EFT of LSS talk at Galaxies meet QCD

Matthew Lewandowski[1]

[1] *Institut fur Theoretische Physik, ETH Zurich, 8093 Zurich, Switzerland*

Contents

1 Notation and conventions

These are informal notes from my blackboard presentation at the Galaxies meet QCD workshop in 2024. They are intended to remind the participants about the subjects discussed, not to be a complete review of the subject, nor to contain a complete list of references. A much more complete set of references to relevant and important work on this subject can be found in the citations contained here.

Our Fourier conventions are

$$f(\vec{x},t) = \int_{\vec{k}} f(\vec{k},t)\, e^{i\vec{k}\cdot\vec{x}} \,, \tag{1.1}$$

and in this work, we use the following notation

$$\int_{\vec{k}_1,\ldots,\vec{k}_n} \equiv \int \frac{d^3k_1}{(2\pi)^3}\cdots\frac{d^3k_n}{(2\pi)^3} \ , \quad \int^{\vec{k}}_{\vec{k}_1,\ldots,\vec{k}_n} \equiv \int_{\vec{k}_1,\ldots,\vec{k}_n} (2\pi)^3\delta_D(\vec{k}-\sum_{i=1}^{n}\vec{k}_i) \ , \tag{1.2}$$

where δ_D is the Dirac delta function.

For a three-dimensional vector $\vec{k}$, we write $k \equiv |\vec{k}|$ for the magnitude, and $\hat{k} \equiv \vec{k}/k$ for the unit vector parallel to $\vec{k}$. We use Latin letters like i, j, k, l to denote spatial indices, in general we do not distinguish between upper and lower spatial indices, and repeated indices imply summation.

- background ΛCDM expansion, with metric $ds^2 = -dt^2 + a(t)^2 d\vec{x}^2$, where $a(t)$ is the scale factor, $a \to 0$ is early times, and $a = a_0 = 1$ is the current time
- the scale factor a is related to the observed redshift z by $a = 1/(1+z)$
- the background expansion is driven by a non-relativistic, time-dependent, background mass density $\bar{\rho}(t)$ which is given by

  $$\bar{\rho}(t) = \bar{\rho}_0 \left(\frac{a(t)}{a_0}\right)^{-3} \ , \tag{1.3}$$

  where subscripts $_0$ refer to current-day values
- $H(t) \equiv \dot{a}(t)/a(t)$ is the Hubble expansion rate
- $M_{\rm Pl}$ is the Planck mass, related to Newton's constant G_N by $M_{\rm Pl}^2 = 1/(8\pi G_N)$.
- Einstein equations give $H(t)^2 = \frac{1}{3M_{\rm Pl}^2}\left(\bar{\rho}(t) + \rho_\Lambda\right)$
- we often use a as the time variable: $da = a(t)H(t)dt$ - prime $'$ is for derivative with respect to a, dot $\dot{}$ is derivative with respect to t
- time-dependent matter fraction: $\Omega_{\rm m}(a) \equiv \bar{\rho}(a)/(3M_{\rm Pl}^2 H(a)^2)$
- in flat ΛCDM, the Hubble rate can be parameterized by $H(a)^2/H_0^2 = \Omega_{\rm m,0}(a/a_0)^{-3} + (1 - \Omega_{\rm m,0})$
- $\mathcal{H}(a) \equiv aH(a)$
- the dark-matter field is described in terms of the mass density $\rho(\vec{x}, a)$ and the velocity field $v^i(\vec{x}, a)$. will often use $\delta(\vec{x}, a) \equiv (\rho(\vec{x}, a) - \bar{\rho}(a))/\bar{\rho}(a)$ and $\pi^i(\vec{x}, a) \equiv \rho(\vec{x}, a)v^i(\vec{x}, a)$
- $\theta \equiv -\partial_i v^i/(faH)$ is normalized such that $\theta \sim D(a)^n G_n$ in Fourier space, where G_n are kernels defined below
- scalar perturbations: $ds^2 = -(1+2\Phi)dt^2 + a(t)^2(1-2\Psi)d\vec{x}^2$, and $\Psi = \Phi$ for us (no anisotropic stress)

2 Dark-matter clustering

2.1 Equations of motion

We are in the Newtonian limit, and the dark matter is non-relativistic (i.e. cold dark matter, or CDM). This regime is described by

$$v \ll 1 \tag{2.1}$$

$$\frac{k}{aH(a)} \gg 1 \tag{2.2}$$

$$k_{\rm NL}^{-1} \sim \frac{v}{aH(a)} \tag{2.3}$$

$$k/k_{\rm NL} \ll 1 \tag{2.4}$$

One can think of the equations of motion (EOM) as coming from the Einstein equations [1–3], or one can construct them from the bottom up from symmetries [4]. Then, the EOM are

$$\begin{aligned} &\dot\delta + a^{-1}\partial_i((1+\delta)v^i) = 0 \ , \\ &\dot v^i + Hv^i + a^{-1}\partial_i\Phi + a^{-1}v^j\partial_j v^i = -a^{-1}\frac{\partial_j\tau^{ij}}{\bar\rho(1+\delta)} \ , \end{aligned} \tag{2.5}$$

along with the Poisson equation

$$a^{-2}\partial^2\Phi = \frac{3}{2}\Omega_{\rm m}H^2\delta \ . \tag{2.6}$$

Above, τ^{ij} is the EFT of LSS stress tensor (i.e. counterterms), which we will describe in detail soon. Importantly, the EOM are nonlinear.

This system has some important symmetries. The continuity equation implies mass conservation in the from

$$\partial_t \int d^3x\, a(t)^3\rho(\vec x,t) = 0 \ , \tag{2.7}$$

and the Euler equation implies momentum conservation in the form

$$\partial_t \int d^3x\, a(t)^4\pi^i(\vec x,t) = 0 \ , \tag{2.8}$$

(which is much more obvious if the EOM are written in terms of π^i instead of v^i).[1] Another symmetry is Galilean invariance, which is inherited as the non-relativistic limit of diffeomorphism invariance. In the the non-relativistic limit that we present here, a Galilean transformation takes the form

$$\begin{aligned} &\partial_i \to \partial_i \ , \quad \partial_t \to \partial_t - \dot n^i(t)\partial_i \ , \quad v^i(x^j,t) \to v^i(x^j,t) + a\dot n^i(t) \ , \quad \tau^{ij}(x^k,t) \to \tau^{ij}(x^k,t) \\ &\Phi(x^j,t) \to \Phi(x^j,t) - a^2(\ddot n^i(t) + 2H\dot n^i(t))x^i \ , \quad \delta(x^j,t) \to \delta(x^j,t) \ , \end{aligned} \tag{2.9}$$

for any time dependent vector $n^i(t)$. We will see later what this symmetry means for the solutions. We also have the equivalence principle, which says that physical measurements can only depend on

[1] Note that these are dynamical statements that are true for any initial conditions. They are simply the field version of total mass and total momentum conservation in multi-particle Newtonian mechanics.

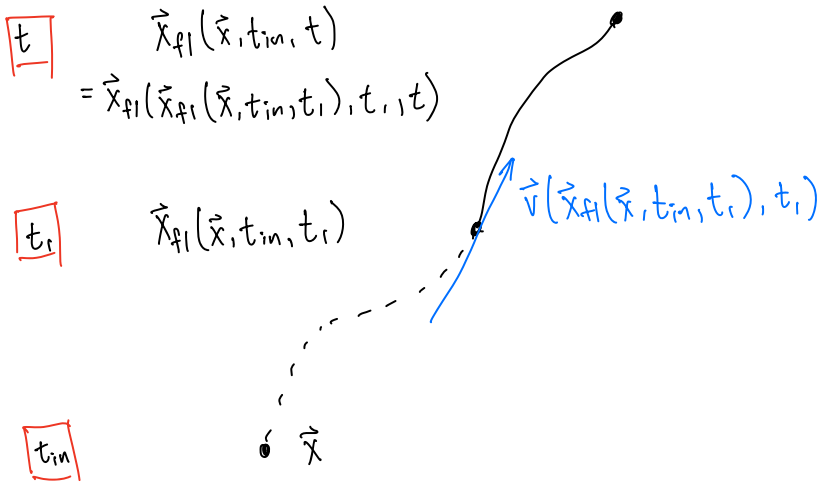

Figure 1: Diagram for the fluid line element showing how the composition rule works.

second spatial derivatives of the metric $\partial_i\partial_j\Phi$ and gradients of the velocity $\partial_i v^j$. Finally, we have statistical homogeneity and isotropy, which is translation and rotation invariance of the correlation functions.

A useful quantity to talk about Galilean invariance is the fluid element $\vec{x}_{\rm fl}$, which traces the fluid velocity, see Fig. 1. Since the velocity is a tangent, we have

$$\frac{d}{dt}\vec{x}_{\rm fl}(\vec{x}, t_{\rm in}, t) = \frac{1}{a(t)}\vec{v}(\vec{x}_{\rm fl}(\vec{x}, t_{\rm in}, t), t) \ , \tag{2.10}$$

which can be integrated to give the recursive solution

$$\vec{x}_{\rm fl}(\vec{x}, t_{\rm in}, t) = \vec{x} + \int_{t_{\rm in}}^{t} \frac{dt'}{a(t')}\vec{v}\left(\vec{x}_{\rm fl}(\vec{x}, t_{\rm in}, t'), t'\right) \ . \tag{2.11}$$

2.2 EFT counterterms

Let us write the EOM for δ in a different way

$$a^2\delta'' + \left(2 + \frac{a\mathcal{H}'}{\mathcal{H}}\right)a\delta' - \frac{3}{2}\Omega_{\rm m}\delta = \frac{\partial_i\partial_j}{\mathcal{H}^2\bar{\rho}}\left(2M_{\rm Pl}^2 a^{-2}\left(\partial_i\Phi\partial_j\Phi - \frac{1}{2}\delta_{ij}(\partial\Phi)^2\right) + \frac{\pi^i\pi^j}{\rho} + \tau^{ij}\right) \ . \tag{2.12}$$

From here, we obviously see that for δ, the stress tensor enters with two spatial derivatives.

The EFT of LSS is local in space, but non-local in time [3]. The most general expansion of the stress tensor satisfying all of the symmetries above is

$$\tau^{ij}(\vec{x}, t) = \int^t dt' H(t') \sum_\alpha \kappa_\alpha(t, t') T_\alpha^{ij}(\vec{x}_{\rm fl}(\vec{x}, t, t'), t') \ , \tag{2.13}$$

where $\{T_\alpha^{ij}(\vec{x}, t)\}_\alpha$ are all local-in-time Galilean scalars (and tensors under rotations on the i and j indices), and the $\kappa_\alpha(t, t')$ are unknown EFT kernels describing the non-locality in time. The

$T^{ij}_\alpha(\vec{x},a)$ are then organized in a local spatial-derivative expansion of the long-wavelength fields ($\partial_i\partial_j\Phi(\vec{x},a)$, $\partial_i v^j(\vec{x},a)$, etc.) and of stochastic fields $\epsilon^{ij}(\vec{x},a)$.

Let us quickly comment on the stochastic fields. Physically, they capture the fact that short modes can randomly combine to produce long modes, and as we will see, they are needed for renormalization. They are uncorrelated with the matter fields

$$\langle \epsilon^{ij}(\vec{x},a)\delta(\vec{y},a)\rangle = 0 \ . \tag{2.14}$$

In momentum space, we define the correlation of the stochastic fields $\epsilon_n^{ij\cdots}$ as an expansion in powers of $\vec{k}$ of all of the terms allowed by rotation invariance [5], for example

$$\begin{aligned}\langle \epsilon_a^{ij}(\vec{k})\epsilon_b^{kl}(\vec{k}')\rangle =&(2\pi)^3\delta_D(\vec{k}+\vec{k}')\Big(c^{(1)}_{a,b}\delta^{ij}\delta^{kl}+c^{(2)}_{a,b}(\delta^{ik}\delta^{jl}+\delta^{il}\delta^{jk})\\ &+k_{\rm NL}^{-2}\left(c^{(3)}_{a,b}\delta^{ij}k^kk^l+c^{(4)}_{a,b}\delta^{kl}k^ik^j+c^{(5)}_{a,b}(\delta^{ik}k^jk^l+\delta^{il}k^jk^k)\right)+\dots\Big) \ .\end{aligned} \tag{2.15}$$

2.3 Solving the EOM

The equations of motion in the form that we will use for perturbation theory are

$$\begin{aligned}&a\delta'-\Theta=\delta\Theta+\partial_i\delta\frac{\partial_i\Theta}{\partial^2} \ ,\\ &a\Theta'+\left(1+\frac{a\mathcal{H}'}{\mathcal{H}}\right)\Theta-\frac{3}{2}\Omega_{\rm m}\delta=\partial_i\Theta\frac{\partial_i\Theta}{\partial^2}+\frac{\partial_i\partial_j\Theta}{\partial^2}\frac{\partial_i\partial_j\Theta}{\partial^2}+\frac{1}{\mathcal{H}(a)^2}\partial_i\left(\frac{\partial_j\tau^{ij}}{\bar\rho(1+\delta)}\right) \ ,\end{aligned} \tag{2.16}$$

where we have introduced the rescaled velocity divergence

$$\Theta=-\frac{\partial_i v^i}{aH} \ , \tag{2.17}$$

and to a high degree of accuracy, we only need the divergence of the velocity (can ignore the vorticity).

As we saw above, the stress tensor enters the solution for δ as a higher-derivative contribution, so it makes sense to treat it as a perturbation. So for now, we focus on the EOM with $\tau^{ij}=0$, and include that contribution later. This is sometimes called 'standard perturbation theory,' or SPT. To solve perturbatively, we expand

$$\delta(\vec{k},a)=\sum_n\delta^{(n)}(\vec{k},a) \ , \tag{2.18}$$

where $\delta^{(n)}\sim[\delta^{(1)}]^n$.

Linearizing the equations above, we find a scale-independent linear equation for δ, so we can write in momentum space

$$a^2\delta^{(1)}(\vec{k},a)''+\left(2+\frac{a\mathcal{H}'}{\mathcal{H}}\right)a\delta^{(1)}(\vec{k},a)'-\frac{3}{2}\Omega_{\rm m}\delta^{(1)}(\vec{k},a)=0 \ . \tag{2.19}$$

There are of course two independent solutions to this second-order differential equation, so we can write

$$\delta^{(1)}(\vec{k},a)=D_+(a)\tilde\delta^{(1)}_+(\vec{k})+D_-(a)\tilde\delta^{(1)}_-(\vec{k}) \ , \tag{2.20}$$

where $\tilde{\delta}_{\pm}^{(1)}(\vec{k})$ are the initial conditions for the field. One can check that we can choose these solutions to be [6]

$$D_+(a) = \frac{5}{2}\Omega_{\mathrm{m},0}\mathcal{H}_0^2\frac{\mathcal{H}(a)}{a}\int_0^a \frac{da_1}{\mathcal{H}(a_1)^3} \ , \quad \text{and} \quad D_-(a) = \frac{H(a)}{H_0\Omega_{\mathrm{m},0}^{1/2}} \ , \tag{2.21}$$

where we have chosen the normalizations for future convenience. It turns out that, in our expanding universe, $D_+(a)$ grows with time, and $D_-(a)$ decays, so much so that we can ignore the decaying mode for our purposes here (and we will now drop the $+$ and write $D(a) \equiv D_+(a)$ and $\tilde{\delta}^{(1)}(\vec{k}) \equiv \tilde{\delta}_+^{(1)}(\vec{k})$). Thus, we take

$$\delta^{(1)}(\vec{k}, a) = D(a)\tilde{\delta}^{(1)}(\vec{k}) \ . \tag{2.22}$$

Now we can also get the linear solution for the velocity. Looking back at Eq. (2.16), this means that the solution for the velocity divergence is

$$\Theta^{(1)}(\vec{x}, a) = f(a)\delta^{(1)}(\vec{x}, a) \ , \tag{2.23}$$

where the logarithmic derivative of the growth factor $D(a)$ is called the growth function $f(a)$

$$f(a) \equiv \frac{aD'(a)}{D(a)} \ . \tag{2.24}$$

Now let's look for the non-linear solutions. The EOM Eq. (2.16) (with $\tau^{ij} = 0$) are, in Fourier space,

$$\begin{aligned} a\delta(\vec{k}, a)' - \Theta(\vec{k}, a) &= \int_{\vec{k}_1,\vec{k}_2}^{\vec{k}} \alpha(\vec{k}_1, \vec{k}_2)\Theta(\vec{k}_1, a)\delta(\vec{k}_2, a) \ , \\ a\Theta(\vec{k}, a)' + \left(1 + \frac{a\mathcal{H}'}{\mathcal{H}}\right)\Theta(\vec{k}, a) - \frac{3}{2}\Omega_{\mathrm{m}}\delta(\vec{k}, a) &= \int_{\vec{k}_1,\vec{k}_2}^{\vec{k}} \beta(\vec{k}_1, \vec{k}_2)\Theta(\vec{k}_1, a)\Theta(\vec{k}_2, a) \ , \end{aligned} \tag{2.25}$$

where α and β are the dark matter interaction vertices,

$$\alpha(\vec{q}_1, \vec{q}_2) = 1 + \frac{\vec{q}_1 \cdot \vec{q}_2}{q_1^2} \qquad \text{and} \qquad \beta(\vec{q}_1, \vec{q}_2) = \frac{|\vec{q}_1 + \vec{q}_2|^2 \vec{q}_1 \cdot \vec{q}_2}{2q_1^2 q_2^2} \ . \tag{2.26}$$

We then make the following ansatz for the non-linear solutions

$$\begin{aligned} \delta^{(n)}(\vec{k}, a) &= D(a)^n \int_{\vec{k}_1,\ldots,\vec{k}_n}^{\vec{k}} F_n(\vec{k}_1, \ldots, \vec{k}_n)\tilde{\delta}^{(1)}(\vec{k}_1)\cdots\tilde{\delta}^{(1)}(\vec{k}_n) \ , \\ \Theta^{(n)}(\vec{k}, a) &= f(a)D(a)^n \int_{\vec{k}_1,\ldots,\vec{k}_n}^{\vec{k}} G_n(\vec{k}_1, \ldots, \vec{k}_n)\tilde{\delta}^{(1)}(\vec{k}_1)\cdots\tilde{\delta}^{(1)}(\vec{k}_n) \ , \end{aligned} \tag{2.27}$$

and also use the tilde to denote the time-independent fields

$$\tilde{\delta}^{(n)}(\vec{k}) \equiv \frac{\delta^{(n)}(\vec{k}, a)}{D(a)^n} \ , \quad \text{and} \quad \tilde{\Theta}^{(n)}(\vec{k}) \equiv \frac{\Theta^{(n)}(\vec{k}, a)}{f(a)D(a)^n} \ , \tag{2.28}$$

specifically for the ansatz in Eq. (2.27). Plugging Eq. (2.27) into Eq. (2.25) and matching perturbative orders, we have

$$\begin{aligned} D(a)^n f(a) \left(n\tilde{\delta}^{(n)}(\vec{k}) - \tilde{\Theta}^{(n)}(\vec{k}) \right) &= D(a)^n f(a) \int_{\vec{k}_1,\vec{k}_2}^{\vec{k}} \alpha(\vec{k}_1,\vec{k}_2) \sum_{m=1}^{n-1} \tilde{\Theta}^{(m)}(\vec{k}_1)\tilde{\delta}^{(n-m)}(\vec{k}_2) \ , \\ D(a)^n f(a)^2 \left((n-1)\tilde{\Theta}^{(n)}(\vec{k}) + \frac{3}{2}\frac{\Omega_{\rm m}(a)}{f(a)^2} \left(\tilde{\Theta}^{(n)}(\vec{k}) - \tilde{\delta}^{(n)}(\vec{k}) \right) \right) &= \\ D(a)^n f(a)^2 \int_{\vec{k}_1,\vec{k}_2}^{\vec{k}} \beta(\vec{k}_1,\vec{k}_2) \sum_{m=1}^{n-1} \tilde{\Theta}^{(m)}(\vec{k}_1)\tilde{\Theta}^{(n-m)}(\vec{k}_2) \ . \end{aligned} \tag{2.29}$$

where we have used the linear equation Eq. (2.19) which gives

$$af'(a) = -f(a) - f(a)^2 + \frac{3}{2}\Omega_{\rm m}(a) - f(a)\frac{a\mathcal{H}'(a)}{\mathcal{H}(a)} \ . \tag{2.30}$$

We notice that the time dependence cancels from the equations if $\Omega_{\rm m}(a) = f(a)^2$, which is not true in general, but turns out to be approximately true in our universe. Using the approximation

$$\Omega_{\rm m}(a) \approx f(a)^2 \tag{2.31}$$

is called the EdS approximation. This also is exactly true in the so-called Einstein de Sitter (EdS) universe, which has

$$\Omega_{\rm m}(a) = 1 \ , \quad f(a) = 1 \ , \quad D(a) = a \ , \quad \text{and} \quad \frac{a\mathcal{H}'(a)}{\mathcal{H}(a)} = -\frac{1}{2} \ . \tag{2.32}$$

So, using the EdS approximation, the time dependence cancels from Eq. (2.29), so we have the following time-independent equations for the F_n and G_n kernels

$$\begin{aligned} nF_n(\vec{k}_1,\ldots,\vec{k}_n) - G_n(\vec{k}_1,\ldots,\vec{k}_n) &= \sum_{m=1}^{n-1} \alpha(\vec{k}_{1;m},\vec{k}_{m+1;n}) \\ &\qquad \times G_m(\vec{k}_1,\ldots,\vec{k}_m)F_{n-m}(\vec{k}_{m+1},\ldots,\vec{k}_n) \ , \\ -F_n(\vec{k}_1,\ldots,\vec{k}_n) + \left(n+\frac{1}{2}\right) G_n(\vec{k}_1,\ldots,\vec{k}_n) &= \sum_{m=1}^{n-1} \beta(\vec{k}_{1;m},\vec{k}_{m+1;n}) \\ &\qquad \times G_m(\vec{k}_1,\ldots,\vec{k}_m)G_{n-m}(\vec{k}_{m+1},\ldots,\vec{k}_n) \ , \end{aligned} \tag{2.33}$$

where $\vec{k}_{i;j} \equiv \vec{k}_i + \cdots + \vec{k}_j$, which can be algebraically solved to give the famous dark-matter recursion

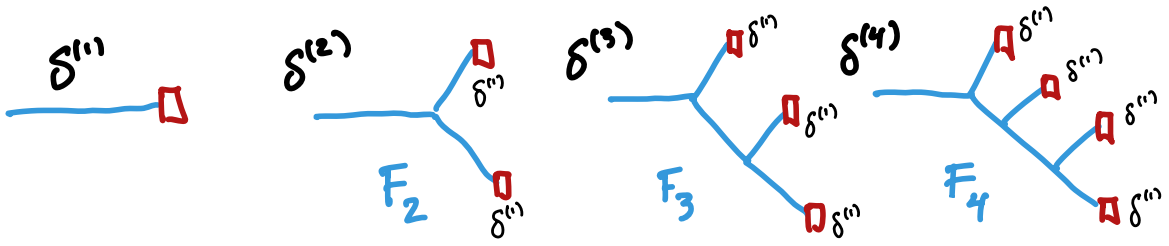

Figure 2: Diagrammatic representation of solutions.

relations [7, 8]

$$\begin{aligned} F_n(\vec{k}_1,\dots,\vec{k}_n) &= \operatorname*{sym.}_{\vec{k}_1,\dots,\vec{k}_n} \sum_{m=1}^{n-1} \frac{G_m(\vec{k}_1,\dots,\vec{k}_m)}{(2n+3)(n-1)} \Big((1+2n)\alpha(\vec{k}_{1;m},\vec{k}_{m+1;n}) F_{n-m}(\vec{k}_{m+1},\dots,\vec{k}_n) \\ &\qquad + 2\beta(\vec{k}_{1;m},\vec{k}_{m+1;n}) G_{n-m}(\vec{k}_{m+1},\dots,\vec{k}_n) \Big) \,, \\ G_n(\vec{k}_1,\dots,\vec{k}_n) &= \operatorname*{sym.}_{\vec{k}_1,\dots,\vec{k}_n} \sum_{m=1}^{n-1} \frac{G_m(\vec{k}_1,\dots,\vec{k}_m)}{(2n+3)(n-1)} \Big(3\alpha(\vec{k}_{1;m},\vec{k}_{m+1;n}) F_{n-m}(\vec{k}_{m+1},\dots,\vec{k}_n) + \\ &\qquad + 2n\beta(\vec{k}_{1;m},\vec{k}_{m+1;n}) G_{n-m}(\vec{k}_{m+1},\dots,\vec{k}_n) \Big) \,, \end{aligned} \tag{2.34}$$

where $F_1 = G_1 = 1$, and the F_n and G_n are always symmetrized over $\vec{k}_1,\dots,\vec{k}_n$, and we take

$$\operatorname*{sym.}_{\vec{k}_1,\dots,\vec{k}_n} f(\vec{k}_1,\dots,\vec{k}_n) \equiv \frac{1}{n!} \sum_{\sigma\in S_n} f(\sigma(\vec{k}_1,\dots,\vec{k}_n)) \,. \tag{2.35}$$

A diagrammatic representation of the solution is given in Fig. 2.

2.4 Observables, loops, and diagrams

The observables that we compute are correlation functions. For example, the two-point and three-point correlation functions in Fourier space, which are called the power spectrum and bispectrum

$$\begin{aligned} \langle \delta(\vec{k},a)\delta(\vec{k}',a)\rangle &= (2\pi)^3 \delta_D(\vec{k}+\vec{k}') P(k) \,, \\ \langle \delta(\vec{k}_1,a)\delta(\vec{k}_2,a)\delta(\vec{k}_3,a)\rangle &= (2\pi)^3 \delta_D(\vec{k}_1+\vec{k}_2+\vec{k}_3) B(k_1,k_2,k_3) \,. \end{aligned} \tag{2.36}$$

The dependencies of P and B on the wavenumbers are due to imposing translation and rotation invariance of the correlation functions.

In perturbation theory, we expand the fields as in Eq. (2.27). We then assume Gaussian initial conditions (coming from inflation), and use Wick's theorem to compute the connected correlations. For example, let's look at the one-loop power spectrum. We have

$$\langle \delta(\vec{k},a)\delta(\vec{k}',a)\rangle = \langle \delta^{(1)}(\vec{k},a)\delta^{(1)}(\vec{k}',a)\rangle + 2\langle \delta^{(1)}(\vec{k},a)\delta^{(3)}(\vec{k}',a)\rangle + \langle \delta^{(2)}(\vec{k},a)\delta^{(2)}(\vec{k}',a)\rangle + \dots \,. \tag{2.37}$$

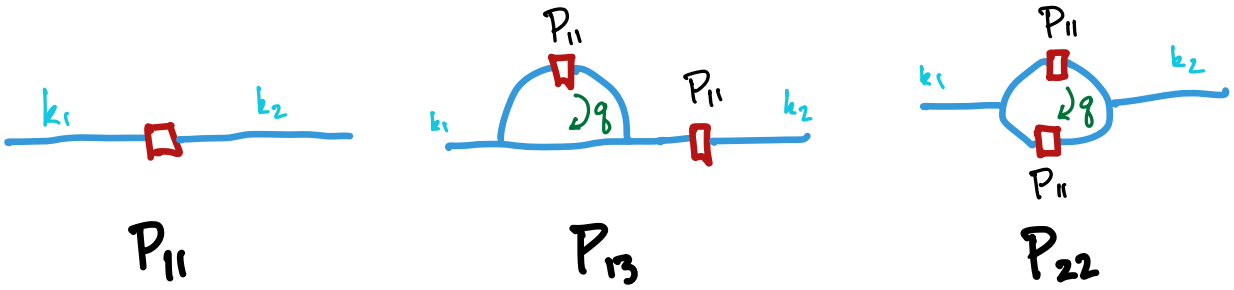

Figure 3: Diagrammatic representation of the power spectrum.

The linear power spectrum is like the propagator

$$\langle\tilde{\delta}^{(1)}(\vec{k})\tilde{\delta}^{(1)}(\vec{k}')\rangle = (2\pi)^3\delta_D(\vec{k}+\vec{k}')P_{11}(k)\ . \tag{2.38}$$

Then, we have the 13 diagram

$$\begin{aligned}\langle\delta^{(1)}(\vec{k},a)\delta^{(3)}(\vec{k}',a)\rangle &= D(a)^4\int_{\vec{k}_1,\vec{k}_2,\vec{k}_3}^{\vec{k}'} F_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)\langle\tilde{\delta}^{(1)}(\vec{k})\tilde{\delta}^{(1)}(\vec{k}_1)\tilde{\delta}^{(1)}(\vec{k}_2)\tilde{\delta}^{(1)}(\vec{k}_3)\rangle \\ &= 6D(a)^4\int_{\vec{k}_1,\vec{k}_2,\vec{k}_3}^{\vec{k}'} F_3(\vec{k}_1,\vec{k}_2,\vec{k}_3)\langle\tilde{\delta}^{(1)}(\vec{k})\tilde{\delta}^{(1)}(\vec{k}_1)\rangle\langle\tilde{\delta}^{(1)}(\vec{k}_2)\tilde{\delta}^{(1)}(\vec{k}_3)\rangle \\ &= (2\pi)^3\delta_D(\vec{k}+\vec{k}')6D(a)^4P_{11}(k)\int_{\vec{q}} F_3(\vec{q},-\vec{q},\vec{k})P_{11}(q)\end{aligned} \tag{2.39}$$

which also defines P_{13}. For the 22 diagram, we do something similar to get

$$P_{22}(k) = 2\int_{\vec{q}} F_2(\vec{q},\vec{k}-\vec{q})^2P_{11}(q)P_{11}(|\vec{k}-\vec{q}|)\ . \tag{2.40}$$

For example, these kernels look like

$$F_2(\vec{k}_1,\vec{k}_2) = \frac{5}{7} + \frac{\hat{k}_1\cdot\hat{k}_2}{2}\left(\frac{k_1}{k_2}+\frac{k_2}{k_1}\right) + \frac{2}{7}(\hat{k}_1\cdot\hat{k}_2)^2 \tag{2.41}$$

and

$$\begin{aligned}F_3(\vec{q},-\vec{q},\vec{k}) = &-\frac{97}{1512} + \frac{|\vec{k}-\vec{q}|^2}{24k^2} + \frac{1195k^2}{6552|\vec{k}-\vec{q}|^2} - \frac{19|\vec{k}-\vec{q}|^4}{504q^4} + \frac{|\vec{k}-\vec{q}|^2k^2}{14q^4} - \frac{5k^4}{168q^4} \\ &- \frac{k^6}{252|\vec{k}-\vec{q}|^2q^4} + \frac{211|\vec{k}-\vec{q}|^2}{1512q^2} - \frac{|\vec{k}-\vec{q}|^4}{72k^2q^2} - \frac{187k^2}{1512q^2} - \frac{k^4}{504|\vec{k}-\vec{q}|^2q^2} \\ &- \frac{19q^2}{504|\vec{k}-\vec{q}|^2} - \frac{q^2}{24k^2} + \frac{q^4}{72|\vec{k}-\vec{q}|^2k^2}\end{aligned} \tag{2.42}$$

So ultimately, these are the kinds of integrals that we have to evaluate. The diagrammatic representation is shown in Fig. 3.

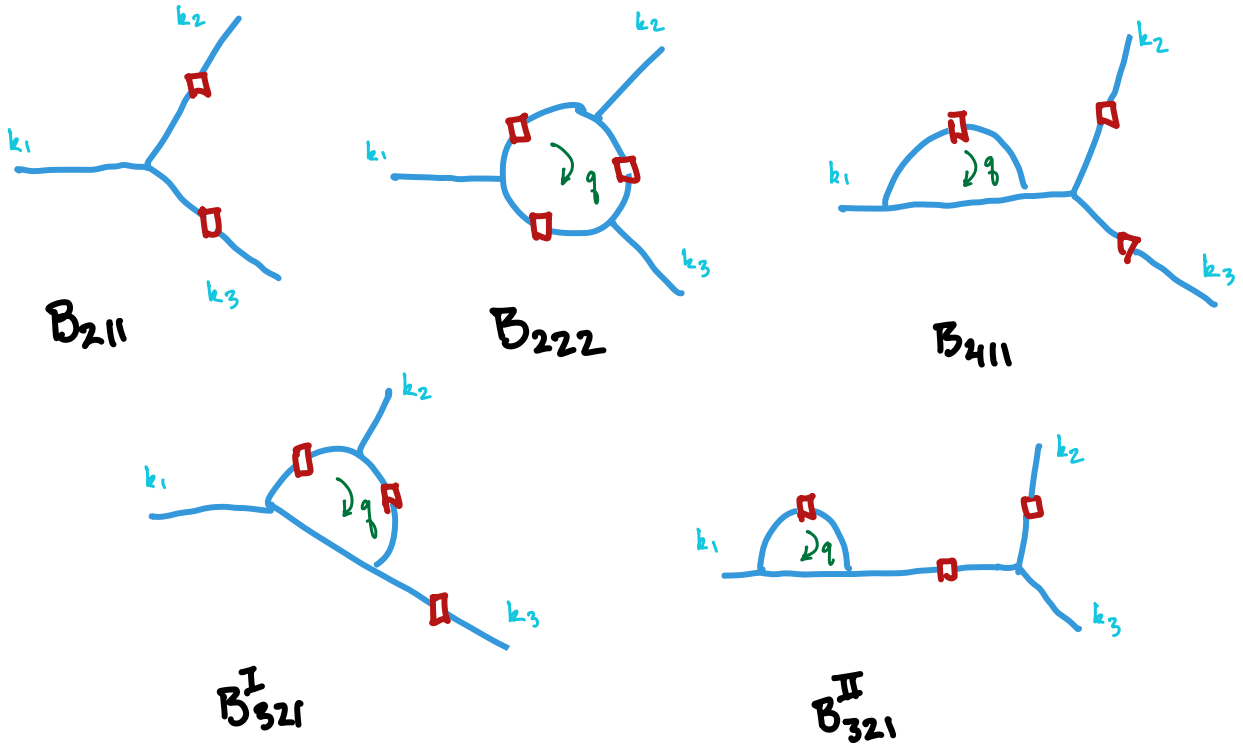

Figure 4: Diagrammatic representation of bispectrum.

There are four types of diagrams for the one-loop bispectrum, and one needs up to $\delta^{(4)}$ to do that. The total one-loop bispectrum is

$$B_{\text{1-loop tot.}} = D(a)^4 B_{211} + D(a)^6 \left(B_{222} + B^{(I)}_{321} + B^{(II)}_{321} + B_{411} \right) \ , \tag{2.43}$$

where the tree-level bispectrum is

$$B_{211}(k_1, k_2, k_3) = 2F_2(\vec{k}_1, \vec{k}_2) P_{11}(k_1) P_{11}(k_2) + \ 2 \text{ perms.} \ , \tag{2.44}$$

and the one-loop contributions are

$$\begin{aligned}
B_{222}(k_1, k_2, k_3) &= 8 \int_{\vec{q}} P_{11}(q) P_{11}(|\vec{k}_2 - \vec{q}|) P_{11}(|\vec{k}_1 + \vec{q}|) \\
&\qquad \times F_2(-\vec{q}, \vec{k}_1 + \vec{q}) F_2(\vec{k}_1 + \vec{q}, \vec{k}_2 - \vec{q}) F_2(\vec{k}_2 - \vec{q}, \vec{q}) \ , \\
B^{(I)}_{321}(k_1, k_2, k_3) &= 6 P_{11}(k_1) \int_{\vec{q}} P_{11}(q) P_{11}(|\vec{k}_2 - \vec{q}|) \\
&\qquad \times F_3(-\vec{q}, -\vec{k}_2 + \vec{q}, -\vec{k}_1) F_2(\vec{q}, \vec{k}_2 - \vec{q}) + \ 5 \text{ perms.} \ , \\
B^{(II)}_{321}(k_1, k_2, k_3) &= 6 P_{11}(k_1) P_{11}(k_2) F_2(\vec{k}_1, \vec{k}_2) \int_{\vec{q}} P_{11}(q) F_3(\vec{k}_1, \vec{q}, -\vec{q}) + \ 5 \text{ perms.} \ , \\
B_{411}(k_1, k_2, k_3) &= 12 P_{11}(k_1) P_{11}(k_2) \int_{\vec{q}} P_{11}(q) F_4(\vec{q}, -\vec{q}, -\vec{k}_1, -\vec{k}_2) + \ 2 \text{ perms.} \ .
\end{aligned} \tag{2.45}$$

The diagrammatic representation is shown in Fig. 4.

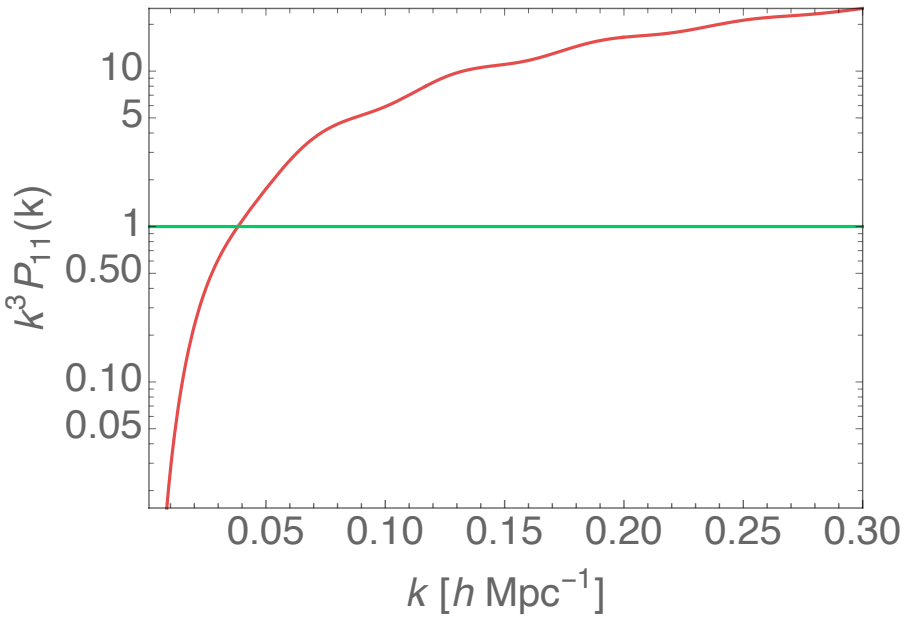

Figure 5: Plot of P_{11}

2.5 What is P_{11}?

We mentioned the early universe, Boltzmann codes, linear equations, the Dodelson book [6], etc. A plot of $P_{11}(k)$ for a typical set of cosmological parameters A_s, n_s, $\Omega_{m,0}$, and H_0 is given in Fig. 5.

2.6 Counterterms

Notice that the dimensionless power spectrum becomes bigger than $\mathcal{O}(1)$ at quite a low scale, but the integral in Eq. (2.39) goes all the way to infinity. This means that we are integrating over modes where we do not have perturbative control. It turns out, however, that the integral is convergent. But this is just an accident. There is no way that this integral will give us the right number, since it is outside of the perturbative regime. Specifically, look at the UV limit

$$\begin{aligned} P_{13}(k) &\to -\frac{61}{630\pi^2} k^2 P_{11}(k) \int^{\Lambda_{\rm UV}} dq\, P_{11}(q)\ , \\ P_{22}(k) &\to \frac{9}{196\pi^2} k^4 \int^{\Lambda_{\rm UV}} dq \frac{P_{11}(q)^2}{q^2}\ , \end{aligned} \tag{2.46}$$

The values of these integrals, even if not formally UV divergent, depend on $\Lambda_{\rm UV}$, and in any case cannot be correct. This is where the counterterms come in.

Let's go back to the stress tensor Eq. (2.13). To find the most general expression up to the desired order, we first start by listing all of the T^{ij}_α that we can construct from $\partial_i\partial_j\Phi$, $\partial_i v^j$, and

ϵ^{ij}. Then, we expand the $\vec{x}_{\rm fl}$ in terms of the velocity, and since all of the time dependencies of the solutions are simply powers of $D(a)$, the time integrals over the kernels κ_α can be done. Then, one can find the minimal basis of operators. We will detail this procedure later.

At first order (again using tilde to mean that we stripped away all of the time dependent factors), we have

$$\frac{1}{\bar{\rho}\mathcal{H}^2}\partial_i\partial_j\tau^{ij}_{(1)} = \frac{\partial_i\partial_j}{k^2_{\rm NL}}\left(D(a)^3c_s^2\delta^{ij}\tilde{\delta}^{(1)} + D(a)^2\tilde{\epsilon}^{ij}\right) \ , \tag{2.47}$$

which when plugging into the EOM, leads to

$$\delta^{\rm ct}(\vec{k},a) \sim -D(a)^3c_s^2\frac{k^2}{k^2_{\rm NL}}\tilde{\delta}^{(1)}(\vec{k}) - D(a)^2\frac{k^ik^j}{k^2_{\rm NL}}\tilde{\epsilon}^{ij}(\vec{k}) \tag{2.48}$$

which ultimately leads to

$$\begin{aligned}\Delta P_{13}(k) &\propto c_s^2\frac{k^2}{k^2_{\rm NL}}P_{11}(k) \ , \\ \Delta P_{22}(k) &\propto c_\epsilon\frac{k^4}{k^4_{\rm NL}} \ .\end{aligned} \tag{2.49}$$

Notice that these are exactly the forms of the UV limits of the loops. The values of c_s^2 and c_ϵ remove the $\Lambda_{\rm UV}$ dependence and allow us to set the correct value that matches the true universe. It is interesting to look back at Eq. (2.42), where we see that only one term there corresponds to the ΔP_{13} counterterm, meaning all of the other coefficients are predictions of the theory.

To renormalize the bispectrum, we have to go up to second order in the counterterms. The expressions are longer, but the idea is exactly the same. There are 4 response and 6 stochastic free coefficients. A nice check is that the UV limits of all of the loops can be absorbed by the available free EFT coefficients.

2.7 Symmetries and kernel identities

Given that our observables are in terms of the K_n and G_n kernels, let's look at some general properties. First of all, the forms of the UV limits of the loops are determined by the symmetries, and so have to be equivalent to local counterterms that one can add to τ^{ij}. For dark matter, where we have mass and momentum conservation, this means that the UV limit of P_{13} is k^2P_{11}, and the UV limit of P_{22} is k^4. Similar limits for the bispectrum and higher loops also apply. In general, mass and momentum conservation implies that [7, 9]

$$\lim_{|\vec{k}_1+\cdots+\vec{k}_n|\to 0} F_n(\vec{k}_1,\ldots,\vec{k}_n) \propto |\vec{k}_1+\cdots+\vec{k}_n|^2 \ . \tag{2.50}$$

Furthermore, from the equation for δ Eq. (2.12) (whose structure is also a consequence of mass and momentum conservation), we can also see that

$$\lim_{|\vec{k}_1+\cdots+\vec{k}_{n-2}|\ll\vec{q}} F_n(\vec{k}_1,\ldots,\vec{k}_{n-2},\vec{q},-\vec{q}) \propto \frac{|\vec{k}_1+\cdots+\vec{k}_{n-2}|^2}{q^2} \ . \tag{2.51}$$

This comes from the fact that

$$\delta \sim \partial_i \partial_j \left(\frac{\partial_i \delta}{\partial^2} \frac{\partial_j \delta}{\partial^2} \right) \ , \tag{2.52}$$

and taking a $\vec{q}$ in the first δ on the right-hand side and a $-\vec{q}$ in the second δ on the right-hand side (taking $\vec{q}$ and $-\vec{q}$ in the same factor of δ will make it just cancel out in the denominator). We will see below that this is no longer the case for galaxies, which do not have mass and momentum conservation.

Another interesting limit is when internal momenta go to zero. These limits are determined by Galilean invariance (which one can see because it involves a space-independent transformation, which is when momentum is going to zero). A consequence of that is [10]

$$\lim_{\vec{q}_1,\ldots,\vec{q}_m \to 0} F_{n+m}(\vec{k}_1,\ldots,\vec{k}_n,\vec{q}_1,\ldots,\vec{q}_m) \to \frac{n!}{(n+m)!} \frac{\vec{q}_1 \cdot \sum_i \vec{k}_i}{q_1^2} \ldots \frac{\vec{q}_m \cdot \sum_i \vec{k}_i}{q_m^2} F_n(\vec{k}_1,\ldots,\vec{k}_n) \ . \tag{2.53}$$

2.8 c_s^2 running

This is a quite nice test of the validity of our EFT approach [2], and indeed shows that it is correct. Imagine that we have a good N-body simulation, which exactly solves the equations of motion deep into the UV, $\Lambda \to \infty$. Now, let us do the two following procedures to measure the counterterm c_s^2.

First, we can smooth the simulation with various cutoffs, say $\Lambda_1 = 1/6\,h\,\mathrm{Mpc}^{-1}$ and $\Lambda_2 = 1/3\,h\,\mathrm{Mpc}^{-1}$. This directly gives us the long-wavelength fields. We can then take various correlations of these long-wavelength fields to extract $c_s^2(\Lambda)$. Schematically, we have

$$\langle \partial_i \partial_j [\tau^{ij}]_\Lambda \partial_k \partial_l [\tau^{kl}]_\Lambda \rangle \supset c_s^2(\Lambda) \langle \partial_i \partial_j [\tau^{ij}]_\Lambda \partial^2 [\delta]_\Lambda \rangle + \ldots \ . \tag{2.54}$$

So, doing this on two simulations, we have two different values $c_s^2(\Lambda_1)$ and $c_s^2(\Lambda_2)$.

Second, we can match to some observable, like the power spectrum, at a renormalization scale $k_{\text{ren.}}$. This is typical of what we do in QFT. We say

$$c_{s,\text{bare}}^2(\Lambda, k_{\text{ren.}}) = c_{s,\text{ren.}}^2(k_{\text{ren.}}) + c_{s,\text{ctr.}}^2(\Lambda) \ . \tag{2.55}$$

The counterterm contribution $c_{s,\text{ctr.}}^2(\Lambda)$ has to exactly cancel the Λ dependence of the loop, e.g. Eq. (2.46). Even if the loop is finite, this still has to cancel the dependence on Λ. Then, we determine the 'finite' piece $c_{s,\text{ren.}}^2(k_{\text{ren.}})$ by matching the power spectrum to data at the scale $k_{\text{ren.}}$. One can of course use any sufficiently large Λ. However, it is convenient to take $\Lambda \to \infty$ because this means that we can safely ignore higher-derivative operators (suppressed by k/Λ) when matching to the observables. Then, one can run to finite values of Λ using the fact that the total power spectrum does not depend on the cutoff Λ. Defining

$$P_{13}^{\text{tot.}}(k, \Lambda, k_{\text{ren.}}) \equiv P_{13}^{\Lambda}(k) - c_{s,\ \text{bare}}^2(\Lambda, k_{\text{ren.}}) \frac{k^2}{k_{\text{NL}}^2} P_{11}(k) \ , \tag{2.56}$$

we have

$$P_{13}^{\text{tot.}}(k, \Lambda, k_{\text{ren.}}) - P_{13}^{\text{tot.}}(k, \infty, k_{\text{ren.}}) = 0 \ , \tag{2.57}$$

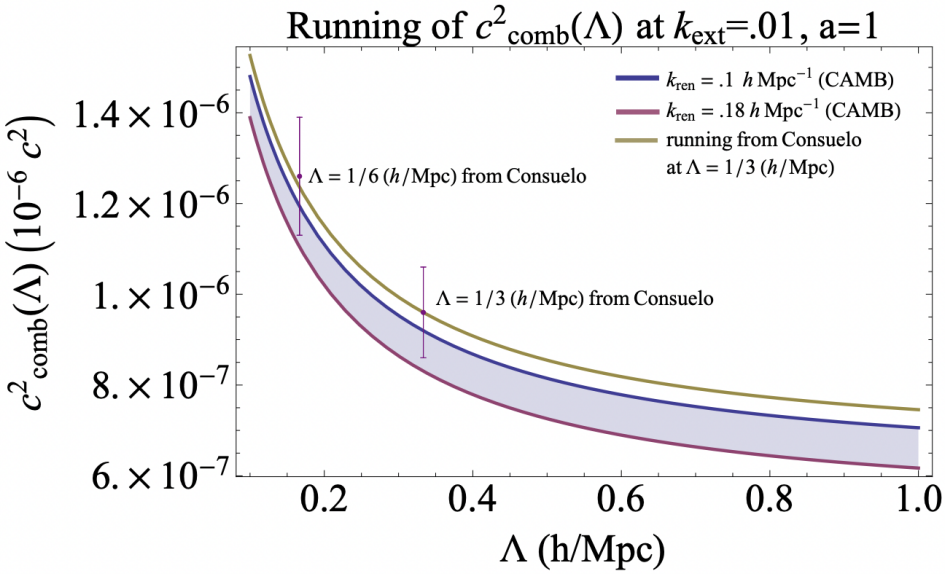

Figure 6: Plot of running of c_s^2. Here, $c^2_{\rm comb}(\Lambda)$ is $c^2_{s,\rm bare}(\Lambda, k_{\rm ren.})$ in Eq. (2.55), and $k_{\rm ext}$ is k in Eq. (2.59).

which implies that

$$P_{13}^{\Lambda}(k) - P_{13}^{\infty}(k) - \left(c^2_{s,\ \rm bare}(\Lambda, k_{\rm ren.}) - c^2_{s,\ \rm bare}(\infty, k_{\rm ren.})\right) \frac{k^2}{k_{\rm NL}^2} P_{11}(k) = 0 \ , \tag{2.58}$$

and finally

$$c^2_{s,\rm bare}(\Lambda, k_{\rm ren.}) = c^2_{s,\rm bare}(\infty, k_{\rm ren.}) + \frac{P_{13}^{\Lambda}(k) - P_{13}^{\infty}(k)}{(k/k_{\rm NL})^2 P_{11}(k)} \ , \tag{2.59}$$

or

$$c^2_{s,\rm ctr.}(\Lambda) = c^2_{s,\rm ctr.}(\infty) + \frac{P_{13}^{\Lambda}(k) - P_{13}^{\infty}(k)}{(k/k_{\rm NL})^2 P_{11}(k)} \tag{2.60}$$

for $k \ll \Lambda$, and where $P_{13}^{\Lambda}(k)$ is the SPT loop with cutoff Λ. Fig. 6 shows how these two approaches agree.

3 Biased tracers

Above we discussed the solution for CDM clustering, which as we saw is connected to early universe physics by P_{11}. CDM is the main component of the universe, so this is a crucial component to understand. However, we do not directly observe CDM, so we have to connect it to other observables. Here, we are concerned with the density of galaxies, which we do in fact observe. So now we will connect the CDM distribution to the distribution of galaxies.

3.1 Bias expansion

Some general references are [11–17].

The general expression for the density of galaxies is something like this:

$$\delta_g(\vec{x},t) = \int^t dt' H(t') f_g\left(\partial_i\partial_j\Phi(\vec{x}_{\rm fl},t'), \partial_i v^j(\vec{x}_{\rm fl},t'), \epsilon^{ij}(\vec{x}_{\rm fl},t'); m_e, m_p, \alpha; \ldots; t, t'\right) \ , \tag{3.1}$$

i.e. given by some complicated function f_g of many different variables, integrated over the past fluid element. It is impossible to know this function exactly, so what we do is Taylor expand this function for the small fluctuations $\partial_i\partial_j\Phi$, $\partial_i v^j$, and ϵ^{ij}.

So, we must now write down the most general expression for a Galilean scalar integrated over time, in an expansion in the number of fields and derivatives. For now, we will focus on the lowest order in derivatives, and ignore the stochastic field ϵ^{ij}. The building blocks of Galilean scalars are the dimensionless tensors

$$r_{ij} \equiv \frac{2\partial_i\partial_j\Phi}{3\Omega_{\rm m}a^2H^2} \ , \quad \text{and} \quad p_{ij} \equiv -\frac{D}{a\dot{D}}\partial_i v^j \ , \tag{3.2}$$

such that their traces are the familiar fields from before $\delta^{ij}r_{ij} = \delta$ and $\delta^{ij}p_{ij} = \theta$.

Overall, this means that at n-th order, we have

$$\delta_g^{(n)}(\vec{x},t) = \sum_{\mathcal{O}_m} \int^t dt' H(t') c_{\mathcal{O}_m}(t,t') [\mathcal{O}_m(\vec{x}_{\rm fl}(\vec{x},t,t'),t')]^{(n)} \ , \tag{3.3}$$

where $\{\mathcal{O}_m\}$ is the set of all contractions of r_{ij} and p_{ij} *up to* n-th order (the subscript m means that the field starts at m-th order), and the coefficients $c_{\mathcal{O}_m}(t,t')$ come from functionally Taylor expanding Eq. (3.1), i.e. they contain the information on UV physics. We have this expansion so far up to fifth order [17]. So, for example, for $\delta_g^{(5)}$, we need to find all of the contractions up to fifth order

$$\{\mathcal{O}_m\} = \{\delta, \theta, \delta^2, \delta\theta, \theta^2, r\cdot r, r\cdot p, p\cdot p, \ldots, p\cdot r\cdot p\cdot r, \ldots\} \ . \tag{3.4}$$

There are 130 such contractions. However, since we only go up to fifth order, we can set $r_{ij}^{(1)} = p_{ij}^{(1)}$ in the terms that start at fifth order. This leaves us with 63 contractions. Then, we will expand the $\mathcal{O}_m$ and the $\vec{x}_{\rm fl}$ in each term and take the n-th order terms. After doing this, the expression in Eq. (3.3) has 151 terms.

Let's look at them more closely. For each $\mathcal{O}_m$, we can write the expansion of the fluid element in the form

$$[\mathcal{O}_m(\vec{x}_{\rm fl}(\vec{x},t,t'),t')]^{(n)} = \sum_{\alpha=1}^{n-m+1} \left(\frac{D(t')}{D(t)}\right)^{\alpha+m-1} \mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}(\vec{x},t) \ , \tag{3.5}$$

where we have used the EdS approximation for the time dependence. The term with $\alpha = 1$ has the most expansion of the velocity in $\vec{x}_{\rm fl}$, while the term with $\alpha = n-m+1$ contains the term with no expansion of $\vec{x}_{\rm fl}$, but also has $\vec{x}_{\rm fl}$ terms.

Then, defining new time-dependent coefficients as the integrals

$$c_{\mathcal{O}_m,\alpha}(t) \equiv \int^t dt' H(t') c_{\mathcal{O}_m}(t,t') \left(\frac{D(t')}{D(t)}\right)^{\alpha+m-1} \ , \tag{3.6}$$

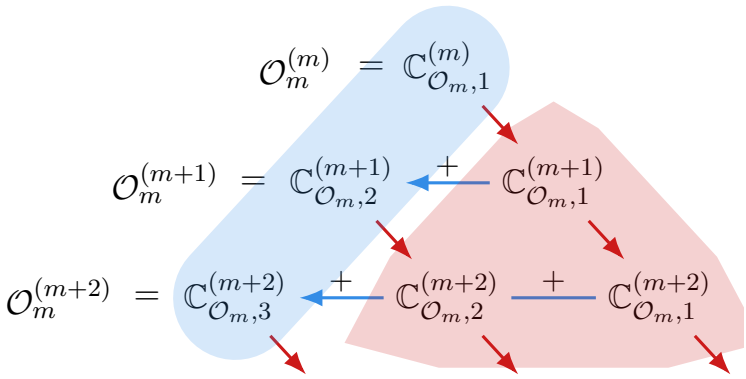

Figure 7: Diagrammatic representation of one way of using the recursion relations Eq. (3.8) and Eq. (3.9) to determine the full set of bias functions $\mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}$ in the fluid expansion of a seed function $\mathcal{O}_m$. The red arrows indicate the use of the fluid recursion Eq. (3.9), while the blue arrows indicate the use of the completeness relation Eq. (3.8). Thus, the terms in the red shading ($\alpha < n - m + 1$) are determined by the fluid recursion Eq. (3.9) and the terms in the blue shading ($\alpha = n - m + 1$) are determined by the completeness relation Eq. (3.8).

we get the most general bias expansion

$$\delta_g^{(n)}(\vec{x},t) = \sum_{\mathcal{O}_m} \sum_{\alpha=1}^{n-m+1} c_{\mathcal{O}_m,\alpha}(t)\, \mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}(\vec{x},t) \ . \tag{3.7}$$

So the functions $\mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}(\vec{x},t)$ essentially describe all of the possible types of signals at order n. But we should check if they are actually independent functions. We'll return to that soon, but first let's look at how we can use a recursion relation to compute these functions in an easy way (such that we formally expand in $\vec{x}_{\rm fl}$, but do not have to do it manually).

The recursion relation comes in two parts. The first is the *equal-time completeness relation*

$$\mathcal{O}_m^{(n)}(\vec{x},t) = \sum_{\alpha=1}^{n-m+1} \mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}(\vec{x},t) \ , \tag{3.8}$$

which is trivially obtained by setting $t = t'$ in Eq. (3.5), and where $\mathcal{O}_m^{(n)}$ is the standard expansion of $\mathcal{O}_m$ at n-th order in perturbations. The second, which captures the consequences of expanding $\vec{x}_{\rm fl}$ in Eq. (3.5), is the *fluid recursion*

$$\mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}(\vec{x},t) = \frac{1}{n-\alpha-m+1} \sum_{\ell=m+\alpha-1}^{n-1} \partial_i \mathbb{C}^{(\ell)}_{\mathcal{O}_m,\alpha}(\vec{x},t) \frac{\partial_i}{\partial^2} \theta^{(n-\ell)}(\vec{x},t) \ , \tag{3.9}$$

which is valid for $n - \alpha - m + 1 > 0$. This recursion is reminiscent of the famous dark-matter recursion relations [7], and provides, for the first time, a full generalization to generic biased tracers. We give a diagrammatic representation of this recursion relation in Fig. 7.

A proof of this is given in detail in [17], but basically, we take d/dt of both sides of Eq. (3.5) and use properties of the fluid element. We can see how to use the recursion relation in Fig. 7,

and we see that we only have to compute $\mathcal{O}_m^{(n)}$, which is made up of the dark-matter solutions of the last lecture, and the rest are determined by simple contractions with $\partial_i\theta/\partial^2$.

Our expansion Eq. (3.7) is fully non-local in time, since the functions $\mathbb{C}^{(n)}_{\mathcal{O}_m,\alpha}(\vec{x},t)$ came from expanding the fluid element that was integrated over all past times. We can compare this with the local-in-time expansion, which has $c_{\mathcal{O}_m}(t,t') = c_{\mathcal{O}_m}(t)\delta_D(t-t')/H(t)$, which gives

$$\delta^{(n)}_{g,\text{loc}}(\vec{x},t) = \sum_{\mathcal{O}_m} c_{\mathcal{O}_m}(t)\,\mathcal{O}_m^{(n)}(\vec{x},t)\ . \tag{3.10}$$

3.2 Bias bases and kernels

Now that we have the most general bias expansion, we should ask if all of the 151 functions $\mathbb{C}^{(5)}_{\mathcal{O}_m,\alpha}(\vec{x},t)$ are actually independent. To answer this, we simply take a brute force approach. It would be nice to have a more clever way to get directly to a complete basis. What we do is convert to momentum space, where we obtain the kernels $K_5^{\mathcal{O}_m,\alpha}(\vec{k}_1,\ldots,\vec{k}_5)$. We then look to see if, as functions of $\vec{k}_i$, any kernels are linear combinations of other kernels. Specifically, what we do is plug in 151 sets of $\vec{k}_i$ and form a 151×151 matrix, and then find the degeneracies of this matrix. This isn't too bad, but perhaps at higher orders, it is overkill.

The bottom line is that we end up with 29 basis elements for the non-local in time expansion,

$$\delta_g^{(5)}(\vec{x},t) = \sum_{j=1}^{29} b_j(t)\mathbb{B}_j^{(5)}(\vec{x},t)\ , \tag{3.11}$$

and only 26 for the local-in-time expansion. This means that it is possible to detect non-local-in-time evolution using the static pictures of galaxy clustering.

All in all, the above procedure allows us to define the kernels K_n^g, which are analogues to the F_n of dark matter, such that

$$\delta_g^{(n)}(\vec{k},a) = D(a)^n \int_{\vec{k}_1,\ldots,\vec{k}_n}^{\vec{k}} K_n^g(\vec{k}_1,\ldots,\vec{k}_n)\tilde{\delta}^{(1)}_{\vec{k}_1}\cdots\tilde{\delta}^{(1)}_{\vec{k}_n}\ . \tag{3.12}$$

These kernels depend on the bias parameters, e.g.

$$K_1^g[b_1]\ ,\quad K_2^g[b_1,b_2,b_5]\ ,\quad K_3^g[b_1,b_2,b_3,b_5,b_6,b_8,b_{10}]\ ,\quad \text{and}\quad K_4^g[b_1,\ldots,b_{15}]\ . \tag{3.13}$$

Since galaxies do not satisfy mass or momentum conservation, the kernels are less suppressed in the UV limit, i.e.

$$\lim_{\vec{k}_i\ll\vec{q}} K_n^g(\vec{k}_1,\ldots,\vec{k}_{n-2},\vec{q},-\vec{q}) \propto \frac{k^0}{q^0}\ . \tag{3.14}$$

However, since δ_g is still a Galilean scalar, the kernels satisfy the same IR limits [10]

$$\lim_{\vec{q}_1,\ldots,\vec{q}_m\to 0} K^g_{n+m}(\vec{k}_1,\ldots,\vec{k}_n,\vec{q}_1,\ldots,\vec{q}_m) \to \frac{n!}{(n+m)!}\frac{\vec{q}_1\cdot\sum_i\vec{k}_i}{q_1^2}\cdots\frac{\vec{q}_m\cdot\sum_i\vec{k}_i}{q_m^2}K_n^g(\vec{k}_1,\ldots,\vec{k}_n)\ . \tag{3.15}$$

3.3 Bias renormalization

We compute loops in the same way that we did for DM, just with the F_n kernels replaced by K_n^g. However, now that the UV limits of the kernels are different, this means that the UV limits of the loops will be different.

For concreteness, let's focus on fourth order [16]. The first thing we should check is that we did not shift the average of the field from zero. To the order relevant here, we have

$$\langle\delta_g(\vec{x})\rangle \approx \langle\delta_g^{(2)}(\vec{x})\rangle = \int_{\vec{q}} K_2^g(\vec{q},-\vec{q})P_{11}(q) = \frac{-b_1+b_2+b_5}{2\pi^2}\int dq\, q^2 P_{11}(q)\ , \tag{3.16}$$

so we should really use the field $\delta_g - \langle\delta_g^{(2)}\rangle$. This is a zero mode, and so is irrelevant for our present purposes, but it becomes important in redshift space. We also note that since number and momentum density are not conserved for tracers, the loops P_{13}^g, B_{411}^g, and $B_{321}^{g,(II)}$ start at k^0 (as opposed to k^2 for dark matter) as $k\to 0$. As described in [18, 14], this is best understood as the renormalization of lower-order bias parameters. For example, consider the power spectrum up to one loop

$$P^g(k) = P_{11}^g(k) + P_{13}^g(k) + P_{22}^g(k)\ . \tag{3.17}$$

In particular, we have $P_{11}^g(k) = b_1^2 P_{11}(k)$, and the UV limits

$$P_{13}^g(k) \to -\frac{b_1}{21\pi^2}\left(13b_1 - 63b_{10} + 34b_2 - 47b_3 + 42b_5 - 110b_6 - 82b_8\right)P_{11}(k)\int^{\Lambda_{\rm UV}} dq\, q^2 P_{11}(q)\ , \tag{3.18}$$

$$P_{22}^g(k) \to \frac{(b_1-b_2-b_5)^2}{\pi^2}\int^{\Lambda_{\rm UV}} dq\, q^2 P_{11}(q)^2\ . \tag{3.19}$$

Since they have the same form as a function of $P_{11}(k)$, we see that the UV limit of P_{13}^g can be absorbed by redefining b_1. You can think of this as a β-function for the bias parameters, i.e. how they change as a function of the UV cutoff. Renormalization of P_{22}^g is done by stochastic terms.

3.4 Higher derivative renormalization

To add the higher-derivative counterterms we write our renormalized field $[\delta_g]_R$ as

$$[\delta_g]_R = \delta_g - \langle\delta_g^{(2)}\rangle + \mathcal{O}_{\delta_g}\ , \tag{3.20}$$

where $\mathcal{O}_{\delta_g}$ contains all of the higher derivative counterterms. Similar to what we did above in expanding τ^{ij}, we now expand a general non-local-in-time Galilean scalar. We find

$$\begin{aligned} k_{\rm NL}^2\mathcal{O}_{\delta_g} = D^3 c_1^g\partial^2\tilde{\delta}^{(1)} + D^4\Bigg(& c_1^g\partial_i\partial^2\tilde{\delta}^{(1)}\frac{\partial_i\tilde{\delta}^{(1)}}{\partial^2} + c_2^g\partial^2\tilde{\delta}^{(2)} - c_2^g\partial_i\partial^2\tilde{\delta}^{(1)}\frac{\partial_i\tilde{\delta}^{(1)}}{\partial^2} \\ &+ c_3^g\partial^2(\tilde{\delta}^{(1)}\tilde{\delta}^{(1)}) + c_4^g\partial^2\left(\frac{\partial_i\partial_j\tilde{\delta}^{(1)}}{\partial^2}\frac{\partial_i\partial_j\tilde{\delta}^{(1)}}{\partial^2}\right) + c_5^g\partial_i\tilde{\delta}^{(1)}\partial_i\tilde{\delta}^{(1)}\Bigg)\ , \end{aligned} \tag{3.21}$$

is the minimal set of possible terms. There are of course stochastic terms, but we'll skip those for lack of time. In this way, we renormalize $\mathcal{O}(k^2P_{11})$ for P_{13}^g, $\mathcal{O}(k^2P_{11}^2)$ for B_{411}^g and $B_{321}^{g,(II)}$. A nice

way to check everything is to do the UV matching, which is determining the parameters c_i^g that cancel the UV limits of the loops.

In some ways, galaxy renormalization is easier than DM renormalization. For DM, all renormalization is done through τ^{ij}, so we had to correctly track its contribution through the equations of motion. For galaxies, we just write down the EFT expansion directly at the level of the field (since there are no equations of motion).

3.5 Redshift space

The distribution of galaxies is homogeneous and isotropic in the coordinates determined by the Hubble flow $\vec{x}(a)$. However, since the redshift a that we observe is also affected by the motion of the galaxies, the distribution is anisotropic in the measured coordinate

$$\vec{x}_r(a_{\text{obs.}}(a)) \approx \vec{x}(a) + \frac{\hat{z}\cdot\vec{v}}{aH}\hat{z} \ , \tag{3.22}$$

where $\hat{z}$ is the line of sight direction (i.e. where the telescope is pointed).

Ultimately, what this means is that the distribution of galaxies that we measure is related to the one discussed above by a non-linear transformation depending on the velocity and the line of sight

$$\delta_{r,g}(\vec{k},\hat{z}) = \delta_g(\vec{k}) + \int d^3x\, e^{-i\vec{k}\cdot\vec{x}} \left(\exp\left[-i\frac{(\hat{z}\cdot\vec{k})}{aH}(\hat{z}\cdot\vec{v}(\vec{x})) \right] - 1 \right) (1+\delta_g(\vec{x})) \ . \tag{3.23}$$

This is basically just a coordinate transformation using Eq. (3.22). The main thing to notice is that because the density depends on $\hat{z}$ now, isotropy is broken, and correlations can depend on external wave vectors dotted with the line of sight. Expanding this out perturbatively, we have

$$\begin{aligned} \delta_{r,g} = \delta_g &- \frac{\hat{z}^i\hat{z}^j}{aH\bar{\rho}_g}\partial_i\pi_g^j + \frac{\hat{z}^i\hat{z}^j\hat{z}^k\hat{z}^l}{2(aH)^2\bar{\rho}_g}\partial_i\partial_j(\pi_g^k v^l) \\ &- \frac{\prod_{a=1}^6 \hat{z}^{i_a}}{3!(aH)^3\bar{\rho}_g}\partial_{i_1}\partial_{i_2}\partial_{i_3}(\pi_g^{i_4}v^{i_5}v^{i_6}) + \frac{\prod_{a=1}^8 \hat{z}^{i_a}}{4!(aH)^4\bar{\rho}_g}\partial_{i_1}\partial_{i_2}\partial_{i_3}\partial_{i_4}(\pi_g^{i_5}v^{i_6}v^{i_7}v^{i_8}) + \dots \ . \end{aligned} \tag{3.24}$$

This is now the new expression that we will compute loops with. This leads to new UV limits of the loops that are not captured by the renormalization that we have done so far. To renormalize this expression, we notice that the transformation involves products of fields at the same point, which we normally call contact operators in QFT. These need to be separately renormalized. So, we write

$$\begin{aligned} [\delta_g]_R &= \delta_g + \mathcal{O}_{\rho_g}/\bar{\rho}_g \ , \\ [\pi_g^i]_R &= \rho_g v^i + v^i\mathcal{O}_{\rho_g} + \mathcal{O}^i_{\pi_g} \ , \\ [\pi_g^i v^j]_R &= \rho_g v^i v^j + v^i v^j\mathcal{O}_{\rho_g} + v^i\mathcal{O}^j_{\pi_g} + v^j\mathcal{O}^i_{\pi_g} + \mathcal{O}^{ij}_{\pi_g v} \ , \\ [\pi_g^i v^j v^k]_R &= \rho_g v^i v^j v^k + v^i v^j v^k\mathcal{O}_{\rho_g} + (v^i v^j\mathcal{O}^k_{\pi_g} + 2 \text{ perms.}) + (v^i\mathcal{O}^{jk}_{\pi_g v} + 2 \text{ perms.}) + \mathcal{O}^{ijk}_{\pi_g v^2} \ , \\ [\pi_g^i v^j v^k v^l]_R &= \rho_g v^i v^j v^k v^l + v^i v^j v^k v^l\mathcal{O}_{\rho_g} + (v^i v^j v^k\mathcal{O}^l_{\pi_g} + 3 \text{ perms.}) \\ &\quad + (v^i v^j\mathcal{O}^{kl}_{\pi_g v} + 5 \text{ perms.}) + (v^i\mathcal{O}^{jkl}_{\pi_g v^2} + 3 \text{ perms.}) + \mathcal{O}^{ijkl}_{\pi_g v^3} \ , \end{aligned} \tag{3.25}$$

where all of the $\mathcal{O}^{ijk\cdots}$ are Galilean scalars, and we use these renormalized quantities in Eq. (3.24). We get this form by demanding that the renormalized operators satisfy the same Galilean transformation rules as the non-renormalized operators, and the form is important to insure overall Galilean invariance. The bottom line is that we expand each $\mathcal{O}^{ijk\cdots}$ in the same way that we did before, with free parameters in front of each independent term. Again, a good consistency check is doing the UV matching. All of this is described in detail in [16].

3.6 Loop integrals

$$B_{222}^{r,g} \sim 8 \int_{\vec{q}} P_{11}(q) P_{11}(|\vec{k}_2 - \vec{q}|) P_{11}(|\vec{k}_1 + \vec{q}|) \times \left(\frac{(\vec{k}_1 \cdot \hat{z})^2 (\vec{k}_2 \cdot \hat{z})^2 ((\vec{k}_1 + \vec{k}_2) \cdot \hat{z})^2 ((\vec{k}_2 - \vec{q}) \cdot \hat{z})^2 (\vec{q} \cdot \hat{z})^2 ((\vec{k}_1 + \vec{q}) \cdot \hat{z})^2}{q^4 |\vec{k}_1 + \vec{q}|^4 |\vec{k}_2 - \vec{q}|^4} + \dots \right) ,$$

Doing Expand in Mathematica, there are 23,000 terms in this diagram.

3.7 Non-Gaussianity

For a quick example of why we are doing any of this, let's look at inflation. We'll skip all of the details, but roughly, if there are interactions in the action for the inflaton π [19],

$$\mathcal{L}_{\text{inflation}} \sim \dot{\pi}^2 - c_s^2 (\partial_i \pi)^2 + f_{\text{NL}}^{(1)} \dot{\pi}^3 + f_{\text{NL}}^{(2)} \dot{\pi} (\partial_i \pi)^2 + \dots \tag{3.26}$$

then the statistics of π will be non-Gaussian, described by these parameters f_{NL}. The inflaton eventually sources the density perturbations that get imprinted on dark matter, such that the initial conditions for dark matter change

$$\tilde{\delta}_{\text{NG}}^{(1)}(\vec{k}) = \tilde{\delta}^{(1)}(\vec{k}) + f_{\text{NL}} \int_{\vec{k}_1, \vec{k}_2}^{\vec{k}} W(\vec{k}_1, \vec{k}_2) \frac{T(|\vec{k}_1 + \vec{k}_2|)}{T(k_1) T(k_2)} \tilde{\delta}^{(1)}(\vec{k}_1) \tilde{\delta}^{(1)}(\vec{k}_2) \ . \tag{3.27}$$

Here, the W function comes from the derivative coupling in Eq. (3.26), and the T function (called a transfer function) tells us how to convert from π to δ. The bottom line is that we now use this new $\tilde{\delta}_{\text{NG}}^{(1)}$ as the initial conditions in all of the expressions for the solutions that we showed above. It will give new possible shapes in the observables, but in order to be able to measure f_{NL}, we have to understand the gravitational part very well. That is what the EFT allows us to do.

References

[1] D. Baumann, A. Nicolis, L. Senatore and M. Zaldarriaga, *Cosmological Non-Linearities as an Effective Fluid*, *JCAP* **1207** (2012) 051, [1004.2488].

[2] J. J. M. Carrasco, M. P. Hertzberg and L. Senatore, *The Effective Field Theory of Cosmological Large Scale Structures*, *JHEP* **09** (2012) 082, [1206.2926].

[3] J. J. M. Carrasco, S. Foreman, D. Green and L. Senatore, *The Effective Field Theory of Large Scale Structures at Two Loops*, *JCAP* **1407** (2014) 057, [1310.0464].

[4] D. P. L. Bragança, M. Lewandowski, D. Sekera, L. Senatore and R. Sgier, *Baryonic effects in the Effective Field Theory of Large-Scale Structure and an analytic recipe for lensing in CMB-S4*, 2010.02929.

[5] M. Lewandowski, L. Senatore, F. Prada, C. Zhao and C.-H. Chuang, *EFT of large scale structures in redshift space*, *Phys. Rev.* **D97** (2018) 063526, [1512.06831].

[6] S. Dodelson, *Modern Cosmology*. Academic Press, 2003.

[7] M. H. Goroff, B. Grinstein, S. J. Rey and M. B. Wise, *Coupling of Modes of Cosmological Mass Density Fluctuations*, *Astrophys. J.* **311** (1986) 6–14.

[8] B. Jain and E. Bertschinger, *Second order power spectrum and nonlinear evolution at high redshift*, *Astrophys. J.* **431** (1994) 495, [astro-ph/9311070].

[9] F. Bernardeau, S. Colombi, E. Gaztanaga and R. Scoccimarro, *Large scale structure of the universe and cosmological perturbation theory*, *Phys. Rept.* **367** (2002) 1–248, [astro-ph/0112551].

[10] G. D'Amico, M. Marinucci, M. Pietroni and F. Vernizzi, *The large scale structure bootstrap: perturbation theory and bias expansion from symmetries*, *JCAP* **10** (2021) 069, [2109.09573].

[11] P. Coles, *Galaxy formation with a local bias*, *Mon. Not. Roy. Astron. Soc.* **262** (1993) 1065–1075.

[12] J. N. Fry and E. Gaztanaga, *Biasing and hierarchical statistics in large scale structure*, *Astrophys. J.* **413** (1993) 447–452, [astro-ph/9302009].

[13] V. Desjacques, D. Jeong and F. Schmidt, *Large-Scale Galaxy Bias*, *Phys. Rept.* **733** (2018) 1–193, [1611.09787].

[14] P. McDonald and A. Roy, *Clustering of dark matter tracers: generalizing bias for the coming era of precision LSS*, *JCAP* **0908** (2009) 020, [0902.0991].

[15] L. Senatore, *Bias in the Effective Field Theory of Large Scale Structures*, *JCAP* **1511** (2015) 007, [1406.7843].

[16] G. D'Amico, Y. Donath, M. Lewandowski, L. Senatore and P. Zhang, *The one-loop bispectrum of galaxies in redshift space from the Effective Field Theory of Large-Scale Structure*, 2211.17130.

[17] Y. Donath, M. Lewandowski and L. Senatore, *Direct signatures of the formation time of galaxies*, 2307.11409.

[18] P. McDonald, *Clustering of dark matter tracers: Renormalizing the bias parameters*, *Phys. Rev. D* **74** (2006) 103512, [astro-ph/0609413].

[19] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan and L. Senatore, *The Effective Field Theory of Inflation*, *JHEP* **03** (2008) 014, [0709.0293].