Advanced Cosmology
Statistics, Non-Gaussianity and Non-Linearity

Tobias Baldauf

January 29, 2019

In the Michaelmas course you focused on the history and composition of the homogeneous Universe, described linear fluctuations to the metric, their evolution and their creation in vanilla slow-roll inflation. This term, we will be focusing on making contact with observables in our Universe. Due to the stochastic origin of the fluctuations we will be requiring a statistical approach to compare theory and observations, and the relevant tools will be the topic of the first part of these lectures. You will then learn about Gaussian fluctuations in the Cosmic Microwave Background (CMB) temperature and polarization signals. Afterwards we will be considering modifications of inflation and their statistical signatures. In the last part of the lectures we will be concerned with non-linear structure formation in the late time Universe.

So far, the observed state of the Universe is compatible with the concordance ΛCDM model, where around 70% of the energy content of the Universe are in the form of dark energy, 25 % in the form of dark matter and only 5% in the form of baryonic matter. The fluctuations follow a spectrum resembling the one predicted by inflation and are very Gaussian. However, there are still many open questions, which we would like to answer using all possible observables available to us:

- What is causing the current accelerated expansion of the Universe? Is it simply a cosmological constant, is it a scalar field or do we need to modify Einstein's gravity?
- What is the nature of Dark Matter?
- What are the dynamics and field content of inflation?
- What is the mass of the neutrinos, which of the two hierarchies applies?
- How did the rich structures in our Universe arise from the small initial perturbations and how can we understand the process analytically?

In contrast to particle physics, where a particle collision can be repeated many times, we have only one Universe at our disposal and need to extract as much information as possible from the observables that are available to us:

- CMB (in particular lensing, polarization, small scales, also spectral distortions)
- Large-Scale Structure (LSS) - galaxy positions and weak gravitational lensing of the photons emitted by background galaxies or the CMB
- higher order statistics of LSS or the CMB
- 21 cm
- Ly-alpha forest

Unfortunately there is no textbook covering all aspects of this course, but here are some overview texts that should provide insightful additional reading for the first two parts of the course:

- R. Durrer *The Cosmic Microwave Background*
- S. Dodelson *Modern Cosmology*
- P. Peebles *The Large-Scale Structure of the Universe*
- F. Bernardeau, S. Colombi, E. Gaztanaga, R. Scoccimarro *Large-Scale Structure of the Universe and Cosmological Perturbation Theory* [arXiv:astro-ph/0112551]

This list will be updated and extended as we go along.

These lecture notes are evolving and I would be grateful if you could report mistakes or typos to t.baldauf@damtp.cam.ac.uk.

Contents

1 Statistics

The currently favoured model for creating the fluctuations in the Universe is cosmological inflation, a quantum mechanical process. Due due it's quantum nature, it is impossible to predict the realization of fluctuations in the Universe, but only it's statistical properties. There are many different realisations of the Universe that are statistically equivalent. To compare theory and observation, we will need to measure statistical properties of the Universe and compare the result to the same statistics evaluated in the model of the Universe under consideration.

1.1 Random fields in 3D space

The cosmological principle asserts that on large enough scales the Universe is isotropic and homogeneous. We will demand that the statistics are:

- invariant under rotations ⇒ **statistical isotropy**,
- invariant translations ⇒ **statistical homogeneity**.

We will start to consider random fields in three dimensional space. As these fields describe deviations from or fluctuations around the background Universe, we will only consider mean zero fields $\langle f \rangle = 0$. For instance we can consider the temperature fluctuations in the CMB

$$\theta = \frac{T}{\bar{T}} - 1 = \frac{\Delta T}{T} \tag{1.1}$$

or the the **overdensity**

$$\delta(\boldsymbol{x}) = \frac{\rho(\boldsymbol{x})}{\bar{\rho}} - 1 = \frac{n(\boldsymbol{x})}{\bar{n}} - 1\,, \tag{1.2}$$

where ρ is a density and $n(\boldsymbol{x}) = \delta N(\boldsymbol{x})/\delta V$ is a number density as appropriate for discrete objects.

1.1.1 Power Spectrum

Correlators are expectation values of products of field values at different spatial locations (or different Fourier modes). In the first part of the cosmology course you were exposed to the **power spectrum**[1]

$$\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle = (2\pi)^3 \delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}')P(|\boldsymbol{k}|) \tag{1.3}$$

By statistical isotropy the power spectrum may depend only on the magnitude of $\boldsymbol{k}$. The Dirac delta function ensures that the power spectrum is invariant under spatial translations. We will show this in more detail below. Distinct Fourier modes are uncorrelated and thus statistically independent. The above definition gives the power per unit Fourier space volume $\mathrm{d}^3k/(2\pi)^3$, sometimes one defines the power in a shell

$$\Delta^2(k) = \frac{k^3 P(k)}{2\pi^2} \tag{1.4}$$

such that $\Delta^2(k)\mathrm{d}\ln k = P(k)\mathrm{d}^3k/(2\pi)^3$ after integrating out the azimuthal angle. For real valued configuration space fields we have $\langle \delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\rangle = \langle \delta(\boldsymbol{k})\delta(-\boldsymbol{k}')\rangle$, which is why you might sometimes encounter a different definition of the power spectrum

$$\langle \delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\rangle = (2\pi)^3 \delta^{(\mathrm{D})}(\boldsymbol{k}-\boldsymbol{k}')P(|\boldsymbol{k}|)\,. \tag{1.5}$$

[1] The prefactor of $(2\pi)^3$ is due to the Fourier convention used here, for other conventions it might be absent or arise with an exponten of 3/2 instead of 3.

1.1.2 Correlation Function

The real space equivalent of the power spectrum is the **correlation function**

$$\langle \delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r})\rangle = \xi(|\boldsymbol{r}|) \tag{1.6}$$

By statistical homogeneity the correlation function can only depend on the difference of the positions $\boldsymbol{x}+\boldsymbol{r}$ and $\boldsymbol{x}$ and statistical isotropy enforces dependence on the magnitude only. Correlation function and power spectrum are related by a simple Fourier transformation and thus contain the same information. They do however have their respective advantages and disadvantages, both from a theoretical point of view regarding how apparent certain features are but also in terms of estimating the statistic.

An alternative interpretation of the correlation function defined above can be found in terms of the multi-point probability distribution functions. We will consider the background density field to be traced by a certain species with number density $\bar{n}$ and consider infinitesimally small volumes δV, which either host or don't host one tracer particle. This setup is sketched in Fig. 1.1. The one point probability for finding a particle in the small volume δV_1 is $\mathbb{P}_{1\text{pt}}(1) = \bar{n}\delta V_1$. If we had a purely random field the joint or two point probability of finding particles both in volumes δV_1 and δV_2 separated by $r_{12} = |\boldsymbol{x}_1 - \boldsymbol{x}_2|$ would be given by the product of the independent probabilities

$$\mathbb{P}_{2\text{pt}}(1,2) = \mathbb{P}_{1\text{pt}}(1)\mathbb{P}_{1\text{pt}}(2) = \bar{n}^2\delta V_1\delta V_2. \tag{1.7}$$

For a correlated sample the probabilities will no longer be independent and the correlation function can now be defined as the excess over random probability of finding two particles in volumes δV_1 and δV_2 separated by r_{12}

$$\mathbb{P}_{2\text{pt}}(1,2) = \bar{n}^2\left[1+\xi(r_{12})\right]\delta V_1\delta V_2. \tag{1.8}$$

Since the probability of having a particle in δV_1 is given by $\bar{n}\delta V_1$, we can write the conditional probability to find a particle in δV_2 given there is one in δV_1

$$\mathbb{P}_{1\text{pt}}(2|1) = \frac{\mathbb{P}_{2\text{pt}}(1,2)}{\mathbb{P}_{1\text{pt}}(1)} = \bar{n}\left[1+\xi(r_{12})\right]\delta V_2, \tag{1.9}$$

where we used Bayes theorem for the conditional probability. So we see that for correlated samples ($\xi(r_{12}) > 0$) the probability of finding a second particle is enhanced over random, whereas it is suppressed over random for the anti-correlated case ($\xi(r_{12}) < 0$). One can straightforwardly deduce a correlation function estimator from the above equation: count the number of neighbours in a shell of volume V around a given particle and compare to the expected number of pairs in a random field $\bar{n}V$.

We can also consider moments of the fields, which are products of the field at the same spatial location, for instance the **variance** of the field

$$\sigma_R^2 = \langle\delta_R^2\rangle = \xi_R(r=0). \tag{1.10}$$

Here the subscript R symbolizes a smoothing on a spatial scale R. We will relate the variance to the power spectrum shortly but note here that it is an important cosmological parameter describing the typical amplitude of fluctuations, which by convention is quantified in spheres of radius $R = 8\,h^{-1}$Mpc and amounts to about $\sigma_8 \approx 0.8$.

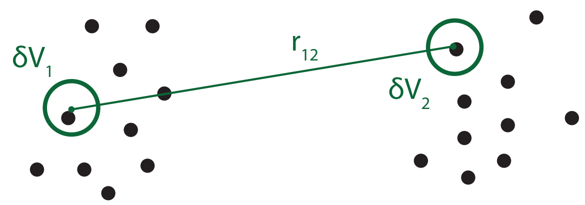

Figure 1.1: Clustering in a point distribution with a preferred distance, i.e., correlation length.

We understand the Universe to be a realization of an ensemble of universes and thus expectation values are understood as averages over many realizations. As we have only one Universe at our disposal to perform measurements we need to make use of **ergodic theorem** to replace ensemble averages by spatial averages. That the two agree is non-trivial and only the case if the correlation of field values vanishes in the large distance limit.

1.1.3 Higher order Correlators

In the simple case of linearly evolved fluctuations from single field inflation, the power spectrum or two-point function is sufficient to describe the statistics. However, deviations from the simple inflationary models, non-linear evolution or the consideration of tracers of the cosmic density distribution will require higher order correlators. The simplest of those is the **bispectrum**

$$\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\rangle = (2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B(k_1,k_2,k_3),. \tag{1.11}$$

Statistical homogeneity leads to the delta function that forces the three wavevectors to form a triangle and due to statistical isotropy the triangle can be fully described by three lengths or two lengths and one enclosed angle. The Fourier transform of the bispectrum is the three point correlation function $\zeta = \mathrm{FT}(B)$, which in the above probabilistic interpretation can be defined as

$$\mathbb{P}_{3\mathrm{pt}}(1,2,3) = \bar{n}^3\left[1+\xi(r_{12})+\xi(r_{23})+\xi(r_{31})+\zeta(r_{12},r_{23},r_{31})\right]\delta V_1\delta V_2\delta V_3\,. \tag{1.12}$$

Adding yet another field we obtain the **trispectrum**

$$\langle\delta(\boldsymbol{k}_1)\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\delta(\boldsymbol{k}_4)\rangle = (2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3+\boldsymbol{k}_4)T(k_1,k_2,k_3,k_4,|\boldsymbol{k}_1+\boldsymbol{k}_2|,|\boldsymbol{k}_2+\boldsymbol{k}_3|)\,, \tag{1.13}$$

which we will encounter again when discussing cosmic variance below.

1.2 Fourier Space

It will prove convenient to build up the actual density field from a superposition of modes that describe the behaviour on a certain scale.
We introduce the following Fourier convention:

$$\delta(\boldsymbol{k}) \quad = \quad \int \mathrm{d}^3r\,\exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right]\delta(\boldsymbol{r}), \tag{1.14}$$

$$\delta(\boldsymbol{r}) \quad = \quad \int \frac{\mathrm{d}^3k}{(2\pi)^3}\exp\left[-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right]\delta(\boldsymbol{k}). \tag{1.15}$$

The k-space representation of the nabla operator is given by $\boldsymbol{\nabla}\to -\mathrm{i}\boldsymbol{k}$. More often than not, the configuration space fields will be real, leading to $\delta^*(\boldsymbol{k}) = \delta(-\boldsymbol{k})$. Under a spatial shift $\boldsymbol{x}\to\boldsymbol{x}+\boldsymbol{\Delta x}$ the Fourier modes transform as

$$\delta(\boldsymbol{k})\to\exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{\Delta x}\right]\delta(\boldsymbol{k}) \tag{1.16}$$

and the power spectrum thus transforms as

$$\delta(\boldsymbol{k})\delta(\boldsymbol{k}')\to\delta(\boldsymbol{k})\delta(\boldsymbol{k}')\exp\left[\mathrm{i}(\boldsymbol{k}+\boldsymbol{k}')\cdot\boldsymbol{\Delta x}\right], \tag{1.17}$$

thus invariance under translations obviously requires $\boldsymbol{k}' = -\boldsymbol{k}$ and thus $\delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}')$. An equivalent argument can be made for higher n-point functions.

The Dirac delta function is thus given by

$$\delta^{(\mathrm{D})}(\boldsymbol{x}+\boldsymbol{x}') = \int\frac{\mathrm{d}^3q}{(2\pi)^3}\exp\left[\mathrm{i}(\boldsymbol{x}+\boldsymbol{x}')\boldsymbol{q}\right]. \tag{1.18}$$

In particular, for finite volumes this leads to

$$\delta^{(\mathrm{D})}(\boldsymbol{k}-\boldsymbol{k}') = \frac{V}{(2\pi)^3}\delta^{(\mathrm{K})}_{\boldsymbol{k},\boldsymbol{k}'}\,. \tag{1.19}$$

An important advantage of working in Fourier space is that convolutions in real space become simple multiplications in k-space

$$f(\boldsymbol{x}) = \int \mathrm{d}^3 y g(\boldsymbol{y})h(\boldsymbol{x}-\boldsymbol{y}) \Rightarrow f(\boldsymbol{k}) = g(\boldsymbol{k})h(\boldsymbol{k}). \tag{1.20}$$

This is of particular advantage, when **smoothing** operations are considered.

$$f_R(\boldsymbol{x}) = \int \mathrm{d}^3 y W_R(\boldsymbol{y})f(\boldsymbol{x}-\boldsymbol{y}) \Rightarrow f_R(\boldsymbol{k}) = f(\boldsymbol{k})W_R(\boldsymbol{k}). \tag{1.21}$$

For the variance of spheres in real space we consider a spatial **top-hat filter** $W_{\mathrm{TH},R}(r) = 3\theta(R-r)/4\pi R^3$ leading to

$$W_{\mathrm{TH},R}(k) = 3\frac{\sin(kR)-(kR)\cos(kR)}{(kR)^3} \tag{1.22}$$

In case of spherical symmetry we can perform the angular integration in the definition of the Fourier transform

$$f(r) = \frac{1}{2\pi^2}\int \mathrm{d}k\; k^2\frac{\sin kr}{kr}f(k) = \frac{1}{2\pi^2}\int \mathrm{d}k\; k^2 j_0(kr)f(k) \tag{1.23}$$

where j_0 is the spherical Bessel function. For the inverse transform this yields

$$f(k) = 4\pi\int \mathrm{d}r\; r^2 j_0(kr)f(r). \tag{1.24}$$

$$\begin{aligned}\xi(r) &= \int\frac{\mathrm{d}^3k}{(2\pi)^3}\int\frac{\mathrm{d}^3k'}{(2\pi)^3}\langle\delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle\exp\left[-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}\right]\exp\left[-\mathrm{i}\boldsymbol{k}'\cdot(\boldsymbol{x}+\boldsymbol{r})\right]\\ &= \int\frac{\mathrm{d}^3k}{(2\pi)^3}\int\frac{\mathrm{d}^3k'}{(2\pi)^3}(2\pi)^3P(k)\delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}')\exp\left[-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}\right]\exp\left[-\mathrm{i}\boldsymbol{k}'\cdot(\boldsymbol{x}+\boldsymbol{r})\right]\\ &= \int\frac{\mathrm{d}^3q}{(2\pi)^3}P(k)\exp\left[-\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}\right] = \frac{1}{2\pi^2}\int \mathrm{d}kk^2P(k)j_0(kr)\end{aligned} \tag{1.25}$$

In the other direction we have

$$\begin{aligned}\langle\delta(\boldsymbol{k})\delta(\boldsymbol{k}')\rangle &= \int \mathrm{d}^3x\int \mathrm{d}^3x'\exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}\right]\exp\left[\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{x}\right]\langle\delta(\boldsymbol{x})\delta(\boldsymbol{x}')\rangle\\ &= \int \mathrm{d}^3x\exp\left[\mathrm{i}\boldsymbol{x}\cdot(\boldsymbol{k}+\boldsymbol{k}')\right]\int \mathrm{d}^3r\exp\left[\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}\right]\xi(r)\\ &=(2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}')\int \mathrm{d}^3r\exp\left[\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}\right]\xi(r)\end{aligned} \tag{1.26}$$

since the last line has the same form as the definition of the power spectrum, the power spectrum is in turn related to the correlation function by

$$P(k) = \int \mathrm{d}^3r\xi(r)\exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right] = 4\pi\int \mathrm{d}rr^2\xi(r)j_0(kr). \tag{1.27}$$

Let us check that the transformations actually close.

$$P(k) = 4\pi\int\frac{\mathrm{d}kk^2}{2\pi^2}P(k')\int \mathrm{d}rr^2j_0(kr)j_0(k'r) = P(k)\,, \tag{1.28}$$

where we used the closure relation for spherical Bessel functions

$$\int_0^\infty \mathrm{d}xx^2j_\alpha(ux)j_\alpha(vx) = \frac{\pi}{2u^2}\delta^{(\mathrm{D})}(u-v)\,. \tag{1.29}$$

Let us finally consider the **variance** of the smoothed density field

$$\sigma_R^2 = \left\langle\delta^2_{\mathrm{TH},R}\right\rangle = \frac{1}{2\pi^2}\int \mathrm{d}kk^2P(k)W^2_{\mathrm{TH},R}(k)\,. \tag{1.30}$$

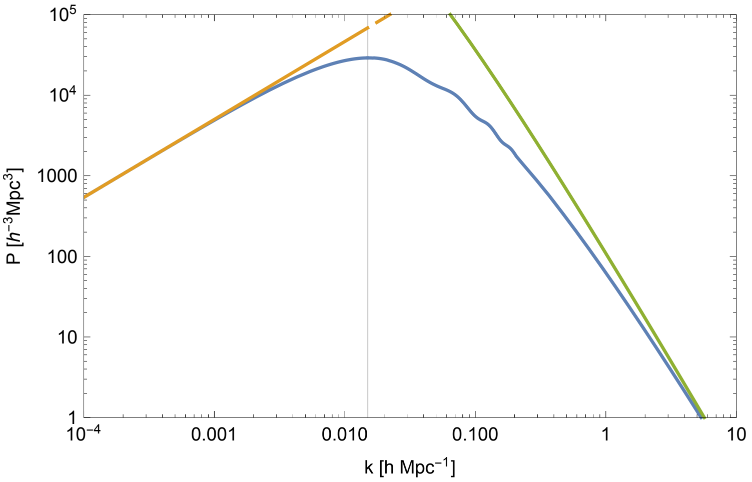

Figure 1.2: Linear matter power spectrum $P_{\delta\delta}$ of the ΛCDM model (blue) and its low- and high-k approximations.

1.3 Recap: Shape of the Power Spectrum

The shape of the matter power spectrum is given by linear transformations of the power spectrum generated by inflation. In the cosmology course it was shown that the power spectrum of the metric perturbations scales as

$$P_{\zeta\zeta}(k) \propto k^{n_s - 4} \tag{1.31}$$

where n_s is proportional to the slow-roll parameters and observed to be close to unity $n_s \approx 0.967$. The modes seeded during inflation are subject to different total growth rates depending on whether they reentered the horizon before or after matter-radiation equality. To account for this fact it is convenient to introduce the **transfer function** $T(k)$, which can be implicitly defined as a relation between the late matter power spectrum after recombination and the seed fluctuations produced by inflation

$$P_{\delta\delta}(k) \propto T^2(k) k^4 P_{\zeta\zeta}(k) \propto T^2(k) k^{n_s} \tag{1.32}$$

where we used that inside the horizon $k^2\zeta \propto k^2\phi \propto \delta$

- outside the horizon potential fluctuations are constant $\Rightarrow$ transfer function constant on large scales
- modes that enter during radiation domination have a supressed growth $\Rightarrow$ transfer function is supressed on small scales

As you have discussed in the cosmology course last term, the transfer function can be parametrized as

$$T^2(k) \propto \begin{cases} 1 & k < k_{\rm eq}\,, \\ \frac{k_{\rm eq}^4}{k^4}\left(1 + \ln\left(\frac{k}{k_{\rm eq}}\right)\right)^2 & k > k_{\rm eq}\,. \end{cases} \tag{1.33}$$

The expansion factor at equality is given by

$$a_{\rm eq} = \frac{\Omega_{\rm r,0}}{\Omega_{\rm m,0}} \tag{1.34}$$

and we can use $\rho_{\rm r} = \frac{\pi^2}{15} T_{\rm CMB}^4$ to calculate the radiation density parameter $\Omega_{\rm r,0} \approx 8 \times 10^{-5}$ yielding $a_{\rm eq} \approx 3 \times 10^{-4}$. The horizon wavenumber at equality is then given by $k_{\rm eq}/\mathcal{H}_{\rm eq} = 1$, giving $k_{\rm eq} \approx 0.015 h^{-1}$Mpc. An example of the power spectrum as well as its low- and high-k limits is given in Fig. 1.3. The power spectrum thus contains information about the initial conditions, which are unperturbed on large scales as well as information on the various components affecting the suppression of growth on small scales. The mass of neutrinos for instances directly affects the matter transfer function and allows us to put constraints on neutrino mass that are close to the minimal mass required from neutrino oscillation experiments.

Let us focus on the impact and relevance of the **Baryon Acoustic Oscillation** (BAO) feature in the correlation function and power spectrum. In the correlation function the BAO shows up as a distinct feature at $r_{\mathrm{BAO}} \approx 100\ h^{-1}$Mpc. Let us for simplicity consider the BAO to be a Dirac delta function in one spatial dimension and calculate the corresponding power spectrum

$$P_{1\mathrm{D}}(k) = \int \mathrm{d}x \ \left[\delta^{(\mathrm{D})}(x - r_{\mathrm{BAO}}) + \delta^{(\mathrm{D})}(x + r_{\mathrm{BAO}})\right] \exp\left[\mathrm{i}kx\right] = 2\cos(kr_{\mathrm{BAO}})\,. \tag{1.35}$$

We thus see that the wiggles in the power spectrum correspond to a spatially localized feature in the correlation function. In reality the BAO gets broadended at a scale w in the correlation function and the corresponding wiggles in the power spectrum are suppressed at high wavenumbers

$$\begin{aligned} P(k) &= \int \mathrm{d}x \ \frac{1}{\sqrt{2\pi}w} \left\{\exp\left[-\frac{(x - r_{\mathrm{BAO}})^2}{2w^2}\right] + \exp\left[-\frac{(x + r_{\mathrm{BAO}})^2}{2w^2}\right]\right\} \exp\left[\mathrm{i}kx\right] \\ &= 2\cos(kr_{\mathrm{BAO}}) \exp\left[-\frac{1}{2}k^2 w^2\right]. \end{aligned} \tag{1.36}$$

The BAO scale is very well known from CMB observations and thus forms a standard ruler in LSS. When observing this scale in transverse galaxy clustering we are probing the angular diameter distance and when observing it along the line of sight, we are probing the Hubble rate. Both of these quantities depend on the expansion history of the Universe and thus on the dark energy equation of state $w = p/\rho$.Constraining the dark energy equation of state at the 10% level requires 1% precision in the measurement of distances or $H(z)$.

1.4 Gaussian Random Fields

1.4.1 Probability and Characteristic Function

A vector $\boldsymbol{f} = [f_1, \ldots, f_N]$ of random variables is called Gaussian, if the joint probability density function (PDF) is a multivariate Gaussian

$$\mathbb{P}(\boldsymbol{f}) = \frac{1}{\sqrt{(2\pi)^N |C|}} \exp\left[-\frac{1}{2} f_i C^{-1}_{ij} f_j\right], \tag{1.37}$$

where the positive definite, symmetric $N \times N$-matrix $C_{ij} = \langle f_i f_j \rangle$ is called the **covariance matrix**.

A random field $f : \mathbb{R}^3 \to \mathbb{R}$ is a **Gaussian random field** (GRF) if for arbitrary collections of field points $(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$ the variables $[f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_N)]$ are joint Gaussian variables. For the GRF, the PDF can be expressed as a Gaussian functional of f, for practical purposes we will often work with a finite set of tracer points or pixels and denote $f_i = f(\boldsymbol{x}_i)$.

For a GRF, by statistical homogeneity the covariance matrix C can only depend on the separation $\boldsymbol{x}_i - \boldsymbol{x}_j$ and by statistical isotropy only on the magnitude $|\boldsymbol{x}_i - \boldsymbol{x}_j|$. As $f(\boldsymbol{k})$ is linear in $f(\boldsymbol{x})$, the PDF of the Fourier modes $f(\boldsymbol{k})$ is Gaussian as well. In particular, the canonical commutation relations and the resulting momentum conserving delta function ensure a diagonal covariance matrix for the Fourier modes. As you can show yourself, irrelevant degrees of freedom can be integrated out from the PDF

$$\mathbb{P}(f_1, \ldots, f_{N-M}) = \left\{\prod_{i=N-M+1}^{N} \int \mathrm{d}f_i\right\} \mathbb{P}(\boldsymbol{f})\,. \tag{1.38}$$

The simplest inflationary models predict Gaussian primordial fluctuations, but there are distinct models that can predict non-Gaussian features. Those will be studied in the third part of this course and for now we will work under the assumption of Gaussian seeds. The Gaussian property is conserved by linear evolution. As the (primary) CMB is very nearly linear in the initial fluctuations the fluctuations look very Gaussian, there are however small non-Gaussian deviations imprinted by gravitational lensing of CMB photons by the large-scale structure of the Universe. Non-linear structure formation at late times leads to strong deviations from Gaussianity that will be discussed in the last part of this course.

Let us first note that the multivariate Gaussian is appropriately normalized, i.e., that the probability integrates to unity $\int \mathrm{d}^n f \mathbb{P}(\boldsymbol{f}) = 1$

$$\int \mathrm{d}^n f \; \mathbb{P}(\boldsymbol{f}) = \prod_i \Big\{ \int_{-\infty}^{\infty} \mathrm{d}f_i \Big\} \frac{1}{\sqrt{(2\pi)^N |C|}} \exp\left[-\frac{1}{2} f_i C^{-1}_{ij} f_j \right] . \tag{1.39}$$

Let us orthogonalize $C^{-1} = O^{-1} D O$ (where $O^{-1} = O^{\mathrm{T}}$) and $y = O^{-1} f$, i.e., we are performing a rotation in the Euclidean space we are integrating over. With $D = \mathrm{diag}\left\{1/\sigma_i^2\right\}$ we have

$$\prod_i \Big\{ \int_{-\infty}^{\infty} \mathrm{d}y_i \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left[-\frac{1}{2} \frac{y_i^2}{\sigma_i^2} \right] \Big\} = 1 \, . \tag{1.40}$$

For the density field at one point we have thus

$$\mathbb{P}(\delta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2} \frac{\delta^2}{\sigma^2} \right] . \tag{1.41}$$

This can be used to calculate the probability of a point in a density field smoothed on scale R to exceed a certain density threshold δ_{c}

$$\mathbb{P}(\delta_R > \delta_{\mathrm{c}}) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\delta_{\mathrm{c}}}^{\infty} \mathrm{d}\delta \; \exp\left[-\frac{1}{2} \frac{\delta^2}{\sigma^2} \right] = \frac{1}{2} \mathrm{erfc} \left(-\frac{\delta_{\mathrm{c}}}{\sqrt{2}\sigma_R} \right) \tag{1.42}$$

We will return to this towards the end of this course, when we calculate the abundance of collapsed objects (dark matter haloes), which can be identified with regions of mass $M \propto R^3$ whose linear overdensity is exceeding the threshold $\delta_{\mathrm{c}} = 1.686$.

We can now calculate the **characteristic function** of the Gaussian PDF

$$\mathcal{M}(\mathrm{i}\boldsymbol{J}) = \langle \exp\left[\mathrm{i}\boldsymbol{J} \cdot \boldsymbol{f}\right] \rangle = \int \mathrm{d}^N f \frac{1}{\sqrt{(2\pi)^N |C|}} \exp\left[-\frac{1}{2} f_i C^{-1}_{ij} f_j + \mathrm{i} J_i f_i \right] = \exp\left[-\frac{1}{2} J_i C_{ij} J_j \right] \tag{1.43}$$

This Gaussian integral can be easily evaluated by completing the squares. Generic moments of the field f can now be generated from derivatives of the characteristic function at $J = 0$

$$\langle f_{i_1} \dots f_{i_n} \rangle = (-\mathrm{i})^n \frac{\partial}{\partial J_{i_1}} \dots \frac{\partial}{\partial J_{i_n}} \langle \exp\left[\mathrm{i}\boldsymbol{J} \cdot \boldsymbol{f}\right] \rangle \Big|_{J=0} \tag{1.44}$$

Let us first show that the correlation function is indeed $\langle f_i f_j \rangle = C_{ij}$. For $N = 2$ we obtain

$$\begin{aligned} \langle f_m f_n \rangle =& (-\mathrm{i})^2 \frac{\partial}{\partial J_m} \frac{\partial}{\partial J_n} \langle \exp\left[\mathrm{i}\boldsymbol{J} \cdot \boldsymbol{f}\right] \rangle \Big|_{J=0} \\ =& (-\mathrm{i}) \frac{\partial}{\partial J_m} \left(J_j C_{jm} \exp\left[-\frac{1}{2} J_i C_{ij} J_j \right] \right) \Big|_{J=0} = C_{mn} \, , \end{aligned} \tag{1.45}$$

which proves that the components of the covariance matrix are indeed given by the correlation function.

1.4.2 Gaussian Random Fields in Fourier Space

Note that the real and imaginary part of the complex density $\delta(\boldsymbol{k}) = a(\boldsymbol{k}) + \mathrm{i}b(\boldsymbol{k})$ field are independent Gaussian random fields with variance $P/2$

$$\begin{aligned}
\langle a(\boldsymbol{k})a(\boldsymbol{k}')\rangle =& \frac{1}{4}\langle(\delta+\delta^*)(\delta+\delta^*)\rangle = (2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}-\boldsymbol{k}')\frac{P}{2} + (\boldsymbol{k}'\to-\boldsymbol{k}') && (1.46)\\
\langle b(\boldsymbol{k})b(\boldsymbol{k}')\rangle =& -\frac{1}{4}\langle(\delta-\delta^*)(\delta-\delta^*)\rangle = (2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}-\boldsymbol{k}')\frac{P}{2} - (\boldsymbol{k}'\to-\boldsymbol{k}') && (1.47)\\
\langle a(\boldsymbol{k})b(\boldsymbol{k}')\rangle =& \frac{1}{4\mathrm{i}}\langle(\delta+\delta^*)(\delta-\delta^*)\rangle = 0 && (1.48)
\end{aligned}$$

Thus realisations can be generated by drawing real and imaginary parts from independent Gaussian distributions with mean zero and variance $P/2$. While the real and imaginary parts are independent Gaussian distributed, the magnitude of δ follows a Rayleigh distribution. This can be seen as follows: Both real and imaginary parts are Gaussian distributed with variance $P/2$

$$\mathbb{P}_a(x) = \mathbb{P}_b(x) = \frac{1}{\sqrt{\pi P}}\exp\left[-\frac{x^2}{P}\right] \tag{1.49}$$

The probability has to be invariant under a reparametrization and thus going to polar coordinates $\delta(\boldsymbol{k}) = a(\boldsymbol{k}) + \mathrm{i}b(\boldsymbol{k}) = r\exp[\mathrm{i}\phi]$ yields

$$\mathbb{P}_{|\delta|}(r)\mathbb{P}_\varphi(\varphi)\mathrm{d}r\mathrm{d}\varphi = \mathbb{P}_a(a)\mathbb{P}_a(b)\mathrm{d}a\mathrm{d}b = \frac{r}{\pi P}\exp\left[-\frac{r^2}{P}\right]\mathrm{d}r\mathrm{d}\varphi\,. \tag{1.50}$$

Thus, the probability is uniform for the phase of the mode and Rayleigh for the magnitude.

$$\mathbb{P}_{|\delta|}(r) = \frac{2r}{P}\exp\left[-\frac{r^2}{P}\right], \qquad \mathbb{P}_\varphi(\varphi) = \frac{1}{2\pi}\,. \tag{1.51}$$

1.4.3 Wick Theorem

Wick theorem[2]: for a mean zero Gaussian random field the reduced correlation functions of order higher than two either vanish (odd number of fields) or are expressible in terms of products of two-point functions summed over all possible pairings (even number of fields).

The odd case obviously vanishes, for the even case we take Eq. (1.44) and obtain

$$\begin{aligned}
\left\langle f_{i_1},\ldots,f_{i_{2N+1}}\right\rangle =&\quad 0,\\
\left\langle f_{i_1},\ldots,f_{i_{2N}}\right\rangle =& \sum_{\text{ordered pairings }\mathcal{P}_a}\;\prod_{\text{pairs }(i,j)\text{ in the pairing }\mathcal{P}_a} C_{ij}.
\end{aligned} \tag{1.52}$$

Alternatively one can also consider all $(2N)!$ possible permutations of $(i_1,\ldots,i_{2N})$, cut each of them up in subsequent pairs and remove the redundancies by dividing by appropriate factors

$$\left\langle f_{i_1},\ldots,f_{i_{2N}}\right\rangle = \frac{1}{2^N N!}\underbrace{\left(C_{i_1i_2}C_{i_3i_4}\ldots C_{i_{2N-1}i_{2N}} + \mathrm{perm}\right)}_{(2N)!\text{ permutations}}\,. \tag{1.53}$$

The factor 2^N counts the redundant terms arising from exchanges of indices $C_{ij} \leftrightarrow C_{ji}$ which leaves the correlator unaffected due to the symmetry of the covariance matrix. The factor $N!$ counts possible reorderings of whole pairs $C_{ij}C_{mn} \leftrightarrow C_{mn}C_{ij}$ which leave the correlator unaffected as well. The number of products of correlation functions in the correlator of $2N$ fields is thus $(2N)!/(2^N N!) = (2N-1)!!$. You can convince yourself that this is the case by calculating the four point function explicitly and by induction going from $2N$ to $2N+2$.

In summary, the procedure for calculating correlators of GRF is thus:

[2] In cosmology this theorem was introduced by Isserlis, but is often associated with the QFT version introduced by Wick.

1. We first generate all $(2N-1)!!$ possible ordered pairings of indices from $(i_1, \ldots, i_{2N})$, i.e., we generate $\mathcal{P} = \{[(i_1, i_2), \ldots, (i_{2N-1}, i_{2N})], \ldots, [(i_1, i_{2N}), \ldots, (i_2, i_{2N-1})]\}$. For the sake of definiteness we choose $i < j$ for the pairings (i, j).
2. For each of these pairings in $\mathcal{P}$ calculate the product of N correlators, e.g. for $\mathcal{P}_1$ evaluate $C_{i_1 i_2} \ldots C_{i_{2N-1} i_{2N}}$
3. Sum over all of these products of correlators.

Let us consider the example of the four point function. We expect that there will be $3!! = 3$ contributing terms. Indeed, there are three different fields that f_1 can be correlated with. Once this partner for f_1 has been chosen, the remaining fixed pair is correlated as well.

$$\begin{aligned} \langle f_1 f_2 f_3 f_4 \rangle =& \langle f_1 f_2 f_3 f_4 \rangle + \langle f_1 f_2 f_3 f_4 \rangle + \langle f_1 f_2 f_3 f_4 \rangle \\ =& C_{12} C_{34} + C_{13} C_{24} + C_{14} C_{23} \end{aligned} \tag{1.54}$$

In Fourier space this leads to

$$\begin{aligned} \left\langle \delta(\boldsymbol{k}_{i_1}), \ldots, \delta(\boldsymbol{k}_{i_{2N+1}}) \right\rangle =& \quad 0, \\ \left\langle \delta(\boldsymbol{k}_{i_1}), \ldots, \delta(\boldsymbol{k}_{i_{2N}}) \right\rangle =& \sum_{\text{ordered pairings } \mathcal{P}_a} \; \prod_{\text{pairs } (i,j) \text{ in the pairing } \mathcal{P}_a} \left\langle \delta(\boldsymbol{k}_i), \delta(\boldsymbol{k}_j) \right\rangle . \end{aligned} \tag{1.55}$$

1.4.4 Weakly non-Gaussian Fields

Let us consider the **moment generating function** which is closely related to the characteristic function discussed above

$$\mathcal{M}(J) = \sum_{p=0}^{\infty} \frac{\langle \delta^p \rangle}{p!} J^p = \langle \exp[J\delta] \rangle = \int \mathrm{d}\delta \; \mathbb{P}(\delta) \exp[J\delta] . \tag{1.56}$$

For a Gaussian field we obviously have

$$\mathcal{M}(J) = \exp\left[\frac{1}{2} J^2 \sigma^2\right] = 1 + \frac{\sigma^2}{2!} J^2 + \frac{3\sigma^4}{4!} J^4 + \ldots \tag{1.57}$$

The moment generating function is the Laplace transform of the PDF and thus the PDF can be written as the inverse Laplace transform of the moment generating function

$$\mathbb{P}(\delta) = \int_{-\mathrm{i}\infty}^{\mathrm{i}\infty} \frac{\mathrm{d}J}{2\pi \mathrm{i}} \exp[\delta J] \mathcal{M}(J) \tag{1.58}$$

For a Gaussian PDF we can replace $y = -\mathrm{i}J$ and perform the Gaussian integral

$$\mathbb{P}(\delta) = \int_{-\infty}^{\infty} \frac{\mathrm{d}y}{2\pi} \exp\left[-\frac{1}{2}\sigma^2 y^2 - \mathrm{i}y\delta\right] = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{\delta^2}{\sigma^2}\right] \tag{1.59}$$

Let us now consider a field that has a non-vanishing cubic moment $\left\langle \delta^3 \right\rangle = S_3 \sigma^4$ with $S_3 \ll 1$. Writing the moment generating function as

$$\mathcal{M}(J) = \exp\left[\frac{1}{2} J^2 \sigma^2\right] \left(1 + \frac{\sigma^4 S_3}{3!} J^3\right) , \tag{1.60}$$

we obtain for the PDF

$$\mathbb{P}(\delta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{\delta^2}{\sigma^2}\right] \left(1 + \frac{S_3(\delta^3 - 3\delta\sigma^2)}{3!\sigma^2}\right) . \tag{1.61}$$

This is the leading correction to the Gaussian PDF, a more general form is known as the Edgeworth or Gram-Charlier expansion of the PDF

$$\mathbb{P}(\delta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\nu^2\right] \left[1 + \sigma \frac{S_3 H_3(\nu)}{6} + \sigma^2 \left(\frac{S_4 H_4(\nu)}{24} + \frac{S_3^2 H_6(\nu)}{72}\right) + \ldots\right] , \tag{1.62}$$

where $\nu = \delta/\sigma$ and $H_n(\nu)$ are the Hermite polynomials. Note that the Edgeworth expansion can also be written as a derivative operator acting on the Gaussian PDF

$$\mathbb{P}(\delta) = \left(1 - \frac{S_3\sigma^4}{3!}\frac{\partial^3}{\partial\delta^3}\right)\mathbb{P}_{\rm G}(\delta). \tag{1.63}$$

1.4.5 The simplest form of non-Gaussianity

Non-Gaussian fields can be straightforwardly generated from Gaussian fields by non-linear transformations. Let us consider a very simple model for relating the fluctuations in the galaxy distribution $\delta_{\rm g}$ to the underlying matter Gaussian matter distribution δ, the so called local bias model, which we will discuss in much more detail at the end of this course

$$\delta_{\rm g}(\boldsymbol{x}) = b_1\delta(\boldsymbol{x}) + \frac{b_2}{2}(\delta^2(\boldsymbol{x}) - \sigma^2) + \mathcal{O}(\delta^3). \tag{1.64}$$

Here the subtraction of the variance in the squared term ensures that the galaxy overdensity averages to zero. Let us now consider the above model in Fourier space, as we have seen before, squaring in real space corresponds to convolutions in Fourier space

$$\delta_{\rm g}(\boldsymbol{k}) = b_1\delta(\boldsymbol{k}) + \frac{b_2}{2}\int\frac{{\rm d}^3q}{(2\pi)^2}\delta(\boldsymbol{q})\delta(\boldsymbol{k}-\boldsymbol{q}) = \delta_{\rm g}^{(1)}(\boldsymbol{k}) + \delta_{\rm g}^{(2)}(\boldsymbol{k}) \tag{1.65}$$

We can now write down the definition of the galaxy bispectrum and write down the contributions, which will start at fourth order through a correlator of a second order contribution (quadratic in the fields in the above equation) with two linear contributions.

$$\begin{aligned}(2\pi)^3\delta^{({\rm D})}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)B_{\rm g}(k_1,k_2,k_3) &= \langle\delta_{\rm g}(\boldsymbol{k}_1)\delta_{\rm g}(\boldsymbol{k}_2)\delta_{\rm g}(\boldsymbol{k}_3)\rangle \\ &= \left\langle\delta_{\rm g}^{(2)}(\boldsymbol{k}_1)\delta_{\rm g}^{(1)}(\boldsymbol{k}_2)\delta_{\rm g}^{(1)}(\boldsymbol{k}_3)\right\rangle + 2\text{ cyc.}\end{aligned} \tag{1.66}$$

Using the implicit definition of the linear and quadratic bias contributions in Eq. (1.65) we obtain

$$\left\langle\delta_{\rm g}^{(2)}(\boldsymbol{k}_1)\delta_{\rm g}^{(1)}(\boldsymbol{k}_2)\delta_{\rm g}^{(1)}(\boldsymbol{k}_3)\right\rangle = \frac{1}{2}b_1^2b_2\left\langle\int\frac{{\rm d}^3q}{(2\pi)^3}\delta(\boldsymbol{q})\delta(\boldsymbol{k}_1-\boldsymbol{q})\delta(\boldsymbol{k}_2)\delta(\boldsymbol{k}_3)\right\rangle \tag{1.67}$$

Wick theorem would allow for three different pairings, however the $\langle\delta(\boldsymbol{q})\delta(\boldsymbol{k}_1-\boldsymbol{q})\rangle$ correlator only contributes to the homogenous $\boldsymbol{k}_1 = 0$ mode that is irrelevant for our purposes. The remaining two correlators link the $\boldsymbol{q}$-mode with either $\boldsymbol{k}_2$ or $\boldsymbol{k}_3$. Both of these give the same result as we can exchange $\boldsymbol{q} \leftrightarrow \boldsymbol{k}-\boldsymbol{q}$ in the momentum integral. We obtain

$$\begin{aligned}\left\langle\delta_{\rm g}^{(2)}(\boldsymbol{k}_1)\delta_{\rm g}^{(1)}(\boldsymbol{k}_2)\delta_{\rm g}^{(1)}(\boldsymbol{k}_3)\right\rangle =& b_1^2b_2\int\frac{{\rm d}^3q}{(2\pi)^3}(2\pi)^3\delta^{({\rm D})}(\boldsymbol{q}+\boldsymbol{k}_2)P(k_2)(2\pi)^3\delta^{({\rm D})}(\boldsymbol{k}_1-\boldsymbol{q}+\boldsymbol{k}_3)P(k_3) \\ =& b_1^2b_2(2\pi)^3\delta^{({\rm D})}(\boldsymbol{k}_1+\boldsymbol{k}_2+\boldsymbol{k}_3)P(k_2)P(k_3)\end{aligned} \tag{1.68}$$

One of the delta functions collapsed the momentum integral and the remaining momentum conserving delta function just reproduces the one required in the definition of the bispectrum in Eq. (1.66). Thus, we finally obtain for the bispectrum in the simple quadratic bias model

$$B_{\rm g}(k_1,k_2,k_3) = b_1^2b_2\left[P(k_1)P(k_2) + 2\text{ cyc.}\right]. \tag{1.69}$$

We will discuss a prescription to derive this result from Feynman diagrams later this term.

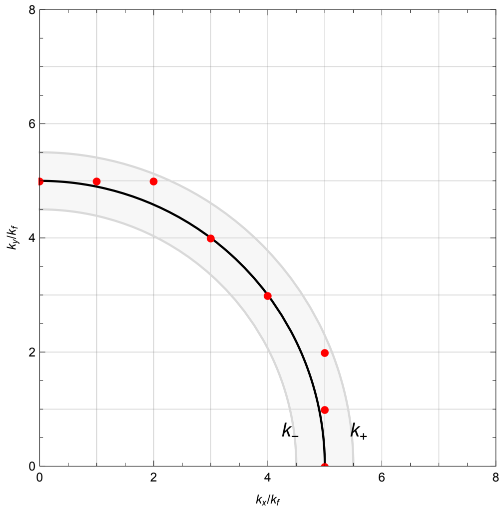

Figure 1.3: Discrete Fourier grid and discrete modes (red points) contributing to a wavenumber bin (gray shaded region) centered around k.

1.5 Estimators and Cosmic Variance

Given a dataset, i.e., a realization of the underlying statistical ensemble, we need to estimate the relevant statistics and the uncertainty of that estimate.

In observations and numerical simulations the volume is limited, which leaves us with finite Fourier modes, the smallest of them given by the **fundamental mode** $k_\mathrm{f} = 2\pi/L$ and the corresponding volume of the **fundamental cell** is $V_\mathrm{f} = (2\pi)^3/V$. The Dirac Delta function rewritten for discrete k as

$$\delta^{(\mathrm{D})}(\boldsymbol{k_i} - \boldsymbol{k_j}) = \delta^{(\mathrm{D})}\left((\boldsymbol{i} - \boldsymbol{j})k_\mathrm{f}\right) = \frac{1}{k_\mathrm{f}^3}\delta^{(\mathrm{D})}(\boldsymbol{i} - \boldsymbol{j}) = \frac{V}{(2\pi)^3}\delta^{(\mathrm{K})}_{\boldsymbol{i},\boldsymbol{j}} \tag{1.70}$$

All the Fourier modes can now be expressed as $\boldsymbol{k_i} = \boldsymbol{i}k_\mathrm{f}$, where $\boldsymbol{i} = (i_x, i_y, i_z)$ is an integer vector.

1.5.1 Power Spectrum estimator and variance

The power spectrum for discrete cells is thus given by

$$V\delta^{(\mathrm{K})}_{\boldsymbol{k},\boldsymbol{k}'}P\left(|\boldsymbol{k}|\right) = \langle\delta(\boldsymbol{k})\delta^*(\boldsymbol{k}')\rangle \tag{1.71}$$

We estimate the power spectrum in bins, i.e. spherical shells of width dk corresponding to an intervall in wavevector magnitude $[k\pm] = [k_-, k_+) = [k - \mathrm{d}k/2, k + \mathrm{d}k/2)$ centered at k as depicted in Fig. 1.5 averaging all possible directions for the wavevector (making use of statistical isotropy)

$$\hat{P}(k) = \frac{1}{N_k V}\sum_{\boldsymbol{k_i}\in[k_-,k_+)}\delta(\boldsymbol{k_i})\delta^*(\boldsymbol{k_i}), \tag{1.72}$$

where N_k is the number of cells in the k-bin. Effectively, we are estimating the variance of δ in a shell using N_k observations. Note that the estimator is unbiased since $\langle\hat{P}\rangle = P$. Obviously there is some freedom in choosing the configuration of the bins. Basically, the above estimator gives the mean power in the shell $[k_\pm]$, and thus it is advisable to average the theory calculations over the same bins for a fair comparison. The number of grid cells in the bin is given by the shell volume

$$V_s = \int_{[k\pm]}\mathrm{d}^3q\int\mathrm{d}^3q'\delta^{(\mathrm{D})}(\boldsymbol{q} + \boldsymbol{q}') \tag{1.73}$$

divided by the volume of the fundamental cell

$$N_k = \frac{V_s}{V_{\rm f}} = \frac{4\pi k^2 {\rm d}k}{V_{\rm f}} = \frac{4\pi k^3 {\rm d}\ln k}{V_{\rm f}}. \tag{1.74}$$

Let us now calculate the variance of the power spectrum estimator

$$\begin{aligned}\left\langle \hat{P}^2(k)\right\rangle - \left\langle \hat{P}(k)\right\rangle^2 =& \frac{1}{N_k^2 V^2} \sum_{\boldsymbol{k_i},\boldsymbol{k_j}\in[k\pm]} \langle \delta(\boldsymbol{k_i})\delta(-\boldsymbol{k_i})\delta(\boldsymbol{k_j})\delta(-\boldsymbol{k_j})\rangle - P^2(k)\\ =& \frac{1}{N_k^2} \sum_{\boldsymbol{k_i},\boldsymbol{k_j}\in[k\pm]} P(\boldsymbol{k_i})P(\boldsymbol{k_j}) + \frac{2}{N_k^2}\sum_{\boldsymbol{k_i}\in[k\pm]} P^2(|\boldsymbol{k_i}|) - P^2(k)\\ =& \frac{2}{N_k}P^2(k) = \frac{2V_f}{4\pi k^3 {\rm d}\ln k}P^2(k)\end{aligned} \tag{1.75}$$

Here we assumed Gaussianity of the underlying field, in which case the covariance matrix is diagonal, i.e., power spectrum estimates for distinct wavenumbers are independent. In the more general case of non-Gaussian fields, the connected trispectrum contributes as well. Note also that the covariance matrix of power spectra of non-overlapping bins is diagonal, i.e., estimates of power in non-overlapping bins are independent.

1.5.2 Bispectrum*

The bispectrum estimator for a fixed configuration $\{k_1, k_2, \mu = \boldsymbol{k}_1 \cdot \boldsymbol{k}_2\}$ can be estimated as

$$\hat{B}(k_1,k_2,\mu) = \frac{1}{N_{\rm tr}V} \sum_{\boldsymbol{k_i}\in[k\pm]} \sum_{\boldsymbol{j}k_{\rm f}\in[k\pm,\mu\pm]} \delta(\boldsymbol{k_i})\delta(\boldsymbol{k_j})\delta(-\boldsymbol{k_i}-\boldsymbol{k_i}) \tag{1.76}$$

the estimator is unbiased since $\left\langle \hat{B}\right\rangle = B$. The number of triangles in the bin is given by the shell volume divided by the volume of the fundamental cell squared

$$N_{\rm tr} = \frac{V_{123}}{V_{\rm f}^2} = \frac{8\pi^2 k_1 k_2 k_3 ({\rm d}k)^3}{V_{\rm f}^2} \tag{1.77}$$

For the variance of the bispectrum estimator we have

$$\begin{aligned}\left\langle \hat{B}^2(k_1,k_2,\mu)\right\rangle - \left\langle \hat{B}(k_1,k_2,\mu)\right\rangle^2 =& \frac{1}{N_{\rm tr}^2 V^2} \sum_{\boldsymbol{i},\boldsymbol{j},\boldsymbol{l},\boldsymbol{m}} \langle \delta(\boldsymbol{k_i})\delta(\boldsymbol{k_j})\delta(-\boldsymbol{k_i}-\boldsymbol{k_j})\delta(\boldsymbol{k_l})\delta(\boldsymbol{k_m})\delta(-\boldsymbol{k_l}-\boldsymbol{k_m})\rangle\\ &- B^2(k_1,k_2,\mu)\\ =& s_{123}\frac{V}{N_{\rm tr}^2}\sum_{\boldsymbol{i},\boldsymbol{j}} P(\boldsymbol{k_i})P(\boldsymbol{k_j})P(-\boldsymbol{k_i}-\boldsymbol{k_j})\\ =& s_{123}\frac{V}{N_{\rm tr}}P(k_1)P(k_2)P(k_3)\\ =& s_{123}\frac{(2\pi)^3 V_{\rm f}}{8\pi^2 k_1 k_2 k_3 ({\rm d}k)^3 {\rm d}\mu}P(k_1)P(k_2)P(k_3)\end{aligned} \tag{1.78}$$

The factor s_{123} takes on values of $6, 2, 1$ for general, isosceles and equilateral triangles. This is a simple consequence of the fact, that for equilateral triangles the k-modes are indistinguishable. We again assumed Gaussianity, for which $B = 0$.

1.6 Random Fields on the Sphere

We have discussed inhomogeneities in 3D space in some detail. Let us now consider the statistics required to describe anisotropies on S^2. These statistics are natural for our observations of the Universe. For instance the CMB photons are released from a spherical shell surrounding us and we naturally observe the temperature distribution on the sky as a function of azimuth and polar angle. More generally, all observations naturally live on shells of varying distance from the observer and a 3D or flat 2D description is only appropriate in small volumes. Brink & Satchler *Angular Momentum* is a good reference for the spherical harmonics and so is R. Durrer *The Cosmic Microwave Background*, in particular it's appendices.

1.6.1 Spherical Harmonics Expansion

Anisotropies on a sphere, for instance CMB fluctuations, can be expanded in **spherical harmonics** $Y_{lm}(\hat{\boldsymbol{n}})$, which are a basis for square integrable functions on S^2

$$f(\hat{\boldsymbol{n}}) = \sum_{l,m} f_{lm} Y_{lm}(\hat{\boldsymbol{n}})\,. \tag{1.79}$$

Here $\hat{\boldsymbol{n}}$ is a unit vector on the sphere $\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}} = 1$. The spherical harmonics are the position space representation of the eigenstates of total angular momentum $\hat{L}^2 = -\nabla^2$ and azimuthal angular momentum $\hat{L}_z = -i\partial_\phi$

$$\hat{L}^2 Y_{lm} = l(l+1) Y_{lm}\,, \tag{1.80}$$

$$\hat{L}_z Y_{lm} = m Y_{lm}\,. \tag{1.81}$$

Note in particular, that the two operators commute $[\hat{L}^2, \hat{L}_z] = 0$. Under parity, the spherical harmonics transform as $Y_{lm}(-\hat{\boldsymbol{n}}) = (-1)^l Y_{lm}(\hat{\boldsymbol{n}})$ and in our convention $Y^*_{l,m} = (-1)^m Y_{l,-m}$.
The spherical harmonics are products of associated Legendre polynomials[3] and a azimuthal phase factor

$$Y_{lm}(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}}\; P_l^m(\cos\theta)\;\exp\left[\mathrm{i}m\phi\right] \tag{1.84}$$

The spherical harmonics are orthonormal on the sphere

$$\int \mathrm{d}^2\Omega\; Y_{lm}(\hat{\boldsymbol{n}}) Y^*_{l'm'}(\hat{\boldsymbol{n}}) = \delta^{(\mathrm{K})}_{ll'}\delta^{(\mathrm{K})}_{mm'}\,, \tag{1.85}$$

such that the **spherical multipole coefficients** of the expansion can be calculated as

$$f_{lm} = \int \mathrm{d}^2\Omega\; f(\hat{\boldsymbol{n}}) Y^*_{lm}(\hat{\boldsymbol{n}})\,. \tag{1.86}$$

For real valued fields we thus have $f^*_{l,m} = (-1)^m f_{l,-m}$. For a practial implementation the HEALPIX[4] formalism is commonly used. A easy to use Python implementation for dealing with CMB maps, calculation of f_{lm} and CMB power spectra is the HEALPY package.
Let us now calculate the angular correlation function between two points $\hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{n}}'$ on the sphere. By statistical isotropy this quantity can only depend on the (cosine of) the enclosed angle $\mu = \cos\theta = \hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}}'$ and thus allows for an expansion in Legendre polynomials

$$C(\theta) = \langle f(\hat{\boldsymbol{n}}), f(\hat{\boldsymbol{n}}')\rangle = \sum_l \frac{2l+1}{4\pi} C_l P_l(\mu) = \sum_{lm} C_l Y_{lm}(\hat{\boldsymbol{n}}) Y^*_{lm}(\hat{\boldsymbol{n}}') \tag{1.87}$$

[3]Legendre polynomials can be derived from Rodrigues' formula

$$P_l(x) = \frac{1}{2^l l!}\frac{\mathrm{d}^l}{\mathrm{d}x^l}(x^2-1)^l \tag{1.82}$$

whereas the associated Legendre polynomials satisfy

$$P_l^m(x) = (-1)^m(1-x^2)^{m/2}\frac{\mathrm{d}^m}{\mathrm{d}x^m}P_l(x) = \frac{(-1)^m}{2^l l!}(1-x^2)^{m/2}\;\frac{\mathrm{d}^{l+m}}{\mathrm{d}x^{l+m}}(x^2-1)^l \tag{1.83}$$

thus $P_l^0(x) = P_l(x)$.

[4]HEALPIX: Hierarchical Equal Area isoLatitude Pixelization https://healpix.jpl.nasa.gov

here we used the addition theorem for spherical harmonics.[5] At the same time we could have written the angular correlation function naively in terms of the expansion in spherical harmonics Eq. (1.79)

$$C(\theta) = \langle f(\hat{\boldsymbol{n}}), f(\hat{\boldsymbol{n}}')\rangle = \sum_{lm}\sum_{l'm'} \langle f_{lm} f^*_{l'm'}\rangle\, Y_{lm}(\hat{\boldsymbol{n}}) Y^*_{l'm'}(\hat{\boldsymbol{n}}') \tag{1.91}$$

From this we can directly deduce that

$$\langle f_{lm} f^*_{l'm'}\rangle = C_l \delta^{(\mathrm{K})}_{ll'} \delta^{(\mathrm{K})}_{mm'} . \tag{1.92}$$

For a more direct proof, we can multiply with $Y^*_{l_1 m_1}(\hat{\boldsymbol{n}}) Y_{l_2 m_2}(\hat{\boldsymbol{n}}')$ and integrate over Ω and Ω'.

1.6.2 Rotations

For statistics on the sphere, statistical isotropy is the relevant symmetry. To understand what independence under spatial rotations entails for the statistics, let us formalize their impact on the f_{lm}. A generic spatial rotation can be described by three Euler angles (α, β, γ) and the rotation $\hat{D}(\alpha, \beta, \gamma)$ is given by *i*) a rotation around the original z-axis by γ, *ii*) a rotation around the original y-axis by β and finally *iii*) another rotation around the original z-axis by α. Obviously the inverse rotation is undoing the above in reverse order, i.e., $\hat{D}^{-1}(\alpha, \beta, \gamma) = \hat{D}(-\gamma, -\beta, -\alpha)$ and $\hat{D} f(\hat{\boldsymbol{n}}) = f(\hat{D}^{-1}\hat{\boldsymbol{n}})$

Under a rotation γ around the z-axis, the spherical harmonics transform as

$$[\hat{D}(0, 0, \gamma) Y_{lm}](\theta, \phi) = Y_{lm}(\theta, \phi - \gamma) = \exp\left[-\mathrm{i} m\gamma\right] Y_{lm}(\theta, \phi) . \tag{1.93}$$

Thus the $\hat{L}_i$ are the generators of rotations about the coordinate axes.

$$\hat{D}(\alpha, \beta, \gamma) = \exp\left[-\mathrm{i}\alpha \hat{L}_z\right] \exp\left[-\mathrm{i}\beta \hat{L}_y\right] \exp\left[-\mathrm{i}\gamma \hat{L}_z\right] \tag{1.94}$$

The total and azimuthal angular momentum operators commute. Thus, the total angular momentum is the same in the original and rotated frame and the rotated Y_{lm} can be expressed as a linear combination of the original $(2l + 1)$ Y_{lm}

$$\hat{D} Y_{lm} = \sum_{m'} D^l_{m'm} Y_{lm'} \tag{1.95}$$

The transformation matrices are the **Wigner D** matrices.

Acting with the rotation operator on a test function f we have

$$\hat{D} f(\hat{\boldsymbol{n}}) = \sum_{l,m} f_{lm} \hat{D} Y_{lm} = \sum_{l,m'}\sum_{m} f_{lm} D^l_{m'm} Y_{lm'}(\hat{\boldsymbol{n}}) = \sum_{l,m'} f_{lm'} Y_{lm'}(\hat{\boldsymbol{n}}) \tag{1.96}$$

Thus the multipole coefficients transform as

$$\tilde{f}_{lm'} = \sum_{m} D_{m'm} f_{lm} , \tag{1.97}$$

[5] The addition theorem for spherical harmonics states

$$P_l(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}') = \frac{4\pi}{2l+1} \sum_m Y_{lm}(\hat{\boldsymbol{n}}) Y^*_{lm}(\hat{\boldsymbol{n}}') \tag{1.88}$$

$$\begin{aligned} \sum_n Y_{ln}(D^{-1}\hat{\boldsymbol{n}}) Y^*_{ln}(D^{-1}\hat{\boldsymbol{n}}') &= \sum_{m,m',n} D^l_{mn} Y_{lm}(\hat{\boldsymbol{n}}) D^{l*}_{m'n} Y_{lm'}(\hat{\boldsymbol{n}}') \\ &= \sum_m Y_{lm}(\hat{\boldsymbol{n}}) Y^*_{lm}(\hat{\boldsymbol{n}}') \end{aligned} \tag{1.89}$$

Let us now move to the special case $D^{-1}\hat{\boldsymbol{n}}' = \hat{\boldsymbol{z}}$. Using that $Y_{lm}(\hat{\boldsymbol{z}}) = \sqrt{2l + 1/4\pi}\,\delta^{(\mathrm{K})}_{m0}$ we have

$$\sum_m Y_{lm}(\hat{\boldsymbol{n}}) Y_{lm}(\hat{\boldsymbol{n}}') = \sqrt{2l + 1/4\pi}\, Y_{l0}(D^{-1}\hat{\boldsymbol{n}}) = \frac{2l+1}{4\pi} P_l(\hat{\boldsymbol{n}} \cdot \hat{\boldsymbol{n}}') \tag{1.90}$$

Here we used that the polar angle of $D^{-1}\hat{\boldsymbol{n}}$ is the enclosed angle of $\hat{\boldsymbol{n}}$ and $\hat{\boldsymbol{n}}'$

where the index of $D_{m'm}$ that is summed over is exchanged wrt Eq. (1.95). The rotation operator is unitary

$$\hat{D}^{-1} = \hat{D}^{\dagger}\,. \tag{1.98}$$

This property is inherited by the Wigner D matrices

$$\begin{aligned}\delta^{(\mathrm{K})}_{mm'} &= \int \mathrm{d}\Omega_n\ Y_{lm}(\hat{\boldsymbol{n}})Y^*_{lm'}(\hat{\boldsymbol{n}}) = \int \mathrm{d}\Omega_s\ Y_{lm}(D^{-1}\hat{\boldsymbol{s}})Y^*_{lm'}(D^{-1}\hat{\boldsymbol{s}}) \\ &= \sum_{n,n'} \int \mathrm{d}\Omega_s\ D^l_{nm}Y_{ln}(\hat{\boldsymbol{s}})D^{l*}_{n'm'}Y^*_{ln'}(\hat{\boldsymbol{s}}) = \sum_n D^l_{nm}D^{l*}_{nm'}\end{aligned} \tag{1.99}$$

$$\begin{aligned}\delta^{(\mathrm{K})}_{mm'} &= \int \mathrm{d}\Omega_n\ Y_{lm}(\hat{\boldsymbol{n}})Y^*_{lm'}(\hat{\boldsymbol{n}}) = \int \mathrm{d}\Omega_s\ Y_{lm}(D\hat{\boldsymbol{s}})Y^*_{lm'}(D\hat{\boldsymbol{s}}) \\ &= \sum_{n,n'} \int \mathrm{d}\Omega_s\ D^{*l}_{mn}Y_{ln}(\hat{\boldsymbol{s}})D^{l}_{m'n'}Y^*_{ln'}(\hat{\boldsymbol{s}}) = \sum_n D^{l*}_{mn}D^{l}_{m'n}\end{aligned} \tag{1.100}$$

The $m = 0$ components of the Wigner matrices are related to the spherical harmonics themselves. To see that, let's consider a unit vector $\hat{\boldsymbol{n}}$ with components (θ, ϕ). This unit vector can obviously be created by rotating a unit vector in z-direction by $D(\phi, \theta, 0)$

$$Y_{lm}(\hat{\boldsymbol{n}}) = Y_{lm}(\hat{D}(\phi,\theta,0)\hat{\boldsymbol{z}}) = D^{l*}_{mm'}Y_{lm'}(\hat{\boldsymbol{z}}) = \sqrt{\frac{2l+1}{4\pi}}D^{l*}_{m0}(\phi,\theta,0) \tag{1.101}$$

Here we used that $P^m_l(1) = \delta^{(\mathrm{K})}_{m0}$ and $Y_{lm}(\hat{\boldsymbol{z}}) = \sqrt{2l+1/4\pi}\delta^{(\mathrm{K})}_{m0}$ and finally see

$$D^l_{m0}(\phi,\theta,0) = \sqrt{\frac{4\pi}{2l+1}}Y^*_{lm}(\hat{\boldsymbol{n}})\,. \tag{1.102}$$

Let us finish up this section by studying the implications of statistical isotropy on the power spectrum of f_{lm}

$$\left\langle \tilde{f}_{lm}\tilde{f}^*_{l'm'}\right\rangle = \sum_{n,n'} D^l_{mn}D^{l'*}_{m'n'}\left\langle f_{ln}f^*_{l'n'}\right\rangle \tag{1.103}$$

This clearly requires $\langle f_{ln}f^*_{l'n'}\rangle = C_l\delta^{(\mathrm{K})}_{ll'}\delta^{(\mathrm{K})}_{nn'}$

$$C_l\delta^{(\mathrm{K})}_{ll'}\delta^{(\mathrm{K})}_{mm'} = \sum_n D^l_{mn}D^{l*}_{m'n}C_l\delta^{(\mathrm{K})}_{ll'} = C_l\delta^{(\mathrm{K})}_{ll'}\delta^{(\mathrm{K})}_{mm'} \tag{1.104}$$

Finally the structure of the angular power spectrum can be directly deduced from the statistical isotropy of the angular correlation function by using the definition of the f_{lm}

$$\left\langle \tilde{f}_{lm}\tilde{f}^*_{l'm'}\right\rangle = \int \mathrm{d}\Omega \int \mathrm{d}\Omega' C(\hat{\boldsymbol{n}}\cdot\hat{\boldsymbol{n}}')Y^*_{lm}(\hat{\boldsymbol{n}})Y_{l'm'}(\hat{\boldsymbol{n}}') \tag{1.105}$$

Let's split the angular integration into an reference angle $\hat{\boldsymbol{n}}$ and a relative angle defined wrt $\hat{\boldsymbol{n}}$. Clearly, the correlation function can only depend on the relative angle θ_r. Defining a rotation by $\hat{\boldsymbol{n}} = \hat{D}\hat{\boldsymbol{z}}$ we have $\hat{\boldsymbol{n}}^\mathrm{T}\hat{\boldsymbol{n}}' = (D^{-1}\hat{\boldsymbol{n}})^\mathrm{T}D^{-1}\hat{\boldsymbol{n}}' = \hat{\boldsymbol{z}}^\mathrm{T}D^{-1}\hat{\boldsymbol{n}}'$

$$\begin{aligned}\left\langle \tilde{f}_{lm}\tilde{f}^*_{l'm'}\right\rangle &= \int \mathrm{d}\Omega \int \mathrm{d}\Omega_\mathrm{r} C(\theta_\mathrm{r})Y^*_{lm}(D\hat{\boldsymbol{z}})Y_{l'm'}(DD^{-1}\hat{\boldsymbol{n}}') \\ &= \sum_{\tilde{m},\tilde{m}'} \int \mathrm{d}\Omega \int \mathrm{d}\Omega_\mathrm{r} C(\theta_\mathrm{r})D^l_{m\tilde{m}}Y^*_{l\tilde{m}}(\hat{\boldsymbol{z}})D^{l*}_{m'\tilde{m}'}Y_{l'\tilde{m}'}(D^{-1}\hat{\boldsymbol{n}}') \\ &= \sqrt{\frac{2l+1}{4\pi}\frac{2l'+1}{4\pi}} \int \mathrm{d}\Omega\ D^{l*}_{m0}D^l_{m'0} \int \mathrm{d}\Omega_\mathrm{r} C(\theta_\mathrm{r})P_{l'}(\cos\theta_\mathrm{r}) \\ &= \delta^{(\mathrm{K})}_{ll'}\delta^{(\mathrm{K})}_{mm'} \int \mathrm{d}\Omega_\mathrm{r} C(\theta_\mathrm{r})P_{l'}(\cos\theta_\mathrm{r})\end{aligned} \tag{1.106}$$

Thus

$$C_l = \int \mathrm{d}\Omega_\mathrm{r} C(\theta_\mathrm{r})P_{l'}(\cos\theta_\mathrm{r}) = 2\pi \int \mathrm{d}\mu C(\mu)P_l(\mu) \tag{1.107}$$

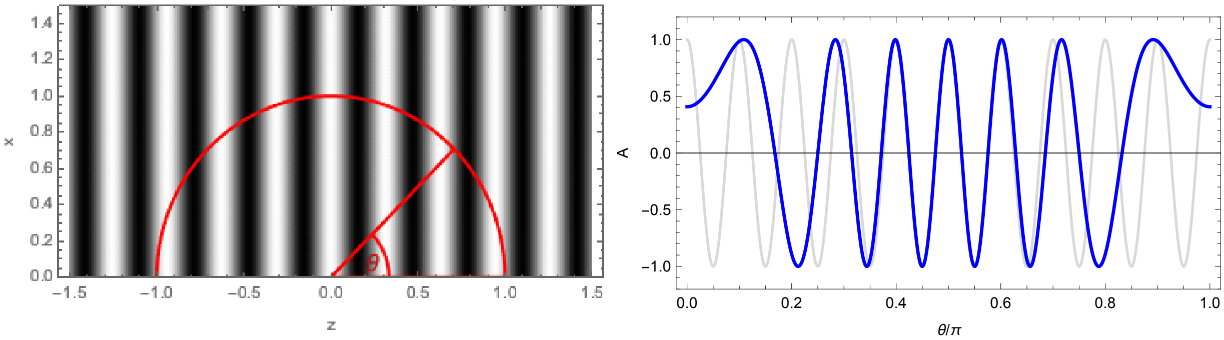

Figure 1.4: *Left panel:* Plane wave in z-direction and intersection with a sphere. *Right panel:* Amplitude of a plane wave with $\boldsymbol{k} \parallel \hat{\boldsymbol{z}}$ on the surface of the sphere as a function of polar angle (blue line). Perpendicilar to the wavevector, the plane wave contributes to angular modes with $l \approx rk$ (gray line), whereas along the wavevector, the plane wave contributes to angular fluctuations of longer wavelength, smaller angular wavenumber $l < kr$.

1.6.3 Projection of 3D fields

Let us now express the C_l's in terms of the 3D power spectrum. Let us consider the three dimensional quantity f on a sphere of radius r surrounding the observer. You can for instance think of it as an expansion of the CMB temperature fluctuations, where r would thus be the distance from us to the last scattering surface

$$f(\hat{\boldsymbol{n}}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} f(\boldsymbol{k}) \exp\left[-\mathrm{i}\hat{\boldsymbol{k}} \cdot \hat{\boldsymbol{n}} kr\right]. \tag{1.108}$$

Using the plane wave or Rayleigh expansion

$$\exp\left[\mathrm{i}\boldsymbol{k} \cdot \boldsymbol{r}\right] = 4\pi \sum_{lm} \mathrm{i}^l j_l(kr) Y_{lm}^*(\hat{\boldsymbol{k}}) Y_{lm}(\hat{\boldsymbol{r}}) = \sum_l (2l+1)\mathrm{i}^l j_l(kr) P_l(\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{k}}) \tag{1.109}$$

we obtain

$$f_{lm} = 4\pi \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\ f(\boldsymbol{k}) \sum_{l'm'} (-\mathrm{i})^{l'} j_{l'}(kr) Y_{l'm'}^*(\hat{\boldsymbol{k}}) \int \mathrm{d}\Omega\ Y_{l'm'}(\hat{\boldsymbol{n}}) Y_{lm}^*(\hat{\boldsymbol{n}}) \tag{1.110}$$

$$= 4\pi \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3}\ f(\boldsymbol{k})(-\mathrm{i})^l j_l(kr) Y_{lm}^*(\hat{\boldsymbol{k}}) \tag{1.111}$$

We are now well prepared to calculate the relation between the angular and the three dimensional power spectrum

$$\langle f_{lm} f_{l'm'}^* \rangle = (4\pi)^2 \int \frac{\mathrm{d}^3\boldsymbol{k}}{(2\pi)^3} \int \frac{\mathrm{d}^3\boldsymbol{k}'}{(2\pi)^3} \langle f(\boldsymbol{k}) f^*(\boldsymbol{k}') \rangle\, Y_{lm}(\hat{\boldsymbol{k}}) Y_{l'm'}^*(\hat{\boldsymbol{k}}') j_l(kr) j_{l'}(k'r)(-\mathrm{i})^l \mathrm{i}^{l'} \tag{1.112}$$

$$= 4\pi \int \frac{\mathrm{d}k k^2}{2\pi^2} P(k) j_l(kr) j_{l'}(kr)(-\mathrm{i})^l \mathrm{i}^{l'} \int \mathrm{d}\Omega_{\boldsymbol{k}}\ Y_{lm}(\hat{\boldsymbol{k}}) Y_{l'm'}^*(\hat{\boldsymbol{k}}) \tag{1.113}$$

$$= 4\pi \delta_{ll'}^{(\mathrm{K})} \delta_{mm'}^{(\mathrm{K})} \int \frac{\mathrm{d}k k^2}{2\pi^2} j_l^2(kr) P(k) \tag{1.114}$$

As above, the structure of the angular power spectrum is revealed to be $C_l \delta_{ll'}^{(\mathrm{K})} \delta_{mm'}^{(\mathrm{K})}$, where

$$C_l = 4\pi \int \frac{\mathrm{d}k k^2}{2\pi^2} j_l^2(kr) P(k) = 4\pi \int \mathrm{d}\ln k\ \Delta^2(k) j_l^2(kr)\,. \tag{1.115}$$

In Fig. 1.4 we consider the contributions of a plane wave with wavenumber k travelling in z-direction to the spherical multipole coefficients on an intersected sphere of radius r. Transverse to the $\hat{\boldsymbol{z}}$-axis the plane wave

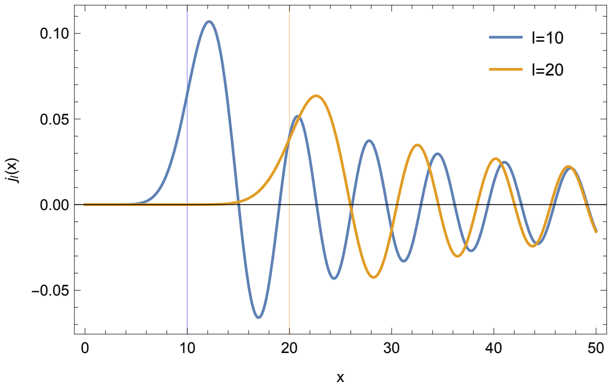

Figure 1.5: Spherical Bessel functions $j_l(x)$ for $l = 10$ and $l = 20$. At low x the Bessel functions vanish indicating that modes with wavelength exceeding l/r do not contribute to the angular fluctuations You can see a pronounced peak at $x = l$ arising from $k \approx l/r$ transverse to the line of sight and oscillatory contributions from modes appearing as longer angular wavelenghts along the line of sight.

contributes to the $l \approx kr$ spherical mode, but along the $\hat{\boldsymbol{z}}$ the plane wave has a less oscillatory contribution on the surface of the sphere, thus contributing to lower angular modes $l < kr$. Correspondingly, the spherical bessel functions in Fig. 1.5 have a pronounced positive peak at $x = kr \approx l$ and oscillatory contributions from plane waves with $k > l/r$.

1.7 CMB Estimators and Covariance

As we discussed for the 3D power spectrum $P(k)$ above in Sec. 1.5.1, we need to estimate statistics given the single realization of the cosmic fields at our disposal using ergodic theorem using an estimator. For the 3D power spectrum we averaged over directions of Fourier modes using that the power spectrum is a function of the magnitude of wavenumber only. Here the angular power spectrum can be estimated as the average over all m-modes given l

$$\hat{C}_l = \frac{1}{2l+1} \sum_m f_{lm} f^*_{lm} \tag{1.116}$$

The estimator is unbiased

$$\left\langle \hat{C}_l \right\rangle = \frac{1}{2l+1} \sum_m \langle f_{lm} f^*_{lm} \rangle = \frac{1}{2l+1} \sum_m C_l = C_l \tag{1.117}$$

Finally, the fact that we have only $2l+1$ modes to estimate the variance of the field leads to an uncertainty of the estimator. The covariance between the power spectrum estimators for l and l' can be estimated as

$$\begin{aligned} \mathrm{cov}(\hat{C}_l, \hat{C}_{l'}) =& \left\langle \hat{C}_l \hat{C}_{l'} \right\rangle - \left\langle \hat{C}_l \right\rangle \left\langle \hat{C}_{l'} \right\rangle \\ =& \frac{1}{(2l+1)(2l'+1)} \sum_{m,m'} \langle f_{lm} f^*_{lm} f_{l'm'} f^*_{l'm'} \rangle - \langle C_l \rangle \langle C_{l'} \rangle \\ =& \frac{1}{(2l+1)(2l'+1)} \sum_{m,m'} 2C_l^2 \delta^{(\mathrm{K})}_{ll'} \delta^{(\mathrm{K})}_{mm'} = \frac{2}{2l+1} C_l^2 \delta^{(\mathrm{K})}_{ll'} \end{aligned} \tag{1.118}$$

Here we used Wick theorem to express the four point function as a sum over three products of two-point functions.[6] One of the three terms cancels the product of the two C_ls. We see that for a Gaussian field

[6] As stated below Eq. (1.86) we have $f^*_{l,m} = (-1)^m f_{l,-m}$ and thus

$$\langle f_{l,m} f_{l',m'} \rangle = (-1)^{m'} \left\langle f_{l,m} f^*_{l',-m'} \right\rangle = (-1)^{m'} C_l \delta^{(\mathrm{K})}_{l,l'} \delta^{(\mathrm{K})}_{m,-m'}$$

the covariance matrix is diagonal, i.e., that the estimate for the power at different angular wavenumbers is independent. When performing data analysis on the CMB, one can thus work with a diagonal covariance matrix. The relative statistical error of a given angular wavenumber $\Delta C_l/C_l = \sqrt{2/(2l+1)}$ decreases with increasing wavenumber. This explains the larger errorbars on the left of the CMB power spectrum plots.

1.8 Angular Bispectrum

We have already discussed a simple model that would introduce a non-vanishing bispectrum in the late time Universe. Modifications to the simplest inflationary models can produce detectable levels of non-Gaussianity in the early universe, which can be detected in the CMB. Let us develop the relevant statistics for quantifying the bispectrum on S^2. The angular bispectrum is given by

$$\langle f_{lm} f_{l'm'} f_{l''m''} \rangle = B_{ll'l''} \begin{pmatrix} l & l' & l'' \\ m & m' & m'' \end{pmatrix} = b_{ll'l''} \mathcal{G}^{ll'l''}_{mm'm''} \tag{1.119}$$

where we introduced the **Wigner 3j symbols**. The second equality is based on the definition of the triple integral over spherical harmonics, the so called **Gaunt integral**

$$\begin{aligned} \mathcal{G}^{ll'l''}_{mm'm''} &= \int \mathrm{d}\Omega \; Y_{lm}(\hat{\boldsymbol{n}}) Y_{l'm'}(\hat{\boldsymbol{n}}) Y_{l''m''}(\hat{\boldsymbol{n}}) \,, \\ &= \sqrt{\frac{(2l+1)(2l'+1)(2l''+1)}{4\pi}} \begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} l & l' & l'' \\ m & m' & m'' \end{pmatrix} . \end{aligned} \tag{1.120}$$

These results should be familiar from angular momentum coupling in quantum mechanics. The product state of two systems $\big|j_1 j_2 m_1 m_2\big\rangle = \big|j_1 m_1\big\rangle \otimes \big|j_2 m_2\big\rangle$ can also be expressed in terms of the total angular momentum $\big|j_1 j_2 J M\big\rangle$. The two representations are related to each other using the identity operator $\mathbb{1} = \sum_x |x\rangle\langle x|$

$$\big|j_1 j_2 m_1 m_2\big\rangle = \sum_{J,M} \big\langle j_1 j_2 J M \big| j_1 j_2 m_1 m_2 \big\rangle \big| j_1 j_2 J M \big\rangle \,, \tag{1.121}$$

$$\big|j_1 j_2 J M\big\rangle = \sum_{m_1,m_2} \big\langle j_1 j_2 m_1 m_2 \big| j_1 j_2 J M \big\rangle \big| j_1 j_2 m_1 m_2 \big\rangle \,. \tag{1.122}$$

The real expansion coefficients are the **Clebsch-Gordan coefficients**. Rotatet states can be expressed in terms of the original states using the Wigner D matrices introduced above

$$\hat{D}\big|j_1 j_2 J M\big\rangle = \sum_N D^J_{NM} \big|j_1 j_2 J N\big\rangle \tag{1.123}$$

$$\hat{D}\big|j_1 j_2 m_1 m_2\big\rangle = \sum_{m_1',m_2'} D^{j_1}_{n_1 m_1} D^{j_2}_{n_2 m_2} \big|j_1 j_2 m_1 m_2\big\rangle \tag{1.124}$$

We can now employ the fact that the rotation operator is unitary $\hat{D}^\dagger \hat{D}$

$$\begin{aligned} \big\langle j_1 j_2 m_1 m_2 \big| j_1 j_2 J M \big\rangle &= \big\langle j_1 j_2 m_1 m_2 \big| \hat{D}^\dagger \hat{D} \big| j_1 j_2 J M \big\rangle \\ &= \sum_{n_1,n_2,N} D^{j_1 *}_{n_1 m_1} D^{j_2 *}_{n_2 m_2} D^J_{NM} \big\langle j_1 j_2 n_1 n_2 \big| j_1 j_2 J N \big\rangle \\ &= \sum_{n_1,n_2,N} D^{j_1}_{n_1 m_1} D^{j_2}_{n_2 m_2} D^{J*}_{NM} \big\langle j_1 j_2 n_1 n_2 \big| j_1 j_2 J N \big\rangle \\ &= \sum_{n_1,n_2,N} D^{j_1}_{n_1 m_1} D^{j_2}_{n_2 m_2} D^{J}_{-N-M} (-1)^{N-M} \big\langle j_1 j_2 n_1 n_2 \big| j_1 j_2 J N \big\rangle \end{aligned} \tag{1.125}$$

$$\left\langle f^*_{l,m} f^*_{l',m'} \right\rangle = (-1)^m \left\langle f_{l,-m} f^*_{l',m'} \right\rangle = (-1)^m C_l \delta^{(\mathrm{K})}_{l,l'} \delta^{(\mathrm{K})}_{-m,m'}$$

The vector addition coefficients (Clebsch-Gordan coefficients) can be expressed in terms of the Wigner 3j symbols.

$$\sqrt{2J+1}\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & M \end{pmatrix} = (-1)^{j_1-j_2-M}\left\langle j_1 j_2 m_1 m_2 | j_1 j_2 J - M \right\rangle \tag{1.126}$$

Using this result in Eq. (1.125), we obtain

$$(-1)^{-M}\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & -M \end{pmatrix} = \sum_{n_1,n_2,N} D^{j_1}_{n_1 m_1} D^{j_2}_{n_2 m_2} D^{J}_{-N-M} (-1)^{N-M}(-1)^{-N}\begin{pmatrix} j_1 & j_2 & J \\ n_1 & n_2 & -N \end{pmatrix} \tag{1.127}$$

We thus have

$$\begin{pmatrix} j_1 & j_2 & J \\ m_1 & m_2 & M \end{pmatrix} = \sum_{n_1,n_2,N} D^{j_1}_{n_1 m_1} D^{j_2}_{n_2 m_2} D^{J}_{NM}\begin{pmatrix} j_1 & j_2 & J \\ n_1 & n_2 & N \end{pmatrix} \tag{1.128}$$

Let us express the expectation value of three angular multipole coefficients in the rotated frame in terms of the original multipole coefficients

$$\left\langle \tilde{f}_{lm}\tilde{f}_{l'm'}\tilde{f}_{l''m''}\right\rangle = \sum_{\tilde{m}\tilde{m}'\tilde{m}''} D^{l}_{m\tilde{m}} D^{l'}_{m'\tilde{m}'} D^{l''}_{m''\tilde{m}''}\left\langle f_{l\tilde{m}} f_{l'\tilde{m}'} f_{l''\tilde{m}''}\right\rangle \tag{1.129}$$

We can now use Eq. (1.119) in the rhs

$$\begin{aligned} \left\langle \tilde{f}_{lm}\tilde{f}_{l'm'}\tilde{f}_{l''m''}\right\rangle =& B_{ll'l''}\sum_{\tilde{m}\tilde{m}'\tilde{m}''} D^{l}_{m\tilde{m}} D^{l'}_{m'\tilde{m}'} D^{l''}_{m''\tilde{m}''}\begin{pmatrix} l & l' & l'' \\ \tilde{m} & \tilde{m}' & \tilde{m}'' \end{pmatrix} \\ =& B_{ll'l''}\begin{pmatrix} l & l' & l'' \\ m & m' & m'' \end{pmatrix} \end{aligned} \tag{1.130}$$

The bispectrum estimator is thus

$$\hat{B}_{ll'l''} = \sum_{mm'm''}\begin{pmatrix} l & l' & l'' \\ m & m' & m'' \end{pmatrix} f_{lm} f_{l'm'} f_{l''m''} = \sum_{mm'm''} h^{-1}_{ll'l''}\mathcal{G}^{l'l''}_{mm'm''} f_{lm} f_{l'm'} f_{l''m''} \tag{1.131}$$

where

$$h_{ll'l''} = \sqrt{\frac{(2l+1)(2l'+1)(2l''+1)}{4\pi}}\begin{pmatrix} l & l' & l'' \\ 0 & 0 & 0 \end{pmatrix} \tag{1.132}$$

Using the orthogonality of the Wigner $3j$ symbols, it is easy to show that this estimator is unbiased.

In the exercises you will show by explicit projection, that the bispectrum in l-space is related to the 3D bispectrum as

$$\left\langle f_{lm} f_{l'm'} f_{l''m''}\right\rangle = \left(\frac{2}{\pi}\right)^3 \int \mathrm{d}x x^2 \prod_{i=1}^{3}\left\{\int \mathrm{d}k_i k_i^2\; j_l(kr) j_l(kx)\right\} B(k_1,k_2,k_3)\mathcal{G}^{ll'l''}_{mm'm''} \tag{1.133}$$

which again confirms the structure to be $\left\langle f_{lm} f_{l'm'} f_{l''m''}\right\rangle = b_{ll'l''}\mathcal{G}^{ll'l''}_{mm'm''}$.