Les Houches Lectures: EFT of Large-Scale Structure
Lecture Notes

Tobias Baldauf

July 24, 2017

Abstract

These are **very preliminary** notes of my lectures given at the Les Houches 2017 School on EFT in particle physics and cosmology. They will be corrected and expanded over the next couple of weeks. Please refrain from distributing these notes and report typos to t.baldauf@tbaweb.de.

Contents

1 Introduction to Large-Scale Structure Cosmology

1.1 Review of homogeneous Cosmology

$$H(a) = \frac{\dot{a}}{a} = H_0\sqrt{\Omega_{\mathrm{m},0}a^{-3} + \Omega_{\mathrm{r},0}a^{-4} + \Omega_\Lambda} \tag{1}$$

The conformal time is defined as $\mathrm{d}t = a\,\mathrm{d}\eta$. Derivatives w.r.t. conformal time will be represented by dashes and derivatives w.r.t. coordinate time by dots.

The horizon describes how far information could propagate

$$\chi(a) = \int_0^a \mathrm{d}\tilde{a}\ \frac{c^2}{\tilde{a}^2 H(\tilde{a})} \tag{2}$$

1.2 Observations

- dark matter
- CMB
- LSS
- Bias
- Redshift Space Distortions

1.3 Statistics

The Universe we live in was seeded by quantum fluctuations, which became classical during inflation and grew by the subsequent evolution to form the highly nonlinear structures we observe today. As a consequence of this quantum mechanical origin, the structures are stochastic with random initial conditions. Due to this fact we can not hope to develop a theory that exactly reproduces the Universe we observe today. Rather, we should consider our Universe as one representation of an ensemble of possible Universes. Therefore, we need to introduce statistical quantities, which can be used to compare theoretical predictions with the observed data.

1.3.1 Overdensities

Starting from a smooth matter density field $\rho(\boldsymbol{r})$ we can define a dimensionless overdensity or density contrast

$$\delta(\boldsymbol{r}) = \frac{\rho(\boldsymbol{r}) - \bar{\rho}}{\bar{\rho}}, \tag{3}$$

which satisfies $\langle\delta(\boldsymbol{r})\rangle = 0$ and should be homogeneous and isotropic in a statistical sense. Here statistical homogeneity means that all multipoint moments remain invariant under coordinate translations, whereas statistical isotropy states that the latter will be true for coordinate rotations. The brackets stand for an averaging process, which can be understood either as an ensemble average over many possible realisations of the Universe or as a spatial average considering all $\boldsymbol{x}$ of the Universe. That these two averages are equivalent is not trivial. But we can assume that points that are far away from each other in the Universe are not causally connected and therefore we can use averages over widely separated regions as an approximation for independent realisations [**?**].[1]

1.3.2 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:

$$\delta(\boldsymbol{k}) = \int \mathrm{d}^3 r \exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right]\delta(\boldsymbol{r}), \tag{4}$$

$$\delta(\boldsymbol{r}) = \int \frac{\mathrm{d}^3 k}{(2\pi)^3} \exp\left[-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right]\delta(\boldsymbol{k}). \tag{5}$$

[1] Fields which satisfy the property that volume average is equivalent to ensemble average are termed ergodic in statistical physics.

such that the k-space representation of the nabla operator is given by $\nabla \to -\mathrm{i}\boldsymbol{k}$. The Dirac Delta function is thus given by

$$\delta^{(\mathrm{D})}(\boldsymbol{x}+\boldsymbol{x}') = \int \frac{\mathrm{d}^3q}{(2\pi)^3} \exp\left[\mathrm{i}(\boldsymbol{x}+\boldsymbol{x}')\boldsymbol{q}\right]. \tag{6}$$

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

$$f(\boldsymbol{x}) = \int \mathrm{d}^3y g(\boldsymbol{y})h(\boldsymbol{x}-\boldsymbol{y}) \Rightarrow f(\boldsymbol{k}) = g(\boldsymbol{k})h(\boldsymbol{k}). \tag{7}$$

This is of particular advantage, when smoothing operations are considered.
In case of spherical symmetry we can perform the angular integration in the definition of the Fourier transform

$$f(r) = \frac{1}{2\pi^2}\int \mathrm{d}k\ k^2\frac{\sin kr}{kr} f(k) = \frac{1}{2\pi^2}\int \mathrm{d}k\ k^2 j_0(kr)f(k) \tag{8}$$

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

$$f(k) = 4\pi \int \mathrm{d}r\ r^2 j_0(kr)f(r). \tag{9}$$

For the spherical Bessel functions we have a closure equation

$$\int_0^\infty \mathrm{d}x x^2 j_\alpha(ux)j_\alpha(vx) = \frac{\pi}{2u^2}\delta^{(\mathrm{D})}(u-v) \tag{10}$$

One of the most important clustering statistics in configuration space is the two-point correlation functioncorrelation function, defined as

$$\xi(\boldsymbol{r}) = \langle \delta(\boldsymbol{x})\delta(\boldsymbol{x}+\boldsymbol{r})\rangle. \tag{11}$$

Due to statistical isotropy the two point correlation only depends on the magnitude of the separation $\xi(\boldsymbol{r}) = \xi(|\boldsymbol{r}|)$. The most important Fourier space statistics used in LSS are the power spectrum and bispectrum defined by

$$\left\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}')\right\rangle = (2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}')P(\boldsymbol{k}) \tag{12}$$

$$\left\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}')\delta(\boldsymbol{k}'')\right\rangle = (2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}'+\boldsymbol{k}'')B(\boldsymbol{k},\boldsymbol{k}',\boldsymbol{k}'') \tag{13}$$

The power spectrum has units of volume and bispectrum has units of volume squared.
Let us see how the correlation function is related to the power spectrum.

$$\xi(r) = \int \frac{\mathrm{d}^3q}{(2\pi)^3}\int \frac{\mathrm{d}^3q'}{(2\pi)^3}\left\langle \delta(\boldsymbol{q})\delta(\boldsymbol{q}')\right\rangle \exp\left[-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}\right]\exp\left[-\mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{x}+\boldsymbol{r})\right] \tag{14}$$

$$= \int \frac{\mathrm{d}^3q}{(2\pi)^3}\int \frac{\mathrm{d}^3q'}{(2\pi)^3}(2\pi)^3P(q)\delta^{(\mathrm{D})}(\boldsymbol{q}+\boldsymbol{q}')\exp\left[-\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{x}\right]\exp\left[-\mathrm{i}\boldsymbol{q}'\cdot(\boldsymbol{x}+\boldsymbol{r})\right] \tag{15}$$

$$= \int \frac{\mathrm{d}^3q}{(2\pi)^3}P(q)\exp\left[-\mathrm{i}\boldsymbol{q}'\cdot\boldsymbol{r}\right] = \frac{1}{2\pi^2}\int \mathrm{d}q q^2 P(q)j_0(qr) \tag{16}$$

In the other direction we have

$$\left\langle \delta(\boldsymbol{k})\delta(\boldsymbol{k}')\right\rangle = \int \mathrm{d}^3x \int \mathrm{d}^3x' \exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{x}\right]\exp\left[\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{x}\right]\left\langle \delta(\boldsymbol{x})\delta(\boldsymbol{x}')\right\rangle \tag{17}$$

$$= \int \mathrm{d}^3x \exp\left[\mathrm{i}\boldsymbol{x}\cdot(\boldsymbol{k}+\boldsymbol{k}')\right]\int \mathrm{d}^3r \exp\left[\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}\right]\xi(r) \tag{18}$$

$$=(2\pi)^3\delta^{(\mathrm{D})}(\boldsymbol{k}+\boldsymbol{k}')\int \mathrm{d}^3r \exp\left[\mathrm{i}\boldsymbol{k}'\cdot\boldsymbol{r}\right]\xi(r) \tag{19}$$

since the last line has the same form as the definition of the power spectrum in Eq. (11), the power spectrum is in turn related to the correlation function by

$$P(k) = \int \mathrm{d}^3r\xi(r)\exp\left[\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}\right] = 4\pi\int \mathrm{d}r r^2\xi(r)j_0(kr). \tag{20}$$

1.3.3 Two-Point Probability Distribution

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 small volumes δV, which either host or don't host one tracer particle. The one point probability for finding a particle in the small volume δV_1 is $p_{\rm 1pt}(1) = \bar{n}\delta V_1$. 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

$$\mathbb{P}_{\rm 2pt}(1,2) = \mathbb{P}_{\rm 1pt}(1)\mathbb{P}_{\rm 1pt}(2) = \bar{n}^2 \delta V_1 \delta V_2. \tag{21}$$

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}

$$\mathbb{P}_{\rm 2pt}(1,2) = \bar{n}^2 \left[1 + \xi(r_{12})\right] \delta V_1 \delta V_2. \tag{22}$$

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

$$\mathbb{P}_{\rm 1pt}(2|1) = \frac{\mathbb{P}_{\rm 2pt}(1,2)}{\mathbb{P}_{\rm 1pt}(1)} = \bar{n}\left[1 + \xi(r_{12})\right] \delta V_2, \tag{23}$$

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 ($\xi(r_{12}) < 0$).
Similarly, we can define a quantity which describes how much two different tracers of the cosmological density field are correlated. The cross-correlation functioncross-correlation!function between two populations A and B is defined by

$$\mathbb{P}_{\rm 2pt}(r) = \bar{n}_{\rm A}\bar{n}_{\rm B}\left[1 + \xi_{\rm AB}(r)\right] \delta V_1 \delta V_2, \tag{24}$$

which can be calculated from the density fields as

$$\xi_{\rm AB}(\boldsymbol{r}) = \left\langle \delta_{\rm A}(\boldsymbol{x})\delta_{\rm B}(\boldsymbol{x}+\boldsymbol{r})\right\rangle. \tag{25}$$

1.3.4 Gaussian Random Fields

The seeds for structure formation are most probably of quantum mechanical origin. Hence we can treat the density field as a noise-like random field, where the phases of the Fourier modes are independent. From central limit theorem we know that the superposition of a large number of independent random fields will tend to a joint normal distribution. Besides δ itself, all quantities that can be expressed by linear sums over the modes will tend to be normally distributed. Since the first moment of δ vanishes, the Gaussian random field is entirely determined by its power spectrum, the variance for a certain Fourier mode.
By Wick theorem the reduced correlation functions of order higher than two either vanish or are expressible in terms of two-point functions [**?**]

$$\langle \delta(\boldsymbol{k}_1), \ldots, \delta(k_{2n+1})\rangle = 0, \tag{26}$$

$$\langle \delta(\boldsymbol{k}_1), \ldots, \delta(k_{2n})\rangle = \sum_{\rm pairs} \prod_{P\{(i,j)\}} \langle \delta(\boldsymbol{k}_i), \delta(\boldsymbol{k}_j)\rangle. \tag{27}$$

The Gaussiannity of the random field is also clear from the commutation relations for the quantum field.

1.4 Cosmic Variance

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 $k_{\rm f} = 2\pi/L$ and the corresponding volume of the fundamental cell is $V_{\rm f} = (2\pi)^3/V$. The Fourier modes are thus given by

$$\delta(\boldsymbol{k}_i) = \frac{V}{N_{\rm c}^3}\sum_j \delta(\boldsymbol{x}_j)\exp\left[{\rm i}\boldsymbol{k}_i \cdot \boldsymbol{x}_j\right], \tag{28}$$

where $\boldsymbol{k}_i = k_\text{f}\boldsymbol{i}$. The Dirac Delta function rewritten for discrete k as

$$\delta^{(\text{D})}(\boldsymbol{k}_i + \boldsymbol{k}_j) = \delta^{(\text{D})}\left((\boldsymbol{i}+\boldsymbol{j})k_\text{f}\right) = \frac{1}{k_\text{f}}\delta^{(\text{D})}(\boldsymbol{i}-\boldsymbol{j}) = \frac{V}{(2\pi)^3}\delta^{(\text{K})}_{\boldsymbol{i},\boldsymbol{j}} \tag{29}$$

The power spectrum for discrete cells is thus given by

$$P\left(|\boldsymbol{i}|k_\text{f}\right) = \frac{1}{V}\left\langle \delta(\boldsymbol{i}k_\text{f})\delta(-\boldsymbol{i}k_\text{f})\right\rangle \tag{30}$$

In simulations we estimate the power spectrum in logarithmic bins $[k\pm]$ centered at k

$$\hat{P}(k) = \frac{1}{N_k V}\sum_{\boldsymbol{i}k_\text{f}\in[k\pm]} \delta(\boldsymbol{k}_i)\delta(-\boldsymbol{k}_i), \tag{31}$$

where N_k is the number of cells in the k-bin. Note that the estimator is unbiased since $\left\langle \hat{P}\right\rangle = P$. The number of grid cells in the bin is given by the shell volume divided by the volume of the fundamental cell

$$N_k = \frac{V_k}{V_\text{f}} = \frac{4\pi k^3 \,\text{d}\ln k}{V_\text{f}} \tag{32}$$

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

$$\hat{B}(k_1,k_2,\mu) = \frac{1}{N_\text{tr}V}\sum_{\boldsymbol{i}k_\text{f}\in[k\pm]}\sum_{\boldsymbol{j}k_\text{f}\in[k\pm,\mu\pm]} \delta(\boldsymbol{k}_i)\delta(\boldsymbol{k}_j)\delta(-\boldsymbol{k}_i-\boldsymbol{k}_i) \tag{33}$$

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

$$N_\text{tr} = \frac{V_{123}}{V_\text{f}^2} = \frac{8\pi^2 k_1^3 k_2^3 (\,\text{d}\ln k)^2\,\text{d}\mu}{V_\text{f}^2} \tag{34}$$

Let us now calculate the variance of the power spectrum

$$\begin{aligned}\left\langle \hat{P}^2(k)\right\rangle - \left\langle \hat{P}(k)\right\rangle^2 =& \frac{1}{N_k^2 V^2}\sum_{\boldsymbol{i}k_\text{f},\boldsymbol{j}k_\text{f}\in[k\pm]}\left\langle \delta(\boldsymbol{k}_i)\delta(-\boldsymbol{k}_i)\delta(\boldsymbol{k}_j)\delta(-\boldsymbol{k}_j)\right\rangle - P^2(k) &(35)\\ =& \frac{1}{N_k^2}\sum_{\boldsymbol{i}k_\text{f},\boldsymbol{j}k_\text{f}\in[k\pm]} P(\boldsymbol{k}_i)P(\boldsymbol{k}_j) + \frac{2}{N_k^2}\sum_{\boldsymbol{i}k_\text{f}\in[k\pm]} P^2(\boldsymbol{k}_i) - P^2(k) &(36)\\ =& \frac{2}{N_k}P^2(k) = \frac{2V_f}{4\pi k^3\,\text{d}\ln k}P^2(k) &(37)\end{aligned}$$

Here we assumed Gaussianity and used the fact that due the reality condition in real space only half of the complex plane is independent. For the variance of the bispectrum estimator we have

$$\begin{aligned}\left\langle \hat{B}^2(k_1,k_2,\mu)\right\rangle - \left\langle \hat{B}(k_1,k_2,\mu)\right\rangle^2 =& \frac{1}{N_\text{tr}^2 V^2}\sum_{\boldsymbol{i},\boldsymbol{j},\boldsymbol{l},\boldsymbol{m}}\left\langle \delta(\boldsymbol{k}_i)\delta(\boldsymbol{k}_j)\delta(-\boldsymbol{k}_i-\boldsymbol{k}_j)\delta(\boldsymbol{k}_l)\delta(\boldsymbol{k}_m)\delta(-\boldsymbol{k}_l-\boldsymbol{k}_m)\right\rangle\\ &- B^2(k_1,k_2,\mu)\\ =& s_{123}\frac{V}{N_\text{tr}^2}\sum_{\boldsymbol{i},\boldsymbol{j}} P(\boldsymbol{k}_i)P(\boldsymbol{k}_j)P(-\boldsymbol{k}_i-\boldsymbol{k}_j)\\ =& s_{123}\frac{V}{N_\text{tr}}P(k_1)P(k_2)P(k_3)\\ =& s_{123}\frac{(2\pi)^3 V_\text{f}}{8\pi^2 k_1^3 k_2^3(\,\text{d}\ln k)^2\,\text{d}\mu}P(k_1)P(k_2)P(k_3) \qquad (38)\end{aligned}$$

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 $B = 0$.

1.5 Motivation

- 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

1.6 Brief History of Fluctuations

- fluctuations are generated as zero point fluctuations of the inflaton field
- they are stretched due to the exponential expansion and become classical
- fluctuations become superhorizon and are frozen
- fluctuations reenter the horizon - depending on whether this happens during matter or radiation domination strongly affects their evolution
- fluctuations are imprinted in the CMB
- fluctuations grow during matter domination, eventually become non-linear and lead to the cosmic web and galaxies

2 Dynamics in Newtonian Regime

2.1 Equations of Motion

Let us now consider the equations governing the cosmological fluid in the Newtonian limit, i.e., for small distances $x \ll H^{-1}$ and small velocities $v \ll 1$. The equation of motion for a particle at physical position $\boldsymbol{r}$ is

$$\ddot{\boldsymbol{r}} = -\boldsymbol{\nabla}_r \Phi \tag{39}$$

Defining comoving coordinates as $\boldsymbol{r} = a\boldsymbol{x}$ we have $\boldsymbol{\nabla}_x = a\boldsymbol{\nabla}_r$. From now on we will use the gradient with respect to comoving coordinates.[2] Let us take the derivative of the physical coordinate with respect to physical time and rewrite in terms of the comoving position

$$\dot{\boldsymbol{r}} = \mathcal{H}\boldsymbol{x} + \boldsymbol{x}'. \tag{40}$$

Likewise we have for the second derivative

$$\ddot{\boldsymbol{r}} = \frac{1}{a}\left(\mathcal{H}'\boldsymbol{x} + \mathcal{H}\boldsymbol{x}' + \boldsymbol{x}''\right) = -\frac{1}{a}\boldsymbol{\nabla}\Phi. \tag{41}$$

the term proportional to the position is peculiar, since it leads to a spatial dependence of the particle acceleration. This term arises from the comoving coordinates and accounts for the expansion of space-time. We can thus bring it to the right hand side of the above equation and define the peculiar potential

$$\Phi = -\frac{1}{2}\mathcal{H}'x^2 + \phi. \tag{42}$$

Thus we have for the equations of motion in terms of physical and conformal time

$$\ddot{\boldsymbol{x}} + 2H\dot{\boldsymbol{x}} = -\frac{\boldsymbol{\nabla}\phi}{a^2}, \qquad \boldsymbol{x}'' + \mathcal{H}\boldsymbol{x}' = -\boldsymbol{\nabla}\phi. \tag{43}$$

The peculiar potential is solely seeded by the energy density fluctuations in the Universe. Since we are assuming that dark energy is homogeneous, at late times these energy density fluctuations are dominated by the fluctuations in the matter density. Hence the Poisson equation for the peculiar potential can be expressed as

$$\nabla_x^2\phi = \frac{3}{2}\mathcal{H}^2\Omega_{\mathrm{m}}(a)\delta = \frac{3}{2}\Omega_{\mathrm{m},0}H_0^2\frac{\delta}{a}. \tag{44}$$

Defining the canonical momentummomentum!canonical

$$\boldsymbol{p} = am\boldsymbol{u}, \tag{45}$$

where $\boldsymbol{u} = \boldsymbol{x}'$ is the comoving velocity we have for the equation of motion

$$\boldsymbol{p}' = -am\boldsymbol{\nabla}_x\phi. \tag{46}$$

Let us finally stress again that these equations are only true in the Newtonian regime.

[2] For an arbitrary function $f(t)$ we have $a\dot{f} = f'$ $a^2\ddot{f} = f'' - \mathcal{H}f'$ Unless otherwise quoted we will refer to dots as derivatives with respect to coordinate time and dashes as derivatives with respect to conformal time.

2.2 The Fluid Equations

The particle distribution in phase space is conveniently described by the distribution function $f(\boldsymbol{x}, \boldsymbol{p}, \eta)$. The number of particles in a infinitesimal phase space volume $\mathrm{d}^3x\,\mathrm{d}^3p$ is thus given by $\mathrm{d}N = f(\boldsymbol{x}, \boldsymbol{p}, \eta)\,\mathrm{d}^3x\,\mathrm{d}^3p$. Note that the distribution function behaves as a scalar under parameter changes $\tilde{f}(\boldsymbol{x}, \boldsymbol{q}, \eta) = f(\boldsymbol{x}, \boldsymbol{p}(q), \eta)$. Louiville theorem asserts the conservation of the phase space density. This conservation of phase space density then yields the collisionless Boltzmann equation, also known as Vlaslov equation

$$\frac{\mathrm{d}f}{\mathrm{d}\eta} =\frac{\partial f}{\partial \eta} + \frac{\mathrm{d}\boldsymbol{x}}{\mathrm{d}\eta}\cdot\frac{\partial f}{\partial \boldsymbol{x}} + \frac{\mathrm{d}\boldsymbol{p}}{\mathrm{d}\eta}\frac{\partial f}{\partial \boldsymbol{p}} \tag{47}$$

$$=\frac{\partial f}{\partial \eta} + \frac{\boldsymbol{p}}{ma}\cdot\frac{\partial f}{\partial \boldsymbol{x}} - am\boldsymbol{\nabla}\phi\cdot\frac{\partial f}{\partial \boldsymbol{p}} = 0 \tag{48}$$

where we have used the equation of motion (45) in the last line.
We will be rarely interested in the full phase space distribution and thus there is no need to solve this non-linear seven dimensional differential equation. Instead we will be mostly concerned with fluid properties such as density, mean streaming velocity and velocity dispersion, which are readily obtained as moments of the distribution function

$$\rho(\boldsymbol{x}, \eta) =ma^{-3}\int \mathrm{d}^3p\; f(\boldsymbol{x}, \boldsymbol{p}, \eta), \tag{49}$$

$$v_i(\boldsymbol{x}, \eta) =\int \mathrm{d}^3p\;\frac{p_i}{am} f(\boldsymbol{x}, \boldsymbol{p}, \eta)/\int \mathrm{d}^3p\; f(\boldsymbol{x}, \boldsymbol{p}, \eta), \tag{50}$$

$$\sigma_{ij}(\boldsymbol{x}, \eta) =\int \mathrm{d}^3p\;\frac{p_i}{am}\frac{p_j}{am} f(\boldsymbol{x}, \boldsymbol{p}, \eta)/\int \mathrm{d}^3p\; f(\boldsymbol{x}, \boldsymbol{p}, \eta) - v_i(\boldsymbol{x})v_j(\boldsymbol{x}). \tag{51}$$

The velocity dispersion is sometimes also referred to as anisotropic stress and describes the deviation from a single coherent flow as is obvious in Eq (50).
The equations of motion for these quantities can now be obtained by taking moments of the Vlaslov equation (47). The zeroth moment of the Vlaslov equation yields the continuity equation. Upon integrating over the momentum, we have to integrate the last term by parts and use that the .

$$\delta' + \boldsymbol{\nabla}\cdot\left[\boldsymbol{v}(1+\delta)\right] = 0 \tag{52}$$

Taking the the first moment and using the continuity equation yields the Euler equation

$$v_i' + \mathcal{H}v_i + \boldsymbol{v}\cdot\boldsymbol{\nabla}v_i = -\nabla_i\phi - \frac{1}{\rho}\nabla_i(\rho\sigma_{ij}) \tag{53}$$

or conservation of momentum. In principle we could have continued to hierarchy of equations, which couples the equation of motion for n-th moment of the Vlaslov equation to the $n+1$-th moment. To close the hierarchy we will postulate that all moments beyond the velocity are vanishing, an assumption that is denoted the pressureless perfect fluid. This assumption is reasonable in the linear regime but needs to be validated numerically at late times, when structures collapse, virialize and shell crossing occurs.
The fluid velocity can be decomposed into a scalar and a vector part $\boldsymbol{v} = \boldsymbol{v}_{\|} + \boldsymbol{v}_{\perp}$, where $\boldsymbol{\nabla}\times\boldsymbol{v}_{\|} = 0$ and $\boldsymbol{\nabla}\cdot\boldsymbol{v}_{\perp} = 0$. The velocity field can thus be described by its vorticity $\boldsymbol{w} = \boldsymbol{\nabla}\times\boldsymbol{v}$ and its divergence $\theta = \nabla\cdot\boldsymbol{v}$.

2.3 Linearized Equations

Let us neglect all the quadratic terms in the continuity and Euler equation and assume that the velocity dispersion vanishes

$$\delta' + \theta =0 \tag{54}$$

$$\boldsymbol{v}' + \mathcal{H}\boldsymbol{v} =-\boldsymbol{\nabla}\phi. \tag{55}$$

The system can be solved straightforwardly after rewriting the Euler equation in terms of velocity vorticity and divergence

$$\theta' + \mathcal{H}\theta =-\Delta\phi \tag{56}$$

$$\boldsymbol{w}' + \mathcal{H}\boldsymbol{w} =0. \tag{57}$$

The solution of the vorticity equation is simply $\boldsymbol{w}\propto a^{-1}$, i.e., any initially present vorticity decays at linear level. To solve the scalar equation, we take the time derivative of Eq. (53) and replace θ' with

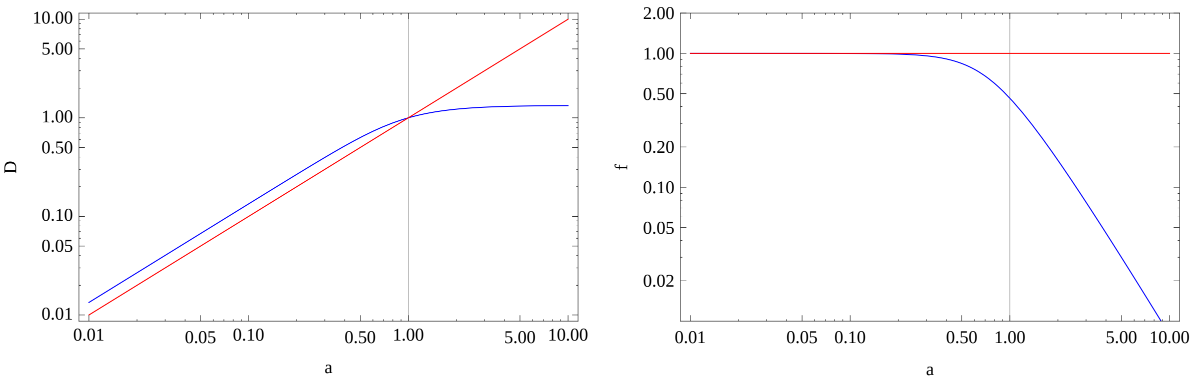

Figure 1: Linear growth of structure for our fiducial ΛCDM(blue) and a matter only EdS Universe (red). *Left panel:* Linear growth factor D. *Right panel:* Logarithmic growth factor f.

Eq. (55). In the resulting equation, we can replace θ using Eq. (55) and $\Delta\phi$ using the Poisson Eq. (43). We obtain

$$\delta''(\boldsymbol{k},\eta)+\mathcal{H}(\eta)\delta'(\boldsymbol{k},\eta)-\frac{3}{2}\Omega_{\mathrm{m}}(\eta)\mathcal{H}^2(\eta)\delta(\boldsymbol{k},\eta)=0. \tag{58}$$

As can be easily confirmed the above differential equation has a growing and a decaying mode solution $\delta(\boldsymbol{k},\eta)=D_+(\eta)\delta_{+,0}(\boldsymbol{k})+D_-(\eta)\delta_{-,0}(\boldsymbol{k})$, where $D_-(\eta)=D_{-,0}H=\mathcal{H}/a$. The growing mode solution can then be obtained as

$$D_+(\eta)=D_{+,0}H(\eta)\int_0^{a(\eta)}\frac{\mathrm{d}a'}{\mathcal{H}^3(a')}, \tag{59}$$

where $D_{+,0}$ is a normalization factor used to achieve $D_+(a=1)=1$. Let us first discuss the solution in a Einstein-de-Sitter (EdS) matter only Universe. Since $a\propto t^{2/3}$ we have $H=a^{-3/2}$ and thus $D_+=a$ and $D_-=a^{-3/2}$. This special case and the solution for our fiducial ΛCDMmodel are shown in the left panel of Fig. 1. We see that the growth in the ΛCDMUniverse stalls at late times when the cosmological constant starts to dominate. In what follows we will concentrate on the growing mode solutions and use $D\equiv D_+$ unless otherwise stated.

2.4 Velocities

In the above subsection we have seen, that the vorticity decays at linear level in the absence of anisotropic stress. At non-linear level we have

$$\boldsymbol{w}+\mathcal{H}\boldsymbol{w}+\boldsymbol{\nabla}\times(\boldsymbol{v}\times\boldsymbol{w})=0. \tag{60}$$

This equation tells us, that even at non-linear level, if there is no initial vorticity, evolution won't generate it. Together with the knowledge that vorticity decays at early times when the fluctuations are still linear we can conclude that vorticity will be negligible throughout evolution in absence of anisotropic stress. However, there is evidence from simulations [**?**] that at late times velocity dispersion and thus vorticity is generated in high density regions. In the following, we will present Standard Perturbation Theory, ignoring the possible vorticity. This assumption will restrict the validity of the solutions to large scales. In case of vanishing curl we have

$$\boldsymbol{v}(\boldsymbol{k})=\mathrm{i}\frac{\boldsymbol{k}}{k^2}\theta(\boldsymbol{k}). \tag{61}$$

From the linearized continuity equation in Fourier space we have

$$\theta(\boldsymbol{k},\eta)=-\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{v}(\boldsymbol{k},\eta)=-\delta'(\boldsymbol{k},\eta)=-\mathcal{H}f\delta(\boldsymbol{k},\eta), \tag{62}$$

where we defined the logarithmic growth factor

$$f=\frac{\mathrm{d}\ln D}{\mathrm{d}\ln a} \tag{63}$$

For the linear growing mode defined above in Eq. (58), we have

$$f(a) = \frac{\mathrm{d}\ln H}{\mathrm{d}\ln a} + \frac{a}{(aH)^3} \frac{1}{\int_0^a \mathrm{d}a' \, [a' H(a')]^3}, \tag{64}$$

which is unity for EdS. The right panel of Fig. 1 shows the logarithmic growth factor for a EdS and our fiducial ΛCDMcosmology. The late time dominance of the cosmological constant is even more apparent in this plot where f decays to $f(a=1) \approx 0.48$ at present time.
We can now employ this result to derive simple velocity statistics, such as the linear velocity dispersion

$$\kappa_{ij} = \langle v_i v_j \rangle - \langle v_i \rangle \langle v_j \rangle = \langle v_i v_j \rangle = \mathcal{H}^2 f^2 \int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{q_i q_j}{q^4} P_{\rm lin}(q), \tag{65}$$

and its trace

$$\tilde{\sigma}_v^2 = \frac{1}{3} \mathrm{Tr}\left(\kappa_{ij}^2\right) = \frac{\mathcal{H}^2 f^2}{6\pi^2} \int \mathrm{d}q \; P_{\rm lin}(q). \tag{66}$$

In the following we will frequently consider the normalized velocity dispersion $\sigma_v = \tilde{\sigma}_v / \mathcal{H} f$, for which we obtain in our fiducial cosmology $\sigma_v \approx 6 \; h^{-1}\mathrm{Mpc}$.

2.5 Fluid Equations in Fourier Space

After having obtained some intuition on the solutions in the linear regime, where the quadratic terms are negligible, we will now return to the full equations. To facilitate the analysis, we will work in Fourier space, where the Euler and continuity equations read as

$$\delta'(\boldsymbol{k}) + \theta(\boldsymbol{k}) = -\int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{\mathrm{d}^3 q'}{(2\pi)^3} \delta^{(\mathrm{D})}(\boldsymbol{k} - \boldsymbol{q} - \boldsymbol{q}') \alpha(\boldsymbol{q}, \boldsymbol{q}') \theta(\boldsymbol{q}) \delta(\boldsymbol{q}'), \tag{67}$$

$$\theta'(\boldsymbol{k}) + \mathcal{H}\theta(\boldsymbol{k}) + \frac{3}{2}\Omega_{\rm m}(a)\mathcal{H}^2 \delta(\boldsymbol{k}) = -\int \frac{\mathrm{d}^3 q}{(2\pi)^3} \frac{\mathrm{d}^3 q'}{(2\pi)^3} \delta^{(\mathrm{D})}(\boldsymbol{k} - \boldsymbol{q} - \boldsymbol{q}') \beta(\boldsymbol{q}, \boldsymbol{q}') \theta(\boldsymbol{q}) \theta(\boldsymbol{q}'). \tag{68}$$

The coupling kernels on the right hand side are defined as

$$\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{\boldsymbol{k}_1 \cdot (\boldsymbol{k}_1 + \boldsymbol{k}_2)}{k_1^2} \tag{69}$$

$$\beta(\boldsymbol{k}_1, \boldsymbol{k}_2) = \frac{1}{2} (\boldsymbol{k}_1 + \boldsymbol{k}_2)^2 \frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_1^2 k_2^2} = \frac{1}{2} \frac{\boldsymbol{k}_1 \cdot \boldsymbol{k}_2}{k_1 k_2} \left(\frac{k_2}{k_1} + \frac{k_1}{k_2} \right) + \frac{(\boldsymbol{k}_1 \cdot \boldsymbol{k}_2)^2}{k_1^2 k_2^2} \tag{70}$$

Note that $\alpha(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is not symmetric in its arguments but $\beta(\boldsymbol{k}_1, \boldsymbol{k}_2)$ is. The fluid Eqs. (66) and (67) are non-linear coupled differential equations for the density and velocity divergence. A closed form solution does in general not exist. One can however try to solve them perturbatively in the regime, where $\delta \ll 1$ and $\theta \ll 1$. We will discuss the perturbative solutions in much more detail in §**??** below.

2.6 Perturbative Treatment of the Fluid Equations

Standard Perturbation Theory aims to solve the fluid equations perturbatively using a power law ansatz. This approach simplifies significantly in an EdS Universe, where $D = a$. Hence we will study this case first and discuss the generalization to ΛCDMlater. We will furthermore neglect the decaying mode. The power law ansatz reads

$$\delta(\boldsymbol{k}, \eta) = \sum_{i=1}^{\infty} a^i(\eta) \delta^{(i)}(\boldsymbol{k}) \qquad \theta(\boldsymbol{k}, \eta) = -\mathcal{H}(\eta) \sum_{i=1}^{\infty} a^i(\eta)\, \theta^{(i)}(\boldsymbol{k}). \tag{71}$$

The expansion is in powers of the linear density field discussed above, i.e., ${}^{(i)}\delta = \mathcal{O}({}^{(1)}\delta^i)$. We can now write the n-th order solutions as convolutions of linear density fields

$$\delta^{(n)}(\boldsymbol{k}) = \prod_{m=1}^{n} \left\{ \int \frac{\mathrm{d}^3 q_m}{(2\pi)^3} \delta^{(1)}(\boldsymbol{q}_m) \right\} F_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \delta^{(\mathrm{D})}(\boldsymbol{k} - \boldsymbol{q}|_1^n) \tag{72}$$

$$\theta^{(n)}(\boldsymbol{k}) = \prod_{m=1}^{n} \left\{ \int \frac{\mathrm{d}^3 q_m}{(2\pi)^3} \delta^{(1)}(\boldsymbol{q}_m) \right\} G_n(\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n) \delta^{(\mathrm{D})}(\boldsymbol{k} - \boldsymbol{q}|_1^n) \tag{73}$$

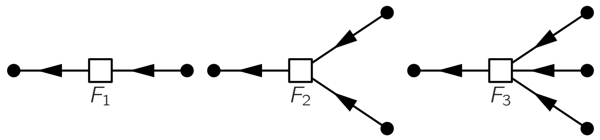

Figure 2: Diagrammatic representation of the series expansion of the density field in Eq. (71). The points on the right side are initial density fields and the points on the left side are n-th order density fields.

A diagrammatic representation of this expansion is shown in Fig. 2. One can derive the following recursion relations for the convolution kernels

$$F_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)}{(2n+3)(n-1)} \Big[(2n+1)\alpha\left(\boldsymbol{q}|_1^m, \boldsymbol{q}|_{m+1}^n\right) F_{n-m}\left(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n\right) + 2\beta\left(\boldsymbol{q}|_1^m, \boldsymbol{q}|_{m+1}^n\right) G_{n-m}\left(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n\right)\Big] \tag{74}$$

$$G_n(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_n) = \sum_{m=1}^{n-1} \frac{G_m(\boldsymbol{q}_1,\ldots,\boldsymbol{q}_m)}{(2n+3)(n-1)} \Big[3\alpha\left(\boldsymbol{q}|_1^m, \boldsymbol{q}|_{m+1}^n\right) F_{n-m}\left(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n\right) + 2n\beta\left(\boldsymbol{q}|_1^m, \boldsymbol{q}|_{m+1}^n\right) G_{n-m}\left(\boldsymbol{q}_{m+1},\ldots,\boldsymbol{q}_n\right)\Big], \tag{75}$$

where $\boldsymbol{q}|_i^j = \sum_{m=i}^j \boldsymbol{q}_l$. For general ΛCDM the series ansatz in Eq. (70) can be generalized to

$$\delta(\boldsymbol{k},\eta) = \sum_{i=1}^{\infty} D^i(\eta)\delta^{(i)}(\boldsymbol{k}) \qquad \theta(\boldsymbol{k},\eta) = -\mathcal{H}(\eta)f(\eta)\sum_{i=1}^{\infty} D^i(\eta)\,\theta^{(i)}(\boldsymbol{k}). \tag{76}$$

The exact solution deviates from the above solution and doesn't allow for a separation of time and space dependence as in Eq. (75). The differences are discussed in detail in [?] and are generally at the sub-percent level. We will thus stick to the approximation Eq. (75), which is sufficiently accurate for our purpose.
Let us come back to the kernels and evaluate them explicitly at second and third order. The second order density kernel is given by

$$F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{5}{7}\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) + \frac{2}{7}\beta(\boldsymbol{k}_1,\boldsymbol{k}_2) \tag{77}$$

$$= \frac{5}{7} + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2}\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right) + \frac{2}{7}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2} \tag{78}$$

$$= \frac{5}{7} + \frac{1}{2}\mu_{12}\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right) + \frac{2}{7}\mu_{12}^2 \tag{79}$$

Here we defined $\boldsymbol{k}_1\cdot\boldsymbol{k}_2 = k_1k_2\mu_{12}$ and symmetrized Eq. (71) over the momenta. The latter can be rewritten as

$$F_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{17}{21} + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2}\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right) + \frac{2}{7}\left[\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2} - \frac{1}{3}\right] \tag{80}$$

where the quadratic density term, the shift term and the anisotropic stress term are more obvious. At the same time, the angular structure of monopole, dipole and quadrupole is more obvious. The second order velocity kernel reads

$$G_2(\boldsymbol{k}_1,\boldsymbol{k}_2) = \frac{3}{7}\alpha(\boldsymbol{k}_1,\boldsymbol{k}_2) + \frac{4}{7}\beta(\boldsymbol{k}_1,\boldsymbol{k}_2) \tag{81}$$

$$= \frac{3}{7} + \frac{1}{2}\frac{\boldsymbol{k}_1\cdot\boldsymbol{k}_2}{k_1k_2}\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right) + \frac{4}{7}\frac{(\boldsymbol{k}_1\cdot\boldsymbol{k}_2)^2}{k_1^2k_2^2} \tag{82}$$

$$= \frac{3}{7} + \frac{1}{2}\mu_{12}\left(\frac{k_2}{k_1}+\frac{k_1}{k_2}\right) + \frac{4}{7}\mu_{12}^2. \tag{83}$$

Note that

$$F_2(\boldsymbol{q}_1,\boldsymbol{q}_2) - G_2(\boldsymbol{q}_1,\boldsymbol{q}_2) = \frac{2}{7}\left[\alpha(\boldsymbol{q}_1,\boldsymbol{q}_2) - \beta(\boldsymbol{q}_1,\boldsymbol{q}_2)\right] \tag{84}$$

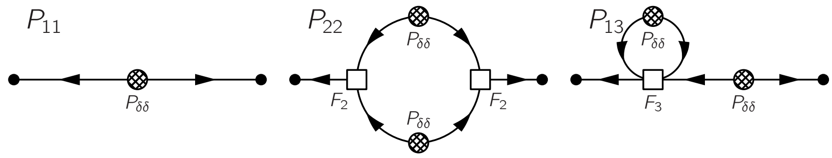

Figure 3: Diagrammatic representation of the one loop matter power spectrum in Eq. (87) where $P_{11} = P_{\rm lin}$.

The explicit expressions for F_3 and G_3 are

$$\begin{aligned} F_3(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) =& \frac{1}{18}\left(7\alpha(\boldsymbol{q}_1,\boldsymbol{q}_2+\boldsymbol{q}_3)F_2(\boldsymbol{q}_2,\boldsymbol{q}_3)+2\beta(\boldsymbol{q}_1,\boldsymbol{q}_2+\boldsymbol{q}_3)G_2(\boldsymbol{q}_2,\boldsymbol{q}_3)\right) \\ &+\frac{G_2(\boldsymbol{q}_1,\boldsymbol{q}_2)}{18}\left(7\alpha(\boldsymbol{q}_1+\boldsymbol{q}_2,\boldsymbol{q}_3)+2\beta(\boldsymbol{q}_1+\boldsymbol{q}_2,\boldsymbol{q}_3)\right) \qquad (85)\\ G_3(\boldsymbol{q}_1,\boldsymbol{q}_2,\boldsymbol{q}_3) =& \frac{1}{18}\left(3\alpha(\boldsymbol{q}_1,\boldsymbol{q}_2+\boldsymbol{q}_3)F_2(\boldsymbol{q}_2,\boldsymbol{q}_3)+6\beta(\boldsymbol{q}_1,\boldsymbol{q}_2+\boldsymbol{q}_3)G_2(\boldsymbol{q}_2,\boldsymbol{q}_3)\right) \\ &+\frac{G_2(\boldsymbol{q}_1,\boldsymbol{q}_2)}{18}\left(3\alpha(\boldsymbol{q}_1+\boldsymbol{q}_2,\boldsymbol{q}_3)+6\beta(\boldsymbol{q}_1+\boldsymbol{q}_2,\boldsymbol{q}_3)\right) \qquad (86)\end{aligned}$$

The above formulae are not symmetrized over the arguments yet. Upon integration over three equivalent density fields $\delta(\boldsymbol{q}_1)\delta(\boldsymbol{q}_2)\delta(\boldsymbol{q}_3)$ we have to symmetrize, accounting both for the cyclic and odd permutations of the arguments in the kernels.

2.7 Power Spectrum & Bispectrum

The discussion in the previous section allowed us to express non-linear density and velocity fields as a sum of products of linear density fields. If we are interested in n-spectra of the fields, we have to correlate two of these non-linear fields with each other. This will again lead to a sum of correlators. As we have seen in Sec. 1.3.4, only even correlators of linear density fields contribute and these correlators can be expressed as products of linear power spectra, whose form is given by the process seeding the fluctuations and the subsequent linear growth. Thus we were able to reduce the problem of calculating spectra of non-linear fields to convolutions of linear power spectra.
As this calculation becomes more and more tedious order by order, we usually truncate the calculation at next-to-leading order or next-to-next-to-leading order. In this context it is also useful to introduce the notion of loops. The leading order contribution to the power spectrum is second order in the fields and doesn't involve any momentum integrations and is thus formally a zero-loop result. The next to leading order has to be of fourth order in the fields by Wick theorem. There are two possibilities to achieve this, by correlating two second order density fields or by correlating a linear density field with a second order density field

$$\begin{aligned} \langle\delta(\boldsymbol{k})\delta(-\boldsymbol{k})\rangle =& \left\langle{}^{(1)}\delta(\boldsymbol{k}){}^{(1)}\delta(-\boldsymbol{k})\right\rangle+2\left\langle{}^{(1)}\delta(\boldsymbol{k}){}^{(3)}\delta(-\boldsymbol{k})\right\rangle+\left\langle{}^{(2)}\delta(\boldsymbol{k}){}^{(2)}\delta(-\boldsymbol{k})\right\rangle \qquad (87)\\ P_{\rm 1loop}(k) =& P_{\rm lin}(k)+P_{13}(k)+P_{22}(k). \qquad (88)\end{aligned}$$

A diagrammatic representation of the above power spectrum is given in Fig. 3. As we will see in detail below, the next to leading order calculations involve one momentum integral and are thus one loop terms. Up to sixth order in the fields, i.e. two loops, we have

$$\begin{aligned} \langle\delta(\boldsymbol{k})\delta(-\boldsymbol{k})\rangle =& \left\langle{}^{(1)}\delta(\boldsymbol{k}){}^{(1)}\delta(-\boldsymbol{k})\right\rangle+2\left\langle{}^{(1)}\delta(\boldsymbol{k}){}^{(3)}\delta(-\boldsymbol{k})\right\rangle+2\left\langle{}^{(1)}\delta(\boldsymbol{k}){}^{(5)}\delta(-\boldsymbol{k})\right\rangle \\ &+\left\langle{}^{(2)}\delta(\boldsymbol{k}){}^{(2)}\delta(-\boldsymbol{k})\right\rangle+\left\langle{}^{(3)}\delta(\boldsymbol{k}){}^{(3)}\delta(-\boldsymbol{k})\right\rangle+2\left\langle{}^{(2)}\delta(\boldsymbol{k}){}^{(4)}\delta(-\boldsymbol{k})\right\rangle \qquad (89)\end{aligned}$$

here we separated the terms correlating non-linear density field and linear density field (propagator terms) in the first line and the mode coupling terms in the second line.
In Sec. 1.3.4 we mentioned, that Gaussian random fields have a vanishing bispectrum and that their trispectrum can be written as a product of power spectra. The gravitational evolution changes this behaviour and leads to a non-vanishing bispectrum given by

$$B(\boldsymbol{k}_1,\boldsymbol{k}_2,\boldsymbol{k}_3) = 2F_2(\boldsymbol{k}_1,\boldsymbol{k}_2)P_{\rm lin}(k_1)P_{\rm lin}(k_2)+2\text{ cyc.} \qquad (90)$$

Here cyc. stands for a cyclic permutation of the three k vectors in the arguments of the power spectra and coupling kernels.

3 Effective Field Theory Approach