9

PROBES OF INHOMOGENEITIES

The power spectra we have explored in the previous two chapters — $P(k)$ of the density field and C_l of the anisotropies — are the most obvious first tests of any cosmological model. The most direct way of measuring $P(k)$ is to do a redshift survey, wherein the angular positions and the redshifts (which are a measure of radial distance) of galaxies are recorded. There are, however, a number of problems with redshift surveys and their interpretation. The first is the simple fact that taking redshifts is time consuming: it is much easier to get the angular positions of galaxies than it is to also measure redshifts. In the same time that a 10,000 galaxy redshift survey, say, could be completed, a million angular positions of galaxies could be obtained. With the much greater statistics, angular surveys often compensate for the lack of radial information. Indeed, some claim that the best information we have on large-scale clustering comes from angular surveys. Clearly, then, one skill we must acquire is the ability to make predictions about the angular correlation function $w(\theta)$. In Section 9.1 we will see that the angular correlation function is an integral over the 3D power spectrum.

Redshift surveys suffer from another, more profound problem than the fact that they are time consuming. While is true that the redshift gives a reasonable estimate of radial distance (by *radial* distance I simply mean distance from us), it is not true that this estimate is completely accurate. Recall that a galaxy's velocity is determined solely by the Hubble expansion (and hence redshift is a perfect indicator of distance) only if the galaxy is stationary on the comoving grid. Most galaxies have nonnegligible *peculiar velocities*; that is, they are moving on this grid. A galaxy's total velocity, which is measurable, is

$$\vec{v} = \vec{v}_{\rm pec} + \hat{x} v_{\rm H} \tag{9.1}$$

where the Hubble velocity $v_{\rm H} \equiv \chi da/dt = \chi a H$. Recall that χ is the comoving distance between us and a distant galaxy, so its physical distance from us is $a\chi$. In the absence of peculiar velocities, $\hat{x} \cdot \vec{v}/H$ is a perfect distance indicator. In the real world, though, where peculiar velocities do not vanish, even an accurate

measurement of a galaxy's recession velocity does *not* translate into an unambiguous measurement of its radial distance away from us.

The ambiguity introduced by peculiar velocities offers an opportunity. In linear theory, peculiar velocities are determined by the surrounding density field, so we can correct for the distortions induced by working with redshifts. Indeed, we can go even further and use these distortions to learn about the growth of perturbations, for the precise way in which velocities are related to the density field is determined by the rate at which perturbations grow. Since this rate is determined in cosmology by Ω_m, studying the distribution of galaxies in redshift space is one of the more promising ways to measure Ω_m.

Finally, another way of gleaning information about the underlying mass density is by studying clusters of galaxies. Although strictly speaking, this topic falls in the realm of nonlinear evolution of the density field, and therefore beyond the boundary of this book, the Press–Schechter method of approximating cluster abundances is only a very small step away from linear theory and has been shown to be quite accurate in its predictions. Further, the study of clusters is advancing at an extraordinarily rapid rate, since clusters can now be probed with many different astronomical techniques. Section 9.5 introduces the basic predictions of the Press–Schechter theory and the implications for cold dark matter models.

9.1 ANGULAR CORRELATIONS

Figure 9.1 shows the angular positions of over a million galaxies from the Automated Plate Measuring (APM) Survey. In an angular survey such as this, what statistic can be computed that can be compared with theory? The simplest statistic is the two-point function: in real space it is $w(\theta)$ the *angular correlation function.* In Fourier space, the relevant function is the Fourier transform of w, $P_2(l)$, the two-dimensional power spectrum. In this section we compute both of these very important functions, relating them to the three-dimensional power spectrum.

First let me introduce some notation. Figure 9.2 shows the geometry: a given galaxy is at comoving distance $\chi(z)$ (Eq. (2.42)) away from us. The z-axis is typically chosen so that it points to the center of the distribution of galaxies. In the plane perpendicular to this axis, a galaxy's position is determined by the two-dimensional vector $\vec{\theta} = (\theta_1, \theta_2)$. Therefore, the three-dimensional position vector $\vec{x}$ has components

$$\vec{x}(\chi(z), \vec{\theta}) = \chi(z)(\theta_1, \theta_2, 1). \tag{9.2}$$

The assumption that all galaxies are located near the z-axis clearly breaks down if the survey measures structure on very large angular scales. As an example the APM survey in Figure 9.1 covers roughly 50 by 100 square degrees. From this data, one can measure the correlation function accurately out to about 10° or 0.17 radians, safely smaller than unity.

We measure all galaxies along the line of sight, effectively integrating over $\chi(z)$. Therefore an overdensity at angular position $\vec{\theta}$ is

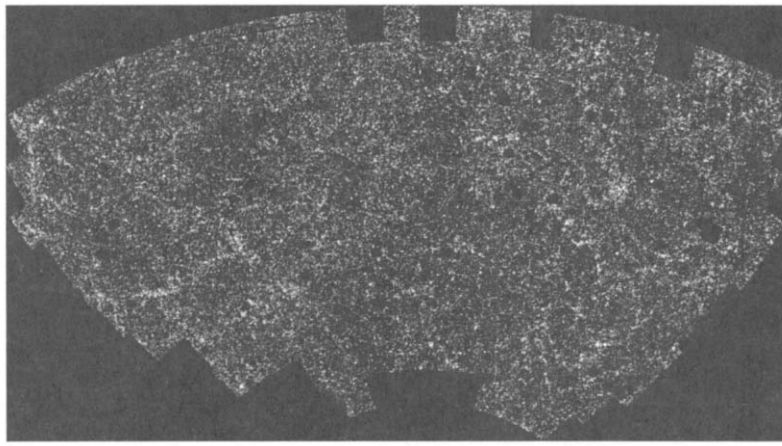

Figure 9.1. The distribution of galaxies in the APM survey. Blacked-out regions were not observed during the survey.

$$\delta_2(\vec{\theta}) = \int_0^{\chi_\infty} d\chi W(\chi)\delta\left(\vec{x}(\chi,\vec{\theta})\right) \tag{9.3}$$

where the subscript $_2$ denotes the fact that δ on the left is the angular — or two-dimensional — overdensity, while δ on the right is the full three-dimensional overdensity. (And I will stick with this convention: for example P_2 is the 2D power spectrum while P denotes the 3D spectrum.) The upper limit on the χ integral corresponds to $z \to \infty$, equal to $\chi_\infty = 2/H_0$ in a flat, matter-dominated universe. In practice, the magnitude-limited surveys that have yielded the most cosmological information to date have probed $z \lesssim 0.5$. The selection function $W(\chi)$ encodes this information: it is the probability of observing a galaxy a comoving distance χ from us. Galaxies at large distances are too faint to be included in a survey, whereas there are relatively few galaxies at very low redshift simply because the volume is small. Since it is a probability, the selection function is normalized so that $\int_0^{\chi_\infty} d\chi W(\chi) = 1$.

The 2D vector conjugate to $\vec{\theta}$ will be $\vec{l}$, so that the Fourier transform of $\delta_2(\vec{\theta})$ is

$$\tilde{\delta}_2(\vec{l}) = \int d^2\theta e^{-i\vec{l}\cdot\vec{\theta}}\delta_2(\vec{\theta}). \tag{9.4}$$

The two-dimensional power spectrum is defined just as was the 3D:

$$< \tilde{\delta}_2(\vec{l})\tilde{\delta}_2^*(\vec{l}') >= (2\pi)^2\delta^2(\vec{l}-\vec{l}')P_2(l); \tag{9.5}$$

here $\delta^2()$ is the 2D Dirac delta function, not to be confused with the overdensities $\delta_2(\vec{\theta})$. Integrating, we can therefore write the 2D power spectrum as

$$P_2(l) = \frac{1}{(2\pi)^2}\int d^2l' \left\langle \delta_2(\vec{l})\delta_2^*(\vec{l}')\right\rangle$$

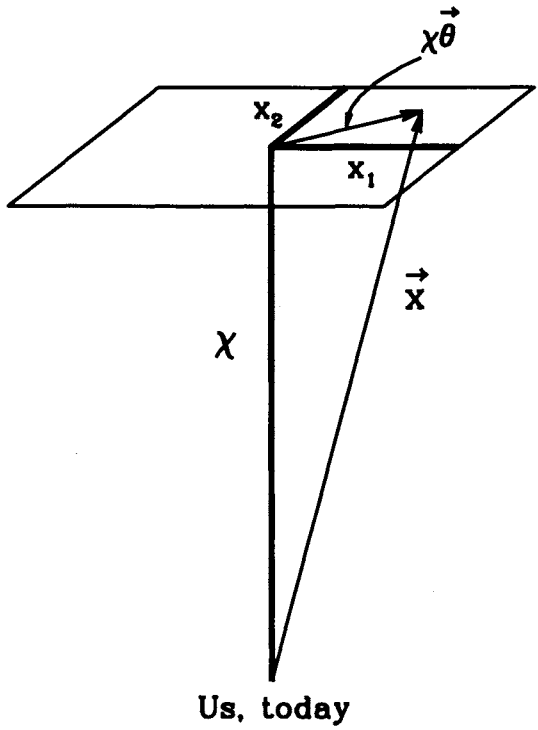

Figure 9.2. A distant galaxy is located at position $\vec{x}$ with respect to us at the origin today. This position can also be expressed in terms of $\chi(z)$, the comoving distance out to the redshift of the galaxy. The $x_1 - x_2$ plane is perpendicular to a suitably chosen x_3-axis. In this plane, a galaxy's position is given by the two-dimensional vector $\chi\vec{\theta}$, so $\vec{x} \simeq \chi(z)(\theta_1, \theta_2, 1)$.

$$= \frac{1}{(2\pi)^2} \int d^2l' \int d^2\theta \int d^2\theta' e^{-i\vec{l}\cdot\vec{\theta}} e^{i\vec{l}'\cdot\vec{\theta}'}$$

$$\times \int_0^{\chi_\infty} d\chi W(\chi) \int_0^{\chi_\infty} d\chi' W(\chi') \Big\langle \delta\left(\vec{x}(\chi, \vec{\theta})\right) \delta\left(\vec{x}'(\chi', \vec{\theta}')\right) \Big\rangle. \quad (9.6)$$

The integral over $\vec{l}'$ gives $(2\pi)^2$ times a Dirac delta function in $\vec{\theta}'$ and the brackets give the 3D correlation function,

$$\xi(\vec{x} - \vec{x}') \equiv \langle \delta(\vec{x})\delta(\vec{x}') \rangle$$

$$= \int \frac{d^3k}{(2\pi)^3} P(k) e^{i\vec{k}\cdot(\vec{x}-\vec{x}')}. \quad (9.7)$$

The average here $\langle \ldots \rangle$ is over all realizations of the density field. At very small distance, we expect galaxies to be clustered strongly as a result of gravity, so ξ is positive. As the distance gets larger, correlations die off and ξ gets smaller and

eventually goes negative. The second line follows since the correlation function is the Fourier transform of the power spectrum (Exercise 1).

Performing the integral over $\vec{\theta}'$ in Eq. (9.6) then leads to

$$P_2(l) = \int d^2\theta e^{-i\vec{l}\cdot\vec{\theta}} \int_0^{\chi_\infty} d\chi W(\chi) \int_0^{\chi_\infty} d\chi' W(\chi') \int \frac{d^3k}{(2\pi)^3} P(k) e^{i\vec{k}\cdot[\vec{x}(\chi,\vec{\theta})-\vec{x}(\chi',0)]}. \tag{9.8}$$

The argument of the exponential at the end here is $i[k_1\chi\theta_1 + k_2\chi\theta_2 + k_3(\chi - \chi')]$, so the integral over angles $\vec{\theta}$ gives Dirac delta functions setting $l_1 = \chi k_1$ and $l_2 = \chi k_2$. We can use these delta functions to do the k_1 and k_2 parts of the d^3k integration, remembering to divide by the derivative of the argument, in this case putting a factor of χ^2 in the denominator. We are thus left with

$$P_2(l) = \int_0^{\chi_\infty} d\chi \frac{W(\chi)}{\chi^2} \int_0^{\chi_\infty} d\chi' W(\chi') \int_{-\infty}^{\infty} \frac{dk_3}{(2\pi)} P\left(\sqrt{k_3^2 + l^2/\chi^2}\right) e^{ik_3[\chi-\chi']}. \tag{9.9}$$

Until now, we have been doing math; to complete the calculation of the power spectrum we need to introduce some physics. I claim that the only 3D Fourier modes that contribute to the integral are those with k_3 very small, much smaller than l/χ. To see why, we first need to estimate l, the variable conjugate to θ. Roughly, l^{-1} is of order the angular scales probed by the survey[1]. Since we are working in the small angle approximation, $l/\chi \sim 1/(\chi\theta) \gg 1/\chi$. Now let's consider Figure 9.3. There

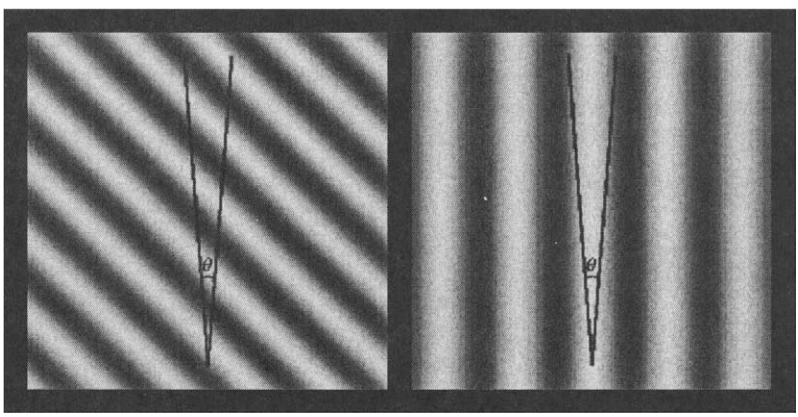

Figure 9.3. Two plane-wave perturbations and their contributions to the 2D power spectrum. Left panel shows a perturbation with longitudinal wavenumber $k_3 \gg \chi^{-1}$ (the $\hat{z}$ direction is vertical). Right panel shows a mode with $k_3 \sim \chi^{-1}$. Angular correlations due to the large k_3 mode (left panel) are negligible since there are many cancellations along the line of sight.

[1]In this sense, l is very similar to the degree of the Legendre polynomials introduced in Chapter 8 to study anisotropies. In the same sense, $P_2(l)$ is very similar to C_l; indeed in Exercise 5 you will show that they are identical on small scales.

we see that modes with longitudinal wavenumber k_3 much greater than χ^{-1} do not give rise to angular correlations because of cancellations along the line of sight. Only modes with k_3 smaller than χ^{-1} lead to angular correlations. Therefore, the relevant transverse wavenumbers l/χ are much larger than the relevant longitudinal wavenumbers, and we can safely set the argument of the 3D power spectrum to l/χ (see Exercise 2 for a more systematic justification). With this approximation, the k_3 integral gives a Dirac delta function in $\chi - \chi'$ so

$$P_2(l) = \int_0^{\chi_\infty} d\chi \frac{W^2(\chi)}{\chi^2} P(l/\chi). \tag{9.10}$$

This is an expression for the 2D power spectrum as an integral over the line of sight. We can change dummy variables from $\chi \to k \equiv l/\chi$ to rewrite the integral as

$$P_2(l) = \frac{1}{l} \int_0^{\infty} dk\ P(k) W^2(l/k). \tag{9.11}$$

The angular correlation function is the Fourier transform of the 2D power spectrum, so

$$w(\theta) = \int \frac{d^2 l}{(2\pi)^2} e^{i\vec{l}\cdot\vec{\theta}} P_2(l). \tag{9.12}$$

Since P_2 depends only on the magnitude of $\vec{l}$, the angular part of the integration over l is $\int_0^{2\pi} d\phi\, e^{il\theta\cos\phi}$, which is proportional to $J_0(l\theta)$, the Bessel function of order zero (Eq. (C.21)). Therefore,

$$\begin{aligned} w(\theta) &= \int_0^{\infty} \frac{dl}{2\pi} l P_2(l) J_0(l\theta) \\ &= \int_0^{\infty} dk\ k\ P(k) F(k,\theta), \end{aligned} \tag{9.13}$$

where the second line follows from changing the order of integration. Here the *kernel* for the angular correlation function is

$$F(k,\theta) \equiv \frac{1}{k} \int_0^{\infty} \frac{dl}{2\pi} J_0(l\theta) W^2(l/k). \tag{9.14}$$

The kernel is plotted in Figure 9.4 for two surveys. Note that it is a function of $k\theta$ (see Exercise 3). The kernel is constant at small $k\theta$ and then begins damped oscillations. The contribution from small k though is suppressed because the integral in Eq. (9.13) is over the kernel weighted by $kP(k)$, and the latter goes to zero as $k \to 0$. Therefore, the modes that contribute most to $w(\theta)$ are typically those with wavenumbers of order the first turnover in the kernel, $k\theta \sim 0.2\,h$ Mpc^{-1} degrees for APM and a factor of 3 smaller for the deeper Sloan Digital Sky Survey (SDSS). This means that the angular correlation function at 5° in APM is most sensitive to power at $k = 0.04\,h$ Mpc^{-1}. The wavenumbers contributing to $w(\theta)$ in a deeper

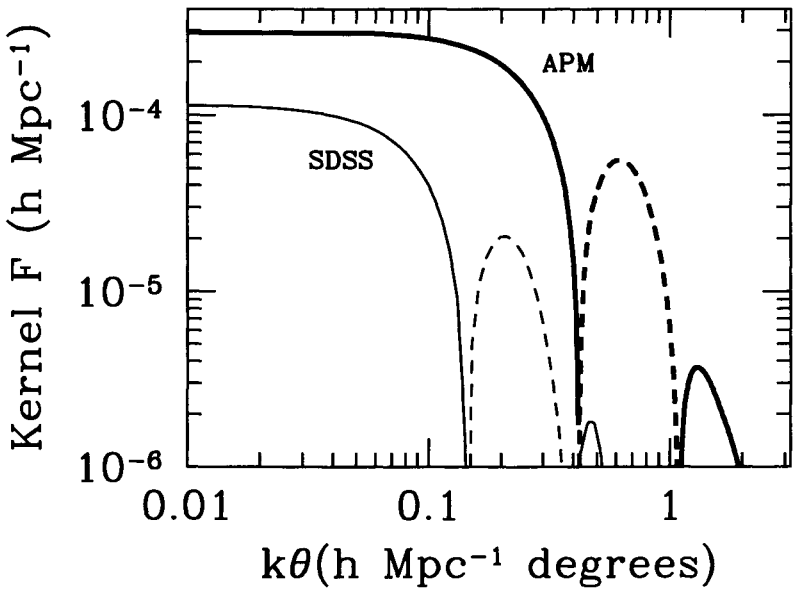

Figure 9.4. The kernel relating the angular correlation function to the 3D power spectrum in two surveys. Kernel is negative when line is dashed, positive otherwise. APM Survey probes galaxies brighter than apparent magnitude $m = 20$, while the Sloan Digital Sky Survey (SDSS) will go much deeper, potentially sensitive to galaxies brighter than $m = 23$.

angular survey are smaller. This makes sense: the same angle probes larger physical scales in a deeper survey.

Figure 9.5 shows the measurements of the angular correlation function from the APM survey. The most important conclusion from the data is that standard Cold Dark Matter — with $\Omega_m = 1$ and $h = 0.5$ — is a bad fit. To quote from the abstract of the Maddox *et al.* (1990) paper which measured the correlation function, "more large-scale clustering than predicted by popular versions of the Cold Dark Matter cosmogony is implied." Although sCDM has died many deaths since its inception in the early 1980s, the death from APM was perhaps its most celebrated. Despite long, hard work by many people trying to find systematic problems with this and other surveys, nothing significant has changed over the past decade to alter the conclusion that the standard Cold Dark Matter model of structure formation fails to predict accurately the observed pattern of large-scale clustering. Having said that, I want to emphasize that the situation is not quite as severe as you might imagine from a cursory examination of Figure 9.5. Consider the prediction from a ΛCDM model, also shown in Figure 9.5. Although the agreement is much better than with sCDM, there are clear discrepancies on both large and small scales. The small-scale discrepancies are completely illusory, though, because I have used the linear power spectrum to compute $w(\theta)$. Nonlinearites become important on scales of order $k \sim 0.2h$ Mpc^{-1} as we saw in Chapter 7 (Exercise 10). These scales

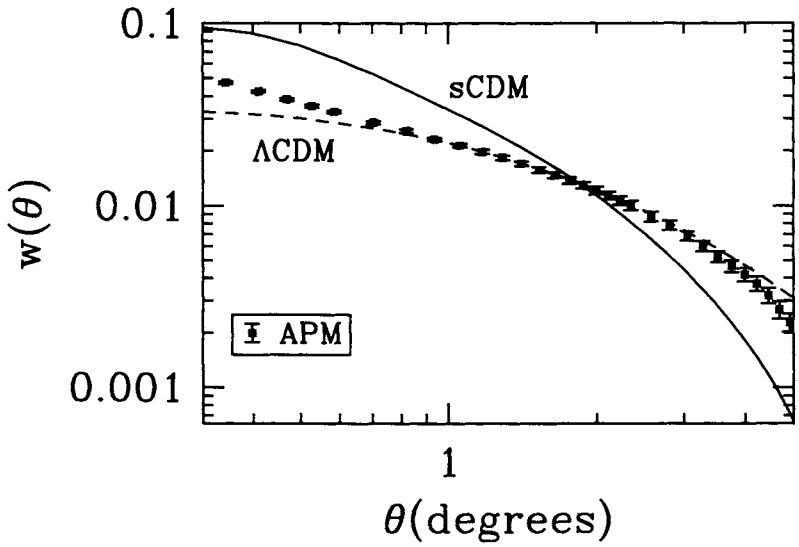

Figure 9.5. Angular correlation function in the APM survey and two theoretical models. Standard CDM (solid curve) is a bad fit to the data, while a model with a cosmological constant (here $\Omega_\Lambda = 0.7$) fits well. The apparent disagreements between the data and ΛCDM on small and large scales are illusory; see text.

contribute to $w(\theta)$ in APM when θ is of order 1°. So to compare to the data fairly, we really need to account for nonlinearities; I have not done this, so we cannot take the small-angle discrepancy seriously. On large angles, people have begun to realize that the data have been overinterpreted. The basic problem is that the points on large angles are highly correlated. So the slight discrepancy between the data and ΛCDM on large angles looks worse than it really is.

The angular correlation function can be inverted to obtain the 3D power spectrum: the results from the APM Survey are shown in the top panel of Figure 9.6. We will discuss inversion techniques in Chapter 11, but you should be aware that the points in Figure 9.6 are the result of a long process (i.e., many papers), in the midst of which error bars were often vastly underestimated. Using only the results on large scales (where nonlinear effects are irrelevant and the relation between mass overdensity and galaxy overdensity is expected to be simple), Efstathiou and Moody (2001) placed constraints on CDM models, shown in the bottom panel of Figure 9.6. We found in Chapter 7 that the CDM transfer function depends only on $k/k_{\rm eq}$. Since $k_{\rm eq}$ scales as $\Omega_m h^2$ and since surveys measure the wavenumber k in units of $h\,{\rm Mpc}^{-1}$, the combination $\Omega_m h$ determines the shape of the power spectrum. For this reason, fits to large-scale structure data are often given in terms of the shape parameter $\Gamma \equiv \Omega_m h$ (sometimes modified to account for baryons: see Exercise 8). Standard CDM with $h = 0.5$ and $\Omega_m = 1$ corresponds to $\Gamma = 0.5$; if the index of the primordial spectrum is $n = 1$, then the bottom right panel of Figure 9.6 shows

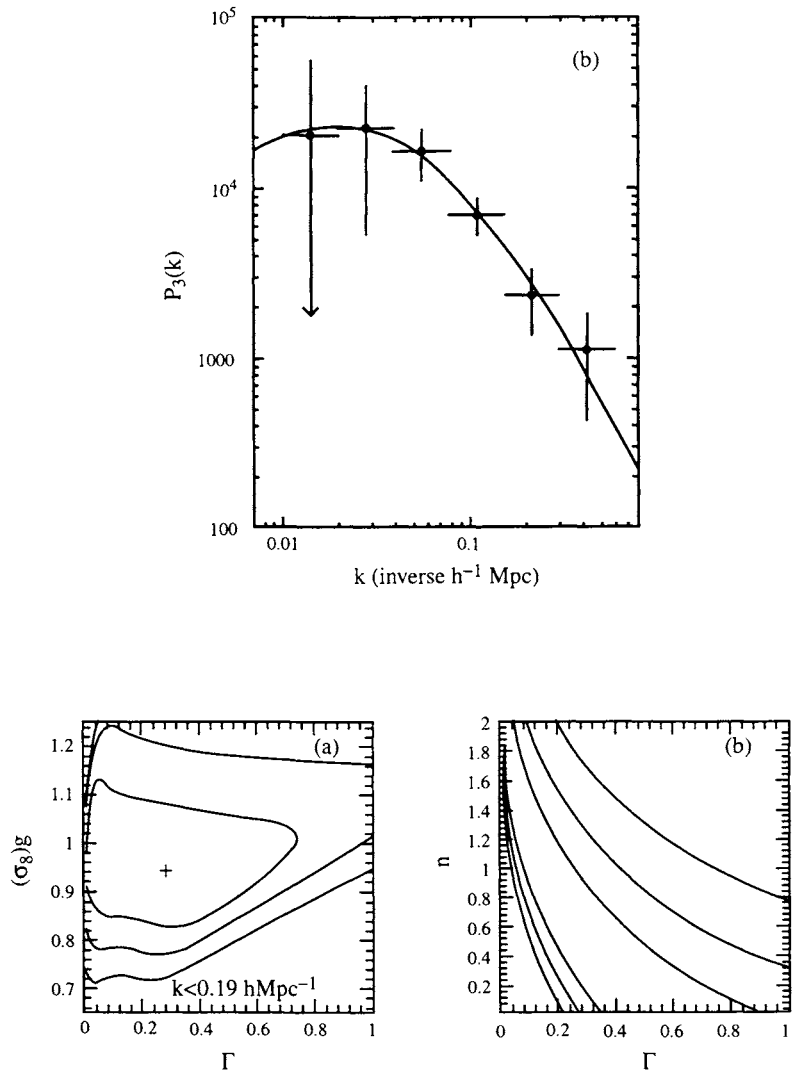

Figure 9.6. Results from the APM Survey (Efstathiou and Moody 2001). *Top panel.* The 3D power spectrum inferred from the angular correlation function. The curve is a CDM model with $\Gamma(\simeq \Omega_m h) = 0.2$, normalized to fit the data. *Bottom Panel.* Constraints on CDM models from the power spectrum. Three parameters — the shape parameter Γ, the amplitude σ_8, and the index of the primordial power spectrum n — were varied. Contours delineate one-, two-, and three-sigma regions. Left panel shows constraints after integrating over n; right after integrating over the amplitude. Standard CDM has $n = 1$; $\Gamma = 0.5$; and (COBE-normalized) $\sigma_8 = 1.15$.

that sCDM is ruled out at the 2-sigma level. ΛCDM has $\Gamma \simeq 0.2$ and is indeed a better fit to the data if n is close to 1.

9.2 PECULIAR VELOCITIES

In linear theory, velocities are related in a simple way to nearby overdensities. We first derive this relation in this section and then consider some of its ramifications. In particular, we will see that by measuring both the peculiar velocity field and the density field, one can infer the present value of the matter density, Ω_m.

In linear theory, we have already derived the equation which determines the velocity field. On scales well within the horizon, the continuity equation (4.103) reduces to

$$\dot{\delta} + ikv = 0. \tag{9.15}$$

At late times, though, we have solved for the evolution of δ: we know that it scales as the growth factor D_1, so

$$v(k,\eta) = \frac{i}{k}\frac{d}{d\eta}\left[\frac{\delta}{D_1}D_1\right] = \frac{i\delta(k,\eta)}{kD_1}\frac{dD_1}{d\eta}. \tag{9.16}$$

A function commonly employed to relate the velocity to the density is the dimensionless linear growth rate,

$$f \equiv \frac{a}{D_1}\frac{dD_1}{da}. \tag{9.17}$$

Since $d/d\eta = a^2 H d/da$, the velocity is related to the density via

$$v(k,a) = \frac{ifaH\delta(k,a)}{k}. \tag{9.18}$$

The linear growth rate can be computed from Eq. (7.77). Most probes of the velocity field to date have been limited to relatively nearby objects, $z \lesssim 0.1$, so it is a reasonable approximation to evaluate f today and neglect its evolution. Figure 9.7 shows the linear growth rate today as a function of the matter density. For small Ω_m there is less growth: mass collapsing into overdense regions has lower velocity than if there was a critical density of matter. This makes sense since an overdensity in a lower density universe has less mass and therefore exerts a weaker gravitational pull on infalling matter. Figure 9.7 shows that, for all practical purposes, the growth rate depends only on the matter density and not, say, on the cosmological constant. Also, the approximation

$$f = \Omega_m^{0.6} \tag{9.19}$$

is seen to work extremely well (see Exercise 7 for a slightly better fit to a flat universe).

There are two important points about the relation between the density and the velocity in Eq. (9.18) that should be emphasized. First, we need to remember that

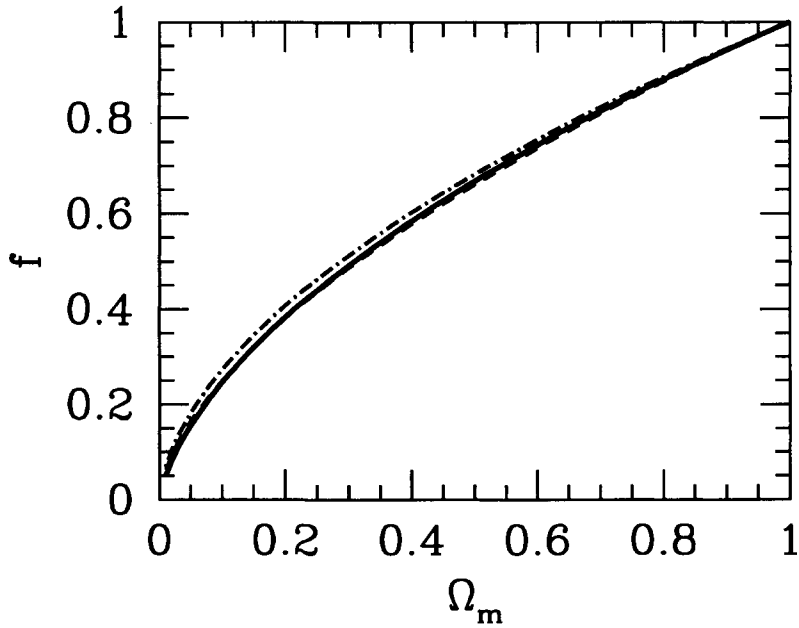

Figure 9.7. The linear growth rate as a function of matter density. There are three curves here, nearly indistinguishable. One (solid) is for an open universe with $\Omega_\Lambda = 0$; another (dot-dashed) for a flat universe with $\Omega_\Lambda = 1 - \Omega_m$; and the last (dashed) is $\Omega_m^{0.6}$.

the velocity is of course a vector and v in Eq. (9.18) is the Fourier component of the velocity parallel to $\vec{k}$. Explicitly, at low redshift we have

$$\vec{v}(\vec{k}) = if H_0 \delta(\vec{k}) \frac{\vec{k}}{k^2}. \tag{9.20}$$

The second point about the relation between the velocity and the density is that it holds only in linear theory. This has turned out to be a big problem for those who have tried to extract information from velocity studies. Velocities are easiest to obtain on small scales, but easiest to compare with theory on large scales.

9.3 DIRECT MEASUREMENTS OF PECULIAR VELOCITIES

A number of surveys have directly measured peculiar velocities. Measuring radial velocities is relatively easy: one just looks for shifted features in the spectrum of the galaxy. The hard part is breaking the radial velocity into the part due to the Hubble expansion and the remainder, the peculiar velocity. Subtracting off the Hubble expansion requires independent (i.e., other than the redshift) knowledge of the galaxy's distance from us. How accurate must a distance indicator be to be useful? Very roughly, typical peculiar velocities are of order 500 km sec^{-1}, while the Hubble velocity is $H_0\chi = 100\,\mathrm{km\ sec^{-1}}\, h\, (\chi/\mathrm{Mpc})$ for a galaxy a distance χ

away. So a galaxy 50 Mpc away has a Hubble velocity roughly 10 times as large as its peculiar velocity. To be useful, therefore, a distance indicator for such galaxies much have an accuracy of order 10%. Indeed this is roughly the best one can hope for, so 50 Mpc is roughly the farthest one can hope to go in a velocity survey.

There are a number of ways of extracting cosmologically useful information from a velocity survey, but I will focus on just one of these: the two point function. With enough velocities, a survey can hope to measure the correlation function

$$\xi_v(\vec{x}_1, \vec{x}_2) \equiv \langle \vec{v}(\vec{x}_1) \cdot \hat{x}_1 \; \vec{v}(\vec{x}_2) \cdot \hat{x}_2 \rangle, \tag{9.21}$$

where the radial components $\vec{v} \cdot \vec{x}$ appear because these are all that can be measured using redshifts. Let us compute this correlation function using linear theory. We will see that it is an integral over the power spectrum, so — just like the angular correlation function of galaxies we considered in Section 9.1 — the radial velocity correlation function is a probe of the power spectrum. The observational obstacles involved in obtaining accurate peculiar velocities are daunting. However, the promise of measuring the *matter* power spectrum (the velocities are due to the matter, which may or may not be aligned with the galaxies) as opposed to the *galaxy* power spectrum sampled by $w(\theta)$ is so alluring that it is likely that peculiar velocity surveys will continue to play an important role in cosmology.

To evaluate the velocity correlation function, we can Fourier transform the velocities appearing in Eq. (9.21) so that

$$\xi_v(\vec{x}_1, \vec{x}_2) = \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}_1} \int \frac{d^3k'}{(2\pi)^3} e^{-i\vec{k}'\cdot\vec{x}_2} \langle \vec{v}(\vec{k}) \cdot \hat{x}_1 \vec{v}^*(\vec{k}') \cdot \hat{x}_2 \rangle. \tag{9.22}$$

Using linear theory for $\vec{v}$ (9.20) and the fact that $\langle \delta(\vec{k})\delta^*(\vec{k}') \rangle = (2\pi)^3 \delta^3(\vec{k} - \vec{k}') P(k)$ leads to

$$\xi_v(\vec{x}_1, \vec{x}_2) = f^2 H_0^2 \int_0^\infty \frac{dk \; k^2}{(2\pi)^3} P(k) \int d\Omega_k e^{i\vec{k}\cdot(\vec{x}_1 - \vec{x}_2)} \frac{\vec{k} \cdot \hat{x}_1 \vec{k} \cdot \hat{x}_2}{k^4}. \tag{9.23}$$

One way to do the angular integral here is to write the occurrences of $\vec{k}$ in the integrand as

$$\vec{k} e^{i\vec{k}\cdot\vec{x}} = \frac{1}{i} \frac{\partial}{\partial \vec{x}} e^{i\vec{k}\cdot\vec{x}} \tag{9.24}$$

where

$$\vec{x} \equiv \vec{x}_1 - \vec{x}_2. \tag{9.25}$$

Then taking the derivatives outside leads to an angular integral over the exponential, which has no azimuthal dependence. The integral of $e^{ikx\mu}$ over μ is is $4\pi j_0(kx)$ (Eq. (C.15)). So we have

$$\xi_v(\vec{x}_1, \vec{x}_2) = -f^2 H_0^2 \hat{x}_{1,i} \hat{x}_{2,j} \int_0^\infty \frac{dk}{2\pi^2 k^2} P(k) \frac{\partial^2}{\partial x_i \partial x_j} j_0(kx). \tag{9.26}$$

The first partial derivative here is

$$\frac{\partial j_0(kx)}{\partial x_j} = \frac{\partial(kx)}{\partial x_j} j_0'(kx) = k\frac{x_j}{x} j_0'(kx) \tag{9.27}$$

where the prime here is derivative with respect to the argument kx. The second derivative then gives

$$\frac{\partial}{\partial x_i}\left[k\frac{x_j}{x} j_0'(kx)\right] = k^2\left[\left\{\delta_{ij} - \hat{x}_i\hat{x}_j\right\}\frac{j_0'(kx)}{kx} + \hat{x}_i\hat{x}_j j_0''(kx)\right]. \tag{9.28}$$

Then, the velocity correlation function is

$$\xi_v(\vec{x}_1, \vec{x}_2) = -f^2 H_0^2 \hat{x}_{1,i}\hat{x}_{2,j}\int_0^\infty \frac{dk}{2\pi^2} P(k)\left[\left\{\delta_{ij} - \hat{x}_i\hat{x}_j\right\}\frac{j_0'(kx)}{kx} + \hat{x}_i\hat{x}_j j_0''(kx)\right]$$

$$\equiv \hat{x}_{1,i}\hat{x}_{2,j}\left\{\delta_{ij} - \hat{x}_i\hat{x}_j\right\}\xi_{v,\perp} + \hat{x}_{1,i}\hat{x}_{2,j}\hat{x}_i\hat{x}_j\xi_{v,\parallel}. \tag{9.29}$$

The definitions here reflect the fact that the first term is sensitive to the component of velocity perpendicular to the line connecting two galaxies, while the second probes the velocity parallel to this line. Finally let's define the angles

$$\cos\theta_1 \equiv \hat{x}_1 \cdot \hat{x} \quad ; \quad \cos\theta_2 \equiv \hat{x}_2 \cdot \hat{x}. \tag{9.30}$$

With the aid of Figure 9.8 we see that $\hat{x}_1 \cdot \hat{x}_2$ is equal to $\cos(\theta_1 - \theta_2)$, so performing the sums over i, j leads to

$$\xi_v(\vec{x}_1, \vec{x}_2) = \sin\theta_1 \sin\theta_2 \xi_{v,\perp} + \cos\theta_1 \cos\theta_2 \xi_{v,\parallel}. \tag{9.31}$$

Both components of ξ_v are integrals over the power spectrum. We can write

$$\begin{pmatrix} \xi(x) \\ \xi_{v,\perp}(x) \\ \xi_{v,\parallel}(x) \end{pmatrix} = \int_0^\infty \frac{dk}{2\pi^2 k} P(k) \begin{pmatrix} k^3 j_0(kx) \\ -f^2 H_0^2 j_0'(kx)/x \\ -f^2 H_0^2 k j_0''(kx) \end{pmatrix}. \tag{9.32}$$

The weighted kernels are shown in Figure 9.9 and compared with the correlation function of the density. The key feature of Figure 9.9 is that, at fixed distance x, the density correlation function probes power on *smaller* scales than do the velocity correlation functions. Put another way, velocity surveys may be limited in how far out they go, but they pack an extra punch since they are sensitive to long wavelength modes. This is a direct result of the fact that $v \propto \delta/k$. The extra factor of $1/k$ gives additional weight to large scales.

Before going further and looking at some results from a velocity survey, I must digress to make note of an important feature in Figure 9.9. The density correlation function, as oppossed to the velocity correlation functions, gets its contribution at a fixed distance from many Fourier modes. The figure shows the contributions to $\xi(50\,h^{-1}\,\mathrm{Mpc})$. Naively, we would expect the main contribution to $\xi(50\,h^{-1}\,\mathrm{Mpc})$ to come from Fourier modes with $k \sim x^{-1} = 0.02\,h\,\mathrm{Mpc}^{-1}$. Since overdensities are small, i.e., still in the linear regime, on these scales, the naive expectation is

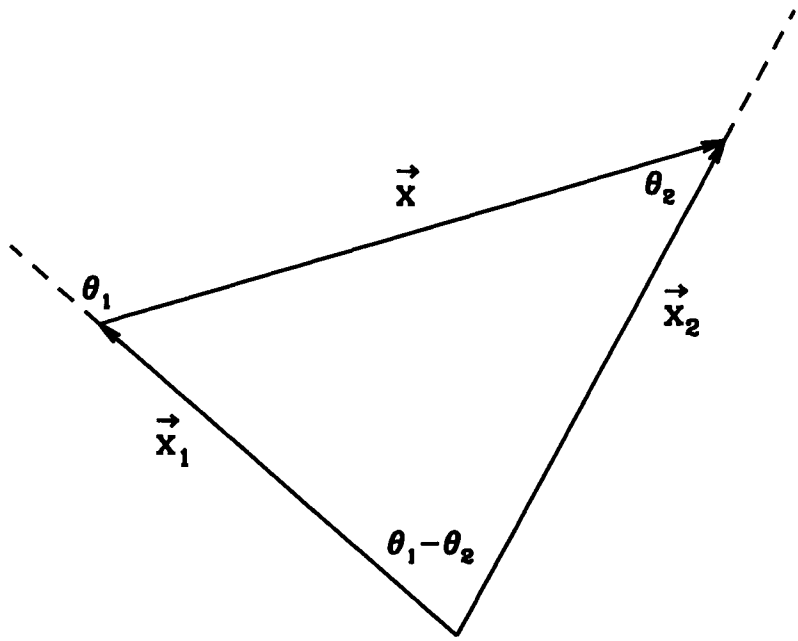

Figure 9.8. The vectors and angles of the the velocity correlation function. $\vec{x}$ is the difference vector; since the angle complementary to θ_1 is $\pi - \theta_1$, and since the three angles in the triangle must sum to π, the angle between the two galaxies is $\theta_1 - \theta_2$.

that $\xi(50\,h^{-1}\,\mathrm{Mpc})$ probes the linear power spectrum. Figure 9.9 shows that this expectation is incorrect. Modes with k as small as $0.02\,h\,\mathrm{Mpc}^{-1}$ do contribute to the correlation function, but contributions extend out to $k \sim 0.3\,h\,\mathrm{Mpc}^{-1}$ and beyond. This means that even on scales you would think would be safely linear, the correlation function depends on the small-scale power. Ultimately we want to compare theory with observations, and we are most confident doing so for modes that are still linear. The correlation function mixes up linear and nonlinear modes, so makes it difficult to compare theory with observations. For this reason, the power spectrum has gained preeminence as the statistic of choice for large-scale structure.

Returning to the velocity correlations, let's consider Figure 9.10. It shows one attempt (Freudling *et al.*1999) to extract cosmological information from a velocity survey, using 1300 velocities in the all-sky SFI catalogue (Haynes *et al.*1999), which goes out to $70\,h^{-1}\,\mathrm{Mpc}$. The power spectrum was parameterized by an amplitude A and the Γ parameter in the BBKS transfer function. For the analysis shown in Figure 9.10, the primordial spectral index was set to 1. The standard COBE-normalized CDM model, with critical matter density and $h = 0.5$, has an amplitude of $A = 0.29A_0$ (Eq. (7.9)) and shape parameter $\Gamma = 0.5$, so the SFI survey rules out this model at many sigma.

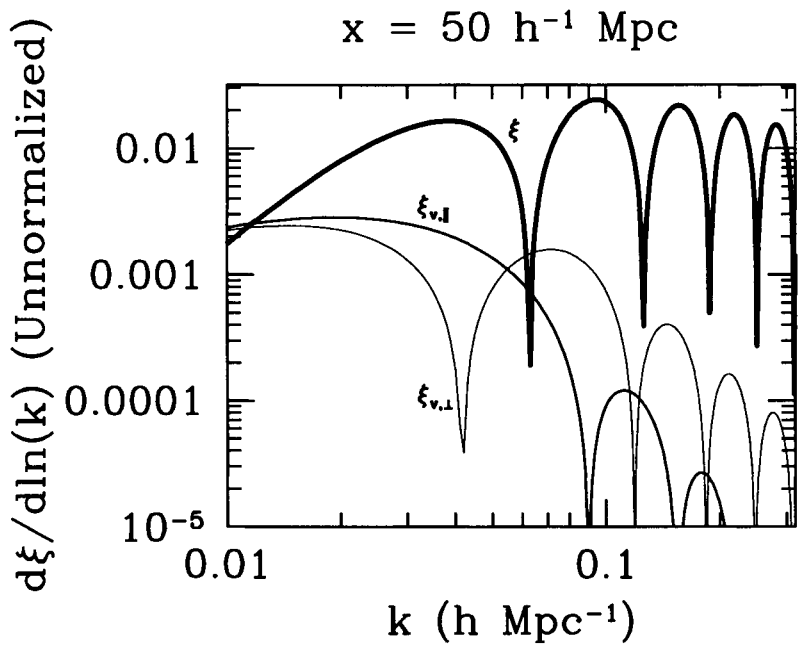

Figure 9.9. Contribution to various correlation functions from wavenumber k. Note that the velocity correlation functions get most of their contribution from smaller k than does the ordinary density correlation function.

9.4 REDSHIFT SPACE DISTORTIONS

Redshift surveys supplement the angular information about galaxies with an estimator for the radial position, the redshift. The simplest guess about the radial position of a galaxy with redshift z is that it lies a distance

$$\chi_s(z) = \frac{z}{H_0} \tag{9.33}$$

away from us, where the subscript $_s$ denotes *redshift space*. Redshift space then corresponds to assigning Cartesian coordinates to a galaxy equal to

$$\vec{x}_s = \frac{z}{H_0}(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)\,. \tag{9.34}$$

This assignment neglects several unpleasant realities. First, the comoving distance out to a galaxy at redshift z is equal to z/H_0 only at relatively low redshifts. A glance back at Figure 2.3 should convince you that this approximation is off by as much as 50% at $z = 1$. Fortunately, this first problem has not yet been much of a problem since most redshift surveys to date have probed $z \lesssim 0.1$.

A second, more pernicious, problem with redshift space is that the estimate for the distance in Eq. (9.33) neglects peculiar velocities. Figure 9.11 illustrates the

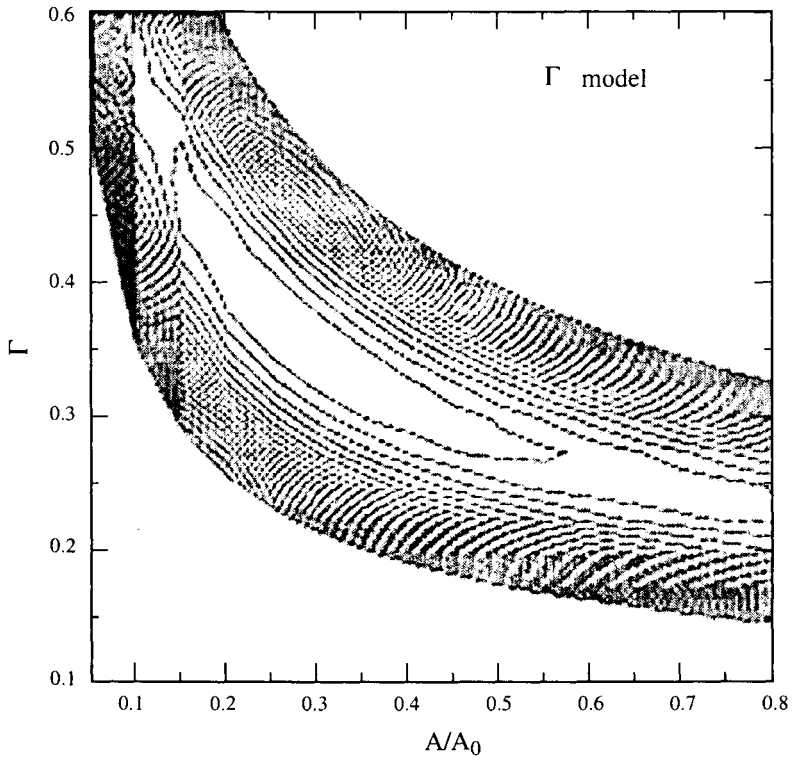

Figure 9.10. Likelihood contours on the amplitude A of the power spectrum and the shape parameter $\Gamma \equiv \Omega_m h$ from the SFI peculiar velocity survey. The amplitude is given in units of $A_0 = 2 \times 10^6 (h^{-1}\ \mathrm{Mpc})^4$, and the contours indicate shifts in the likelihood function by successive factors of $1/e$. The allowed region — delineated by banana-shaped contour in the center — is consistent with ΛCDM, but strongly disfavors standard CDM, which has $\Gamma = 0.5$ and $A = 0.29 A_0$. From Freudling *et al.* (1999).

distortions that appear in redshift space. A slightly overdense region which is just beginning to collapse appears squashed in redshift space: the galaxies closest to us are moving toward the center of the overdense region and hence away from us, so they appear farther from us (and closer to the center of the overdense region) than they actually are. Similarly, galaxies on the "other side" of the perturbation are moving toward us, so they appear closer to us than they actually are. The overall effect is to induce an apparent quadrupole moment in an otherwise circular overdensity.

As a region becomes more overdense, the nature of the redshift space distortion changes. The bottom part of Figure 9.11 shows that a more collapsed object gets distorted in a much different way. It is elongated along the line of sight. More quantitatively, its quadrupole moment has the opposite sign as does a linear overdensity. It is clear then that accounting for redshift space distortions will be a tricky busi-

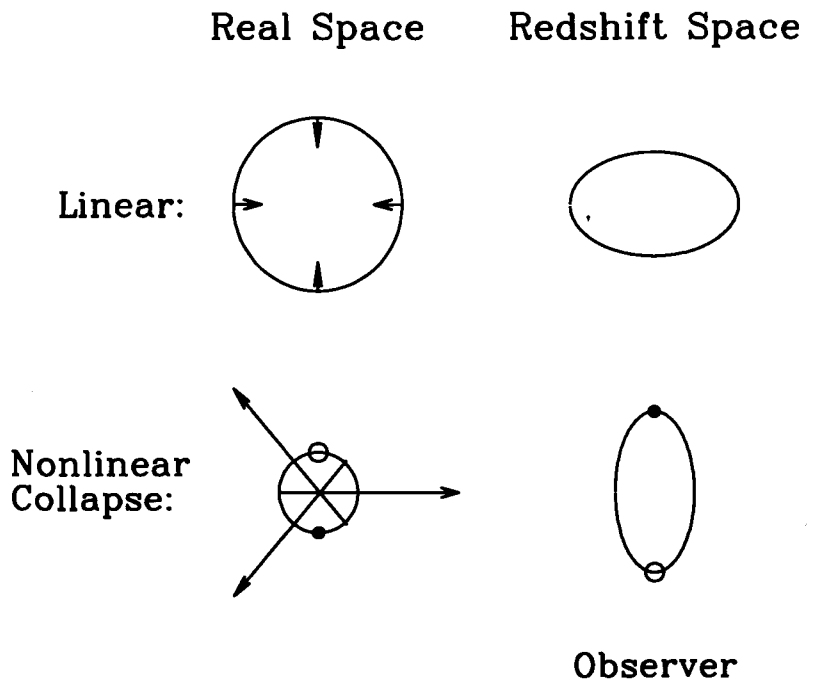

Figure 9.11. Redshift space distortions. In each case, a contour of constant density (circular in real space) is distorted in redshift space so that it looks asymmetric. Arrows denote direction and magnitude of velocity. In the case of nonlinear collapse, the velocities are so large that a point on "our side" (the bottom) of the center is mapped onto a point on the opposite side (compare the position of the solid dot on the bottom left and right).

ness, requiring careful treatment not only of linear overdensities, but also of the much more complicated effects of nonlinearities. We will content ourselves with a quantitative treatment of linear distortions, since this applies on large scales and is the starting point for all further work.

Suppose we measure the power spectrum in redshift space. How is this distorted power spectrum related to the underlying true spectrum in real space? Kaiser (1987) first solved this problem, working within the context of linear theory. The starting point is the realization that the number of galaxies in a particular region is the same, whether we use redshift-space or real-space coordinates. Therefore,

$$n_s(\vec{x}_s)d^3x_s = n(\vec{x})d^3x \tag{9.35}$$

where n is the density of galaxies at $\vec{x}$ in real space, and n_s is the density in redshift space. The infinitesimal volume around a point in redshift space is $d^3x_s = dx_s x_s^2 \sin\theta d\theta d\phi$, while the volume around a point in real space is $d^3x = dx x^2 \sin\theta d\theta d\phi$. The angular volume elements are identical, so

$$n_s(\vec{x}_s) = n(\vec{x})J \tag{9.36}$$

where the Jacobian J is given by

$$J \equiv \left| \frac{d^3x}{d^3x_s} \right| = \frac{dx}{dx_s} \frac{x^2}{x_s^2}. \tag{9.37}$$

To compute the Jacobian, we use the fact that the observed redshift is the sum of two terms:

$$z = H_0 x + \vec{v} \cdot \hat{x}. \tag{9.38}$$

The first term is the standard Hubble law, which says that redshift is proportional to distance; the second is the velocity along the line of sight. Recalling that redshift space corresponds to equating a galaxy's redshift with its distance from us, we see that, after dividing by H_0, this equation becomes

$$x_s = x + \frac{\vec{v} \cdot \hat{x}}{H_0}. \tag{9.39}$$

The Jacobian can be now be read off as

$$J = \left(1 + \frac{\partial}{\partial x}\left[\frac{\vec{v} \cdot \hat{x}}{H_0}\right]\right)^{-1} \left(1 + \frac{\vec{v} \cdot \hat{x}}{H_0 x}\right)^{-2}. \tag{9.40}$$

Kaiser realized that the correction term due to the derivative of the velocity is much more important than the $\vec{v} \cdot \hat{x}/H_0 x$ term. The argument goes as follows. For a plane wave perturbation, the term with the derivative of the velocity is of order $v'/H_0 \sim kv/H_0$, while the other correction is of order $v/H_0 x$. That is, the first correction term is larger than the second by a factor of order kx. Why do I say it's *larger*? That is, why is kx larger than unity? Kaiser's argument is that x is of order the size of the survey, while k is of order the Fourier modes we can hope to measure in the survey. Perturbations on the largest scale probed by the survey $k \sim x^{-1}$ are very poorly determined, since there are only a handful of Fourier modes with wavelength of order the survey size. Modes with smaller wavelength are much easier to measure since there are many such modes, and we effectively average over all of them to get an estimate of the power spectrum. Therefore, we are really interested only in modes with $kx \gg 1$. Expanding the remaining denominator about $v = 0$, we see that

$$J \simeq \left(1 - \frac{\partial}{\partial x}\left[\frac{\vec{v} \cdot \hat{x}}{H_0}\right]\right). \tag{9.41}$$

The number densities in real and redshift space are $n = \bar{n}(1 + \delta)$ and $n_s = \bar{n}(1 + \delta_s)$, respectively, with $\bar{n}$ the average number density. In light of Eq. (9.36), the overdensity in redshift space is

$$1 + \delta_s = [1 + \delta]\left(1 - \frac{\partial}{\partial x}\left[\frac{\vec{v} \cdot \hat{x}}{H_0}\right]\right). \tag{9.42}$$

Expanding to first order, we see that the overdensity in redshift space is actually a sum of the overdensity in real space and a correction due to peculiar velocity,

$$\delta_s(\vec{x}) = \delta(\vec{x}) - \frac{\partial}{\partial x}\left[\frac{\vec{v}(\vec{x})\cdot\hat{x}}{H_0}\right]. \tag{9.43}$$

Now I want to introduce the *distant observer* approximation. The idea is that, in most cases of interest, the direction vector $\vec{x}$ is fixed, varying little from galaxy to galaxy. To see this, go back to Figure 9.2: $\vec{x}$ is mostly radial, with only small components in the transverse direction (proportional to the θ_1 and θ_2). As long as the galaxies are relatively close to each other in this plane, we can approximate $\hat{x}\cdot\vec{v} \to \hat{z}\cdot\vec{v}$, where $\hat{z}$ is a radial vector pointing to the center of the galaxies of interest.

In the distant observer approximation, we can compute the Fourier transform of the redshift space overdensity (here denoting the Fourier-transformed density by $\tilde{\delta}$ to avoid confusion),

$$\begin{aligned}\tilde{\delta}_s(\vec{k}) &= \int d^3x\, e^{-i\vec{k}\cdot\vec{x}}\left[\delta(\vec{x}) - \frac{\partial}{\partial x}\left[\frac{\vec{v}(\vec{x})\cdot\hat{z}}{H_0}\right]\right] \\ &= \tilde{\delta}(\vec{k}) - if\int d^3x\, e^{-i\vec{k}\cdot\vec{x}}\frac{\partial}{\partial x}\left[\int\frac{d^3k'}{(2\pi)^3}e^{i\vec{k}'\cdot\vec{x}}\tilde{\delta}(\vec{k}')\frac{\vec{k}'}{k'^2}\cdot\hat{z}\right],\end{aligned} \tag{9.44}$$

the first equality following from our Fourier convention and Eq. (9.43) and the second from Eq. (9.20) for linear velocities. The derivative with respect to the length x acts on the exponential, bringing down a factor of $i\vec{k}'\cdot\hat{x}$, which we again set to $i\vec{k}'\cdot\hat{z}$, so

$$\tilde{\delta}_s(\vec{k}) = \tilde{\delta}(\vec{k}) + \int\frac{d^3k'}{(2\pi)^3}\;\tilde{\delta}(\vec{k}')\left[f\left(\hat{k}'\cdot\hat{z}\right)^2\right]\int d^3x e^{i(\vec{k}'-\vec{k})\cdot\vec{x}}. \tag{9.45}$$

The $\vec{x}$ integral gives a Dirac delta function, equating $\vec{k}'$ with $\vec{k}$. Therefore, in the distant observer approximation,

$$\tilde{\delta}_s(\vec{k}) = \left[1 + f\mu_{\mathbf{k}}^2\right]\tilde{\delta}(\vec{k}). \tag{9.46}$$

Here $\mu_{\mathbf{k}}$ is defined to be $\hat{z}\cdot\hat{k}$, the cosine of the angle between the line of sight and the wavevector $\hat{k}$. Equation (9.46) quantifies what we should have anticipated about (linear) redshift space distortions. First of all, since $f\mu_{\mathbf{k}}^2 \geq 0$, the apparent overdensity in redshift space is *larger* than in real space. This is clear from Figure 9.12. The central region of the galaxies is clearly more overdense in redshift space than in real space, the enhancement due to the illusion that infalling galaxies are located close to the center. The second feature of Eq. (9.46) worth noting is that the enhancement is for waves with wavevector parallel to the line of sight. A plane wave perturbation with $\vec{k}$ perpendicular to the line of sight — one in which the density along the line of sight is constant — experiences no redshift space distortion.

The power spectrum in redshift space depends not only on the magnitude of $\vec{k}$ but also on its direction, which we are parameterizing with μ. It follows immediately from Eq. (9.46) that

Figure 9.12. A hundred galaxies in real space squashed in redshift space due to linear velocities. The apparent overdensity in redshift space is much larger near the center than it is in real space. We, the observers, are sitting at the bottom of the page.

$$P_s(\vec{k}) = P(k)\left[1+\beta\mu_{\mathbf{k}}^2\right]^2. \tag{9.47}$$

Here I have introduced the parameter β, which you might think is simply equal to f, the linear growth rate. There is an additional factor in β, though, due to the fact that the mass overdensity δ is not necessarily equal to the overdensity in galaxies, δ_g. The velocity field samples the mass overdensity. So if we define the *bias*

$$b \equiv \frac{\delta_g}{\delta} \tag{9.48}$$

then $\vec{v} \propto \delta \propto \delta_g/b$. Therefore, the correction due to redshift space distortions in Eq. (9.47) is proportional to

$$\beta = \frac{f}{b} \simeq \frac{\Omega_m^{0.6}}{b}. \tag{9.49}$$

The redshift space distortion in the power spectrum, encoded in Eq. (9.47), is both good news and bad news. The good news is that by measuring the distortion in the redshift space power spectrum, we can hope to measure β, a combination of the density and bias. A quantitative way to do this is to measure the ratio of the quadrupole to the monopole of the power spectrum:

$$\frac{P_s^{(2)}(k)}{P_s^{(0)}(k)} \equiv \frac{5\int_{-1}^{1}\frac{d\mu_{\mathbf{k}}}{2}\mathcal{P}_2(\mu_{\mathbf{k}})P_s(\vec{k})}{\int_{-1}^{1}\frac{d\mu_{\mathbf{k}}}{2}\,\mathcal{P}_0(\mu_{\mathbf{k}})\,P_s(\vec{k})}. \tag{9.50}$$

Recall that $\mathcal{P}_l$ is the Legendre polynomial of order l, while P_s is the redshift-space power spectrum. Since (Exercise 11)

$$(1+\beta\mu_{\mathbf{k}}^2)^2 = \left[1+\frac{2}{3}\beta+\frac{1}{5}\beta^2\right]\mathcal{P}_0(\mu_{\mathbf{k}}) + \left[\frac{4}{3}\beta+\frac{4}{7}\beta^2\right]\mathcal{P}_2(\mu_{\mathbf{k}}) + \frac{8}{35}\beta^2\mathcal{P}_4(\mu_{\mathbf{k}}), \tag{9.51}$$

the orthogonality of the Legendre polynomials (Eq. (C.2)) implies that the ratio of the quadrupole to the monopole in linear theory is

$$\frac{P_s^{(2)}(k)}{P_s^{(0)}(k)} = \frac{\frac{4}{3}\beta+\frac{4}{7}\beta^2}{1+\frac{2}{3}\beta+\frac{1}{5}\beta^2}. \tag{9.52}$$

Figure 9.13 shows Hamilton's efforts to measure the quadrupole-to-monopole ratio in two different redshift surveys. In both cases, nonlinearities are very important

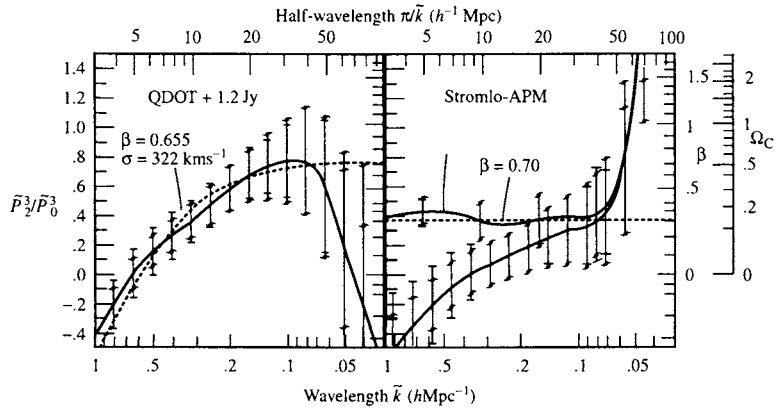

Figure 9.13. The quadrupole-to-monopole ratio for two redshift surveys (Hamilton, 1998). Left panel shows data from two redshift surveys which selected galaxies from the *Infrared Astronomical Satellite* (IRAS): QDOT (Lawrence *et al.*, 1999) picked one out of six galaxies brighter than 0.6 Jansky and the 1.2Jy survey (Strauss *et al.*, 1992) picked all galaxies above that brightness limit. Stromlo-APM survey (Loveday *et al.*, 1996) in right panel contains redshifts for 1 in 20 galaxies seen in APM (Figure 9.1).

and must be handled carefully. In the infrared-selected surveys depicted at left in the figure, Hamilton modeled the nonlinearities with a parameter measuring the small-scale velocity dispersion, σ. Only on scales of order $k \sim 0.1\,h$ Mpc^{-1} does the ratio begin to asymptote to ~ 0.75, implying a value of $\beta \sim 0.7$. The right panel shows a survey of optically selected galaxies. For these, the ratio does not seem to asymptote at all (lower curve) unless nonlinear structures are removed by hand (upper curve). In that case, Hamilton finds a quadrupole-to-monopole ratio closer to 0.4, implying $\beta = 0.3$. Figure 9.14 gives a broader view of the spread in measures of β from redshift and peculiar velocity surveys.

That was the good news. The bad news is that, even if we were to give up hope of measuring β from redshift surveys, we still need to account for redshift space distortions if we want to measure the power spectrum. Blindly measuring the power spectrum by averaging over all directions $\mu_{\mathbf{k}}$ is actually a measure of

$$P_s^{(0)}(k) = \left[1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right] P(k), \tag{9.53}$$

the equality holding only in linear theory. That is, $P_s^{(0)}$ overestimates the power spectrum by as much as a factor of 2. On even moderate scales, as suggested by Figure 9.13, nonlinear effects must be taken into account. The Peacock and Dodds compilation in Figure 7.11 uses a model of the small-scale velocities to do this.

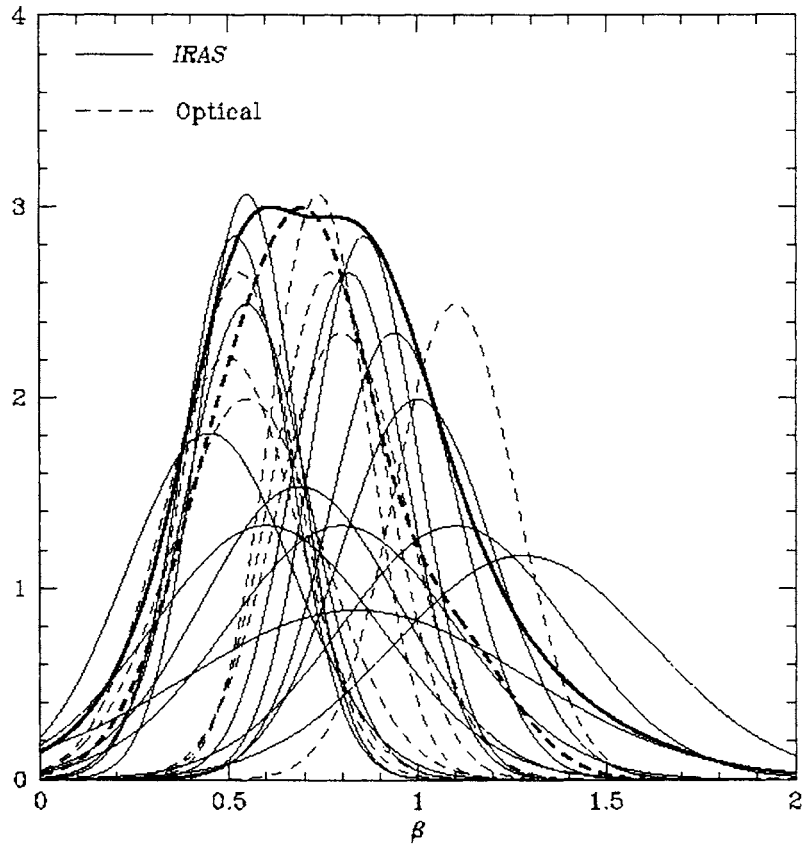

Figure 9.14. Compilation of the likelihood of β from redshift and peculiar velocity surveys (Strauss and Willick, 1995).

9.5 GALAXY CLUSTERS

Until now, we have focused solely on the two-point function: the angular correlation function, the velocity covariance, and the power spectrum. You might have wondered why little has been said about one-point functions, the number density of galaxies for example. To answer this question, first consider an extreme example. How many people are there in the universe? This clearly is an impossible question to answer with the tools we have developed. Putting aside the thorny question of the definition of a "person," we still would have to develop theories of star formation out of the gas in galaxies, then planet formation around stars, then the evolution of life via various biological processes. A prediction of the "person density" in the universe is beyond the scope of ... this book, to say the least.

In a similar, but less extreme, way, a prediction of the galaxy density of the universe from what we have learned about the distribution of matter in the universe

involves complicated issues we do not have the tools to address. What fraction of the matter has collapsed into nonlinear structures? How do these nonlinear structures evolve? Do galactic-size nonlinear structures merge? If so, how often? Upon collapse, how do stars form? How are stars distributed? In spiral patterns? Elliptical? One of the exciting developments of the 1990s was the evolution of a number of techniques to answer these questions. In addition to the brute-force approach of numerical simulations, several groups (e.g., Kauffman *et al.*, 1999; Somerville and Primack, 1999; Colberg *et al.*, 2000; Benson *et al.*, 2001; Cooray and Sheth, 2002) developed *semianalytic* techniques which have been remarkably successful at predicting properties and abundances of different galaxy types. Although we will not study these techniques directly here, the one technique we will encounter — the Press–Schechter formalism — forms the basis for much of this work.

Perhaps the fundamental difficulty encountered by one attempting to make predictions about the number density of galaxies is that the galactic scale has already gone nonlinear. Recall from Exercise 7.9 that scales smaller than $\sim 10\,h$ Mpc^{-1} have gone non-linear. What scale in the unperturbed universe encloses the mass of a typical galaxy, $M = 10^{12} M_\odot$? The density in a spherical region of radius R is

$$\rho_{\rm m} = \frac{M}{4\pi R^3/3}. \tag{9.54}$$

Since $\rho_{\rm m} = \Omega_m \rho_{\rm cr}$, we can invert this to find

$$R = 0.951\, h^{-1}\,{\rm Mpc} \left(\frac{Mh}{10^{12}\,\Omega_m\,M_\odot} \right)^{1/3}. \tag{9.55}$$

So a galaxy comes from matter within a radius of about 1 Mpc, corresponding to fluctuations on scales of order $k \sim 1\,h\,{\rm Mpc}^{-1}$, well into the nonlinear regime.

This answers a question ("Why not try to predict the number density of galaxies?") but begs another: Are there objects, corresponding to scales closer to the linear regime, for which one might be able to make reliable predictions about abundances? If we invert Eq. (9.55) to get the mass enclosed within a sphere of radius R,

$$M = 1.16 \times 10^{15} \Omega_m h^{-1} M_\odot \left(\frac{R}{10\ h^{-1}\ {\rm Mpc}} \right)^3, \tag{9.56}$$

then we see that clusters of galaxies — with masses up to $10^{15} M_\odot$ — arise from perturbations on just the right scales.

How, then, to predict the abundance of galaxy clusters? The basic insight comes from a paper by Press and Schechter (1974), and the resulting framework is called Press–Schechter theory. To understand their argument, consider the one-dimensional density field in Figure 9.15. The average inhomogeneity is zero, of course. There are regions with relatively large excursions in both the positive and negative direction. Underdensities cannot get smaller than -1 (when the density is zero), but there are some regions in the figure with densities more than three times the average density. It is these rare regions of large overdensity that we are

interested in: they *collapse*, accumulate so much excess matter that local gravity becomes more important than the Hubble flow. Particles in this region stop expanding away from each other and are trapped in the local gravitational field.

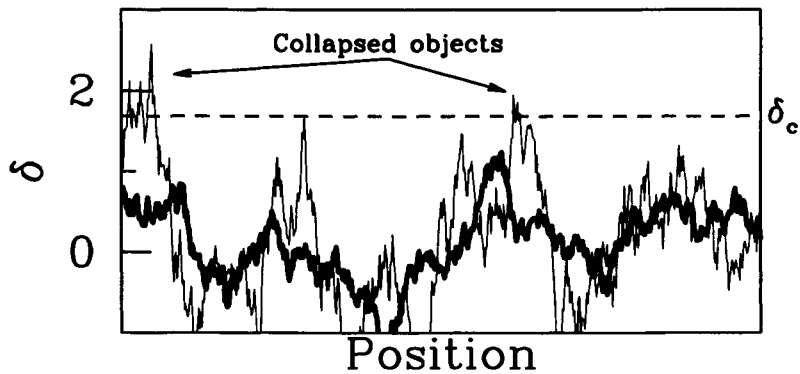

Figure 9.15. Inhomogeneities as a function of 1D position. Dark curve is the same field smoothed on larger scales. Several small-scale fluctuations have collapsed, while the large scale density field does not have any overdensities greater than δ_c, the critical value for collapse.

The Press–Schechter theory predicts the fraction of the volume that has collapsed,

$$f_{\rm coll}(M(R),z) = \frac{2}{\sqrt{2\pi}\sigma(R,z)} \int_{\delta_c}^{\infty} d\delta \; e^{-\delta^2/2\sigma^2(R,z)}. \qquad (9.57)$$

Here R is the radius over which the density field has been smoothed. This radius is used to compute $\sigma(R,z)$, the rms of the smoothed density field (Exercise 7.9). As you can see from Figure 9.15, the smoothing scale matters. Typically, inhomogeneities on large scales are smaller in magnitude than those on small scales, so small scales collapse first. As time evolves, overdensities grow (proportional to the growth function), so eventually some large-scale inhomogeneities will also collapse. The right-hand side of Eq. (9.57) counts all parts of the Gaussian distribution for which the overdensity is greater than some critical density δ_c. There are several pieces of magic in this formula. First, it assumes that the distribution of inhomogeneities is Gaussian. This is impossible since δ, by definition, never gets smaller than -1. And indeed, it is possible to show that gravity skews an initially Gaussian distribution, producing more underdense regions and a nonnegligible tail of highly overdense regions. Second, the normalization is a bit of a cheat: one would naively not include the factor of 2. Finally, the rms σ in the formula is the *linear* rms, specifically ignoring nonlinear effects. On small scales, there is a huge difference between σ calculated with the linear power spectrum and that with the true nonlinear spectrum. Press–Schechter tells us that the collapsed fraction can be obtained using the linear σ. These peculiarities of the Press–Schechter formalism do not detract from its effectiveness. Numerical simulations (e.g., White, Efstathiou and Frenk, 1993) have shown that it works extremely well. Further, a number of groups (Peacock and

Heavens, 1990; Bond *et al.*, 1991) have justified theoretically some of the aspects of the formula that initially appeared ad-hoc.

To get the collapsed fraction into a form more comparable with observations, first differentiate $f_{\rm coll}$ with respect to M and multiply by a small interval dM. This gives the fraction of the volume collapsed into objects with mass between M and $M+dM$. Multiply this by the average number density of such objects $\rho_{\rm m}/M$ to get the the number density of collapsed objects with mass between M and $M+dM$,

$$dn(M,z) = -\frac{\rho_{\rm m}}{M}\frac{df_{\rm coll}(M(R),z)}{dM}dM. \tag{9.58}$$

The minus sign appears here since $f_{\rm coll}$ is a decreasing function of the mass M. Carry out the derivative using the fact that $dM/dR = 3M/R$. Then,

$$\frac{dn(M,z)}{dM} = \sqrt{\frac{2}{\pi}}\frac{\rho_{\rm m}\delta_c}{3M^2\sigma}e^{-\delta_c^2/2\sigma^2}\left[-\frac{R}{\sigma}\frac{d\sigma}{dR}\right]. \tag{9.59}$$

The term in brackets here, the logarithmic derivative, is close to 1 for most models of interest. The dominant factor in Eq. (9.59), at least for large masses, is the exponential. If σ on a given scale is small, then the number density of collapsed objects on that scale is exponentially suppressed.

Until now, I have sidestepped the question of the numerical value of δ_c, the critical overdensity above which objects collapse. There are two approaches to obtaining δ_c; fortunately, both appear to agree. The first is to rely on a simple model of collapse, a model in which the overdensity is perfectly spherical. One can show that collapse occurs in such a model when $\delta = 1.686$ (for $\rho_{\rm m} = \rho_{\rm cr}$). The other way is simply to treat δ_c as a free parameter and calibrate it with numerical simulations (e.g., Eke, Cole and Frenk, 1996). The δ_c obtained this way is close enough to the spherical value that typically one simply adopts $\delta_c = 1.686$.

Measuring the cluster abundance, and therefore testing theories with Eq. (9.59), is a subtle business. One class of difficulties is identifying a cluster. Sophisticated algorthims have been developed to find clusters in a galaxy survey. The second set of difficulties revolves around determining the mass of the cluster. There are several methods of mass determination:

- **X-Ray Temperatures** Hot ionized gas in a cluster emit radiation with a cutoff frequency which is determined by the temperature of the gas. This temperature can be related to the mass of the cluster under certain assumptions.
- **Sunyaev–Zeldovich Distortion** Photons from the CMB pasing through clusters get scattered by the hot gas. This scattering distorts the CMB spectrum as a function of frequency, inducing a decrement at low frequency and excess at high frequency (low-energy photons gain energy from the hot electrons). The shape of the distortion is fixed; the amplitude is another measure of the temperature of the gas, which, again under certain assumptions, can be translated into a measurement of the mass.

- **Weak Lensing** Images of background galaxies are distorted by a foreground cluster. The larger the mass of the cluster, the larger the distortions. Weak lensing is therefore becoming a fabulous tool for measuring masses of clusters without using the temperature.

At least in the first two techniques, the direct measurement is of the cluster temperature. So let's work through a relation between the mass and temperature of a cluster under a simple set of assumptions. Suppose a cluster has virialized so that its kinetic energy is equal to minus half its potential energy. Suppose also that the cluster is spherical with radius $R_{\rm vir}$ with potential (gravitational) energy is equal to $-3GM^2/5R_{\rm vir}$. Then,

$$\frac{1}{2}Mv^2 = \frac{3}{10}\frac{GM^2}{R_{\rm vir}}. \tag{9.60}$$

The overdensity of the cluster $\Delta_{cl} \equiv \rho_{cl}/\rho_{\rm m}$ then allows us to eliminate the radius, since

$$\Delta_{cl} \equiv \frac{M}{4\pi R_{\rm vir}^3 \rho_{\rm m}/3}. \tag{9.61}$$

The temperature is equally apportioned among three directions, so the average velocity squared is $v^2 = 3T/m_p$, where m_p is the proton mass. The temperature can now be expressed in terms of the total mass of the system,

$$T = \frac{m_p}{5}\left[GMH_0\sqrt{\frac{\Delta_c}{2}}\right]^{2/3} \tag{9.62}$$

when $\Omega_m = 1$. Invert to get

$$M = 8.2 \times 10^{13} h^{-1} M_\odot \left(\frac{T}{\rm keV}\right)^{3/2} \sqrt{\frac{178}{\Delta_{cl}}}. \tag{9.63}$$

Here I have normalized Δ_{cl} by its value in the spherical collapse model (with $\rho_{\rm m} = \rho_{\rm cr}$), although again simulations have verified the numerical value.

Figure 9.16 shows the cluster density as a function of temperature for the standard CDM model with $\Omega_m = 1$. One of the most important points made immediately after the COBE detections of anisotropies in 1992 was that this plain-vanilla model predicts too many clusters. Indeed, the abundance of clusters today is often used as an excellent way of normalizing a power spectrum. A typical value for σ_8 from cluster abundances (e.g., Wang and Steinhardt, 1998) is

$$\sigma_8 = 0.5\Omega_m^{-.33-.35\Omega_m} \tag{9.64}$$

with error estimates ranging from just a few percent up to 20%, the latter probably more accurately reflecting uncertainties in the mass determinations.

Another exciting application of the Press–Schechter prediction for cluster abundances is the evolution with redshift. The basic point stems from the exponential

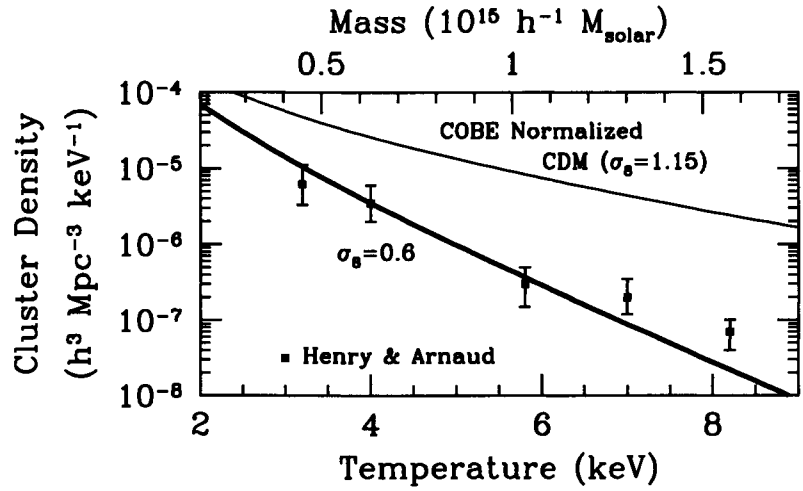

Figure 9.16. The cluster density as a function of temperature. Data from Henry and Arnaud (1991). The two theoretical curves are the Press–Schechter estimates (Eq. (9.59)) of models with $\Omega_m = 1, h = 0.5, \Omega_b = .05$. The only difference between the two is the normalization.

dependence in Eq. (9.59). At high redshift, $\sigma(R, z)$ is necessarily smaller, but how much smaller depends on the underlying cosmology. The growth is fastest in a model with $\Omega_m = 1$ (see Figure 7.12), so for a fixed abundance today, one expects many fewer clusters in such a model. Figure 9.17 illustrates this effect. An $\Omega_m = 1$ model predicts a factor of 700 fewer clusters with masses greater than $3.5 \times 10^{14} M_\odot$ at redshift 0.5 than at redshift zero. By contrast, ΛCDM predicts just a factor of 4 decline. With the observational assault on clusters just getting off the ground, we can expect strong constraints on cosmology to emerge in the coming decade.

SUGGESTED READING

The Large Scale Structure of the Universe (Peebles) has a good description and derivation of the angular correlation function, using Limber's (1953) original derivation. The derivation given in Section 9.1 is based on the Appendix of Kaiser's (1992) work on weak lensing. This derivation has the advantage of being physically intuitive and also generally applicable. It will allow us to easily compute the weak lensing correlation functions we encounter in Chapter 10.

The data discussed in Section 9.1 come from the APM Survey (Maddox *et al.*, 1990). Other recent angular results of note include the Edinburgh/Durham Southern Galaxy Catalogue (EDSGC, Collins, Nichol, and Lumsden, 1992) and the Sloan Digital Sky Survey (York *et al.*, 2000), the first analysis of which is presented in Scranton *et al.* (2002). As I hinted in the text, analysis of these data sets has gotten progressively more sophisticated over time. Baugh and Efstathiou (1993) first inverted the APM angular correlation function and extracted the 3D

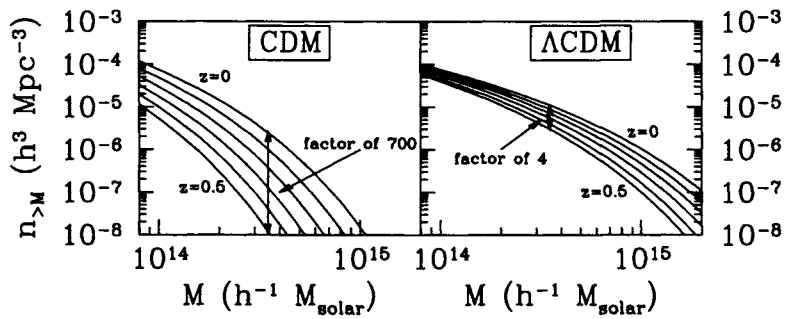

Figure 9.17. The number density of clusters with mass greater than M at different redshifts. *Left panel:* CDM model with $\Omega_m = 1, h = 0.5, \Omega_b = 0.05$ but normalized to give the correct abundance at $z = 0$ ($\sigma_8 = 0.5$). Curves give the abundances in redshift increments of $\delta z = 0.1$ starting from $z = 0$ going out to $z = 0.5$. *Right panel:* ΛCDM model which fits CMB and other data ($\Omega_m = 0.35, h = 0.7, \Omega_b h^2 = 0.02$). Note the relatively slow evolution of clusters with mass $\sim 3 \times 10^{14} M_\odot$ as compared with the critical density model.

power spectrum, confirming that it was quite different from the standard CDM power spectrum on large scales. Dodelson and Gaztanaga (2000) pointed out that the resultant errors on the power spectrum are correlated and that an accurate treatment would also account for the fact that the errors in $w(\theta)$ are also correlated. Eisenstein and Zaldarriaga (2001) and then Efstathiou and Moody (2001) accounted for these correlations leading to the softened conclusions summarized in Figure 9.6. The Eisenstein and Zaldarriaga paper also contains a clear discussion of the relation between $w(\theta)$ and the 2D power spectrum, or the C_l's. They observed that errors on the C_l's are much less correlated, which leads me to believe that C_l's will ultimtely replace $w(\theta)$ as the statistic of choice for angular surveys. Indeed, Huterer, Knox, and Nichol (2001) have analyzed the EDSGC survey with C_l's, and Tegmark *et al.* (2002) obtained C_l's from early SDSS data.

Cosmological Physics (Peacock) is a good resource for peculiar velocities and their effect on redshift surveys. Two important and informative review articles are Strauss and Willick (1995) and Hamilton (1998), the former particularly good for experimental issues involved in determining peculiar velocities and the latter for analyzing galaxy surveys in the presence of redshift space distortions. Another leader in the field, Dekel (1997), has written a good review of the cosmological implications of the peculiar velocity field.

The seminal work on redshift space distortions is by Kaiser (1987), who solved the problem working in Fourier space for linear distortions in the distant-observer, low redshift approximation. Hamilton (1992) found the analogue of this solution in real space. Recently, with large-area, relatively deep surveys coming on line, the generalization for cosmological corrections (i.e., distance is *not* equal to cz/H_0 as z gets large) and all-sky analysis has been carried out by a number of authors. As examples of the work currently going on in the field, Szalay, Matsubara, and Landy (1998) generalized Kaiser's work to large angles while Magira, Ying, and Suto (2000)

accounted for cosmological distortions and also nonlinearities and evolution. Peacock and Dodds (1994) analyzed a variety of surveys, arguing that — accounting for redshift space distortions, nonlinearities, and bias properly — the power spectrum has been well measured. Undoubtedly, with the upcoming Sloan Digital Sky Survey, Two Degree Field, and others, more will be learned about the power spectrum in coming years. The semianalytic work referred to on Page 283 is based on the seminal papers of White and Rees (1978) and White and Frenk (1991). Related, but separate form these semianalytic models, is the halo model (reviewed by Cooray and Sheth, 2002) which postulates that all the dark matter is in halos, thereby reducing the clustering problem to (i) the clustering of the halos and (ii) the distribution of matter and galaxies within the halo. Good descriptions of what the halo model is and how it can be used to compare theories with redshift surveys can be found in White (2001); Seljak (2000); and Berlind and Weinberg (2001).

The prediction of the cluster abundance is treated nicely in *Cosmological Physics* (Peacock) and *Cosmological Inflation and Large Scale Structure* (Liddle and Lyth). *Structure Formation in the Universe* (Padmanabhan) has a detailed section on the spherical collapse model, which is the source of the numbers 1.686 and 178 in Section 9.5. In addition to the papers cited in the text, some important cluster normalization papers are Viana and Liddle (1996, 1999) and Pierpaoli, Scott, and White (2001).

EXERCISES

Exercise 1. Suppose the correlation function is defined as

$$\xi(\vec{r}) \equiv \langle \delta(\vec{x})\delta(\vec{x}+\vec{r}) \rangle. \tag{9.65}$$

By Fourier expanding each of the δ's and using Eq. (C.20), show that this definition implies that the correlation function is the Fourier transform of the power spectrum.

Exercise 2. Expand the 3D power spectrum in the integral of Eq. (9.9) about $k_3 = 0$. The leading term is the one we considered. Show that the next term is of order $(1/l)^2$, compared with the leading term.

Exercise 3. Rewrite the kernel in Eq. (9.14) as an integral over χ. Show that F is a function of $k\theta$ only.

Exercise 4. Give an order-of-magnitude estimate for the kernel of the angular correlation function.
(a) Consider a shell in Fourier space with radius k and width dk. What fraction of the volume of this shell has $|k_3| < \chi^{-1}$?
(b) Argue that only Fourier modes with $|k_3| < \chi^{-1}$ contribute to the angular correlation function with a weight $\Delta^2(k) = k^3 P(k)/2\pi^2$. Combine this argument

with the fraction computed in part **(a)** to estimate the kernel relating $w(\theta)$ to $P(k)$. Compare this estimate with Eq. (9.13).

Exercise 5. Decompose the angular correlation function into a sum over spherical harmonics,

$$w(\theta) = \sum_l \frac{2l+1}{4\pi} C_l^{\text{matter}} \mathcal{P}_l(\cos\theta), \tag{9.66}$$

where the superscript $^{\text{matter}}$ distinguishes these C_l's from the ones characterizing anisotropies in the CMB, and $\mathcal{P}_l$ here are the Legendre polynomials. Express C_l^{matter} as an integral over the 3D power spectrum. Show that on small scales $C_l^{\text{matter}} = P_2(l)$, where P_2 is the 2D power spectrum introduced in Section 9.1.

Exercise 6. In Section 9.1 we implicitly neglected the evolution of the power spectrum. That is, we assumed that $P(k)$ remains constant with time. Allow $P(k)$ to scale as $(1+z)^\beta$. What is β for linear modes in a flat, matter-dominated universe? Rewrite the kernel in terms of an integral over z, accounting for this evolution.

Exercise 7. Compute (numerically) the linear growth rate f today in an open universe and compare with the approximation $\Omega_m^{0.6}$. What is the fractional error between the approximation and the exact result? Now assume that the universe is flat, with $\Omega_m + \Omega_\Lambda = 1$. Again compare the exact linear growth rate with $\Omega_m^{0.6}$. Show that

$$f = \Omega_m^{0.6} + \frac{\Omega_\Lambda}{70}\left(1 + \frac{\Omega_m}{2}\right) \tag{9.67}$$

is a better approximation, with no worse than 4% accuracy for $\Omega_m > 0.025$.

Exercise 8. Using CMBFAST, compute the transfer function for standard CDM ($\Omega_m = 1; h = 0.5$) with $\Omega_b = 0.01, 0.05$, and 0.1. Show that the BBKS transfer function is still a reasonable fit as long as

$$\Gamma = \Omega_m h \rightarrow \Omega_m h e^{-2\Omega_b}. \tag{9.68}$$

Exercise 9. Using CMBFAST or the BBKS transfer function, compute COBE-normalized σ_8 for ΛCDM with $h = 0.7$, $\Omega_\Lambda = 0.7$, $\Omega_m = 0.26$, and $\Omega_b = 0.04$. Locate the model on the bottom left panel of Figure 9.6. What does this imply about the relation between σ_8 (of the mass, which you have just computed) and $(\sigma_8)_g$ (of the galaxies, which APM is sensitive to)?

Exercise 10. Assume that the universe is flat with matter and a cosmological constant. Expand the comoving distance out to a galaxy at redshift z (neglecting peculiar velocities) about $z = 0$. The first-order term in the expansion should give back the redshift space answer. What is the second-order term, the leading correction to redshift space? Express your answer in terms of Ω_m.

Exercise 11. Derive (9.51), using the fact that $\mathcal{P}_4(x) = 35x^4/8 - 15x^2/4 + 3/8$. Show that the definition of the moments in Eq. (9.50) — $P_s^{(l)}(k) = (2l + 1)\int_{-1}^{1}(d\mu_{\mathbf{k}}/2)\mathcal{P}_l(\mu_{\mathbf{k}})P_s(k,\mu_{\mathbf{k}})$ — means that

$$P_s(k,\mu_{\mathbf{k}}) = \sum_l \mathcal{P}_l(\mu_{\mathbf{k}})P_s^{(l)}(k). \tag{9.69}$$

Exercise 12. In the text we showed how redshift space distortions affect the power spectrum. Show how the redshift space distortions affect the correlation function. Assume linear theory. You will probably need to consult Hamilton (1992), which transforms Kaiser's result to the correlation function in a single (dense) paragraph.