

Chapter 4

Spherical Redshift Distortions

Redshift surveys give a 3-dimensional picture of the galaxy distribution in which the redshift of a galaxy is taken as a measure of its distance, through Hubble's law. In addition to redshift, due to cosmological expansion, galaxies have "peculiar" velocities caused by gravitational (and perhaps other) interactions. This causes distortions in the catalog's redshift patterns. Kaiser (1987) pointed out that the distortions have a simple form in Fourier space. He showed that a wave with amplitude $\delta(\mathbf{k})$ has an amplification in redshift space of $1 + \beta\mu^2$. So

$$\delta^s(\mathbf{k}) = (1 + \beta\mu^2)\delta^r(k). \quad (4.1)$$

where the superscripts, s and r , refer to redshift and real space respectively (as they will throughout this chapter). Here μ is the cosine of the angle between the line of sight and the wavevector \mathbf{k} . $\beta = f/b$ is the ratio of the dimensionless linear growth rate (f) to the linear light-to-mass bias (b). In the linear regime, peculiar velocities (v) are related to overdensities (δ) by the continuity equation $\nabla \cdot \mathbf{v} + \beta\delta = 0$ (where $H = 1$). In standard pressure-less Friedmann cosmologies, the growth rate is just a function of Ω ,

$$f \approx \Omega_m^{\frac{4}{7}} + \frac{\Omega_\Lambda}{70} \left(1 + \frac{\Omega_m}{2}\right) \quad (4.2)$$

(Lahav et al., 1991) From equation 4.1 Kaiser concluded that since the Fourier transform of the correlation function, the power spectrum or $P(k) = \langle \delta(k'), \delta^*(\mathbf{k} + \mathbf{k}') \rangle$ is also amplified. In the plane-parallel limit this becomes:

$$P^s(\mathbf{k}) = (1 + \beta\mu^2)^2 P^r(k). \quad (4.3)$$

Simply stated, the plane-parallel approximation is the assumption that the angular separation of the galaxies is small enough such that their line-of-sight displacements are effectively parallel. This assumption was applied by many subsequent papers (Lilje and Efstathiou 1989; McGill 1990; Loveday et al. 1992; Hamilton 1992, 1993; Gramann et al. 1994; Bromley 1994; Fry and Gaztañaga 1994; Fisher et al. 1994; Cole et al. 1994, 1995). In order for the plane parallel approximation to be close to accurate one must only take into consideration pairs of galaxies which are separated by a small angle on the sky. Unfortunately, in order to obtain good statistics (to include the largest number of galaxy pairs), one would like to have this angle be as large as possible. Also, linear theory of clustering is most accurate in regions with small overdensities: the largest regions possible. Therefore, methods involving the plane parallel approximation must strike a compromise between angles which are too large and separations which are not large enough. For example, Hamilton (1993) and Cole et al. (1994) both took this angle to be ≤ 50 deg. Cole et al. (1994, fig. 8) show from simulations that the plane parallel approximation causes errors of about 5 per cent if the analysis is cut off at 50° .

In reality the distortions are not parallel but rather radially directed towards (or away from) the observer. A proper treatment of the spherical distortions should be superior to the plane parallel treatment in several ways. First, the same catalog can provide better statistics for large wavelengths. Second, larger wavelengths (regions where the linear clustering is more valid) can be included in the analysis. Finally, the systematic errors introduced by the plane parallel approximation can be eliminated completely.

Progress on spherical distortions began with Fisher et al. (1994). They expanded the density field of the 1.2-Jy survey into spherical harmonics, windowing the density in the radial direction with Gaussian windows at several depths. Heavens and Taylor (1995) expanded the radial direction in a complete set of spherical waves. Both of these methods required a prior shape of the power spectrum. Ballinger et al. (1995) allowed the power spectrum to vary in six bins thereby eliminating the prior assumption of the shape for the power spectrum. Hamilton and Tegmark (2000b) introduced several improvements, allowing them to measure a full, high

resolution, power spectrum and also measure β as a function of k . These brute force techniques have been used to analyze most of the current available catalogs (*e.g.*, Tadros et al. 1999, Hamilton and Tegmark 2000). The drawback to brute force techniques are that they require the assumption of Gaussian fluctuations (not true in the nonlinear regime) and computing costs constrain the brute force techniques to work on only the largest scales.

Hamilton and Culhane (1996) began working on a method for measuring the power spectrum and β using a fully spherically symmetric analysis that had the flexibility to be used in the nonlinear regime. In particular we derived a spherical distortion operator which when applied to the undistorted correlation function gives the distorted correlation function:

$$\xi^s(r_{12}, r_1, r_2) = D_{\text{dist}} \xi^r(r_{12}) \quad (4.4)$$

where

$$D_{\text{dist}} = \left[1 + \beta \left(\frac{\partial^2}{\partial r_1^2} + \frac{\alpha(r_1)}{r_1} \frac{\partial}{\partial r_1} \right) \nabla_1^{-2} \right] \left[1 + \beta \left(\frac{\partial^2}{\partial r_2^2} + \frac{\alpha(r_2)}{r_2} \frac{\partial}{\partial r_2} \right) \nabla_2^{-2} \right] \quad (4.5)$$

Here,

$$\alpha(r) \equiv \frac{\partial \ln r^2 \bar{n}^s(r)}{\partial \ln r} \quad (4.6)$$

where $\bar{n}^s(r)$ is the galaxy selection function in redshift space for the catalog in question. Looking at the distortion operator in equation 4.5 notice that the distortions depend only on β and the selection function of the catalog. One would like to use the properties of the distortion operator to find a way to measure β for each catalog. Hamilton and Culhane (1996) outlined a method for doing exactly that. However, in applying this method we found that some changes must be made in order to account for the properties of the individual catalog.

4.1 Spherical Shape Functions

In order to measure the real space power spectrum and β from a redshift survey it is necessary to have a model for the redshift distortions. In the Appendix, it is shown that, in

linear theory, the redshift space correlation function can be written as

$$\xi^s(r_{12}) = \sum_i g_i(\beta) B_i(r_{12}, r_1, r_2) \Xi_i(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12})) \quad (4.7)$$

with

$$\begin{aligned} \check{\xi} &\equiv \frac{2}{r_{12}^2} \int_0^{r_{12}} \xi(r) r dr, \\ \bar{\xi} &\equiv \frac{3}{r_{12}^3} \int_0^{r_{12}} \xi(r) r^2 dr, \\ \bar{\bar{\xi}} &\equiv \frac{5}{r_{12}^5} \int_0^{r_{12}} \xi(r) r^4 dr. \end{aligned} \quad (4.8)$$

Each of the B 's is a function only of the shape of the triangle formed by r_1 , r_2 , and r_{12} . Each of the measurable quantities

$$\check{\xi}_i(r_{12}) \equiv g_i(\beta) \Xi_i(r_{12}) \quad (4.9)$$

can be separated into a function of the separation (which can be transformed into $P(k)$) times a function only of β . The choice of B 's and $\check{\xi}$'s is not unique, however the sum in Equation 4.7 must be correct.

4.1.1 Importance of α

Hamilton and Culhane (1996) assumed that α could be treated as a constant, in particular, that

$$\int \alpha^i B_j(r_{12}, r_1, r_2) B_k(r_{12}, r_1, r_2) d^3 r_1 d^3 r_2 = \alpha^i M_{j,k} \quad (4.10)$$

for all i, j, k . In practice, the true α does not satisfy this equation. So then the question becomes: what is the meaning of α in the first place? We can see from the definition that α (Equation 4.6) measures the distortions caused by selection function of the catalog. Since each catalog has a different $\bar{n}(r)$ each catalog has a different $\alpha(r)$. Ideally, the distortions can be calculated exactly. In the above description of the method α is not specifically mentioned. However, in order to treat the effect of α properly it must be allowed to vary over r_1 or r_2 . In the method outlined

above the only functions allowed to vary over r_1 and r_2 are the shape functions ($B_i(r_{12}, r_1, r_2)$). So in order to treat α properly the shape functions must include α . Notice that the inclusion of α forces us to use a selection function with no radial jumps (except for when the selection function goes to zero).

4.1.2 Shape Functions

For our calculations we have chosen the following shape functions:

$$B_0 \equiv 1,$$

$$B_1 \equiv \left[\frac{\alpha(r_1)}{r_1} r_{12} \mu_1 + \frac{\alpha(r_2)}{r_2} r_{12} \mu_2 \right],$$

$$B_2 \equiv \frac{1}{2} [P_2(\mu_1) + P_2(\mu_2)],$$

$$B_3 = -\frac{2}{3} \mu_{12}^2 + \frac{2}{3} + \frac{1}{6} \frac{\alpha(r_1) \alpha(r_2) \mu_{12} r_{12}^2}{r_1 r_2} + \frac{1}{6} \frac{\alpha(r_1) \alpha(r_2) \mu_1 \mu_2 r_{12}^2}{r_1 r_2} - \frac{1}{2} \left[\frac{\alpha(r_1) r_{12}}{r_1} \mu_1 \mu_2^2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1^2 \mu_2 \right] - \frac{1}{3} \left[\frac{\alpha(r_1) r_{12}}{r_1} \mu_{12} \mu_2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1 \mu_{12} \right],$$

$$B_4 \equiv \frac{1}{8} (35 \mu_1^2 \mu_2^2 - 15 \mu_1^2 - 15 \mu_2^2 + 3),$$

$$B_5 = \frac{4}{5} \mu_{12}^2 - \frac{4}{5} - \frac{1}{30} \frac{\alpha(r_1) \alpha(r_2) \mu_{12} r_{12}^2}{r_1 r_2} - \frac{1}{10} \frac{\alpha(r_1) \alpha(r_2) \mu_1 \mu_2 r_{12}^2}{r_1 r_2} + \frac{1}{2} \left[\frac{\alpha(r_1) r_{12}}{r_1} \mu_1 \mu_2^2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1^2 \mu_2 \right] + \frac{1}{5} \left[\frac{\alpha(r_1) r_{12}}{r_1} \mu_{12} \mu_2 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_1 \mu_{12} \right] - \frac{1}{10} \left[\frac{\alpha(r_1) r_{12}}{r_1} \mu_1 + \frac{\alpha(r_2) r_{12}}{r_2} \mu_2 \right],$$

$$B_6 = -\frac{\alpha(r_1) \alpha(r_2) \mu_{12} r_{12}^2}{r_1 r_2}. \quad (4.11)$$

This choice of shape functions leads to the following expressions for $\check{\xi}_i$:

$$\check{\xi}_0 \equiv \left(1 + \frac{2}{3}\beta + \frac{1}{5}\beta^2\right)\xi(r_{12}),$$

$$\check{\xi}_1 \equiv \left[\frac{1}{3}\beta + \frac{1}{6}\beta^2\right]\bar{\xi}(r_{12}), \quad (4.12)$$

$$\check{\xi}_2 \equiv \left(\frac{4}{3}\beta + \frac{4}{7}\beta^2\right) [\xi(r_{12}) - \bar{\xi}(r_{12})], \quad (4.13)$$

$$\check{\xi}_3 \equiv \beta^2\bar{\xi}(r_{12}), \quad (4.14)$$

$$\check{\xi}_4 \equiv \frac{8}{35}\beta^2 \left[\xi(r_{12}) + \frac{5}{2}\bar{\xi}(r_{12}) - \frac{7}{2}\bar{\bar{\xi}}(r_{12}) \right], \quad (4.15)$$

$$\check{\xi}_5 \equiv \beta^2\bar{\bar{\xi}}(r_{12}), \quad (4.16)$$

$$\check{\xi}_6 \equiv \frac{\beta^2}{6}\check{\xi}(r_{12}). \quad (4.17)$$

4.2 Generalized Fourier Transform Windows

In Cole et al. (1995), Hamilton (1992) and Hamilton and Culhane (1996) it was stated that the functions $\check{\xi}_l^s(r_{12})$ (where $l = 0, 2, 4$) can be transformed into a function of β times $P(k)$ using Bessel functions:

$$P(k) = 4\pi \int \Xi_{0,2,4}^s(r_{12}) j_{0,2,4}(kr_{12}) r_{12}^2 dr_{12}. \quad (4.18)$$

Unfortunately, in the spherically symmetric case a single Bessel function is not sufficient for some of the $\Xi_l^s(r_{12})$ (where $l \neq 0, 2, 4$). However, there are functions which will serve the same purpose.

The first thing to notice is that the window over k will not effect β . This means that the function for $\check{\xi}^s$ needs to be a function of β times the function of the averages of $\xi^r(r)$. This

is the reason for setting $\check{\xi}_i = g_i(\beta)\Xi_i$. In other words, take the function $g_i(\beta)$ outside of the windowing integral. Assume that

$$4\pi \int \Xi_i^s(r_{12})\check{j}_l(kr_{12})r_{12}^2 dr_{12} = 4\pi \int f_l(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12}))\check{j}_l(kr_{12})r_{12}^2 dr_{12} = P(k) \quad (4.19)$$

in order to find $\check{j}_l(kr_{12})$. In this equation $\check{j}_l(kr_{12})$ is the generalized window. In the cases of ($l = 0, 2, 4$) $\check{j}_l(kr_{12}) = 4\pi i^l j_l(kr_{12})$ where $j_l(kr_{12})$ is the usual spherical Bessel function. Since this is not true for all i we want to find $\check{j}_l(kr_{12})$ such that:

$$4\pi \int f_l(\xi(r_{12}), \check{\xi}(r_{12}), \bar{\xi}(r_{12}), \bar{\bar{\xi}}(r_{12}))\check{j}_l(kr_{12})r_{12}^2 dr_{12} = 4\pi \int \xi(r_{12})j_0(kr_{12})r_{12}^2 dr_{12}. \quad (4.20)$$

To find the appropriate windows solve this equation for each value of l . First, choose functions f_l in such a way to make the calculations of the windows as simple as possible. The above separation leaves $f_l, l \neq 0, 2, 4$ to be a function of only one of the averages of $\xi^r(r_{12})$.

$$\int_0^\infty \left[\frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1} \xi^r(r) dr \right] \check{j}_l(kr_{12})r_{12}^2 dr_{12} = \int_0^\infty \xi^r(r_{12})j_0(kr_{12})r_{12}^2 dr_{12}. \quad (4.21)$$

Now switch the order of integration of the left hand side of equation 4.21:

$$\int_0^\infty n r^{n-1} \left[\int_r^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} \right] \xi^r(r) dr = 4\pi \int_0^\infty \xi^r(r_{12})j_0(kr_{12})r_{12}^2 dr_{12}. \quad (4.22)$$

Rename the integration variables so that the semi-infinite integrals are over the same variable:

$$\int_0^\infty n s^{n-1} \left[\int_s^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} \right] \xi^r(s) ds = \int_0^\infty \xi^r(s)j_0(ks)s^2 ds. \quad (4.23)$$

One solution, perhaps the most restrictive solution, to this equation is found by equating the integrands of the two sides of the equation:

$$n s^{n-1} \left[\int_s^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} \right] \xi^r(s) = \xi^r(s)j_0(ks)s^2. \quad (4.24)$$

This can be further simplified to

$$\int_s^\infty r_{12}^{2-n} \check{j}_l(kr_{12}) dr_{12} = \frac{s^{3-n}}{n} j_0(ks). \quad (4.25)$$

To find the new window differentiate equation 4.25 with respect to s . This gives:

$$-s^{2-n}\check{j}_l(ks) = (-ksj_1(ks) + (3-n)j_0(ks))\frac{s^{2-n}}{n}. \quad (4.26)$$

The final form of the windows for functions of just one of the averages of ξ^r (we have employed recursion relations for spherical Bessel functions so that only even functions remain) is:

$$\check{j}_l(ks) = \frac{1}{n} \left(\frac{k^2 s^2}{3} j_0(ks) + \frac{k^2 s^2}{3} j_2(ks) + (n-3)j_0(ks) \right). \quad (4.27)$$

So estimates of $\xi_\alpha(r)$ can be transformed into estimates of $f_i(\beta)\xi(k)$.

It turns out that the inverse relation is equally (if not more) important. Fortunately, the inverse window functions are even easier to calculate. The equation

$$\frac{4\pi}{(2\pi)^3} \int_0^\infty \tilde{j}_n(kr_{12})P(k)k^2 dk = \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1}\xi(r)dr \quad (4.28)$$

needs to be solved for $\tilde{j}_n(kr_{12})$. However, because

$$\frac{4\pi}{(2\pi)^3} \int_0^\infty j_0(kr)P(k)k^2 dk = \xi(r) \quad (4.29)$$

equation 4.28 can be rewritten

$$\frac{4\pi}{(2\pi)^3} \int_0^\infty \tilde{j}_n(kr_{12})P(k)k^2 dk = \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1} \left[\frac{4\pi}{(2\pi)^3} \int_0^\infty j_0(kr)P(k)k^2 dk \right] dr \quad (4.30)$$

or

$$\tilde{j}_n(kr_{12}) = \frac{n}{r_{12}^n} \int_0^{r_{12}} r^{n-1} j_0(kr) dr. \quad (4.31)$$

Now the Ξ 's can be transformed to and from $P(k)$.

4.3 Pair Weight Compression with Spherical Redshift Distortions

In chapter 2 we showed how to extract power spectra from redshift surveys disregarding redshift distortions. However, the pair weight compression method becomes more powerful when redshift distortions are explicitly included. Not only can the real space power spectrum be extracted, but it can also yield an estimate of β .

First of all, look at the resulting estimate of $\hat{\xi}_\gamma$ (Equation 2.20):

$$\hat{\xi}_\gamma = M_{\gamma\alpha}^{-1} B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} \delta_k \delta_l \quad (4.32)$$

where

$$M_{\alpha\gamma} \equiv B_{\alpha ij} W_{aij} K_{ab}^{-1} W_{bkl} B_{\gamma kl} \quad (4.33)$$

and

$$K_{ab} \equiv W_{aij} \langle \mathbf{e}_{ijkl} \rangle W_{bkl} \approx 2 W_{aij} C_{ik} C_{jl} W_{bkl}. \quad (4.34)$$

No major modifications of these formulae are necessary to explicitly include redshift distortions. However, the meanings of some of these symbols change.

4.3.1 Shape Functions

The most necessary change comes from the definition of B . Recall from chapter 2 that B comes from taking the derivative of $\langle \delta_i \delta_j \rangle$ with respect to the parameters we wish to measure. In the case where redshift distortions are ignored, the only parameters of interest were the value of the power spectrum at various values of k . Now that redshift distortions are included

$$\langle \delta_i \delta_j \rangle = \xi_{ij} + \bar{n}^{-1}(\mathbf{r}_i) \delta_{ij} = \sum_{\alpha} \check{\xi}_{\alpha} B_{\alpha ij} + \bar{n}^{-1}(\mathbf{r}_i) \delta_{ij}. \quad (4.35)$$

Now each value of $\check{\xi}_{\alpha}$ (after transforming into Fourier space) gives a measurement of the power spectrum at a particular wavenumber times some function of β . This means that the information about P_k and $\beta_{(k)}$ is contained within the parameters $\check{\xi}_{\alpha}$. Of course, one would like to go straight to measuring the actual values of P_k and β with no intermediate steps. However, because the data lie in real space, the direct measurement of P_k and β is infeasible. One can, however, measure the parameters $\check{\xi}_{\alpha}$, transform these parameters into Fourier space, and then do a maximum likelihood calculation using these parameters (and the associated error bars) to find the values of P_k and β .

So differentiating Equation 4.35 with respect to $\check{\xi}_\alpha$ yields just the shape functions $B_{\alpha ij}$. Actually, the differentiation occurs at a particular separation so the shape functions also include a delta function:

$$\frac{\partial \langle \delta_i \delta_j \rangle}{\partial \check{\xi}_\alpha} = \delta(r_\alpha - |\mathbf{r}_i - \mathbf{r}_j|) B_{\alpha ij}. \quad (4.36)$$

Notice that the subscript α stands for the subscript of the particular shape function at the particular separation.

4.3.2 Window Functions

At each separation, instead of measuring one parameter, there are now seven parameters to measure. Because there are seven parameters to measure it makes sense to use at least seven window functions at each separation. In the case of the plane-parallel approximation the window functions to use were obvious. The three different Ξ 's were orthogonal to one another. Each window was tuned to measure exactly one of the parameters of interest. This meant that the data needed to be weighted by the Legendre polynomials of order 0, 2 and 4. In the case of spherical distortions the parameters are not orthogonal to one another. However, we can still make an analogous choice of window functions. In other words, use window functions like those described in chapter 2 weighted by the shape functions. This should be a good choice because each window function will be picking out a majority of the information about a particular parameter. Adding additional windows will again raise the amount of information available to the parameter estimation. However, because the computation is so lengthy as it is, we want to use the smallest possible number of parameters.

4.3.3 Better Final Parameters

The seven Ξ 's are the natural parameters to measure. However, the physical meaning of some of the Ξ 's is not so clear. The Fisher matrix representation allows us to take a linear combination of the measured parameters and form a more physically motivated set of parameters.

To do this notice that

$$F_{\theta_1, \theta_2} = \frac{\partial \phi_1}{\partial \theta_1} F_{\phi_1, \phi_2} \frac{\partial \phi_2}{\partial \theta_2} \quad (4.37)$$

and

$$\Theta_i = \frac{\partial \theta_i}{\partial \phi_j} \Phi_j. \quad (4.38)$$

However, it is common to have Φ as a function of Θ rather than the other way around. In this case, the quantity $\frac{\partial \theta_i}{\partial \phi_j}$ is unknown. In fact, if there are fewer values of Θ than Φ then Θ may be over-determined. This means that the best choice is the one where the values of Θ , as determined by Φ , make the best use of the data.

The best use of the data can be obtained by looking at the log-likelihood function

$$-\ln \mathcal{L} = \frac{1}{2} (\Phi_i - \hat{\Phi}_i) F_{\phi_i \phi_j} (\Phi_j - \hat{\Phi}_j) \quad (4.39)$$

which can be rewritten in terms of θ :

$$-\ln \mathcal{L} = \frac{1}{2} (\Phi_i - \frac{\partial \phi_i}{\partial \theta_k} \hat{\Theta}_k) F_{\phi_i \phi_j} (\Phi_j - \frac{\partial \phi_j}{\partial \theta_l} \hat{\Theta}_l). \quad (4.40)$$

Setting the first derivative of $-\ln \mathcal{L}$ with respect to $\hat{\Theta}$ equal to zero gives the best estimate of $\hat{\Theta}$:

$$\frac{\partial(-\ln \mathcal{L})}{\partial(\hat{\Theta}_k)} = -\frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} (\Phi_j - \frac{\partial \phi_j}{\partial \theta_l} \hat{\Theta}_l) = 0. \quad (4.41)$$

Rearranging this equation yields:

$$\frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} \Phi_j = \frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} \frac{\partial \phi_j}{\partial \theta_l} \hat{\Theta}_l. \quad (4.42)$$

Using equation 4.37 gives:

$$\frac{\partial \phi_i}{\partial \theta_k} F_{\phi_i \phi_j} \Phi_j = F_{\theta_k \theta_l} \hat{\Theta}_l. \quad (4.43)$$

This means that

$$\Theta_i = F_{\theta_i \theta_j}^{-1} \frac{\partial \phi_k}{\partial \theta_j} F_{\phi_k \phi_l} \Phi_l \quad (4.44)$$

is the best representation of Θ given Φ . Notice that if there are as many values of Θ as there are of Φ then equation 4.44 collapses to the expected:

$$\Theta_i = \left(\frac{\partial\phi_j}{\partial\theta_i}\right)^{-1}\Phi_j. \quad (4.45)$$

Using this formalism, each of the spherical distortion parameters Ξ can be expressed as a function of $P(k)$, $\beta P(k)$ and $\beta^2 P(k)$. These are parameters that make more physical sense. The first is just the galaxy-galaxy power spectrum. The second is the galaxy-velocity power spectrum. The third is the velocity-velocity power spectrum Hamilton and Tegmark (2000b). Not only are these more physically motivated, but since we are ultimately trying to measure the power spectrum and β , this gives us a small number of parameters with which to do a non-linear χ^2 fit.

4.3.4 Discreteness Correction

In Chapter 3, we discussed the fact that a discrete representation of the correlation function can cause errors in the measurement of the power spectrum. This is true for the case with redshift distortions as well. Upon transforming the prior value of the Ξ 's (using the proper generalized FFT window) the result should be the prior power spectrum. Unfortunately, for any realistic prior power spectrum, a discrete FFT of the Ξ 's does not yield exactly the prior power spectrum. To correct for this, it is necessary to multiply the result of each FFT by the discreteness correction (ζ_i):

$$\zeta_i(k) = \frac{P_{\text{prior}}(k)}{\sum_i \Xi_{i\text{prior}}(r) \tilde{j}_i(r)}. \quad (4.46)$$

This means that, after the correction, if the measurement of Ξ corresponds exactly to the prior value of Ξ then the measurement of $P(k)$ will be the prior value of the power spectrum, as expected.

4.3.5 Correlation and Decorrelation

As in Chapter 3, the result from equation 4.32 is a measurement of the parameter in question with no width in k . This result actually anti-correlates neighboring points. That is to say, the cross-covariance of two neighboring points will be negative. This is not an ideal way to present the data, particularly because we expect neighboring points to have similar values. The two ways that are sensible are the correlated and decorrelated versions. The decorrelated version has the advantage of having each estimate being uncorrelated with any other estimate. This means that the decorrelated points can be used in a least squares fit as a simple sum of the deviation from each estimate. For obtaining the best least-squares result, the decorrelated estimates are clearly the best choice.

The correlated estimates give results which are easy to visually compare one set of results with another. That is to say, that when using correlated points, the scatter is small enough that the most likely value of the parameter is easily determined from a graph.

The method for obtaining the correlated and decorrelated estimates were given in Chapter 3. There is one major additional concern when using the redshift distortion analysis. This is that the decorrelation (correlation) matrix naturally combines some of the $P(k)$ estimate in with pieces of the $\beta P(k)$ and the $\beta^2 P(k)$ terms. A representative row of the Fisher matrix is shown in Figure 4.1. In going from the anti-correlated version of the estimates, it would be nice to avoid entangling one type of estimate with another.

To disentangle the three estimates from one another first define \hat{q} (here the hat denotes a measurement):

$$\langle \hat{q}_\alpha \rangle = \mathbf{F}_{\alpha\gamma} \xi_\gamma. \quad (4.47)$$

where \mathbf{F} is just the Fisher matrix and ξ_γ is one of the new parameters $P(k)$, $\beta P(k)$, or $\beta^2 P(k)$. Notice that the Fisher matrix smoothes over all three parameters over all wavenumbers. To

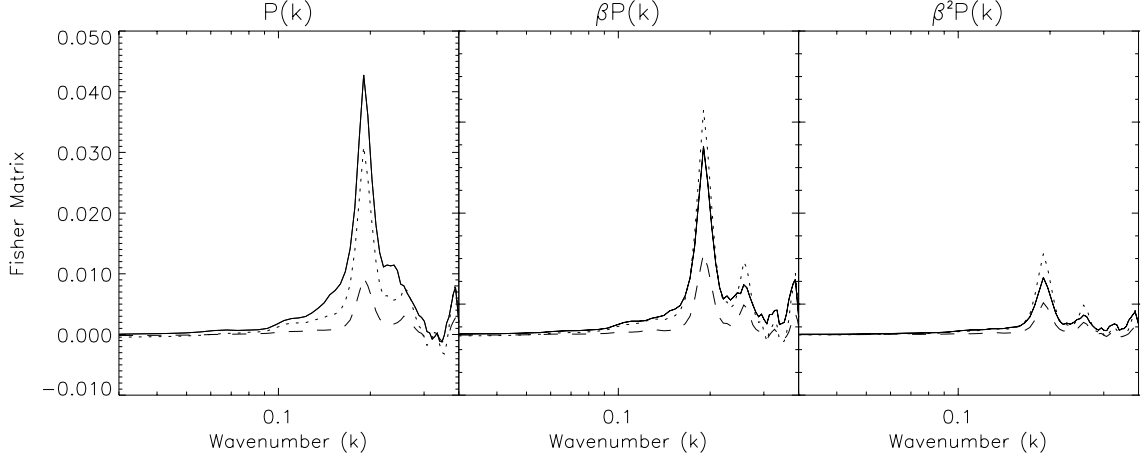


Figure 4.1: The rows of the Fisher matrix are one of the smoothing windows. Plotted here are rows corresponding to the $P(k)$ term (solid) the $\beta P(k)$ term (dotted) and the $\beta^2 P(k)$ term (dashed). Each row is peaked about the nominal wavelength and contains contributions from each of the three terms.

explicitly disentangle the three parameters rewrite equation 4.47:

$$\langle \hat{q}_a(k) \rangle = \sum_b \left[\frac{\langle \hat{\xi}_b(k) \rangle}{\xi_b(k)} \right] \sum_k \mathbf{F}_{ab}(k, k') \xi_b(k'). \quad (4.48)$$

Notice that the Greek indices have been expanded into Latin indices over the type of parameter and k the wavenumber. Also the summations are now explicitly included because some of the following equations are not summed. The term in square brackets is added to form the definition of \mathcal{F}_{ab} :

$$\mathcal{F}_{ab}(k) \equiv \frac{1}{\xi_b(k)} \sum_k \mathbf{F}_{ab}(k, k') \xi_b(k') \quad (4.49)$$

with no summation over b . This allows 4.48 to be rewritten

$$\langle \hat{q}_a(k) \rangle = \sum_b \mathcal{F}_{ab}(k) \langle \hat{\xi}_b(k) \rangle \quad (4.50)$$

with no summation over k . Now, $\mathcal{F}_{ab}(k)$ is a 3×3 matrix at each value of k . To obtain a measurement of each of the ξ 's multiply both sides of equation 4.50 by the inverse of $\mathcal{F}_{ab}(k)$:

$$\xi_b(k) = \sum_a \mathcal{F}_{ba}^{-1}(k) \langle \hat{q}_a(k) \rangle \quad (4.51)$$

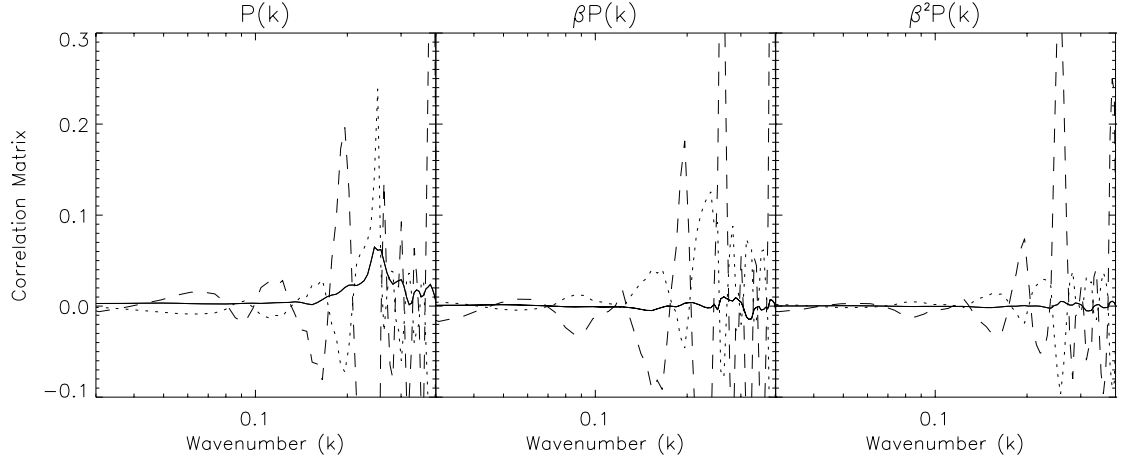


Figure 4.2: Representative rows of the correlation matrix. Each line shows the window (one of the rows of \mathcal{C}) over which one of the parameters has been smoothed to obtain a correlated estimate. These are the same rows that are plotted in Figure 4.1. The $P(k)$, $\beta P(k)$ and $\beta^2 P(k)$ rows are shown by the solid, dotted and dashed lines respectively.

or

$$\hat{\xi}_b(k) = \sum_a \mathcal{F}_{ba}^{-1}(k) \hat{q}_a. \quad (4.52)$$

Equation 4.52 gives the estimate of $\xi_b(k)$ which has an expectation value of the prior power spectrum at the given wavelength. Furthermore, each of the parameters has no contribution from the others (provided that the true power spectrum is similar in shape to the prior power spectrum). When starting with the anti-correlated estimate of $\hat{\xi}_b$ (i.e., $\mathbf{F}_{ab}^{-1} \hat{q}_b$), to obtain the untangled and correlated measurement of ξ use the band-power window matrix \mathcal{C}_{ab} :

$$\hat{\xi}_a^{\text{corr}}(k) = \sum_b \mathcal{F}_{ab}^{-1}(k) \sum_{c,k'} \mathbf{F}_{bc}(k, k') \hat{\xi}_c(k') \equiv \sum_{c,k'} \mathcal{C}_{ac}(k, k') \hat{\xi}_c(k') \quad (4.53)$$

where

$$\mathcal{C}_{ac}(k, k') \equiv \sum_b \mathcal{F}_{ab}^{-1}(k) \mathbf{F}_{bc}(k, k'). \quad (4.54)$$

Figure 4.2 shows rows of the correlation matrix (\mathcal{C}). Notice that each of the parameters is nicely peaked about the nominal wavenumber within the same type of parameter. Similarly,

the $P(k)$ term is nicely zero in the other parameters. However, the contributions to the terms $\beta P(k)$ and $\beta^2 P(k)$ from the $P(k)$ term can be quite large if the power spectrum is not similar in shape to the prior power spectrum.

To obtain the covariance matrix of the correlated estimates take

$$\langle \Delta \hat{\xi}_a^{\text{corr}}(k) \Delta \hat{\xi}_b^{\text{corr}}(k') \rangle = C_{ac}(k, k'') \langle \Delta \hat{\xi}_c(k'') \Delta \hat{\xi}_d(k''') \rangle C_{bd}(k', k'''). \quad (4.55)$$

Next recognize that by the definition of the Fisher matrix

$$\langle \Delta \hat{\xi}_c(k'') \Delta \hat{\xi}_d(k''') \rangle = \mathbf{F}_{cd}^{-1}(k'', k'''), \quad (4.56)$$

and from the definition of \mathcal{C} (Equation 4.54):

$$\langle \Delta \hat{\xi}_a^{\text{corr}}(k) \Delta \hat{\xi}_b^{\text{corr}}(k') \rangle = \mathcal{F}_{ae}^{-1}(k) \mathbf{F}_{fe}(k, k') \mathcal{F}_{bf}^{-1}(k') \quad (4.57)$$

with implicit summation over each pair of Latin indices and no summation over k or k' .

Notice that if the smoothing matrix is instead defined as (Hamilton and Tegmark, 2000b)

$$\mathbf{M} = \mathbf{F}^{1/2} \quad (4.58)$$

then the covariance matrix is then the unit matrix. Here the square-root of a matrix is any matrix that satisfies:

$$\mathbf{F} = \mathbf{M}^\top \mathbf{M}. \quad (4.59)$$

For example, let

$$\hat{z}_\alpha \equiv \mathbf{M}_{\alpha\beta} \hat{\xi}_\beta \quad (4.60)$$

then

$$\langle \Delta \hat{z}_\alpha \Delta \hat{z}_\beta \rangle = \mathbf{1}_{\alpha\beta} \quad (4.61)$$

where $\mathbf{1}$ is the unit matrix. This allows the estimates to be decorrelated from one another.

Once again, however, it would be better if the estimate of a particular parameter were disentangled from the other parameters. To do this, rewrite

$$\langle \hat{z}_a(k) \rangle = \sum_b \left[\frac{\langle \hat{\xi}_b(k) \rangle}{\xi_b(k)} \right] \sum_k \mathbf{M}_{ab}(k, k') \xi_b(k'). \quad (4.62)$$

So defining \mathcal{M} :

$$\mathcal{M}_{ab}(k) \equiv \frac{1}{\xi_b(k)} \sum_k \mathbf{M}_{ab}(k, k') \xi_b(k') \quad (4.63)$$

again with no summation over b . This allows 4.62 to be rewritten

$$\langle \hat{z}_a(k) \rangle = \sum_b \mathcal{M}_{ab}(k) \langle \hat{\xi}_b(k) \rangle. \quad (4.64)$$

Now to disentangle the decorrelated parameters multiply each side by the inverse of \mathcal{M} :

$$\hat{\xi}_\beta^{\text{decorr}}(k) = \mathcal{M}_{\beta\alpha}^{-1}(k) \hat{z}_\alpha(k). \quad (4.65)$$

Notice that by disentangling the different parameters at each wavenumber the three parameters at each wavelength are no longer decorrelated from one another. However, all parameters at each wavenumber remain decorrelated from all parameters at every other wavenumber. Specifically,

$$\langle \Delta \hat{\xi}_a^{\text{decorr}}(k) \Delta \hat{\xi}_b^{\text{decorr}}(k) \rangle = \mathcal{M}_{ac}^{-1} \mathcal{M}_{bc}^{-1} \quad (4.66)$$

with summation over c only.

Figure 4.3 shows the representative decorrelated smoothing windows. Notice that in each case the smoothing window of the relevant parameter is sharply peaked about the nominal wavenumber. Also the contribution to $P(k)$ from the $\beta P(k)$ and $\beta^2 P(k)$ terms is minimal as is the contribution to the $\beta P(k)$ term from the $\beta^2 P(k)$ term.

4.4 Information in LCRS

Figure 4.4 shows the information contained within each of the decorrelated parameters. LCRS has the highest amount of information for the redshift space power spectrum. It is clear

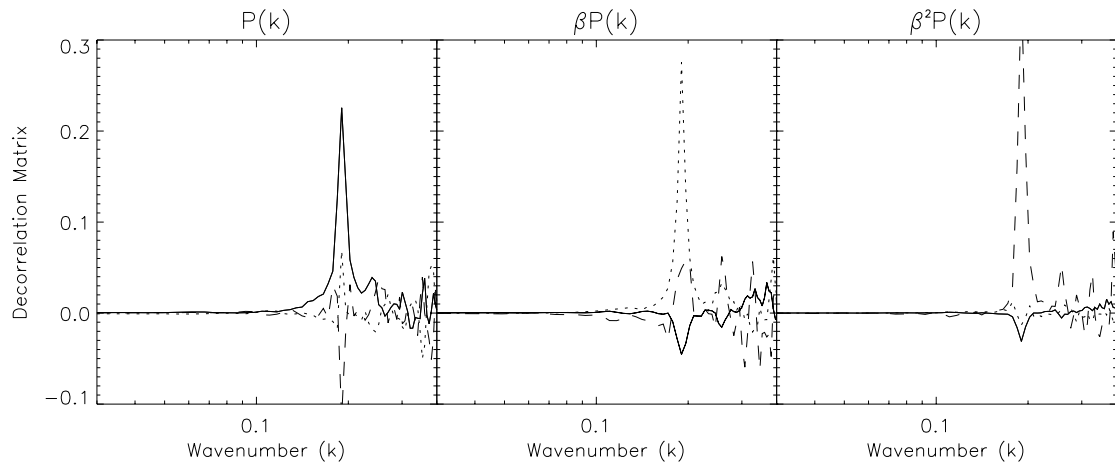


Figure 4.3: Representative rows of the decorrelation matrix. Each line shows the window (one of the rows of \mathcal{D}) over which one of the parameters has been smoothed to obtain a decorrelated estimate. These are the same rows that are plotted in Figure 4.1 and Figure 4.2. The $P(k)$, $\beta P(k)$ and $\beta^2 P(k)$ rows are shown by the solid, dotted and dashed lines respectively.

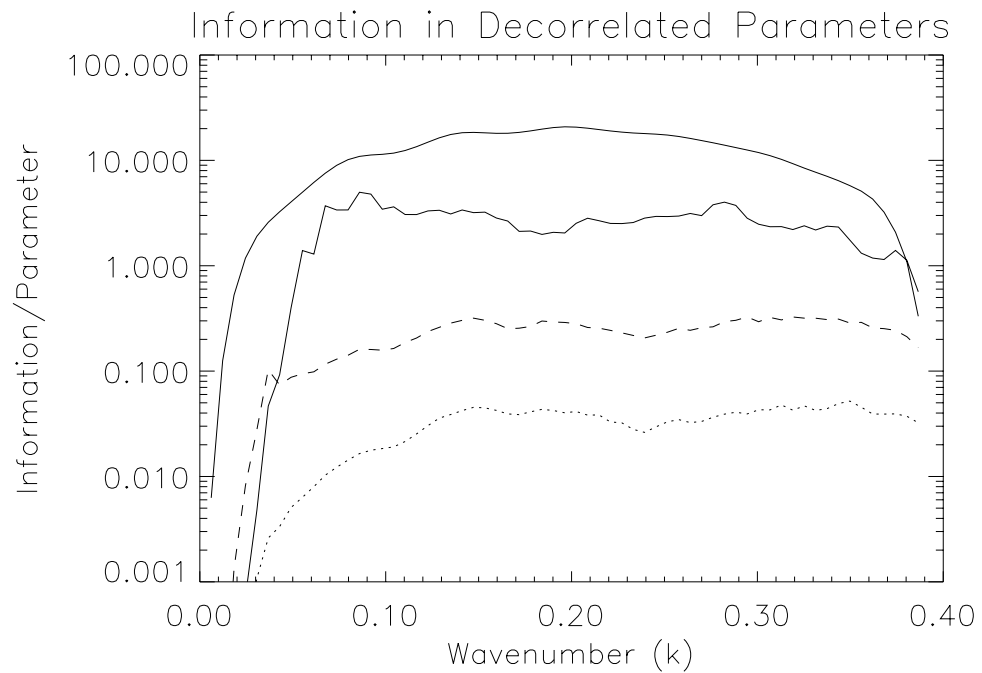


Figure 4.4: Information in the decorrelated parameters. The upper solid line is for the redshift space power spectrum. The lower solid line is for the real space power spectrum. The dashed line is for $\beta P(k)$ and the dotted line is for $\beta^2 P(k)$.

that this should always be the case. Upon adding additional measured parameters, the information per parameter should go down.

In this case the total amount of information appears to be reduced by at least a factor of two. Clearly, the information has not been apportioned out to the various terms. It appears that the terms containing β do not contain a large fraction of the information. There has been an actual leak in the analysis. This is largely due to the inability to measure the covariance between parameters exactly. The off diagonal elements of the covariance matrix are typically calculated to about 5 per cent expected errors. This means that if one were able to measure the off-diagonal elements of the covariance matrix to better precision, the information retained would go up by as much as a factor of a few.

This Figure also shows that, to get an appreciable amount of signal for the $\beta P(k)$ term or the $\beta^2 P(k)$ term, one must use coarse gridding.

4.5 Discussion

In this chapter, we discuss the extension of the pair weight method to the case of spherical redshift distortions. In linear theory, the spherical redshift distortions can be modeled as the sum of seven shape functions times functions of the correlation function (the Ξ 's) and β (the g 's). Using the generalized Fourier transform functions, each of the Ξ 's can be transformed into a measurement of the real space power spectrum. Using the properties of the Fisher matrix the seven estimates can be combined into three physically meaningful quantities: $P(k)$, $\beta P(k)$ and $\beta^2 P(k)$.

Using the Las Campanas Redshift Survey, we showed that the estimates of $P(k)$ can be disentangled from the estimates of $\beta P(k)$ and $\beta^2 P(k)$. The Fisher matrix, the disentangled correlated band-powers and the disentangled decorrelated band-powers are all sharply peaked about the wavenumber of interest. Thus, any smoothing will leave an estimate that is representative of the nominal wavenumber.

It is clear that the measurement of the real space galaxy-galaxy power spectrum is a

more difficult measurement than that of the redshift space galaxy-galaxy power spectrum. For LCRS, using the pair weight compression method, the expected errors for the real galaxy-galaxy power spectrum are only about a factor of two as large as those for the redshift space power spectrum.