

Chapter 9

Cosmic Structures

9.1 Quantifying structures

9.1.1 Introduction

- We have seen before that there is a very specific prediction for the power spectrum of density fluctuations in the Universe, characterised by (2.23). Recall that its shape was inferred from the simple assumption that the mass of density fluctuations entering the horizon should be independent of the time when they enter the horizon, and from the fact that perturbation modes entering during the radiation era are suppressed until matter begins dominating.
- Given the simplicity of the argument, and the corresponding strength of the prediction, it is very important for cosmology to find out whether the actual power spectrum of matter density fluctuations does in fact have the expected *shape*, and furthermore to determine the only remaining parameter, namely the *normalisation* of the power spectrum (this is discussed later in chapter 11).
- Since the location k_0 of the maximum in the power spectrum is determined by the horizon radius at matter-radiation equality (2.21),

$$k_0 = \frac{2\pi}{r_{\text{eq}}} = \frac{2\sqrt{2}\pi H_0}{c} \frac{\sqrt{\Omega_{\text{m}0}}}{a_{\text{eq}}^{3/2}}, \quad (9.1)$$

and the scale factor at equality is

$$a_{\text{eq}} = \frac{\Omega_{\text{r}0}}{\Omega_{\text{m}0}}, \quad (9.2)$$

the peak scale provides a measure of the matter-density parameter, $k_0 \propto \Omega_{\text{m}0}$. A measurement of k_0 would thus provide an independent and very elegant determination of $\Omega_{\text{m}0}$.

- Since the power spectrum is defined in Fourier space, it is not obvious how it can be measured. In a brief digression, we shall first summarise the relation between the power spectrum and the correlation function in configuration space, and clarify the meaning of the correlation function.

9.1.2 Power spectra and correlation functions

- The definition (2.22) shows that the power spectrum is given by an average over the Fourier modes of the density contrast. This average extends over all Fourier modes with a wave number k , i.e. it is an average over all directions in Fourier space keeping k constant. In other words, Fourier modes are averaged within spherical shells of radius k .
- In configuration space, structures can be quantified by the (two-point) correlation function

$$\xi(x) \equiv \langle \delta(\vec{y})\delta(\vec{x} + \vec{y}) \rangle , \quad (9.3)$$

where the average is now taken over all positions \vec{y} and all orientations of the separation vector \vec{x} , assuming homogeneity and isotropy.

- Inserting the Fourier expansion

$$\delta(\vec{x}) = \int \frac{d^3k}{(2\pi)^3} \hat{\delta}(\vec{k}) e^{-i\vec{k}\vec{x}} \quad (9.4)$$

of the density contrast into (9.3), using the definition (2.22) of the power spectrum and taking into account that the Fourier transform $\hat{\delta}$ must obey $\hat{\delta}(-\vec{k}) = \hat{\delta}^*(\vec{k})$ because δ is real, it is straightforward to show that the correlation function ξ is the Fourier transform of the power spectrum,

$$\xi(x) = \int \frac{d^3k}{(2\pi)^3} P(k) e^{-i\vec{k}\vec{x}} . \quad (9.5)$$

- Assuming isotropy, the integral over all relative orientations between \vec{x} and \vec{k} can be carried out, yielding

$$\xi(x) = \frac{1}{2\pi^2} \int_0^\infty P(k) \frac{\sin kx}{kx} k^2 dk , \quad (9.6)$$

whose inverse transform is

$$P(k) = 4\pi \int_0^\infty \xi(x) \frac{\sin kx}{kx} x^2 dx . \quad (9.7)$$

This indicates one way to determine the power spectrum via measuring the correlation function $\xi(x)$.

9.1.3 Measuring the correlation function

- How can the correlation function be measured? Obviously, we cannot measure the correlation function of the density field directly. All we can do is using galaxies as tracers of the underlying density field and use their correlation function as an estimate for that of the matter.
- Suppose we divide space into cells of volume dV small enough to contain at most a single galaxy. Then, the probability of finding one galaxies in dV_1 and another galaxy in dV_2 is

$$dP = \langle n(\vec{x}_1)n(\vec{x}_2) \rangle dV_1 dV_2, \quad (9.8)$$

where n is the number density of the galaxies as a function of position.

- If we introduce a density contrast for the galaxies in analogy to the density contrast for the matter,

$$\delta n \equiv \frac{n}{\bar{n}} - 1, \quad (9.9)$$

and assume for now that $\delta n = \delta$, we find from (9.8) with $n = \bar{n}(1 + \delta)$

$$dP = \bar{n}^2 \langle (1 + \delta_1)(1 + \delta_2) \rangle dV_1 dV_2 = \bar{n}^2 [1 + \xi(x)] dV_1 dV_2, \quad (9.10)$$

where x is the distance between the two volume elements. This shows that the correlation function quantifies the excess probability above random for finding galaxy pairs at a given distance.

- Thus, the correlation function can be measured by counting galaxy pairs and comparing the result to the Poisson expectation, i.e. to the pair counts expected in a *random* point distribution. Symbolically,

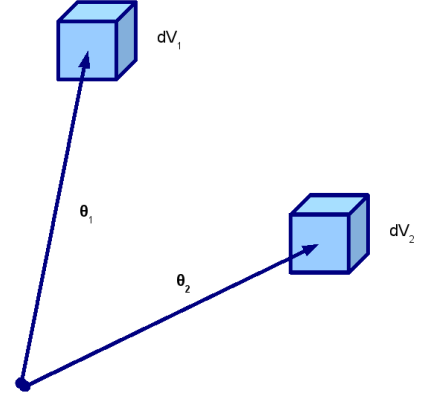
$$1 + \xi_1 = \frac{\langle DD \rangle}{\langle RR \rangle}, \quad (9.11)$$

where D and R represent the data and the random point set, respectively.

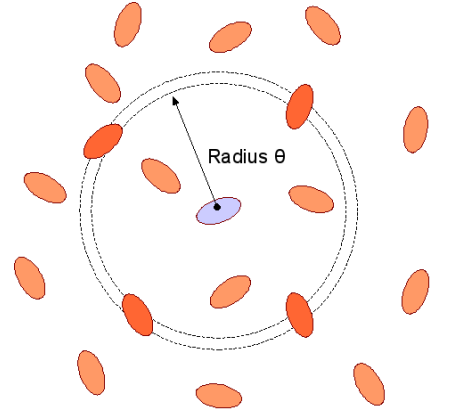
- Several other ways of measuring ξ have been proposed, such as

$$\begin{aligned} 1 + \xi_2 &= \frac{\langle DD \rangle}{\langle DR \rangle}, \\ 1 + \xi_3 &= \frac{\langle DD \rangle \langle RR \rangle}{\langle DR \rangle^2}, \\ 1 + \xi_4 &= 1 + \frac{\langle (D - R)^2 \rangle}{\langle RR \rangle^2}, \end{aligned} \quad (9.12)$$

which are all equivalent in the ideal situation of an infinitely extended point distribution. For finite point sets, ξ_3 and ξ_4 are superior to ξ_1 and ξ_2 due to their better noise properties.



The correlation function quantifies the probability to find a galaxy in the small volume dV_2 if there is a galaxy in the small volume dV_1 , a distance $r = |\vec{r}_2 - \vec{r}_1|$ away.



Correlations between points can be determined by counting pairs.

- The recipe for measuring $\xi(x)$ is thus to count pairs separated by x in the data D and in the random point set R , or between the data and the random point set, and to use one of the estimators given above.
- The obvious question is then how accurately ξ can be determined. The simple expectation in the absence of clustering is

$$\langle \xi \rangle = 0, \quad \langle \xi^2 \rangle = \frac{1}{N_p}, \quad (9.13)$$

where N_p is the number of pairs found. Thus, the Poisson error on the correlation function is

$$\frac{\Delta \xi}{1 + \xi} = \frac{1}{\sqrt{N_p}}. \quad (9.14)$$

- This is a lower limit to the actual error, however, because the galaxies are in fact correlated. It turns out that the result (9.14) should be multiplied with $1 + 4\pi\bar{n}J_3$, where J_3 is the volume integral over ξ within the galaxy-survey volume. The true error bars on ξ are therefore hard to estimate.
- Having measured the correlation function, it would in principle suffice to carry out the Fourier transform (9.7) to find $P(k)$, but this is difficult in reality because of the inevitable sample limitations. Consider (9.6) and an underlying power spectrum of CDM shape, falling off $\propto k^{-3}$ for large k , i.e. on small scales. For fixed x , the integrand in (9.6) falls off very slowly, which means that a considerable amount of small-scale power is mixed into the correlation function. Since ξ at large x is small and most affected by measurement errors, this shows that any uncertainty in the large-scale correlation function is propagated to the power spectrum even on small scales.
- A further problem is the uncertainty in the mean galaxy number density \bar{n} . Since $1 + \xi \propto \bar{n}^{-1}$ according to (9.10), the uncertainty in ξ due to an uncertainty in \bar{n} is

$$\frac{\Delta \xi}{1 + \xi} \approx \Delta \xi = \frac{\Delta \bar{n}}{\bar{n}}, \quad (9.15)$$

showing that ξ cannot be measured with an accuracy better than the relative accuracy of the mean galaxy density.

9.1.4 Measuring the power spectrum

- Given these problems with real data, it seems appropriate to estimate the power spectrum directly. The function to be transformed

is the density field sampled by the galaxies, which can be represented by a sum of Dirac delta functions centred on the locations of the N galaxies,

$$n(\vec{x}) = \sum_{i=1}^N \delta_D(\vec{x} - \vec{x}_i) . \quad (9.16)$$

- The Fourier transform of the density contrast is then

$$\hat{\delta}(\vec{k}) = \frac{1}{N} \sum_{i=1}^N e^{i\vec{k}\vec{x}_i} . \quad (9.17)$$

In the absence of correlations, the Fourier phases of the individual terms are independent, and the variance of the Fourier amplitude for a single mode becomes

$$\langle \hat{\delta}(\vec{k}) \hat{\delta}^*(\vec{k}) \rangle = \frac{1}{N^2} \sum_{i=1}^N e^{i\vec{k}\vec{x}_i} e^{-i\vec{k}\vec{x}_i} = \frac{1}{N} . \quad (9.18)$$

This is the so-called *shot noise* present in the power spectrum due to the discrete sampling of the density field.

- The shot-noise contribution needs to be subtracted from the power spectrum of the real, correlated galaxy distribution,

$$P(k) = \frac{1}{m} \sum |\hat{\delta}(\vec{k})|^2 - \frac{1}{N} , \quad (9.19)$$

where the sum extends over all m modes contained in the survey with wave number k .

- This is not the final result yet, because any real survey typically covers an irregularly shaped volume from which parts need to be excised because they are overshadowed by stars or unusable for any other reasons. The combined effect of mask and irregular survey volume is described by a window function $f(\vec{x})$ which multiplies the galaxy density,

$$n(\vec{x}) \rightarrow f(\vec{x})n(\vec{x}) , \quad (1 + \delta) \rightarrow f(\vec{x})(1 + \delta) , \quad (9.20)$$

implying that the Fourier transform of the mask needs to be subtracted.

- Moreover, the Fourier convolution theorem says that the Fourier transform of the *product* $f(\vec{x})\delta(\vec{x})$ is the *convolution* of the Fourier transforms $\hat{f}(\vec{k})$ and $\hat{\delta}(\vec{k})$,

$$\widehat{f\delta} = \hat{f} * \hat{\delta} \equiv \int \hat{f}(\vec{k}') \hat{\delta}(\vec{k} - \vec{k}') d^3k' . \quad (9.21)$$

If the Fourier phases of \hat{f} and $\hat{\delta}$ are uncorrelated, which is the case if the survey volume is large enough compared to the size $2\pi/k$ of the density mode, this translates to a convolution of the power spectrum,

$$P_{\text{obs}} = P_{\text{true}} * |\hat{f}(\vec{k})|^2 . \quad (9.22)$$

- This convolution typically has two effects; first, it smooths the observed compared to the true power spectrum, and second, it changes its amplitude. The corresponding correction is given by

$$P(k) \rightarrow P(k) \frac{(\int f d^3x)^2}{\int f^2 d^3x \int d^3x} . \quad (9.23)$$

- If the Poisson error dominates in the survey, the different modes $\hat{\delta}(\vec{k})$ can be shown to be uncorrelated, and the standard deviation after summing over the m modes with wave number k is $\sqrt{2m}/N$, which yields the minimal error bar to be attached to the power spectrum.
- Thus, the shot noise contribution and the Fourier transform of the window function need to be subtracted, the window function needs to be deconvolved, and the amplitude needs to be corrected for the effective volume covered by the window function before the measured power spectrum can be compared to the theoretical expectation.
- Finally, it is usually appropriate to assign weights $0 \leq w_i \leq 1$ to the individual galaxies to account for their varying density. The optimal weight for the i th galaxy sampling a Fourier mode with wave number k has been determined to be

$$w_i(k) = \frac{1}{1 + \bar{n}_i P(k)} , \quad (9.24)$$

where \bar{n}_i is the local mean density around the i th galaxy, and $P(k)$ is the power spectrum. If the density is low, the galaxies are weighted equally, and less if the local density is very high, because the many galaxies from a dense environment might otherwise suppress information from galaxies in less dense regions.

- Including weights, eqs. (9.17) and (9.18) become

$$\hat{\delta}(\vec{k}) = \frac{\sum w_i e^{i\vec{k}\vec{x}_i}}{\sum w_i} , \quad \langle |\hat{\delta}(\vec{k})|^2 \rangle = \frac{\sum w_i^2}{(\sum w_i)^2} . \quad (9.25)$$

- A final problem due to the finite size of the survey regards the normalisation of the power spectrum. The mean density estimate within the survey volume does not necessarily equal the true mean

density. Since, by definition, the mean of the density contrast δ' within the survey vanishes, we must have

$$\delta' = \delta - \int f(\vec{x})\delta(\vec{x})d^3x, \quad (9.26)$$

where δ is the true density contrast. Thus, the constant mean value of δ within the (masked) survey volume is subtracted.

- Subtracting a constant gives rise to a delta-function peak at $k = 0$ in the Fourier-transformed density contrast, and thus also in the power spectrum P' estimated from the survey.
- The observed power spectrum, however, is a convolution of the true power spectrum, as shown in (9.22). Thus, the delta-function peak caused by the misestimate of the mean density also needs to be convolved, giving rise to a contribution $P(0) * |\hat{f}(\vec{k})|^2$ in the observed power spectrum.
- Since the mean density contrast δ' within the survey is zero, the observed power spectrum at $k = 0$ must vanish, thus

$$P'_{\text{obs}}(k) = P_{\text{obs}}(k) - P_{\text{obs}}(0) * |\hat{f}(\vec{k})|^2. \quad (9.27)$$

9.1.5 Biasing

- What we have determined so far is the power spectrum of the *galaxy number-density contrast* δn rather than that of the *matter density contrast* δ . Simple models for the relation between both assume that there is a so-called *bias factor* $b(k)$ between them, such that

$$\widehat{\delta n}(\vec{k}) = b(k)\hat{\delta}(\vec{k}), \quad (9.28)$$

where $b(k)$ may or may not be more or less constant as a function of scale.

- Clearly, different types of objects sample the underlying matter density field in different ways. Galaxy clusters, for instance, are much more rare than galaxies and are thus expected to have a substantially higher bias factor than galaxies.
- Obviously, the bias factor enters squared into the power spectrum, e.g.

$$P_{\text{gal}} = b_{\text{gal}}^2(k) P(k). \quad (9.29)$$

It constitutes a major uncertainty in the determination of the matter power spectrum from the galaxy power spectrum.

9.1.6 Redshift-space distortions

- Of course, for the estimate (9.17) of the Fourier-transformed (galaxy) density contrast, the three-dimensional positions \vec{x}_i of the galaxies in the survey need to be known. Distances can be inferred only from the galaxy redshifts and thus from galaxy velocities.
- These, however, are composed of the Hubble velocities, from which the distances can be determined, and the peculiar velocities,

$$v = v_{\text{Hubble}} + v_{\text{pec}} , \quad (9.30)$$

which are caused by local density perturbations and are unrelated to the galaxy densities.

- Since observations of individual galaxies do not allow any separation between the two velocity components, distances are inferred from the total velocity v rather than the Hubble velocity as it should be,

$$D = \frac{v}{H_0} = \frac{v_{\text{Hubble}} + v_{\text{pec}}}{H_0} = D_{\text{true}} + \Delta D , \quad (9.31)$$

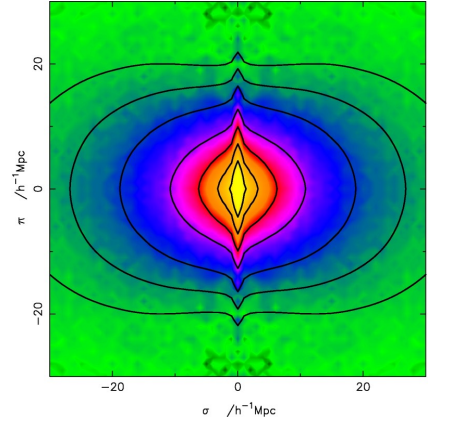
giving rise to a distance error $\delta D = v_{\text{pec}}/H_0$, the so-called *redshift-space distortion*.

- Fortunately, the redshift-space distortions have a peculiar pattern through which they can be corrected. Consider a matter overdensity such as a galaxy cluster, containing galaxies moving with random virial velocities in it. The virial velocities of order 1000 km s^{-1} scatter around the systemic cluster velocity and thus widen the redshift distribution of the cluster galaxies. In redshift space, therefore, the cluster appears stretched along the line-of-sight, which is called the *finger-of-god* effect.
- In addition, the cluster is surrounded by an infall region, in which the galaxies are not virialised yet, but move in an ordered, radial pattern towards the cluster. Galaxies in front of the cluster thus have higher, and galaxies behind the cluster have lower recession velocities compared to the Hubble velocity, leading to a flattening of the infall region in redshift space.
- A detailed analysis shows that the redshift-space power spectrum P_z is related to the real-space power spectrum P by

$$P_z(k) = P(k) (1 + \beta \mu^2)^2 , \quad (9.32)$$

where μ is the direction cosine between the line-of-sight and the wave vector \vec{k} , and β is related to the bias parameter b through

$$\beta \equiv \frac{f(\Omega_m)}{b} , \quad (9.33)$$



Peculiar velocities give rise to redshift-space distortions, whose characteristic shape constrains the bias.

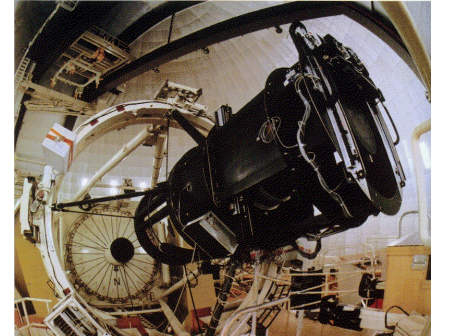
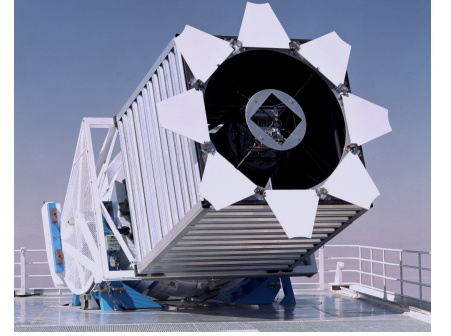
and $f(\Omega_m)$ is the logarithmic derivative of the growth factor $D_+(a)$,

$$f(\Omega_m) \equiv \frac{d \ln D_+(a)}{d \ln a} \approx \Omega_m^{0.6}. \quad (9.34)$$

- Thus, the characteristic pattern of the redshift-space distortions around overdensities allows a measurement of the bias factor. Another way of measuring b is based upon gravitational lensing. Corresponding measurements of b show that it is in fact almost constant or only weakly scale-dependent, and that it is very close to unity for “ordinary” galaxies.

9.1.7 Baryonic acoustic oscillations

- As we have seen in the discussion of the CMB, acoustic oscillations in the cosmic fluid have left density waves in the cosmic baryon distribution. Their characteristic wave length is set by the sound horizon at decoupling (7.26), $r_s \approx 63$ kpc. By now, this was increased by the cosmic expansion to 1280×63 kpc ≈ 80.6 Mpc, or $k_0 \approx 0.078$ Mpc $^{-1}$.
- This must be compared to the horizon size at matter-radiation equality (2.21). With $a_{\text{eq}} \approx 6200$ from (7.7), we find $r_{\text{eq}} \approx 11.0$ kpc, which was stretched by now to 6200×11.0 kpc ≈ 68.3 Mpc, or $k_s \approx 0.092$ Mpc $^{-1}$.
- Thus, the peak scale of the power spectrum and the wavelength of the fundamental mode of the baryonic acoustic oscillations are of comparable size. Near the peak of the power spectrum, we thus expect a weak wave-like imprint on top of the otherwise smooth dark-matter power spectrum.



Top: The telescope dedicated to the *Sloan Digital Sky Survey*. Bottom: The two-degree field camera in the prime focus of the Anglo-Australian Telescope.

9.2 Measurements and results

9.2.1 The power spectrum

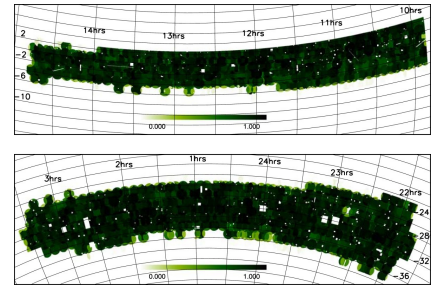
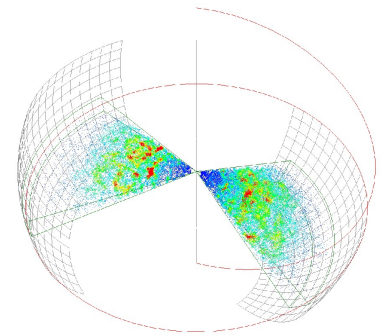
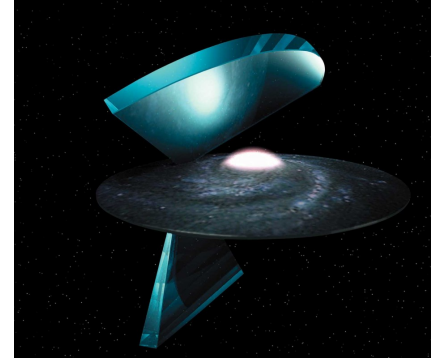
- Spectacularly successful measurements of the power spectrum became recently possible with the two largest galaxy surveys to date, the *Two-Degree Field Galaxy Redshift Survey* (2dFGRS) and the *Sloan Digital Sky Survey* (SDSS).
- As expected from the preceding discussion, an enormous effort has to be made to identify galaxies, measure their redshifts, selecting homogeneous galaxy subsamples as a function of redshift by luminosity and colour so as not to compare apples with oranges, estimating the window function of the survey,

determining the average galaxy number density, correcting for the convolution with the window function and for the bias, and so forth.

- Moreover, calibration experiments have to be carried out in which all measurement and correction techniques are applied to simulated data in the same way as to the real data to determine reliable error estimates and to test whether the full sequence of analysis steps ultimately yields an unbiased result.
- Based on 221,414 galaxies, the 2dFGRS consortium derived a power spectrum of superb quality. First and foremost, it confirms the power-spectrum shape expected for cold dark matter on the small-scale side of the peak. On its own, this is a highly remarkable result.
- Next, the 2dFGRS power spectrum clearly shows a turn-over towards larger scales, signalling the peak. The survey is still not quite large enough to show the peak, but the peak location can be estimated from the flattening of the power spectrum. Its proportionality to Ω_{m0} allows an independent determination of the matter density parameter.
- Finally, and most spectacularly, the power spectrum shows the baryonic acoustic oscillations, whose amplitude allows an independent determination of the ratio between the density parameters of baryons and dark matter.
- Apart from the fact that the CDM shape of the power spectrum is confirmed on small scales, the results obtained from the 2dFGRS can be summarised as follows:

$$\begin{aligned}\Omega_{m0} & 0.233 \pm 0.022 \\ \Omega_{b0}/\Omega_{m0} & 0.185 \pm 0.046\end{aligned}$$

The Hubble constant of $h = 0.72$ is assumed here. Indirectly, the baryon density is constrained to be $\Omega_{b0} \approx 0.04$, which is in perfect agreement with the value derived from primordial nucleosynthesis and the measured abundances of the light elements; the SDSS gives a value of $\Omega_{m0} = 0.24 \pm 0.02$, in agreement with 2dF and WMAP5.



Top: Geometry of the 2dFGRS survey volume. *Middle:* Galaxy distribution therein. *Bottom:* The area covered by the 2dFGRS on the sky.

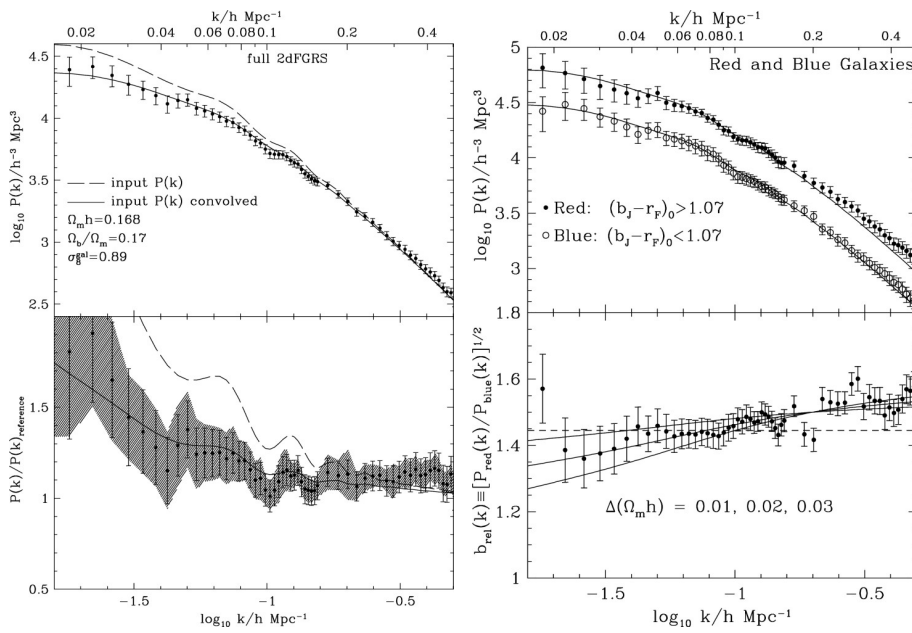


Figure 9.1: *Left*: Power spectrum of the 2dFGRS galaxy distribution (top) and after division by the smooth Λ CDM expectation (bottom). *Right*: Separate power spectra of red and blue galaxies (top) and their ratio (bottom).

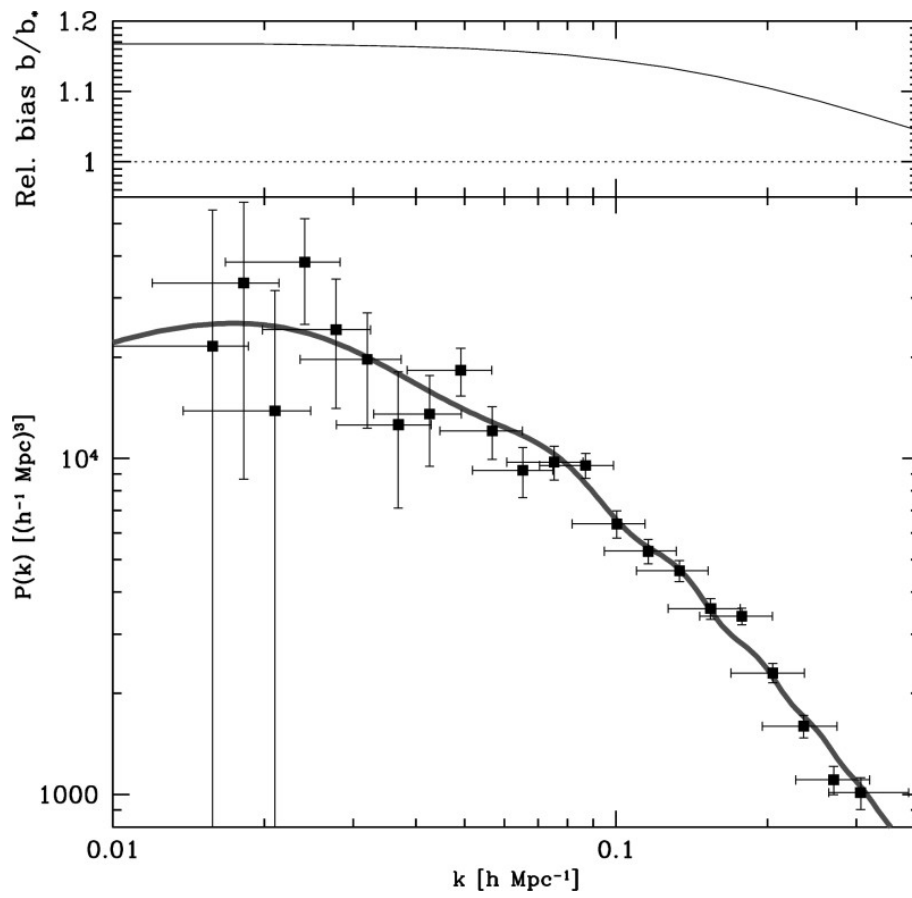


Figure 9.2: The galaxy power spectrum obtained from the SDSS (bottom), and the ratio between the power spectra of red and blue galaxies (top).