of ψ can be neglected. All nonlinear terms must be included, however, unless one is only interested in the linear regime of the peak correlation function. The peak-background split, equation (6.41), is somewhat easier to use, but only becomes accurate at a larger separation. The two-point correlation functions were determined assuming a global threshold exists for galaxy formation which is found by normalizing to the galaxy density. The choice of smoothing radius is debatable, so a range was selected. The principal results (Table 6.1 and Figure 6.1) show that (1) in the large separation limit the peak correlation function is proportional to the (normalized) mass density correlation ψ_ρ, but is amplified by the square $(\langle\tilde{\nu}\rangle^2)$ of an effective threshold (equation 6.14) which is significantly less than ν_t^2; (2) in the nonlinear regime, $\xi_{pk} > 1$, the spectrum is quite close to a power law. What is remarkable is that these statistical correlations lead, for the adiabatic spectrum, to a fairly good $r^{-1.8}$ power law. The slope is somewhat sensitive to the choice of filtering radius. Without including any dynamics, the isocurvature model with the smaller filtering radii already gives ξ_{pk} nearly one at $r_0 = 5\ h^{-1}Mpc$, the point where the observed galaxy correlation function is unity. The adiabatic ξ_{pk} is a factor ~ 5 smaller.

The (connected) three-point (statistical) correlation function of the galactic peaks gives values of Q defined by equation (6.49) near to that observed (~ 1) for the adiabatic CDM model. However, a significant amount of dynamical evolution of the three-point function over the scales where it has been reliably determined would be expected, and this may modify the good agreement between theory and observation of Q.

8.2.6 *Dynamical Correlations and the Redshift of Galaxy Formation*: Inclusion of dynamics is very complicated because the peak number density per physical (as opposed to Lagrangian) volume at the Eulerian position $\vec{x}$ at time t depends not only on the $1+F$ enhancement associated with the transition

from Lagrangian to Eulerian volume, but also on the complex statistical field $\vec{r}(\vec{x},t)$ describing the motion of the peaks through Eulerian space. Only by ignoring the latter effect has the problem proved tractable. The approximation procedure adopted in §6.6 involved using the peak-background split, requiring F_b to be in the linear regime, and neglecting this intrinsic motion of the peaks. On large scales, this should be adequate, and the result is given by equation (6.63). For $\xi_{pk} < 1$, we find $\xi_{pk}(r,t) \approx b^2(t)\xi_\rho(r,t)$, where $b = \langle\tilde{\nu}\rangle/\sigma_{0s}(t)+1$, which measures the amplitude of biasing, is typically ~ 2.

Requiring $\xi_{pk}(r_0,t_0) = 1$ sets σ_0 on galaxy scales at the present, giving a normalization to the power spectrum which fixes when structure collapsed on all other scales. Together with the galaxy formation threshold, σ_0 determines the redshift of galaxy formation $1+z_g = f_c^{-1}\langle\nu\rangle\sigma_0$ (Table 5.2). For definiteness, we discuss the $\Omega = 1$, $h = 0.5$, $q = 8$ *adiabatic* CDM example with $R_s = 0.356\ h^{-1}Mpc$ and $f_c = 1.69$. Then $\nu_t \approx \langle\nu\rangle \approx 2.8$ and $\sigma_0 \approx 2.4$. Since objects on this scale with $\nu > f_c\sigma_0^{-1} = 0.7$ will have collapsed by the present, we must suppose 'bright' galaxies were not made below the redshift $z_g = 2.9$. It is not clear what these 'failed' galaxies should be identified with observationally; they could just be low surface brightness objects.

The prediction that galaxy formation occurred as late as the epoch $z_g \approx 3-4$ may already be in trouble with limits from primeval galaxy searches. Koo (1985) estimates that $z_g > 5$ is required, though, with slow steady star formation rather than a burst at galaxy birth or dust re-radiation of the starburst energy in the infrared, such a strong limit can be avoided. We may rather regard a late galaxy formation epoch as one of the exciting testable predictions of biased galaxy formation. The collapse of halos at late times apparently does lead to appropriate halo profiles ($\sim r^{-2}$) and velocity dispersions for spiral galaxy models in the CDM scenario (Ryden and Gunn 1985, Carlberg and Lake 1985, Miller 1985, Primack *etal.* 1985, Frenk *etal.* 1985).

The amplitude of the statistical correlations in the *isocurvature* CDM picture implies there is very little room left for dynamical evolution. Even for the larger galactic filtering radii in Table 6.1b, the strongly inadmissable value $z_g \sim 0.1$ was obtained. In addition, the isocurvature models give large angle microwave background anisotropies and large scale velocity fields in excess of observational constraints (Efstathiou and Bond 1985).

8.2.7 *The Cluster Threshold Problem*: Fixing σ_0 on galactic scales also fixes the threshold for peaks on rich cluster scales which will have collapsed by the present. Consider the adiabatic CDM example of §8.2.6. Scaling $\sigma_0(0.356\ h^{-1}Mpc) \approx 2.4$ for $h = 0.5$ to (Gaussian) cluster smoothing scales, we find $\sigma_0(5\ h^{-1}Mpc) \approx 0.4$. Using a collapse parameter $f_c = 1.69$ leads to the threshold $\nu_t \approx 4.4$; in §8.2.3, we found $\nu_t \approx 3$ was required to reproduce the cluster number density. This special case illustrates a general problem with biased galaxy formation: pushing the epoch of galactic scale nonlinearity σ_0 down due to the galaxy correlation function bias also pushes down the amount of dynamical evolution on large scales so that clusters which have collapsed might be rarer than observed. Since $\sigma_0(R_f)$ falls off significantly with increasing R_f, larger filtering radii (which may be appropriate for Abell clusters of higher richness class) lead to more severe problems. Another prediction of the model is that the threshold would be a factor of 2 higher at redshift 1, implying collapsed clusters would be exceedingly rare then. However, an enhancement in the number of galaxies is predicted where rich clusters will eventually form, so 'statistical' clusters may be seen. If an abundance of collapsed (virialized) clusters similar to that observed at $z \sim 0$ can be shown to exist at $z \sim 1$, the simple biasing hypothesis in the adiabatic CDM scenario can be ruled out.

Though the cluster density problem indicates the adiabatic CDM spectrum may lack sufficient power on large scales with the biasing normalization to adequately account for the observed large scale structure, we do not yet regard it as a fatal flaw for the model. Uncertainties in cluster smoothing scales and collapse parameters due, for example, to subclustering, asymmetric collapse and lack of virialization make our estimates imprecise. This large scale structure issue certainly warrants further investigation. The adiabatic CDM spectrum with $\Omega = 1$ and a global threshold imposed for galaxy formation does offer a promising explanation for the clustering properties of galaxies and a possible

reconciliation of low M/L ratios in clusters with a global $\Omega = 1$ mass density.

We wish to thank the Institute for Theoretical Physics at Santa Barbara for its hospitality during the workshop on Large Scale Structure where this collaboration began and Rocky Kolb for bringing us all together for a few days at Fermilab. JMB was supported in part by the DOE under contract DE-AC06-81ER40048 at Washington, JRB by a Sloan Research Fellowship and by NASA grant NAGW-299 at Stanford, NK by the SERC, AS by the NASA/DOE astrophysics program at Fermilab and by DOE grant 84-ER-40161 at Berkeley. At ITP, this research was supported by the NSF under grant PHY77-270843, supplemented by funds from NASA.

APPENDIX 1: Derivation of $\mathcal{N}_{pk}(\nu)$

We follow the Rice notation for the derivatives of the random field $F(\vec{r},t)$: $\eta_i = \nabla_i F$, $\zeta_{ij} = \nabla_i \nabla_j F$. The correlations of these fields at an arbitrary point are:

$$\begin{aligned}
&\langle FF\rangle = \sigma_0^2, \qquad \langle \eta_i \eta_j\rangle = \frac{\sigma_1^2}{3}\delta_{ij},\\
&\langle F\zeta_{ij}\rangle = -\frac{\sigma_1^2}{3}\delta_{ij}, \qquad \langle \zeta_{ij}\zeta_{kl}\rangle = \frac{\sigma_2^2}{15}(\delta_{ij}\delta_{kl}+\delta_{ik}\delta_{jl}+\delta_{il}\delta_{jk})\\
&\langle F\eta_i\rangle = 0, \qquad \langle \eta_i\zeta_{jk}\rangle = 0.
\end{aligned} \tag{A1.1}$$

The joint Gaussian probability distribution for the variables F, η_i, and ζ_{ij} is given by equation (2.2). Due to the symmetry of ζ_{ij}, only 6 components are independent. We label them by ζ_A, where the $A = 1,2,3,4,5,6$ components of the 6-dimensional vector refer to the $ij = 11,22,33,23,13,12$ components of the tensor. The covarince matrix M thus has dimension 10; however 6 of the dimensions are already diagonalized ($\{\eta_i, \zeta_A,\ A = 4,5,6\}$). To diagionalize the remaining 4 dimensions, we transform to a new set of variables: $\{\zeta_A,\ A = 1,2,3\} \rightarrow \{x,y,z\}$, where

$$\begin{aligned}
&\sigma_2 x = -\nabla^2 F = -(\zeta_1+\zeta_2+\zeta_3)\\
&\sigma_2 y = -(\zeta_1-\zeta_3)/2, \qquad \sigma_2 z = -(\zeta_1-2\zeta_2+\zeta_3)/2.
\end{aligned} \tag{A1.2}$$

We also introduce $\nu = F/\sigma_0$. With these choices,

$$\langle \nu^2\rangle = 1,\ \langle x^2\rangle = 1,\ \langle x\nu\rangle = \gamma,\ \langle y^2\rangle = 1/15,\ \langle z^2\rangle = 1/5, \tag{A1.3}$$

and all other correlations are zero. Thus, the matrix is now diagonal in y and z. The quadratic form Q appearing in $P \propto e^{-Q}$ is then simply

$$2Q = \nu^2 + \frac{(x-x_*)^2}{(1-\gamma^2)} + 15y^2 + 5z^2 + \frac{3\vec{\eta}\cdot\vec{\eta}}{\sigma_1^2} + \sum_{A=4}^{6}\frac{15\zeta_A^2}{\sigma_2^2} \tag{A1.4}$$

Here,

$$\begin{aligned}
&x_* \equiv \gamma\nu,\\
&\gamma \equiv \frac{\sigma_1^2}{\sigma_2\sigma_0} = \frac{\langle k^2\rangle}{\langle k^4\rangle^{1/2}}.
\end{aligned} \tag{A1.5}$$

At the moment, the variables x, y, and z are defined for an arbitrary choice of axes. The correlations given by equation (A1.5) are independent of this choice. We now restrict our attention to the principal axes, and introduce the eigenvalues λ_i of $(-\zeta_{ij})$; therefore, $\zeta_A = -\lambda_A$, $A = 1,2,3$, and x, y and z are now defined in terms of the $\lambda's$. The other 3 degrees of freedom in the matrix can be expressed in terms of the Euler angles α,β,γ required to define the orientation of the triad of orthonormal eigenvectors of ζ_{ij}.

In Appendix 2, we prove that the volume element for the space of symmetric real 3×3 matrices can be expressed in terms of the eigenvalues and Euler angles (α,β,γ)

$$\begin{aligned}
&\prod_{A=1}^{6} d\zeta_A = |(\lambda_1-\lambda_2)(\lambda_2-\lambda_3)(\lambda_1-\lambda_3)|d\lambda_1 d\lambda_2 d\lambda_3 \frac{d\Omega_{S^3}}{6},\\
&d\Omega_{S^3} = \sin\beta d\beta d\alpha d\gamma.
\end{aligned} \tag{A1.6}$$

Here, $d\Omega_{S^3}$ is the volume element on the surface of the 3-sphere. Since Q is independent of the Euler angles, an expression of all triad orientations being equally probable, we can integrate over the

3-sphere, whose volume is $2\pi^2$. The factor 3! arises if the eigenvalues are not ordered. A first set of rotations of one of the coordinate axes to one of the principal axes requires only $4\pi/(2\cdot 3)$ steradians of solid angle; the 2 comes from not caring whether the rotated axis points in the positive or negative direction; the 3 comes from the one chosen axis necessarily being within $4\pi/3$ steradians of a principal axis, unless one of the principal axes is chosen for some special property (e.g. largest eigenvalue), and the principal axes are labelled according to this ordering. Having fixed one of the principal axes, a rotation in the plane perpendicular by an angle at most $2\pi/(2\cdot 2)$ will align the other 2 axes with the principal ones; again, 2 comes from not caring about the direction of pointing, and 2 comes from not trying to align labelled axes. The total volume of S^3 needed is therefore $2\pi^2/3!$. To transform from λ_i to x, y, z, we also need

$$\prod_{A=1}^{3} d\lambda_A = \frac{2}{3}\sigma_2^3 dxdydz$$

This 2/3 is compensated by a 2/3 that arises in transforming the determinant in the ς's to one in the x, y, z system.

The relevant joint probability now becomes

$$P(\nu,\vec{\eta},x,y,z)d\nu d^3\vec{\eta}dxdydz = N|2y(y^2-z^2)|e^{-Q}d\nu dxdydz\frac{d^3\eta}{{\sigma_0}^3},$$
$$N = \frac{(15)^{5/2}}{32\pi^3}\frac{\sigma_0^3}{\sigma_1^3(1-\gamma^2)^{1/2}}. \qquad \text{(A1.7)}$$

No assumptions have been made so far about the ordering of the eigenvalues. We now enforce the order

$$\lambda_1 \geq \lambda_2 \geq \lambda_3. \qquad \text{(A1.8)}$$

Since there are 5 other possible orderings we could have chosen, and P is invariant under changes of ordering, we must multiply the probability expression by 6 to account for this. This compensates for the 3! lowering of the available S^3 volume due to the identical nature of the axes. Thus, if the eigenvalues are ordered, the entire volume of S^3 is available for triad rotation to the principal. With this ordering, the constraints $y \geq z \geq -y, \quad y \geq 0$ are imposed. These are the only constraints necessary if we are considering all extrema, for which the λ_i can be positive or negative. A further constraint does arise if we require all eigenvalues to be positive - as is necessary for maxima. With our ordering convention, this amounts to the requirement that $\lambda_3 = \sigma_2\big((x+z)/3 - y\big)$ be positive. This is the extra constraint on the $x-y-z$ domain of integration needed to deal with maxima.

We now consider the consequences of requiring that $\vec{\nabla}F$ vanish at an extremum. The shape of the random field $F(\vec{r})$ in the neighborhood of the peak point $\vec{r}_p$ is determined by a Taylor series expansion:

$$F(\vec{r}) \approx F(\vec{r}_p) + \frac{1}{2}\sum_{ij}\varsigma_{ij}(\vec{r}_p)(r-r_p)_i(r-r_p)_j, \qquad \eta_i(\vec{r}) \approx \sum_j \varsigma_{ij}(\vec{r}_p)(r-r_p)_j. \qquad \text{(A1.9)}$$

The *full* random density field for the maxima of height between ν_0 and $\nu_0 + d\nu$ is

$$\begin{aligned} n_{pk}(\vec{r},\nu_0)d\nu &= \sum_p \delta^{(3)}(r-r_p) \\ &= \delta^{(3)}[\ (\varsigma^{-1})\vec{\eta}\ ]\ \theta(\lambda_1)\theta(\lambda_2)\theta(\lambda_3)\ \delta(\nu-\nu_0)d\nu \\ &= |det(\varsigma)|\ \delta^{(3)}[\vec{\eta}]\ \theta(\lambda_1)\theta(\lambda_2)\theta(\lambda_3)\ \delta(\nu-\nu_0)d\nu. \end{aligned} \qquad \text{(A1.10)}$$

Here, θ is the Heaviside unit function. Ideally, we would like to know the probability density functional of this random field. Notice that such a functional would be that of a point distribution, since there

are only countably many zeros of $\vec{\eta}$, each separated from the other (Adler 1981). This general problem appears to be intractable. In this paper, we only determine the average of this expression.

The average peak density for maxima of height ν_0 is therefore given by

$$\mathcal{N}_{pk}(\nu_0)d\nu \equiv \langle n_{pk}(\vec{r},\nu_0)\rangle d\nu = \langle |\lambda_1\lambda_2\lambda_3|\ \theta(\lambda_3)\delta(\nu-\nu_0)\rangle\ d\nu. \tag{A1.11}$$

Of course this mean density is independent of the position due to the homogeneity of the underlying random density field, so we suppress the $\vec{r}$ variable in the following. We also denote the average by $\mathcal{N}_{pk}$.

To calculate A1.11, we need to derive a number of results along the way which are useful in their own right. By introducing extra δ-functions in A1.10, we can further restrict the class of maxima we are considering to be those with the parameters $x-y-z$ as well as ν in specific infinitessimal ranges. The density of this class of peaks is then

$$\begin{aligned}\mathcal{N}_{pk}(\nu,x,y,z)\ d\nu dxdydz &= \frac{5^{5/2}3^{1/2}}{(2\pi)^3}(\frac{\sigma_2}{\sigma_1})^3\frac{1}{(1-\gamma^2)^{1/2}}e^{-\tilde{Q}}\ F(x,y,z)\chi\ d\nu dxdydz,\\ \tilde{Q} &= \frac{\nu^2}{2}+\frac{(x-x_*)^2}{2(1-\gamma^2)}+\frac{5}{2}(3y^2+z^2),\end{aligned} \tag{A1.12}$$

where

$$\begin{aligned}F(x,y,z) &= \frac{27}{2}\sigma_2^{-6}\lambda_1\lambda_2\lambda_3(\lambda_1-\lambda_2)(\lambda_2-\lambda_3)(\lambda_1-\lambda_3)\\ &= (x-2z)((x+z)^2-(3y)^2)y(y^2-z^2),\end{aligned} \tag{A1.13}$$

and χ is a characteristic function which is 1 if the constraints in the $x-y-z$ domain are satisfied, and is 0 if they are not. This density is used in Appendix 3 to get the probability distribution of the asymmetry variables y and z.

We can integrate A1.12 first over z then over y to get the density

$$\mathcal{N}_{pk}(\nu,x)\ d\nu dx = \frac{e^{-\nu^2/2}}{(2\pi)^2R_*^3}\ f(x)\ \frac{exp(-\frac{(x-x_*)^2}{2(1-\gamma^2)})}{(2\pi(1-\gamma^2))^{1/2}}\ d\nu dx. \tag{A1.14}$$

The integrals over z and y are contained in the function

$$f(x) = \frac{3^2 5^{5/2}}{\sqrt{2\pi}}\left(\int_0^{x/4} e^{-\frac{1}{2}\frac{5}{2}y^2}dy\int_{-y}^{y}F(x,y,z)\ e^{-\frac{5}{2}z^2}dz + \int_{x/4}^{x/2} e^{-\frac{1}{2}\frac{5}{2}y^2}dy\int_{3y-x}^{y}F(x,y,z)\ e^{-\frac{5}{2}z^2}dz\right).$$

The computation of $f(x)$ is tedious, but the result can be expressed in closed form:

$$\begin{aligned}f(x) &= (x^3-3x)(\ erf(\ (\frac{5}{2})^{1/2}x\ ) + erf(\ (\frac{5}{2})^{1/2}x/2\ )\ )/2\\ &+ (\frac{2}{5\pi})^{1/2}\ \left((31x^2/4+8/5)e^{-5x^2/8}+(x^2/2-8/5)e^{-5x^2/2}\right).\end{aligned} \tag{A1.15}$$

The asymptotic limits of this function include a remarkable cancellation to eighth order at small x and the x^3 law expected for spherical peaks at large x:

$$\begin{aligned}f(x) &\to \frac{3^5 5^{3/2}}{7\ 2^{11}\sqrt{2\pi}}\ x^8\ (1-5x^2/8) \qquad \text{as} \quad x\to 0\\ &\to x^3\ -3x \qquad \text{as} \quad x\to\infty.\end{aligned} \tag{A1.16}$$

A reasonable approximation to this function is given by

$$\begin{aligned} f_{ap}(x) &= x^3 - 3x + 4.08/x^2 \qquad \text{for} \quad x \geq 1.5 \\ &= \frac{x^8}{13.2(1+5x^2/8)} \qquad \text{for} \quad x < 1.5. \end{aligned} \tag{A1.17}$$

Another fit appropriate to the large x limit is given by equation (6.19). Whereas the function $\exp(-\frac{(x-x_*)^2}{2(1-\gamma^2)})$ is rapidly falling as x increases, f is monotonically and rapidly rising: the product gives a relatively symmetric function with a clear maximum. (See Figure 7.1 for the plot of a function proportional to this product.)

We must integrate A1.14 numerically over x to obtain the differential density of peaks in the range ν to $\nu + d\nu$:

$$N_{pk}(\nu)\, d\nu = \frac{1}{(2\pi)^2}(\frac{\sigma_2}{\sqrt{3}\sigma_1})^3 e^{-\nu^2/2} G(\gamma, x_*) d\nu. \tag{A1.18}$$

The function

$$G(\gamma, x_*) = \int_0^\infty dx\; f(x)\; \frac{exp(-\frac{(x-x_*)^2}{2(1-\gamma^2)})}{(2\pi(1-\gamma^2))^{1/2}} \tag{A1.19}$$

is very accurately fit by equation (4.5) if the coefficients are given by equation (4.6). The fit was obtained by determining the asymptotic large x_* expansion of G (equation (4.5) with all $C_i = 0$), then adding the appropriate nonzero C_i to get accurate results at low x_*.

To obtain the number of peaks in excess of ν, $n_{pk}(\nu)$, a further integration is required. Again this integral must generally be done numerically, even using the approximate formula for G.

The evaluation of the density of upcrossing points (4.20) and of the Euler characteristic (4.15) is straightforward. The $\vec{\eta}$ part of the calculation is trivial since the η_i are statistically independent of the other variables. The rest of the calculation is most easily accomplished if the variable combinations

$$a = (\varsigma_{11} + \varsigma_{22})/\sigma_2, \quad b = \left((\varsigma_{11} - \varsigma_{22})^2 + (2\varsigma_{12})^2\right)^{1/2}/(2\sigma_2), \quad \phi = arctan(2\varsigma_{12}/(\varsigma_{11} - \varsigma_{22}))$$

are used in place of the ς variables, for all 3 are independent of each other and only the first is correlated with ν:

$$\begin{aligned} P(\nu, a, b, \phi)\; d\nu da db d\phi = &\frac{(15/8)^{1/2}}{2\pi(1-5\gamma^2/6)^{1/2}} exp[-\frac{(\nu^2 - 5\gamma\nu a/2 + 15a^2/8)}{2(1-5\gamma^2/6)}]\; d\nu da \\ &\times\; \theta(b) exp(-15b^2/2) 15b db\; \frac{d\phi}{(2\pi)}. \end{aligned} \tag{A1.20}$$

ϕ has a random phase distribution and can be immediately integrated. Also, since $\varsigma_{11}\varsigma_{22} - \varsigma_{11}^2 = a^2/4 - b^2$, the b integration is straightforward in spite of the absolute value constraint. The simple analytic forms (4.15) and (4.20) follow because the range of the a integration is $(-\infty, \infty)$.

APPENDIX 2: The Volume Element in the Space of Symmetric Matrices

In this Appendix, we sketch the proof that the volume element in the 6-dimensional space of symmetric real matrices is given by

$$d\ vol = \prod_{A=1}^{6} d\varsigma_A = |(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)| d\lambda_1 d\lambda_2 d\lambda_3\; d\ vol\ (SO(3)) \tag{A2.1}$$

where

$$d\ vol\ (SO(3)) = \sin\beta d\beta d\alpha d\gamma \tag{A2.2}$$

is the volume element of the 3-dimensional rotation group $SO(3)$ and also of the 3-sphere, the space which the Euler angles ($\alpha, \gamma \in [0, 2\pi]$, $\beta \in [0, \pi]$) coordinatize.

We define the inner product of two symmetric matrices S_1 and S_2 to be $Tr(S_1 S_2)$. The metric in the space of symmetric matices is $ds^2 = Tr((dS)^2)$ where S is a symmetric matrix. Now S can be diagonalized by a rotation R to a diagonal matrix $\lambda = diag(\lambda_1, \lambda_2, \lambda_3)$: $S = R^\dagger \lambda R$. Hence, $dS = R^\dagger(\lambda + [\lambda, R^\dagger dR])R$, where [,] denotes the commutator. Thus,

$$ds^2 = Tr((d\lambda)^2) + Tr[\lambda, R^\dagger dR]^2.$$

The cross terms disappear since they involve the trace of a product of a symmetric and an antisymmetric matrix. Now, $(R^\dagger dR)_{ij} = \epsilon_{ijk}\omega_k$ is antisymmetric, where ω_k is a vector (infinitessimal) and ϵ_{ijk} is the alternating symbol, so $[\lambda, R^\dagger dR]_{ij} = (\lambda_i - \lambda_j)\epsilon_{ijk}\omega_k$. Thus the metric on the symmetric tensors is

$$ds^2 = \sum_i (d\lambda_i)^2 + (\lambda_2 - \lambda_3)^2\omega_1^2 + (\lambda_3 - \lambda_1)^2\omega_2^2 + (\lambda_1 - \lambda_2)^2\omega_3^2$$

with orthonormal basis

$$\{d\lambda_1, d\lambda_2, d\lambda_3, |\lambda_2 - \lambda_3|\omega_1, |\lambda_3 - \lambda_1|\omega_2, |\lambda_1 - \lambda_2|\omega_3\}.$$

Since the volume element is the 'wedge product' of orthonormal basis elements, we have

$$d\ vol = |(\lambda_1 - \lambda_2)(\lambda_2 - \lambda_3)(\lambda_1 - \lambda_3)| d\lambda_1 \wedge d\lambda_2 \wedge d\lambda_3 \wedge \omega_1 \wedge \omega_2 \wedge \omega_3.$$

However, the matrices $R^\dagger dR$ form the Lie algebra of SO(3), the ω_i form an orthonormal basis, so the volume element of $SO(3)$ is given by

$$d\ vol\ (SO(3)) = \omega_1 \wedge \omega_2 \wedge \omega_3.$$

That the (unnormalized) volume element of the rotation group is in fact the same as the volume element of the 3-sphere and is given by equation (A2.2) is standard.

APPENDIX 3: Conditional Probability for Ellipticity and Prolateness

The conditional probability for the parameters y and z subject to the constraint that the point is a peak with given values of ν and x is simply obtained by taking the ratio of A1.12 and A1.14:

$$P(y, z|\nu, x)dydz = \frac{3^2 5^{5/2}}{\sqrt{2\pi}} \frac{F(x, y, z)\chi}{f(x)}\ exp(-\frac{5}{2}(3y^2 + z^2)) \qquad \text{(A3.1)}$$

where F is defined by equation (A1.13) and χ is the constraint characteristic function. In §7.1 we introduced the 'ellipticity' and 'prolateness' parameters

$$e = y/x, \qquad p = z/x \qquad \text{(A3.2)}$$

in terms of which

$$\begin{aligned} \chi(e,p) &= 1\ if\ 0 \le e \le \frac{1}{4}\ and - e \le p \le e \\ &= 1\ if\ \frac{1}{4} \le e \le \frac{1}{2}\ and - (1 - 3e) \le p \le e \\ &= 0\ otherwise. \end{aligned} \qquad \text{(A3.3)}$$

Notice that equation (A3.1) is independent of ν. To obtain equation (7.6) for the conditional probability $P(e,p|x)$, we introduce the polynomial

$$\begin{aligned} W(e,p) &\equiv x^{-8}\ F(x, y=ex, z=px)\ \chi(e,p) \\ &= e(e^2-p^2)(1-2p)((1+p)^2-9e^2)\chi(e,p). \end{aligned} \tag{A3.4}$$

The nature of the contour plots of $P(e,p|x)$ given in Fig.7.2 suggest that a Gaussian approximation centered about the most probable values e_m and p_m may be a good approximation. If we approximate $V \equiv W\ e^{-x^2(3e^2+p^2)}$ by

$$\ln V \approx -(p^2+3e^2)5x^2/2 + 3\ln e - p^2/e^2 - 3p^2 - 2p^3 - 9e^2 + 18e^2p$$

then use the zeros of the gradient of V to obtain the most probable values, we get equation (7.7). Similarly, we define

$$\sigma_e^{-2} \equiv \left(\frac{\partial^2 \ln V}{\partial e^2}\right)(e_m, p_m)$$

with an equivalent definition for σ_p. There should generally be a cross correlation in this Gaussian approximation, but this term only adds

$$\left(\frac{\partial^2 \ln V}{\partial e \partial p}\right)(e_m, p_m) = 60 e_m$$

The eigenvalues of the quadratic are only modified by a term $\propto e_m^3$, and we can ignore this term provided x is large.

APPENDIX 4: Density Profiles

To illustrate the methods we use for calculating probabilities, we present details of the derivations required for §7.2. If we denote by $\mathcal{N}_{pk}(\nu, \varsigma, F(r))d\nu d^6\varsigma dF(r)$ the density of maxima with height ν in the range ν to $\nu + d\nu$, ς_A, $A = 1,...,6$ in the range ς_A to $\varsigma_A + d\varsigma_A$ which have the field in the range $F(r)$ to $F(r) + dF(r)$ a distance r away, then the conditional probability that $F(r)$ falls in this range subject to the constraint that there is a peak at $\vec{r} = 0$ of height ν with second derivative matrix ς_A is

$$P_F(F(r)|\nu, \varsigma_A)dF(r) = \frac{\mathcal{N}_{pk}(\nu, \varsigma, F(r))d\nu d^6\varsigma dF(r)}{\mathcal{N}_{pk}(\nu, \varsigma)d\nu d^6\varsigma}. \tag{A4.1}$$

See the discussion in §3 for a justification of this. Here,

$$\mathcal{N}_{pk}(\nu, \varsigma, F(r))d\nu d^6\varsigma dF(r) = |det\ \varsigma|\ P(F(0)/\sigma_0 = \nu, \vec{\eta}(0) = 0, \varsigma, F(r))\ d\nu d^6\varsigma dF(r). \tag{A4.2}$$

There is also a constraint condition that ς be negative definite if the extremum is to be a maximum. The probability distribution appearing here is a multivariate Gaussian. The equation for $\mathcal{N}_{pk}(\nu, \varsigma)d\nu d^6\varsigma$ is similar to A4.2. In particular, when the ratio in equation (A4.1) is taken, the $|det\ \varsigma|$ terms cancel, leaving the constrained probability A4.1 as a ratio of multivariate Gaussian distributions. There is a general theorem which is extremely useful when working with such ratios.

Theorem: If $\vec{Y} = (\vec{Y}_A, \vec{Y}_B)$, $\vec{Y}_A$ and $\vec{Y}_B$ are all Gaussian-distributed, where $\vec{Y}_A$ and $\vec{Y}_B$ are 2 vectors of arbitrary length d_A and d_B, then the conditional probability

$$P(\vec{Y}_B|\vec{Y}_A) = P(\vec{Y}_A, \vec{Y}_B)/P(\vec{Y}_A),$$

is a Gaussian with mean

$$\langle \vec{Y}_B|\vec{Y}_A\rangle = \langle \vec{Y}_B \otimes \vec{Y}_A\rangle \langle \vec{Y}_A \otimes \vec{Y}_A\rangle^{-1} \vec{Y}_A^{\dagger}$$

and covariance matrix

$$\langle \Delta\vec{Y}_B \otimes \Delta\vec{Y}_B | \vec{Y}_A \rangle = \langle \vec{Y}_B \otimes \vec{Y}_B \rangle - \langle \vec{Y}_B \otimes \vec{Y}_A \rangle \langle \vec{Y}_A \otimes \vec{Y}_A \rangle^{-1} \langle \vec{Y}_A \otimes \vec{Y}_B \rangle,$$

where $\Delta\vec{Y}_B \equiv \vec{Y}_B - \langle \vec{Y}_B | \vec{Y}_A \rangle$. (We have assumed both $\langle \vec{Y}_B \rangle$ and $\langle \vec{Y}_A \rangle$ vanish.) The tensor product notation $\vec{Y}_B \otimes \vec{Y}_A$ just makes a $d_B \times d_A$ matrix out of the 2 vectors in the obvious way; $\dagger$ denotes transpose. The proof is straightforward. (See, for example, Adler 1981.)

In our case, $\vec{Y}_B = (F(r))$ is a one-dimensional vector and $\vec{Y}_A$ contains all the other variables $(\nu, \vec{\eta}, \varsigma_A)$. The constraint values of η vanish of course. As in Appendix 1, we find it convenient to introduce the x, y and z variables instead of ς_A, $A = 1,2,3$, for then many of the cross correlations vanish and the correlation matrix of $\vec{Y}_A$ is largely diagonal (see Eq.A1.3). We also choose to work in the principal axes system, so $\varsigma_A = 0$, $A = 4,5,6$. The explicit dependence on the Euler angles disappears. These degrees of freedom were used up in defining the orientation of the axes. The application of the theorem then gives us equation (A4.1) as a Gaussian with mean

$$\begin{aligned} \langle F(r) | C \rangle = & \frac{\nu}{(1-\gamma^2)} \Big(\langle F\nu \rangle - \gamma \langle Fx \rangle \Big) + \\ & \frac{x}{(1-\gamma^2)} \Big(\langle Fx \rangle - \gamma \langle F\nu \rangle \Big) + 15\langle Fy \rangle y + 5\langle Fz \rangle z \end{aligned} \tag{A4.3}$$

and variance

$$\begin{aligned} \langle (\Delta F(r))^2 | C \rangle = & \sigma_0^2 - \frac{\langle F\nu \rangle^2}{1-\gamma^2} + 2\frac{\gamma}{1-\gamma^2} \langle F\nu \rangle \langle Fx \rangle \\ & - \frac{\langle Fx \rangle^2}{1-\gamma^2} - 15\langle Fy \rangle^2 - 5\langle Fz \rangle^2 - \sum_{A=4}^{6} \frac{15}{\sigma_2^2} \langle F\varsigma_A \rangle^2 - \sum_{i=1}^{3} \frac{3}{\sigma_1^2} \langle F\eta_i \rangle^2. \end{aligned} \tag{A4.4}$$

The identity

$$3\sum_i a_i b_i - \sum_i a_i \sum_i b_i = 6\frac{(a_1 - a_3)}{2}\frac{(b_1 - b_3)}{2} + 2\frac{(a_1 - 2a_2 + a_3)}{2}\frac{(b_1 - 2b_2 + b_3)}{2},$$

where a_i and b_1 are 2 vectors, is useful in transforming back and forth between λ_i and x, y, z. It aids in proving the result

$$15\langle Fy \rangle^2 + 5\langle Fz \rangle^2 + \sum_{A=4}^{6} \frac{15}{\sigma_2^2} \langle F\varsigma_A \rangle^2 = \frac{15}{2\sigma_2^2} \Big(2(\xi'/r)^2 + (\xi'')^2 \Big) - \frac{5}{2} \langle Fx \rangle^2.$$

Here, $\xi = \langle F(r)F(0) \rangle \equiv \sigma_0^2 \psi(r)$ is the correlation function. Also, $\sum \langle F\eta_i \rangle^2 = (\xi')^2$ is required for the variance.

The mean and variance are therefore given by equations (7.8) and (7.9). Notice that the dispersion does not depend upon the parameters characterizing the peak. Also, it only depends upon the magnitude of r. Far away from the point r, all derivative terms die quickly, giving the asymptotics

$$\langle F(r) | C \rangle \to \frac{\nu + x\gamma}{(1-\gamma^2)} \psi, \quad \langle (\Delta F(r))^2 | C \rangle / \sigma_0^2 \to 1 - \frac{\psi^2}{1-\gamma^2} \; as \; r \to \infty. \tag{A4.5}$$

Thus, a long distance away from the peak the asymmetry is forgotten, and the average profile is proportional to the density correlation function. This behavior is only due to the outer regions becoming uncorrelated with the inner. This is reflected in the large dispersion beyond a few filtering

radii (Fig. 7.3). An expansion of equation (7.8) about $r = 0$ yields the triaxial result, equation (7.4), to $O(r^4)$.

If we form the conditional probability for the density at $F(r)$ and the Euler angles defining the principal axes orientation, $P(F(r), \alpha, \beta, \gamma | \nu, x, e, p)$, then this takes precisely the same form as equation (7.8), since the distribution of the Euler angles is simply that appropriate to all orientations being equally likely (equation A1.6). If we now form the conditional average over this distribution, the average over the Euler angle distribution just amounts to an integration over θ and ϕ, which results in the A-term integrating to zero. Therefore, we only need ν and κ to specify this mean which is orientation-averaged and thus spherically symmetric:

$$\langle F(r)|\nu, x, e, p\rangle \equiv \langle F(r)|\nu, x\rangle = \frac{\nu}{(1-\gamma^2)}(\psi + \gamma^2\nabla^2\psi/3) - \frac{x\gamma}{(1-\gamma^2)}(\psi + \nabla^2\psi/3) \qquad \text{(A4.6)}$$

The dispersion equation (7.9) is already orientation-independent, so this is the result upon averaging over the Euler angles.

The distribution

$$P(F(r)|\nu,\ peak) = \int P(F(r), x, \alpha, \beta, \gamma|\nu,\ peak) dx \sin\beta d\beta d\alpha d\gamma$$

is a non-Gaussian one, involving an x-integration. Nonetheless, it is easy to show that the dispersion is still given by equation (7.9) and the mean by equation (A4.6) with $\langle x|\nu,\ peak\rangle$, the average of x subject to the constraint of being at a peak of height ν (equation 6.13), replacing x in equation (A4.6).

APPENDIX 5: Peaks in Background Fields

In this appendix, we derive in §A5.1 the probability distribution for the background field to have a value F_b at a given point subject to the constraint that there is a peak at that point in the field smoothed on some smaller scale, F_s. In §A5.2, we obtain $\mathcal{N}_{pk}(\nu_s|\nu_b)d\nu_s$, the number density of peaks in the F_s field of height between ν_s and $\nu_s + d\nu_s$ subject to the constraint that the background field has height $F_b = \nu_b\sigma_{0b}$ at the peak point.

§A5.1: CALCULATION OF $P(\nu_b|\nu_s)$

According to the theorem given in Appendix 4, the probability that the background field has height ν_b subject to the constraint that there is a peak characterized by the data $\mathcal{C}$ of §7.2 and Appendix 4 is a *Gaussian*. The mean and dispersion can be calculated in the same way that the profile was calculated. Since the correlations of $F_b(0)$ with η_i, y, z and ς_A, $A = 4,5,6$ all vanish, these are especially simple:

$$\begin{aligned}
\bar{\nu}_b &\equiv \langle \nu_b|\mathcal{C}\rangle = \langle \nu_b|\nu_s,\ x_s\rangle \\
&= \frac{\epsilon}{1-\gamma_s^2}\left(\nu_s(1-\gamma_s^2 r_1) - \gamma_s x(1-r_1)\right) \\
(\Delta\nu_b)^2 &\equiv \langle(\Delta\nu_b)^2|\mathcal{C}\rangle = \langle(\Delta\nu_b)^2|\nu_s,\ x_s\rangle \\
&= 1 - \frac{\epsilon^2}{1-\gamma_s^2}(1 - 2\gamma_s^2 r_1 + \gamma_s^2 r_1^2).
\end{aligned} \tag{A5.1}$$

We have introduced the following notation for various spectral averages:

$$\begin{aligned}
\epsilon &\equiv \langle \nu_s\nu_b\rangle = \frac{\sigma_{0h}^2}{\sigma_{0s}\sigma_{0b}} \\
r_1 &\equiv \frac{\langle k^2\rangle_h}{\langle k^2\rangle_s} = \frac{\sigma_{1h}^2\sigma_{0s}^2}{\sigma_{0h}^2\sigma_{1s}^2} \\
\sigma_{jh}^2 &\equiv \int \frac{d^3k}{(2\pi)^3}\, k^{2j} C_b(k) C_s(k) \langle |F(k)|^2\rangle, \quad j = 0,1,2.
\end{aligned} \tag{A5.2}$$

Quantities with the subscript s and b are defined as in §4.2. C_s and C_b denote general filtering functions defining the F_s and F_b fields.

If both C_s and C_b are Gaussian then the quantities in equation (A5.2) can be determined from equation (4.10). In that case, the F_b, F_s cross correlations are also given by equation (4.10) with the filtering scale being the *rms* average $R_h \equiv \left((R_b^2 + R_s^2)/2\right)^{1/2}$. If we also assume that the unfiltered spectrum is a power law, the parameters which enter into this expression are

$$\begin{aligned}
\epsilon &= \left(\frac{R_s R_b}{R_h^2}\right)^{\frac{n+3}{2}} \to (2R_s/R_b)^{\frac{n+3}{2}} \\
r_1 &= (R_s/R_h)^2 \to 2(R_s/R_b)^2 \\
\gamma_s &= \left(\frac{n+3}{n+5}\right)^{1/2}.
\end{aligned} \tag{A5.3}$$

The limits are valid for $R_s \ll R_b$. In that case, the dispersion deviates from unity only by a term of order ϵ^2 and the mean $\bar{\nu}_b \to \epsilon\nu_s$ with dispersion 1 for high peaks. This average should be compared with the value of ν_b obtained if the one peak of height ν_s on scale R_s is just smoothed to the scale R_b, for that contributes a term $\sim (R_s/R_b)^3(\sigma_{0s}/\sigma_{0b})\nu_s \sim (R_s/R_b)^{(3-n)/2}\nu_s$. For steep spectra ($n > 0$) this may be larger than $\bar{\nu}_b \sim (R_s/R_b)^{(3+n)/2}\nu_s$, whereas for shallow spectra ($n < 0$) with a significant amount of power between the two smoothing scales, it is smaller.

It is not surprising that this conditional probability is independent of the orientation of the principal axes and the asymmetry of the peak. Thus, the conditional probability subject only to the constraint of peaks with fixed height ν_s and curvature parameter x is

$$P(\nu_b|\nu_s,\ x_s) \equiv \frac{\mathcal{N}_{pk}(\nu_b,\nu_s,x_s)}{\mathcal{N}_{pk}(\nu_s,x_s)} = \frac{1}{\sqrt{2\pi}\Delta\nu_b}\ exp(-\frac{(\nu_b-\bar{\nu}_b)^2}{2(\Delta\nu_b)^2}). \tag{A5.4}$$

We can use equation (A5.4) to obtain the joint differential number density of points which are peaks of the smoothed field of height ν_s and have the background height ν_b:

$$\mathcal{N}_{pk}(\nu_b,\nu_s) = \int_0^\infty dx\ P(\nu_b|\nu_s,\ x_s)\ \mathcal{N}_{pk}(\nu_s,x_s) = \frac{G(\tilde{\gamma},\tilde{\gamma}\tilde{\nu})}{(2\pi)^2 R_{*s}^3}\ \frac{exp\left(-\frac{(\nu_s-\epsilon\nu_b)^2}{2(1-\epsilon^2)}\right)}{(1-\epsilon^2)^{1/2}}\ \frac{e^{-\nu_b^2/2}}{\sqrt{2\pi}}. \tag{A5.5}$$

Here,

$$\tilde{\nu} = \frac{\gamma_s}{\tilde{\gamma}}\frac{(1-r_1)}{(1-\epsilon^2)}\left(\nu_s\frac{(1-\epsilon^2 r_1)}{(1-r_1)}\ -\ \epsilon\nu_b\right),$$
$$\tilde{\gamma}^2 = \gamma_s^2\left(1+\epsilon^2\frac{(1-r_1)^2}{(1-\epsilon^2)}\right). \tag{A5.6}$$

The *tilde* variables are introduced only for convenience in expressing the result. In the limit in which the background scale is sufficiently larger than the smoothing scale R_s that $r_1 \approx 0$, these parameters become

$$\tilde{\nu} \to \nu_p,$$
$$\tilde{\gamma} \to \gamma_p \quad \text{as } r_1 \to 0. \tag{A5.7a}$$

where we have defined the 'peak-field' parameters

$$\nu_p \equiv (\nu_s - \epsilon\nu_b)/(1-\epsilon^2)^{1/2}$$
$$\gamma_p \equiv \gamma_s/(1-\epsilon^2)^{1/2}. \tag{A5.7b}$$

We have not required that ϵ be small. For most spectra, r_1 falls off faster than ϵ (Eq.A5.3). In this limit, the joint density is simply

$$\mathcal{N}_{pk}(\nu_b,\nu_s)\ d\nu_b d\nu_s \to \mathcal{N}_{pk}(\nu_p;\gamma_p)d\nu_p\ P(\nu_b)d\nu_b. \tag{A5.7c}$$

Recall that F_b is Gaussian-distributed:

$$P(\nu_b) = exp(-\nu_b^2/2)/\sqrt{2\pi}. \tag{A5.7d}$$

Equation A5.7c is usually an excellent approximation, for it is only in G that the deviations occur. The relations between the *tilde* and peak arguments of this function can be used to estimate when the full formula A5.5 should be used:

$$\tilde{\gamma}^2/\gamma_p^2 = 1 - 2\epsilon^2 r_1(1-r_1/2)$$
$$\tilde{\gamma}\tilde{\nu} = \gamma_p\nu_p(1-\epsilon^2 r_1) + \gamma_s\epsilon r_1\nu_b. \tag{A5.7e}$$

We can form the conditional probability that the background field has height ν_b given that the peak has height ν_s:

$$\begin{aligned} P(\nu_b|\nu_s)d\nu_b &\equiv \frac{\mathcal{N}_{pk}(\nu_b,\nu_s)}{\mathcal{N}_{pk}(\nu_s)}d\nu_b \\ &= \frac{exp\left(-\frac{(\nu_b-\epsilon\nu_s)^2}{2(1-\epsilon^2)}\right)}{(2\pi(1-\epsilon^2))^{1/2}}\ d\nu_b\ \frac{G(\tilde{\gamma},\tilde{\gamma}\tilde{\nu})}{G(\gamma_s,\gamma_s\nu_s)}. \end{aligned} \tag{A5.8}$$

In the $R_s \ll R_b$ limit, this becomes

$$P(\nu_b|\nu_s)d\nu_b \approx P(\nu_b)d\nu_b\mathcal{N}_{pk}(\nu_s-\epsilon\nu_b)/\mathcal{N}_{pk}(\nu_s). \tag{A5.9}$$

The constrained probability that the background field has height ν_b given that the smoothed field has height ν_s if the point is *not* required to be a peak of the F_s-field is

$$P(\nu_b|\nu_s,\ no\ peak\ )d\nu_b = \frac{1}{(2\pi(1-\epsilon^2))^{1/2}}\ exp\left(-\frac{(\nu_b-\epsilon\nu_s)^2}{2(1-\epsilon^2)}\right)\ d\nu_b. \tag{A5.10}$$

This yields the average $\langle\nu_b|\nu_s,\ no\ peak\ \rangle = \epsilon\nu_s$, the high ν_s limit of the result with peaks. Equation (A5.10) has been used by Blumenthal, Faber and Primack (1985) with top hat filtering. In the limit that ν_s is high, the deviation from equation (A5.8) is small.

§A5.2: CALCULATION OF $\mathcal{N}_{pk}(\nu_s|\nu_b)$

The number density of peaks in the F_s field at points with the background field of some specified height also follows from the joint density A5.5:

$$\begin{aligned} \mathcal{N}_{pk}(\nu_s|\nu_b)d\nu_s &\equiv \frac{\mathcal{N}_{pk}(\nu_b,\nu_s)}{P(\nu_b)}d\nu_s \\ &= \frac{G(\tilde{\gamma},\tilde{\gamma}\tilde{\nu})}{(2\pi)^2R_{*s}^3}e^{-\nu_p^2/2}\ d\nu_p. \end{aligned} \tag{A5.11}$$

In the $R_s \ll R_b$ limit, this becomes

$$\mathcal{N}_{pk}(\nu_s|\nu_b) \approx \mathcal{N}_{pk}(\nu_p)/(1-\epsilon^2)^{1/2}. \tag{A5.12}$$

Notice that $\nu_p \equiv \left(\nu_s - \langle\nu_s\nu_b\rangle\nu_b\right)\left(1-\langle\nu_s\nu_b\rangle^2\right)^{-1/2}$ is uncorrelated with ν_b, and is *Gaussian-distributed* with variance 1 and zero mean. The following simple approximation, valid in the limit $R_s \ll R_b$, is useful: $\mathcal{N}_{pk}(\nu_s|\nu_b)\ d\nu_s \approx \mathcal{N}_{pk}(\nu_p)\ d\nu_p$ with γ_p and R_{*p} evaluated from the power spectrum of ν_p in the differential density.

If we drop the restriction that the F_s point be a peak, then the conditional probability that its height be ν_s given that the height of F_b is ν_b is identical to equation (A5.10) if s and b are interchanged:

$$\begin{aligned} P(\nu_s,\ no\ peak\ |\nu_b)d\nu_s &= \frac{1}{(2\pi(1-\epsilon^2))^{1/2}}\ exp\left(-\frac{(\nu_s-\epsilon\nu_b)^2}{2(1-\epsilon^2)}\right)\ d\nu_s \\ &= \frac{1}{\sqrt{2\pi}}\ e^{-\nu_p^2/2}\ d\nu_p. \end{aligned} \tag{A5.13}$$

Once again, it is the field ν_p which enters.

APPENDIX 6: Asymptotic Peak Correlation Functions

In this appendix, we determine the n-point correlation function of peaks in the limit that terms involving gradients of $\psi(r)$ can be neglected in comparison with terms involving ψ alone. Since $d^n\psi/dr^n$ falls off a factor of r^{-n} faster than ψ for large r, this approximation should prove accurate in the small ψ regime. These asymptotic results are based on work in progress by Bardeen, Bond, Jensen and Szalay (1985) where the 2-point function of peaks is calculated for all r, and small and large distance expansions are given.

We label the positions of the peak points by r_i, $i = 1, ..., n$. The (unreduced) n-point correlation function of the peak densitites is

$$\langle \prod_i n_{pk}(i) \rangle = \langle \prod_i \delta(\nu(i) - F(r_i)/\sigma_0)\ |\lambda_1(i)\lambda_2(i)\lambda_3(i)|\ \theta(\lambda_3(i))\ \delta(\vec{\eta}(i)) \rangle. \tag{A6.1}$$

To evaluate this, we need to know the joint Gaussian probability density for $10n$ variables

$$P(\{\nu(i), \vec{\eta}(i), \varsigma_A(i), A = 1, ..., 6 | i = 1, ...n\}).$$

Generally, this requires knowledge of $\langle \nu(i)\nu(j) \rangle$, $\langle \nu(i)\vec{\eta}(j) \rangle$, $\langle \nu(i)\varsigma_A(j) \rangle$ and the $\eta - \varsigma$ cross correlations for $i \neq j$. Since all but $\langle \nu(i)\nu(j) \rangle \equiv \psi(ij) \equiv \psi(|\vec{r}_i - \vec{r}_j|)$ involve derivatives of ψ, we need only retain this spatial correlation connecting i to j.

At each point, we rotate to the principal axes and use the $x(i), y(i), z(i)$ variables of Appendix 1. The correlation matrix between all $10n$ variables reduces to two pieces: (1) The $8n \times 8n$ matrix involving the $\vec{\eta}(i), y(i), z(i)$ and $\varsigma_A(i), A = 4, 5, 6$ variables is diagonal and independent of ψ. The corresponding contribution to P is the product $\prod_i P(\vec{\eta}(i), y(i), z(i), \varsigma_A(i))$ of independent terms. We can therefore immediately integrate over these $8n$ variables:

$$\langle \prod_i n_{pk}(r_i, \nu(i), x(i)) \rangle = \langle \prod_i \left(\frac{f(x(i))}{(2\pi)^{3/2} R_*^3} \right) \rangle. \tag{A6.2}$$

The average is now over the remaining $2n$ variables $\{(\nu(i), x(i))\}$; $f(x)$ is given by equation (A1.15). (2) To evaluate this average, we need the $2n \times 2n$ correlation matrix in these variables. It is convenient to consider the matrix as a tensor product of a 2×2 and an $n \times n$ matrix. The tensor product of A and B, $m_A \times n_A$ and $m_B \times n_B$ matrices respectively, is an $m_A m_B \times n_A n_B$ matrix $A \otimes B$. We make use of the properties

$$(A \otimes B)(C \otimes D) = AB \otimes CD,\ (A \otimes B)^{-1} = A^{-1} \otimes B^{-1},\ (A \otimes B)^\dagger = A^\dagger \otimes B^\dagger,\ TrA \otimes B = TrA \cdot TrB. \tag{A6.3}$$

The correlation matrix M can be written

$$M = A \otimes I + u \otimes \Psi, \tag{A6.4}$$

where the (ν, x) correlation matrix is

$$A = \begin{pmatrix} 1 & \gamma \\ \gamma & 1 \end{pmatrix}, \quad u = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \tag{A6.5}$$

I is the $n \times n$ identity matrix, and $\Psi = \left(\psi(ij)(1 - \delta_{ij})\right)$ is an $n \times n$ matrix with components $\psi(ij)$ if $i \neq j$ and 0 if $i = j$.

The inverse of M is required for $P(\{\nu(i), x(i)\})$:

$$M^{-1} = A^{-1} \otimes I - A^{-1} u A^{-1} \otimes \mathbf{C}, \tag{A6.6}$$

where

$$\mathbf{C} = \Psi\ (1 + \Psi/(1-\gamma^2))^{-1}, \tag{A6.7}$$

and

$$A^{-1} = \frac{1}{1-\gamma^2}\begin{pmatrix} 1 & -\gamma \\ -\gamma & 1 \end{pmatrix}, \quad A^{-1}uA^{-1} = \frac{1}{(1-\gamma^2)^2}\begin{pmatrix} 1 & -\gamma \\ -\gamma & \gamma^2 \end{pmatrix}. \tag{A6.8a}$$

The properties A6.3 and the relation

$$(A^{-1}u)^r = A^{-1}u/(1-\gamma^2)^{r-1},\ \textit{for } r \textit{ an integer} \tag{A6.8b}$$

have been used.

The quadratic $2Q = y^\dagger M^{-1} y$, where $y = \sum_i (\nu(i), x(i))^\dagger \otimes e_i^\dagger$ is a $2n$ dimensional vector. Here, e_i denotes the n- dimensional row vector with 1 in the i^{th} position and 0 elsewhere. For $\psi = 0$ the quadratic reduces to a sum of independent terms

$$Q_0 = \sum_i (\nu(i)\ x(i)) A^{-1} (\nu(i)\ x(i))^\dagger = \sum_i \left(\nu(i)^2/2 + \frac{(x(i) - x_*(i))^2}{2(1-\gamma^2)} \right), \tag{A6.9}$$

leading to a product of n independent $\nu(i) - x(i)$ probabilities,

$$P(\nu, x) d\nu dx = \frac{exp(-\nu^2/2)}{\sqrt{2\pi}} \frac{exp(-\frac{(x-x_*)^2}{2(1-\gamma^2)})}{(2\pi(1-\gamma^2))^{1/2}}\ d\nu dx, \tag{A6.10}$$

a combination which has appeared repeatedly throughout the paper. The full quadratic can be written

$$Q = Q_0 - \frac{1}{2} \sum_{ij} C_{ij} \tilde{\nu}_i \tilde{\nu}_j, \tag{A6.11a}$$

where

$$\tilde{\nu}_i \equiv (\nu_i - \gamma x(i))/(1-\gamma^2). \tag{A6.11b}$$

The joint probability also involves

$$det\ M = det\ (A \otimes I)\ det\ (I + \Psi/(1-\gamma^2)). \tag{A6.12}$$

The unreduced *differential* correlation function for fixed $\nu(i)$ and curvature parameters $x(i)$ follows:

$$\langle \prod_i n_{pk}(r_i, \nu(i), x(i)) \rangle = \prod_i \left(P(\nu(i), x(i)) \frac{f(x(i))}{(2\pi)^{3/2} R_*^3} \right) \frac{exp(\tilde{\nu}^\dagger \mathbf{C} \tilde{\nu}/2)}{(det(I + \Psi/(1-\gamma^2)))^{1/2}}. \tag{A6.13}$$

The next step is to integrate equation (A6.13) over the $x(i)$. The integrals would factor if it were not for the cross terms $\tilde{\nu}(i) C_{ij} \tilde{\nu}(j)/2$, $i \neq j$ in the exponents. One way around this is to replace $\tilde{\nu}(i)$ by

$$\tilde{\nu}(i) = \hat{\nu}(i) - \gamma y(i)/(1-\gamma^2), \tag{A6.14}$$

where

$$\hat{\nu}(i) = (\nu(i) - \gamma \langle x(i) \rangle)/(1-\gamma^2), \quad y(i) = x(i) - \langle x(i) \rangle. \tag{A6.15}$$

Write

$$Q = Q' - \frac{1}{2} \sum_{i \neq j} \gamma^2 C_{ij} y(i) y(j)/(1-\gamma^2)^2. \tag{A6.16}$$

Define $\langle x(i)\rangle$ to be the average of $x(i)$ weighted by the integrand with Q replaced by Q'. This modified integrand does factor into terms of the form

$$f(x(i))\ exp[-\frac{1}{2}(x(i)-x_*(i))^2/(1-\hat{\gamma}(i)^2)], \tag{A6.17}$$

with

$$1-\hat{\gamma}(i)^2 \equiv (1-\gamma^2)/(1-\gamma^2 C_{ii}/(1-\gamma^2)) \tag{A6.18}$$

and, in matrix notation,

$$x_* = [I+\Psi]^{-1}\ [\gamma\nu + \frac{\gamma^2}{1-\gamma^2}(I+\Psi/(1-\gamma^2))C\Theta], \tag{A6.19}$$

where Θ is a vector with components

$$\theta(i) \equiv \langle x(i)\rangle - x_*(i). \tag{A6.20}$$

The interpolation formula given in equation (6.14) can be used to evaluate $\theta(i)$ if $\gamma(i)$ replaces γ and $x_*(i)$ replaces $\gamma\nu$. Equation (A6.19) is then an implicit equation for the $x_*(i)$ which must be solved by iteration, though the iteration process does converge fairly rapidly.

Since the integrand is fairly sharply peaked about $\langle x(i)\rangle$ as a function of $x(i)$, the error in replacing Q by Q' is of order $[\gamma^2\psi/(1-\gamma^2)^2]^2$, as can be seen by expanding $exp[-\frac{1}{2}(Q-Q')]$ in a power series. The error is small unless some of the ψ_{ij} are close to one, regardless of how large the $\nu(i)$ are.

Reexpressing Q' in terms of the $x(i)$ gives, in matrix notation,

$$\begin{aligned} Q' = &\frac{1}{2}\sum_i [x(i)-x_*(i)]^2/[1-\gamma^2] + \frac{1}{2}\nu^\dagger[I+\Psi]^{-1}\nu \\ &+ \frac{1}{2}\frac{\gamma^2}{(1-\gamma^2)^2}\Theta^\dagger[\tilde{\mathbf{C}} + \frac{\gamma^2}{(1-\gamma^2)}\tilde{\mathbf{C}}(I+\Psi/(1-\gamma^2))(I+\Psi)^{-1}\tilde{\mathbf{C}}]\ \Theta\ , \end{aligned} \tag{A6.21}$$

where $\tilde{\mathbf{C}}$ is the matrix containing just the *off-diagonal* components of $\mathbf{C}$. The integrals over the $x(i)$ have now just the same form as in the number density integral, with $\gamma\nu(i)$ replaced by $x_*(i)$ and γ replaced by $\tilde{\gamma}(i)$, so they can be expressed in terms of the function $G(\nu(i),\tilde{\gamma}(i),x_*(i))$, for which we have the accurate interpolation formula given by equations (4.4) and (4.5).

The procedure outlined above is quite accurate as long as $\psi_{ij} < 1/2$ (or for even larger ψ_{ij} if $\gamma < 0.5$), but is obviously rather cumbersome, For this reason, the Gaussian approximation to the integrand in the $x(i)$ is the recommended procedure in §6.2. Comparing the approximation schemes offers cross-checks on their accuracy. A straightforward Gaussian approximation using x_m gives a result for the two-point correlation amplitude $\xi^{(2)}_{pk}$ about 4% larger than the result of the above procedure when ψ_{12} is small and about 10% larger when $\psi_{12} \approx 1/2$ (for $\gamma = 0.555$). The fudge of replacing x_m by $\langle x\rangle$ reduces the discrepency to about 1% and 2.5%, respectively. Of course, $\psi_{12} \approx 1/2$ is also where the approximation of neglecting derivatives of ψ is beginning to break down. The accuracy of neglecting the $y(i)y(j)$ cross-terms can be checked by dropping these terms in the Gaussian approximation. This lowers $\xi^{(2)}_{pk}$ by about 4% at $\psi_{12} \approx 1/2$, but the discrepency decreases rapidly (as ψ^2_{12}) for smaller ψ_{12}. Dropping the derivatives of ψ probably raises $\xi^{(2)}_{pk}$, since the exact $\xi^{(2)}_{pk}$ should become negative once $r_{12}/R_s < 1$, due to the smoothing.

Kaiser (1984) introduced the effective field $\theta(\nu - F(r)/\sigma_0)$ which describes those regions with densities in excess of the height ν. Politzer and Wise (1984) calculated the n-point correlation function

of this field. An equivalent result can be obtained very concisely by considering instead the effective field

$$D_\nu(r) \equiv \delta(\nu - F(r)/\sigma_0) \tag{A6.22}$$

which is non-vanishing only on the contour surfaces $F(r) = \nu\sigma_0$ used in §4.4 where the Euler characteristic of these surfaces per unit volume is given. At low ν, these surfaces are multiply connected, and the correlation function of D_ν will tell us nothing about peak clustering. However, at high ν the contour surfaces split up into disconnected 'bags', each surrounding a peak of height in excess of ν. For distances large compared with the scale of an individual bag, the correlation function should then equal that for peaks. Since the n-point (unreduced) correlation function of D_ν is *exactly* the joint distribution for $F(r_1), ..., F(r_n)$, the calculation is trivial:

$$\begin{aligned} 1 + \xi_n^D(r_1, \nu(1), ..., r_n, \nu(n)) &\equiv \frac{\langle \prod_i D_{\nu(i)}(r_i) \rangle}{\prod_i \langle D_{\nu(i)}(r_i) \rangle} = \frac{P(\{F(r_i) = \nu(i)\sigma_0,\ 1 = 1, ..., n\})}{\prod_i P(F(r_i) = \nu(i)\sigma_0)} \\ &= \frac{exp(\nu^\dagger \Psi (I + \Psi)^{-1} \nu / 2)}{(\ det\ (I + \Psi)\ )^{1/2}} \\ &\to exp(\nu^\dagger \Psi \nu / 2)\ as\ \psi \to 0. \end{aligned} \tag{A6.23}$$

In the high ν limit, the $\theta(\nu - F(r)/\sigma_0)$ field also gives this result, as Politzer and Wise demonstrated. The degree to which it fails to provide a description of the clustering of peaks is the degree to which $\hat{\nu}$ differs from ν. In most of the cosmologically interesting regime, the deviation is substantial.

APPENDIX 7: Power Spectra and Dark Matter

The only information required to specify our homogeneous isotropic Gaussian random fields is their power spectra. We define the transfer function for linear perturbations which takes the initial Fourier components of the density $F(k,t_i)$ at some very early time t_i to the final ones $F(k,t_f)$ at some late time t_f by

$$T(k,t_f) \equiv \frac{b(t_i)}{b(t_f)} \frac{F(k,t_f)}{F(k,t_i)}, \tag{A7.1}$$

where $b(t)$ specifies the linear growth law of long wavelength perturbations. The difficulty with this definition of the transfer function is that it depends upon the choice of gauge and the choice of hypersurface upon which the density fluctuations are measured. Gauge-invariant quantities can easily be constructed to alleviate the first problem. In equations (A7.9) and (A7.10) below, we give a hypersurface-independent way of treating the transfer function which we feel is superior to the usual definition, equation (A7.1). The density perturbations should be taken to be those defined, for example, in the popular synchronous gauge (Peebles 1980) referred to some specific choice of synchronous hypersurfaces. Referring the density fluctuations to comoving hypersurfaces would give the same result for T. In the matter-dominated regime in an Einstein-de Sitter universe, $b = a$, where $a(t)$ is the expansion factor. For most cosmological models, T is approximately time independent below a redshift $z \sim 100$. The power spectrum of density fluctuations evolves according to:

$$P(k,t_f) = \left(\frac{b(t_i)}{b(t_f)}\right)^2 T^2(k,t_f)\ P(k,t_i). \tag{A7.2}$$

The following fitting formulas drawn from the work of Bardeen (1984), Bond and Szalay (1983), Bond and Efstathiou (1984) and Efstathiou and Bond (1985) accurately reproduce the transfer functions at late times for universes dominated by collisionless relics of the Big Bang in which the baryon density Ω_B is much smaller than the density of dark matter Ω_X.

(1) Cold Dark Matter, Adiabatic Fluctuations:

$$\begin{aligned} T_{CDM,ad,X}(k) &= \frac{ln\ (1+2.34q)}{2.34q}(1+3.89q+(16.1q)^2+(5.46q)^3+(6.71q)^4)^{-1/4}, \\ q &\equiv k\theta^{1/2}/(\Omega_X h^2\ Mpc^{-1}). \end{aligned} \tag{A7.3}$$

Here, $\theta = \rho_{cr}/(1.68\rho_\gamma)$ is a measure of the ratio of the energy density in relativistic particles (photons plus neutrinos) to that in photons. $\theta = 1$ corresponds to 3 flavors of relativistic neutrinos plus the photons. Other CDM fits appropriate to the case when Ω_B is not small are given in Bond and Efstathiou (1984). (See also equation (A7.6) below.) The transfer function of the baryons is 'filtered' on small scales due to the finite baryon pressure around recombination:

$$T_{CDM,ad,B}(k) = T_{CDM,ad,X}(k)(1+(kR_{Jr})^2/2)^{-1}, \quad R_{Jr} = 1.6(\Omega_X h^2)^{-1/2}\ kpc. \tag{A7.4}$$

(2) Massive Neutrinos (one species), Adiabatic Fluctuations:

$$\begin{aligned} T_{\nu,ad}(k) &= exp(-0.16(kR_{f\nu}) - (kR_{f\nu})^2/2)(1+1.6q+(4.0q)^{3/2}+(0.92q)^2)^{-1}, \\ q &\equiv k/(\Omega_\nu h^2\ Mpc^{-1}), \\ R_{f\nu} &= 2.6(\Omega_\nu h^2)^{-1} Mpc \\ T_{\nu,ad}(k) &\approx exp(-3.9q - 2.1q^2). \end{aligned} \tag{A7.5}$$

The first expression gives an exponential damping factor times the cold dark matter spectrum appropriate to a universe with 2 species of relativistic neutrinos ($\theta = 0.87$). $R_{f\nu}$ is a characteristic damping

length for neutrinos. The damping is dominated by the Gaussian part of the filter. If we define a characteristic Gaussian filtering length by the radius at which the filtering function drops to $1/e$, then this radius is $1.1R_{f\nu}$. In §7.3, we just use $R_{f\nu}$ which gives a fairer estimate of the short wavelength falloff. The second expression is a straight fit to the results. Note that $\Omega_\nu h^2 = 0.31 m_\nu/(30\ eV)$.

(3) Warm Dark Matter, Adiabatic Fluctuations:

$$\begin{aligned} T_{warm,ad}(k) &\approx exp(-kR_{fw}/2 - (kR_{fw})^2/2) T_{CDM,ad,X}(k), \\ T_{CDM,ad,X}(k) &= (1 + 1.7q + (4.3q)^{3/2} + q^2)^{-1}, \\ q &\equiv k/(\Omega_X h^2\ Mpc^{-1}), \\ R_{fw} &= 0.2(\frac{g_{Xdec}}{100})^{-4/3}(\Omega_X h^2)^{-1}. \end{aligned} \tag{A7.6}$$

Here, g_{Xdec} is the effective number of particle degrees of freedom when the X-particles decoupled; values in the range $60 - 300$ are typical of minimal grand unified theories over the range of decoupling temperature $T \sim 1$ GeV $- 10^{18}$ GeV. R_{fw} involves the free-streaming length for the warm dark matter as in the neutrino case. Note that $\Omega_X h^2 = 1.0(g_{Xdec}/100)^{-1}(m_X/keV)$ for warm dark matter. The same form for the exponential as in the massive neutrino case was chosen for simplicity, at the expense of goodness of fit. Notice that the damping is effectively gentler than in the neutrino case. The $1/e$ Gaussian filtering length is $1.9R_{fw}$, but this gives far too much damping on larger scales. The exponential damping term multiplies the $\theta = 1$ cold particle spectrum. The CDM transfer function given in A7.6 provides an alternative simpler fit to the CDM transfer function, though it is not quite as accurate as equation (A7.3). In both the neutrino and warm dark matter cases, the baryons have the same power spectrum as the collisionless relics.

To get the final power spectra, the initial spectrum is required. This is usually assumed to be a power law with index n. In particular, the index $n = 1$ defines the Zeldovich spectrum that arises in inflation. The asymptotic form for the output power spectrum of cold dark matter density fluctuations is therefore $\sim Ak$ for small k and $\sim Ak^{-3}(ln(k))^2/16$ for large k, with A a normalization amplitude. Hot (massive neutrinos) and warm dark matter have almost no power on small scales due to collisionless damping.

(4) Cold Dark Matter, Isocurvature Fluctuations:

$$\begin{aligned} T_{CDM,isoc,X}(k) &= (5.6q)^2\left(1 + \frac{(40q)^2}{1 + 215q + (16q)^2(1 + 0.5q)^{-1}} + (5.6q)^{8/5}\right)^{-5/4} \\ &= (5.6q)^2(1 + (15q + (0.86q)^{3/2} + (5.6q)^2)^{1.24})^{-0.807}, \\ q &\equiv k/(\Omega_X h^2\ Mpc^{-1}). \end{aligned} \tag{A7.7}$$

Two excellent fits are given. The asymptotic behavior, $\rightarrow (5.6k)^2$ and $\rightarrow 1$ for large and small k respectively, differs considerably from that appropriate to the adiabatic case. However, for the isocurvature axion mode, an $n = -3$ initial spectrum is appropriate, which yields the $\sim k$ and $\sim k^{-3}$ limits for the evolved power spectrum as for the $n = 1$ adiabatic case. If primordial black holes form the dark matter, $n = 0$ might be the appropriate initial spectral index ($\xi = 0$ for $r \neq 0$). This implies little power ($\sim k^4$) on large scales.

In addition, an overall amplitude must be specified to normalize these spectra. In §6.6, we discussed two ways to do this using the observed galaxy correlation function and the galaxy correlation function we predict using the threshold hypothesis. The first is simply to normalize ξ to unity at $r_0 = 5\ h^{-1}Mpc$. The second is to use $J_3(r)$ normalization for $r \gg r_0$ using the spatially uniform biasing factor b. Another standard procedure would be to require that the power spectrum top-hat-filtered on scale R_{TH} satisfy $\sigma_0(R_{TH} = 8\ h^{-1}Mpc, t_0) = b^{-1}(t_0)$, since this is the scale when the *rms* fluctuations in the number of bright galaxies equals unity. Deviations from the simple b^2 proportionality on scales which contribute significantly to $\langle(\Delta N_{gal}(R_{TH})/N_{gal})^2\rangle$ imply that J_3 normalization is preferred.

Another way to characterize this overall amplitude is to relate the final density to an initial measure of the perturbation amplitude (so that t_f in A7.2 is t_0, the age now). In this case, it is useful to specify the initial amplitude in terms of quantities that are constant when the perturbation is outside the horizon. For adiabatic perturbations,

$$\varsigma = 3\phi + \delta\rho_{tot}/(\rho_{tot} + p_{tot}) \tag{A7.8}$$

is such a measure which is hypersurface-invariant. Here p_{tot} is the total pressure and ϕ is related to the spatial scalar curvature perturbation by $\delta R = -4\nabla^2\phi - 12C\phi/a^2$, where C is the background spatial curvature constant. (See Bardeen 1980 and Bardeen, Steinhardt and Turner 1983 for notation.) If F_X is extrapolated linearly to the present according to the flat universe growth law ($\sim a$), then

$$F_X(k,t_0) = 2.0 \times 10^6(\Omega_X h^2/\theta)\ T(k,t_0)\ q^2\varsigma(k,t_i) \quad (adiabatic). \tag{A7.9}$$

We require $t_i \ll t_{eq}$, where t_{eq} is the time that the energy in relativistic particles equals that in nonrelativistic ones. To use this for $\Omega < 1$, one should back track F_X to a time before a redshift Ω^{-1}, then evolve forward using the correct growth law for $\Omega \neq 1$ universes (e.g. Peebles 1980). The $n = 1$ initial spectrum has $\langle|\varsigma(k,t_i)|^2\rangle \propto k^{-3}$.

For isocurvature CDM perturbations, the hypersurface-invariant initial amplitude is measured by the fractional relative perturbation $\delta s_X/s_X$ of s_X, the comoving entropy per X particle:

$$F_X(k,t_0) = 6.3 \times 10^4(\Omega_X h^2/\theta)\ T(k,t_0)\left(-\frac{\delta s_X(k,t_i)}{s_X(t_i)}\right) \quad (isocurvature). \tag{A7.10}$$

We advocate F_X/ς and $F_X/(-\delta s_X/s_X)$ as better 'transfer functions' than equations (A7.3-A7.7) for adiabatic and isocurvature cosmological perturbations respectively, since, for waves that are inside the horizon at the present time, they are hypersurface-invariant characterizations. Notice that both converge to the same value, $2 \times 10^6 q^2$, as $k \to 0$.

Table 5.1

Peak Enhancement Factor Coefficients for $\gamma = 0.619$

	$q = \infty$		$q = 16$		$q = 8$	
ν_t	α	β	α	β	α	β
2	1.408	0.956	1.373	0.903	1.260	0.799
3	2.425	1.061	2.215	0.941	1.835	0.767
4	3.500	1.082	2.937	0.868	2.203	0.659

Table 5.2

Enhancement Factors for Galaxies in Rich Clusters

$R_s(h^{-1}Mpc)$	q	ν_t	$E(F_b)$	ν_b	z_t	z_g
0.178	16	3.47	9.2	4.4	3.7	3.6
	8	3.90	5.2	3.7	5.3	4.0
0.356	8	2.75	3.9	3.4	2.9	2.9
	4	3.22	2.7	3.0	4.3	3.0

Table 6.1a

Asymptotic Correlations for the Adiabatic Spectrum

h	$R_s(h^{-1}\mathrm{Mpc})$	q	ν_t	$\langle\tilde{\nu}\rangle$	$\langle\tilde{\nu}\rangle^2 \xi_{\rho/\sigma_0^2}$ $(r=5h^{-1}\mathrm{Mpc})$
0.4	0.178	16	3.47	2.62	0.271
		8	3.90	2.22	0.194
	0.445	16	2.23	1.59	0.225
		8	2.29	1.47	0.191
0.5	0.143	16	3.74	2.80	0.202
		8	4.31	2.31	0.137
	0.356	16	2.62	1.92	0.214
		8	2.75	1.73	0.173
		4	3.21	1.36	0.107

Table 6.1b

Asymptotic Correlations for the Isocurvature Spectrum

h	$R_s(h^{-1}\mathrm{Mpc})$	q	ν_t	$\langle\tilde{\nu}\rangle$	$\langle\tilde{\nu}\rangle^2 \xi_{\rho/\sigma_0^2}$ $(r=5h^{-1}\mathrm{Mpc})$
0.4	0.178	16	3.12	2.47	0.970
		8	3.43	2.17	0.745
	0.445	16	1.79	1.35	0.444
		8	1.80	1.28	0.394
		4	1.87	1.10	0.295
0.5	0.143	16	3.40	2.67	0.854
		8	3.83	2.29	0.626
	0.356	16	2.23	1.71	0.536
		8	2.30	1.58	0.458
		4	2.57	1.32	0.317

REFERENCES

Adler, R.J. 1981, The Geometry of Random Fields (Chichester: Wiley).

Bahcall, N. and Soneira, R. 1983, Ap. J. 270, 20.

Bardeen, J.M., Steinhardt, P. and Turner, M.S. 1983, Phys. Rev. D28, 679.

Bardeen, J.M. 1984, Proceedings Inner Space Outer Space Conference, Fermilab

Bardeen, J.M., Bond, J.R., Jensen, L. and Szalay, A.S. 1985, in preparation.

Barnes, J., Dekel, A., Efstathiou, G.P. and Frenk, C. 1984, preprint.

Blumenthal, G.R. and Primack, J.R. 1984, preprint.

Blumenthal, G.R., Faber, S.M., Primack, J.R. and Rees, M.J. 1984, Nature 311, 517.

Blumenthal, G.R., Faber, S.M. and Primack, J.R. 1985, preprint.

Bond, J.R. and Szalay, A.S. 1983, Ap. J. 277, 443.

Bond, J.R. and Efstathiou, G. 1984, Ap. J. Lett. 285, L45.

Bond, J.R., Szalay, A.S. and Silk, J. 1985, in preparation.

Carlberg, R. and Lake, G. 1985, preprint.

Davis, M. and Huchra, J. 1982, Ap. J. 254, 437.

Dekel, A. 1981, Astron. Ap. 101, 79.

Dekel, A.S. and Silk, J. 1985, preprint.

Doroshkevich, A.G. 1970, Astrophysica 6, 320.

Doroshkevich, A.G. and Shandarin, S. 1978, Sov. Astron. 22, 653.

Doroshkevich, A.G. and Shandarin, S. 1978, M.N.R.A.S. 182, 27.

Dressler, A. 1978, Ap.J. 226, 55.

Efstathiou, G., Davis, M., Frenk, C. and White, S. 1984, preprint.

Efstathiou, G. and Jedrezejewski, R. 1984, Adv. Space. Res. 3, 379.

Efstathiou, G. and Bond, J.R. 1985, preprint.

Faber, S.M. 1982, Astrophysical Cosmology, ed. H.A.Bruck, G.V. Coyne, and M.S.Longair, (Vatican).

Frenk, C., Davis, M., Efstathiou, G. and White, S.D.M. 1985, preprint.

Gott, J.R. and Rees, M.J. 1975, Astron. Ap. 45, 365.

Groth, E. and Peebles, P.J.E. 1977, Ap. J.

Gunn, J.E. and Gott, J.R. 1972, Ap.J. 176, 1.

Hawking, S. 1982, Phys. Lett. 115B, 295.

Hoffman, Y. and Shaham, J. 1984, preprint.

Hogan, C. 1983, M.N.R.A.S. 202, 1101.

Hogan, C. and Kaiser, N. 1983, Ap.J. 274, 7.

Ikeuchi, S. 1981, Publ. Astr. Soc. Japan 33, 211.

Kaiser, N. 1984a, Ap. J. Lett. 284, L9.

Kaiser, N. 1984b, Proceedings Inner Space Outer Space Conference, Fermilab

Kent, S. and Gunn, J. 1982, Ap.J. 87, 945.

Koo, D. 1985, preprint.

Miller, R. 1985, preprint.

Lin, C.C., Mestel, L. and Shu, F.H. 1965, Ap. J. 142, 1431.

Linde, A.D. 1984, Pisma 40, 453; JETP Lett. 40, 1333.

Longuet-Higgins, M.S. 1957, Phil. Trans. A249, 321.

Ostriker, J.P and Cowie, L.L. 1981, Ap.J. Lett. 243, L127.

Peebles, P.J.E. 1980, *The Large Scale Structure of the Universe* (Princeton).

Peebles, P.J.E. 1982a, Ap. J. 258, 415.

Peebles, P.J.E. 1982b, Ap. J. Lett. 263, L1.

Peebles, P.J.E. 1984, Ap. J. 277, 470.

Pi, S.-Y. and Guth, A. 1982, Phys. Rev. Lett. 49, 1110.

Politzer, D. and Wise, M. 1984, Ap. J. Lett. 285, L1.

Primack, J. *etal.* 1985, preprint.

Rees, M.J. and Ostriker, J.P. 1977, M.N.R.A.S. 179, 541.

Rees, M. 1984, Proceedings Trieste meeting on Large Scale Structure.

Rees. M.J. 1985, preprint.

Rice, S.O. 1944, Bell System Tech. J. 23, 282.

Rice, S.O. 1945, Bell System Tech. J. 24, 41.

Ryden, B.S. and Gunn, J.E. 1985, Proc. IAU Symposium on Dark Matter, Princeton.

Silk, J.I. 1983, Nature 301, 574.

Silk, J.I. 1985, preprint.

Starobinskii, A.A. 1982, Phys. Lett. 117B, 175.

Steinhardt, P. and Turner, M.S. 1983, Phys. Lett. 129B, 51.

Vilenkin, A. 1984, Proceedings Inner Space Outer Space Conference, Fermilab

White, S.D.M. and Rees, M.J. 1978, M.N.R.A.S. 183, 341.

Zeldovich, Ya.B. 1970, Astron. Ap. 5, 84.

FIGURE CAPTIONS

FIGURE 4.1:

The spectral parameters for the adiabatic and isocurvature cold dark matter models evolved from Zeldovich initial conditions as a function of Gaussian filtering scale R_f. γ, R_* and σ_0 are defined by equation (4.6). The conditions chosen were $h = 0.5$, $\Omega = 1$, $\Omega_B \ll \Omega$. The scaling of R_f approximately follows $(\Omega h^2)^{-1}$. (See appendix §7.) An equivalent top-hat filtering scale is $R_{TH} \approx 1.6R_f$. Normalization for σ_0 at the present time on galactic scales is discussed in §6.6.

FIGURE 4.2:

The differential number density $\mathcal{N}_{pk}(\nu)d\nu$ of peaks between ν and $\nu + d\nu$ for various values of γ. The spectral parameters γ and R_* are defined by equation (4.6).

FIGURE 4.3:

The cumulative number density $n_{pk}(\nu)$ of peaks with height in excess of $\nu\sigma_0$.

FIGURE 4.4:

The product of the selection function equation (4.13) and the number differential peak number density for $\nu_t = 3.5$ and $q = 8$ and 16 demonstrates the effect of fuzzing out the sharp threshold.

FIGURE 6.1:

Two-point statistical peak-peak correlation functions of peaks of galactic scale R_s for the adiabatic (a) and isocurvature (b) cold dark matter models with $\Omega = 1$, $h = 0.4$. Given R_s and the threshold sharpness parameter q (16 and 8 for the small and large R_s), the threshold is determined by equating $\langle n_{pk}(\nu_t)\rangle$ to the observed density of bright galaxies: (a) $\nu_t = 3.5,\ 2.3$, (b) $\nu_t = 3.1,\ 1.8$ for $R_s = 0.18$ and 0.45 respectively for the two cases. The direct calculation (§6.2) is ξ_{pk} and the one utilizing the peak-background split (§6.3) is ξ_{pb}. For comparison, the mass correlation function ξ_ρ, amplified by a constant factor determined from the large scale limit of ξ_{pk} is also shown. The peak-peak correlation is much closer to a power law than that of the mass. Dynamical evolution of these statistical correlation functions must be included to reach the correct current amplitude of ξ_{pk}.

FIGURE 7.1:

The constrained probability $P(x|\nu)$ of equation (7.5) giving the distribution of $x = -\nabla^2 F/\sigma_2$ for peaks. (The x used here is 1.58 times that in the text.)

FIGURE 7.2:

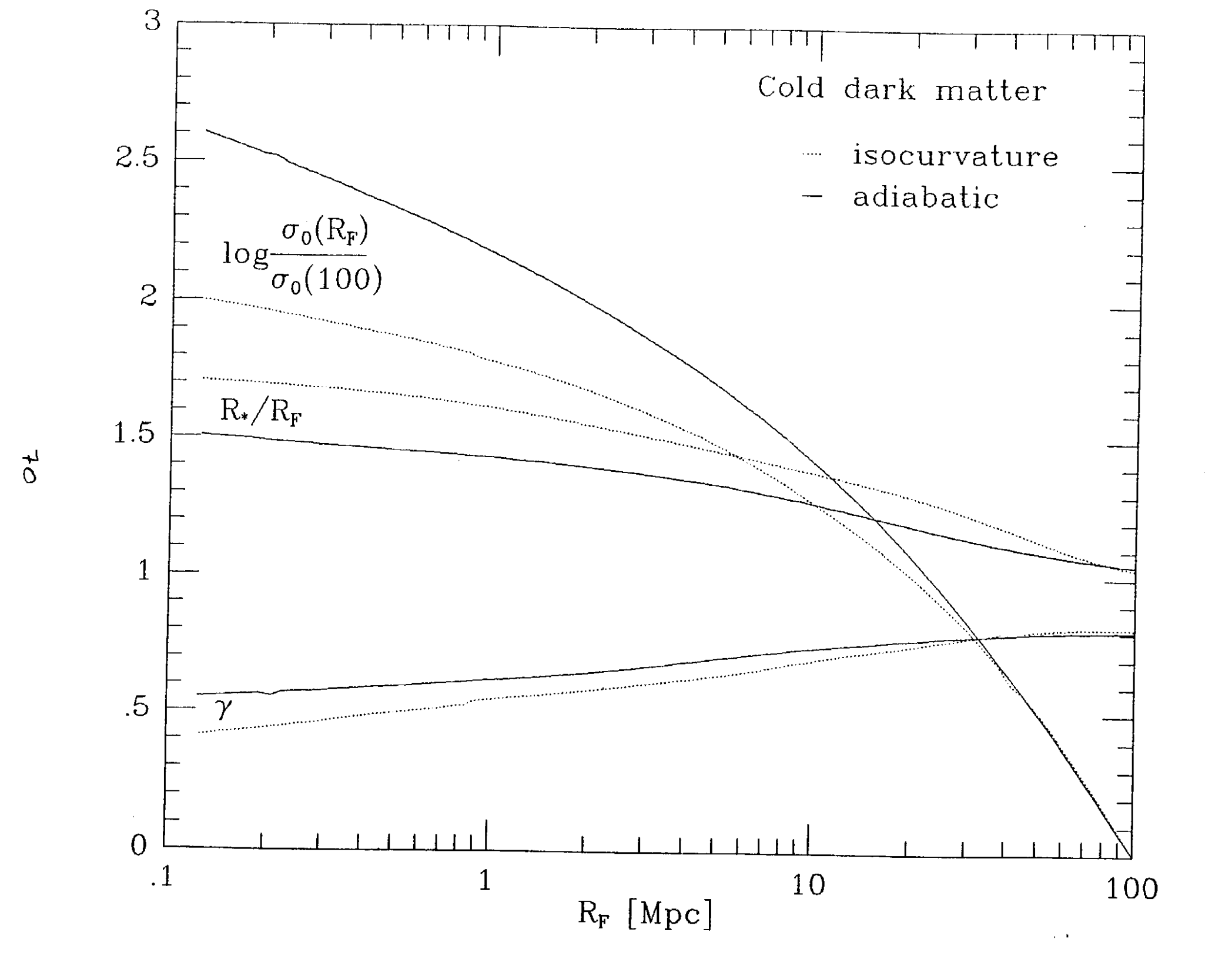
The 95%, 90% and 50% contours of the conditional probability for ellipticity $e = y/x$ and prolateness $p = z/x$ subject to the constraint of given x for peaks (equation 7.6). (The x and x_* used here are $1.58 \approx \gamma^{-1}$ times those used in the text, so $\nu = 1, 2, ..., 6$ corresponds to the different curves.) This figure demonstrates that, even for high ν, the shapes are triaxial. The values of e and p are constrained to lie in the triangle.

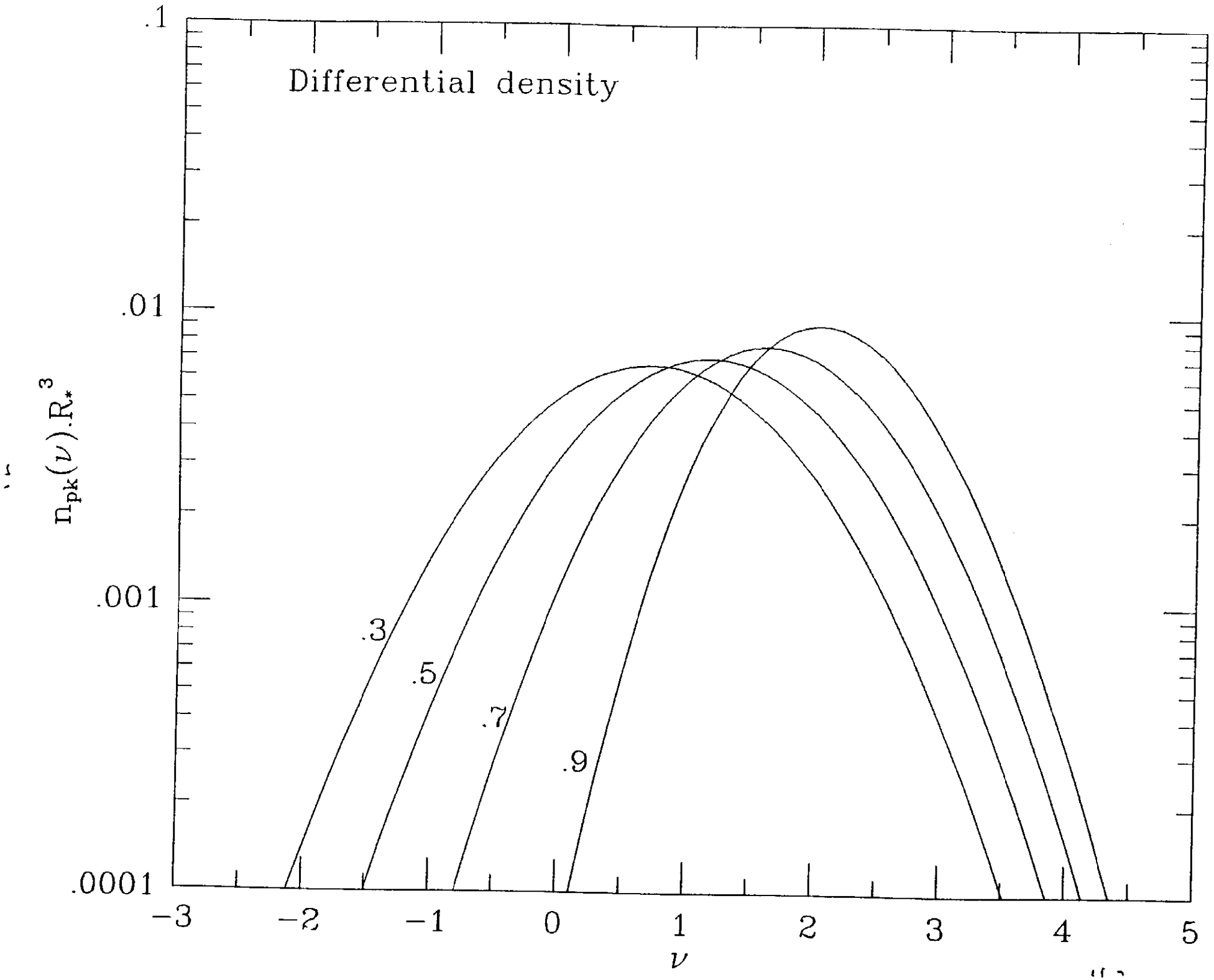
FIGURE 7.3:

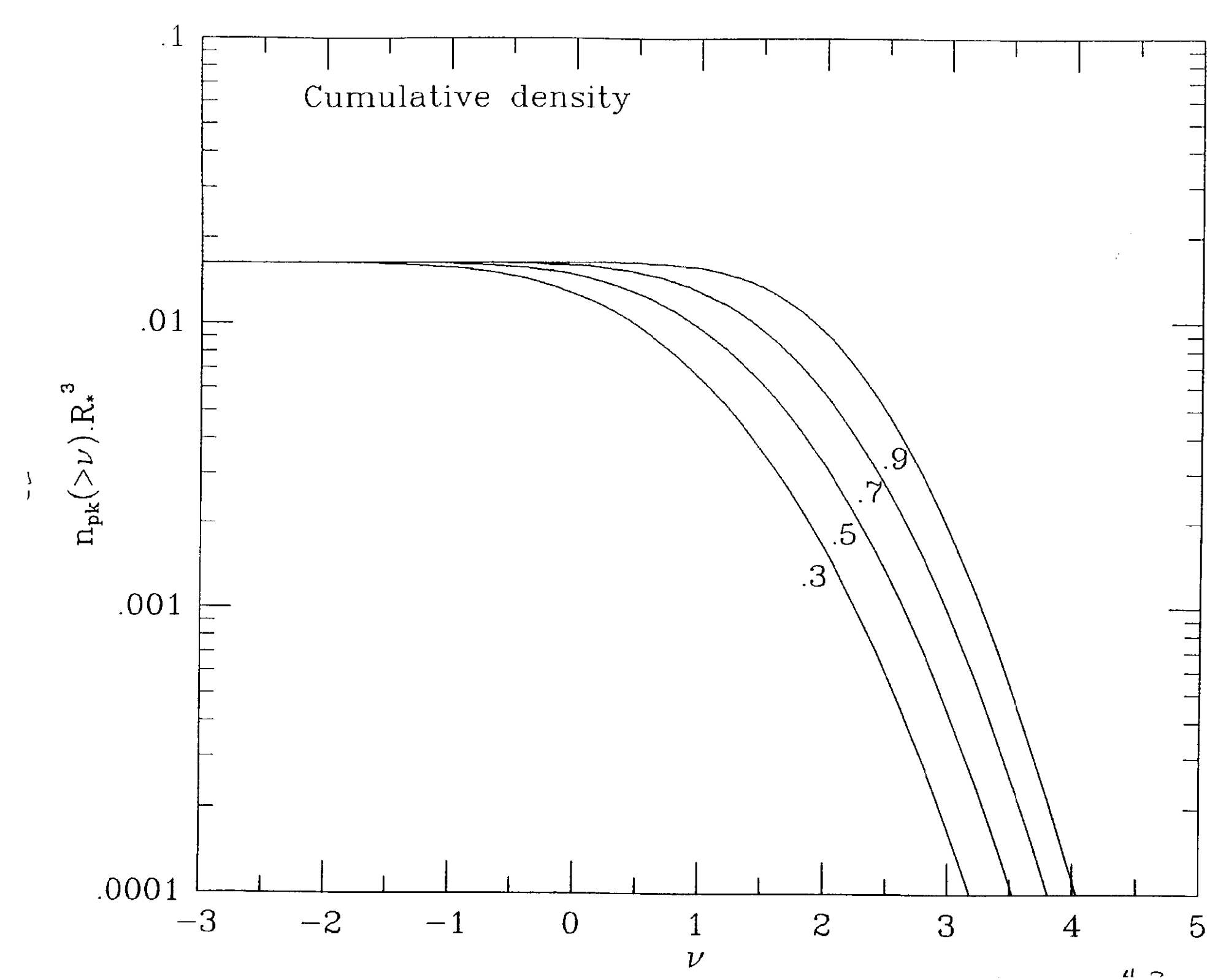
The orientation-averaged density profile about a peak of height $\nu = 2.7$ and its $\pm 1 - sigma$ deviations are compared with the average profile for an ambient field point. The parameters are typical of CDM models of galactic scale structure. The mean spacings of peaks filtered on this scale in units of R_f for the $n = -2.5$ and adiabatic examples respectively are: 18 and 13 for $\nu > 2.68$; 7.5 and 6.3 for $\nu > 1$, which is also just about the spacing for peaks of arbitrary height.

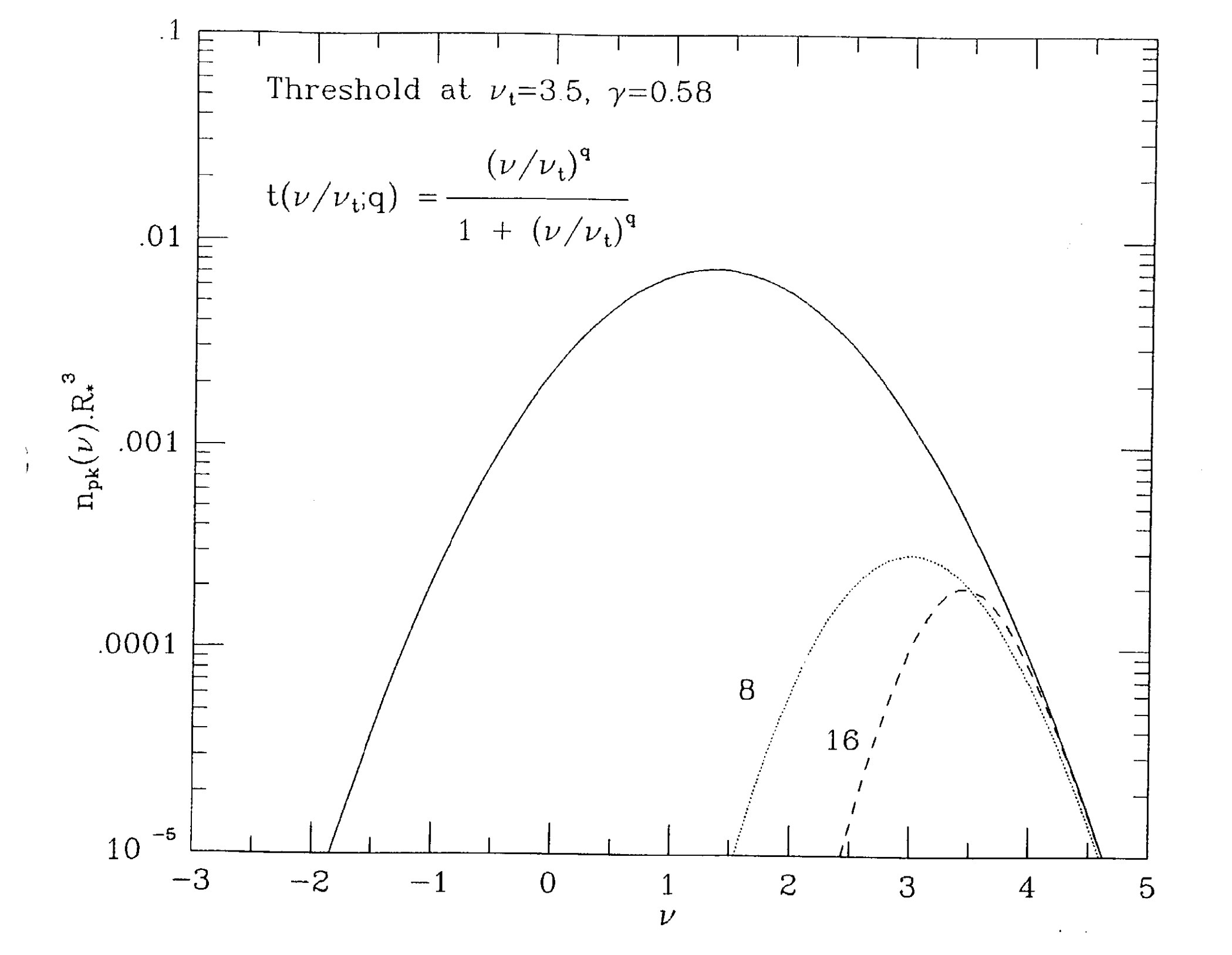
FIGURE 7.4:

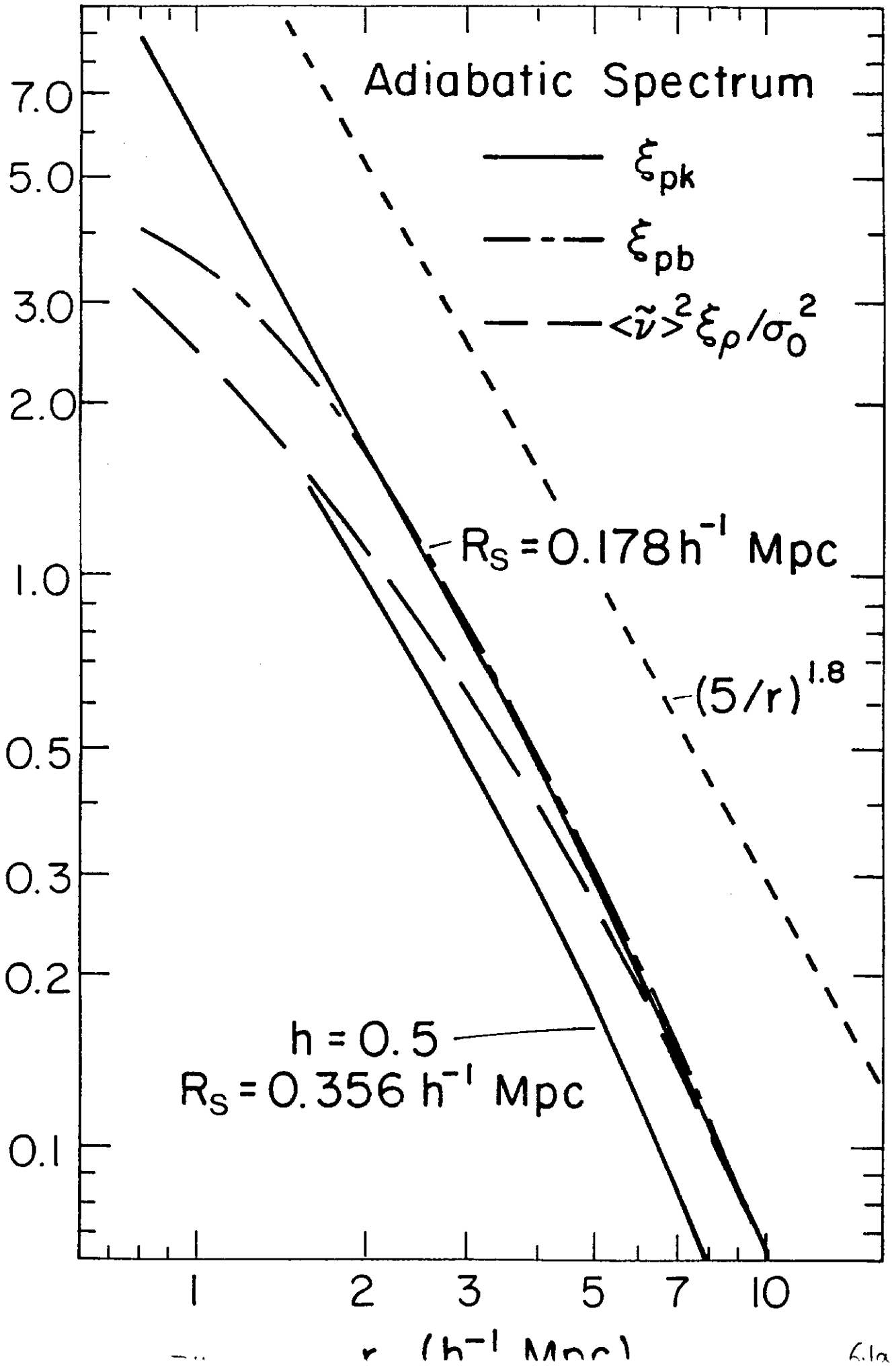
The asymmetry of the $\nu = 2.7$ peak of the previous example with $e \sim 0.2, p = 0$, the most probable values. The density profile along the 1-axis is compared with the sphericalized mean ($e = 0$) of Fig. 7.3 and the profile along the 3-axis. An intermediate asymmetry is also shown.


Cold dark matter
isocurvature
adiabatic
$\log\frac{\sigma_0(R_F)}{\sigma_0(100)}$
R_*/R_F
γ
3
2.5
2
1.5
1
.5
0
.1
1
10
100
R_F [Mpc]
σ_8

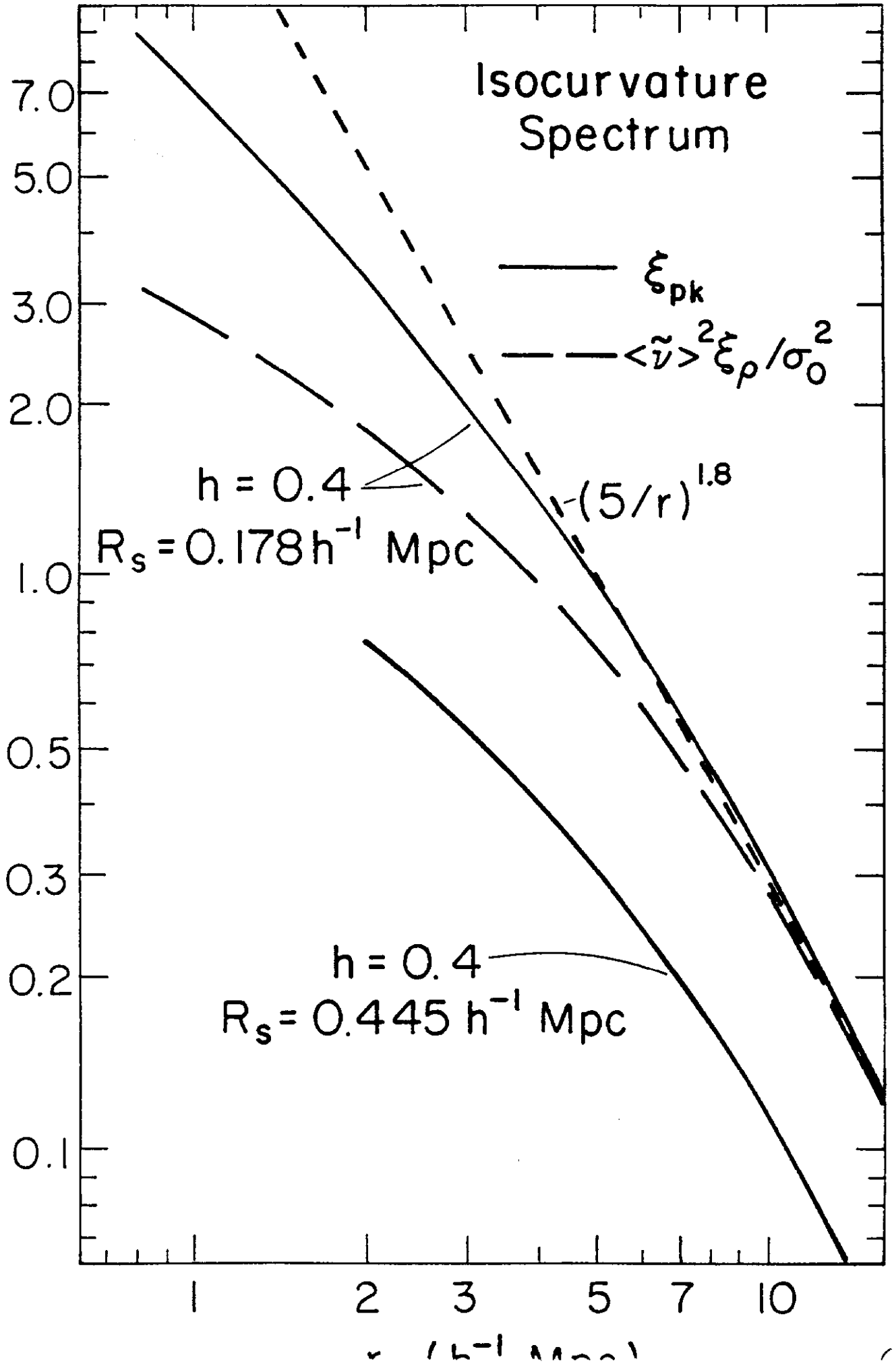

Differential density
.1
.01
.001
.0001
$n_{pk}(\nu).R_*^3$
.3
.5
.7
.9
−3
−2
−1
0
1
2
3
4
5
ν

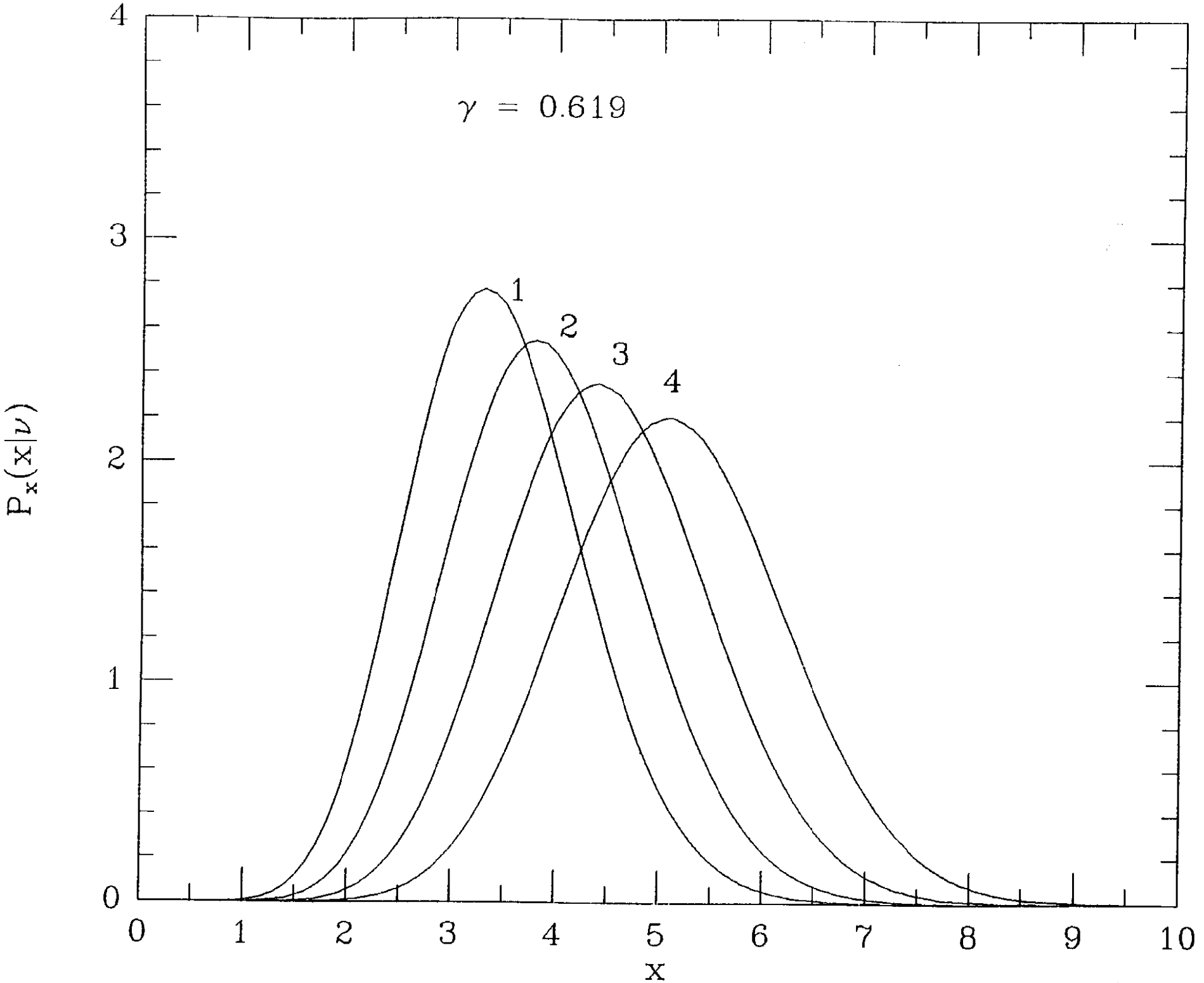

Cumulative density
.1
.01
.001
.0001
$n_{pk}(>\nu).R_*^3$
−3
−2
−1
0
1
2
3
4
5
ν
.9
.7
.5
.3

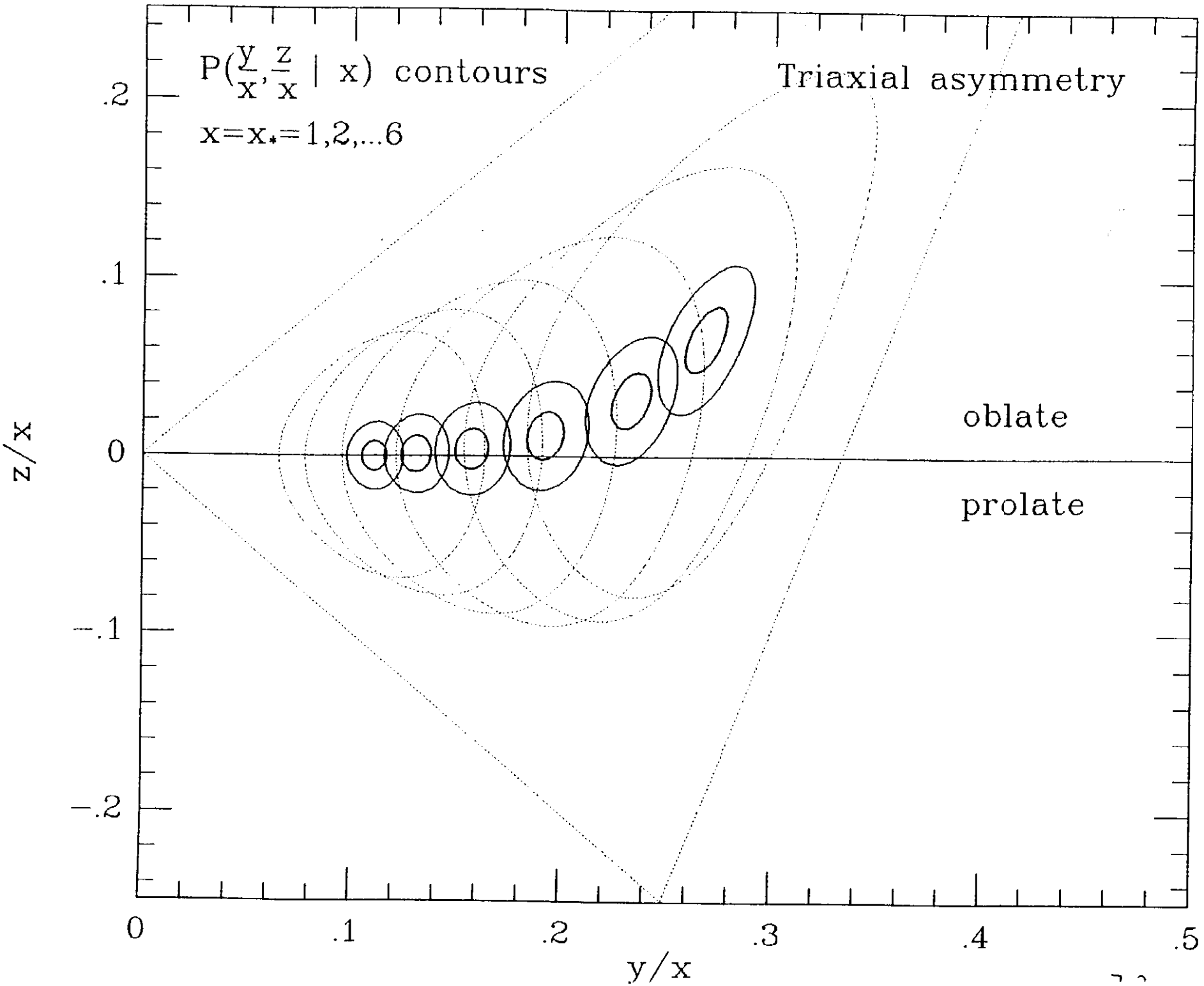

Threshold at ν_t=3.5, γ=0.58
$t(\nu/\nu_t;q) = \frac{(\nu/\nu_t)^q}{1 + (\nu/\nu_t)^q}$
$n_{pk}(\nu).R_*^3$
.1
.01
.001
.0001
10^{-5}
8
16
−3
−2
−1
0
1
2
3
4
5
ν

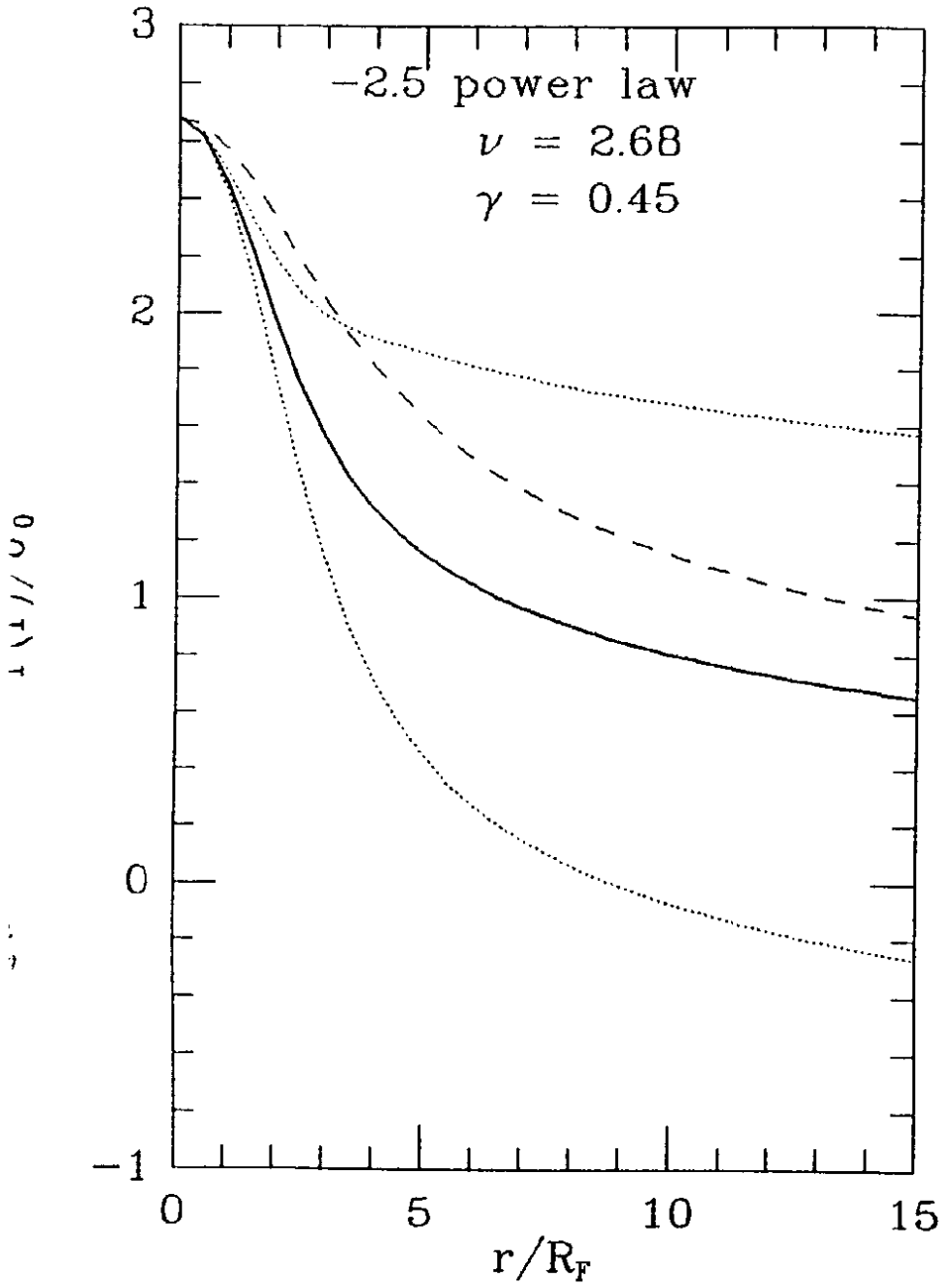

Adiabatic Spectrum
ξ_{pk}
ξ_{pb}
$\langle\tilde{\nu}\rangle^2\xi_\rho/\sigma_0^2$
$R_S = 0.178\,h^{-1}$ Mpc
$(5/r)^{1.8}$
$h = 0.5$
$R_S = 0.356\,h^{-1}$ Mpc
7.0
5.0
3.0
2.0
1.0
0.5
0.3
0.2
0.1
1
2
3
5
7
10
$r\ (h^{-1}$ Mpc)

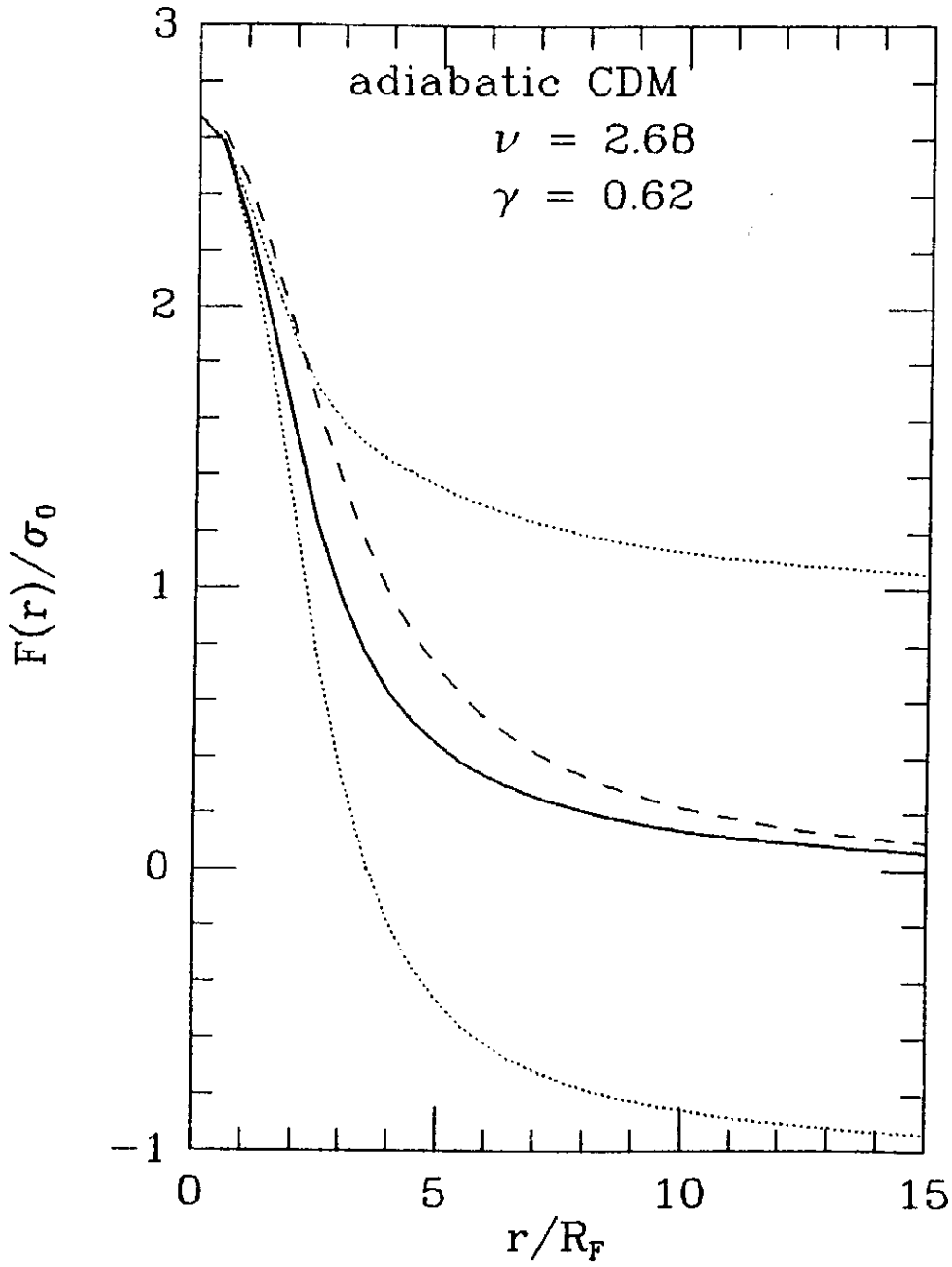
6.1a

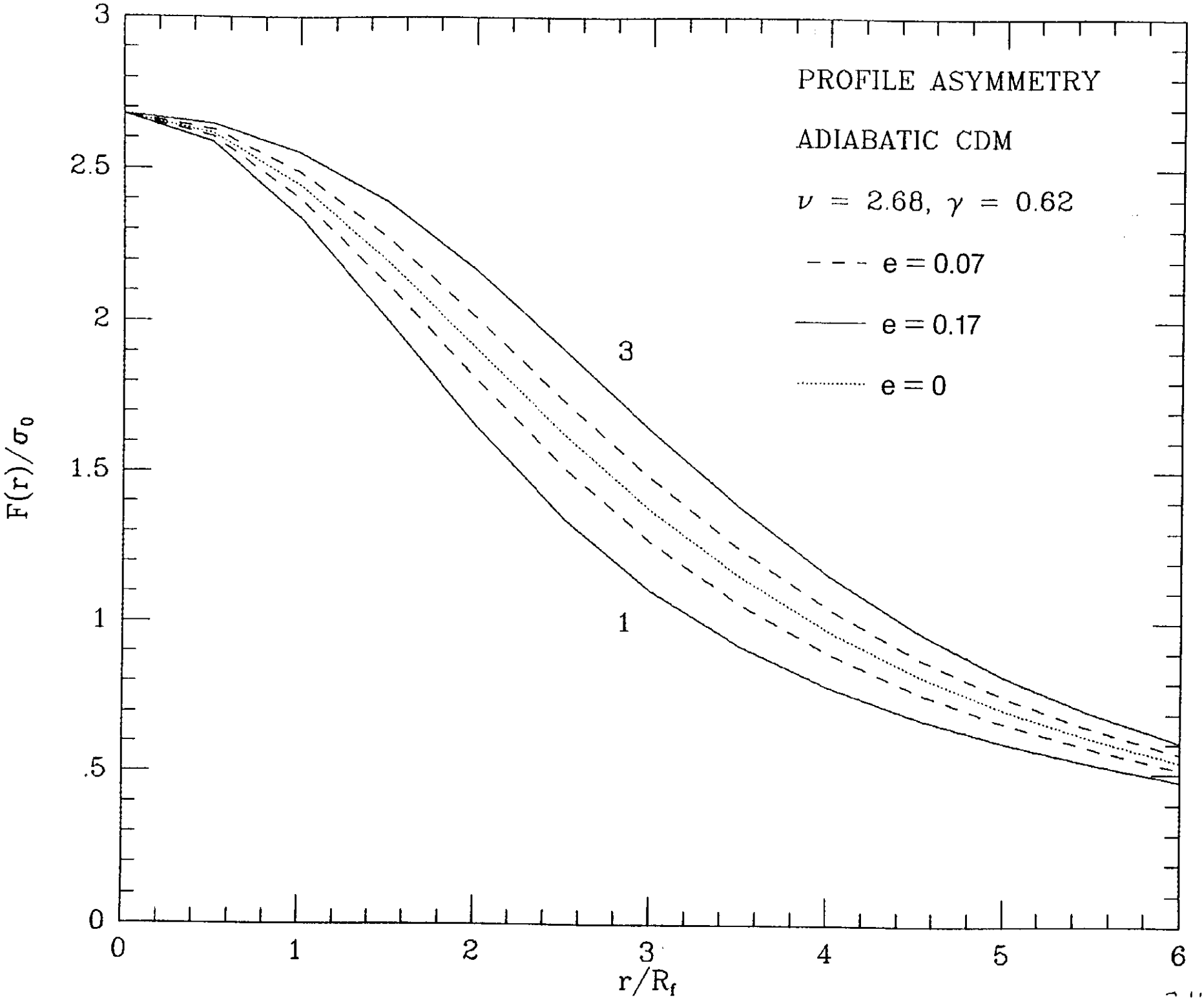

Isocurvature
Spectrum
ξ_{pk}
$\langle\tilde{\nu}\rangle^2\xi_\rho/\sigma_0^2$
$(5/r)^{1.8}$
h = 0.4
$R_s = 0.178\,h^{-1}$ Mpc
h = 0.4
$R_s = 0.445\,h^{-1}$ Mpc
7.0
5.0
3.0
2.0
1.0
0.5
0.3
0.2
0.1
1
2
3
5
7
10


$\gamma = 0.619$
1
2
3
4
$P_x(x|\nu)$
x
0
1
2
3
4
5
6
7
8
9
10


P(y/x, z/x | x) contours
x=x*=1,2,...6
Triaxial asymmetry
oblate
prolate
z/x
y/x
.2
.1
0
-.1
-.2
0
.1
.2
.3
.4
.5


−2.5 power law
ν = 2.68
γ = 0.45
3
2
1
0
−1
0
5
10
15
r/R_F


adiabatic CDM
ν = 2.68
γ = 0.62
F(r)/σ_0
3
2
1
0
−1
0
5
10
15
r/R_F


PROFILE ASYMMETRY
ADIABATIC CDM
ν = 2.68, γ = 0.62
e = 0.07
e = 0.17
e = 0
3
1
$F(r)/\sigma_0$
r/R_f
0
.5
1
1.5
2
2.5
3
0
1
2
3
4
5
6