Field Theory in Cosmology (Part III)

Enrico Pajer

Contents

Overview

The Field Theory in Cosmology course will cover three related topics: quantum field theory and inflation, statistical field theory and large scale structures, classical field theory and the cosmic microwave background. These notes cover the first part. Notes for the second part will be distributed by T. Baldauf. Notes for the third part will be made available on Moodle.

Estimated list of lectures for the first part:

1. Review: big picture, motivations and slow-roll inflation,
2. $P(X,\phi)$, massless scalar;
3. massive scalar; [particle creation]; graviton; amplitudes
4. correlators, in-in, example ϕ^3
5. $P(X,\phi)$ at quadratic and cubit order; Gravity as an Effective Field Theory
6. ADM formalism and constraints. Decoupling limit.
7. SVT decomposition. Gauge transformations. $\mathcal{R}$ correlators.
8. Adiabatic modes
9. Soft theorems

Notation, units and conventions I use units in which $\hbar = c = k_b = 1$. Therefore energy is temperature and inverse time or inverse length. On the other hand, I will try to keep the reduced Planck mass explicit, $M_{\rm Pl} = (8\pi G_N)^{-1/2}$. Beware that some authors use $M_{\rm Pl}$ to indicate the "full" Planck mass $G_N^{-1/2} \simeq 1.2 \times 10^{19}$GeV. The necessary conversion factors can be added using dimensional analysis and

$$c = 3\times 10^8 \frac{\rm m}{\rm sec}\,, \quad {\rm pc} = 3.2\,{\rm lightyears}\,, \quad {\rm year} = \pi \times 10^7\,{\rm sec}\,, \tag{0.1}$$

$$\hbar c = 0.2\,{\rm eV}\,\mu{\rm m}\,, \quad M_{\rm Pl} \simeq 2.4 \times 10^{18}{\rm GeV}\,. \tag{0.2}$$

I use the mostly plus signature $(-,+,+,+)$. Latin indices indicate space, $i,j,\dots = \{1,2,3\}$, while greek indices run over spacetime, $\mu,\nu,\dots = \{0,1,2,3\}$. 3D vectors are in boldface, e.g. $\mathbf{k}$ and $\mathbf{x}$. Unless otherwise specified, all tensors are expressed in terms of the FLRW coordinates

$$ds^2 = -dt^2 + a^2 dx^2\,. \tag{0.3}$$

. Standard derivatives are represented with a comme and covariant derivatives with a semi-column

$$T^{\cdots}_{\cdots,\mu} \equiv \partial_\mu T^{\cdots}_{\cdots}\,, \quad T^{\cdots}_{\cdots;\mu} \equiv \nabla_\mu T^{\cdots}_{\cdots}\,. \tag{0.4}$$

Symmetrization and anti-symmetrization of a pair of indices is indicated with $(\dots)$ and $[\dots]$ respectively and is defined to have weight 1

$$A_{(\mu\nu)} \equiv \frac{1}{2}\left(A_{\mu\nu} + A_{\nu\mu}\right)\,, \quad A_{[\mu\nu]} \equiv \frac{1}{2}\left(A_{\mu\nu} - A_{\nu\mu}\right)\,. \tag{0.5}$$

My convention for the Fourier transform are

$$F(\mathbf{x}) = \int_{\mathbf{k}} F(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad F(\mathbf{k}) = \int_{\mathbf{x}} F(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}}\,, \tag{0.6}$$

with

$$\int_{\mathbf{k}} \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3}\,, \quad \int_{\mathbf{x}} \equiv \int d^3\mathbf{x}\,. \tag{0.7}$$

I sometimes use the shorthand notations: D for Dodelson's Modern Cosmology book [16], W for Weinberg's Cosmology book [49]. For example D 3 is Chapter 3 of Donelson's book, while W appB is appendix B of Weinberg's book.

Introduction and Motivations

Field theory has proven to be an extremely useful language to describe nature. We are very familiar with the role fields play at the classical level from thinking about the gravitational field in Newton's theory. In Maxwell's description of electromagnetic phenomena, we learn that the electric and magnetic fields are not just a crutch to compute the forces acting on changed particles. The electromagnetic field is an entity on its own: it carries energy and momentum, it can be set up in a variety of initial conditions and it evolves according to precise laws.

At the classical level, fields are often contrasted with particles: fields permeate space while particles are localized; fields are continuous functions while particles are discrete. This distinction between particles and field already becomes blurred when we enter the realm of statistical physics. Even though a given system might be in a specific microstate, where all particles occupy precise positions, if we can only observe macroscopic properties of the system this point of view is useless. We are then forced to ascribe to the system a macroscopic state, i.e. an ensemble of many distinct microscopic configurations. In a macroscopic state, particles don't occupy precise positions anymore. Rather they can be characterized by a probability of being here or there. So, in statistical physics, the concept of a field emerges again. For example, density, temperature or pressure are statistical fields that describe macroscopic properties of a system or properties of an average particle.

In quantum mechanics, all particles have an associated "probability" wave (actually an amplitude, with $|\text{amplitude}|^2 \sim$ probability) and fields emerge again as the right language to describe the dynamics. Indeed, it is not a coincidence that the very idea of quantization emerged form the study of statistical systems, as in Planck's solution of the ultraviolet catastrophe. The particle-wave duality completely shatters the distinction between particles and fields. Finally, when we face the daunting task of marrying quantum mechanics with special relativity and its demands of causality and locality, we are forced once again to resort to fields. *Quantum field theory* (QFT) is the framework within which quantum mechanics can describe interactions that respect the observed locality of natural phenomena. Fields in QFT are not probability waves nor amplitudes. Fields are the tool by which locality and oftentimes Lorentz invariance is imposed onto the Hilbert space of quantum mechanics.

In cosmology, namely the study of the evolution of the universe, we encounter fields playing all the roles discussed above. In trying to describe the first fraction of a second of the big bang, we are forced to use QFT because both relativity and quantum mechanics induce large deviations from a classical behavior. As time goes on and the universe expands, the quantum perturbations generated during the primordial universe become classical in the sense that all practically measurable observables commute with each other to a large degree of accuracy. But our theory is still not deterministic because the initial conditions are only provided now as classical probabilities. We then need *statistical field theory* to describe observations. This is particularly important in the description of the formation of structures in the universe, where our inability to predict initial conditions in a deterministic way changes the dynamics qualitatively. Finally, in the quasi-linear regime that is relevant to describe how light moves in the universe and creates the Cosmic Microwave Background, we revert back to classical, deterministic field theory, with statistical effects playing only a marginal role.

Field theory and cosmology is therefore a match made in heaven. For those interested in cosmology and its phenomenology, this course will attempt to provide a solid theoretical foundation and a field theory toolkit to tackle the hardest problems. For those interested in field theory, this course will provide a point of view complementary to that of courses that focus on particle physics (such as QFT and advanced QFT in Part III). In the expanding spacetime provided by cosmology, fields behave differently from we observe in the flat spacetime of particle physics. New phenomena take place such as the ambiguity to define a vacuum state and particle creation; the fundamental observables change too: we set aside scattering amplitudes and focus on cosmological correlators.

This course discusses applications of classical, statistical and quantum field theory to cosmology. The course comprises of three interconnected topics:

- Cosmological inflation and primordial quantum perturbations (QFT in curved spacetime)
- The matter and galaxy distribution in the Large Scale Structure of the Universe (statistical field theory)
- The physics of the Cosmic Microwave Background (classical and statistical field theory)

The goals of the course are: to discuss open problems in cosmology and describe their intimate relation to fundamental high energy physics; to provide the basic knowledge

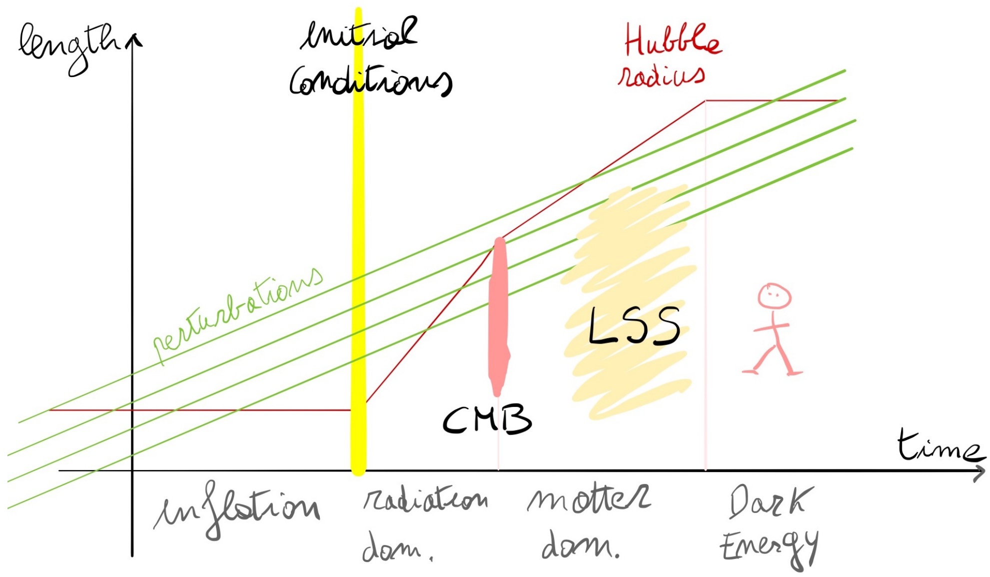

Figure 1: The big picture of cosmology. Inflation sets up the initial condition for perturbations which subsequently re-enter the Hubble radius and source perturbations in the Cosmic Microwave Background (CMB) and Large Scale Structures (LSS).

to understand modern research literature in cosmology; to explore how field theory provides a unifying formalism to describe disparate physical processes from the birth of the Universe to the highly non-linear cosmic web.

The course is aimed at students that have already had a first course in quantum field theory (QFT) and general relativity (GR). Some knowledge of cosmology is useful, but not necessary, as the relevant material will be reviewed when needed.

1 A quick review of background cosmology ref

In this section, I start by reviewing some relevant facts about cosmological backgrounds. Later on, this will simplify our task of describing quantum fields on these backgrounds and will allow us to make contact with observations. Most of the material in this section is also covered in any introductory cosmology course and in particular in Cosmology Part III.

1.1 Classical cosmological backgrounds

In general relativity (GR), the spacetime metric $g_{\mu\nu}(\mathbf{x},t)$ is a dynamical degree of freedom that obeys a set of ten, coupled, second order, non-linear partial differential equations known as Einstein Equations (EE's)

$$R_{\mu\nu}-\frac{1}{2}g_{\mu\nu}R=\frac{1}{M_{\rm Pl}^2}T_{\mu\nu}\,. \tag{1.1}$$

A few definitions are in order. $M_{\rm Pl}\equiv(8\pi G_N)^{-1/2}$ is the reduced Planck mass (in units such that $\hbar=c=k_B=1$), $R_{\mu\nu}$ is the Ricci tensor given in terms of the Christoffel symbols by[1]

$$R_{\mu\nu} \equiv 2\Gamma^{\rho}_{\mu[\nu,\rho]} + 2\Gamma^{\rho}_{\lambda[\rho}\Gamma^{\Lambda}_{\beta]\alpha}\,, \tag{1.2}$$

$$\Gamma^{\mu}_{\alpha\beta} \equiv \frac{1}{2}g^{\mu\gamma}\left(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}\right)\,. \tag{1.3}$$

Finally $T_{\mu\nu}$ is the energy momentum tensor, which can be extracted from a given action S by

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g_{\mu\nu}}\,, \tag{1.4}$$

Notice that it is symmetric by definition. It is painfully clear from (1.1) that exact solutions of GR are in general hard to come by. Nevertheless many highly symmetric solutions are relatively easy to find. The simplest possible solution of the Einstein Equations is Minkowski spacetime,

$$ds^2 = -dt^2 + dx^i\delta_{ij}dx^j\,, \tag{1.5}$$

which requires $T_{\mu\nu} = 0$. This solution of GR is closest to our intuition of space and time. It is also *maximally symmetric*, i.e. it possesses the largest amount of symmetry: in $(3+1)$-dimensions this is the ten isometries that combine to form the Poicaré group ($ISO(3,1)$): three rotations and three boosts, forming the Lorentz group ($SO(3,1)$) plus one time and three spatial translations. Minkowski is so symmetric that any point can be related to any other point by a symmetry transformation. There are actually two other spacetimes that have this property, *de Sitter* (dS) and *Anti-de Sitter* (AdS) spacetime, and they will be discussed at the end of this section.

FLRW spacetime In a maximally symmetric universe such as Minkowski, there cannot be any beginning or end of time; there cannot even be a history because every time is equivalent to any other time. While such an eternal universe might be appealing from a philosophical or aesthetical point of view, it is in contradiction with the last century of cosmological observations. In particular, the observation of the expansion of the universe and of the Cosmic Microwave Background radiation show that the universe in the past was much denser and hoter than it is now. In such a universe time translations should not be a symmetry, or more precisely there should be no time-like Killing vectors.

The simplest possibility is the Friedmann-Lemaitre-Robertson-Walken (FLRW) metric

$$ds^2 = -dt^2 + a(t)^2\frac{dx^i dx^j\delta_{ij}}{(1+K\mathbf{x}^2/4)^2} \tag{1.6}$$

where the metric is written in "quasi-cartesian" coordinates, the parameter K is known as the *spatial curvature* and $a(t)$ is the *scale factor*, whose dynamic will be dictated by the EE's. If we so wish, by rescaling x and a we can always normalize curvature to one of three values: $K=-1$ for an open universe, $K=0$ for a flat universe and $K=1$ for a closed universe. The spatial coordinates x_i are called *comoving coordinates.* Notice that the physical distance between two points at fixed comoving coordinates changes with time because of the scale factor $a(t)$. The time t corresponds to the *proper*

[1]Indices right after a comma indicate partial derivative, as in $g_{\mu\nu,\rho} = \partial_\rho g_{\mu\nu}$, while indices right after a semi-colon indicate covariant derivatives, as in $A_{\mu;\nu} = \nabla_\nu A_\mu$.

time of observers that are at rest with respect to the comoving coordinates. It is often convenient to use a different time coordinate, $a d\tau = dt$, known as *conformal time*, so that

$$ds^2 = a^2(\tau)\left[-d\tau^2 + \frac{dx^i dx^j \delta_{ij}}{(1+K\mathbf{x}^2/4)^2}\right] . \tag{1.7}$$

This FLRW metric has six isometries, which locally can be thought of as three space translations and three rotations. It describes a *homogeneous and isotropic* universe that looks the same at every point in space in every direction. This metric is simple enough that it provides a class of exact, fully non-linear solutions of EE's. A most remarkable fact that should blow your mind is that this most simple spacetime for $K = 0$ is in fact a very good description of our own universe on distances much larger than average distance between galaxies, about a few Megaparsec[2] (Mpc). There are small deviations from perfect homogeneity in our universe: you, me, the galaxy we live in and the other billion of galaxies out there. We will come to discuss those in the second part of this course.

As compared with Minkowski spacetime, in an FLRW spacetime time translation and the three Lorentz boosts fail to be isometries because of the time dependence of $a(t)$, which is conveniently capture by the so-called *Hubble parameter*

$$H(t) \equiv \frac{\dot{a}(t)}{a(t)} . \tag{1.8}$$

All new surprising cosmological phenomena that you have not encountered in particle physics, will have an H appearing somewhere. The absence of time translations has profound implications for constructing a QFT. First, energy is not conserved and particles can be created or destroyed as the universe expands. We believe that the structures we observe in the universe were created precisely by this process. Second, unlike in Minkowski, there is in general no unique choice for a vacuum state. It will be only under certain specific conditions and assumptions that we will be able to choose a preferred vacuum. Third, the expansion of space changes the energy of a given state hence mixing different scales, which requires care when discussing Effective Field Theories (EFT's).

Continuity equation To have a chance to solve the EE's we need to specify an energy momentum tensor that has the same symmetries as the spacetime metric. The most generic $T_{\mu\nu}$ that is invariant under rotations and translations is[3]

$$T^{\mu}{}_{\nu} = \text{Diag}\left\{-\rho, p, p, p\right\} , \tag{1.10}$$

where the *energy density* ρ (with units $[\rho] = E/L^3$) and the *pressure* p (with units $[p] = M/(T^2 L) = E/L^3$) are only functions of time $\rho = \rho(t)$, $p = p(t)$. We can interpret this as the energy-momentum tensor of a homogeneous perfect fluid in its rest frame

$$T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + g_{\mu\nu} p , \tag{1.11}$$

[2]One Megaparsec, or 10^6 parsec is a useful unit of distance in cosmology. One parsec is the distance at which one astronomical unit (Au), the average distance Earth-Sun, subtends and angle of one arcsecond, which is 1/3600 of a degree. A parcec is about 3.26 lightyears.

[3]The indices are raised and lowered with the FLRW metric. So for example, for $K = 0$

$$T_{\mu\nu} = \text{Diag}\left\{\rho, a^2 p, a^2 p, a^2 p\right\} . \tag{1.9}$$

where u_μ is the normalized fluid velocity ($u_\mu u^\mu = -1$), which in the rest frame would be $u_\mu = \{1, 0, 0, 0\}$. EE's imply that the energy-momentum tensor is covariantly conserved, namely

$$T^{\mu\nu}{}_{;\nu} = \partial_\nu T^{\mu\nu} + \Gamma^\mu_{\alpha\nu} T^{\alpha\nu} + \Gamma^\nu_{\alpha\nu} T^{\mu\alpha} = 0\,. \tag{1.12}$$

Of these four equations, the only non-vanishing one is $\nu = 0$ (a consequence of rotation invariance). Using the FLRW metric to compute the Christ-awful symbols and (1.10) one finds the *continuity equation*

$$\boxed{\dot\rho + 3H\,(\rho + p) = 0}\,. \tag{1.13}$$

In words, this tells us that the energy density changes in time only if the universe expands or contracts, $H \neq 0$. The EE's will not tell us what type of matter permeates the universe. To specify that we need to specify an *equation of state*, namely how the pressure is related to the density and possibly other thermodynamical variables. Most systems of interest in cosmology can be described to good approximation by a simple, linear equation of state

$$p = w\rho\,, \tag{1.14}$$

where the constant w is the equation-of-state parameter. For these linear equations of state, it is easy to solve the continuity equation for any give expansion history $a(t)$:

$$\dot\rho + 3\frac{\dot a}{a}(1+w)\rho = 0 \qquad \Rightarrow \qquad \rho(t) = \rho(t_0)\left[\frac{a(t)}{a(t_0)}\right]^{-3(1+w)}\,. \tag{1.15}$$

For example:

- For non-relativistic matter, also known as *dust*, the velocity is much smaller than the speed of light, $v \ll c$ and so the energy $E = m\sqrt{c^2 + v^2} \simeq mc$ is much larger than the pressure $E \sim mc \gg mv$. Then the pressure, which is a measure of the average momentum of particles is negligible compared with the energy density, which is proportional to the mass density, $p \ll \rho$ or $w \ll 1$. In an expanding universe dust dilutes as $\rho \propto a^{-3}$.

- For relativistic matter, also known as *radiation*, momentum and energy are equal and therefore the pressure is similar to the energy density. From statistical mechanics we find out that the precise proportionality constant is[4] $p = \rho/3$ and so $w = 1/3$. In an expanding universe radiation dilutes as $\rho \propto a^{-4}$, which is faster than dust.

- For a *cosmological constant*, also known as vacuum energy, $T_{\mu\nu} = -\Lambda g_{\mu\nu}$ and therefore $p = -\rho = -\Lambda$ or $w = -1$. Since now $\rho + p = 0$, from (1.13) we learn that a cosmological constant does not dilute as the universe expands, $\rho \propto a^0 \sim$ *const*. We could have expected this from its name.

A simple interpretation of the above scalings is that of an expanding box of linear size $a(t)$. Non-relativistic matter density dilutes with the volume, i.e. a^{-3}. Relativistic matter, a.k.a. radiation, also dilutes with the volume as a^{-3}, but it has an extra a^{-1} suppression due to the redshift of the momentum of each particle (and the mass is negligible).

[4] A good mnemonic for this is to recall that the theory of electromagnetism does not have any scale and is therefore conformal invariant. This in turns demands that the energy-momentum tensor is traceless, $T^\mu_\mu = 0$, from which $w = 1/3$ follows immediately.

Friedmann equation Let solve the EE's for an FLRW metric. Using the definition of the Ricci tensors in (1.2), and the FLRW metric (1.6), a lengthy but straightforward computation shows

$$R^0{}_0 = 3\frac{\ddot{a}}{a}\,, \quad R^i{}_j = \delta_{ij}\,\frac{2K + 2\dot{a}^2 + a\ddot{a}}{a^2}\,, \quad R = 6\left[\frac{\ddot{a}}{a} + \left(\frac{\dot{a}}{a}\right)^2 + \frac{K}{a^2}\right]. \tag{1.16}$$

The 00-component of the EE's (most conveniently computed with one upper and one lower index) in (1.1) is then easily derived

$$\boxed{3M_{\rm Pl}^2\left(H^2 + \frac{K}{a^2}\right) = \sum_a \rho_a\,,} \tag{1.17}$$

where i runs over all constituents of the universe (for example radiation, Dark Matter, neutrinos and baryons). This is the *Friedmann equation.* Notice that since an FLRW metric has only one free function $a(t)$, we need only one of the ten EE's. It is sometimes convenient to make the Friedman equation dimensionless by dividing it by the *critical density* (a function of time)

$$\rho_c \equiv 3M_{\rm Pl}^2 H^2\,. \tag{1.18}$$

This leads to

$$1 - \Omega_k = \sum_a \Omega_a\,, \quad \text{with} \quad \Omega_k \equiv -\frac{K}{H^2 a^2}\,, \quad \Omega_a \equiv \frac{\rho_a}{\rho_c}\,. \tag{1.19}$$

The Ω's are called *fractional energy densities* and are manifestly dimensionless. From this form of the equation we see that curvature parameter K tells us whether the energy density of the constituents of the universe is smaller or larger than the critical one.

Single-component flat universe To develop some intuition let's focus on a simple universe that has only one type of stuff, i.e. with a single component, and zero curvature $K = 0$. Using a to parameterize time we can solve the Friedman equation as follows:

$$H = \frac{\dot{a}}{a} = \sqrt{\frac{\rho}{3M_{\rm Pl}^2}} = H_0\left(\frac{a_0}{a}\right)^{3(1+w)/2} \Rightarrow a(t) = \left[\frac{3}{2}\left(1+w\right)H_0 t\right]^{\frac{2}{3(1+w)}}\,, \tag{1.20}$$

where $w \neq 1$ was assumed and I fixed the integration constant requiring that a vanishes at past infinity.

Important solutions for the scale factor are then

- For non-relativistic matter, or *dust*, $w \simeq 0$ so $a \propto t^{2/3}$.
- For relativistic matter, or *radiation*, $w = 1/3$ so $a \propto t^{1/2}$.
- For a cosmological constant, or *vacuum energy*, $w = -1$ this expressions is singular. Solving this particular case separately one finds $a \propto e^{H_0 t}$.

Notice that if a is a monomial in t one finds always $H \propto t^{-1}$, or more precisely

$$H(t) = \frac{2}{3(1+w)}\,\frac{1}{t}\,. \tag{1.22}$$

Box 1.1 Null Energy Condition (NEC) A certain form of matter with energy-momentum tensor $T_{\mu\nu}$ satisfies the Null Energy Condition iff for ever null vector $N^\mu N_\mu = 0$ one has

$$T_{\mu\nu} N^\mu N^\nu \geq 0 \quad \text{(NEC)} \,. \tag{1.21}$$

For an FLRW universe we can choose $N^\mu = \{1, -1, 0, 0\}$ and find $\rho + p \geq 0$. Violations of the NEC are often associated with pathologies such as ghosts instabilities (i.e. field with the wrong-sing kinetic term that can be created at will by decreasing the energy of the system) or tachyon instabilities [18]. Yet, more exotic theories with non-standard kinetic terms, such as the ghost condensate, are known to safely violate the NEC, see e.g. [11,38].

This gives the *age of the universe* for a single-component universe

$$t_{age} = \frac{2}{3(1+w)} \frac{1}{H(t_{age})} \,. \tag{1.23}$$

There are two other combinations of EE's that come in handy. First, subtracting the 00 EE from the (summed) *ii* EE's, one finds the *acceleration equation*

$$M_{\rm Pl}^2 \frac{\ddot{a}}{a} = -\frac{1}{6} \left(\rho + 3p\right) \,. \tag{1.24}$$

Second, by taking the time derivative of the Friedmann equation and using the continuity equation to get rid of $\dot{\rho}$, we can find the variation of the Hubble parameter

$$-\dot{H} M_{\rm Pl}^2 = \frac{1}{2} \left(\rho + p\right) \,. \tag{1.25}$$

Most cosmological "stuff" obeys the so-called Null Energy Condition, namely $\rho + p \geq 0$ (see Box 1), and so H decreases during the expansion of the universe.

1.2 Motivations for Inflation

In this section[5], I recall several problems with any cosmological model in which the universe undergoes decelerated expansion[6] all the way until the Big Bang, as it is for example the case for radiation or matter domination. I will refer to this class of models collectively as "Hot Big Bang" model, where "hot" refers to the temperature of radiation. First, I discuss two old "background" problems, namely the horizon and curvature problems, which can be stated already for the unperturbed FLRW universe. These problems were originally formulated in the 80's and have not changed much since. Second, I mention two new "perturbation" problems, namely scale invariance and phase-coherence problems, which have to do with the large amount of new data we have collected in the past 30 years, especially from the Cosmic Microwave Background (CMB).

[5] All details in this Section 1.2 are non-examinable.

[6] In the current standard cosmological model known as ΛCDM, an accelerated expansion is induced by the cosmological constant Λ at late times, $z \simeq 0.5$. At any earlier time the expansion is decelerated and so these problems also affect ΛCDM.

The curvature problem The curvature problem is the fact that we do not observe any spatial curvature in our universe, $K \simeq 0$, despite the fact that curvature dilutes more slowly than radiation and matter and so grows with time relatively to them. Let us see this in formulae.

Current bounds tell us that [1]

$$\Omega_K \equiv \left(\frac{K}{a^2 H^2}\right) , \quad \Omega_{K,0} = 0.000 \pm 0.005 \,. \tag{1.26}$$

On the other hand, as we saw in around (1.6), the most general homogeneous and isotropic space can have spatial curvature, i.e. $K \neq 0$. From Eq. (1.26) we see that Ω_K grows with time in an decelerated ($\ddot{a} < 0$) expanding ($\dot{a} > 0$) universe

$$\dot{\Omega}_K = -\ddot{a}\frac{2K}{\dot{a}^3} \propto -\ddot{a} \propto (\rho + 3p) \propto (1 + 3w) \,, \tag{1.27}$$

where in the second step I used the acceleration equation (1.24) to show that in an expanding universe ($\dot{a} > 0$) the fact that $\rho + 3p > 0$ implies deceleration (this is also known as the Strong Energy Condition (SEC)). Since at early times in ΛCDM the universe is dominated by radiation, $w = 1/3$, we conclude that Ω_K must have been even smaller in the past. In other words, extrapolating closer and closer to the Big Bang singularity at $a \to$ and $\rho \to \infty$, we are forced to assume that the initial curvature was tiny, $\Omega_K(a_i) \to 0$, or equivalently the initial total density of the universe was extremely close to the critical one, $\sum_i \rho_i \to \rho_c$ (defined in 1.18). There are only two logical possibilities:

1. The initial conditions of the universe, as it emerged from some yet unknown non-perturbative theory of quantum gravity[7], were extremely fined tuned close to $\Omega_K = 0$. In this scenario, the existence of our universe is a very rare fluctuation, since any larger initial value of $\Omega_K(t_i)$ would have grown to dominate the energy density and would have prevented the formation of galaxies and therefore of life as we know it. Not a great option, in the opinion of many.

2. The early expansion history of our universe is modified to stop Ω_K from growing as we move back in time. From (1.27) we see that this requires either $\ddot{a}, \dot{a} < 0$, i.e. an early phase of decelerated contraction, or $\ddot{a}, \dot{a} > 0$, i.e. an early phase of accelerated expansion. Since we know the current universe is expanding (recall Hubble's law), the first of these options requires to *bounce* i.e. to transition from $\dot{a} \propto H < 0$ to $\dot{a} \propto H > 0$. Achieving the bounce in a controlled construction is still an open problem and the many proposed models have a series of pathologies. Therefore we will henceforth assume an early phase of accelerated expansion, a.k.a. *inflation*.

Summarizing, to explain the spatial flatness of the observed universe we postulate the existence of a primordial phase of accelerated expansion, $\ddot{a}, \dot{a} > 0$, called inflation.

[7]Strictly within GR, K is just a parameter, not a dynamical variable, and so there in no physical perturbation that can make $\Omega_K = 0$ unstable. On the other hand, GR is most likely just a low-energy (subPlanckian) effective description of some UV-complete theory of quantum gravity, and it is at least plausible that $\Omega_K = 0$ might be unstable within that larger, yet unknown theory. Perhaps a more concrete example is bubble nulceation. instanton solutions are known in which a new universe nucleates from a single point []. To respect the isometries of the system the new universe must have some negative curvature. It is not known whether bubble nucleation and the ensuing ideas about the multiverse play a role in the history of our own universe, and the discussion among experts continues.

Horizon problem In general, we would like to have

$$\left(\begin{array}{c}\text{Distance between regions}\\ \text{of space that look similar}\end{array}\right) \ll \left(\begin{array}{c}\text{Distance travelled by light}\\ \text{since the beginning of time}\end{array}\right), \tag{1.28}$$

so that we can explain why two regions look similar in a way that is compatible with causality. However, in any Hot Big Bang model this inequality is dramatically violated. More precisely, cosmological observations of far away objects allow us to see regions in the past that are separated by much more that than the particle horizon at the time, which is the furthest a signal can travel. Any mechanism attempting to explain homogeneity across these regions then necessarily violates causality and/or locality, leading to the horizon problem.

To see this quantitatively, recall that the *comoving distance* between two generic times t_1 and t_2 with $a_1 = a(t_1) < a(t_2) = a_2$ is found to be

$$\chi(a_1, a_2) \equiv \int_{a_1}^{a_2} \frac{da}{a^2 H} = \frac{1}{a_1 H_1} \frac{2}{3w+1} \left[\left(\frac{a_2}{a_1} \right)^{(3w+1)/2} - 1 \right], \tag{1.29}$$

where I assumed $w \neq -1/3$. Then, the distance of an object at redshift $1 + z = a^{-1}$ from us at $a = a_0 = 1$ is given by

$$\chi(a, 1) \equiv \int_{a}^{a_0} \frac{d \log a}{aH} = \frac{1}{H_0} \frac{2}{3w+1} \left[1 - a^{(3w+1)/2} \right], \tag{1.30}$$

Imagine now to look out in the night sky in opposite directions and detect a pair of antipodal objects, each sending us radiation with the same redshift z. The relative comoving distance $\Delta\chi$ between the objects is just $2\chi(a, 1)$. To simplify the algebra, let us neglect Dark Energy and so $w > -1/3$ and assume $a \ll 1$. Then

$$\Delta\chi(a, 1) \simeq 2 \times \frac{1}{H_0} \frac{2}{3w+1} \simeq \frac{\mathcal{O}(1)}{H_0}, \tag{1.31}$$

Recall that the redshift of these objects is $1 + z = 1/a$, and so we conclude that high redshift objects $z \gg 1$ are at a distance of order the Hubble radius today H_0^{-1}, almost independently of z. Since this is a comoving distance between objects at fixed comoving position (i.e. far away object are in the Hubble flow), it does not depend on time.

Let us compare now this distance with the *comoving particle horizon* in a Hot Big Bang model, i.e. extrapolating radiation domination all the way to $a_i = 0$. Recall that the comoving particle horizon $x_{\text{p.h.}}$ is the comoving distance traveled by light since the beginning of time τ_i, namely $x_{\text{p.h.}}(a) \equiv \chi(a_i, a)$. Recall also that for $w > -1/3$, or equivalently decelerated expansion $\ddot{a} < 0$ (as it is the case for radiation and dust), one can safely take $a_i \to 0$ and so $x_{\text{p.h.}}(a)$ equals the comoving Hubble radius[8] times an order one number

$$x_{\text{p.h.}}(a) = \frac{1}{aH} \frac{2}{3w+1} \simeq \frac{1}{aH} \mathcal{O}(1) \simeq r_H(a) \mathcal{O}(1) \quad \text{(decelerated)}. \tag{1.32}$$

[8]In the literature, r_H is often referred to as *Hubble "horizon"*. This is a misnomer since neither $(aH)^{-1}$ nor its physical cousin H^{-1} represent a horizon in the usual sense of GR. This nomenclature is widely spread and not harmful as long as one is aware of the subtleties.

Assuming decelerated expansion since the Big Bang, and combining (1.32) with (1.31) one finds

$$\frac{\Delta\chi(a)}{x_{\text{p.h.}}(a)} \simeq 2\,\frac{aH}{a_0 H_0} \simeq 2\left(\frac{1}{a}\right)^{(3w+1)/2} \gg 1 \quad \text{(decelerated)}\,. \tag{1.33}$$

This means that, in an ever decelerating universe, by observing far away objects ($1/a = 1+z \gg 1$) we are actually probing scales much larger than the particle horizon at that time. In practice, one can reach $a = (1+z)^{-1} \sim 0.1$ with quasar and $a \sim z^{-1} \sim 10^{-3}$ with Cosmic Microwave Background (CMB) photons. In both cases, the observed physical properties (e.g. density of quasars, temperature and polarization of the CMB) are the same in average in all directions. These observables said to be "statistically isotropic". We conclude that, in the absence of accelerated expansion in our past, the mechanism responsible for this observed statistical isotropy must violate causality. This is the *particle horizon problem.*

Conversely, for a phase of accelerated expansion, $\ddot{a} > 0$ or $w < -1/3$ (such as during Dark Energy or inflation) during a period $a \in \{a_i, a_f\}$, the result is divergent as $a_i \to 0$:

$$x_{\text{p.h.}}(a_f) = \frac{1}{a_f H_f}\frac{2}{|3w+1|}\left[\left(\frac{a_f}{a_i}\right)^{|3w+1|/2} - 1\right] \tag{1.34}$$

$$\simeq \frac{1}{a_f H_f}\frac{2}{|3w+1|}\left(\frac{a_f}{a_i}\right)^{|3w+1|/2} \gg r_H \quad \text{(accelerated)}\,. \tag{1.35}$$

In the extreme case $w \simeq -1$ (inflation), H is approximately constant and $x_{\text{p.h.}}$ asymptotes to

$$x_{\text{p.h.}} \to \frac{1}{a_i H_i} \quad \text{(inflation)}\,. \tag{1.36}$$

Combining this with (1.31) we see that we can make $\Delta\chi(a)/x_{\text{p.h.}}(a)$ as small as we want by taking a_i sufficiently small. This allows to try and find an explanation for the observed statistical isotropy that respects causality. Yet again, we are lead to postulate a phase of accelerated expansion $\ddot{a}, \dot{a} > 0$ in the early universe.

Phase coherence problem Our universe has perturbations on all observed scales. A remarkable fact is that these perturbations are observed to oscillate in exact synchronicity. This occurs even on very large scales. In any Hot Big Bang model, distant regions oscillate with precisely the same phase even thought they lie outside of each other's particle horizon. This is the *phase coherence* of cosmological perturbations. This is a problem because on these super-horizon scales no causal mechanism can be devised to "synchronize" the phases and so their coherence becomes a very unlikely coincidence.

This strongly suggests that there was a primordial phase, before the hot Big Bang, during which perturbations were generated and synchronized. For more detail see e.g. my lecture notes [32]. This is a crucial point. It tells us that the seeds of the distribution of everything we see in the night sky today were actually sown during the first fraction of a second of the Big Bang. Cosmological observations are then a probe of *primordial perturbations* and we can use them to learn about the laws of physics during this primordial time.

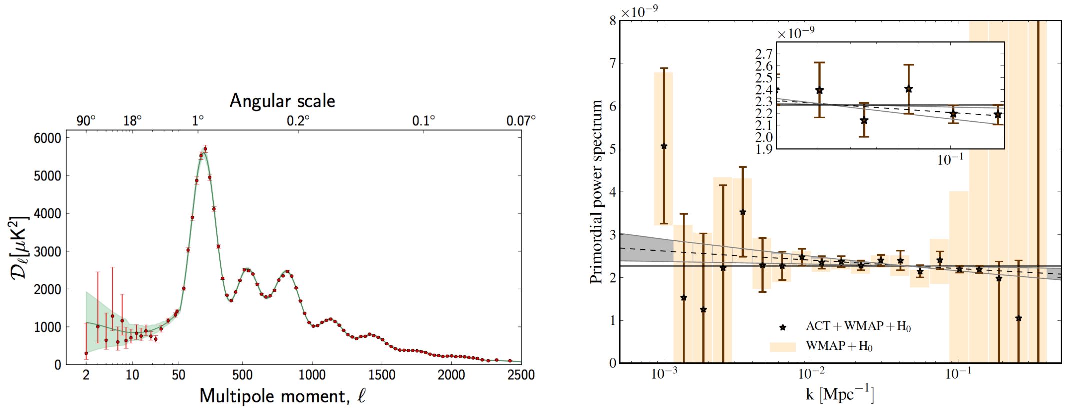

Figure 2: The left plot shows the CMB temperature anisotropy power spectrum as function of multipoles l or equivalently inverse angular scale. For $l \lesssim 50$ the angular power spectrum approximately coincides with the power spectrum of the primordial initial conditions, showing that it is indeed approximately scale invariant, i.e. constant as function of l. The right plot shows the power spectrum of the initial condition obtained by the same measurements as on the left hand side but evolved back in time using the ΛCDM evolution equations. The primordial power spectrum is scale invariant up to a 4% "red" tilt.

Scale invariance problem The last problem with the Hot Big Bang is the surprising fact that the amplitude of perturbations observed in our universe is approximately the same (within 4%) on all cosmological scales (about 3 orders of magnitudes, between 10^{-4} and 10^{-1}Mpc^{-1}). This remarkable feature of what we can now call primordial perturbations goes under the name of (approximate) *scale invariance*. The mathematical statement is that for every $\lambda \in \mathbb{R}$ and $n \in \mathbb{N}^+$, a field ϕ obeys scale invariance iff

$$\langle \phi(\mathbf{x}_1)\phi(\mathbf{x}_2)\dots\phi(\mathbf{x}_n)\rangle = \langle \phi(\lambda\mathbf{x}_1)\phi(\lambda\mathbf{x}_2)\dots\phi(\lambda\mathbf{x}_3)\rangle\,, \tag{1.37}$$

where all the fields are evaluated at the same time[9]. Scale invariance is most evident in the large scales ($l \lesssim 50$) of the CMB temperature angular power spectrum on the left hand side of Figure 2 (from [1]), but a detailed analysis shows that the initial conditions for the CMB were scale invariant on all scale, see right-hand side of Figure 2 (from [25]).

One would like to see scale invariance emerging from some (scaling) symmetry of the primordial physics that generated perturbations. A very simple and elegant solution is found by assuming that, during some primordial era, the background spacetime was well approximated by *de Sitter spacetime* (dS) (in so-called flat slicing)

$$ds^2 = \frac{-d\tau^2 + dx^i dx^j \delta_{ij}}{\tau^2 H^2} = -dt^2 + e^{2Ht} dx^i dx^j \delta_{ij}\,, \tag{1.38}$$

for some constant Hubble parameter H. This is a flat FLRW spacetime with scale factor

$$a = -\frac{1}{H\tau} = e^{Ht} \quad \text{(de Sitter)}\,. \tag{1.39}$$

[9]Beware that this is Cosmology lingo. In other fields, such as Conformal Field Theory, the term scale invariance is refers to the rescaling of time as well as space in the correlators. Also, ϕ could have a non-vanishing conformal dimension Δ, so that $\phi(x) \to \lambda^\Delta \phi(x)$. In cosmology, scale invariance usually refers to $\Delta = 0$, as in (1.37).

For time in the interval $-\infty < t < +\infty$ (equivalently $-\infty < \tau < 0$), this scale factor describes an expanding ($\dot{a} = Ha > 0$) accelerated ($\ddot{a} = H^2 a > 0$) universe. Like Minkowski spacetime, de Sitter is also maximally symmetric. One of the ten isometries is the *dilation* symmetry

$$\tau \to \lambda \tau\,, \quad \mathbf{x} \to \lambda \mathbf{x}\,. \tag{1.40}$$

If all other non-gravitation background quantities depend very weakly on time, then Eq. (1.40) is an approximate symmetry of the full theory and primordial correlators must be invariant under it. This in turn implies scale invariance. We will come back to this property after having learned to compute correlators in curved spacetime.

1.3 A prolonged phase of quasi-de Sitter expansion **ref**

The problems encountered in the previous section suggest that we need a prolonged phase of accelerated expansion (curvature, horizon and phase coherence problem), with a background close to dS (scale invariance), which we will call *inflation* [21,28,42]. Let's see this in detail.

De Sitter spacetime is a solution of Einstein equations in the presence of a cosmological constant

$$S_2 = \int d^4x \sqrt{-g}\left[\frac{M_{\rm Pl}^2}{2}R - \Lambda\right] \quad \Rightarrow \quad R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = 0\,. \tag{1.41}$$

The trace of this expression (for $d > 2$) tells us $R = \Lambda 2d/(d-2)$ and therefore dS is an Einstein manifold, namely the Ricci tensor is proportional to the metric[10]

$$R_{\mu\nu} = \frac{2\Lambda}{d-2} g_{\mu\nu}\,. \tag{1.43}$$

However, as the name suggest, the cosmological *constant* does not change with time and an exact dS spacetime is eternal, and cannot be connected to the universe as we know it. There is an easy fix: we introduce a "clock" ϕ that "turns off" the cosmological constant Λ after some time so that the dS phase can indeed stop when desired. I will call this clock-dependent cosmological non-constant $V(\phi)$, to avoid confusing it with the cosmological constant Λ. We will describe the dynamics of ϕ shortly.

The horizon, curvature and phase coherence problems taught us that we should postulate the existence of an early phase of accelerated expansion $\ddot{a}, \dot{a} > 0$, which we call inflation. Let's reformulate this as

$$\frac{\ddot{a}}{a} = \dot{H} + H^2 = H^2(1-\epsilon) > 0\,, \tag{1.44}$$

where I have introduced the *first Hubble slow-roll parameter*

$$\boxed{\epsilon \equiv -\frac{\dot{H}}{H^2}}\,, \tag{1.45}$$

[10] Actually, the full Riemann tensor is also given in terms of the metric

$$R_{\mu\nu\rho\sigma} = \frac{R}{12}\left(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho}\right)\,. \tag{1.42}$$

which is a dimensionless measure of the time variation of H. From (1.44), we recognise that acceleration requires $\epsilon < 1$. Also, as long as the matter sector satisfies the Null Energy Condition, which we will assume in the following (see Box 1), $\epsilon > 0$. Observations of the Cosmic Microwave Background (CMB) and of Large Scale Structures (LSS) probe cosmological scales over roughly three orders of magnitudes, and observe approximate scale invariance up to percent corrections. We gave a heuristic argument that scale invariance follows when the background is close to de Sitter spacetime, i.e. H is approximately constant. Quantitatively, we will therefore be interested in small deviations from dS, namely

$$0 < \epsilon \ll 1 \,. \tag{1.46}$$

To address the horizon and curvature problems, we need the phase of quasi de Sitter expansion to last for some "time". The requirement of generating a large and flat universe gives us a lower bound on how much the scale factor has to grow during inflation. This is best expressed in terms of the *number of efoldings* N, defined as

$$dN \equiv d\log a = H dt \quad \Rightarrow \quad N - N_0 = \log\left(\frac{a}{a_0}\right) \,. \tag{1.47}$$

For inflation to solve the Hot Big Bang problems we require

$$\Delta N_{\text{infl}} \equiv N_i - N_e > 50 + \log\left(\frac{T_{\text{reh}}}{10^{10}\text{GeV}}\right) \,, \tag{1.48}$$

where N_i denotes the beginning and N_e the end of inflation and T_{reh} is the reheating temperature, namely the temperature of radiation at the beginning of the Hot Big Bang that followed inflation. For phenomenologically viable reheating temperatures $\Delta N_{\text{infl}} \in \{25 - 60\}$. I'll often use $\Delta N_{\text{infl}} \sim 50$ for numerical estimates.

We observe approximate scale invariance for about 7 of the total ΔN_{infl} efoldings of expansion, but it is natural to assume that $\epsilon \ll 1$ remains valid during the whole of inflation. To quantify this, let us re-write the definition of ϵ and generalise it to the second and higher order Hubble slow-roll parameters (all dimensionless)

$$\epsilon \equiv -\frac{\dot{H}}{H^2} = -\partial_N \ln H \,, \tag{1.49}$$

$$\eta \equiv \frac{\dot{\epsilon}}{H\epsilon} = \partial_N \ln(\epsilon) \,, \tag{1.50}$$

$$\xi_{n\geq 3} \equiv \partial_N \ln \xi_{n-1} \,, \tag{1.51}$$

with $\xi_2 \equiv \eta$ and where I used $dN = Hdt$ from (1.47). Then, the Taylor expansion of ϵ around some reference time N_* is

$$\epsilon(N) - \epsilon(N_*) = \left.\frac{\partial \epsilon}{\partial N}\right|_{N_*} (N - N_*) + \left.\frac{\partial^2 \epsilon}{\partial N^2}\right|_{N_*} \frac{(N-N_*)^2}{2} + \mathcal{O}\left(\partial_N^3 \epsilon\right) \tag{1.52}$$

$$= \epsilon \left[\eta\,(N - N_*) + \eta\xi_3 \frac{(N-N_*)^2}{2} + \mathcal{O}\left(\eta^3, \eta^2\xi_3, \eta\xi_3\xi_4, \epsilon\right)\right] , \tag{1.53}$$

where all the slow-roll parameters are evaluated at N_*. The requirement that ϵ does not change much during inflation is then $\eta \Delta N_{\text{infl}}, \xi_n \eta \Delta N_{\text{infl}} < 1$ and so

$$\epsilon, \eta, \xi_n \ll 1 \quad \text{(slow-roll inflation)} \,. \tag{1.54}$$

1.4 Single field inflation in $P(X, \phi)$ theories

In the previous subsection, we have characterised the expansion history during inflation. We now want to ask how such an expansion history can emerge dynamically, from solving the equations of motion.

Since spatial curvature decays quickly during inflation, from now on we will assume $K = 0$. To try to mimic a cosmological constant, we were led to consider the action of scalar field coupled to gravity. We will assume the rather general action, which I will call[11] a $P(X, \phi)$ (read "P of X and ϕ" or simply "P of X") theory,

$$S = \int \sqrt{-g} \left[\frac{1}{2} M_{\rm Pl}^2 R + P(X, \phi) \right] , \tag{1.55}$$

where X represents the standard canonical kinetic term,

$$X \equiv -\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi = \frac{1}{2} (\dot{\phi}^2 - (\partial_i \phi)^2) , \tag{1.56}$$

and $P(X, \phi)$ is a generic function of its two arguments. For example, a so-called *canonical* scalar field corresponds to the simple choice

$$P = X - V(\phi) \quad \text{(canonical scalar)} , \tag{1.57}$$

where $V(\phi)$ is some potential. The coupling between the scalar field and gravity is called *minimal*, because it simply arises from writing down a Lorentz-invariant action in Minkowski spacetime and substituting $d^4x \to d^4x\sqrt{-g}$ and $\eta_{\mu\nu} \to g_{\mu\nu}$. I will not discuss here non-minimal couplings, such as for example $R\phi^2$ or $R^{\mu\nu\rho\sigma} \partial_\mu \phi \partial_\nu \phi \partial_\rho \phi \partial_\sigma \phi$.

The energy-momentum tensor (1.4) is then

$$T_{\mu\nu} = P_{,X} \partial_\mu \phi \partial_\nu \phi + g_{\mu\nu} P(X, \phi) . \tag{1.58}$$

This takes the same form as the energy-momentum tensor of a perfect fluid, (1.11), under the following identifications

$$\rho = 2XP_{,X} - P , \qquad p = P , \qquad u_\mu = -\frac{\partial_\mu \phi}{\sqrt{2X}} . \tag{1.59}$$

Let us start by focussing on the homogeneous background dynamics, $\phi = \bar{\phi}(t)$. The homogeneous equation of motion is

$$\ddot{\bar{\phi}} (P_{,X} + 2XP_{,XX}) + 3H\dot{\bar{\phi}} P_{,X} + (2XP_{,X\phi} - P_{,\phi}) = 0 , \tag{1.60}$$

while the Friedmann and acceleration equations read

$$3M_P^2 H^2 = 2XP_{,X} - P , \qquad -M_P^2 \dot{H} = XP_{,X} . \tag{1.61}$$

The specific choice of P is irrelevant for us as we will not be solving any of these equations. It suffices to notice that

$$\epsilon = \frac{3XP_{,X}}{2XP_{,X} - P} , \tag{1.62}$$

and so there are (many) choices of the function P that support a prolonged phase of slow-roll inflation, i.e. $\epsilon, \eta \ll 1$ (a necessary condition is $P_{,\phi} \neq 0$ [19]).

[11]This class of theories goes under many different names in the literature depending on the context: P-of-X theory, k-inflation [4], k-essence [5] and, specifically when $P(X, \phi) = P(X)$, superfluid [41].

2 Free fields on curved backgrounds

ref

According to our current leading paradigm, the *quantum* fluctuations on top of the classical inflationary background are the seeds of the structures that we see in our universe today. Hence we would like to quantize the inflationary model discussed above. To this end, we start from the homogeneous background $\bar{g}_{\mu\nu}$ and $\bar{\phi}$ that we discussed in the previous section, where $\bar{g}_{\mu\nu}$ is the FLRW spacetime whose scale factor $a(t)$ is given by the solution of the Friedmann and acceleration equations (1.61), and $\bar{\phi}$ obeys (1.60). Then we want to promote fluctuations to quantum operators

$$g_{\mu\nu}(t,\mathbf{x}) = \bar{g}_{\mu\nu}(t) + \hat{h}_{\mu\nu}(t,\mathbf{x}) \,, \quad \phi(t,\mathbf{x}) = \bar{\phi}(t) + \hat{\varphi}(t,\mathbf{x}) \,. \tag{2.1}$$

We will achieve this in steps. First, in this section we will quantize free theories. Second, in Section 3 we will learn the formalism to quantize general interacting theories, perturbatively. Third, we will apply this formalism to a scalar field theory in Section 4 and finally to gravity in Section 5.

2.1 Massless scalar in de Sitter

A good starting point to understand more realistic models of inflation is a massless scalar field ϕ in de Sitter spacetime without any classical background $\bar{\phi}(t) = 0$. This can arise for example by simply taking $P = X - V$ with $\phi = 0$ a minimum of V. Consider the action

$$S = -\int \sqrt{-g}\frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi = \int d^3x dt\, a^3 \frac{1}{2}\left(\dot{\varphi}^2 - \frac{1}{a^2}\partial_i\varphi\partial_i\varphi\right) \,, \tag{2.2}$$

In Fourier space, this free theory reduces to an infinite sum of decoupled harmonic oscillators

$$S = \int \frac{d^3k}{(2\pi)^3} dt\, a^3 \frac{1}{2}\left[\dot{\varphi}(\mathbf{k})\dot{\varphi}(-\mathbf{k}) - \frac{k^2}{a^2}\varphi(\mathbf{k})\varphi(-\mathbf{k})\right] \,, \tag{2.3}$$

where

$$\varphi(\mathbf{x}) = \int_{\mathbf{k}} \varphi(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \varphi(\mathbf{k}) = \int_{\mathbf{x}} \varphi(\mathbf{x}) e^{-i\mathbf{k}\cdot\mathbf{x}} \,, \tag{2.4}$$

with

$$\int_{\mathbf{k}} \equiv \int \frac{d^3\mathbf{k}}{(2\pi)^3} \,, \quad \int_{\mathbf{x}} \equiv \int d^3\mathbf{x} \,. \tag{2.5}$$

To quantize the theory, we promote φ to an operator (but I will omit the hat). As for the harmonic oscillator, we write φ in terms of creation and annihilation operators[12]

$$\varphi(\mathbf{k},t) = f_k(t) a_{\mathbf{k}} + f_k^*(t) a^\dagger_{-\mathbf{k}} \,, \tag{2.6}$$

which satisfy

$$[a_{\mathbf{k}}, a_{\mathbf{k}'}] = 0 \,, \qquad [a_{\mathbf{k}}, a^\dagger_{\mathbf{k}'}] = (2\pi)^3 \delta_D^3\left(\mathbf{k} - \mathbf{k}'\right) \,. \tag{2.7}$$

[12] Notice that $\varphi^*(\mathbf{k}) = \varphi(-\mathbf{k})$, as required by the fact that $\varphi(\mathbf{x})$ is a real field.

Here we have chosen to work in the Heisenberg picture, in which operators depend on time while the states do not, so $\varphi = \varphi(\mathbf{k}, t)$ (but I'll omit the time argument when no ambiguity arises). All the time dependence of $\varphi(t, \mathbf{k})$ has been collected in $f_k(t)$ and $f_k^*(t)$, which are known as *mode functions*. They are determined by requiring that φ solves the equations of motion[13] derived from (2.3)

$$\ddot{f}_k + 3H\dot{f}_k + \frac{k^2}{a^2} f_k = 0\,. \tag{2.9}$$

Because of the isotropy of the background f_k depends only on the norm of $\mathbf{k}$, as suggested by the notation. This equation becomes more familiar if we use conformal time ($' \equiv \partial_\tau$)

$$(af_k)'' + \left(k^2 - \frac{a''}{a}\right)(af_k) = 0\,, \tag{2.10}$$

where it looks like a harmonic oscillator with a time dependent mass $a''/a = 2/\tau^2$. The two linearly independent solutions are the complex conjugate of each other

$$f_k = \alpha(1 + ik\tau)e^{-ik\tau} + \beta(1 - ik\tau)e^{ik\tau}\,. \tag{2.11}$$

The quickest way to determine the integration constants α and β is to notice that in the far past, i.e. for $k\tau \gg 1$, (2.10) reduces to the Klein-Gordon equation for the field (af_k), since $k^2 \gg a''/a$. In other words, the field $(a\varphi(\mathbf{k}))$ at early times lives effectively in Minkowski spacetime. In this limit we expect to recover the (Heisenberg picture) free scalar field that we learn about in introductory QFT,

$$\varphi(\mathbf{x}, t) = \int \frac{d^3k_p}{(2\pi)^3} \frac{e^{i\mathbf{k}_p\cdot\mathbf{x}}}{\sqrt{2k_p}} \left[e^{-ik_pt} a_{\mathbf{k}_p} + e^{ik_pt} a^\dagger_{-\mathbf{k}_p}\right] \quad \text{(Minkowski)}\,, \tag{2.12}$$

where k_p is the physical wave number, related to the comoving one at some time by $k_p = k/a$. Recall that this choice of time dependence means that $\varphi(\mathbf{k}_p) \supset e^{+ik_pt} a^\dagger_{\mathbf{k}_p}$ *creates particles of positive energy*, as can be check from

$$\hat{H}\varphi(\mathbf{k})\left|0\right\rangle = [\hat{H}, \varphi(\mathbf{k})]\left|0\right\rangle + \varphi(\mathbf{k})\hat{H}\left|0\right\rangle = -i\dot{\varphi}(\mathbf{k})\left|0\right\rangle = +k\varphi(\mathbf{k})\left|0\right\rangle\,, \tag{2.13}$$

where the hat distinguishes the Hamiltonian $\hat{H}$ from the Hubble parameter H, and in the second step I used the Heisenberg equation and $\hat{H}\left|0\right\rangle = 0$.

To find α and β we thefore match the solution (2.11) of the de Sitter equation of motion and its time derivative to the Minkowski one, (2.12), at some early time τ_* such that $k\tau_* \gg 1$

$$af_k\Big|_{\tau=\tau_*} = \frac{e^{ik_pt}}{\sqrt{2k_p}}\Big|_{t=t_*}\,, \tag{2.14}$$

$$\partial_t(af_k)\Big|_{\tau=\tau_*} = \partial_t\left(\frac{e^{ik_pt_*}}{\sqrt{2k_p}}\right)\Big|_{t=t_*}\,. \tag{2.15}$$

[13] Or equivalently the Heisenberg equation

$$\dot{\varphi}(\mathbf{k}) = i[H, \varphi(\mathbf{k})]\,. \tag{2.8}$$

where H is the Hamiltonian derived from the action (2.3).

Solving the linear system for α and β one finds

$$\alpha = ie^{ik\tau_*(1+Ht_*)}\frac{H}{\sqrt{2k^3}}\left[1+\frac{i}{k\tau_*}-\frac{1}{2(k\tau_*)^2}\right], \tag{2.16}$$

$$\beta = ie^{-ik\tau_*(1-Ht_*)}\frac{H}{\sqrt{2k^3}}\frac{1}{2(k\tau_*)^2}. \tag{2.17}$$

If the matching is done in the infinite past, $k\tau_* \to -\infty$, this reduces simply to

$$\lim_{\tau\to-\infty}|\alpha| = \frac{H}{\sqrt{2k^3}}, \qquad \lim_{\tau\to-\infty}\beta = 0. \tag{2.18}$$

The normalization of α is fixed only up to an overall phase because one can always shift t in Minkowski and so the value of t_* is arbitrary. The dS mode functions that create positive-energy particles in the infinite past therefore are

$$\boxed{f_k = \frac{H}{\sqrt{2k^3}}(1+ik\tau)e^{-ik\tau}} \quad \text{(dS mode functions)}. \tag{2.19}$$

The dS mode functions (2.19) are very different from the Minkowski counterpart when the physical wavenumber becomes smaller than the comoving Hubble parameter

$$k < k_{\text{H.c.}} = aH = \frac{1}{|\tau|} \quad \text{(Hubble crossing)}, \tag{2.20}$$

where "H.c." stands for *Hubble crossing*, sometime also called horizon crossing. In physical length scales, this means the physical wavelength $\lambda_p = a/k$ is stretched by the expansion to be become larger than the Hubble radius $1/H$. Since k and H are (approximately) constant, while $a = e^{Ht}$ grows with time, all modes cross the Hubble radius as time proceeds and become "superHubble" modes. Unlike "subHubble" modes, $k \gg aH$, which oscillate, superHubble modes *freeze out* and asymptotes a constant. The canonical commutation relations with the momentum conjugate

$$\pi(\mathbf{x}) \equiv \frac{\partial\mathcal{L}}{\partial\dot\varphi(\mathbf{x})} = a^3\dot\varphi(\mathbf{x}), \tag{2.21}$$

impose the constraint

$$[\varphi(\mathbf{k}), \pi(\mathbf{k}')] = a^2\left(f_k f_k'^* - f_k^* f_k'\right)(2\pi)^3\delta_D^3\left(\mathbf{k}+\mathbf{k}'\right) \overset{!}{=} i(2\pi)^3\delta_D^3\left(\mathbf{k}+\mathbf{k}'\right), \tag{2.22}$$

Using (2.11) we find the condition

$$|\alpha_k|^2 - |\beta_k|^2 = \frac{H^2}{2k^3}, \tag{2.23}$$

which is indeed satisfied by (2.18).

More rigorous ways to derive (2.19) include Hamiltonian minimization, i.e. choosing as vacuum the lowest energy state in the asymptotic past and matching to the uniquely defined Euclidean vacuum of the Wick rotated Euclidean field theory. Now that we have related φ to creation and annihilation operators, we can specify the "vacuum state" $|0\rangle$ by the usual condition $a_{\mathbf{k}}|0\rangle = 0$ for all $\mathbf{k}$'s. This is often called the Bunch-Davies vacuum or Hartle-Hawking state. Excited states are then obtained by acting with creation operators on this vacuum.

Box 2.1 Free theories and Gaussianity All free theories can be understood by analogy with the most famous free theory, the quantum harmonic oscillator. In quantum mechanics you learn that the probability of finding a particle at position x is a Gaussian distribution

$$\text{Prob}(x) \sim |\psi(x)|^2 \propto e^{-x^2/(2\sigma^2)}\,, \tag{2.25}$$

where $\psi(x)$ is the position space wavefunction and $\sigma^2 = \langle x^2 \rangle$ describes the width of the Gaussian and is fixed by the parameters of the theory ($\sigma^2 = (2m\omega)^{-1}$). It is clear by parity that

$$\langle x^{2n+1} \rangle \propto \int_{-\infty}^{\infty} dx\, |\psi(x)|^2 x^{2n+1} \sim \int dx\, e^{-x^2/(2\sigma^2)} x^{2n+1} = 0\,. \tag{2.26}$$

While with repeated integration by parts we can always rewrite $\langle x^{2n} \rangle$ in terms of $\langle x^2 \rangle^n$. Because of this, the expression free theory and Gaussian theory are often used interchangeably. In the next section we will study interacting theories, where we will compute *non-Gaussianities*, i.e. deviations from a Gaussian wavefunction. This discussion applies to quantum*field* theory (as opposed to quantum mechanics) by thinking of $\phi(\mathbf{k})$ as an infinite collection of decoupled harmonic oscillators, as depicted in Fig. 3

What observables can we compute for this theory? As familiar from Quantum Mechanics, observables are given by the expectation value of operators. In cosmology, we have observational access only to these expectation values in the infinite future $k\tau \to 0$. In this limit, observables become approximately constant and so we will only be interested in the expectation value of product of correlators at equal time, or simply cosmological correlators for short

$$\lim_{\tau \to 0} \langle \mathcal{O}(\mathbf{k}_1, \tau) \dots \mathcal{O}(\mathbf{k}_n, \tau) \rangle\,, \tag{2.24}$$

for some local operators $\mathcal{O}$. Because we are studying a *free* theory, all information is contained in the two-point correlators of φ and its conjugate momentum π. All odd-point correlators vanish by the symmetry $\varphi \to -\varphi$ and all higher even-point correlators can be reduced to the two-point one using Wick's theorem.

Let's compute the two-point correlation function of φ:

$$\lim_{\tau \to 0} \langle \varphi(\mathbf{k}) \varphi(\mathbf{k}') \rangle = |f_k|^2 \langle a_{\mathbf{k}} a^\dagger_{-\mathbf{k}'} \rangle \tag{2.27}$$

$$= (2\pi)^3 \delta_D^3 \left(\mathbf{k} + \mathbf{k}'\right) P(k) \quad \text{with} \quad P(k) = \frac{H^2}{2k^3}\,. \tag{2.28}$$

Here, I have introduced the *power spectrum* $P(k)$, which is just the two-point correlator stripped of the Dirac delta and its accompanying factor of $(2\pi)^3$. A few comments are in order:

- The Dirac delta reminds us that momentum is conserved as a consequence of the homogeneity of the background. In Example Sheet 1 you will show that this delta appears in all correlators. Pictorially we can imagin that perturbations in this state must exist in pairs of opposite wavenumber $\mathbf{k}$ and $-\mathbf{k}$.

- $P(k)$ does not depend on the direction of $\mathbf{k}$ as consequence of the isotropy of the background.

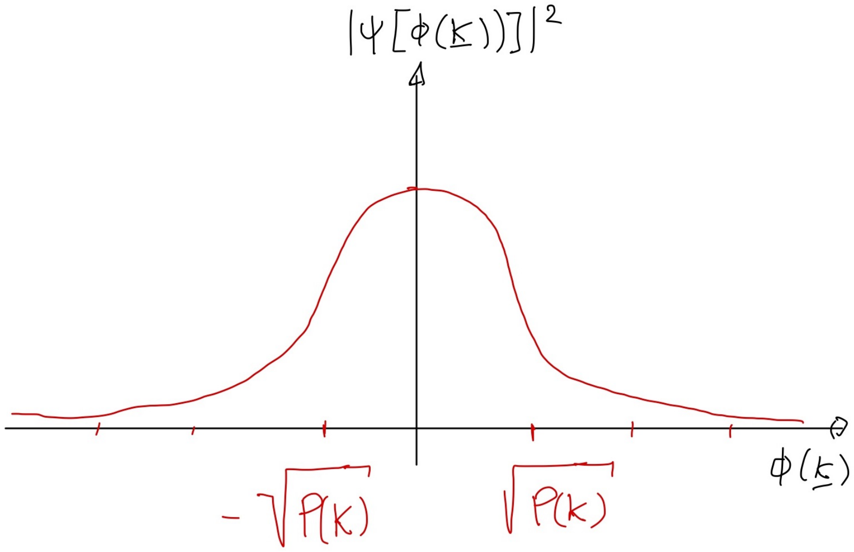

Figure 3: The figure show that the probability distribution function of a free field theory, which is proportional to the norm square of the wavefunction, is a multivariate Gaussian in the infinitely many decoupled harmonic oscillators $\phi(\mathbf{k})$, each with variance $P(k)$.

- As we will see shortly, the fact that the power spectrum asymptotes some (non-vanishing) constant value as $\tau \to 0$ is related to the absence of a mass.
- The k-dependence $P \propto k^{-3}$ is the one corresponding to *scale invariance*, as defined in (1.37). To see this, we can Fourier transform to the real-space *correlation function*,[14]

$$\langle \varphi(\mathbf{x})\varphi(\mathbf{0})\rangle = \int_{\mathbf{k}\mathbf{k}'} \langle \varphi(\mathbf{k})\varphi(\mathbf{k}')\rangle \sim H^2 \,, \tag{2.30}$$

and notice that the correlation does not depend on distance. In particular, it doesn't change if we rescale $\mathbf{x} \to \lambda \mathbf{x}$.

It is worth comparing the power spectrum in dS, (2.28), with that in Minkowski

$$\langle \varphi(\mathbf{k})\varphi(\mathbf{k}')\rangle = (2\pi)^3 \delta_D^3 \left(\mathbf{k} + \mathbf{k}'\right) \frac{1}{2k} \quad (\text{Minkowski})\,. \tag{2.31}$$

and the associated real-space correlation function

$$\langle \varphi(\mathbf{x})\varphi(\mathbf{0})\rangle = \int_{\mathbf{k}\mathbf{k}'} \langle \varphi(\mathbf{k})\varphi(\mathbf{k}')\rangle \sim \frac{1}{x^2} \quad (\text{Minkowksi})\,. \tag{2.32}$$

So in dS the correlation function is independent of the distance, while in Minkowski it decays as $1/x^2$, as expected for a massless particle. In the Example Sheet 1, you will derive the correlators involving the momentum conjugate.

[14]Actually the dS correlation function at separated points, $\mathbf{x} \neq 0$, is IR divergent. The physical reason is that dS is eternal. This divergence can be regularized either with a small tilt of the power spectrum $k^{-(3+\delta)}$, with $0 < \delta \ll 1$ or with an IR cutoff k_{IR} of the integral ($\tilde{k} = xk$ is dimensionless)

$$\langle \varphi(\mathbf{x})\varphi(\mathbf{0})\rangle = \frac{H^2}{(2\pi)^2} \int_{xk_{\mathrm{IR}}}^{\infty} d\tilde{k}\, \frac{\sin \tilde{k}}{\tilde{k}^2} \xrightarrow{xk_{\mathrm{IR}} \to 0} \gamma_E - 1 + \log(xk_{\mathrm{IR}}) + \mathcal{O}((xk_{\mathrm{IR}})^2)\,. \tag{2.29}$$

.

2.2 Massive scalar in de Sitter

It is interesting to ask what changes if the scalar field has a mass m,

$$S = -\int \sqrt{-g}\frac{1}{2}\left[\partial_\mu\varphi\partial^\mu\varphi - m^2\varphi^2\right] . \tag{2.33}$$

You will quantize this theory in Example Sheet 1. The relation to creation and annihilation is the same

$$\varphi(\mathbf{k}, t) = f_k(t)a_{\mathbf{k}} + f_k^*(t)a_{-\mathbf{k}}^\dagger , \tag{2.34}$$

but now the mode functions are modified

$$f_k(\tau) = \frac{\sqrt{\pi}H}{2}(-\tau)^{3/2}H_\nu^{(1)}(-k\tau) , \quad \nu = \sqrt{\frac{9}{4} - \frac{m^2}{H^2}} , \tag{2.35}$$

where $H^{(1)}$ is the Hankel function of the first kind. Hankel functions are solutions of the Bessel's differential equation and are linear combination of Bessel functions. For $\tau \to 0$ this becomes

$$f_k(\tau \to 0) = H\tau^{3/2}\left[\frac{\sqrt{\pi}(-k\tau)^\nu}{2^{1+\nu}\Gamma[\nu+1]}(1 + i\cot(\pi\nu)) - \frac{i(-k\tau)^{-\nu}}{\sqrt{\pi}2^{1-\nu}\Gamma[\nu]}\right] + \dots . \tag{2.36}$$

Now it is useful to distinguish two cases. The first case is when the mass square is small or negative, $m^2 < 9H^2/4$. Then ν is real and positive. In this case, the first term in brackets approaches zero faster than the second and can be neglected. So the power spectrum now becomes

$$P(k) = |f_k|^2 = \frac{H^2}{\pi 2^{2(\nu-1)}\Gamma(\nu)^2}\frac{(-k\tau)^{3-2\nu}}{k^3} \quad (\text{for } m^2 < \frac{9}{4}H^2) . \tag{2.37}$$

Because of the mass, the power spectrum is not scale invariant anymore, $P \propto k^{-2\nu}$. Also, P has acquired a time dependence. For positive $m^2 > 0$, one finds $3 - 2\nu > 0$ and the power spectrum decays with time and vanish at future infinity. This is to be expected because the quadratic potential pushes the field towards $\varphi = 0$. For negative m^2 we would expect an instability and indeed the power spectrum grows with time and diverges at future infinity.

The second case is when the mass square is large and positive, $m^2 > 9H^2/4$, then ν becomes complex and the two terms in the brackets of (2.36) are of the same order. The power spectrum oscillates while decaying as τ^3. In cosmology, we are mostly interested in massless or almost massless fields, which do not create large instability and whose perturbations survive long enough to be observable at late times.

2.3 Particle creation*

For QFT in Minkowski, we can think of excitations generated by the creation operators as particles. However, in curved spacetime particle and particle number are more subtle concepts. Let's see this in detail. We found field excitations in dS oscillate as $f_k \sim e^{-ik\tau}$, where the comoving wavenumber k is related to the Minkowski energy by

$$E = \sqrt{k_i k_j g^{ij}} = \frac{k}{a} . \tag{2.38}$$

Let us now Taylor expand the time-dependent phase of f_k in time around some time t_*:

$$
\begin{aligned}
-ik\tau &= i\frac{k}{aH} = i\frac{k}{a_*}\frac{a_*}{aH} \\
&= i\frac{E}{H}e^{-H(t-t_*)} \simeq iE\left[\frac{1}{H} - (t-t_*) + \frac{1}{2}H(t-t_*)^2 + \dots\right] .
\end{aligned} \tag{2.39}
$$

The first term in brackets is an irrelevant phase. The second term is precisely the time dependence of particles in Minkowski, namely e^{-iEt}. So field excitations in dS have a chance to look like particles only for a time interval $(t-t_*) \ll 1/H$, during which we can neglect the higher order terms in brackets. Moreover, we must demand that during this interval, the wavefunction oscillates many times, so $E(t-t_*) \gg 1$. Using again (2.39) this requires $E/H = -k\tau \gg 1$. When this condition is not satisfied, the energy and momentum of the state are redshifted by the expansion before a single oscillation of the wavefunction.

Even when particles can be defined, the expansion of the universe can create particles (unless there is a conserved quantum number) at a rate controlled again by H. This is to be expected as the expansion of the universe breaks time translations and so energy is not conserved. Let's see this in more detail.

In the previous section, we found the mode functions by demanding that φ creates positive-energy particles at $k\tau \to -\infty$. Let's instead require that φ creates positive-energy particles at some finite τ_*, still satisfying $|k\tau_*| \gg 1$. By matching to the Minkowski vacuum at τ_*, we find α and β as given in (2.16). The quantized field then takes the form

$$
\varphi(\mathbf{k}) = g_k b_{\mathbf{k}} + g_k^* b_{-\mathbf{k}}^\dagger \tag{2.40}
$$

with[15]

$$
\begin{aligned}
g_k &= \alpha f_k(\tau) + \beta f_k^*(\tau) \\
&= \frac{H}{\sqrt{2k^3}}\left[1 + \frac{i}{k\tau_*} - \frac{1}{2(k\tau_*)^2}\right] f_k(\tau) + e^{-2ik\tau_*}\frac{H}{\sqrt{2k^3}}\frac{1}{2(k\tau_*)^2} f_k^*(\tau) ,
\end{aligned} \tag{2.41}
$$

and $\{b_{\mathbf{k}}, b_{\mathbf{k}}^\dagger\}$ a new set of creation and annihilation operators, which define a new vacuum state by $b_{\mathbf{k}}|\tilde{0}\rangle = 0$. By matching the two expressions for $\varphi(\mathbf{k})$, (2.6) and (2.40), we see that the two sets of ladder operators are related,

$$
a_{\mathbf{k}} = \frac{\sqrt{2k^3}}{H}\left(\alpha b_{\mathbf{k}} + \beta^* b_{-\mathbf{k}}^\dagger\right) , \qquad a_{\mathbf{k}}^\dagger = \frac{\sqrt{2k^3}}{H}\left(\beta b_{-\mathbf{k}} + \alpha^* b_{\mathbf{k}}^\dagger\right) , \tag{2.42}
$$

This relation is called a *Bogoliubov transformation.* It can be inverted to give

$$
b_{\mathbf{k}} = \frac{\sqrt{2k^3}}{H}\left(\alpha^* a_{\mathbf{k}} + \beta^* a_{-\mathbf{k}}^\dagger\right) , \qquad b_{\mathbf{k}}^\dagger = \frac{\sqrt{2k^3}}{H}\left(\beta a_{-\mathbf{k}} + \alpha a_{\mathbf{k}}^\dagger\right) , \tag{2.43}
$$

Now we want to ask what a detector that measures $b_k^\dagger$ excitations would measure in the Bunch Davies vacuum. To this end, we define the "b-particle" number operator

$$
N_b(\mathbf{k}) = b_{\mathbf{k}}^\dagger b_{\mathbf{k}} . \tag{2.44}
$$

[15] Again, I used the arbitrariness in t_* to multiply g_k by a convenient phase.

As expected, this operator has a vanishing expectation value in the $\left|\tilde{0}\right\rangle$ state. But if we compute its expectation value in the Bunch-Davies vacuum $|0\rangle$ we find

$$\langle 0|\, N_b(\mathbf{k})\, |0\rangle = \frac{2k^3}{H^2}|\beta_k|^2(2\pi)^3\delta_D^3(\mathbf{0}) = \frac{1}{4(k\tau)^4}(2\pi)^3\delta_D^3(\mathbf{0})\,. \tag{2.45}$$

The singular factor $\delta_D^3(\mathbf{0})$ is a reminder that we are working with an infinite volume

$$(2\pi)^3\delta_D^3(\mathbf{0}) = \lim_{V\to\infty}\int_V d^3x e^{-i\mathbf{0}\cdot\mathbf{x}} = \lim_{V\to\infty} V\,. \tag{2.46}$$

It is therefore wise to compute the number density of particles, $n_b(\mathbf{k}) \equiv N_b(\mathbf{k})/V$, instead of the total number $N_b(\mathbf{k})$. We find

$$\langle 0|\, n_b(\mathbf{k})\, |0\rangle = \frac{1}{4(k\tau)^4} \neq 0\,. \tag{2.47}$$

In words, the Bunch-Davies "vacuum" state is found to contain some b-particles, defined with respect to the "vacuum" state at some large but finite $|\tau_*|$. As we take $\tau_* \to -\infty$, $\left|\tilde{0}\right\rangle$ approaches $|0\rangle$ and indeed the number density of particles vanishes as expected. We can say that the expansion of the universe creates particles. These particles are always created in pairs of opposite wavenumber, to conserve momentum, $n(-\mathbf{k}) = n(\mathbf{k})$.

2.4 Gravitons in de Sitter

By following a very similar procedure as for the scalar field, it is straightforward to quantize metric fluctuations as well. We divide the metric into a classical background $\bar{g}_{\mu\nu}$ and small quantum fluctuations $h_{\mu\nu}$:

$$g_{\mu\nu}(x,t) \;\;=\;\; \bar{g}_{\mu\nu}(t) + h_{\mu\nu}(x,t)\,, \tag{2.48}$$

A priori, there are ten independent components of $h_{\mu\nu}$. Four of them, namely $h_{0\mu}$ obey constraint equations, which are at most first order in time derivatives, and therefore are not dynamical. To see this, recall the Bianchi identity

$$\nabla^\mu G_{\mu\nu} \equiv \nabla^\mu\left(R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R\right) = 0\,. \tag{2.49}$$

Writing out the covariant derivative we find

$$\partial_t G^{t\nu} = -\partial_k G^{k\nu} - \Gamma^\alpha_{\alpha\gamma}G^{\nu\gamma} - \Gamma^\nu_{\alpha\gamma}G^{\alpha\gamma}\,. \tag{2.50}$$

Since the right-hand side has at most second derivatives of the metric, we conclude that $G^{t\nu}$ has at most one time derivative. But then the metric must appear with just one time derivative in four of the Einstein equations, namely $G^{t\nu} = T^{t\nu}$. These must then be constraint equations that limit the set of consistent initial data $h_{\mu\nu}, \dot{h_{\mu\nu}}\}$ that one can specify. It takes a bit more work and the ADM formalism to specify which components of the metric appear with at most one derivative. Later we will see that these are precisely $g_{\mu 0}$.

Other four components in δg_{ij} can be set to zero by a gauge transformation $x^\mu \to x^\mu + \epsilon^\mu(x)$, which changes the metric by

$$g_{\mu\nu} \to g_{\mu\nu} + 2\nabla_{(\mu}\epsilon_{\nu)}\,, \tag{2.51}$$

where the symmetrization of indices is defined in (0.5). The remaining two components are dynamical and describe the two helicities of the graviton, $h = \pm 2$. A convenient gauge choice to study the linear dynamics of gravitons on an FLRW spacetime is

$$ds^2 = -dt^2 + a^2 (\delta_{ij} + \gamma_{ij}) \, dx^i dx^j \,, \tag{2.52}$$

where γ_{ij} is transverse, $\partial_i \gamma_{ij} = 0$, and traceless, $\gamma_{ii} = 0$, and so has indeed only $6-3-1 = 2$ independent components. For the moment $a(t)$ is arbitrary. We should now expand the Einstein-Hilbert action,

$$S = \frac{M_{\rm Pl}^2}{2} \int d^4x \sqrt{-g} R \,, \tag{2.53}$$

to quadratic order in γ_{ij}. At the end of a long and tedious calculation, one finds

$$S_2 = \frac{M_{\rm Pl}^2}{8} \int d^3x d\tau \, a^2 \left[\gamma'_{ij}\gamma'_{ij} - \partial_k \gamma_{ij} \partial_k \gamma_{ij} \right] \,. \tag{2.54}$$

This action could have been easily guessed as it contains the only two terms allowed by the symmetries of the problem[16]. As we do for the photon, we can expand the graviton in plane waves by writing[17]

$$\gamma_{ij}(x) = \int_{\mathbf{k}} \sum_{s=+,\times} \epsilon^s_{ij}(\mathbf{k}) \gamma_s(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x}} \,, \tag{2.55}$$

where $\epsilon^s_{ij}(\mathbf{k})$ are *polarization tensors*, which are generally complex and satisfy

$$\epsilon^s_{ii}(\mathbf{k}) = k^i \epsilon^s_{ij}(\mathbf{k}) = 0 \qquad \text{(transverse and traceless)} \,, \tag{2.56}$$

$$\epsilon^s_{ij}(\mathbf{k}) = \epsilon^s_{ji}(\mathbf{k}) \qquad \text{(symmetric)} \,, \tag{2.57}$$

$$\epsilon^s_{ij}(\mathbf{k}) \epsilon^s_{jk}(\mathbf{k}) = 0 \qquad \text{(lightlike)} \,, \tag{2.58}$$

$$\epsilon^s_{ij}(\mathbf{k}) \epsilon^{s'}_{ij}(\mathbf{k})^* = 2\delta_{ss'} \qquad \text{(normalization)} \,, \tag{2.59}$$

$$\epsilon_{ij}(\mathbf{k})^* = \epsilon_{ij}(-\mathbf{k}) \qquad (\gamma_{ij}(x) \text{ is real}) \,. \tag{2.60}$$

Explicit expressions for ϵ^s_{ij} are derived in Appendix A. Let's re-write the action using this decomposition:

$$S_2 = \frac{M_{\rm Pl}^2}{4} \int_{\mathbf{k}} d\tau \, a^2 \sum_{s=+,\times} \left[\gamma'_s(\mathbf{k}) \gamma'_s(-\mathbf{k}) - \frac{k^2}{a^2} \gamma_s(\mathbf{k}) \gamma_s(-\mathbf{k}) \right] \,. \tag{2.61}$$

Now this action consists of two independent copies of the action for a massless scalar field (2.3), up to an overall factor of $M_{\rm Pl}^2/2$. The two polarizations $\gamma_{+,\times}$ are now canonically normalized. To quantize the theory we can then proceed exactly as we did in Section

[16] By rotational invariance one has to contract the two indices of γ_{ij}. Any contraction with δ_{ij} or ∂_i gives zero, so the only possibility is contracting with another γ_{ij}. The Ricci scalar contains two derivatives, which can act on the background (e.g. on $a(t)$) or on γ_{ij}. The only terms with two derivatives on perturbations are those in S_2. The relative factor is what we call *speed of light* and has been set to unity here. Terms with one time derivative can be integrated by part into terms without any derivatives. Finally, terms without any derivatives cannot be invariant under diffs, (2.51), so they must all cancel out.

[17] Notice that $\gamma_{ij}(x)$ and ϵ^s_{ij} are dimensionless, while $[\gamma^s(\mathbf{k})] \sim M^{-3}$.

2.1. We promote $\gamma_s(\mathbf{k})$ to an operator and write it in terms of creation and annihilation operators

$$\gamma_s(\mathbf{k}) = \frac{\sqrt{2}}{M_{\rm Pl}}\left(f_k a^s_{\mathbf{k}} + f^*_k a^{s\dagger}_{-\mathbf{k}}\right) . \tag{2.62}$$

where the commutation relations are the usual ones,

$$[a^s_{\mathbf{k}}, a^{s'}_{\mathbf{k}'}] = 0 \qquad\qquad [a^s_{\mathbf{k}}, a^{s'\dagger}_{\mathbf{k}'}] = (2\pi)^3\delta^3_D\left(\mathbf{k}-\mathbf{k}'\right)\delta_{ss'} . \tag{2.63}$$

If we assume a dS background, i.e. $a = e^{Ht}$, the mode functions f_k are the same as for the massless scalar field, (2.19). The graviton power spectrum, often called the *tensor power spectrum* $P_T(k)$ can be easily computed

$$\langle\gamma_{ij}(\mathbf{k})\gamma_{ij}(\mathbf{k}')\rangle = \sum_{s,s'}\epsilon^s_{ij}(\mathbf{k})\epsilon^{s'}_{ij}(\mathbf{k}')\langle\gamma_s(\mathbf{k})\gamma_{s'}(\mathbf{k}')\rangle \tag{2.64}$$

$$= \frac{2}{M^2_{\rm Pl}}\sum_{s,s'}\epsilon^s_{ij}(\mathbf{k})\epsilon^{s'}_{ij}(\mathbf{k}')(2\pi)^3\delta^3_D(\mathbf{k}+\mathbf{k}')|f_k|^2 \tag{2.65}$$

$$= \frac{2}{M^2_{\rm Pl}}\frac{H^2}{2k^3}(2\pi)^3\delta^3_D(\mathbf{k}+\mathbf{k}')\sum_{s,s'}2\delta_{ss'} \tag{2.66}$$

$$= (2\pi)^3\delta^3_D\left(\mathbf{k}+\mathbf{k}'\right)P_T \quad \text{with} \quad P_T = \frac{4}{k^3}\frac{H^2}{M^2_{\rm Pl}} . \tag{2.67}$$

3 Interacting fields and the in-in formalism

Now that we understand free fields, we can describe weakly interacting fields as we do in particle physics, i.e. in a perturbative expansion. Indeed, we already know that gravity is a non-linear theory and so at least we should see gravitational interactions at play in the early universe. Moreover, we don't know the laws of physics at very high energies, so it is possible that the inflaton had other, non-gravitational interactions as well. While free theories are fully characterized by their power spectra, interacting theories have an enormously richer phenomenology, which we will start exploring here. In particular interactions leads to correlators that are not fixed by the power spectrum. Since the wavefunction of is non-Gaussian (see Box 1), these correlators are often called *non-Gaussianities*. In this section, I'll set up the general formalism to compute correlators in interacting theories in Section 3.1 and 3.2 and provide two explicit examples in Section 3.3.

3.1 Particle physics and scattering amplitudes

A highly effective way to study an object is to throw things at it and see how they bounce off. This describes mundane activities such as looking at things by scattering photons. But it also applies to more advanced "imaging" techniques such as X-ray radiography, electron microscopes and particle accelerators, just to name a few. In the quantum mechanical context the main object of study are scattering amplitudes, namely quantum mechanical amplitudes for the schematic process

$$S_{\alpha\beta} \equiv \langle\alpha, out|\beta, in\rangle \equiv \langle\alpha; +\infty|\beta; -\infty\rangle_S = \langle\alpha|\, S\,|\beta\rangle_H \tag{3.1}$$

where $|\alpha\rangle$ and $|\beta\rangle$ are eigenstates of the free Hamiltonian, and the subscripts S and H refer to the Schrödinger and Heisenberg pictures, respectively[18]. S is a unitary operator, $SS^\dagger = \mathbb{I}$, while $S_{\alpha\beta}$ is a unitary matrix with complex entries, $S_{\alpha\beta}S^\dagger_{\beta\gamma} = \delta_{\alpha\gamma}$.

Given a Hamiltonian $\hat{H}$, one finds the largest number of operators that commute with $\hat{H}$, i.e. a subset of the symmetries of theory, and uses their eigenvalues to label the α and β states. For example, in particle physics single particles states are irreducible representations of the Poincaré group and are classified by their four-momentum p^μ and their *spin* if massive or *helicity* if massless (i.e. the irreps of the associated little group, see e.g. Chapter 3 of [48]). For example, the scattering of 2 into $(n-2)$ particles has amplitude

$$S_{\alpha\beta} = 2^{n/2}\sqrt{E_1E_2\dots E_n}\,\langle\Omega|\,a_{\mathbf{p}_3}(\infty)\dots a_{vp_n}(\infty)a^\dagger_{\mathbf{p}_1}(-\infty)a^\dagger_{\mathbf{p}_2}(-\infty)\,|\Omega\rangle\;, \tag{3.4}$$

where $|\Omega\rangle$ is the Minkowski vacuum, $E(\mathbf{p}) = \sqrt{\mathbf{p}^2+m^2}$, and I used the relativistic normalization of states. Probabilities are obtained by squaring amplitudes

$$\text{Prob} \sim |\,\langle\alpha|\,S\,|\beta\rangle\,|^2\,. \tag{3.5}$$

For most systems of physical interest, the S-matrix can only be computed in perturbation theory. Let's assume that the Hamiltonian of the theory can be divided into a free Hamiltonian $\hat{H}_0$ and an interaction Hamiltonian

$$\hat{H} = \hat{H}_0 + \hat{H}_{int}\,, \tag{3.6}$$

which induces a small perturbation. To study the effect of $\hat{H}_{int}$ it is convenient to introduce the *interaction picture*, labelled by I, where operators evolve with the free Hamiltonian $\hat{H}_0$ and states evolve with the interaction Hamiltonian:

$$|\psi,t\rangle_I = e^{+i\int\hat{H}_0dt}|\psi,t\rangle_S\,, \tag{3.7}$$

$$\mathcal{O}_I(t) = e^{+i\int\hat{H}_0dt}\mathcal{O}_Se^{-i\int\hat{H}_0dt}\,, \tag{3.8}$$

The interaction picture is related to the Heisenberg picture by introducing the interaction-picture evolution operator $U_I(t,t_i)$ between some initial time t_i and some time t:

$$|\psi,t\rangle_I = U_I(t,t_i)\,|\psi\rangle_H\;, \tag{3.9}$$

$$\mathcal{O}_I(t) = U_I(t,t_i)\mathcal{O}_H(t)U_I^\dagger(t,t_i)\,. \tag{3.10}$$

From the Heisenberg equation for $\mathcal{O}_H$ or the Schrödinger equation for $|\psi\rangle_S$, one finds that the evolution operator in the interaction picture U_I obeys

$$\frac{d}{dt_2}U_I(t_2,t_1) = -i\hat{H}_{int}(t_2)U_I(t_2,t_1)\,, \tag{3.11}$$

$$\frac{d}{dt_1}U_I(t_2,t_1) = iU_I(t_2,t_1)\hat{H}_{int}(t_1)\,, \tag{3.12}$$

[18]Recall that in the Schrödinger (S) and Heisenberg (H) pictures

$$|\psi,t\rangle_S = e^{-i\int\hat{H}dt}\,|\psi,t_i\rangle_S\;, \qquad \mathcal{O}_S(t) = \mathcal{O}_S(t_i)\equiv\mathcal{O}_S\,, \tag{3.2}$$

$$|\psi\rangle_H = |\psi,t_i\rangle_S = e^{+i\int\hat{H}dt}|\psi,t\rangle_S\,, \qquad \mathcal{O}_H(t) = e^{i\int\hat{H}dt}\mathcal{O}_S\,e^{-i\int\hat{H}dt}\,. \tag{3.3}$$

for some reference initial time t_i.

as well as $U_I(t,t) = 1$. For $t_2 \geq t_1$, the solution of these equation is concisely given by Dyson's formula

$$U_I(t_2, t_1) = T \exp\left(-i \int_{t_1}^{t_2} dt' \hat{H}_{int}(t')\right) , \tag{3.13}$$

$$U_I^\dagger(t_2, t_1) = \bar{T} \exp\left(i \int_{t_1}^{t_2} dt' \hat{H}_{int}(t')\right) , \tag{3.14}$$

where the (anti) time-ordered operator ($\bar{T}$) T arranges the operators from left to right in order of (increasing) decreasing time. In the perturbative expansion, this takes the form

$$U(t_2, t_1) = 1 - i \int_{t_1}^{t_2} dt' \hat{H}_{int}(t') - \int_{t_1}^{t_2} dt' \int_{t_1}^{t'} dt'' \hat{H}_{int}(t') \hat{H}_{int}(t'') + \ldots . \tag{3.15}$$

When the arguments are not in the right order, the solution of (3.11) and (3.12) is instead given by (using the shorthand $U_{21} = U_I(t_2, t_1)$, etc.)

$$U_{12} \equiv U_{21}^\dagger = U_{21}^{-1} \qquad\qquad U_{12}^\dagger \equiv U_{21} , \tag{3.16}$$

in such a way that

$$U_{12} U_{21} = U_{21} U_{12} = 1 , \qquad\qquad U_{32} U_{21} = U_{31} , \tag{3.17}$$

for any ordering of $t_{1,2,3}$. Dyson's formula then gives us the useful representation

$$S = U_I(\infty, -\infty) = T \exp\left[-i \int_{-\infty}^{+\infty} dt' \hat{H}_{int}(t')\right] . \tag{3.18}$$

Notice that $\hat{H}_{int}$ and hence U_I in the interaction picture are written in terms of *free fields* (2.12).

3.2 Cosmology and correlators

The situation in cosmology is different from that in particle physics in three major respects:

- **Broken Poincaré symmetry:** As already mentioned, the FLRW background on which cosmological perturbations propagate has four isometries less than the maximally-symmetric Minkowski spacetime. In particular, Lorentz boosts and time translations are spontaneously broken.
- **In-in vs in-out:** At early times, cosmological perturbations were effectively in flat space and we can define an initial state, just as in the discussion of Section 2. Conversely, at late times, cosmological perturbations in general evolve and interact with each other and we cannot assume that the state of the universe at late times is a superposition of free states, as we did for particle scattering. So instead of "in-out" amplitudes, we will be interested in "in-in" expectation values.
- **Cosmic variance:** In an expanding universe with a finite age, causality imposes that there is only a finite volume that we can access observationally. If the expansion decelerates, $\ddot{a} < 0$, we can wait long enough and observe any other

spacetime point. Instead, the expansion of our universe is currently accelerating $\ddot{a}/a \sim (10^{17}\text{sec})^{-2}$ (pretty slowly). If this acceleration continues in the future, the largest spatial volume we can ever observe is of order the Hubble volume today, $H_0^{-3} \sim (4\text{Gpc})^3$. Hence we cannot observe fluctuations in the whole universe and so our measurements have an intrinsic sample variance, known in this context as *cosmic variance.*

Let's define an *in-in correlator* as the expectation value of some operator $\mathcal{O}$ on some state Ω

$$\langle \mathcal{O} \rangle = \langle \Omega | \, \mathcal{O} \, | \Omega \rangle \ . \tag{3.19}$$

In this discussion, $\mathcal{O}$ will always be the *equal-time* product of operators at different space points. Time ordering is therefore irrelevant. As familiar from quantum mechanics, correlators of Hermitian operators are observable and must be real (unlike scattering amplitudes)

$$\langle \Omega | \, \mathcal{O} \, | \Omega \rangle^* = \langle \Omega | \, \mathcal{O}^\dagger \, | \Omega \rangle = \langle \Omega | \, \mathcal{O} \, | \Omega \rangle \in \mathbb{R} \, . \tag{3.20}$$

For formal manipulations, the Heisenberg picture is very convenient. But explicit calculations are most easily performed in the *interaction picture.* We already know from (4.6) that

$$\mathcal{O}_H(t) = U_I^\dagger(\tau, -\infty) \mathcal{O}_I(\tau) U_I(\tau, -\infty) \, , \tag{3.21}$$

where all the fields on the right-hand side are in the interaction picture, i.e. they are just the free fields we studied in Section 2. The last thing we need is to define $|\Omega\rangle$. We will only be interested in the case in which $|\Omega\rangle$ is the "vacuum" of the interacting theory, which in the far past asymptotes the free theory vacuum $|0\rangle$, defined by $a_{\mathbf{k}} \, |0\rangle = 0$ in Section 2:

$$\lim_{\tau \to -\infty} |\Omega\rangle = |0\rangle \ . \tag{3.22}$$

For adiabatic evolution energy levels never cross, so $|\Omega\rangle$ must be the lowest energy state of the full theory, just as $|0\rangle$ is the lowest energy level of the free theory. Also, $|\Omega\rangle$ must minimize both $\hat{H}_0$ and $\hat{H}_{int}$ separately. We can then relate $|\Omega\rangle$ to $|0\rangle$ by the following heuristic argument. Let us expand $|\Omega\rangle$ in terms of energy eigenstates $|n\rangle$ of the interaction Hamiltonian $\hat{H}_{int}$

$$e^{-i\hat{H}_{int}(\tau - \tau_i)} \, |\Omega\rangle = \sum_n e^{-i\hat{H}_{int}(\tau - \tau_i)} \, |n\rangle \, \langle n | \Omega \rangle \tag{3.23}$$

$$= e^{-iE_0(\tau - \tau_i)} \, |0\rangle \, \langle 0 | \Omega \rangle + \sum_{n \neq 0} e^{-iE_n(\tau - \tau_i)} \, |n\rangle \, \langle n | \Omega \rangle \ . \tag{3.24}$$

For $\tau_i \to -\infty$, we want $|\Omega\rangle \to |0\rangle$ and so all the terms in the sum over $n \neq 0$ must drop out. To achieve this, we choose to add to τ a small and negative imaginary part

$$\tau \to \tau(1 - i\epsilon) \, , \tag{3.25}$$

where $0 < \epsilon \ll 1$ is some real number (not the homonimous slow-roll parameter). The expression $e^{-iE_n(\tau - \tau_i)}$ then acquires a factor $e^{-\epsilon E_n(\tau - \tau_i)}$. All states with energy larger than E_0 are then exponentially suppressed in the limit $\tau_i \to -\infty$ and we recover the result. A similar argument applies to the conjugate of this expression relating $\langle\Omega|$ to $\langle 0|$. In this case one finds that the opposite sign is needed for the $i\epsilon$ shift to project onto the vacuum of the free theory. From now on we will therefore always assume that the time integral in the evolution operator has been slightly rotated

$$U_I(\tau, -\infty) = T \exp\left(-i \int_{-\infty(1-i\epsilon)}^{\tau} d\tau' \hat{H}_{int}(\tau')\right) , \tag{3.26}$$

$$U_I^\dagger(\tau, -\infty) = \bar{T} \exp\left(i \int_{-\infty(1+i\epsilon)}^{\tau} d\tau' \hat{H}_{int}(\tau')\right) , \tag{3.27}$$

so that we can write

$$\langle \mathcal{O} \rangle = \langle\Omega| U_I^\dagger(\tau, -\infty) \mathcal{O}_I(\tau) U_I(\tau, -\infty) |\Omega\rangle \tag{3.28}$$

$$= \langle 0| U_I^\dagger(\tau, -\infty) \mathcal{O}_I(\tau) U_I(\tau, -\infty) |0\rangle \, |\langle 0|\Omega\rangle|^2 . \tag{3.29}$$

But taking the expectation value of the unit operator $\mathcal{O} = \mathbb{I}$, we find that

$$|\langle 0|\Omega\rangle|^2 = \frac{\langle\Omega|\Omega\rangle}{\langle 0|0\rangle} = 1 . \tag{3.30}$$

We come therefore to our final formula for correlators

$$\langle \mathcal{O}(\tau) \rangle = \langle 0| \left[\bar{T} e^{\left(i \int_{-\infty(1+i\epsilon)}^{\tau} d\tau' \hat{H}_{int}(\tau')\right)}\right] \mathcal{O}(\tau) \left[T e^{\left(-i \int_{-\infty(1-i\epsilon)}^{\tau} d\tau' \hat{H}_{int}(\tau')\right)}\right] |0\rangle , \tag{3.31}$$

where all fields appearing in $\mathcal{O}(\tau)$ and $\hat{H}_{int}$ are the free fields we introduced in Section 2. There is an equivalent version of this formula that is more useful when performing perturbative calculations [47]:

$$\langle \mathcal{O}(\tau) \rangle = \sum_{N=0}^{\infty} i^N \int_{-\infty}^{\tau} d\tau_N \int_{-\infty}^{\tau_N} d\tau_{N-1} \dots \int_{-\infty}^{\tau_2} d\tau_1 \tag{3.32}$$
$$\times \langle 0| [\hat{H}_{int}(\tau_1), [\hat{H}_{int}(\tau_2), \dots [\hat{H}_{int}(\tau_N), \mathcal{O}(\tau)] \dots]] |0\rangle .$$

Sometimes people refer to (3.31) as the *factorized form* and to (3.32) as the *commutator form*. To prove that (3.31) and (3.32) are indeed equivalent, we proceed by induction. To zeroth and first order in $\hat{H}_{int}$ they obviously agree. Starting from (3.31) we expand

$$\langle \mathcal{O}(\tau) \rangle_{0^{\text{th}}} = \langle 0| \mathcal{O}(\tau) |0\rangle , \tag{3.33}$$

$$\langle \mathcal{O}(\tau) \rangle_{1^{\text{st}}} = \langle 0| \left[i \int \hat{H}_{int}(\tau') d\tau'\right] \mathcal{O}(\tau) + \mathcal{O}(\tau) \left[-i \int \hat{H}_{int}(\tau') d\tau'\right] |0\rangle \tag{3.34}$$

$$= i \int_{-\infty}^{\tau} \langle 0| [\hat{H}_{int}(\tau'), \mathcal{O}(\tau)] |0\rangle . \tag{3.35}$$

Now assume (3.31) and (3.32) give the same result up to order $(N-1)$. Then take the time derivative of each expression at order N. They can be re-written as the expectation value of some other operator at order $(N-1)$ and so they must agree up to a constant. Since they both give the same result for $\tau \to -\infty$, at arbitrary order, the constant must be zero. You will go through the details in Example Sheet 1. Notice that terms coming from U and $U^\dagger$ combine to form the commutation in (3.32). So one has to be careful in keeping track of the correct $i\epsilon$ prescription to project Ω onto $|0\rangle$.

As an aside, there are two other formalisms to compute correlators that are useful in different applications. One is the path integral or Schwinger-Keldysh formalism, in which the correlator is expressed as a path integral from some initial time to the time at which the operators are evaluated and back to the initial time (see e.g. [10]). The second is the Schrödinger picture of quantum mechanics, where the wave function is a functional of the fields and it is often referred to as the *wave function* of the universe (see e.g. [2, 22, 29]).

3.3 Examples: Contact correlators and Wick's theorem

We just introduced a very general formalism to compute cosmological correlators. Let's see it in action for the simple example of de Sitter spacetime with a scalar field. This will turn out to be a good approximation of realistic inflationary models. For the moment the metric will be fixed, non-dynamical, so we are neglecting the effect of the scalar field perturbations on the geometry. We will amend this in Section 5. The simplest calculation to perform are *contact interactions*, which contribute to correlators already at linear order in H_{int}. An example is depicted on the left-hand side of Figure 4.

Example: cubic interaction The simplest interaction one can think of in particle physics is a cubic potential term $V = \mu\varphi(x)^3$. For this term, we can write the Hamiltonian as

$$H_{\rm int} = -L_{\rm int} = \int d^3x\,\sqrt{-g}\mu\varphi(\mathbf{x},\tau)^3 \tag{3.36}$$

$$= a^4\mu\int_{\mathbf{q_1 q_2 q_3}} \varphi(\mathbf{q_1},\tau)\varphi(\mathbf{q_2},\tau)\varphi(\mathbf{q_3},\tau)\,(2\pi)^3\delta_D^3\,(\mathbf{q_1}+\mathbf{q_2}+\mathbf{q_3})\,. \tag{3.37}$$

where

$$\varphi(\mathbf{q},\tau) = f_q(\tau)a_{\mathbf{q}} + f_q^*(\tau)a^\dagger_{-\mathbf{q}}\,, \tag{3.38}$$

$$f_q(\tau) = \frac{H}{\sqrt{2q^3}}(1+iq\tau)e^{-iq\tau}\,. \tag{3.39}$$

This interaction induces a non-vanishing three-point correlator or *bispectrum*, as it is often called. We use the commutator-form of the in-in formula (3.32) to leading non-trivial order

$$\langle\varphi(\mathbf{k}_1,\tau)\varphi(\mathbf{k}_2,\tau)\varphi(\mathbf{k}_3,\tau)\rangle = i\int_{-\infty}^{\tau} d\tau'\langle[H_{int}(\tau'),\varphi(\mathbf{k}_1,\tau)\varphi(\mathbf{k}_2,\tau)\varphi(\mathbf{k}_3,\tau)]\rangle\,. \tag{3.40}$$

We will eventually be interested in the correlators at late time, $\tau\to 0$. For the time being, I will keep τ general. For any Hermitian operator $\mathcal{O}^\dagger = \mathcal{O}$, we can re-write the commutator as

$$\begin{aligned}\langle[H_{int},\mathcal{O}]\rangle &= \langle H_{int}\mathcal{O}\rangle - \langle\mathcal{O}H_{int}\rangle = \langle H_{int}\mathcal{O}\rangle - \langle\mathcal{O}^\dagger H^\dagger_{int}\rangle\\ &= \langle H_{int}\mathcal{O}\rangle - \langle(H_{int}\mathcal{O})^\dagger\rangle = \langle H_{int}\mathcal{O}\rangle - \langle H_{int}\mathcal{O}\rangle^* = 2i\,\mathrm{Im}\langle H_{int}\mathcal{O}\rangle\,,\end{aligned}$$

where I used that also H_{int} is Hermitian. All equal-time products of fields in real space are Hermitian

$$(\varphi(\mathbf{x}_1)\ldots\varphi(\mathbf{x}_n))^\dagger = \varphi(\mathbf{x}_1)\ldots\varphi(\mathbf{x}_n)\,, \tag{3.41}$$

because $\varphi(\mathbf{x})$ is Hermitian and the equal-time fields commute. By taking the Fourier transform we find

$$(\varphi(\mathbf{k}_1)\dots\varphi(\mathbf{k}_n))^\dagger = \left(\int_{\mathbf{x}_1\dots\mathbf{x}_n} e^{-ik_a x_a}\varphi(\mathbf{x}_1)\dots\varphi(\mathbf{x}_n)\right)^\dagger \tag{3.42}$$

$$= \varphi(-\mathbf{k}_1)\dots\varphi(-\mathbf{k}_n)\,. \tag{3.43}$$

If the theory is symmetric under spatial parity[19], which we will assume in the following, then we can flip the sign of momenta again and find that also the product of field in Fourier space is an Hermitian operator.

Our correlator becomes

$$-2\,\mathrm{Im}\int_{-\infty}^{\tau} d\tau' a^4(\tau')\int_{\mathbf{q}_1\mathbf{q}_2\mathbf{q}_3}(2\pi)^3\delta_D\left(\mathbf{q}_1+\mathbf{q}_2+\mathbf{q}_3\right)\times \tag{3.44}$$
$$\langle\varphi(\mathbf{q}_1,\tau')\varphi(\mathbf{q}_2,\tau')\varphi(\mathbf{q}_3,\tau')\varphi(\mathbf{k}_1,\tau)\varphi(\mathbf{k}_2,\tau)\varphi(\mathbf{k}_3,\tau)\rangle\,,$$

where I dropped the prime in the integration variable. This kind of expressions are most easily computed using (a variant of) *Wick's theorem.* Let us define the *contraction* of two fields as

$$\varphi^\bullet(\mathbf{q},\tau')\varphi^\bullet(\mathbf{k},\tau) = \varphi(\mathbf{q},\tau')\varphi(\mathbf{k},\tau) - :\varphi(\mathbf{q},\tau')\varphi(\mathbf{k},\tau):\,, \tag{3.45}$$

where $:\cdots:$ denotes *normal ordering* (all creation operators to the left of all annihilation operators) and the bullets $^\bullet$ mark the fields to be contracted. Since all normal ordered products vanish inside an expectation value, we find

$$\langle\varphi^\bullet(\mathbf{q},\tau')\varphi^\bullet(\mathbf{k},\tau)\rangle = \langle\varphi(\mathbf{q},\tau')\varphi(\mathbf{k},\tau)\rangle = f_q(\tau')f_k^*(\tau)(2\pi)^3\delta_D^3(\mathbf{q}+\mathbf{k})\,. \tag{3.46}$$

This is the Fourier space *propagator.* Notice that the order matters, namely

$$\langle\varphi(\mathbf{q},\tau')\varphi(\mathbf{k},\tau)\rangle = \langle\varphi(\mathbf{k},\tau)\varphi(\mathbf{q},\tau')\rangle^*\,. \tag{3.47}$$

This propagator is related to but distinct from the Feynman, advanced and retarded propagators. Wick's theorem then states that

$$\prod_{a=1}^{n}\varphi(\mathbf{k}_a,\tau_a) = \sum_{\substack{\text{pairwise}\\ \text{contr's}}} :\prod_a\varphi(\mathbf{k}_a,\tau_a):\,, \tag{3.48}$$

where the sum runs over all possible ways to *pairwise* contract any subset of the fields in the product. Since $\langle:\mathcal{O}:\rangle = 0$, inside an expectation value the only surviving term is that in which all fields have been contracted,

$$\langle\prod_{a=1}^{2n}\varphi_a\rangle = \sum_{\text{perm's}}\left[\langle\varphi_1^\bullet\varphi_2^\bullet\rangle\dots\langle\varphi_{2n-1}^\bullet\varphi_{2n}^\bullet\rangle\right]\,, \tag{3.49}$$

$$= \sum_{\text{perm's}}\left[\langle\varphi_1\varphi_2\rangle\dots\langle\varphi_{2n-1}\varphi_{2n}\rangle\right]\,, \tag{3.50}$$

where I used the shorthand notation $\varphi(\mathbf{k}_a,\tau_a) = \varphi_a$.

[19]Notice that specifically for the three-point correlator, invariance under rotations implies parity. This is because all three vectors must lie on a plane, which can be rotated by 180° to invert all vectors.

Using Wick's theorem in (3.44) we have in principle many possible pairs to contract. But here we are only interested in those contractions between one φ in $H_{int}(\tau')$ and one in $\varphi^3(\tau)$. The sum of all and only such contractions is called a *connected* correlator. We will discuss connected correlators in full generality later on. There are only 3! terms contributing to the connected correlator and they all give the same result, so we pick up a factor of 6. From (3.44), our correlator becomes

$$-2\times 3!\times \mu\,\mathrm{Im}\left[\prod_{a=1}^{3} f^*_{k_a}(\tau)\right]\int_{-\infty}^{\tau} d\tau'\,\frac{1}{H^4\tau'^4}\,f_{k_1}(\tau')f_{k_2}(\tau')f_{k_3}(\tau') \tag{3.51}$$

$$=-\frac{3}{2}\frac{\mu H^2}{(k_1k_2k_3)^3}\,\mathrm{Im}\left[\prod_{a=1}^{3}(1-ik_a\tau)\right]\int_{-\infty}^{\tau}\frac{d\tau'}{\tau'^4}\,e^{-ik_T(\tau'-\tau)}\left[\prod_{a=1}^{3}(1+ik_a\tau')\right], \tag{3.52}$$

where I introduce the "total energy" $k_T=k_1+k_2+k_3$. The integral is a bit complicated. First we notice that, thank to the rotation into the lower complex plain for the anti-time ordered factors, the integral converges at $\tau\to-\infty(1+i\epsilon)$ because of the exponential suppression. The interaction is shutting off in the infinite past, just as we wanted. We can then focus on the upper limit of integration. Upon expanding the product in the integrand one finds integrals of the form

$$\int_{-\infty}^{\tau}\frac{d\tau'}{\tau'^n}\,e^{-ik_T\tau'}, \tag{3.53}$$

for $n=1,2,3,4$. The strategy is then to use integration by parts to reduce each term to the exponential integral Ei defined by

$$\mathrm{Ei}(x)=-\int_{-x}^{\infty}\frac{e^{-t}}{t}\,dt\,. \tag{3.54}$$

The result of the integral is

$$\begin{aligned}\int_{-\infty}^{\tau}\frac{d\tau'}{\tau'^4}\,e^{-ik_T\tau'}\left[\prod_{a=1}^{3}(1+ik_a\tau')\right] &= -\frac{i}{3}\sum_{a=1}^{3}\left(k_a^3\right)\mathrm{Ei}(-ik_T\tau)+ \\ &\quad -\frac{e^{-ik_T\tau}}{3\tau^3}\left[1+ik_T\tau+\left(\sum_{a=1}^{3}k_a^2-\sum_{a\neq b}k_ak_b\right)\tau^2\right].\end{aligned} \tag{3.55}$$

Using the asymptotic

$$\mathrm{Ei}(-ik_T\tau)\simeq\gamma_E+\log(k_T\tau)-i\frac{\pi}{2}, \tag{3.56}$$

where $\gamma_E\simeq 0.577$ is the Euler-Mascheroni constant, we can take the $\tau\to0$ limit of (3.52) and find

$$\begin{aligned}\langle\varphi(k_1)\varphi(k_2)\varphi(k_3)\rangle &= (2\pi)^3\delta_D(\mathbf{k}_1+\mathbf{k}_2+\mathbf{k}_3)\frac{\mu H^2}{2k_1^3k_2^3k_3^3} \\ &\quad\times\left[\sum_a k_a^3\left(\gamma_E-1+\ln(-k_T\tau)\right)+k_1k_2k_3-\sum_{a\neq b}k_a^2k_b\right].\end{aligned} \tag{3.57}$$

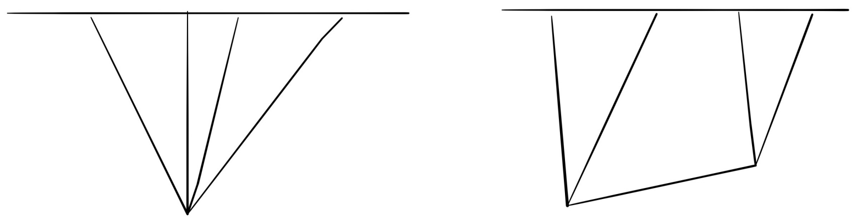

Figure 4: The two diagrams that contribute at tree level to the four-point correlator: the contact diagram on the left-hand side and the exchange diagram on the right-hand side.

Some comments on this result are in order. First, we immediately recognize the ubiquitous momentum-conserving delta functions. It is common to suppress this factor by appending a prime to the correlator or to define B_n as

$$\langle\varphi(\mathbf{k}_1)\dots\varphi(\mathbf{k}_n)\rangle = (2\pi)^3\delta_D^3\left(\sum_{a=1}^n \mathbf{k}_a\right) B_n(\mathbf{k}_1,\dots,\mathbf{k}_n) \tag{3.58}$$

$$\langle\varphi(\mathbf{k}_1)\dots\varphi(\mathbf{k}_n)\rangle' = B_n(\mathbf{k}_1,\dots,\mathbf{k}_n)\,. \tag{3.59}$$

Second, the coupling constant μ unsurprisingly appears linearly. Third we see that B_3 only depends on the norm of the momenta but not on their orientation. This will turn out to be a consequence of rotation and translation invariance. The overall scaling with $k \sim k_a$ is $B_n \sim k^{-6}$. We will soon see that this is a consequence of scale invariance. Fourth, the correlator is fully symmetric under any permutations of $\{k_1, k_2, k_3\}$. Finally, the limit $\tau \to 0$ turned out to be log-divergent! This is one of many divergences that show up in dS spacetime. We will see later on that the gauge invariant observables in the problems are actually finite.

Example: quartic derivative interaction Let us compute another correlator. This time we will choose to compute a four-point function B_4. At tree level, this can be generated by a quartic contact interaction as on the left-hand side of Figure 4 or from two cubic interaction as in the *exchange* diagram on the right-hand side. Let us compute the contribution from the contact interaction, which we assume to come from

$$H_{int} = \int_{\mathbf{x}} a^4\,\frac{1}{4!\Lambda^4}\left(\partial_\tau\varphi g^{\tau\tau}\partial_\tau\varphi\right)^2 \tag{3.60}$$

$$= \int_{\mathbf{q}_1\dots\mathbf{q}_4}\frac{1}{4!\Lambda^4}\prod_{a=1}^4\varphi'(\mathbf{q}_a)\delta_D^3\left(\sum_{a=1}^4\mathbf{q_a}\right)\,. \tag{3.61}$$

where $\varphi' = \partial_\tau\varphi$, Λ is a coupling constant with dimension of mass and the 4! is for later convenience. Using the same trick as in (3.41), The four-point correlator is

$$B_4 = -\frac{2}{4!\Lambda^4}\,\mathrm{Im}\int_{-\infty}^{\tau}d\tau'\int_{\mathbf{q}_1\dots\mathbf{q}_4}\langle\left[\prod_{a=1}^4\varphi'(\mathbf{q}_a,\tau')\right]\left[\prod_{a=1}^4\varphi(\mathbf{k}_a,\tau)\right]\rangle\,. \tag{3.62}$$

From the dS mode functions, (2.19), we find

$$f'_k(\tau) = \frac{H}{\sqrt{2k^3}}\partial_\tau\left[(1+ik\tau)e^{-ik\tau}\right] \tag{3.63}$$

$$= \frac{H}{\sqrt{2k^3}}k^2\tau e^{-ik\tau}\,. \tag{3.64}$$

The relevant propagators now are

$$\langle\varphi'(\mathbf{q},\tau')\varphi(\mathbf{k},\tau)\rangle = f'_q(\tau')f^*_k(\tau)(2\pi)^3\delta^3_D(\mathbf{q}+\mathbf{k})\,. \tag{3.65}$$

The correlator becomes

$$B_4 = -\frac{2\times 4!}{4!\Lambda^4}\times \mathrm{Im}\left[\prod_{a=1}^{4} f_{k_a}(\tau)\right]\int_{-\infty}^{\tau} d\tau'\,\left[\prod_{a=1}^{4} f'_{q_a}(\tau')\right]\,. \tag{3.66}$$

The integral can again be reduced to an exponential integral, but it is now completely finite (again thanks to the $i\epsilon$ rotation of the past infinite boundary). The master integral is

$$\lim_{\tau\to 0}\int_{-\infty(1+i\epsilon)}^{\tau} d\tau' e^{-ik_T\tau'}\left(\tau'\right)^p = -\lim_{\tau\to 0}\eta^{1+p}\,\mathrm{Ei}[-p, ik_T\eta] \tag{3.67}$$

$$= -\frac{(-i)^{p+1}p!}{k_T^{p+1}} \quad \text{for } p\geq 0\,, \tag{3.68}$$

which we will use for $p=4$. Because this has no divergent terms, we can simply take the leading term from $f_k(\tau)$

$$\lim_{\tau\to 0} f_k(\tau) = \frac{H}{\sqrt{2k^3}}\,. \tag{3.69}$$

Finally we find

$$B_4(\tau\to 0) = -\frac{2}{\Lambda^4}\frac{H^8}{2^4(k_1k_2k_3k_4)^{3-2}}\,\mathrm{Im}\left[-\frac{24\,(-i)^5}{k_T^5}\right] \tag{3.70}$$

$$= -\frac{3H^8}{\Lambda^4}\frac{1}{k_T^5k_1k_2k_3k_4}\,. \tag{3.71}$$

Notice that the overall scaling is now $B_4\sim k^{-9}$.

4 Correlators from $P(X,\phi)$ theories

In the previous section we learned how to compute cosmological correlators and went through two examples in detail. Let us now apply these results to inflation. First, we will need to expand in perturbations the class of $P(X,\phi)$ theories introduced in Section 1.4 and then use the in-in formalism to compute correlators of the corresponding interactions.

4.1 $P(X,\phi)$ at quadratic order and the speed of sound

Let us assume we have found some solution to the background equations of motion for some $P(X,\phi)$ theory. We will allow for perturbations around the background solution $\bar{\phi}(t)$ by writing

$$\phi(\mathbf{x},t) = \bar{\phi}(t) + \varphi(\mathbf{x},t)\,, \tag{4.1}$$

and treating $\varphi \ll \bar{\phi}$ perturbatively. Let us expand the Lagrangian in φ:

$$\begin{aligned} L &= P(\bar{X}+\delta X, \bar{\phi}+\varphi) && (4.2)\\ &= P + P_{,\phi}\varphi + P_{,X}\delta X + \frac{1}{2}\left[P_{,XX}\delta X^2 + 2P_{,X\phi}\delta X\varphi + P_{,\phi\phi}\varphi^2\right] + \dots\,, && (4.3) \end{aligned}$$

where P and its derivatives are evaluated on the background $P = P(\bar{X},\bar{\phi})$ and we defined

$$\delta X = X - \bar{X} = \dot{\bar{\phi}}\dot{\varphi} - \frac{1}{2}\partial_\mu\varphi\partial^\mu\varphi\,. \tag{4.4}$$

We are going to massage this ugly looking Lagrangian into something nice that we studied before, namely (2.2). The coefficient of the terms linear in φ is nothing but the background equations of motion up to a total derivative. Since these are satisfied by assumption, we can focus directly on the quadratic terms

$$L_2 = -\frac{P_{,X}}{2}\partial_\mu\varphi\partial^\mu\varphi + \frac{1}{2}\left[P_{,XX}\left(\dot{\bar{\phi}}\dot{\varphi}\right)^2 + 2P_{,X\phi}\dot{\bar{\phi}}\dot{\varphi}\varphi + P_{,\phi\phi}\varphi^2\right]\,. \tag{4.5}$$

The term $\dot{\varphi}\varphi$ can be integrated by part in the action into a φ^2 term. Collecting all terms one finds

$$S_2 = \int d^3x dt\, a^3\, \frac{1}{2}\left[\left(P_{,X} + 2P_{,XX}\bar{X}\right)\dot{\varphi}^2 - P_{,X}\partial_i\varphi\partial^i\varphi - m^2\varphi^2\right]\,, \tag{4.6}$$

where

$$m^2 = 3HP_{,X\phi}\dot{\bar{\phi}} + \partial_t\left(P_{,X\phi}\dot{\bar{\phi}}\right) - P_{,\phi\phi}\,, \tag{4.7}$$

and I used $\dot{\bar{\phi}} = 2\bar{X}$. Despite the ugly coefficients, the action (4.6) has the same terms as that for the massive scalar field we studied in Section 2.2. As discussed there, in cosmology we are interested in almost massless scalar fields, whose correlation functions survive until late times. So we will assume that the mass term is negligible as compared to the others and drop it henceforth. This step can be justified rigorously. It can be shown that all background quantities involving derivatives with respect to ϕ, e.g. $P_{,X\phi}$ or $P_{,\phi\phi}$, are suppressed by slow-roll parameters. The algebra is long and tedious in general, but it's quite simple if we look at a specific model $P = X - V$. Then, the Friedman equations become

$$\begin{aligned} 3M_{\rm Pl}^2H^2 &= 2XP_{,X} - P = X + V && (4.8)\\ -M_{\rm Pl}^2\dot{H} &= XP_{,X} = X\,. && (4.9) \end{aligned}$$

Combining them we find

$$\Rightarrow \quad V = H^2M_{\rm Pl}^2(3-\epsilon)\,. \tag{4.10}$$

Taking a time derivative on each side and using the chain rule on the left-hand side, $\partial_t = \dot{\phi}\partial_t$ we find

$$\begin{aligned}
V'(\bar{\phi}) &= M_{\rm Pl}H^2\left[-\sqrt{\frac{\varepsilon}{2}}\eta - 3\sqrt{2\varepsilon} + \sqrt{2\varepsilon}\varepsilon\right] , &(4.11)\\
V''(\bar{\phi}) &= H^2\left[-\frac{3}{2}\eta + \frac{5}{2}\varepsilon\eta - \frac{1}{4}\eta^2 - \frac{1}{2}\frac{\dot{\eta}}{H} - 2\varepsilon^2 + 6\varepsilon\right] , &(4.12)\\
V'''(\bar{\phi}) &= \frac{H^2}{\sqrt{2\varepsilon}M_{\rm Pl}}\left[-\frac{3}{2}\frac{\dot{\eta}}{H} - \frac{\ddot{\eta}}{2H^2} - \frac{\eta\dot{\eta}}{2H} + 9\varepsilon\eta + 3\frac{\varepsilon\dot{\eta}}{H} + 3\varepsilon\eta^2 - 9\varepsilon^2\eta + 4\varepsilon^3 - 12\varepsilon^2\right] .
\end{aligned}$$

Here we see that all ϕ derivatives of P are suppressed by one or more slow-roll parameters and are therefore negligible to leading order. Later on, I will give an argument of why this is the case using gauge transformations.

After neglecting the mass term, we focus on the two remaining terms and rewrite the action as

$$S_2 \simeq \int d^3x dt\, a^3\, \frac{1}{2}P_{,X}\left[\frac{\left(P_{,X} + 2P_{,XX}\bar{X}\right)}{P_{,X}}\dot{\varphi}^2 - \partial_i\varphi\partial^i\varphi\right] . \tag{4.13}$$

We can get rid of the overall factor by rescaling φ into a canonically normalized φ_c,

$$\varphi_c = \sqrt{P_{,X}}\varphi\,. \tag{4.14}$$

This generates some other mass term when the time derivatives act on P, but we drop those as well. But the relative coefficient between $\dot{\varphi}^2$ and $\partial_i\varphi^2$ cannot be removed. What is it then? By dimensional analysis it must have dimension length^2/time^2, which looks just like a speed squared. Indeed, in particle physics this coefficient is the speed of light squared, c^2, which we usually set to one. So, let us call it *speed of sound* and give it a symbol,

$$c_s^2 = \frac{P_{,X}}{\left(P_{,X} + 2P_{,XX}\bar{X}\right)} , \tag{4.15}$$

and see what it does. From the now much better looking action

$$S_2 \simeq \int d^3x dt\, a^3\, \frac{1}{2}\left[c_s^{-2}\dot{\varphi}_c^2 - \partial_i\varphi\partial^i\varphi\right] , \tag{4.16}$$

we derive the classical equations of motion

$$\ddot{\varphi} + 3H\dot{\varphi} - \frac{c_s^2}{a^2}\partial_i\partial_i\varphi = 0\,. \tag{4.17}$$

We know how to solve this equation when $c_s k/a \gg H$, as it reduces to a wave equation with solution

$$\ddot{\varphi} - \frac{c_s^2}{a^2}\partial_i\partial_i\varphi \simeq 0 \quad \Rightarrow \quad \varphi(x) \sim e^{\pm i c_s k_p t - i\mathbf{k}_p\mathbf{x}}\,. \tag{4.18}$$

This shows that the dispersion relation is $\omega^2 = c_s^2 k_p^2$ and so c_s indeed describe the velocity at which perturbations φ propagate on the $\bar{\phi}(t)$ background.

What is the power spectrum of a massless field with a speed of sound c_s? We could derive it simply by the same steps used for the $c_s = 1$ case in Section 2.1. The only difference would be that the mode functions are now

$$f_k = \frac{H}{\sqrt{2c_s k^3}}(1 + ic_s k\tau)e^{-ic_s k\tau} . \tag{4.19}$$

Notice that $f_k(\tau)$ and so φ stops oscillating at some time $-c_s k\tau \sim 1$, and freezes out. This is often referred to as the *crossing of the sound-horizon*, where $c_s/H)$ is the sound horizon. For small speed of sound, $c_s \ll 1$, this happens much earlier than the crossing of the Hubble radius H^{-1}, discussed around (2.20).

But since c_s is the only constant with dimension of velocity in the action (4.16), it is instructive to use dimensional analysis. We will still use $\hbar = 1$, but we should carefully account for the distinction between space and time $c_s T \sim L$, momentum and energy $c_s k \sim E$:

$$[S] = 1 \quad \Rightarrow \quad [\phi(\vec{x})] = L^{-1/2}T^{-1/2} \quad \Rightarrow \quad [\phi(\mathbf{k})] = L^3[\phi(\mathbf{x})] = L^{5/2}T^{-1/2} . \tag{4.20}$$

The power spectrum has units

$$[P(k)] = \frac{[\langle\phi(\mathbf{k})^2\rangle]}{[\delta_D^3(\mathbf{k})]} = \frac{(L^{5/2}T^{-1/2})^2}{L^3} = \frac{L^2}{T} . \tag{4.21}$$

In (2.28) we found the power spectrum for $c_s = 1$. By requiring that it has the right units we find

$$\frac{L^2}{T} = [P(k)] \overset{!}{=} \frac{[c_s^n H^2]}{[k^3]} = \frac{(L/T)^n T^{-2}}{L^{-3}} \quad \Rightarrow \quad n = -1 . \tag{4.22}$$

We conclude that a field with action (4.16) has a power spectrum

$$P(k) = \frac{H^2}{2c_s k^3} . \tag{4.23}$$

4.2 Cubic interactions in $P(X, \phi)$ theories

To know the interactions that appear in a $P(X, \phi)$ theory we have to expand the Lagrangian to cubic order

$$L_3 = P(\bar{X} + \delta X, \bar{\phi} + \varphi)|_3 \tag{4.24}$$

$$= \frac{1}{2}\left[\frac{1}{3}P_{,XXX}\delta X^3 + P_{,XX}\delta X^2 + 2P_{,X\varphi}\delta X\varphi + P_{,X\varphi\varphi}\delta X\varphi^2 \right. \tag{4.25}$$
$$\left. + P_{,XX\varphi}\delta X^2\varphi + \frac{1}{3}P_{,\phi\phi\phi}\varphi^3\right]_3$$

$$= \frac{1}{6}P_{,XXX}\dot{\bar{\phi}}^3\dot{\varphi}^3 - \frac{1}{2}P_{,XX}\dot{\bar{\phi}}\dot{\varphi}(\partial_\mu\varphi)^2 \tag{4.26}$$
$$- \frac{1}{2}P_{,X\varphi}\varphi(\partial_\mu\varphi)^2 + \frac{1}{2}P_{,X\varphi\varphi}\dot{\bar{\phi}}\dot{\varphi}\varphi^2 + \frac{1}{2}P_{,XX\varphi}\dot{\bar{\phi}}^2\dot{\varphi}^2\varphi + \frac{1}{6}P_{,\phi\phi\phi}\varphi^3 .$$

It is a long expression, but it is conceptually very simple. All cubic interactions appear that are allowed by the symmetries. In this case, the only symmetry is rotation invariance, which enforces that every spatial derivative ∂_i is contracted with another spatial derivative. Notice that time and space derivatives appear separately, as opposed to in the Lorentz-invariant combination $(\partial_\mu\phi)^2$, which is familiar from particle physics. We started from the Lorentz invariant Lagrangian $P(X,\phi)$, but we broke boosts and time translations *spontaneously* by choosing a time-dependent vacuum $\phi=\bar\phi(t)$. In fact, we can check that by setting $\dot{\bar\phi}$ to zero, all non-Lorentz invariant operators disappear.

Again we can use the fact that all background terms with ∂_ϕ derivatives are slow-roll suppressed (I'll give an argument later). We will therefore focus on the only two operators that do not have this suppression, namely $\dot\varphi^3$ and $\dot\varphi(\partial_i\varphi)^2$. The operators stand out from other because φ appears always with a derivative, and they are invariant under a shift symmetry $\varphi\to\varphi+$const (see [19, 20] for an in depth discussion). The calculation of the bispectrum induced by these two operators is completely analogous to that of the previous section, with the exception that we should use the mode functions in (4.19). The results are

$$B_{\varphi'^3}=\frac{H^5\left(P_{,XXX}\dot{\bar\phi}^3+3P_{,XX}\dot{\bar\phi}\right)}{2k_1k_2k_3k_T^3}\,, \tag{4.27}$$

$$\begin{aligned}B_{\varphi'(\partial_i\varphi)^2}=-\frac{1}{8}\frac{H^5P_{,XX}\dot{\bar\phi}}{(k_1k_2k_3)^3\,k_T^3}\Bigg[&24\,(k_1k_2k_3)^2-8k_T\,(k_1k_2k_3)\left(\sum_{a<b}k_ak_b\right)\\&-8\,k_T^2\left(\sum_{a<b}k_ak_b\right)^2+22\,k_T^3\,(k_1k_2k_3)-6k_T^4\left(\sum_{a<b}k_ak_b\right)+2k_T^6\Bigg]\,.\end{aligned} \tag{4.28}$$

5 Gravity

So far we have discussed a quantum scalar field on a classical, fixed spacetime. We will now learn how to account for quantum behavior of gravity as well.

5.1 Effective Field Theory

ref

Before we proceed we have to exorcise a demon: quantum gravity. There is no obstacle whatsoever in *perturbatively* quantizing gravity. One can also compute all kind of observables in pretty much the same way as for non-Abelian gauge theories. In fact, we know since the early 70's that pure general relativity is even finite at one loop [43]. When people talk about the difficulties of "quantum gravity" they have in mind a full non-perturbative treatment. That is a very hard problem. String theory per se or via the gauge-gravity duality might be a way forward. Instead, here we will only need to study the theory perturbatively, for small deviations from a classical background. In this regime, there is a very well-defined procedure to compute observable. More generally, gravity fits perfectly well in the paradigm that has dominated field theory research in the past half a century: Effective Field Theory (EFT).

To understand the idea behind EFT, consider a theory with a characteristic scale E_0. For example, for gravity this would be the Planck scale $M_{\rm Pl}$. Suppose that we are interested in making an experiments at some energy E. If there is a *separation of scales* such that $E\ll E_0$, we can dramatically simplify our description of the system. For these lecture notes E is the Hubble scale H during inflation, which is bounded by the non-observations of gravitational wave to $H<10^{-5}M_{\rm Pl}$. The idea of EFT goes as follows. Choose a *cutoff* Λ well above E and close to, but below E_0

$$E \ll \Lambda \lesssim E_0 \,. \tag{5.1}$$

Since the choice of Λ is arbitrary, Λ better cancel out in the final result. Sometimes people use the word "cutoff" to refer to E_0, which is the highest scale that Λ can be pushed to. Now, divide the fields into a low (L) and a high (H) frequency part

$$\phi = \phi_L + \phi_H \,, \tag{5.2}$$

such that ϕ_L vanishes when its frequency is high, $\omega > \Lambda$, while ϕ_H vanishes when its frequency is low $\omega < \Lambda$. Now the full theory can be formulated in terms of a path integral. Imagine being able to perform the path integral over ϕ_H

$$\int \mathcal{D}\phi_H \mathcal{D}\phi_L e^{iS(\phi_L,\phi_H)} = \int \mathcal{D}\phi_L \left(\int \mathcal{D}\phi_H e^{iS(\phi_L,\phi_H)} \right) \tag{5.3}$$

$$\equiv \int \mathcal{D}\phi_L e^{iS_\Lambda(\phi_L)} \,. \tag{5.4}$$

The new quantity $S_\Lambda(\phi_L)$ is known as *Wilsonian effective action*. Observables computed from $S_\Lambda(\phi_L)$ are UV-finite because ϕ_L vanishes at high energies. But what good does this do us if we cannot compute it? The key insight is that we can Taylor expand this unknown functional in the low-frequency fields

$$S_\Lambda(\phi_L) = \int d^4x \sum g_a \mathcal{O}_a \,, \tag{5.5}$$

where g_a are some coupling constants and $\mathcal{O}_a$ are *all possible local operators compatible with the symmetries of the problem*. The $\mathcal{O}$'s are build from products of fields and their derivatives at the same spacetime point[20]. The sum contains an infinite number of terms and it is useful only if there is some regime in which we can truncate it by making a negligible mistake. Remarkably, this is precisely what happens at low energies, $E \ll \Lambda \lesssim E_0$. To see this, we must learn how to compare different operators. We will do this using dimensional analysis. For example, say $\mathcal{O}_a$ has mass dimension Δ_a. Then, since the action is dimensionless (in units $\hbar = 1$), we have

$$[\mathcal{O}_a] = \Delta_a \quad \Rightarrow \quad [g_a] = 4 - \Delta_a \,. \tag{5.6}$$

It is convenient to make the dimension of g_a explicit by redefining $g_a = \Lambda^{4-\Delta_a}\lambda_a$, where λ_a are dimensionless. It is generally expected that λ_a are order one numbers unless there is some approximate symmetry that we failed to account for. The last ingredient we need is that we want S_Λ to define a weakly coupled theory, which, by definition, gives predictions that are close to the free theory. In this case, we can use the free kinetic term to estimate how large the fields and their derivatives are. For example, consider a massless scalar field[21]

$$S_\Lambda = \int d^4x \, \frac{1}{2} \partial_\mu \phi \partial^\mu \phi \,. \tag{5.7}$$

[20]Some non-locality does emerge at space or time distances of order Λ, but at $E \ll \Lambda$ this can be approximated by local interactions

[21]Dimensional analysis becomes much simpler if the free kinetic term does not contain any dimensionful constant, just as in this scalar field action. This can always be achieved by an appropriate rescaling of the fields.

The free action (5.7) fixes the mass dimension of the field to be $[\phi] = (4-2)/2 = 1$. So when probing the theory at energy E, we can estimate $\phi \sim E$ and $\partial_\mu \sim E$. For example, we estimate

$$\phi^n \sim E^n\,, \quad (\partial\phi)^{2n} \sim E^{4n}\,, \quad (\partial^2\phi)^n \sim E^{3n}\,, \tag{5.8}$$

and so on. We are finally in the position to estimate the size of the terms in the infinite sum (5.5)

$$\int d^4x g_a \mathcal{O}_a = \int d^4x \frac{\lambda_a}{\Lambda^{\Delta_a-4}}\mathcal{O}_a \sim \lambda_a \left(\frac{E}{\Lambda}\right)^{\Delta_a-4}\,. \tag{5.9}$$

This is an important result. It tells us that if $\Delta > 4$, then the operator is very small at low energies $E \ll \Lambda$. These are called *irrelevant operators*. Conversely, for $\Delta < 4$ the operator is called *relevant* and indeed becomes large at low energies. *Marginal operators* with $\Delta = 4$ are in between. Their faith depends on whether loop corrections push their dimension above or below four.

Most field theories you can think of are EFT's. For example the Fermi theory of weak interaction is an EFT below $E_0 \sim m_{W,Z} \sim 80$ GeV. The chiral Lagrangian that describes the interaction of pions is an EFT below the confinement scale of QCD, $E_0 \sim$ GeV. The standard model of particle physics, when extended to include neutrino masses is an EFT.

In the old days, theories that include irrelevant operators used to be called "non-renormalizable" theories. This is a misnomer. It comes from the observations that, if an irrelevant operator is present, it can be shown to generate infinitely many other irrelevant operators via loop corrections. Naively one would then need to know/measure the infinitely many coupling constant with ever increasing dimension and the theory seems to be doomed. But now we understand this is not the case! Operators with large dimension give very small correction to low-energy processes. To be more precise, imagine you want to make a prediction for an experiment at energy $E \ll \Lambda$ that has precision δ. For example $\delta = 10^{-2}$ for percent level predictions and so on. By the estimate (5.9), you only need to include operators up to dimension $\Delta_{\max}$ such that

$$\lambda_a \left(\frac{E}{\Lambda}\right)^{\Delta_{\max}-4} < \delta\,. \tag{5.10}$$

There is a finite number of such operators and so you only need a finite number of couplings. In this precise sense there is no problem in renormalizing and EFT (see [36] for a rigorous proof).

5.1.1 Gravity as an Effective Field Theory

ref

So what about gravity? We might recall hearing our nursery friends saying that "General Relativity is non-renormalizable", because there are irrelevant operators. Let us see how this works. The starting point is the Einstein-Hilbert action

$$S = \frac{M_{\text{Pl}}^2}{2}\int d^4x\,\sqrt{-g}R\,. \tag{5.11}$$

Let us expand it in small perturbations around some background $g_{\mu\nu} = \bar{g}_{\mu\nu} + h_{\mu\nu}$. Dropping all indices, the expanded action looks like

$$S = \frac{M_{\text{Pl}}^2}{2}\int d^4x\,\left[\partial h\partial h + h\partial h\partial h + \dots\right]\,. \tag{5.12}$$

where the dots contain terms with more powers of h. To connect to our discussion in the previous section, we need to normalize the free action so that no dimensionful constants appear. This is achieved by

$$h_{\mu\nu} \to \tilde{h}_{\mu\nu} = h_{\mu\nu} M_{\rm Pl} \,. \tag{5.13}$$

In terms of the canonically normalized field $\tilde{h}_{\mu\nu}$, the action looks like

$$S = \frac{1}{2} \int d^4x \left[\partial\tilde{h}\partial\tilde{h} + \frac{1}{M_{\rm Pl}} \tilde{h}\partial\tilde{h}\partial\tilde{h} + \dots \right] , \tag{5.14}$$

which has now the same schematic form as (5.7). We recognize that the second term has dimension five and so it's an irrelevant operator. The infinitely many other terms hidden in the dots have even larger dimension and are even more irrelevant. We also recognize that $M_{\rm Pl}$ plays the role of the cutoff, $\Lambda < E_0 \sim M_{\rm Pl}$. Therefore, we know that we can quantize GR and make predictions at energies well-below the Planck scale. This is just what we will do in the remainder of this section.

In exorcising the demon of quantum gravity, I have made a few simplifying assumptions, some of which have actually important consequences. Let me briefly mention them. A theory might have more than one scale. Estimating the size of operators then requires more care. I was very vague as to how high- and low-frequency fields should be separated. In fact, the frequency itself is not a Lorentz invariant concept. Even worse, in a time dependent background, both energy and momentum are red- or blu-shifted. Finally, there exists important non-perturbative effects, such as for example tunneling in quantum mechanics. These can often be computed within the EFT, but important subtleties arise.

5.2 Constraints from the ADM formalism

As discussed in Section 2.4, General Relativity (GR) has only two degrees of freedom: one left-handed graviton with helicity $h = -2$ and one right-handed graviton with helicity $h = +2$. Four of the ten $g_{\mu\nu}$ components can be removed by a change of coordinates, while four obey constraint equations. Here we will introduce the ADM formalism due to Arnowitt, Deser and Misner, which makes the appearance of constraint equations manifest.

The main idea is to separate the 4-coordinates x^μ into one time t and three spacial coordinates x^i. To this end, let us foliate spacetime with a family of spatial hypersurfaces $\Sigma(t)$ and parameterize these hypersurfaces with x^i and a 3-dimensional metric h_{ij}. Then, we can write the most generic line element as

$$ds^2 = -N^2 dt^2 + h_{ij}(dx^i + N^i dt)(dx^j + N^j dt) \,, \tag{5.15}$$

where we have introduced the *lapse* $N(x)$ and the *shift* $N^i(x)$. The 4-dimensional metric is then decomoposed into time-time, time-space and space-space parts

$$g_{\mu\nu} = \begin{pmatrix} -N^2 + N_i N^i & N_i \\ N_i & h_{ij} \end{pmatrix} , \quad g^{\mu\nu} = \begin{pmatrix} -1/N^2 & N^i/N^2 \\ N^i/N^2 & h^{ij} - N^i N^j/N^2 \end{pmatrix} , \tag{5.16}$$

where spatial indices are lowered and raised with the spatial metric h_{ij}. The determinant of the $g_{\mu\nu}$ takes the simple form

$$\sqrt{-g} = \sqrt{h} N \,. \tag{5.17}$$

We can define the time-like four-vector n_μ that is normal to Σ as

$$n_\mu = (-N, 0, 0, 0)\,, \tag{5.18}$$

so that it is normalized, $n^\mu n_\mu = -1$. Then the *extrinsic curvature* is the change of the spatial components of this vector as one moves along Σ,

$$K_{ij} \equiv n_{i;j} = n_{i,j} - \Gamma^\lambda_{ij} n_\lambda = N\Gamma^0_{ij} = \frac{1}{2} N g^{0\mu} \left(g_{\mu j,i} + g_{i\mu,j} - g_{ij,\mu}\right) \tag{5.19}$$

$$= -\frac{1}{2N}\left(g_{0j,i} + g_{i0,j} - g_{ij,0}\right) + \frac{1}{2N} N_l h^{lm} \left(h_{mj,i} + h_{im,j} - h_{ij,m}\right) \tag{5.20}$$

$$= \frac{1}{2N}\left(\dot{h}_{ij} - N_{i,j} - N_{j,i}\right) + \frac{1}{N}{}^{(3)}\Gamma^l_{i,j} N_l = \frac{1}{2N}\left(\dot{h}_{ij} - {}^{(3)}\nabla_{(i} N_{j)}\right)\,, \tag{5.21}$$

where ${}^{(3)}\Gamma^i$ is the connection for the 3-dimensional metric h_{ij} and ${}^{(3)}\nabla_i$ the related 3-dimensional covariant derivative. The *Gauss-Codazzi equation* relates the 4-dimensional Ricci scalar R to the three dimensional one ${}^{(3)}R$ as

$$R = {}^{(3)}R + \left(K_{ij}K^{ij} - K^2\right) - 2\nabla_\alpha\left(n^\beta \nabla_\beta n^\alpha - n^\alpha \nabla_\beta n^\beta\right)\,. \tag{5.22}$$

Notice that ${}^{(3)}R$ depends on h_{ij} but not on N or N_i. The last term leads to a total derivative in the action and so drops out[22]. This formula allows us to re-write the action in terms of the ADM variables

$$S = \frac{M_{\rm Pl}^2}{2}\int d^4x\sqrt{h}N\left[{}^{(3)}R + K_{ij}K^{ij} - K^2\right]\,. \tag{5.23}$$

We see explicitly that N appears in the action without any derivatives, while N_i appears with spatial but no time derivatives. This remains true as we couple (minimally) a matter sector to gravity. To see this, let's also write the $P(X,\phi)$ action in the ADM formalism:

$$S = \int d^4x\,\sqrt{-g}\,P(X,\phi) = \int d^3x dt\, N\sqrt{h}\,P(X,\phi)\,, \tag{5.24}$$

where of course there are some N and N_i hiding in X. For example

$$\frac{\partial g^{\mu\nu}}{\partial N} = -\frac{2}{N}\left(g^{\mu\nu} - \delta_{\mu i}\delta_{\nu j}h^{ij}\right)\,, \tag{5.25}$$

As we vary the action with respect to $\{N, N^i\}$ we obtain so called *constraint equations*, in which N and N^i appear without time derivatives and h_{ij} with at most first time derivatives:

$$\frac{\delta S}{\delta N} = 0 \qquad \Rightarrow \qquad {}^{(3)}R - K_{ij}K^{ij} + K^2 + \frac{2}{M_{\rm Pl}^2}\left(P - 2XP_{,X}\right) = 0\,, \tag{5.26}$$

$$\frac{\delta S}{\delta N^j} = 0 \qquad \Rightarrow \qquad \nabla_j[K^j_i - \delta^j_i K] + \frac{P_{,X}}{M_{\rm Pl}^2 N}\partial_i\phi\left(N^j\partial_j\phi - \dot{\phi}\right) = 0\,. \tag{5.27}$$

[22] The boundary term cancels the Gibbons-Hawking-York boundary term

5.3 Scalar-Vector-Tensor decomposition

Before solving these complicated equations, let's discuss two important tools that we can employ to simplify them. The first one is the Scalar-Vector-Tensor (SVT) decomposition. The main idea is the usual one: choose your variables according to the symmetries of the problems. Since all FLRW backgrounds are homogeneous and isotropic, it is a good idea to work with objects that transform nicely under spatial rotations and translations. Mathematically, these are the irreducible representations of the ISO(3) isometry group, which can be obtained using the same method of induced representations, which is also used to define particles in particle physics. This more formal discussion is summarized in Section 10.9 of my cosmology notes [32]. Here we follow a more pedestrian approach. Spatial translations are easily diagonalized by working in Fourier space. For rotations, we separate objects with zero, one and two spatial indices and call them rotation-scalars, rotation-vectors and rotation-tensor, respectively. For example, the field perturbation φ has no spatial indices and already transforms as a scalar under rotations. In particular, for $x^i \to x'^{i'} = R_i^{i'} x^i$, it transforms as

$$\varphi'(x') = \varphi(x) . \tag{5.28}$$

The metric perturbation $h_{\mu\nu}$ instead is more complicated. It is a symmetric 4×4 matrix with 10 independent entries. These can be separated into rotation-scalars, rotation-vectors and rotation-tensors with the following definitions

$$h_{i0} = N_i \equiv a^2 \partial_i \psi + N_i^V \tag{5.29}$$

$$h_{ij} \equiv a^2 \left[\delta_{ij} A + \partial_{ij} B + \partial_{(i} C_{j)} + \gamma_{ij} \right] , \tag{5.30}$$

where all the rotation-vectors are also transverse, in the sense that $\partial_i N_i^V = \partial_i C_i = 0$ and the rotation-tensor is both transverse and traceless $\gamma_{ii} = \partial_i \gamma_{ij} = 0$. Let's check that the number of variables matches the 10 independent entries of $h_{\mu\nu}$. We have four rotation-scalars h_{00}, A, B and ψ, accounting for 4 variables; two transverse rotation-vectors C_i and N_i^V, which with their two "polarizations" each account for $2 + 2 = 4$ variables; finally one transverse traceless rotation-tensor γ_{ij} with its two polarizations accounts for the remaining 2 variables. A similar decomposition can be performed for all other variables in the problem, e.g for the energy-momentum tensor, but we will not need this here.

Now the crucial point: rotation-scalars, transverse rotation-vectors and transverse traceless rotation-tensors decouple from each other at linear order, meaning that in solving the equations of motion for one I can set the others to zero. After finding solutions, I can simply add them up. In the rest of these notes, I will drop the word "rotation-" and simply call the various components scalars, vectors or tensors. You should be aware though that the word "scalar" sometimes refers to a Lorentz scalar, such as ϕ, while sometimes is refers to a rotation-scalar, such as φ or h_{00}.

5.4 Gauge transformations

ref

Let's move on to gauge transformations. In GR, we can always perform a coordinate transformation to simplify the equations. Consider the coordinate transformation

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x) , \tag{5.31}$$

for arbitrary $\epsilon^\mu(x)$. We know that $g_{\mu\nu}$ and ϕ transform as a two-tensor and a scalar respectively, namely as

$$\phi'(x') = \phi(x)\,, \quad g'_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x'^{\mu'}}\frac{\partial x^\nu}{\partial x'^{\nu'}} g_{\mu\nu}\,. \tag{5.32}$$

But how do the perturbations $h_{\mu\nu}$ and φ transform? We have the freedom to specify how the background part and the perturbations transform separately, while keeping their sum covariant. A convenient and very common way to resolve this ambiguity is to work with so called *gauge transformations*, in which case the background is kept fixed and all the transformation of the full tensor is attributed to the perturbations. More in detail, the rules are the following

1. Transform the full tensor covariantly, as in (5.32), and keep the background unchanged
2. Drop the prime from the new coordinates
3. Attribute all the transformation to the perturbations

For example for a scalar field $\phi(x) = \bar\phi + \varphi$, one find the transformation $\Delta\varphi$ to be

$$\Delta\varphi \equiv \phi'(x) - \phi(x) = \phi(x-\epsilon) - \phi(x) = -\epsilon^\mu\partial_\mu\phi(x) + \mathcal{O}(\epsilon^2)\,, \tag{5.33}$$

where I used

$$\phi'(x') = \phi(x) \quad \Rightarrow \quad \phi'(x) = \phi(x-\epsilon)\,. \tag{5.34}$$

For a homogeneous background, $\bar\phi(x) = \bar\phi(t)$, to linear order, this simplifies to

$$\Delta\varphi = -\epsilon^0\dot{\bar\phi} \quad \text{(linear order)}\,. \tag{5.35}$$

The same rules apply tensors

$$\begin{aligned}\Delta h_{\mu\nu}(x) &\equiv g'_{\mu\nu}(x) - g_{\mu\nu}(x) \\ &= g'_{\mu\nu}(x') - \epsilon^\lambda\partial_\lambda g_{\mu\nu}(x) - g_{\mu\nu}(x) \\ &= g_{\lambda\kappa}(x)\frac{\partial x^\lambda}{\partial x'^\mu}\frac{\partial x^\kappa}{\partial x'^\nu} - \epsilon^\lambda\partial_\lambda g_{\mu\nu}(x) - g_{\mu\nu}(x) \\ &= -g_{\lambda\mu}\partial_\nu\epsilon^\lambda - g_{\lambda\nu}\partial_\mu\epsilon^\lambda - \epsilon^\lambda\partial_\lambda g_{\mu\nu}(x) \\ &= -\nabla_\mu\epsilon_\nu - \nabla_\nu\epsilon_\mu = -2\nabla_{(\mu}\epsilon_{\nu)}\,.\end{aligned} \tag{5.36}$$

In differential geometry, the above transformation are known as Lie derivatives (up to a sign). How do the SVT components transforms? Using Eq. (5.36) and the SVT decomposition Eq. (5.30), we find the following linear gauge transformations of the SVT components for the metric[23]

$$\begin{aligned}\Delta A &= 2H\epsilon_0\,, \quad \Delta B = -\frac{2}{a^2}\epsilon^S\,, \\ \Delta C_i &= -\frac{1}{a^2}\epsilon_i^V\,, \quad \Delta\gamma_{ij} = 0\,, \quad \Delta h_{00} = -2\delta N = 2\dot\epsilon^0\,, \\ \Delta\psi &= \frac{1}{a^2}\left(-\epsilon_0 - \dot\epsilon^S + 2H\epsilon^S\right)\,, \quad \Delta N_i^V = -\dot\epsilon_i^V + 2H\epsilon_i^V\,,\end{aligned} \tag{5.37}$$

[23]Notice that $\epsilon_0 = -\epsilon^0$.

where we have SVT-decomposed gauge parameter

$$\epsilon^{\mu} = \{\epsilon^0, \partial^i \epsilon^S + \epsilon^i_V\}\,, \tag{5.38}$$

with $\partial_i \epsilon^i_V = 0$. It is important to notice that to derive these transformations from the general transformation of $h_{\mu\nu}$ in (5.36), one needs to use inverse Laplacians and this is a valid step only if we assume that $\epsilon^{\mu}(x)$ vanishes for $|\mathbf{x}| \to \infty$.

Two comments are in order. First, perturbations do *not* transform covariantly, rather they all shift by something linear in ϵ^{μ}. This means that by carefully choosing ϵ^{μ} we can set to zero some of the perturbations. Second, unlike scalars and vectors, tensor perturbations γ_{ij} are gauge invariant to linear order. The intuitive reason is that the gauge parameter ϵ^{μ} has only scalar and a vector components. We now can proceed in two ways. We can work only with gauge-invariant variables, namely specific linear combinations of the perturbations for which the gauge transformations (5.37) cancel. We will do this in Section 5.5.2. Alternatively, we can *fix the gauge*, namely work with a particular set of coordinates. This second approach is quite convenient and we will use it in the following.

5.4.1 Different gauges

Since vectors decay in cosmology, we will neglect them henceforth. The idea of fixing the gauge is to choose coordinates that correspond to the constant hypersurfaces of some of the perturbations, so that those perturbations appear constant. In other words, we can choose ϵ^0 and ϵ^S in Eq. (5.38) in such a way to cancel whatever profile of some of the scalar perturbations, using the transformation properties in (5.37). Notice that the gauge parameters ϵ^{μ} need to vanish at spatial infinity in the same way as the physical perturbations they need to cancel. In this sense these are *small gauge transformations*. Large gauge transformations will be discussed in Section 6.1. There are infinitely many choices of gauge, but only a few are commonly used. Let's discuss two that we will use.

Newtonian gauge Using the gauge transformations in (5.37), we see that

$$\begin{cases} \epsilon^S = a^2 B/2 \\ \epsilon_0 = a^2\psi - \frac{a^2}{2}\dot{B} \end{cases} \Rightarrow \quad \begin{cases} B' = B + \Delta B = B - B = 0 \\ \psi' = \psi + \Delta\psi = \psi - \psi = 0\,. \end{cases} \tag{5.39}$$

In a more compact form, we will simply write the gauge condition as

$$B = 0 \qquad \psi = 0\,. \tag{5.40}$$

Notice that these two conditions determine ϵ^0 and ϵ^S completely, so small scalar gauge transformations are fully fixed by these requirements. The scalar part of the metric has then only diagonal perturbations, namely in h_{00} and h_{ii}. Traditionally these perturbations are called Φ and Ψ and collectively referred to as *Newtonian potentials*. So, with the identification $h_{00} = -2\Phi$ and $A = -2\Psi$, we find [24]

$$ds^2 = -\left(1 + 2\Phi\right) dt^2 + a^2 dx^i dx^j \left[\left(1 - 2\Psi\right)\delta_{ij} + \gamma_{ij}\right] \quad \text{(Newtonian gauge)}\,. \tag{5.41}$$

This is the perturbed metric in Newtonian gauge. This is particularly useful in the study of the formation of Large Scale Structures.

[24] Be aware that this is possibly the least universal convention in physics. You might find references where the definitions of Φ and Ψ as well as their signs are exchanged. Here I follow Weinberg's notation, which differ from Dodelson's notation by $\Phi_W = \Psi_D$ and $\Psi_W = -\Phi_D$.

Spatially-flat gauge In the study inflation, it is sometimes useful to choose coordinates such that the spatial part of the metric is free from any scalar perturbation,

$$A = B = 0\,. \tag{5.42}$$

Then

$$g_{ij} = a^2\left(\delta_{ij} + \gamma_{ij}\right) \quad \text{(flat gauge)}\,, \tag{5.43}$$

which has only tensor perturbations. When tensors are neglected, this is just the metric of *flat* FLRW background, hence the name. Of course $h_{0\mu}$ does not vanish in this gauge and can be written in terms of the lapse and shift. Since these are fixed by constraints, in this gauge there are no dynamical degrees of freedom in the metric. The only dynamical scalar degrees of freedom is φ. We will shortly use this gauge to solve the constraints (5.26).

Comoving gauge Another option, often employed to study inflation, is comoving gauge[25], sometime also called "ζ-gauge":

$$\varphi = 0 \quad \text{and} \quad B = 0\,. \tag{5.44}$$

In this gauge the metric takes the form

$$ds^2 = (-1 - 2\delta N)\,dt^2 + 2N_i dx^i dt + a^2 dx^i dx^j\left[\delta_{ij}(1 + A) + \gamma_{ij}\right]\,. \tag{5.45}$$

It is straightforward to check that $\varphi = 0$ fixes ϵ^0, while ϵ^S is completely fixed by the condition $B = 0$. This gauge was employed by Maldacena in his seminal paper on primordial non-Gaussianity [29].

Synchronous gauge* An alternative choice of gauge makes the temporal scalar part of the metric $h_{0\mu}$ vanish identically, namely one chooses ϵ^0 and ϵ^S such that

$$g_{00} = -1 \qquad\qquad g_{0i} = 0\,. \tag{5.46}$$

The perturbed metric takes the form

$$ds^2 = -dt^2 + a^2 dx^i dx^j\left[\delta_{ij}(1 + A) + \partial_i\partial_j B + \gamma_{ij}\right]\,. \tag{5.47}$$

This gauge is sometimes used in the study of the Cosmic Microwave Background.

5.5 The bispectrum from inflation

We can finally compute the predictions from inflation. We'll do this first in flat gauge and then re-formulate our results in a gauge invariant manner.

[25] This is not the same as comoving orthogonal gauge, where one imposes $\psi = 0$ instead of $B = 0$.

5.5.1 Flat gauge and the decoupling limit

The constraints in (5.26) become manageable in perturbation theory. For these notes, we want to discuss bispectra and so we need to keep up to cubic perturbations in the action. It can be proven[26] that for this purpose it is sufficient to solve the constraints to linear order. To do this, we will work in flat gauge (5.43).

In flat gauge, to linear order in scalar perturbations

$$
{}^{(3)}R \simeq 0\,, \qquad N \simeq 1+\delta N\,, \qquad N_i \simeq a^2\partial_i\psi \tag{5.48}
$$

$$
NK_{ij} \simeq a^2\left(H\delta_{ij}-\partial_i\partial_j\psi\right)\,, \qquad NK \simeq 3H-\partial_i\partial_i\psi\,. \tag{5.49}
$$

Expanding the constraints in (5.26) to linear order one can solve for δN and ψ. Using repeatedly the background equations of motion, (1.60) and (1.61), the solution can be written as

$$
\partial_i\partial_i\psi = -\frac{\epsilon}{c_s^2}\partial_t\left(\frac{H\varphi}{\dot{\phi}}\right)\,, \tag{5.50}
$$

$$
\delta N = \epsilon H\frac{\varphi}{\dot{\phi}}\,. \tag{5.51}
$$

These equations tell us how spacetime is deformed by the presence of the scalar field perturbations φ, to linear order. This so-called *backreaction* is suppressed by $\epsilon \ll 1$. One can keep these terms and substitute them into the action, (5.23) and (5.24), and see what cubic interactions are generated. This was first done by Maldacena in the seminal paper [29]. As one might expect, the size of the interaction is suppressed by ϵ. This leads to small effects that are beyond observational reach, for at least the next century. In these lectures, we neglect these slow-roll suppressed terms. Moreover, if we are only interested in computing the bispectrum of φ, we can also neglect the tensor perturbations in the metric (5.43). The reason is that tensors don't mix with scalars at linear order. Interactions between φ and γ_{ij} only appear in the cubic action (or higher), e.g. in the form $\gamma_{ij}\partial_i\varphi\partial_j\varphi$. But these terms cannot contribute at tree level to the bispectrum of φ.

Let's summarize: all interactions of φ induce by gravity are slow-roll suppressed. The leading interactions are therefore those coming from the scalar action $P(X,\phi)$, which we studied in Section 4.2 on a fixed-spacetime background.

5.5.2 Curvature perturbations

We realized in Section 5.4 that φ is not gauge invariant, and so neither are its power spectrum $P(k)$ in (2.28), nor its bispectrum B_3 in (4.27). In the previous section we showed that $P(k)$ and B_3 are good approximation to the prediction of inflationary provided we work in flat gauge. Now we would like to express our results in a gauge invariant way, so that they can be used or checked in any gauge.

To achieve this, let's introduce a new variable[27]

$$
\mathcal{R} \;\equiv\; \frac{A}{2}-\frac{H}{\dot{\phi}}\varphi\,, \tag{5.52}
$$

[26] To leading order this was noticed in [29]. More generally, the n^{th} order solution of the constraints is sufficient to obtain the $(2n+1)^{\text{th}}$ order action (see App A.3 of [34]).

[27] Unfortunately, different conventions for the names of these variables exists. For example, in [29], $\mathcal{R}$ is called ζ. This has produces a schism in the subsequent literature. A useful summary of the many possible choices in the literature is given in App A of [44].

From the gauge transformations, it is straightforward to check that $\mathcal{R}$ is gauge invariant at linear order. We will refer to $\mathcal{R}$ as *curvature perturbations on comoving hypersurfaces*, because in comoving gauge it looks like a spatial curvature in the metric:

$$g_{ij} = a^2 dx^i \delta_{ij} dx^j (1 + 2\mathcal{R}) \quad \text{(comoving gauge)}\,. \tag{5.53}$$

Since $\mathcal{R}$ is gauge invariant, we can compute it in any gauge we want. In particular, in flat gauge $A = 0$, and so

$$\mathcal{R} = -\frac{H}{\dot{\bar{\phi}}}\varphi \quad \text{(flat gauge)}\,. \tag{5.54}$$

This relation teaches us something interesting. Notice that $H/\dot{\bar{\phi}}$ is a time dependent function. Its time derivative is slow-roll suppressed but it is not zero. This implies that $\mathcal{R}$ and φ cannot be both constant in time. In the next section we will prove a very general theorem stating that $\mathcal{R}$ is constant on superHubble scales and so φ must evolve. This is another important reason why the predictions of the early universe should always be computed in terms of superHubble correlators of $\mathcal{R}$, which are conserved, rather than those of ϕ, which evolve with time away from exact dS. To this end, recall that our scalar becomes canonical with the rescaling $\varphi = \varphi_c/\sqrt{P_{,X}}$. Using this, the power spectrum for a canonical scalar in dS, (2.28), and (5.54), we can compute the power spectrum of $\mathcal{R}$:

$$P_{\mathcal{R}}(k) = \frac{H^2}{\dot{\bar{\phi}}^2 P_{,X}} P_{\varphi_c}(k) = \frac{1}{4\epsilon c_s}\left(\frac{H}{M_{\rm Pl}}\right)^2 \frac{1}{k^3}\,. \tag{5.55}$$

This an important result. It tells us that the measured amplitude of primordial perturbations

$$\Delta^2_{\mathcal{R}} \equiv \frac{k^3 P(k)}{2\pi^2} = 3.047 \pm 0.014 \quad (68\%\ \text{CL}) \tag{5.56}$$

is a measurement of the scale of inflation H in Planck units, divided by ϵ and c_s.

What about the bispectrum of $\mathcal{R}$? You'd be temped to use again the relation (5.54). But recall that $\mathcal{R}$ was gauge invariant only to linear order. For the bispectrum we clearly need something that is gauge invariant to second order. With a bit of work we could find a second-order version of (5.54), but this involves a lot of algebra as we have to recompute all the gauge transformations to second order. Instead, we'll try to be clever. We'll define the gauge-invariant variable $\mathcal{R}$ to be the quantity that in comoving gauge appears in the metric as[28]

$$g_{ij} = a^2 e^{2\mathcal{R}} \delta_{ij} \quad \text{(comoving gauge)}\,. \tag{5.57}$$

In other gauges, $\mathcal{R}$ is given by its value in comoving gauge plus all the terms induced by the gauge transformation to the required order. This agrees with the previous definition (5.54) in comoving gauge, because $\varphi = 0$ and $2\mathcal{R} = A$. The intuitive meaning of the perturbations $\mathcal{R}$ is depicted in Fig 5.

This figure

[28]This assumes we set tensors to zero, $\gamma_{ij} = 0$. Later one we will see that γ_{ij} can be included by the substitution $\delta_{ij} \to \exp(\gamma_{ij})$.

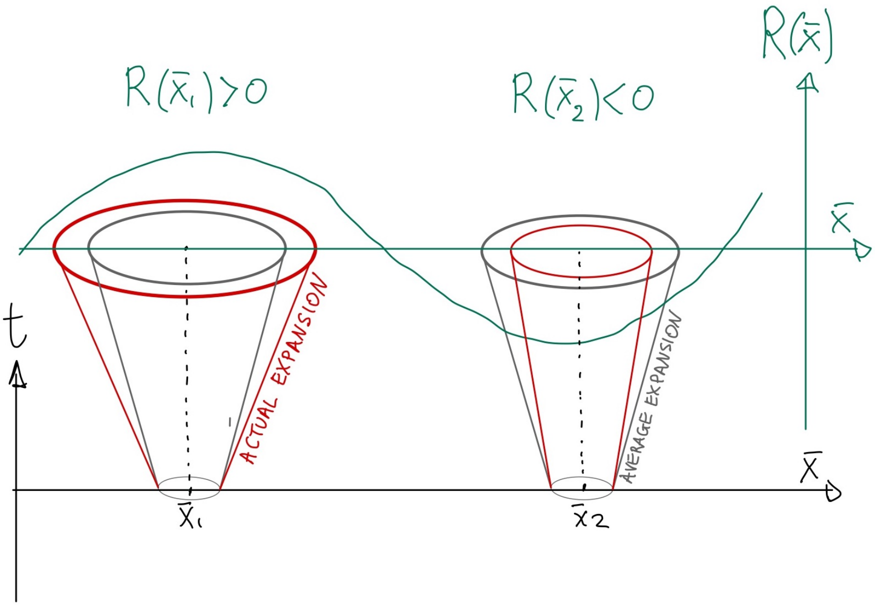

Figure 5: The figure shows the intuitive meaning of $\mathcal{R}(\vec{x})$ perturbations in comoving gauge. The unperturbed homogeneous universe expands in time (along the vertical axis) by the same amount at every point (the "average expansion" with gray lines). But in the perturbed universe, points with different values of $\mathcal{R}(\vec{x})$ experience a larger (smaller) amount of expansion if $\mathcal{R}(x) > 0$ ($\mathcal{R}(x) < 0$) (the "actual expansion" with red lines), as indicated in (5.57).

At second order, the calculation is conceptually the same but it becomes algebraically more involved. The final result is (see app A of [29] for details)

$$\mathcal{R} = -H\frac{\varphi}{\dot{\phi}} + \frac{H\dot{\varphi}\varphi}{\dot{\phi}^2} + \frac{1}{2}\frac{\varphi^2}{\dot{\phi}^2}\left(\dot{H} - \frac{H\ddot{\phi}}{M_{\rm Pl}^2}\right) - \left(1 - \frac{\partial_i\partial_j}{\partial^2}\right)\left[\frac{\partial_i\varphi\partial_i\varphi}{4a^2\dot{\phi}^2} - \frac{1}{2}\frac{\partial_i\psi\partial_i\varphi}{\dot{\phi}}\right].$$

Remarkably, all quadratic terms turn out to be small. The term $\dot{\varphi}\varphi$ decays on superHubble scales because φ becomes approximately constant, up to slow-roll corrections. The term $(\partial_i\varphi)^2$ decays as $a^{-2} \propto \tau^2$. The term $\partial_i\varphi\partial_i\psi$ is slow-roll suppressed because $\partial_i\psi$ is given by (5.50). Finally, the term φ^2 is also slow-roll suppressed because $\dot{H} = -H^2\epsilon$ and $\ddot{\phi} \propto \eta$.

In summary, we can simply use the linear order relation (5.54) and write

$$\langle\mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\mathcal{R}(\mathbf{k}_3)\rangle \simeq -\left(\frac{H}{\dot{\phi}}\right)^3 \langle\varphi(\mathbf{k}_1)\varphi(\mathbf{k}_2)\varphi(\mathbf{k}_3)\rangle + \mathcal{O}(\epsilon,\eta,\dots) \quad \text{for} \quad \tau \to 0\,, \tag{5.58}$$

where the bispectra of φ were given in (4.27). It is interesting to notice that for a canonical inflaton, with $P = X - V$, all scalar interactions in (4.26) vanish except for $P_{,\phi\phi\phi} = -V'''$. This interaction is slow-roll suppressed since $V''' \sim \xi_3$. Since also gravitational interactions are slow-roll suppressed, we conclude that *primordial non-Gaussianity from canonical-field inflation are expected to be very small*, of order $\mathcal{O}(\epsilon,\eta)$. Since we haven't seen any evidence for primordial non-Gaussianity, canonical single-field inflation is still a very successful model for the early universe.

6 Symmetries and soft theorems

So far we focus on computing the correlators directly for a given class of theories. But there is also much we can learn about correlators from the symmetries of the problem alone. This model independent approach leads to some very powerful results. First we will discuss the concept of adiabatic modes and then use it to derive soft theorems that dictate the behavior of cosmological correlators when one of the momenta is soft, namely much smaller than the others.

6.1 Adiabatic modes

We found some gauge invariant predictions for primordial correlations from inflation to leading order in slow roll. To see how these manifest themselves in observables we need to evolve them in time until today. The problem though is that we don't know the constituents of the universe at energies much larger than those probed at colliders, say above 10 TeV. So we don't even know what the right equations to solve are. Luckily for us, under very general conditions $\mathcal{R}$, γ_{ij} and their correlators are conserved in time. This result, which we will prove in this section, is one of the most important in cosmology: It tells us that we can use the sub-eV photons of the CMB to learn something about the laws of physics tens of orders of magnitude higher. This remarkable connection of low-energy observables to high-energy physics has been a tremendous drive for the field of cosmology and has open new possibility to probe fundamental physics.

We are now ready to state an important theorem [46]: *Whatever the constituents of the universe and outside the sound horizon, $c_s k \ll aH$, there is always at least*[29] *one conserved scalar "adiabatic" mode, i.e. $\dot{\mathcal{R}} = 0$ and one conserved tensor mode, i.e. $\dot{\gamma}_{ij} = 0$.* This theorem is valid to all orders in perturbation theory around a flat FLRW spacetime, but we will prove it only at linear order. Also, the theorem applies to gravity coupled to a $P(X, \phi)$ theory, but also holds much more generally. So, in this section we consider a general matter sector, which is described by a generic energy-momentum tensor $T_{\mu\nu}$, with SVT decomposition

$$\begin{aligned}
\delta T_{00} &= -\bar{\rho} h_{00} + \delta\rho\,, \\
\delta T_{i0} &= \bar{p} h_{0i} - (\bar{\rho} + \bar{p})\left[\partial_i \delta u + \delta u_i^V\right]\,, \\
\delta T_{ij} &= \bar{p} h_{ij} + a^2 \left[\delta_{ij}\delta p + \partial_{ij}\pi_{ij}^S + \partial_{(i}\pi_{j)}^V + \pi_{ij}^T\right]\,,
\end{aligned} \tag{6.1}$$

where π^S, π_i^V and π_{ij}^T are known as *anisotropic inertia* and depend on the substance under consideration. For example, all anisotropic inertia vanishes for a scalar field or for a perfect fluid (see (1.11)). In the above decomposition, we recognize four scalars ($\delta\rho$, δp, δu and π^S), two transverse vectors ($\partial_i \pi_i^V = 0 = \partial_i \delta u_i^V$) and one transverse traceless tensor ($\pi_{ii}^T = \partial_i \pi_{ij}^T = 0$), adding up again to 10 components. Notice that we SVT-decomposed the fluid velocity with a *lower* index:

$$u_\mu = \{-1 + \delta u_0, \partial_i \delta u + \delta u_i^V\}\,, \tag{6.2}$$

Our scalar field can be written in this language using the identifications (1.59). We will prove the theorem working in comoving gauge, as in [12, 23, 33]. A derivation in

[29] Actually one can prove the existence of many other decaying modes, including vector modes, see [12, 23, 33] for details.

Newtonian and synchronous gauge was presented in the original paper [46]. In this context, comoving gauge is defined by the gauge condition

$$\delta u = 0\,, \qquad\qquad B = 0\,, \tag{6.3}$$

with the metric given in (5.45). With the identification $\delta u = \varphi/\dot{\bar\phi}$, this is coincides with our definition of comoving gauge in Section 5.4.1, namely $\varphi = 0 = B$.

Consider the change of coordinates

$$\epsilon^\mu = \left\{0, \omega_{ij}(t)x^j\right\}\,, \tag{6.4}$$

with ω_{ij} some time-dependent 3×3 symmetric matrix, $\omega_{ij} = \omega_{ji}$. We have chosen $\epsilon^0 = 0$ so that we don't spoil the gauge condition $\delta u = 0$. Notice that ϵ^μ doesn't vanish at spatial infinity. Therefore, the transformation of $h_{\mu\nu}$ is still given by (5.36) but the transformations of the SVT components (5.37) cannot be used, since they were derived under the assumption that ϵ vanishes at spatial infinity. If we start from an unperturbed, flat FLRW universe, after this gauge transformation we find some non-trivial perturbations given by

$$\begin{aligned} h_{00} &= -2\delta N = 0\,, & \mathcal{R} &= \frac{A}{2} = \frac{1}{3}\omega_{ii}\,, \\ N^i &= \partial_i\psi + N^i_V = -\dot\omega_{ij}x^j & \psi &= f(t) - \frac{1}{6}\dot\omega_{kk}x^i x^j\delta_{ij} & (6.5) \\ N^i_V &= -\dot\omega_{<ij>}x^j & \gamma_{ij} &= -2\omega_{<ij>}\,, & (6.6) \end{aligned}$$

where $< .. >$ indicated the symmetric traceless part

$$A_{<ij>} \equiv \frac{1}{2}\left(A_{ij} + A_{ji}\right) - \frac{1}{3}\delta_{ij}A_{kk}\,, \tag{6.7}$$

and $f(t)$ is an arbitrary time-dependent integration constant. All other perturbations $\{\delta\rho, \delta p, \delta u, B, \pi^{S,V,T}\}$ vanish and so does $\delta T_{\mu\nu}$. Notice that since we still have $B = 0 = \delta u$ after the gauge transformation, so we are still in comoving gauge. The above transformations are completely different from those valid for small gauge transformations, Eq. (5.37). For example, the tensor perturbations γ_{ij} now do change. What do the perturbations in (6.6) represent? Since GR is a covariant theory and we started from and unperturbed FLRW, which is a solution of GR, the perturbations in Eq. (6.6) must also be solution. But because ϵ^μ did not vanish at spatial infinity, this solution is an unusual one: perturbations are constant in space and don't vanish at spatial infinity either. This is depicted in the blue line in Fig. 6. In fact, this solution is just an unperturbed FLRW written in silly coordinates!

The clever insight of Weinberg is to ask when the above solution can be *extended to a physical solution*, with perturbations that do vanish at spatial infinity and can hence arise dynamically (see red line in Fig 6). To answer this, it's easiest to work in Fourier space, where the perturbations in Eq. (6.6), being all constant or power-law in $\mathbf{x}$, are proportional to $\delta_D(\vec{k})$ or its derivative. In particular they have support only at $\mathbf{k} = 0$. Any physical perturbation must vanish at spatial infinity and so its Fourier transform must be continuous at $\mathbf{k} = 0$. So, any non-vanishing perturbation in Fourier space must have support on $\mathbf{k} \neq 0$ as well. When $\mathbf{k} \neq 0$, we are not guaranteed that Eq. (6.6) is still a solution. We have to check. For those equations of motion that do not have an

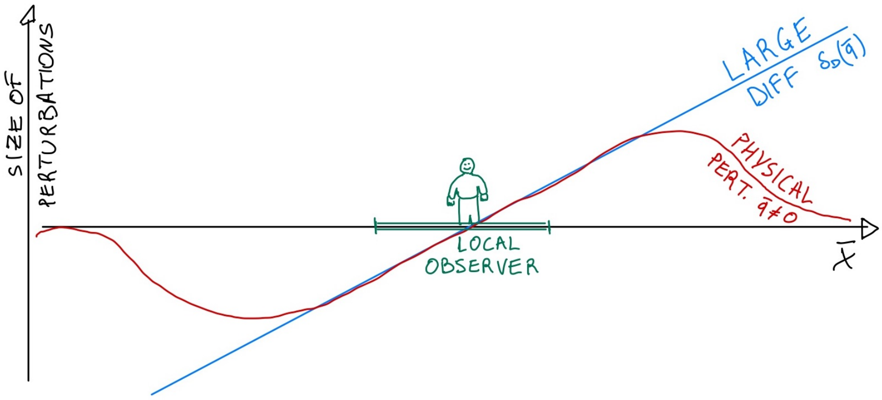

Figure 6: The figure summarized the construction of adiabatic modes. We start from an unperturbed FLRW spacetime a perform a large diff. This generates perturbations to the metric and matter fields (blue line) that solve Einstein's equations but do not vanish at spatial infinity. We can find some physical modes (red line) that locally, i.e. nearby some observer's location, look the same as the large diff but do vanish at $|\vec{x}| \to \infty$. These are adiabatic modes, namely physical solutions that, to the best of my drawing abilities, are locally indistinguishable from a change of coordinates.

overall factor of $\mathbf{k}$, Eq. (6.6) is still a solution up to an arbitrary small correction. For example, for the tensor perturbations,

$$\ddot{\gamma}_{ij} + 3H\dot{\gamma}_{ij} + \frac{k^2}{a^2}\gamma_{ij} = 0\,. \tag{6.8}$$

The solution with $k^2 = 0$ and with k^2 small but non-vanishing are very similar: they differ only at order $k^2/(Ha)^2$, which can be made arbitrarily small for superHubble perturbations, $k \ll Ha$. Then, by continuity we know that physical solutions $\gamma_{ij}(k,t)$ exist, which in the limit $k \to 0$ look like (6.6), namely are constant in time, up to $\mathcal{O}(k^2)$ corrections. On the other hand, the extension to a physical, non-constant solution can be obstructed when some equations of motion vanish identically for $\mathbf{k} = 0$. For example, the ii- and $0i$-components of Einstein's equation in comoving gauge and Fourier space are

$$k_i k_j \left(\delta N + \mathcal{R} + \dot{\psi} + H\psi\right) = 0\,, \tag{6.9}$$

$$k_i \left(H\delta N - \dot{\mathcal{R}}\right) = 0\,, \tag{6.10}$$

$$k_i \left(\dot{N}_j^V + HN_i^V\right) = 0\,. \tag{6.11}$$

For the solution in (6.6), these equations were trivially solved because $k_i = 0$. But if we want physical solutions with $\mathbf{k} \neq 0$, we need to impose that these equations are non-trivially solved. Let us first focus on the trace part ω_{kk} of ω_{ij}, which appear only in the first two of the above equations. The "physicality condition" that (6.10) is satisfied for $k_i \neq 0$ implies

$$\dot{\mathcal{R}} = H\delta N = 0 \quad \Rightarrow \quad \mathcal{R} = \frac{1}{3}\omega_{kk} = \text{const}\,, \tag{6.12}$$

which in turn gives the solution of (6.9) as

$$\psi = -\frac{\omega_{kk}}{3a}\int^t dt' a(t')\,. \tag{6.13}$$

We conclude that a physical solution with $\mathcal{R}$ non-vanishing and constant must always exist in the superHubble limit as consequence of diffeomorphism invariance. As mentioned previously, this is the reason why we want to give the predictions of inflation in terms of $\mathcal{R}$, rather than other fields, such as φ, which continue to evolve in time.

We can also look at the constraint imposed by (6.11) on the traceless part $\omega_{<ij>}$ of ω_{ij}:

$$\ddot{\omega}_{<ij>} + 3H\dot{\omega}_{<ij>} = 0 \quad \Rightarrow \quad \omega_{<ij>} = \bar{\omega}^{(1)}_{<ij>} + \bar{\omega}^{2}_{<ij>}\int^t \frac{dt'}{a(t)^3}\,. \tag{6.14}$$

where $\bar{\omega}^{(1,2)}_{<ij>}$ are integration constant traceless matrices. Notice that the above equation is precisely the same as that of tensor modes in (6.8) for $k \to 0$. We conclude that, whatever the constituents of the universe, there is always a solution to the equations of motion with a constant, non-vanishing $\gamma_{ij} = \bar{\omega}^{(1)}_{<ij>}$, up to corrections that vanish in the superHubble limit. This solution represent the conservation of superHubble *primordial gravitational waves*. This conservation gives us a unique opportunity to use measurements of the late universe, such the CMB, to probe GR in the early universe and its perturbative quantization. The second solution $\bar{\omega}^{(2)}_{<ij>}$ is also general, but it decays with time, so it is not very relevant observationally.

To summarize, we have demonstrated the existence of adiabatic modes, namely physical solutions (which vanish at spatial infinity) that are locally indistinguishable from a change of coordinates (see Fig 6). In Section 6.4 we will see that the existence of adiabatic modes implies the existence of a symmetry for perturbations, which in turn will lead to soft theorems. Before we get there, let us remind ourselves of the role of symmetry in field theory.

6.2 Symmetry symmetry symmetry

Recall that symmetries in field theory are transformations $\Delta\phi$ of the fields ϕ (used in this section to denote collectively fields of any spin) that leave the action invariant, or equivalently that change the Lagrangian by a total derivative

$$\Delta\mathcal{L} = \partial_\mu F^\mu\,. \tag{6.15}$$

What symmetries do for a living is to take some solution ϕ_{sol} of the dynamics and generate another, different one $\phi'_{sol} = \phi_{sol} + \Delta\phi_{sol}$. If one imposes that two states that differ by a symmetry transformation are the same physical state, i.e. all observables give precisely the same values in both states, then the symmetry is called a *gauge symmetry*. A familiar example is electrodynamics, where A^μ and $A^\mu + \partial_\mu\alpha$ represent the same physical state[30] (with appropriate boundary conditions on α). If ϕ_{sol} and ϕ'_{sol} are physically distinguishable, the transformation is called a *global symmetry*. In the following

[30] Often gauge symmetries have parameters that are functions of spacetime as e.g. $\alpha(x)$ in electrodynamics. But this does not have to be the case in general. For example, consider a quantum mechanical particle on a circle of length L. I can describe the system using $x \in \{0, L\}$ but it is sometimes convenient to use the variable $x \in \{-\infty, +\infty\}$ with the identification $x \approx x + nL$ with $n \in \mathbb{N}$. The transformation $x \to x + nL$ is a gauge symmetry even if n is not time dependent.

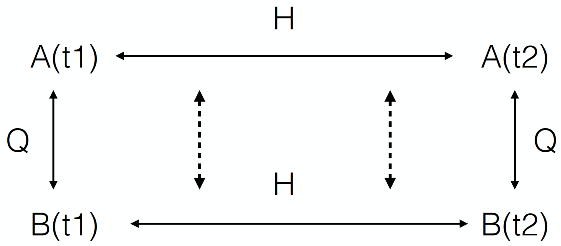

Figure 7: The equivalence of two definitions of symmetry: a transformation that generates new solutions and a transformation that commutes with the Hamiltonian H. Some solution A at time t_1 can be evolved to time t_2 and then transformed by Q into $B(t_2) \neq A(t_2)$. This gives the same result as first transforming to $B(t_1)$ and then evolving because $[Q, H] = 0$. By doing this process at every time t from a solution $A(t)$ one can the generate a new solution $B(t)$.

I'll focus on global symmetries unless otherwise stated.

The fact that Q generates new solutions is equivalent to saying that symmetries commute with the Hamiltonian $[Q, H] = 0$ and so the diagram in Fig. 7 commutes. By Nöther theorem there exists a conserved current

$$J^\mu = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} \Delta\phi - F^\mu \quad \text{with} \quad \partial_\mu J^\mu = 0\,. \tag{6.16}$$

If you wish, you can make things look more covariant by defining $\tilde{J}^\mu \equiv J^\mu (-g)^{-1/2}$ so then $\nabla_\mu \tilde{J}^\mu = (\partial_\mu J^\mu)\,(-g)^{-1/2} = 0$. If the current vanishes sufficiently fast at spatial infinity, then one can define a conserved current Q by

$$Q \equiv \int \sqrt{h} J^\mu n_\mu d^3x\,, \tag{6.17}$$

where n^μ is a time-like vector field that defines some "constant-time" hypersurface over which we integrate. The conservation of J^μ implies $\dot{Q} \equiv n^\mu \partial_\mu Q = 0$. What Q does for a living is to generate the transformations of the fields from which it originally was derived through Nöther theorem:

$$i[Q, \phi] = \Delta\phi\,. \tag{6.18}$$

Since Q is Hermitian, $Q = Q^\dagger$, we can exponentiate this generator to define a unitary symmetry operator

$$\text{Finite unitary transformation:} \quad U \equiv e^{i\alpha Q}\,, \tag{6.19}$$

for some parameter α of the transformation. We say that the symmetry generated by Q is *unbroken* in the state $|\Omega\rangle$ iff

$$\text{Unbroken symmetry:} \quad \langle\Omega|\,[Q, \phi]\,|\Omega\rangle = 0\,, \tag{6.20}$$

otherwise it is spontaneously-broken

$$\text{Spontaneously-broken symmetry:} \quad \langle\Omega|\,[Q, \phi]\,|\Omega\rangle \neq 0\,. \tag{6.21}$$

In words, the laws of nature are invariant under a given symmetry (i.e. $[Q,H]=0$), but the solution of those laws is not ($U|\Omega\rangle \neq |\Omega\rangle$). If Q annihilates $|\Omega\rangle$, namely $Q|\Omega\rangle = 0$, so that $U|\Omega\rangle = |\Omega\rangle$ then Q is unbroken. Conversely, if Q is *spontaneously broken*[31], $|\Omega\rangle$ is not invariant under Q namely $Q|\Omega\rangle \neq 0$. For spontaneously-broken symmetries $\Delta\phi$ must contain a constant term, when expanded in power of ϕ (assuming we are working with fields such that $\langle\phi\rangle = 0$, which is always true up to a field redefinition $\phi \to \phi - \langle\phi\rangle$). So *a spontaneously broken symmetry must always be non-linearly realized*[32]:

$$\text{Non-linearly realized symmetry:} \quad i[Q,\phi] = \Delta\phi = \text{const} + \mathcal{O}(\phi)\,. \tag{6.22}$$

6.3 Correlators and linearly-realized symmetries*

In this section, I'll discuss the observational consequences of linearly realized symmetries in cosmology. I will focus on spacetime symmetries and discuss translations and rotations for FLRW spacetime and then dilations and special conformal transformations for de Sitter spacetime.

6.3.1 FLRW: translations and rotations

If we assume a Lorentz-invariant theory and expand around a flat FLRW background, all primordial correlators must be invariant under translations and rotations. To see this more formally, consider the generators of spatial translations P^i and spatial rotations L^i, acting on scalar[33] operators

$$i[P^i, \phi(\mathbf{x})] = -\partial_i \phi(\mathbf{x})\,, \tag{6.24}$$

$$i[L^i, \phi(\mathbf{x})] = -\epsilon^{ijk} x_j \partial_k \phi(\mathbf{x})\,. \tag{6.25}$$

If these generators commute with the Hamiltonian then the same expressions hold for the Heisenberg operators at any time. These generators exponentiate to finite translations and rotations as in

$$U^{-1}(\vec{\alpha},\vec{\omega})\phi(\mathbf{x})U(\vec{\alpha},\vec{\omega}) = \phi(R^{ij}x^j + \alpha_i)\,, \tag{6.26}$$

with

$$R_{ij} = \exp\left(\epsilon_{ijk}\omega^k\right)\,, \qquad U(\vec{\alpha},\vec{\omega}) = \exp\left(iP^i\alpha_i\right)\exp\left(iL^i\omega_i\right)\,. \tag{6.27}$$

and $U^\dagger U = 1$. Then we see that

$$\langle\Omega|\prod_a \phi(x_a)|\Omega\rangle = \langle\Omega|UU^{-1}\phi(x_1)UU^{-1}\phi(x_2)\ldots\phi(x_n)UU^{-1}|\Omega\rangle \tag{6.28}$$

$$= \langle\Omega|U^{-1}\phi(x_1)UU^{-1}\phi(x_2)\ldots\phi(x_n)U|\Omega\rangle \tag{6.29}$$

$$= \langle\Omega|\prod_a \phi(t_a, R\mathbf{x}_a + \vec{\alpha})|\Omega\rangle\,, \tag{6.30}$$

[31] This should not be confused with *explicit symmetry breaking*, which describes a situation in which the transformation is just not a symmetry anymore.

[32] To avoid confusion, let us stress that the commutator is a linear operation on ϕ and so $[Q,\lambda\phi] = \lambda\Delta\phi$ for any constant λ. By "non-linearly realized" we mean that the transformation acts non-linearly on the solutions of the theory, namely given two solutions $\phi_{sol,1} = \lambda\phi_{sol,2}$ one finds $\Delta\phi_{sol,1} \neq \Delta\phi_{sol,2}$.

[33] For generic operators with spin, the action or rotations is simply replaced by

$$i[L^i, \mathcal{O}^A_S(\mathbf{x})] = -D(L)^A_B \epsilon^{ijk} x_j \partial_k \mathcal{O}^B_S(\mathbf{x})\,, \tag{6.23}$$

where $D(L)^A_B$ is the representation of the algebra $\mathfrak{so}(3)$ relevant for $\mathcal{O}$.

where in the second step I used the invariance of the vacuum and in the last that U commutes with the Hamiltonian. It is useful to re-write this expression as an operator annihilating the correlation function. To this end, we expand (6.30) to linear order in $\vec{\alpha}$ and $\vec{\omega}$ and cancel the zeroth order piece with the left hand side. The remaining term is

$$\sum_{a=1}^{n} \frac{\partial}{\partial \mathbf{x}_a} \langle \phi(x_1)\phi(x_2)\dots\phi(x_n)\rangle \stackrel{!}{=} 0\,, \tag{6.31}$$

$$\sum_{a=1}^{n} \left(x_a^i \frac{\partial}{\partial x_a^j} - x_a^j \frac{\partial}{\partial x_a^i} \right) \langle \phi(x_1)\phi(x_2)\dots\phi(x_n)\rangle \stackrel{!}{=} 0\,. \tag{6.32}$$

These relations must be obeyed by all cosmological correlators. The general solution of the first condition is that the correlator only depends on the distance among points, i.e. only on $n-1$ of the n point appearing. For example, this can be chosen to be $\mathbf{x}_a - \mathbf{x}_1$ for $a = 2, \dots n$. The second condition implies that the correlator must be a function of scalar products $\mathbf{x}_a \cdot \mathbf{x}_b$. The full n-correlator then depends on $3n-3-3$ variables.

It is easier to deal with translation invariance in Fourier space

$$\phi(t,\mathbf{k}) \equiv \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \phi(t,\mathbf{x})\,, \qquad \phi(t,\mathbf{x}) \equiv \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \phi(t,\mathbf{k})\,. \tag{6.33}$$

The generators acting on Fourier space correlators are then

$$P_i : -ik_i \quad \text{and} \quad R_{ij} : -i\left(k_i\partial_j - k_j\partial_i\right)\,, \tag{6.34}$$

and therefore

$$\sum_{a=1}^{n} \mathbf{k}_a \langle \phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\dots\phi(\mathbf{k}_n)\rangle \stackrel{!}{=} 0\,, \tag{6.35}$$

$$\sum_{a=1}^{n} \left(k_a^i \frac{\partial}{\partial k_a^j} - k_a^j \frac{\partial}{\partial k_a^i} \right) \langle \phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\dots\phi(\mathbf{k}_n)\rangle \stackrel{!}{=} 0\,. \tag{6.36}$$

The first condition is satisfied if the correlator is proportional to a Dirac delta function of the sum of all momenta

$$\langle \phi(\mathbf{k}_1)\phi(\mathbf{k}_2)\dots\phi(\mathbf{k}_n)\rangle \equiv (2\pi)^3 \delta_D^3 \left(\sum_a \mathbf{k}_a \right) B_n(\mathbf{k}_1,\dots,\mathbf{k}_n)\,, \tag{6.37}$$

where the prime denotes the *stripped* correlator, i.e. with the delta function and $(2\pi)^3$ removed. The second condition implies again that the correlator only depends on the rotational invariant contractions $\mathbf{k}_a \cdot \mathbf{k}_b$.

6.3.2 De Sitter spacetime: dilations and special conformal transformations

Cosmological observations tell us that primordial perturbations are not only translation and rotation invariant, but also scale invariant. This can be seen for example in the large scale behavior of the CMB temperature anisotropy angular power spectrum C_l^{TT}, where the transfer function is just approximately constant for $l \ll 50$ (the so-called Sachs-Wolfe approximation). On these large scales one finds $C_l^{TT} \propto l^{-2}$, which in angular space implies that the correlation of anisotropies is approximately independent of angle. The leading paradigm to explain such scale invariance is to postulate a phase of quasi-de

Sitter expansion in the very early universe. De Sitter spacetime in flat slicing is given by

$$ds^2 = -dt^2 + e^{2Ht} d\mathbf{x}^2 = \frac{-d\tau^2 + d\mathbf{x}^2}{\tau^2 H^2}\,, \tag{6.38}$$

for some constant Hubble parameter H and with $\tau = -e^{-Ht}/H$. This is a maximally symmetric spacetime with ten isometries, arranged according to the group SO(4,1) (the Lorentz group in (4+1)-dimensions or equivalently the conformal group in 3 euclidean dimension). Besides spatial rotations and translations, de Sitter is also invariant under dilations and special conformal transformations (SCT):

$$\text{dilation:} \quad \tau \to \tau(1+\lambda)\,, \qquad \mathbf{x} \to \mathbf{x}(1+\lambda)\,, \tag{6.39}$$

$$\text{SCT:} \quad \tau \to \tau(1-2\mathbf{b}.\mathbf{x})\,, \qquad \mathbf{x} \to \mathbf{x} - 2(\mathbf{b}.\mathbf{x})\,\mathbf{x} + (\mathbf{x}^2 - \tau^2)\,\mathbf{b}\,, \tag{6.40}$$

If all other non-gravitation background quantities also respect this symmetry, as it is the case for example in the limit of $\dot{H} \ll H^2$, then these additional isometries lead to new symmetry that further constrain cosmological correlators (see e.g. [?, ?, 3, 34]).

By following the same procedure as in the previous section, we can again compute the operators that must annihilate correlation functions as consequence of the full de Sitter isometry group. In real space, these generators are found to be[34]

$$D: \; -\tau\partial_\tau - x^i\partial_i \qquad \text{(dilation)}\,, \tag{6.42}$$

$$\mathbf{b}\cdot\mathbf{K}: \; -2\mathbf{b}\cdot\mathbf{x}\left(\tau\partial_\tau - x^i\partial_i\right) - \left(\tau^2 - |\mathbf{x}|^2\right) b^i\partial_i \qquad \text{(SCT)}\,, \tag{6.43}$$

for an arbitrary constant three-vector $\mathbf{b}$. As before, the sum of D and $\mathbf{K}$ acting on *each* operator in the correlator must vanish by symmetry:

$$\sum_{a=1}^{n} D_a \langle \phi(\tau_1, \mathbf{x}_1)\phi(\tau_2, \mathbf{x}_2)\ldots\phi(\tau_n, \mathbf{x}_n)\rangle \stackrel{!}{=} 0\,, \tag{6.44}$$

$$\sum_{a=1}^{n} \mathbf{b}\cdot\mathbf{K}_a \langle \phi(\tau_1, \mathbf{x}_1)\phi(\tau_2, \mathbf{x}_2)\ldots\phi(\tau_n, \mathbf{x}_n)\rangle \stackrel{!}{=} 0\,. \tag{6.45}$$

The solutions of these equations have been studied for half a century in an attempt to better understand Conformal Field Theories (see e.g. online reviews [31, 39, 40]). For example, the 2 and 3 point functions are completely fixed up to an overall multiplicative constant, while higher n-point functions can only depends on specific invariants called cross ratios.

When acting on a single Fourier-space operator $\phi(\tau, \mathbf{k})$, the generators become (Exercise)

$$D: \quad (3 - \tau\partial_\tau) + k^i\partial_{k^i}\,, \tag{6.46}$$

$$\mathbf{b}\cdot\mathbf{K}: \quad (3 - \tau\partial_\tau)\,2b^i\partial_{k^i} - \mathbf{b}\cdot\mathbf{k}\partial_{k^i}\partial_{k^i} + 2k^i\partial_{k^i}b^j\partial_{k^j}\,. \tag{6.47}$$

[34]Check that indeed $\epsilon^\mu = \{-\tau, -x^i\}$ is a Killing vector for the dS metric in (6.38), namely it solves

$$\mathcal{L}_\epsilon g_{\mu\nu} = -\left(\nabla_\mu\epsilon_\mu - \nabla_\mu\epsilon_\mu\right) = 0\,. \tag{6.41}$$

where $\mathcal{L}$ is the Lie derivative.

It is the combination $-3+\sum_a D_a$ and $\sum_a \mathbf{b}\cdot\mathbf{K}_a$ that annihilates the stripped correlators

$$\left[-3+\sum_{a=1}^{n} D_a\right]\langle\phi(\tau_1,\mathbf{k}_1)\phi(\tau_2,\mathbf{k}_2)\dots\phi(\tau_n,\mathbf{k}_n)\rangle' \stackrel{!}{=} 0\,, \tag{6.48}$$

$$\left[\sum_{a=1}^{n} \mathbf{b}\cdot\mathbf{K}_a\right]\langle\phi(\tau_1,\mathbf{k}_1)\phi(\tau_2,\mathbf{k}_2)\dots\phi(\tau_n,\mathbf{k}_n)\rangle' \stackrel{!}{=} 0\,. \tag{6.49}$$

6.4 Soft theorems **ref**

Soft theorems are constraints on correlators in the limit in which one of the momenta goes to zero, $\mathbf{k}_a \to 0$. In this section we will derive the following soft theorem

$$\lim_{\mathbf{q}\to 0}\langle\mathcal{R}(\mathbf{q})\mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\rangle' = (1-n_s)P_{\mathcal{R}}(k)P_{\mathcal{R}}(q) + \mathcal{O}(q^2)\,. \tag{6.50}$$

for any single-field model of inflation. This results is a consequence of a non-linearly realized symmetry, which is related to adiabatic modes.

In Section 6.1, working in comoving gauge, we found that the change of coordinates (6.51) generates a physical solution with a new $\mathcal{R}$ and γ_{ij}, as long as ω_{ij} is constant. Let us focus on the trace part of ω_{ij}, which is the only one relevant for $\mathcal{R}$. A similar discussion for γ_{ij}can be carried over using the traceless part $\omega_{<ij>}$. So let's drop the out-of-diagonal terms and consider

$$\epsilon^\mu = \left\{0, \lambda x^i\right\}\,, \tag{6.51}$$

which looks like a constant isotropic rescaling. Since this diff maps a physical solution into another physical solution, it implies the existence of a symmetry of the action of cosmological perturbations. We would like to find this symmetry to linear order in ϵ^μ but to all orders in perturbations. To this end, recall

$$\begin{aligned} g_{ij}(x) \to g'_{ij}(x') &= g_{\mu\nu}(x)\frac{\partial x^\mu}{\partial x'^i}\frac{\partial x^\nu}{\partial x'^j} \\ &= g_{\mu\nu}(x)\delta_i^\mu(1-\lambda)\delta_j^\nu(1-\lambda) \\ &= g_{ij}(x)(1-2\lambda) + \mathcal{O}(\lambda^2)\,. \end{aligned} \tag{6.52}$$

Using the form of the spatial metric in comoving gauge we find

$$e^{2\mathcal{R}'(x')} = e^{2\mathcal{R}(x)}(1-2\lambda) \quad \Rightarrow \quad \mathcal{R}'(x) = \mathcal{R}(x) - \lambda x^i\partial_i\mathcal{R}(x) - \lambda\,. \tag{6.53}$$

So in real and Fourier space the symmetry transformation is

$$\Delta\mathcal{R}(\mathbf{x}) = -\lambda - \lambda x^i\partial_i\mathcal{R}(\mathbf{x})\,, \tag{6.54}$$

$$\Delta\mathcal{R}(\mathbf{k}) = -\lambda(2\pi)^3\delta^3(\mathbf{k}) - \lambda\left(3+\mathbf{k}\cdot\partial_{\mathbf{k}}\right)\mathcal{R}(\mathbf{k})\,. \tag{6.55}$$

We already found the shift $-\lambda$ when discussing adiabatic modes, while the linear transformation

$$\Delta_{lin}\mathcal{R}(\mathbf{k}) = -\lambda\left(3+\mathbf{k}\cdot\partial_{\mathbf{k}}\right)\mathcal{R}(\mathbf{k}) \tag{6.56}$$

was neglected there because it has one more perturbation. In the following we will keep both terms. The charge that generates this transformation needs to satisfy (6.18) and so can be written as

$$Q = Q_S + Q_{lin} \tag{6.57}$$

$$Q_S \equiv -\lambda \int d^3x \Pi(t, \mathbf{x}) \,, \tag{6.58}$$

$$Q_{lin} \equiv \frac{1}{2} \int d^3x \{\Pi(t, \mathbf{x}), \Delta_{lin}\mathcal{R}(t, \mathbf{x})\} \,, \tag{6.59}$$

where the parenthesis indicate the anti-commutator and are used to make Q hermitian, while Π is the conjugate momentum of $\mathcal{R}$, namely

$$[\mathcal{R}(t, \mathbf{x}), \Pi(t, \mathbf{y})] = i\delta_D^3(\mathbf{x} - \mathbf{y}) \,. \tag{6.60}$$

By the definition, the charge Q must satisfy

$$i\langle[Q, \mathcal{O}]\rangle = \langle\Delta\mathcal{O}\rangle \,, \tag{6.61}$$

which is known as Ward-Takahashi (WT) identity. Here, $\mathcal{O}$ denotes collectively the product of n curvature perturbations $\mathcal{R}$ and the variation is

$$\mathcal{O} = \prod_{a=1}^{n} \mathcal{R}(\mathbf{k}_a) \quad \Rightarrow \quad \Delta\mathcal{O} = \sum_{a=1}^{n} \mathcal{R}(\mathbf{k}_1) \ldots \Delta\mathcal{R}(\mathbf{k}_a) \ldots \mathcal{R}(\mathbf{k}_n) \,. \tag{6.62}$$

The idea is to compute the left- and right-hand sides of (6.61) in different ways.

The left-hand side On the left-hand side, Q_{lin} has one more perturbation than Q_S and it is higher order. If parity is a symmetry, $\mathcal{O}$ is Hermitian and we can use

$$i\langle[Q, \mathcal{O}]\rangle = 2\mathrm{Im}\langle\mathcal{O}Q\rangle \,. \tag{6.63}$$

We can compute $Q\,|\Omega\rangle$ in perturbation theory, where the free $\mathcal{R}$ and Π fields are given by

$$\mathcal{R}(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \left[a_{\mathbf{k}} f_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}} f^*_k(t) e^{-i\mathbf{k}\cdot\mathbf{x}} \right] \,, \tag{6.64}$$

$$\Pi(\mathbf{x}) = \int \frac{d^3k}{(2\pi)^3} \left[a_{\mathbf{k}} g_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}} g^*_k(t) e^{-i\mathbf{k}\cdot\mathbf{x}} \right] \,, \tag{6.65}$$

with $f_k(t)$ and $g_k(t)$ the solutions of the classical linearized equations of motion. In fact, $g_k(t) = a^3\epsilon(t)\dot{f}_k(t)$ with ϵ the Hubble slow-roll parameter, but we will not need this relation here. The canonical quantization (6.60) fixes the so-called Wronskian

$$f_k g^*_k - f^*_k g_k = i \,. \tag{6.66}$$

We can then write

$$Q_S\,|\Omega\rangle \simeq Q_S\,|0\rangle = \int d^3x \Pi(x)\,|0\rangle = g^*_0(t) a^\dagger_{\mathbf{0}}\,|0\rangle \tag{6.67}$$

$$= \frac{g^*_0(t)}{f^*_0(t)} f^*_0(t) a^\dagger_{\mathbf{0}}\,|0\rangle = \frac{g^*_0(t)}{f^*_0(t)} \mathcal{R}(\mathbf{0})\,|0\rangle \,, \tag{6.68}$$

where $\mathcal{R}(\mathbf{0}) = \mathcal{R}(\mathbf{k} = \mathbf{0})$ is the Fourier space field. So we need to compute

$$i\langle [Q, \mathcal{O}] \rangle = 2\mathrm{Im}\left[\langle \mathcal{O}\mathcal{R}(\mathbf{0}) \rangle \frac{g_0^*(t)}{f_0^*(t)} \right] . \tag{6.69}$$

Since we care about the bispectrum, let us take

$$\mathcal{O} = \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}') , \tag{6.70}$$

where the implicit time argument is $\tau \to 0$. By Hermiticity $\langle \mathcal{O}\mathcal{R} \rangle$ is real:

$$\begin{aligned} \langle \mathcal{O}\mathcal{R}(\mathbf{0}) \rangle^* &= \langle \mathcal{R}^\dagger(\mathbf{0})\mathcal{O}^\dagger \rangle = \langle \mathcal{R}(\mathbf{0})\mathcal{R}(-\mathbf{k}')\mathcal{R}(-\mathbf{k}) \rangle \\ &= \langle \mathcal{R}(\mathbf{0})\mathcal{R}(\mathbf{k}')\mathcal{R}(\mathbf{k}) \rangle = \langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\mathcal{R}(\mathbf{0}) \rangle = \langle \mathcal{O}\mathcal{R}(\mathbf{0}) \rangle , \end{aligned} \tag{6.71}$$

where I used $\mathcal{R}^\dagger(\mathbf{k}) = \mathcal{R}^\dagger(-\mathbf{k})$ and that all equal time $\mathcal{R}$ commute with each other. For the other factor in (6.69) we can use the Wronskian condition

$$\mathrm{Im}\, \frac{g_0^*(t)}{f_0^*(t)} = \frac{\mathrm{Im}\,[g_0^*(t)f_0(t)]}{|f_0(t)|^2} = -\frac{i}{2}\frac{[g_0^*(t)f_0(t) - g_0(t)f_0^*(t)]}{|f_0(t)|^2} \tag{6.72}$$

$$= \frac{1}{2|f_0(t)|^2} = \frac{1}{2P_{\mathcal{R}}(0)} . \tag{6.73}$$

Our calculation of the left-hand side is complete

$$i\langle [Q, \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')] \rangle = \frac{1}{P_{\mathcal{R}}(0)} \langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\mathcal{R}(\mathbf{0}) \rangle . \tag{6.74}$$

The right-hand side The right-hand side of (6.61) depends only on the linear transformation Δ_{lin}. The reason is that we are interested in computing *connected* diagrams, namely diagrams that are proportional to one overall delta function. Instead the shift only contributes to disconnected diagrams, that are proportional to the product of two or more delta functions:

$$\langle \Delta\mathcal{O} \rangle \supset C \sum_{a=1}^{n} \delta_D^3(\mathbf{k}_a) \langle \mathcal{O}(\mathbf{k}_1) \dots \mathcal{O}(\mathbf{k}_{a-1})\mathcal{O}(\mathbf{k}_{a+1}) \dots \mathcal{O}(\mathbf{k}_n) \rangle \propto \delta_D^3(\mathbf{k}_a)\delta_D^3\left(\sum_{b \neq a} \mathbf{k}_b \right) ,$$

So, the right-hand side of the (6.61) with the choice (6.70) becomes

$$\langle \Delta\mathcal{O} \rangle = -\left(3 + \mathbf{k}\cdot\partial_{\mathbf{k}} + 3 + \mathbf{k}'\cdot\partial_{\mathbf{k}'}\right) \langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}') \rangle . \tag{6.75}$$

One can eliminate the Dirac delta function picking up a -3 and express this in terms of the tilt of the power spectrum

$$\langle \Delta\mathcal{O} \rangle' = -\left(3 + k\partial_k\right) P_{\mathcal{R}}(k) = (1 - n_s)P_{\mathcal{R}}(k) . \tag{6.76}$$

We conclude with the WT identity in its final form

$$\lim_{\mathbf{q}\to 0} \langle \mathcal{R}(\mathbf{q})\mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}') \rangle' = (1 - n_s)P_{\mathcal{R}}(k)P_{\mathcal{R}}(q) + \mathcal{O}(q^2) . \tag{6.77}$$

A few comments are in order:

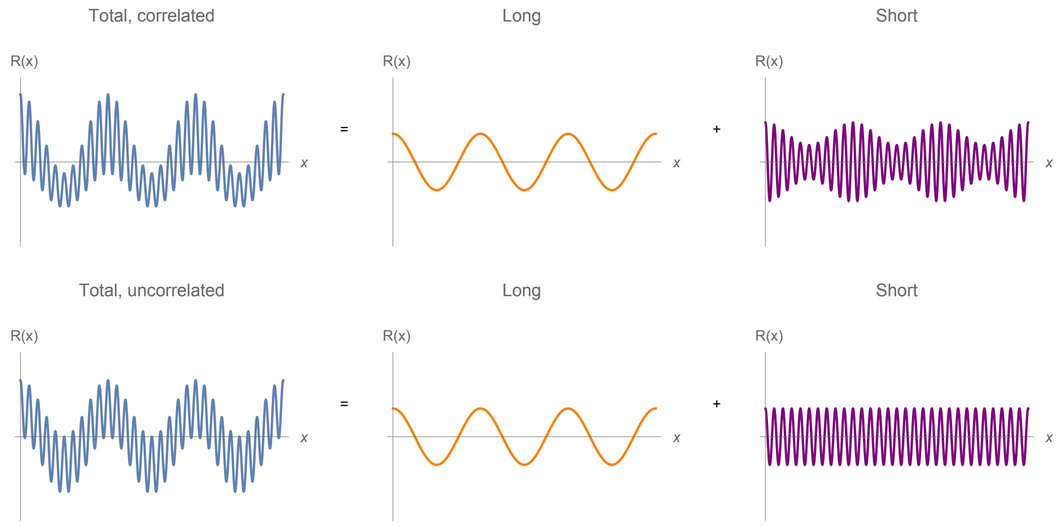

Figure 8: The figure show a profile of $\mathcal{R}(x)$ composed of a long mode (orange) and a short mode (purple). In the top panel they are uncorrelated, while in the bottom they amplitude of the short mode correlates with the long mode. This is the type of non-Gaussianity described by the squeezed limit of the bispectrum, which is fixed by the soft theorem.

- It is the soft limit, a.k.a. squeezed limit of the correlator that is fixed by the theorem. This represents the correlation (a.k.a. mode coupling) between one long wavelength mode $\lambda_{\text{long}} \sim 1/q$ and two short wavelength modes $\lambda_{\text{short}} \sim 1/k \ll \lambda_{\text{long}}$, as depicted in Fig. 8

- This gives the slow-roll suppressed bispectrum. Indeed one can check that the two bispectra in (4.27), which are not slow-roll suppressed, are subleading in this soft limit, $q \to 0$. If one keeps all slow-roll suppressed terms that we have neglected, one can indeed check the validity of this result via direct calculation [29].

- This relation is valid for all single field models in which $\mathcal{R}$ becomes constant (i.e. adiabatic) on superHubble scales, but it is in general violated in multifield models. Observing any deviation from this relation, e.g. in the CMB temperature anisotropy bispectrum would rule out the leading class of inflationary models.

- We derived the relation using comoving momentum $\mathbf{k}$. After relating $\mathbf{k}$ to the physical momentum $\mathbf{k}_p$ using the perturbed metric this result reduces to [35]

$$\lim_{\mathbf{q}_p \to 0} \langle \mathcal{R}(\mathbf{q}_p) \mathcal{R}(\mathbf{k}_p) \mathcal{R}(\mathbf{k}'_p) \rangle' = \mathcal{O}(q^2) \,. \tag{6.78}$$

  This is to be expected since by definition adiabatic modes are locally equivalent to a change of coordinates and so cannot affect the physics. A more formal and precise derivation of this fact uses (conformal) Fermi Coordinates [8, 14, 15]. The $\mathcal{O}(q^2)$ term is model dependent but has a lower bound of order η [9, 13].

- Many other soft theorems exist, also with soft tensor and vectors [24, 26, 29, 33].

7 Phenomenology

It is useful to recap all of our results in a way that can be effectively communicated to a late universe observer, who tries to measure signals from the primordial universe. Our results are predictions for the correlation function of the gauge-invariant variables $\mathcal{R}$ and γ_{ij}.

7.1 Primordial non-Gaussianity

We can write the most generic scalar bispectrum as

$$\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\mathcal{R}(\mathbf{k}_3)\rangle = (2\pi)^3\delta_D(\sum_a \mathbf{k}_a) f_{NL} B(k_1,k_2,k_3)\,, \tag{7.1}$$

where f_{NL} gives us the overall *size* of the bispectrum and B, which is normalized to[35] we impose the conventional normalization

$$k^6 B(k,k,k) = -\frac{18}{5}(2\pi)^4\Delta_{\mathcal{R}}^4\,, \tag{7.2}$$

gives us the *shape*, i.e. the dependence on the momenta. Above I used the fact that B is scale invariant and so $B(k,k,k)\propto k^{-6}$. Notice that in principle the bispectrum depends on 9 variables, namely 3 vectors, each with 3 components. But translation and rotation invariance each remove 3 of them, in such a way that B only depends on $9-3-3=3$ variables, which we have here chosen to be $k_{1,2,3}$. Intuitively the delta function forces $\mathbf{k}_{1,2,3}$ to form a triangle, which is fully determines by the length of it 3 sides. More generally, an n-point function depends on $3(n-2)$ variables, for $n\geq 3$.

We quoted the shape of the bispectrum induced by the interactions $\dot\varphi^3$ and $\dot\varphi\partial_i\varphi^2$ in (4.27) to be

$$\begin{aligned} B_{\varphi'^3} &= -\frac{18}{5}(2\pi)^4\Delta_{\mathcal{R}}^4\,\frac{1}{2k_1k_2k_3k_T^3}\,, \qquad (7.3)\\ B_{\varphi'(\partial_i\varphi)^2} &= -\frac{18}{5}(2\pi)^4\Delta_{\mathcal{R}}^4\,\frac{1}{102}\frac{1}{(k_1k_2k_3)^3k_T^3}\left[24\,(k_1k_2k_3)^2 - 8k_T\,(k_1k_2k_3)\left(\sum_{a<b}k_ak_b\right)\right.\\ &\quad\left. -8\,k_T^2\left(\sum_{a<b}k_ak_b\right)^2 + 22\,k_T^3\,(k_1k_2k_3) - 6k_T^4\left(\sum_{a<b}k_ak_b\right) + 2k_T^6\right]. \qquad (7.4)\end{aligned}$$

The slow-roll suppressed bispectrum induced by gravity for a canonical scalar field, $P=X-V$, which we did not compute in these lectures, is found to be the sum of two terms [29]

$$\begin{aligned} B_\epsilon(k_1,k_2,k_3) &= -\frac{18}{5}(2\pi)^4\Delta_{\mathcal{R}}^4\cdot\frac{1}{5}\frac{1}{\prod k_i^3}\left[-3\sum_i k_i^3 + \sum_{i\neq j}k_ik_j^2 + 8\frac{\sum_{i>j}k_i^2k_j^2}{k_t}\right] \qquad (7.5)\\ B_{\rm loc}(k_1,k_2,k_3) &= -\frac{18}{5}(2\pi)^4\Delta_{\mathcal{R}}^4\cdot\frac{1}{3}\frac{\sum_i k_i^3}{\prod k_i^3}\,, \qquad (7.6)\end{aligned}$$

[35] The strange factor 3/5 comes about because a seminal paper on non-Gaussianity [27] considered the scalar potential Φ instead of $\mathcal{R}$ and during matter domination $\mathcal{R}=-5/3\Phi$.

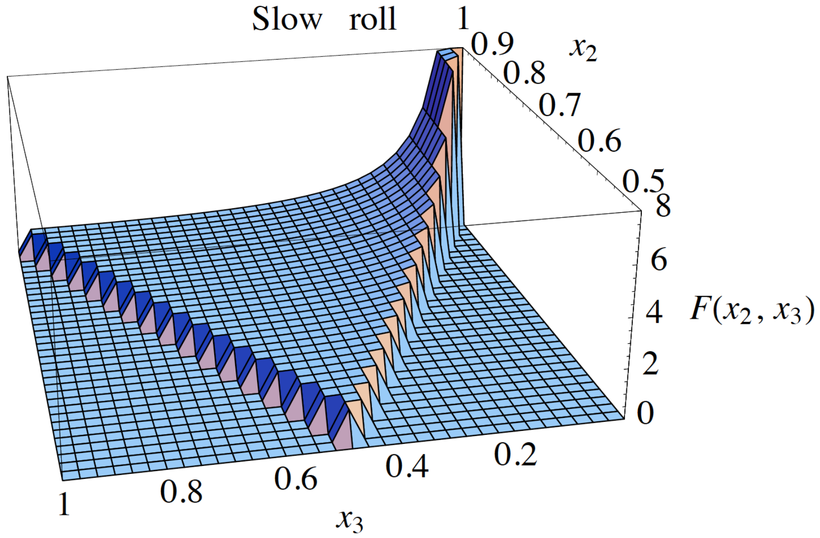

Figure 9: The shape of the bispectrum produced by gravitational non-linearities in canonical, single-field slow-roll inflation.

where the name of "local" in the second shape will become clear in the next lecture. The size of these bispectra are slow-roll suppressed as expected

$$f_{NL}^{\epsilon} = -\left(\frac{5}{12}\right)^2 \epsilon \,, \tag{7.7}$$

$$f_{NL}^{\text{loc}} = -\frac{5}{48}\left(\eta + 2\epsilon\right) = \frac{5}{24}(n_s - 1) \,. \tag{7.8}$$

Because of scale invariance, we can write

$$B(k_1, k_2, k_3) = \frac{1}{k_1^6} B\left(1, \frac{k_2}{k_1}, \frac{k_3}{k_1}\right) \,. \tag{7.9}$$

While all bispectra share the k^{-6} factor, the differ in how they depend on the dimensionless ratios $x_2 \equiv k_2/k_1$ and $x_3 \equiv k_3/k_1$. To see how similar or different two bispectra are, we plot the following function of two variable

$$B(1, x_2, x_3) x_2^2 x_3^2 \,. \tag{7.10}$$

where the additional factor of $x_2^2 x_3^2$ is added to account for the momentum space volume, see [7] for more details. For example, Figure 9 (from [7]) shows the shape of of the bispectrum in canonical slow-roll inflation, which is induced by gravity. Clearly the correlator peaks in squeezed configurations, where $x_2 \sim 1$ and $x_3 \sim 0$, which translates to $k_3 \ll k_1 \sim k_2$.

7.2 Quantum-to-classical transition

As $\mathcal{R}$ perturbations leave the Hubble radius the become effectively classical. A precise derivation of this fact is still a matter of debate in literature, but we will content ourselves with a heuristic argument. Using the dS mode functions and the conversion from a canonical field φ_c to $\mathcal{R}$, we can write the free field $\mathcal{R}$ to leading order in $\tau \to 0$ as

$$\mathcal{R}(\mathbf{k},\tau) \simeq \frac{1}{2\sqrt{\epsilon c_s k^3}}\left(a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}}\right) + \mathcal{O}(\tau^2)\,, \tag{7.11}$$

$$\dot{\mathcal{R}}(\mathbf{k},\tau) \simeq -H\tau^2 k^2 \frac{1}{2\sqrt{\epsilon c_s k^3}}\left(a_{\mathbf{k}} + a^\dagger_{-\mathbf{k}}\right) + \mathcal{O}(\tau^2)\,. \tag{7.12}$$

We notice that $\mathcal{R}$ and $\dot{\mathcal{R}}$ are proportional to each other and so must commute up to $\mathcal{O}(\tau^2)$ corrections. The defining feature of quantum mechanics, namely the non-commutation of operators, becomes harder and harder to measure as time goes on. More quantitatively, we can try to define a classicality parameter C that quantifies how precise our observations need to be to detect the non-commutation of $\mathcal{R}$ and $\dot{\mathcal{R}}$ (see e.g. [6])

$$C \equiv \frac{|\langle[\mathcal{R},\dot{\mathcal{R}}]\rangle|}{\sqrt{\langle\mathcal{R}^2\rangle\langle\dot{\mathcal{R}}^2\rangle}}\,. \tag{7.13}$$

This can be readily computed and expanded for $\tau \to 0$:

$$C = \frac{|f_k \dot{f}_k^* - f_k^* \dot{f}_k|}{|f_k \dot{f}_k|} \tag{7.14}$$

$$\simeq \frac{2H\tau^3 k^3}{1\times H\tau^2 k^2} \simeq 2\tau k \to 0\,. \tag{7.15}$$

In particular, at the end of inflation, when we match to the radiation dominated hot big bang,

$$C \sim \frac{k}{(aH)} = e^{-N} \sim e^{-50} \sim 10^{-21}\,, \tag{7.16}$$

where $N \sim 50$ is the number of efoldings between then end of inflation and when the mode k leaves the Hubble radius during inflation. So, unless we can measure the time evolution of $\mathcal{R}$ with a precision of 10^{-21}, we can safely describe correlators as classical averages, as opposed to quantum ones.

A Graviton polarization tensors

In this appendix, I discuss the graviton polarization tensors. Since all these conditions are invariant under rotations, to find ϵ^s_{ij} explicitly, we can simply choose some convenient $\mathbf{k}$, e.g. $\hat{\mathbf{k}} = \mathbf{k}/k = \hat{\mathbf{z}}$, and then rotate the result. A simple solution to all the conditions in (2.56)-(2.60) is

$$\epsilon^{+2}_{ij}(\hat{\mathbf{z}}) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & i & 0 \\ i & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \epsilon^{-2}_{ij}(\hat{\mathbf{z}}) = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -i & 0 \\ -i & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\,. \tag{A.1}$$

This choice has the advantage of describing eigenvectors of rotations around the $\hat{\mathbf{k}}$ axis, namely

$$\epsilon^{+2}_{ij} \to e^{i2\theta}\epsilon^{+2}_{ij}\,, \qquad \epsilon^{-2}_{ij} \to e^{-i2\theta}\epsilon^{-2}_{ij}\,. \tag{A.2}$$

Indeed they are mapped into each other by parity,

$$\epsilon_{ij}^{+2}(-\mathbf{k}) = \epsilon_{ij}^{+2}(\mathbf{k})^* = \epsilon_{ij}^{-2}(\mathbf{k})\,, \tag{A.3}$$

and viceversa. This is not the only choice since any rotation around $\hat{\mathbf{z}}$ gives a different choice of polarization. More generally, we can use real polarization vectors, which are not eigenstates of rotations. Given wavevector $\mathbf{k}$, we define *real* vectors $\hat{\mathbf{u}}$ and $\hat{\mathbf{v}}$ to form an orthonormal basis with $\hat{\mathbf{k}} = \mathbf{k}/k$. Then

$$\epsilon_{ij}^{+}(\mathbf{k}) = \hat{u}_i\hat{u}_j - \hat{v}_i\hat{v}_j \quad \text{and} \quad \epsilon_{ij}^{\times}(\mathbf{k}) = \hat{v}_i\hat{u}_j + \hat{v}_j\hat{u}_i\,. \tag{A.4}$$

B Useful formulae

Here I collect some useful formulae and their sign conventions

$$S_2 = \int d^4x\sqrt{-g}\left[\frac{M_{\rm Pl}^2}{2}R - \Lambda\right]\,, \tag{B.1}$$

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + g_{\mu\nu}\Lambda = \frac{1}{M_{\rm Pl}^2}T_{\mu\nu}\,, \tag{B.2}$$

$$R_{\mu\nu} \equiv 2\Gamma^{\rho}_{\mu[\nu,\rho]} + 2\Gamma^{\rho}_{\lambda[\rho}\Gamma^{\Lambda}_{\beta]\alpha}\,, \tag{B.3}$$

$$\Gamma^{\mu}_{\alpha\beta} \equiv \frac{1}{2}g^{\mu\gamma}\left(g_{\alpha\gamma,\beta} + g_{\beta\gamma,\alpha} - g_{\alpha\beta,\gamma}\right)\,, \tag{B.4}$$

$$T^{\mu\nu} = \frac{2}{\sqrt{-g}}\frac{\delta S}{\delta g_{\mu\nu}}\,, \tag{B.5}$$

$$ds^2 = -dt^2 + a(t)^2\frac{dx^i dx^j\delta_{ij}}{(1+K\mathbf{x}^2/4)^2}\,, \tag{B.6}$$

$$T^{\mu}_{\ \nu} = \mathrm{Diag}\left\{-\rho, p, p, p\right\}\,, \tag{B.7}$$

$$3M_{\rm Pl}^2\left(H^2 + \frac{K}{a^2}\right) = \sum_a \rho_a\,, \tag{B.8}$$

$$-\dot{H}M_{\rm Pl}^2 = \frac{1}{2}\left(\rho + p\right)\,, \tag{B.9}$$

$$M_{\rm Pl}^2\frac{\ddot{a}}{a} = -\frac{1}{6}\left(\rho + 3p\right)\,, \tag{B.10}$$

$$\epsilon \equiv -\frac{\dot{H}}{H^2}\,, \tag{B.11}$$

$$X \equiv -\frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi \tag{B.12}$$

$$\varphi(\mathbf{x}) = \int_{\mathbf{k}}\varphi(\mathbf{k})e^{i\mathbf{k}\cdot\mathbf{x}}, \quad \varphi(\mathbf{k}) = \int_{\mathbf{x}}\varphi(\mathbf{x})e^{-i\mathbf{k}\cdot\mathbf{x}}\,, \tag{B.13}$$

$$\varphi(\mathbf{k},t) = f_k(t)a_{\mathbf{k}} + f_k^*(t)a^\dagger_{-\mathbf{k}}\,, \tag{B.14}$$

$$f_k = \frac{H}{\sqrt{2k^3}}(1+ik\tau)e^{-ik\tau}\,, \tag{B.15}$$

$$\langle\varphi(\mathbf{k})\varphi(\mathbf{k}')\rangle = (2\pi)^3\delta^3_D\left(\mathbf{k}+\mathbf{k}'\right)P(k)\,, \tag{B.16}$$

$$P(k) = \frac{H^2}{2k^3}\,. \tag{B.17}$$

C Lesson references and further reading

Sec. 1: A quick review of background cosmology This background material can be found in many excellent textbooks, such as the ones by Weinberg [49], Dodelson [16] and Mukhanov [30]. My presentation is based on my own lecture notes for cosmology [32].

Sec. 2: Free fields on curved backgrounds This discussion is based on the very nice review by Yi Wang [45].

Sec. 5.1 A very nice discussion of EFT's, on which this section is based was given by Polchinski in the beautiful lecture notes [37].

Sec. 5.1.1 A nice and pedagogical introduction to GR as an EFT can be found in Donoghue lecture notes [17].

Sec. 5.4 My notation for the SVT decomposition is that of Weinberg [49]

Sec. 6.4 This presentation parallels in spirit that of [24], but instead of the Schrödinger picture of "wave functionals" I use the more standard interaction picture.

References

[1] P. A. R. Ade et al. Planck 2015 results. XIII. Cosmological parameters. 2015.

[2] Nima Arkani-Hamed, Paolo Benincasa, and Alexander Postnikov. Cosmological Polytopes and the Wavefunction of the Universe. 2017.

[3] Nima Arkani-Hamed and Juan Maldacena. Cosmological Collider Physics. 2015.

[4] C. Armendariz-Picon, T. Damour, and Viatcheslav F. Mukhanov. k - inflation. Phys. Lett., B458:209–218, 1999.

[5] C. Armendariz-Picon, Viatcheslav F. Mukhanov, and Paul J. Steinhardt. A Dynamical solution to the problem of a small cosmological constant and late time cosmic acceleration. Phys. Rev. Lett., 85:4438–4441, 2000.

[6] Valentin Assassi, Daniel Baumann, and Daniel Green. Symmetries and Loops in Inflation. JHEP, 02:151, 2013.

[7] Daniel Babich, Paolo Creminelli, and Matias Zaldarriaga. The Shape of non-Gaussianities. JCAP, 0408:009, 2004.

[8] Tobias Baldauf, Uros Seljak, Leonardo Senatore, and Matias Zaldarriaga. Galaxy Bias and non-Linear Structure Formation in General Relativity. JCAP, 1110:031, 2011.

[9] Giovanni Cabass, Enrico Pajer, and Fabian Schmidt. How Gaussian can our Universe be? JCAP, 1701(01):003, 2017.

[10] Xingang Chen, Yi Wang, and Zhong-Zhi Xianyu. Schwinger-Keldysh Diagrammatics for Primordial Perturbations. JCAP, 1712(12):006, 2017.

[11] Paolo Creminelli, Markus A. Luty, Alberto Nicolis, and Leonardo Senatore. Starting the Universe: Stable Violation of the Null Energy Condition and Non-standard Cosmologies. JHEP, 12:080, 2006.

[12] Paolo Creminelli, Jorge Norena, and Marko Simonovic. Conformal consistency relations for single-field inflation. JCAP, 1207:052, 2012.

[13] Paolo Creminelli, Ashley Perko, Leonardo Senatore, Marko Simonovi?, and Gabriele Trevisan. The Physical Squeezed Limit: Consistency Relations at Order q^2. JCAP, 1311:015, 2013.

[14] Liang Dai, Enrico Pajer, and Fabian Schmidt. Conformal Fermi Coordinates. JCAP, 1511(11):043, 2015.

[15] Liang Dai, Enrico Pajer, and Fabian Schmidt. On Separate Universes. JCAP, 1510(10):059, 2015.

[16] Scott Dodelson. Modern Cosmology. Academic Press, Amsterdam, 2003.

[17] John F. Donoghue. General relativity as an effective field theory: The leading quantum corrections. Phys. Rev., D50:3874–3888, 1994.

[18] S. Dubovsky, T. Gregoire, A. Nicolis, and R. Rattazzi. Null energy condition and superluminal propagation. JHEP, 03:025, 2006.

[19] Bernardo Finelli, Garrett Goon, Enrico Pajer, and Luca Santoni. Soft Theorems For Shift-Symmetric Cosmologies. Phys. Rev., D97(6):063531, 2018.

[20] Bernardo Finelli, Garrett Goon, Enrico Pajer, and Luca Santoni. The Effective Theory of Shift-Symmetric Cosmologies. JCAP, 1805(05):060, 2018.

[21] Alan H. Guth. The Inflationary Universe: A Possible Solution to the Horizon and Flatness Problems. Phys. Rev., D23:347–356, 1981.

[22] Jemal Guven, Ben Lieberman, and Christopher T. Hill. Schrodinger Picture Field Theory in Robertson-walker Flat Space-times. Phys. Rev., D39:438, 1989.

[23] Kurt Hinterbichler, Lam Hui, and Justin Khoury. Conformal Symmetries of Adiabatic Modes in Cosmology. JCAP, 1208:017, 2012.

[24] Kurt Hinterbichler, Lam Hui, and Justin Khoury. An Infinite Set of Ward Identities for Adiabatic Modes in Cosmology. JCAP, 1401:039, 2014.

[25] Renee Hlozek et al. The Atacama Cosmology Telescope: a measurement of the primordial power spectrum. Astrophys. J., 749:90, 2012.

[26] Lam Hui, Austin Joyce, and Sam S. C. Wong. Inflationary soft theorems revisited: A generalized consistency relation. JCAP, 1902:060, 2019.

[27] Eiichiro Komatsu and David N. Spergel. Acoustic signatures in the primary microwave background bispectrum. Phys. Rev., D63:063002, 2001.

[28] Andrei D. Linde. A New Inflationary Universe Scenario: A Possible Solution of the Horizon, Flatness, Homogeneity, Isotropy and Primordial Monopole Problems. Adv. Ser. Astrophys. Cosmol., 3:149–153, 1987.

[29] Juan Martin Maldacena. Non-Gaussian features of primordial fluctuations in single field inflationary models. JHEP, 05:013, 2003.

[30] V. Mukhanov. Physical Foundations of Cosmology. Cambridge University Press, Oxford, 2005.

[31] Hugh Osborn. Lectures on Conformal Field Theories in more than two dimensions. 2010.

[32] Enrico Pajer. Cosmology lecture notes.

[33] Enrico Pajer and Sadra Jazayeri. Systematics of Adiabatic Modes: Flat Universes. JCAP, 1803(03):013, 2018.

[34] Enrico Pajer, Guilherme L. Pimentel, and Jaap V. S. Van Wijck. The Conformal Limit of Inflation in the Era of CMB Polarimetry. JCAP, 1706(06):009, 2017.

[35] Enrico Pajer, Fabian Schmidt, and Matias Zaldarriaga. The Observed Squeezed Limit of Cosmological Three-Point Functions. Phys. Rev., D88(8):083502, 2013.

[36] Joseph Polchinski. Renormalization and Effective Lagrangians. Nucl. Phys., B231:269–295, 1984.

[37] Joseph Polchinski. Effective field theory and the Fermi surface. In TASI 92: From Black Holes and Strings to Particles, pages 0235–276, 1992.

[38] V. A. Rubakov. Phantom without UV pathology. Theor. Math. Phys., 149:1651–1664, 2006. [Teor. Mat. Fiz.149,409(2006)].

[39] Slava Rychkov. EPFL Lectures on Conformal Field Theory in D¿= 3 Dimensions. SpringerBriefs in Physics. 2016.

[40] David Simmons-Duffin. The Conformal Bootstrap. In Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in pages 1–74, 2017.

[41] D. T. Son. Low-energy quantum effective action for relativistic superfluids. 2002.

[42] Alexei A. Starobinsky. A New Type of Isotropic Cosmological Models Without Singularity. Adv. Ser. Astrophys. Cosmol., 3:130–133, 1987.

[43] Gerard 't Hooft and M. J. G. Veltman. One loop divergencies in the theory of gravitation. Ann. Inst. H. Poincare Phys. Theor., A20:69–94, 1974.

[44] David Wands. Local non-Gaussianity from inflation. Class. Quant. Grav., 27:124002, 2010.

[45] Yi Wang. Inflation, Cosmic Perturbations and Non-Gaussianities. Commun. Theor. Phys., 62:109–166, 2014.

[46] Steven Weinberg. Adiabatic modes in cosmology. Phys. Rev., D67:123504, 2003.

[47] Steven Weinberg. Quantum contributions to cosmological correlations. Phys. Rev., D72:043514, 2005.

[48] Steven Weinberg. The Quantum theory of fields. Vol. 1: Foundations. Cambridge University Press, 2005.

[49] Steven Weinberg. Cosmology. 2008.

Field Theory in Cosmology: Example Sheet 1

1. For a $P(X,\phi)$ theory

$$S = \int \sqrt{-g} P(X,\phi)\,, \tag{1}$$

compute the equations of motion. Compute the energy-momentum tensor and find the identification upon which it reduces to that of a perfect fluid

$$T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + g_{\mu\nu} p\,. \tag{2}$$

Re-derive the equations of motion by combining the two Friedmann equations, which for a perfect fluid take the general form

$$3M_{\mathrm{P}l}^2 H^2 = \rho\,, \qquad -\dot{H} M_{\mathrm{P}l}^2 = \frac{1}{2}(\rho + p)\,. \tag{3}$$

2. Compute the power spectrum of a massive scalar field in de Sitter. Consider the action

$$S = -\int \sqrt{-g}\frac{1}{2}\left[\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2\right]\,, \tag{4}$$

for some mass m. Write $\phi(\mathbf{k})$ in terms of creation and annihilation operators $\{a_{\mathbf{k}}, a^\dagger_{\mathbf{k}}\}$ and mode functions f_k. Derive the equation that $f_k(\tau)$ has to satisfy from the action (4), using conformal time. To solve this equation, re-write it as an equation for $g_k = (-\tau)^{-3/2} f_k$, and then use the fact that the two linear independent solution of Bessel's differential equation,

$$x^2 \partial_x^2 y + x \partial_x y + (x^2 - \alpha^2) y = 0\,, \tag{5}$$

can be taken to be the two Hankel functions $H^{(1,2)}_\alpha$. Now that you have the most general solution for f_k, with two integration constant, match this solution in the $-k\tau \to \infty$ limit to the flat space solution. You may use the following expansions of the Hankel functions for $x \to \infty$

$$H^{(1)}_\alpha(x) \simeq \sqrt{\frac{2}{\pi}}\frac{e^{ix}}{\sqrt{x}}\,, \qquad H^{(2)}_\alpha(x) \simeq \sqrt{\frac{2}{\pi}}\frac{e^{-ix}}{\sqrt{x}}\,, \tag{6}$$

which are valid up to an irrelevant (α-dependent) phase. You should find

$$f_k(\tau) = \frac{\sqrt{\pi} H}{2}(-\tau)^{3/2} H^{(1)}_\nu(-k\tau)\,, \quad \nu = \sqrt{\frac{9}{4} - \frac{m^2}{H^2}}\,. \tag{7}$$

3. Compute the two-point correlators

$$\lim_{\tau\to 0}\langle \phi(\mathbf{k})\pi(\mathbf{k}')\rangle = (2\pi)^3 \delta^3_D(\mathbf{k} + \mathbf{k}')\frac{1}{2k\tau}\,, \tag{8}$$

$$\lim_{\tau\to 0}\langle \phi(\mathbf{k})\pi(\mathbf{k}')\rangle = \lim_{\tau\to 0}\langle \pi(\mathbf{k})\phi(\mathbf{k}')\rangle\,, \tag{9}$$

$$\lim_{\tau\to 0}\langle \pi(\mathbf{k})\pi(\mathbf{k}')\rangle = (2\pi)^3 \delta^3_D(\mathbf{k} + \mathbf{k}')\frac{k}{2H^2\tau^2}\,. \tag{10}$$

4. Using the power spectrum derived in the lecture, compute the (real space) correlation function at separate points for a massless scalar field in dS and show that it is IR divergent:

$$\langle \phi(\mathbf{x})\phi(\mathbf{0})\rangle = \frac{H^2}{(2\pi)^2}\int_0^\infty d\tilde{k}\,\frac{\sin\tilde{k}}{\tilde{k}^2}\,. \tag{11}$$

5. Compute the amount of particle production in dS. In the lectures, we fixed the mode functions by demanding that φ creates positive-energy particles at $k\tau \to -\infty$. Let's instead require that φ creates positive-energy particles at some finite $|\tau_*| > \infty$, still satisfying $|k\tau_*| \gg 1$. The quantized field then takes the form

$$\varphi(\mathbf{k}) = g_k b_\mathbf{k} + g_k^* b_{-\mathbf{k}}^\dagger \,, \tag{12}$$

where $\{b_\mathbf{k}, b_\mathbf{k}^\dagger\}$ are a new set of creation and annihilation operators. Define the new vacuum state $|\tilde{0}\rangle$. Find g_k by matching to the Minkowski vacuum at τ_* (you may multiply g_k by a convenient phase)

$$g_k = \frac{H}{\sqrt{2k^3}} \left[1 + \frac{i}{k\tau_*} - \frac{1}{2(k\tau_*)^2}\right] f_k(\tau) + e^{-2ik\tau_*} \frac{H}{\sqrt{2k^3}} \frac{1}{2(k\tau_*)^2} f_k^*(\tau) \,,$$

By matching (12) to the expressions for $\varphi(\mathbf{k})$ we found in the lectures (i.e. matching to Minkowski at $|\tau_*| \to \infty$), show that the two sets of ladder operators are related,

$$a_\mathbf{k} = \frac{\sqrt{2k^3}}{H} \left(\alpha b_\mathbf{k} + \beta^* b_{-\mathbf{k}}^\dagger\right) \,, \qquad a_\mathbf{k}^\dagger = \frac{\sqrt{2k^3}}{H} \left(\beta b_{-\mathbf{k}} + \alpha^* b_\mathbf{k}^\dagger\right) \,, \tag{13}$$

This relation is called a *Bogoliubov transformation.* Invert it to give

$$b_\mathbf{k} = \frac{\sqrt{2k^3}}{H} \left(\alpha^* a_\mathbf{k} + \beta^* a_{-\mathbf{k}}^\dagger\right) \,, \qquad b_\mathbf{k}^\dagger = \frac{\sqrt{2k^3}}{H} \left(\beta a_{-\mathbf{k}} + \alpha a_\mathbf{k}^\dagger\right) \,, \tag{14}$$

Now we want to ask what a detector that measures $b_k^\dagger$ excitations would measure in the Bunch Davies vacuum $|0\rangle$, which we defined in the lecture as $a_\mathbf{k}|0\rangle = 0$. To this end, let's define the "b-particle" number operator

$$N_b(\mathbf{k}) = b_\mathbf{k}^\dagger b_\mathbf{k} \,. \tag{15}$$

Compute the expectation value of $N_b(\mathbf{k})$ on the state $|\tilde{0}\rangle$ and on the Bunch-Davies vacuum $|0\rangle$. To understand the singular factor $\delta_D^3(\mathbf{0})$, work at finite volume

$$(2\pi)^3 \delta_D^3(\mathbf{0}) = \lim_{V\to\infty} \int_V d^3x e^{-i\mathbf{0}\cdot\mathbf{x}} = \lim_{V\to\infty} V \,, \tag{16}$$

and define the number density of particles, $n_b(\mathbf{k}) \equiv N_b(\mathbf{k})/V$, instead of the total number $N_b(\mathbf{k})$. You should find that the Bunch-Davies state has a non-vanishing density of b-type particles given by

$$\langle 0| n_b(\mathbf{k}) |0\rangle = \frac{1}{4(k\tau)^4} \neq 0 \,. \tag{17}$$

6. The fact that an FLRW background is invariant under translations, $\mathbf{x} \to \mathbf{x} + \mathbf{b}$, implies that also correlators must be invariant

$$\langle \phi(\mathbf{x}_1) \dots \phi(\mathbf{x}_n) \rangle = \langle \phi(\mathbf{x}_1 + \mathbf{b}) \dots \phi(\mathbf{x}_n + \mathbf{b}) \rangle \,. \tag{18}$$

Using this, prove that momentum space correlators must always be proportional to a delta function of the total momentum

$$\langle \phi(\mathbf{k}_1) \dots \phi(\mathbf{k}_n) \rangle \propto \delta_D^3\left(\sum_{a=1}^n \mathbf{k}_a\right) \,. \tag{19}$$

7. For the metric

$$\begin{aligned} ds^2 &= -dt^2 + a^2 \left(\delta_{ij} + \gamma_{ij}\right) dx^i dx^j && (20) \\ &= \frac{1}{H^2\tau^2} \left[-d\tau^2 + \left(\delta_{ij} + \gamma_{ij}\right) dx^i dx^j\right] \,, && (21) \end{aligned}$$

where $\gamma_{ii} = \partial_i \gamma_{ij} = 0$, we want to expand the Einstein-Hilbert action in de Sitter

$$S_2 = \int d^4x \sqrt{-g} \left[\frac{M_{\rm Pl}^2}{2} R - \Lambda \right] \tag{22}$$

to second order in γ to find the action of a free graviton. You already performed a similar expansion around Minkowski in the General Relativity course and it was a painful calculation. Instead of doing it again, let's use a trick. Start by noticing that the dS metric is proportional to the Minkowski one

$$g_{\mu\nu}^{\rm dS} = a^2 g_{\mu\nu}^{\rm Mink} = \frac{1}{H^2 \tau^2} g_{\mu\nu}^{\rm Mink} \,, \tag{23}$$

with the identification $\tau^{(\rm dS)} = t^{(\rm Mink)}$. Notice that by the Friedmann equation

$$3 M_{\rm Pl}^2 H^2 = \Lambda \tag{24}$$

A metric with this property is called *conformally flat.* Given an arbitrary function Ω of the coordinates, the rescaling

$$g_{\mu\nu} \to \tilde{g}_{\mu\nu} = \Omega^2 g_{\mu\nu} \,, \tag{25}$$

is called a Weyl transformation. Various GR tensors transform quite easily under a Weyl rescaling. For example, the Ricci scalars $\tilde{R} \equiv \tilde{g}^{\mu\nu} \tilde{R}_{\mu\nu}$ and $R \equiv g^{\mu\nu} R_{\mu\nu}$ for the two metrics are related by [this can be proven by direct calculation, if you wish]

$$\tilde{R} = \Omega^{-2} \left[R - 6 \nabla_\mu \nabla^\mu \ln \Omega - 6 (\nabla_\mu \ln \Omega)(\nabla^\mu \ln \Omega) \right] \,. \tag{26}$$

Now recall that in Minkowski, you found

$$S_2^{\rm Mink} = \frac{M_{\rm Pl}^2}{2} \int d^4x \sqrt{-g} R \tag{27}$$

$$= \frac{M_{\rm Pl}^2}{8} \int d^3x dt \left[\dot{\gamma}_{ij} \dot{\gamma}_{ij} - \partial_k \gamma_{ij} \partial_k \gamma_{ij} \right] + \mathcal{O}(\gamma^3) \quad \text{(Minkowski)} \tag{28}$$

Use (26) to rewrite the Einstein-Hilbert action around dS in terms of that around Minkowski, for which you can use the expansion above. You should find that around dS the graviton free action is

$$S_2 = \frac{M_{\rm Pl}^2}{8} \int d^3x d\tau \, a^2 \left[\gamma'_{ij} \gamma'_{ij} - \partial_k \gamma_{ij} \partial_k \gamma_{ij} \right] \,. \tag{29}$$

8. Prove that the two in-in expressions for a generic in-in correlator

$$\langle \mathcal{O}(t) \rangle = \sum_{N=0}^{\infty} i^N \int_{-\infty}^{t} dt_N \int_{-\infty}^{t_N} dt_{N-1} \dots \int_{-\infty}^{t_2} dt_1 \tag{30}$$
$$\times \langle 0 | [\hat{H}_{int}(t_1), [\hat{H}_{int}(t_2), \dots [\hat{H}_{int}(t_N), \mathcal{O}(t)] \dots]] | 0 \rangle \,,$$

$$\langle \mathcal{O}(t) \rangle = \langle 0 | \left[\bar{T} e^{\left(i \int_{-\infty(1+i\epsilon)}^{t} dt' \hat{H}_{int}(t') \right)} \right] \mathcal{O}(t) \left[T e^{\left(-i \int_{-\infty(1-i\epsilon)}^{t} dt' \hat{H}_{int}(t') \right)} \right] | 0 \rangle \,, \tag{31}$$

are indeed equivalent. Proceed by induction. First prove that they are equivalent at order $N = 0$ and $N = 1$. Then, assuming that they agree at order $N-1$, take the time derive of each Nth-order expression and rewrite it as the correlators of some other field to order $N-1$. This proves that the expression agree to order N up to a constant. By taking the limit $t \to -\infty$ show that the constant has to vanish.

9. Using the in-in formalism, compute the bispectrum in a $P(X)$ theory induced by the interactions $\dot{\varphi}^3$ and $\dot{\varphi}(\partial_i \varphi)^2$.

10. The fact that the de Sitter metric,

$$ds^2 = \frac{-d\tau^2 + dx^i dx^j \delta_{ij}}{\tau^2 H^2} \,, \tag{32}$$

is invariant under dilations, $\{\tau, \mathbf{x}\} \to \lambda\{\tau, \mathbf{x}\}$, implies that equal time correlators that do not depend on time, such as for example the power spectrum of a massless scalar field or of the graviton at $\tau \to 0$, must obey

$$\langle \phi(\mathbf{x}_1)\phi(\mathbf{x}_2)\dots\phi(\mathbf{x}_n)\rangle = \langle \phi(\lambda\mathbf{x}_1)\phi(\lambda\mathbf{x}_2)\dots\phi(\lambda\mathbf{x}_3)\rangle\,. \tag{33}$$

Using this, prove that momentum space correlators B_n, defined as

$$\langle \phi(\mathbf{k}_1)\dots\phi(\mathbf{k}_n)\rangle = (2\pi)^3\delta_D^3\left(\sum_{a=1}^{n}\mathbf{k}_a\right)B_n(\mathbf{k}_1,\dots,\mathbf{k}_n)\,. \tag{34}$$

must scale as

$$B_n(\lambda\mathbf{k}_1,\dots,\lambda\mathbf{k}_n) = \frac{1}{\lambda^{3(n-1)}}B_n(\mathbf{k}_1,\dots,\mathbf{k}_n)\,. \tag{35}$$

Field Theory in Cosmology: Example Sheet 2

1. Reproduce the constraint equations by varying the action

$$S = \int d^4x \sqrt{h} N \left\{ \frac{M_{\rm Pl}^2}{2} \left[{}^{(3)}R + K_{ij}K^{ij} - K^2 \right] + P(X, \phi) \right\} . \tag{1}$$

with respect to N and N^i.

2. Derive the linear-order gauge transformations of A, B, ψ and h_{00}, for a generic change of coordinates $x^\mu \to x^\mu + \epsilon^\mu$.

3. Derive the gauge transformation from Newtonian gauge to flat gauge and viceversa. In particular, given some generic perturbations $\{A^N, h_{00}^N, \varphi^N\}$ in Newtonian gauge, determine teh corresponding perturbations $\{h_{00}^f, \psi^f, \varphi^f\}$ in flat gauge.

4. Solve the $\delta S/\delta N^i$ constraint to find δN, working in flat gauge to linear order.

5. In the lecture, we prove the conservation of $\mathcal{R}$ and γ_{ij} on superHubble scales in the presence of a generic energy-momentum tensor $T_{\mu\nu}$ by working in comoving gauge. Prove again the conservation of $\mathcal{R}$ by working in Newtonian gauge. In particular, you might want to start with the change of coordinates

$$\epsilon^\mu = \left\{ \epsilon(t), \lambda x^i \right\} . \tag{2}$$

and show that the gauge transformations are

$$\Phi = -\dot{\epsilon} \,, \qquad \Psi = H\epsilon - \frac{\lambda}{3} .$$

$$\delta\rho = -\dot{\bar{\rho}}\epsilon \,, \qquad \delta u = \epsilon \,, \qquad \pi^S = 0 \,, \tag{3}$$

$$\delta p = -\dot{\bar{p}}\epsilon \,, \qquad \varphi = -\epsilon\dot{\bar{\phi}} . \tag{4}$$

Then use the scalar part of the ij components of the Einstein's equation,

$$k_i k_j (\Phi - \Psi) = 0, \tag{5}$$

to impose the physicality condition on $\epsilon(t)$. Your final result should be

$$\mathcal{R} = \frac{\lambda}{3} \,, \qquad \varphi = -\dot{\bar{\phi}}\frac{\mathcal{R}}{a}\int_T^t a(t')dt' \,, \qquad \Phi = \Psi = \mathcal{R}\left[-1 + \frac{H}{a}\int_T^t a(t')dt' \right] . \tag{6}$$

6.