Lecture Notes on Cosmological Soft Theorems

Enrico Pajer[a]

[a] *Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Wilberforce Road, Cambridge CB3 0WA, U.K.*

E-mail: enrico.pajer@gmail.com

Abstract: In these lecture notes, I review how symmetries constrain correlation functions in Quantum Field Theories on curved spacetime and discuss applications of these results to cosmology. I first discuss linearly realized symmetries in general FLRW spacetimes and in de Sitter spacetime. Then, I review adiabatic modes and show that they imply the existence of non-linearly realized symmetries for any covariant theory on cosmological backgrounds. I then discuss two ways to compute the soft theorems obeyed by correlators: Ward-Takahashi identities and the Operator Product Expansion. The result is that non-linearly realized symmetries relate n-point to $(n-1)$-point function in a model independent way. As concrete examples, I discuss single-field slow-roll and ultra-slow-roll inflation. These examples demonstrate how soft theorems can be used to infer general properties of inflation such as the number of active degrees of freedom and the underlying symmetry breaking pattern.

Contents

1 Invitation to cosmological soft theorems

The laws of nature display invariance under a set of transformations that we call symmetries. Symmetries range from spacetime transformations, such as rotations and translations, to internal symmetries that relate different particles to each other. Symmetries make a physicist's life easier by generating new solutions from old ones. But they do much more. Symmetries constrain the form that the answers to our questions can take. Symmetries also make precise, model independent predictions, which can be directly tested in experiments and observations. These predictions are particularly useful in situations where the number of models is much larger than the number of data points, such as in inflationary cosmology.

There are at least two paths that lead from symmetries to observations. The first one is exemplified by Effective Field Theory (EFT). When investigating physics beyond the standard model, we are low-energy observers in search of a high energy theory, which we cannot probe directly. We are obliged to speculate on possible extensions of our low-energy laws. Often the new physics is sufficiently far away from the regime of our experiments that we have a small parameter, such as the ratio of energy or length scales. In this case there is a very "effective" way proceeding: Begin by organizing high-energy theories in classes that share the same underlying set of symmetries. In each class, write down the most general theory that obeys the given symmetries and rank its effect on the low-energy physics using the small parameter. Finally constrain all free parameters at the appropriate order using experiments. In this EFT approach, the observational consequences of symmetries are build-in from the start. In general, once an EFT has been constructed, one can extract predictions using perturbation theory, which is often the only tool available for a general EFT.

There is a second way to derive the observational consequences of symmetries: obtain constraints directly for the observables, without computing the observables explicitly. In cosmology the main observable are so-called "in-in" correlation functions (see Sec. 2.2), which are simply the expectation values of operators such as curvature or energy perturbations in the quantum state of the universe. This state is usually assumed to be the vacuum state in the infinite past evolved with the full interacting theory and so it is found to contain stuff at late times. The consequences for correlators then depend on whether a given symmetry is *linearly* (and so necessarily unbroken) or *non-linearly realized*. For linearly realized symmetries one finds that correlators are annihilated by some set of linear operators. These relations take the general schematic form (see Sec 4)

$$\sum_{a=1}^{n} L_a \langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\dots\mathcal{O}(\mathbf{k}_n)\rangle = 0\,, \tag{1.1}$$

where $L_a = L(\tau_a, \partial_{\tau_a}, \mathbf{k}_a, \partial_{\mathbf{k}_a})$ is some linear, possibly differential operator made of functions and derivatives of time and momenta (or coordinates if working in real space). For every symmetry generator, there is a different L. A very familiar examples are translations for which L_a is just a multiplication by the a-th momentum $\mathbf{k}_a$.In this case, the solution of the above equation is simply that all correlators are proportional to delta functions of the sum of momenta

$$\langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\dots\mathcal{O}(\mathbf{k}_n)\rangle \equiv (2\pi)^3 \delta_D^3\left(\sum_{a=1}^{n}\mathbf{k}_a\right)\langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\dots\mathcal{O}(\mathbf{k}_n)\rangle'\,, \tag{1.2}$$

where I'll use primed correlators $\langle\dots\rangle'$ to indicate that I've stripped away the omnipresent delta function.

In many cases of interest though, the symmetry is spontaneously broken by the state of the theory and so the symmetry is necessarily non-linearly realized. In this case, the symmetry leads to a *soft theorem*, i.e. a relation between an $(n+1)$- and an n-point correlation function in the limit in which one of the momenta is "soft", i.e. much smaller than any other scale in the problem. The schematic form of a soft theorem is

$$\lim_{\mathbf{q}\to 0} \frac{\langle \mathcal{O}(\mathbf{q})\mathcal{O}(\mathbf{k}_1)\dots\mathcal{O}(\mathbf{k}_n)\rangle'}{\langle \mathcal{O}(\mathbf{q})\mathcal{O}(\mathbf{q})\rangle'} = \sum_{a=1}^{n} L_a \langle \mathcal{O}(\mathbf{k}_1)\dots\mathcal{O}(\mathbf{k}_n)\rangle' \,, \tag{1.3}$$

where again L_a are some linear differential operators that depend on the symmetry.

The non-linearly realized symmetries that play an important role in cosmology, and especially inflation, are associated with the concept of *adiabatic modes* (see Sec. 5). These are specific cosmological perturbation that are locally indistinguishible from a change of coordinates [28] and/or a symmetry transformation [10]. For each adiabatic mode one, can define a non-linearly realized symmetry of the action of cosmological perturbations [6, 7, 9, 29]. The symmetry allows one to trade a field in the correlator for a linear operation on the remaining fields, as happens in (1.3).

There are many different ways of deriving cosmological soft theorems. I will discuss five of them and provide a few example derivations. The most intuitive and least rigorous derivations is the background wave method [1, 34]. Perhaps the most familiar approach for high energy theorists are Ward-Takahashi identities (see e.g. [2, 6]), which are usually discussed in standard QFT classes and textbooks. Slavnov-Taylor identities can also be used and emphasize the role of the quantum effective action [41, 42]. A particularly fast and versatile method uses the Operator Product Expansion (OPE) (see e.g. [3, 5, 10]), also familiar from advanced QFT courses. Finally, one can use the Schrödinger picture [11, 40] and adapt the wave functional technique used in much of the holographic literature.

To make the general discussion more transparent, I'll discuss in detail three main examples: single field inflation both in the slow-roll and ultra-slow-roll regime and solid inflation. Single field slow-roll inflation is by far the most well-known and studied case. This is a great starting point because the algebra and the final results are very simple. But it hides some of the subtleties of the derivation. These subtleties are highlighted in ultra-slow-roll inflation, where perturbations do not necessarily become adiabatic and solid inflation, where the symmetry breaking patters of the theory is very different from standard inflation.

This review would not be completed without a discussion of the late time cosmological observations of the Cosmic Microwave Background (CMB) and large scale structures. The relation between soft theorems for primordial correlators and these late time probes contains a few subtleties. First, primordial perturbations start evolving again in the late universe and their evolution is non-linear. When discussing an n-point correlator, one needs to keep into account up to and including the $(n-1)$-th order in perturbation theory. Second, it is important to carefully consider the choice of coordinates when taking primordial perturbations as initial conditions for the late time evolution. Third, one should be be mindful of the fact that some soft theorems, by their own nature, imply that some effects are *locally* unobservable. When calculating observables in some gauge, this often leads to precise and perhaps unexpected cancellations. In this respect, I'll discuss how (conformal) Fermi Normal coordinates [52, 54, 55] provide a safe method to check the robustness of a calculation.

The goal of these notes is to provide a tool kit to extract observational predictions from symmetries in cosmological setups. The emphasis is on simple examples and basic derivations, which can work as toy models to understand the main underlying ideas and can hopefully be easily generalized to other cases of interest. Much more general results can often also be derived, but I will only mention them briefly and refer the interested reader to the original literature. I'll try to derive most results explicitly and provide as many proofs as possible of the various statements I'll make. The main target for these notes are researchers familiar with standard introductory QFT and cosmology courses. These notes are based on the lectures I gave at the 2019 Asian-Pacific Winter School hosted by the Yukawa Institute of Theoretical Physics in Kyoto, Japan.

2 Cosmological correlators

In this section, I introduce cosmological correlators, which are the main objects of interest in cosmology. I begin with comparing and contrasting correlators in cosmology with scattering amplitudes in particle physics. I then review three perturbative methods to compute cosmological correlators and present some simple application to the generation of primordial perturbations during inflation. Finally, I discuss how linearly realized symmetries constrain their form. As examples, I explicitly discuss spatial rotations and spatial translations for arbitrary FLRW spacetimes and dilation and special conformal transformations for correlators in de Sitter spacetime.

2.1 Scattering amplitudes

Perhaps the most effective way we have to study an object is to throw things at it and study 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

$$S_{\alpha\beta} \equiv \langle \alpha; +\infty | \beta; -\infty \rangle_S = \langle \alpha | S | \beta \rangle_H \tag{2.1}$$

where $|\alpha\rangle$ and $|\beta\rangle$ are eigenstates of the free Hamiltonian, which is assumed to be the same in the asymptotic past and future[1] and the subscripts S and H refer to the Schrödinger and Heisenberg pictures, respectively. Notice that S is an operator, while $S_{\alpha\beta}$ is a matrix with complex entries. In particle physics, which is often studied in Minkowski spacetime

[1]Because of this, scattering amplitudes are also sometimes called "in-out" correlators, but I'll reserve the word "correlator" for the "in-in" correlators we will introduce in Sec 2.2

$\eta_{\mu\nu} = \text{Diag}\{-,+,+,+\}$, the Heisenberg picture is most useful since it keep Lorentz invariance manifest.

One then looks for the largest number of operators that commute with H, i.e. a subset of the symmetries of theory, and uses their eigenvalues to label the α states. For example, in particle physics we typically assume invariance under Poincaré (translations and Lorentz transformations, i.e. $\text{ISO}(3,1) = \mathbb{R}^4 \rtimes \text{SO}(3,1)$). Free single particles states are then defined as irreducible representations of the Poincaré group and are classified by their four-momentum and the representation of the associated little group (see e.g. Chapter 3 of [12]). For example, given some creation an annihilation operators $\{a_p, a_p^\dagger\}$ for a certain particle with four-momentum p^μ and mass $p^2 = -m^2$, one finds for the scattering of 2 into $(n-2)$ particles

$$S_{\alpha\beta} = 2^{n/2}\sqrt{E_1 E_2 \dots E_n}\,\langle\Omega|\, a_{\mathbf{p}_3}(\infty)\dots a_{vp_n}(\infty) a^\dagger_{\mathbf{p}_1}(-\infty) a^\dagger_{\mathbf{p}_2}(-\infty)\,|\Omega\rangle \,, \tag{2.2}$$

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

$$\text{Prob} \sim |\,\langle\alpha|\, S\,|\beta\rangle\,|^2 \,. \tag{2.3}$$

Since the S operators encodes the effect of unitary evolution, the S-matrix is also unitary

$$\sum_\gamma S_{\alpha\gamma}(S^\dagger)_{\gamma\beta} = S_{\alpha\gamma}S^*_{\beta\gamma} = \delta(\alpha-\beta)\,. \tag{2.4}$$

For most systems of physical interest, the S-matrix can only be computed in perturbation theory. There are three popular ways to arrive at a representation of the S-matrix that is amenable to such an approximation scheme:

- Canonical quantization and the Hamiltonian approach, leading to Dyson's time-ordered formula (see e.g. Sec. 7.1 of [13] or Sec. 3.5 of [12]) for the operator S.
- Canonical quantization and the Lagrangian approach, passing by the LSZ reduction formula that relates scattering amplitudes to time-ordered products of fields (a.k.a. Green's function) and using the Schwinger-Dyson equations, which are the quantum equivalent of the classical equations of motion for the fields (see e.g. Section 7.2 of [13]).
- Path integral quantization (see e.g. Ch. 9 of [12]) .

All of these approaches eventually lead to the derivation of Feynman rules. For simplicity, I'll just briefly mention the first approach. For this, we need to introduce the interaction picture[2], labelled by I, where operators evolve with the free Hamiltonian and states evolve

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

$$|\psi,t\rangle_S = e^{-i\int H dt}\,|\psi,t_i\rangle_S\,, \qquad \mathcal{O}_S(t) = \mathcal{O}_S(t_i) \equiv \mathcal{O}_S\,, \tag{2.5}$$

$$|\psi\rangle_H = |\psi,t_i\rangle_S = e^{+i\int H dt}|\psi,t\rangle_S\,, \qquad \mathcal{O}_H(t) = e^{i\int H dt}\mathcal{O}_S\, e^{-i\int H dt}\,. \tag{2.6}$$

for some reference initial time t_i.

with the interaction Hamiltonian $H_{int} = H - H_0$:

$$|\psi, t\rangle_I = U_I(t, t_i)\,|\psi\rangle_H = U_I(t, t_i)\,|\psi, t_i\rangle_S = e^{+iH_0 t}|\psi, t\rangle_S\,, \tag{2.7}$$

$$\mathcal{O}_I(t) = U_I \mathcal{O}_H(t) U_I^\dagger = e^{+iH_0 t}\mathcal{O}_S e^{-iH_0 t}\,, \tag{2.8}$$

In the interaction picture, the evolution operator U_I obeys $U_I(t, t) = 1$ and

$$\frac{d}{dt_2}U_I(t_2, t_1) = -iH_{int}(t_2)U_I(t_2, t_1)\,, \tag{2.9}$$

$$\frac{d}{dt_1}U_I(t_2, t_1) = iU_I(t_2, t_1)H_{int}(t_1)\,. \tag{2.10}$$

For $t_2 > t_1$, the solution of these equation is concisely given by Dyson's formula

$$U_I(t_2, t_1) = T\exp\left(-i\int_{t_1}^{t_2} dt' H_{int}(t')\right)\,, \tag{2.11}$$

$$U_I^\dagger(t_2, t_1) = \bar{T}\exp\left(-i\int_{t_1}^{t_2} dt' H_{int}(t')\right)\,, \tag{2.12}$$

where the (anti) time-ordered operator ($\bar{T}$) T arranges the operators from left to right in order or (increasing) decreasing time. When the arguments are not in the right order, the solution of (2.9) and (2.10) is instead given by[3] (using the shorthand $U_{21} = U_I(t_2, t_1)$ etc.)

$$U_{12} \equiv U_{21}^\dagger = U_{21}^{-1} \qquad\qquad U_{12}^\dagger \equiv U_{21}\,, \tag{2.13}$$

in such a way that

$$U_{12}U_{21} = U_{21}U_{12} = 1\,, \qquad\qquad U_{32}U_{21} = U_{31}\,, \tag{2.14}$$

for any ordering of $t_{1,2,3}$. Here I have assumed the free Hamiltonian H_0 is time independent, but not the interaction part H_{int}. Dyson's formula then gives us the useful representation

$$S = U_I(+\infty, -\infty) = T\exp\left[-i\int_{-\infty}^{+\infty} dt' H_{int}(t')\right]\,. \tag{2.15}$$

Notice that H_{int} and hence U_I in the interaction picture are written in terms of *free fields*, such as for example

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2E}}\left[a_{\mathbf{p}}e^{ipx} + a_{\mathbf{p}}^\dagger e^{-ipx}\right]\,, \tag{2.16}$$

which evolve according to (2.8). (2.9) admits the perturbative expansion

$$U(t_2, t_1) = 1 - i\int_{t_1}^{t_2} dt' H_{int}(t') - \int_{t_1}^{t_2} dt' \int_{t_1}^{t'} dt'' H_{int}(t')H_{int}(t'') + \ldots\,. \tag{2.17}$$

[3]Notice that $U(t_1, t_2)$ is not simply given by (2.11), but rather the time order is replaced by an anti-time order.

2.2 Cosmological correlators

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

- **Broken Poincaré symmetry**. The classical background on which cosmological perturbations propagate has less isometries than Minkowski spacetime where particle physics operate. In particular, a flat [4] FLRW metric with line element

$$ds^2 = -dt^2 + a(t)^2 d\mathbf{x}^2 \tag{2.19}$$

has six space-like Killing vectors corresponding to invariance under rotations a spatial translations ($ISO(3) = \mathbb{R}^3 \rtimes \mathrm{SO}(3)$). This is four isometries less than the maximally symmetric[5] Minkowski spacetime. In particular, Lorentz boosts and time translations are spontaneously broken by the background.

- **Out of equilibrium** There is no asymptotic future in which cosmological perturbations stop interacting and so it is hard to define[6] "in-out" observables as we do for particle physics. In most cosmological models, cosmological perturbations were sub-Hubble at very early times (i.e. the comoving wavenumber k obeyed $k \gg aH$, where H is the Hubble parameter $H \equiv \dot{a}/a$). Then by the equivalence principle, these perturbations were effectively in flat space and so an "in" state in cosmology could be defined just as in particle physics. At late times instead cosmological perturbations can in general evolve and interact with each other, giving rise for example to the ever evolving Large Scale Structures of the universe. As a result of this difficulty, the quantities of interest in cosmology are the "in-in" correlators.

- **Cosmic variance** In an expanding universe with a finite age, causality imposes that there is only a finite volume that we can access at any given time. 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 $\ddot{a} \sim (10^{17}\mathrm{sec})^{-2}$ (pretty slowly). If this acceleration continues in the future, the largest spatial volume we can ever observe is of order the Hubble volume today, $H_0^{-3} \sim (4\mathrm{Gpc})^3$. Hence we cannot observe field fluctuations in the whole universe and so our measurements have an intrinsic sample variance, known in this context as *cosmic variance*.

[4] Everything I am discussing in this note can in principle be generalized to open and closed FLRW spacetimes, which have non-vanishing spatial curvature and the same amount of isometries (6 Killing vectors) as flat FLRW. In practice though there is relatively little literature on this subject because current observations are still compatible with a flat universe:

$$\Omega_k \equiv \left.\frac{K}{a^2 H^2}\right|_{\mathrm{today}} < 0.005\,. \tag{2.18}$$

It would nevertheless be interesting to study cosmological correlators and soft theorems in spatially curved universe because some amount of spatial curvature at the order $\Omega_k \sim 10^{-5}$ is unavoidable from super-Hubble perturbations. In addition, many models of inflation with just enough efoldings to explain current observations do predict that spatial curvature will be observed in the future.

[5] Recall that the maximum number of isometries a D-dimensional space can have is $D(D+1)/2$, which is 10 in $(3+1)$ dimensions.

[6] For an attempt to build an S-matrix in de Sitter see [14].

To define cosmological correlators, a few words on cosmological perturbation theory are in order. One proceeds semi-classically, by postulating a classical background and quantizing small deviations in perturbation theory. For example, for some matter content with an energy-momentum tensor $T_{\mu\nu}$, we define

$$g_{\mu\nu}(x,t) = \bar{g}_{\mu\nu}(t) + h_{\mu\nu}(x,t)\,, \tag{2.20}$$

$$T_{\mu\nu}(x,t) = \bar{T}_{\mu\nu}(t) + \delta T_{\mu\nu}(x,t)\,, \tag{2.21}$$

where barred quantities represent the homogenous and isotropic *background* solutions and we assume small perturbations $|h_{\mu\nu}| \ll |\bar{g}_{\mu\nu}|$, $|\delta T_{\mu\nu}| \ll |\bar{T}_{\mu\nu}|$. In particular, $\bar{g}_{\mu\nu}$ is the flat FLRW metric in (2.19), and

$$\bar{T}^{\mu}{}_{\nu} = \mathrm{Diag}\left\{-\bar{\rho}, \bar{p}, \bar{p}, \bar{p}\right\}\,, \quad \bar{T}_{\mu\nu} = \mathrm{Diag}\left\{\bar{\rho}, a^2\bar{p}, a^2\bar{p}, a^2\bar{p}\right\}\,. \tag{2.22}$$

The evolution of perturbations is determined by some (explicitly time-dependent) Hamiltonian, which we split again in a quadratic or "free" part and the remaining "interaction" part H_{int}. We then define an "in-in" correlator as the expectation value of some operator $\mathcal{O}$ on some state α

$$\langle\mathcal{O}\rangle_{\alpha} = \langle\alpha;t|\,\mathcal{O}\,|\alpha;t\rangle_{\mathrm{S}} = \langle\alpha|\,\mathcal{O}(t)\,|\alpha\rangle_{H}\,. \tag{2.23}$$

Typically, $\mathcal{O}$ is the product of equal-time operators at different space points. Time ordering is therefore irrelevant. The reason why equal time correlators play a central role in cosmology is that variables such as curvature perturbations and gravitons become time-independent on superHubble scales. As familiar from quantum mechanics, in-in correlators of Hermitian operators are real quantities since

$$\langle\alpha|\,\mathcal{O}(t)\,|\alpha\rangle_{H}^{*} = \langle\alpha|\,\mathcal{O}^{\dagger}(t)\,|\alpha\rangle_{H} = \langle\alpha|\,\mathcal{O}(t)\,|\alpha\rangle_{H}\,. \tag{2.24}$$

So, in contrast with scattering amplitudes, in-in correlators are already physical observables. We will be mostly interested in the case in which $|\alpha\rangle_{H}$ is the "vacuum" of the full theory in the far past (to be specified later). In this case I will simply keep it implicit in all subsequent formulae. Explicit calculations are most easily performed in the interaction picture. Using again Dyson's formula for the evolution operator acting on states and (2.7), (2.8), we find

$$\langle\alpha|\,\mathcal{O}(t)\,|\alpha\rangle_{H} = \langle\alpha|\,U_I^{\dagger}(-\infty,t)\mathcal{O}(t)U_I(-\infty,t)\,|\alpha\rangle_{I}\,, \tag{2.25}$$

where both $\mathcal{O}$ and all the fields in U_I are in the interaction picture. Once again, we have been able to reduce the problem to the calculation of a product of interaction picture fields, which are just free fields as for example

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3}\left[a_{\mathbf{k}} f_k(t) e^{i\mathbf{k}\cdot\mathbf{x}} + a_{\mathbf{k}}^{\dagger} f_k^{*}(t) e^{-i\mathbf{k}\cdot\mathbf{x}}\right]\,, \tag{2.26}$$

with $f_k(t)$ given by the solution of the classical, linear equations of motion. For more details see e.g. [15] and references therein. 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. The second is the Schrödinger picture of quantum mechanics, where the wave function is a functional of the fields and it is ofter referred to as the wave function of the universe. Here we will not make explicit calculations using these formalisms.
The (time-ordered) correlators are related to the S-matrix by the celebrated LSZ reduction formula (see e.g. Sec. 6.1 of [13] or Sec. 10.3 of [12]):

$$\langle p_1, \dots, p_n | \, S \, | p_{n+1}, \dots, p_{n+m} \rangle = \prod_{a=1}^{n+m} \left[i \int d^4 x_a e^{i p_a x_a} \left(\Box - m^2 \right) \right] \langle T\{\phi(x_1) \dots \phi(x_{n+m})\} \rangle \, .$$

Since the left-hand side momenta are all on-shell, i.e. $p^2 = -m^2$, the right-hand contains only a subset of all possible correlators, namely only the on-shell ones.

3 Symmetries

Recall that symmetries in field theory are transformations $\Delta\phi$ of the fields[7] ϕ that leave the action invariant, or equivalently change the Lagrangian by to a total derivative

$$\Delta \mathcal{L} = \partial_\mu F^\mu \, . \tag{3.1}$$

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[8] (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 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. 1 commutes. By Nöther theorem there exist a conserved current

$$J^\mu = \frac{\delta \mathcal{L}}{\delta \partial_\mu \phi} \Delta\phi - F^\mu \quad \text{with} \quad \partial_\mu J^\mu = 0 \, . \tag{3.2}$$

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

[7] Here ϕ is used as a collective symbol to indicate all fields in theory, irrespectively of their spin.

[8] 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.

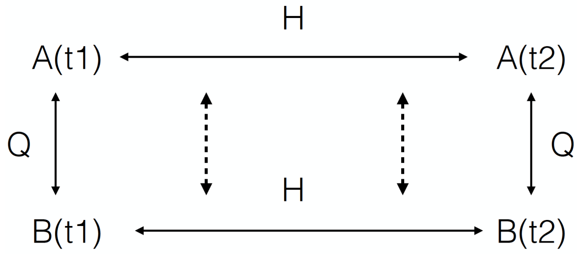

Figure 1. The diagram shows the equivalence of two definitions of symmetries: transformations that generate new solutions and transformations that commute with the Hamiltonian H. Some solution A at time t_1 can be evolved to time t_2 and then transformed by Q into $B(t_2)$. This gives the same result 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 $B(t)$. Notice that we will be always talking about symmetries of the laws of nature, not symmetries of the solutions of those laws (i.e. in general $A \neq B$).

then one can define a conserved current Q by

$$Q \equiv \int \sqrt{h} J^\mu n_\mu d^3x \,, \tag{3.3}$$

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:

$$i[Q, \phi] = \Delta\phi \,. \tag{3.4}$$

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

$$\text{Finite unitary transformation:} \quad U \equiv e^{i\alpha Q} \,, \tag{3.5}$$

for some parameter α of the transformation.

3.1 Unbroken, linearly realized symmetries

Now let us quantize the theory and promote Q and ϕ to operators. We say that the symmetry generated by Q is *unbroken* in the state $|\Omega\rangle$ iff

$$\langle\Omega| [Q, \phi] |\Omega\rangle = 0 \,, \tag{3.6}$$

Notice that a sufficient condition for this equation to be satisfied is that Q annihilates $|\Omega\rangle$, namely $Q|\Omega\rangle = 0$, so that $U|\Omega\rangle = |\Omega\rangle$. Notice that this condition implies that in an expansion in ϕ, $\Delta\phi$ cannot contain a constant term, but needs to start at least linear in ϕ. Assume now that $\Delta\phi$ is exactly linear in ϕ, namely

$$\text{Linearly realized symmetry:} \quad \Delta\phi_a = D_{ab}\phi_b \,, \tag{3.7}$$

with D_{ab} a set of matrices forming the representation of the symmetry group (recall ϕ represent some vector of fields). All linearly realized symmetries map "single-excited" states into single-excited states, where by "single-excited" I mean combinations of states of the form $a|0\rangle$. In particle physics these are "single-particle" states but in cosmology "single-excited" is perhaps more appropriate. For example, if $|0\rangle$ is the vacuum, the one-particle state $\phi_a(x)|0\rangle$ is mapped into

$$Q\phi_a(x)|0\rangle = [Q, \phi_a(x)]|0\rangle = -iD_{ab}\phi_b(x)|0\rangle \,, \tag{3.8}$$

which is also a linear combination of single-excited states. By definition of symmetry, Q must commute with the Hamiltonian, $[H, Q] = 0$, and so the new single-excited state must have the same energy as the original one: single-excited states come in multiples, just like particles do (e.g. the three pions).

If D_{ab} have entries that are (usually complex) numbers, then Q is said to be an internal symmetry. If instead it contains functions or derivatives of spacetime, then Q is a spacetime symmetry. More generally and more precisely, if P^μ and $M^{\mu\nu}$ are the generators of spacetime translations ($\mathbb{R}^4$) and Lorentz transformations (SO(3,1)), then Q is an *internal symmetry* if and only if it commutes with P^μ and $M^{\mu\nu}$:

$$\text{Internal symmetry:} \qquad [P^\mu, Q] = [M^{\mu\nu}, Q] = 0\,. \tag{3.9}$$

All symmetries that do not respect this condition are called *spacetime symmetries*.

3.2 Spontaneous symmetry breaking and non-linearly realized symmetries

In Quantum Field Theory (QFT), it is indeed possible that the vacuum state $|\Omega\rangle$ satisfies[9]

$$\text{Spontaneously broken symmetry:} \quad \langle\Omega|\,[Q, \phi]\,|\Omega\rangle \neq 0\,. \tag{3.10}$$

Notice that in particular $|\Omega\rangle$ is not invariant under Q namely $Q|\Omega\rangle \neq 0$. Also, by performing a field redefinition $\phi \to \phi - \langle\phi\rangle$, we can always work with fields that have vanishing expectation values $\langle\phi\rangle = 0$. These are indeed the fields we use to perturbatively quantize the theory. Then, for (3.10) to be true, a broken symmetry transformation in terms of these fields must contain a constant (i.e. field independent) term. So *a spontaneously broken symmetry must always be non-linearly realized*[10]:

$$\text{Non-linearly realized symmetry:} \quad i[Q, \phi] = \Delta\phi = \text{const} + \mathcal{O}(\phi)\,. \tag{3.11}$$

[9]This cannot happen in ordinary quantum mechanics. The hand-wavy argument is that in quantum mechanics one can always define a linear superposition of states corresponding to all possible images of a given ground state under the symmetry, and this gives a symmetric ground state. In QFT these states are mutually orthogonal because of the large volume limit and therefore belong to different super-selection sectors. See e.g. Sec. 19.1 of [12] or [16] for more details.

[10]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 $\Delta\phi_{sol,1} \neq \Delta\phi_{sol,2}$.

Then there must exist *degenerate vacua* since the state $|\alpha\rangle \equiv U(\alpha)\,|\Omega\rangle$ has the same energy as $|\Omega\rangle$

$$H\,|\alpha\rangle = H\left(U(\alpha)\,|\Omega\rangle\right) = U(\alpha)H\,|\Omega\rangle = E_\Omega\,|\alpha\rangle\ . \tag{3.12}$$

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 ($U\,|\Omega\rangle \neq |\Omega\rangle$). This should not be confused with *explicit symmetry breaking*, which describes a situation in which the transformation is just not a symmetry anymore. Explicit breaking happens when, given a theory with Hamiltonian H and symmetry Q, we add some interaction $\tilde{H}$ that does not commute with the symmetry, $[Q,\tilde{H}]\neq 0$. In this case, any implication of Q being a symmetry is at best only approximate and corrections arise from $\tilde{H}$. We learn something interesting only when $\tilde{H}$ is appropriately small. Spontaneous breaking on the contrary has exact and precise physical implications, without invoking any approximation.
A theorem due to Fabri and Picasso [17] shows that, if the theory is also translational invariant, then Q is divergent:

$$|Q\,|\Omega\rangle\,|^2 = \langle\Omega|\,QQ\,|\Omega\rangle = \int d^3x\,\langle\Omega|\,J^0(\mathbf{x})Q\,|\Omega\rangle = \int d^3x\,\langle\Omega|\,J^0(\mathbf{0})Q\,|\Omega\rangle = \infty\,, \tag{3.13}$$

where in the penultimate step I used translation invariance and in the last step that $Q\,|\Omega\rangle \neq 0$. Despite this subtlety, one can still use Q in commutators to generate a symmetry transformation as the commutator produces a delta function that soaks up the integral. For example for the archetypal generator of ϕ translations

$$Q = \int d^3y\Pi(y)\,, \tag{3.14}$$

with Π the conjugate momentum of ϕ, one finds

$$i[Q,\phi(\mathbf{x})] = i\int d^3y[\Pi(\mathbf{y}),\phi(\mathbf{x})] = \int d^3y\delta^3_D(\mathbf{x}-\mathbf{y}) = 1\,, \tag{3.15}$$

which is not divergent. Spontaneously broken symmetries do not send one-particle states into one-particle states because

$$Q\phi_a(x)\,|0\rangle = \left([Q,\phi_a(x)] + \phi Q\right)|0\rangle = -iD_{ab}\phi_b(x)\,|0\rangle + \phi Q\,|0\rangle\ , \tag{3.16}$$

and the last term in the last expression is not a one-particle state (it is not of the form $\phi_a\,|0\rangle$). If we break an internal symmetry, by Goldstone theorem we get one massless particle per broken generator with the same quantum numbers. This is called a Nambu-Goldstone boson (NGB), or a "goldstino" if we break a supersymmetry charge and the particle is a fermion. Notice that the masslessness of a NGB is a rare and precious fully non-perturbative result in QFT. The symmetry current "interpolates" between the vacuum and a one NGB particle state $|\pi(p)\rangle$ with four-momentum p^μ in the sense that

$$\langle 0|\,J^\mu(x)\,|\pi(p)\rangle = Fp^\mu e^{ipx}\,, \tag{3.17}$$

where F is called the *decay constant* of the NGB and has dimension of mass. In typical models of internal symmetry breaking it is related to the symmetry breaking vev of some field $F\sim\langle\phi\rangle$.

4 Cosmological correlators and linearly realized symmetries

In this section, I will discuss the observational consequences of linearly realized symmetries in cosmology. I will focus on spacetime symmetries and discuss in details translations and rotations, the isometries of all cosmological backgrounds (i.e. of the FLRW spacetime) and then dilations and special conformal transformations, the additional isometries of de Sitter spacetime (with similar results applying to Anti-de-Sitter). Internal symmetries, which change one type of field into another do not play much of a role in cosmology (exceptions are e.g. symmetric multifield models). To the best of my knowledge it is still an open question whether the above list of linearly realized symmetries is actually complete. For Lorentz invariant amplitudes, we have the celebrated Coleman-Mandula theorem [18], which dictates that the most general algebra of symmetries of the flat-space S-matrix is the direct product of Poincaré transformations and internal symmetries (i.e. symmetries that commute with Poincaré generators). But for correlators we don't have an analogous theorem, nor do I know of a counterexample.

4.1 FLRW spacetime: Homogeneity and isotropy

If we assume a Lorentz invariant theory and expand around a flat FLRW background, (2.19), which is homogeneous and isotropic, we find that all primordial correlators must be left invariant by the generators of spatial translations and rotations. The argument goes as follows. Consider the generators of spatial translations P^i and spatial rotations L^i, acting on scalar[11] operators of the theory as

$$i[P^i, \mathcal{O}_S(\mathbf{x})] = -\partial_i \mathcal{O}_S(\mathbf{x})\,, \tag{4.2}$$

$$i[L^i, \mathcal{O}_S(\mathbf{x})] = -\epsilon^{ijk} x_j \partial_k \mathcal{O}_S(\mathbf{x})\,. \tag{4.3}$$

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

$$U^{-1}(\vec{\alpha},\vec{\omega})\mathcal{O}(\mathbf{x})U(\vec{\alpha},\vec{\omega}) = \mathcal{O}(R^{ij}x^j + \alpha_i)\,, \tag{4.4}$$

with

$$R_{ij} = \exp\left(\epsilon_{ijk}\omega^k\right)\,, \qquad U(\vec{\alpha},\vec{\omega}) = \exp\left(iP^i\alpha_i\right)\exp\left(iL^i\omega_i\right)\,. \tag{4.5}$$

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

$$\langle\Omega|\prod_a \mathcal{O}(x_a)|\Omega\rangle = \langle\Omega|UU^{-1}\mathcal{O}(x_1)UU^{-1}\mathcal{O}(x_2)\dots\mathcal{O}(x_n)UU^{-1}|\Omega\rangle \tag{4.6}$$

$$= \langle\Omega|U^{-1}\mathcal{O}(x_1)UU^{-1}\mathcal{O}(x_2)\dots\mathcal{O}(x_n)U|\Omega\rangle \tag{4.7}$$

$$= \langle\Omega|\prod_a \mathcal{O}(t_a, R\mathbf{x}_a + \vec{\alpha})|\Omega\rangle\,, \tag{4.8}$$

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

$$i[L^i, \mathcal{O}_S^A(\mathbf{x})] = -D(L)^A_B \epsilon^{ijk} x_j \partial_k \mathcal{O}_S^B(\mathbf{x})\,, \tag{4.1}$$

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 (5.14) to linear order in $\vec{\alpha}$ and $\vec{\omega}$ and cancel the zeroth order piece with the left hand side. The remaining term is

$$\sum_{a=1}^{n} \frac{\partial}{\partial \mathbf{x}_a} \langle \mathcal{O}(x_1)\mathcal{O}(x_2)\ldots\mathcal{O}(x_n)\rangle \stackrel{!}{=} 0\,, \tag{4.9}$$

$$\sum_{a=1}^{n} \left(x_a^i \frac{\partial}{\partial x_a^j} - x_a^j \frac{\partial}{\partial x_a^i} \right) \langle \mathcal{O}(x_1)\mathcal{O}(x_2)\ldots\mathcal{O}(x_n)\rangle \stackrel{!}{=} 0\,. \tag{4.10}$$

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 $a = 2, \ldots n$. 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

$$\mathcal{O}(t,\mathbf{k}) \equiv \int d^3x e^{-i\mathbf{k}\cdot\mathbf{x}} \mathcal{O}(t,\mathbf{x})\,, \qquad \mathcal{O}(t,\mathbf{x}) \equiv \int \frac{d^3k}{(2\pi)^3} e^{i\mathbf{k}\cdot\mathbf{x}} \mathcal{O}(t,\mathbf{k})\,. \tag{4.11}$$

The generators acting on Fourier space correlators are then (Exercise)

$$P_i : -ik_i \quad \text{and} \quad R_{ij} : -i\left(k_i\partial_j - k_j\partial_i\right)\,, \tag{4.12}$$

and therefore

$$\sum_{a=1}^{n} \mathbf{k}_a \langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\ldots\mathcal{O}(\mathbf{k}_n)\rangle \stackrel{!}{=} 0\,, \tag{4.13}$$

$$\sum_{a=1}^{n} \left(k_a^i \frac{\partial}{\partial k_a^j} - k_a^j \frac{\partial}{\partial k_a^i} \right) \langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\ldots\mathcal{O}(\mathbf{k}_n)\rangle \stackrel{!}{=} 0\,. \tag{4.14}$$

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

$$\langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\ldots\mathcal{O}(\mathbf{k}_n)\rangle \equiv (2\pi)^3 \delta_D^3\left(\sum_a \mathbf{k}_a\right) \langle \mathcal{O}(\mathbf{k}_1)\mathcal{O}(\mathbf{k}_2)\ldots\mathcal{O}(\mathbf{k}_n)\rangle'\,, \tag{4.15}$$

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 $\mathbf{k}_a \cdot \mathbf{k}_b$.

4.2 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

$$ds^2 = -dt^2 + e^{2Ht} d\mathbf{x}^2 = \frac{-d\tau^2 + d\mathbf{x}^2}{\tau^2 H^2}\,, \tag{4.16}$$

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):

$$\text{dilation:} \quad \tau \to \tau(1+\lambda)\,, \qquad \mathbf{x} \to \mathbf{x}(1+\lambda)\,, \tag{4.17}$$

$$\text{SCT:} \quad \tau \to \tau(1-2\mathbf{b}.\mathbf{x})\,, \qquad \mathbf{x} \to \mathbf{x} - 2(\mathbf{b}.\mathbf{x})\,\mathbf{x} + (\mathbf{x}^2 - \tau^2)\mathbf{b}\,, \tag{4.18}$$

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. [19–22]).
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[12]

$$D: \; -\tau\partial_\tau - x^i\partial_i \qquad \text{(dilation)}\,, \tag{4.20}$$

$$\mathbf{b}\cdot\mathbf{K}: \; -2\mathbf{b}\cdot\mathbf{x}\left(\tau\partial_\tau - x^i\partial_i\right) - \left(\tau^2 - |\mathbf{x}|^2\right) b^i\partial_i \qquad \text{(SCT)}\,, \tag{4.21}$$

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:

$$\sum_{a=1}^{n} D_a \langle \mathcal{O}(\tau_1, \mathbf{x}_1)\mathcal{O}(\tau_2, \mathbf{x}_2)\ldots\mathcal{O}(\tau_n, \mathbf{x}_n)\rangle \overset{!}{=} 0\,, \tag{4.22}$$

$$\sum_{a=1}^{n} \mathbf{b}\cdot\mathbf{K}_a \langle \mathcal{O}(\tau_1, \mathbf{x}_1)\mathcal{O}(\tau_2, \mathbf{x}_2)\ldots\mathcal{O}(\tau_n, \mathbf{x}_n)\rangle \overset{!}{=} 0\,. \tag{4.23}$$

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 [23–25]). 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.

[12] Check that indeed $\epsilon^\mu = \{-\tau, -x^i\}$ is a Killing vector for the dS metric in (4.16), namely it solves

$$\mathcal{L}_\epsilon g_{\mu\nu} = -\left(\nabla_\mu \epsilon_\mu - \nabla_\mu \epsilon_\mu\right) = 0\,. \tag{4.19}$$

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

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

$$
\begin{aligned}
D: &\quad (3 - \tau\partial_\tau) + k^i \partial_{k^i}\,, && (4.24)\\
\mathbf{b}\cdot\mathbf{K}: &\quad (3 - \tau\partial_\tau)\, 2b^i \partial_{k^i} - \mathbf{b}\cdot\mathbf{k}\partial_{k^i}\partial_{k^i} + 2k^i \partial_{k^i} b^j \partial_{k^j}\,. && (4.25)
\end{aligned}
$$

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

$$
\left[-3 + \sum_{a=1}^{n} D_a\right] \langle \mathcal{O}(\tau_1, \mathbf{k}_1)\mathcal{O}(\tau_2, \mathbf{k}_2)\ldots\mathcal{O}(\tau_n, \mathbf{k}_n)\rangle' \stackrel{!}{=} 0\,, \qquad (4.26)
$$

$$
\left[\sum_{a=1}^{n} \mathbf{b}\cdot\mathbf{K}_a\right] \langle \mathcal{O}(\tau_1, \mathbf{k}_1)\mathcal{O}(\tau_2, \mathbf{k}_2)\ldots\mathcal{O}(\tau_n, \mathbf{k}_n)\rangle' \stackrel{!}{=} 0\,. \qquad (4.27)
$$

The primordial scalar spectral tilt We can use the above formal results to re-derive the well-known spectral tilt of the primordial curvature perturbation power spectrum $P_{\mathcal{R}}$ from single field slow-roll inflation

$$
1 - n_s \equiv -\frac{\partial \ln\left(k^3 P_{\mathcal{R}}\right)}{\partial \ln k} = 2\epsilon_V - 2\eta_V\,, \qquad (4.28)
$$

where I introduced the potential slow-roll parameters

$$
\epsilon_V \equiv M_{Pl}^2 \left(\frac{V'}{V}\right)^2 \qquad\qquad \eta_V \equiv M_{\rm Pl}^2 \frac{V''}{V}\,. \qquad (4.29)
$$

Since we have assumed exact de Sitter spacetime so far, we will be able to recover this result only in the limit

$$
\epsilon \equiv \frac{\dot{H}}{H^2} \simeq \epsilon_V \to 0\,, \qquad (4.30)
$$

where the Hubble parameter is approximately constant.

To compute n_s, first notice that unlike for translations and rotations, now we have derivatives acting also on the time dependence of the fields. While these constraints are valid at any time, it is particularly simple to evaluate them after Hubble crossing of all the modes, $k_a \ll aH$. For example, consider an inflaton ϕ of mass m in quasi-de Sitter spacetime. Neglecting all terms or order ϵ, ϕ must obey the equation of motion

$$
0 = (-\Box + m^2)\phi \propto \phi'' - \frac{2}{\tau}\phi' + \left(k^2 + \frac{m^2}{\tau^2 H^2}\right)\phi\,, \qquad (4.31)
$$

where a prime indicates a derivative with respect to conformal time and $\mathcal{H} \equiv a'/a = aH$. In the far future $-k\tau \ll 1$ (recall $a \propto e^{Ht}$) and so I can solve

$$
\phi'' - \frac{2}{\tau}\phi' + \frac{m^2}{\tau^2 H^2}\phi \simeq 0\,. \qquad (4.32)
$$

The time dependence of ϕ then is given by two power law solutions

$$
\mathcal{O}(\tau, \mathbf{x}) = \tau^{\Delta}\mathcal{O}_{\Delta}(\mathbf{x}) + \tau^{\tilde{\Delta}}\mathcal{O}_{\tilde{\Delta}}(\mathbf{x})\,, \qquad (4.33)
$$

where

$$\Delta = \frac{3}{2} - \sqrt{\frac{9}{4} - \frac{m^2}{H^2}}\,, \qquad\qquad \tilde{\Delta} = \frac{3}{2} + \sqrt{\frac{9}{4} - \frac{m^2}{H^2}}\,. \tag{4.34}$$

For small mass $m \ll 3H/2$, the $\tilde{\Delta}$ solution decays with time much more quickly than the Δ solution and can therefore be neglected, while

$$\Delta \simeq \frac{m^2}{3H^2} \simeq \eta_V + \mathcal{O}\left(\epsilon\right)\,. \tag{4.35}$$

Then, to compute ϕ correlators I can substitute $\tau\partial_\tau = \Delta$ everywhere in (4.26) and (4.27). For the equal-time two-point function, from (4.26) one finds the superHubble solution (Exercise)

$$\langle \phi(\tau, \mathbf{k})\phi(\tau, -\mathbf{k})\rangle' = A\frac{(-k\tau)^{2\eta_V}}{k^3}\,, \tag{4.36}$$

for some unknown normalization A. We can now compute the power spectrum of curvature perturbations $\mathcal{R} \equiv \phi/(M_{\rm Pl}\sqrt{2\epsilon})$ simply by evaluating the ϕ power spectrum at some fixed constant $\tau = \tau_*$ hypersurface such that $-k\tau_* \ll 1$. From this the spectral tilt follows

$$1 - n_s \equiv -\frac{\partial \ln\left(k^3 P_{\mathcal{R}}\right)}{\partial \ln k} = k\partial_k \ln(-k\tau)^{2\eta_V} = 2\eta_V\,. \tag{4.37}$$

More generally, this argument fixes the overall scaling with k any n-point function. To zeroth order in ϵ and η_V one finds $\Delta \simeq 0$ and the time dependence can be completely neglected. Then (4.26) becomes

$$\left[3(n-1) + \sum_{a=1}^{n} k_a \frac{\partial}{\partial k_a}\right] \langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\ldots\mathcal{R}(\mathbf{k}_n)\rangle' \stackrel{!}{=} 0\,, \tag{4.38}$$

with solution

$$\langle \mathcal{R}(\mathbf{k}_1)\mathcal{R}(\mathbf{k}_2)\ldots\mathcal{R}(\mathbf{k}_n)\rangle' = \frac{S\left(\frac{\mathbf{k}_1}{k_n}, \ldots, \frac{\mathbf{k}_{n-1}}{k_n}\right)}{k_n^{3(n-1)}}\,, \tag{4.39}$$

for some model dependent "shape" function S, which depends only on the displayed ratios. The solutions of the SCT constraint in Fourier space are more difficult to discuss, see e.g. [22, 26, 27].

5 Adiabatic modes

It is time to move on to study non-linearly realized symmetries. By definition, broken symmetries have field transformations that contain a constant term and so are non-linearly realized as in (3.11). Note that linear vs non-linear realization depends on the physical state of the theory and not on the symmetry itself. In other words, the same symmetry can be linearly realized on some states and non-linearly realized on others.

In this section, I discuss a large class of non-linearly realized symmetries that play a crucial role for correlators in presence of dynamical gravity and in particular in cosmology. These symmetries are a special subset of large diffeomorphisms (diffs) that are continuously connected to physical perturbations. When acting on an unperturbed FLRW spacetime, they generate new solutions called *adiabatic modes*. Adiabatic modes are physical perturbations that are locally equivalent to a change of coordinates and possible a global symmetry transformation [28]. In the following, after reviewing gauge transformations in cosmological perturbation theory, I introduce adiabatic modes in the simplest case of a single fluid cosmology (see [6, 7, 9, 29], which is relevant of applications to single field inflation. Then I discuss other, less standard case where adiabatic modes arise, namely theories with a shift symmetry and solids. This section focuses on the classical properties of adiabatic modes, while the quantum properties are discussed in the next section in the form of soft theorems.

5.1 Gauge transformations

In cosmology we have an exact, non-linear solution that describes a homogeneous and isotropic background and we expand in small perturbations as in (2.20). By the rotational invariance of the background, it is convenient to decompose the metric and the matter sectors into scalar, transverse vectors and transverse traceless tensors, or simply vectors and tensors. These are the lowest dimensional representation of SO(2), which is the little group of ISO(3) with respect to some non-vanishing spatial momentum $\mathbf{k}$. These perturbations are the cosmological equivalent of single particle states in particle physics, which are the irreps of the Poincaré symmetry group. Scalar, vectors and tensors have *helicity*[13] zero, one and two respectively. We parameterize the metric as (in the notation of [28, 30])

$$ds^2 = -(1+E)dt^2 + 2a(\partial_i F + G_i)dt dx^i + a^2 \left[(1+A)\delta_{ij} + \partial_i\partial_j B + 2\partial_{(i}C_{j)} + \gamma_{ij}\right] , \quad (5.1)$$

where $\{E, F, A, B\}$ are four scalars, $\{G_i, C_i\}$ two vectors and γ_{ij} a tensor satisfying

$$\gamma_{ii} = \partial_i\gamma_{ij} = \partial_i C_i = \partial_i G_i = 0 . \quad (5.2)$$

To streamline the presentation I will systematically neglect vector modes, $C_i = G_i = 0$. We assume a single perfect fluid, with energy-momentum tensor

$$T_{\mu\nu} = (\rho + p)\, u_\mu u_\nu + p g_{\mu\nu} , \quad (5.3)$$

with normalized velocity $u_\mu u^\mu = -1$, p pressure and ρ energy density. Generalizations to the multiple fluids case are straightforward. For the we assume the fluid has vanishing anisotropic stresses, but we'll relax this assumption later. The fluid velocity is decomposed into a scalar δu and vector δu^i_V as

$$u_\mu = (u_0, u_i) , \quad u_i = \partial_i \delta u + \delta u_i^V , \quad \partial_i \delta u_i^V = 0 , \quad (5.4)$$

[13] This is distinct from spin, which refers to the representations of SO(3), rather than SO(2). For example, a spin one particle has three states $m = -1, 0, 1$, while an helicity one particle is transverse i.e. it has only $m = -1, 1$ for rotations around its momentum $\mathbf{k}$.

and again we set the vector to zero $\delta u_i^V = 0$. By the covariance of general relativity, changes of coordinates

$$x^\mu \to x'^\mu = x^\mu + \epsilon^\mu(x)\,, \tag{5.5}$$

are symmetries of the theory. In cosmological perturbation theory, this is most conveniently expressed in terms of so called *gauge transformations*. For any tensor $T(t,\mathbf{x}) = \bar{T}(t) + \delta T(t, vx)$, with implicit indices, any change of coordinates $\epsilon(x)$ can be expressed as the gauge transformations Δ of the perturbations δT by

$$\Delta\left(\delta T\right) \equiv T'(x) - T(x)\,, \tag{5.6}$$

where T' is the tensor T in x' coordinates. For example, for tensors with zero and two Lorentz indices one finds

$$\phi'(x') = \phi(x)\,, \quad g'_{\mu'\nu'} = \frac{\partial x^\mu}{\partial x'^{\mu'}}\frac{\partial x^\nu}{\partial x'^{\nu'}} g_{\mu\nu}\,. \tag{5.7}$$

To linear order in ϵ and in perturbations, the gauge transformations of these perturbations are then found to be[14]

$$\Delta\delta\phi = -\epsilon^0 \dot{\bar{\phi}} = \epsilon_0 \dot{\bar{\phi}}\,, \tag{5.9}$$

$$\Delta h_{\mu\nu}(x) = -\nabla_\mu \epsilon_\nu - \nabla_\nu \epsilon_\mu\,. \tag{5.10}$$

More explicitly, and separating time and space components[15]:

$$\Delta h_{ij} = 2a^2 H \delta_{ij}\epsilon_0 - 2\epsilon_{(i,j)}\,, \tag{5.12}$$

$$\Delta h_{0i} = -\dot{\epsilon}_i - \partial_i \epsilon_0 + 2H\epsilon_i\,, \qquad \Delta h_{00} = -2\dot{\epsilon}_0\,, \tag{5.13}$$

$$\frac{\Delta\delta\rho}{\dot{\rho}} = \frac{\Delta\delta p}{\dot{p}} = \epsilon_0\,, \qquad \Delta u_i = -\partial_i \epsilon_0\,. \tag{5.14}$$

States of the theory that differ by a gauge transformation are physically indistinguishable. To ensure that we find physically distinct solutions of our theory, rather than gauge transformations of some single solution, it is easiest to compute some *gauge invariant variables* in some convenient gauge. To fix the gauge consider *small gauge transformations*, which are defined by the property

$$\text{Small gauge transformations:} \quad \lim_{|\mathbf{x}|\to\infty} \epsilon^\mu(t,\mathbf{x}) = 0\,. \tag{5.15}$$

[14] The differential geometric reader will have recognize these as (minus) the Lie derivative $\mathcal{L}_\epsilon$ in the ϵ^μ direction. This is true in general (always to linear order in ϵ and perturbations δT)

$$\Delta\delta\text{Tensor} = -\mathcal{L}_\epsilon \text{Tensor} = -\mathcal{L}_\epsilon \overline{\text{Tensor}}\,. \tag{5.8}$$

[15] Spatial indices are lowered and raised with the background metric, $\bar{g}_{00} = -1$ and $\bar{g}_{ij} = a^2\delta_{ij}$ so

$$\epsilon^i = \frac{1}{a^2}\epsilon_i \quad \text{and} \quad \epsilon_0 = -\epsilon^0\,. \tag{5.11}$$

Then the SVT components of the metric and $T_{\mu\nu}$ transform as

$$\Delta A = 2H\epsilon_0\,, \quad \Delta B = -\frac{2}{a^2}\epsilon^S\,, \tag{5.16}$$

$$\Delta C_i = -\frac{1}{a^2}\epsilon_i^V\,, \quad \Delta\gamma_i j = 0\,, \quad \Delta E = 2\dot{\epsilon}_0\,, \tag{5.17}$$

$$\Delta F = \frac{1}{a}\left(-\epsilon_0 - \dot{\epsilon}^S + 2H\epsilon^S\right)\,, \quad \Delta G_i = \frac{1}{a}\left(-\dot{\epsilon}_i^V + 2H\epsilon_i^V\right)\,, \tag{5.18}$$

$$\Delta\delta\rho = \dot{\bar{\rho}}\epsilon_0\,, \quad \Delta\delta p = \dot{\bar{p}}\epsilon_0 \quad \Delta\delta u = -\epsilon_0\,, \tag{5.19}$$

$$\Delta\pi^S = \Delta\pi_i^V = \Delta\pi_{ij}^T = \Delta\delta u_i^V = 0\,,$$

where I used the SVT-decomposition of the gauge parameter

$$\epsilon^\mu = \{\epsilon^0, \partial^i\epsilon^S + \epsilon_V^i\}\,. \tag{5.20}$$

Gauge invariant perturbations are now easy to build and there are many possible choices in the literature. The most commonly used variables are curvature perturbations on comoving ($\mathcal{R}$) and constant density (ζ) hypersurfaces, defined respectively by

$$\mathcal{R} \equiv \frac{A}{2} + H\delta u\,, \tag{5.21}$$

$$\zeta \equiv \frac{A}{2} - H\frac{\delta\rho}{\dot{\rho}}\,. \tag{5.22}$$

Note that these variables are gauge invariant only to linear order, and additional terms must be added for second and higher order. We'll address this in Sec. 7.1. Also, as we will now see, these variables are not invariant under large gauge transformations, with $\epsilon^\mu(\infty) \neq 0$.

5.2 Perfect fluids

In this section I discuss the most generic scalar and tensor adiabatic modes. The algebra is simple but a bit tedious. This discussion is based on [9]. The reader interested in the result can skip directly to Sec. 5.2.2.

5.2.1 The long story

We will execute the following four steps:

1. Fix the small gauge; in this case I'll work in comoving gauge.
2. Find residual large diffs that respect the comoving gauge condition.
3. Find the subset of large diffs that "extend to finite momentum", i.e. that solve Einstein equation non-trivially
4. Acting with these diffs on the unperturbed FLRW metric generate adiabatic modes

Let us see how this works in detail.

Step 1 Using (5.16)-(5.20), one can find diffs such that working in comoving gauge:

$$B' = B + \Delta B = 0 \qquad \delta u' = \delta u + \Delta \delta u = 0\,, \qquad C_i + \Delta C_i = 0\,. \tag{5.23}$$

This fixes small diffs completely, i.e. from the above conditions I can solve for small ϵ^μ *uniquely*. To avoid a pedantic notation, I'll still denote by $\mathcal{R}$ the value of the gauge invariant $\mathcal{R}$ in this gauge

$$\mathcal{R}|_{\text{comoving}} \equiv \mathcal{R} = \frac{A}{2}\,. \tag{5.24}$$

The metric then takes the form

$$ds^2 = -(1+2N_1)dt^2 + 2N_i dt dx^i + a^2\left[(1+2\mathcal{R})\delta_{ij} + \gamma_{ij}\right] dx^i dx^j\,, \tag{5.25}$$

where I renamed $E = 2N_1$ and $\partial_i F + G_i = \frac{1}{a} N_i$ as standard in the literature.

Step 2 By (5.12)- (5.14), a large gauge transformation on the unperturbed FLRW background generates the following perturbations (Exercise)

$$\mathcal{R} = H\epsilon_0 - \frac{1}{3a^2}\partial_k\epsilon_k\,, \qquad N_1 = \dot{\epsilon}_0\,, \tag{5.26}$$

$$N_i = -\partial_i\epsilon_0 + 2H\epsilon_i - \dot{\epsilon}_i \qquad \frac{\delta\rho}{\dot{\rho}} = \epsilon_0\,, \tag{5.27}$$

$$\gamma_{ij} = -2\partial_{<i}\epsilon^{j>} \qquad \delta u_i = -\partial_i\epsilon_0\,. \tag{5.28}$$

where $< \cdots >$ on indices is shorthand notation for the symmetric traceless part:

$$T_{<ij>} \equiv \frac{1}{2}\left(T_{ij} + T_{ji}\right) - \frac{1}{3}T_{kk}\delta_{ij}\,. \tag{5.29}$$

This set of perturbations is a solution of the equations of motion for any large ϵ, e.g. of the form

$$\epsilon^\mu = \sum_n a^\mu_{i_1 i_2 \ldots i_n}(t) x^{i_1} x^{i_2} \ldots x^{i_n}\,. \tag{5.30}$$

In Fourier space, this expression is non-vanishing only at $\mathbf{k} = 0$, since it is just a sum of derivatives of $\delta^3_D(\mathbf{k})$. I will then call this profile a *zero-momentum solution.* But since they come from a change of coordinates, these zero-momentum solutions are nothing but FLRW in unusual coordinates! Now comes the crucial point. We demand that these solutions "extend to finite momentum", i.e. that they can be interpreted as the $\mathbf{k} \to 0$ limit of some perturbation in comoving gauge. For this to be possible we need to impose

$$\partial_i\gamma_{ij} = \gamma_{ii} = 0 \quad \Rightarrow \quad \nabla^2\epsilon_i = -\frac{1}{3}\partial_i\partial_k\epsilon_k\,, \tag{5.31}$$

which is what we mean by a transverse traceless tensor. This in particular implies

$$\nabla^2\partial_i\epsilon^i = 0. \tag{5.32}$$

Step 3 But this is not sufficient. A direct calculation shows that the off-diagonal and ij parts of the Einstein Equations at linear order take the form

$$k_i k_j \left(N_1 + \mathcal{R} + \dot{\psi} + H\psi \right) = 0\,, \tag{5.33}$$

$$k_j \left(\dot{N}_i^V + H N_i^V \right) = 0\,, \tag{5.34}$$

$$k_i \left(H N_1 - \dot{\mathcal{R}} \right) = 0\,, \tag{5.35}$$

where $N_i = \partial_i \psi + N_i^V$. While these equations are automatically satisfied at $\mathbf{k} = 0$, they are not in general at $\mathbf{k}$. So to be able to extend them to finite momentum we demand

$$\left(N_1 + \mathcal{R} + \dot{\psi} + H\psi \right) \stackrel{!}{=} 0\,, \tag{5.36}$$

$$\left(\dot{N}_i^V + H N_i^V \right) \stackrel{!}{=} 0\,, \tag{5.37}$$

$$\left(H N_1 - \dot{\mathcal{R}} \right) \stackrel{!}{=} 0\,. \tag{5.38}$$

Using (5.26) in the constraint (5.38) we find

$$\epsilon_0 = \frac{1}{3\dot{H}} \partial_k \dot{\epsilon}^k \quad \Rightarrow \quad \nabla^2 \epsilon_0 = 0\,. \tag{5.39}$$

Integrating the constraint (5.36) I find

$$\psi = -\epsilon_0 + \frac{1}{3a} \int^t dt' a(t') \partial_k \epsilon^k\,, \tag{5.40}$$

where I set to zero an integration function since it only lead to vector adiabatic modes (see [9] for more details). Using (5.39) and (5.32) we find out that $\nabla^2 \psi = 0$. Comparing with the left equation of (5.27)

$$\partial_i \psi = N_i = -\partial_i \epsilon_0 + 2H \partial_i \epsilon_i - \partial_i \dot{\epsilon}_i\,, \tag{5.41}$$

one finds the solution

$$\epsilon^i(t, x) = \bar{\epsilon}^i(\mathbf{x}) - \partial_i \partial_k \bar{\epsilon}^k \int^t \frac{dt'}{3a(t')^3} \int^{t'} dt'' a(t'')\,. \tag{5.42}$$

Step 4 According to (5.26) these diffs generate the solution (more vector modes can be found in [9])

$$\mathcal{R} = -\frac{1}{3} \partial_k \bar{\epsilon}^k\,, \qquad \psi = \frac{1}{3a} \partial_k \bar{\epsilon}^k \int^t dt' a(t')\,, \tag{5.43}$$

$$\gamma_{ij} = -2\partial_{<i} \bar{\epsilon}^{j>} + 2\partial_i \partial_j \partial_k \bar{\epsilon}^k \int^t \frac{dt'}{3a(t')^3} \int^{t'} dt'' a(t'')\,. \tag{5.44}$$

5.2.2 The short story

Here I discuss only the results for the leading adiabatic mode, for more details see Sec. 5.2.1. Consider the following diff

$$\epsilon^\mu = \{0, \omega_{ij} x^j\} \,. \tag{5.45}$$

According to (5.12)-(5.14), when acting on the unperturbed FLRW background, the diagonal part of this diff, namely ω_{ii}, generates the constant curvature mode

$$\text{Scalar adiabatic mode:} \quad \mathcal{R} = -\frac{\omega_{ii}}{3} \,, \qquad \psi = \frac{\omega_{ii}}{3a} \int a(t')dt' \,, \tag{5.46}$$

The anti-symmetric part $\omega_{[ij]}$ is just a rotation, which does not generate any perturbation because FLRW is rotation invariant. Finally, the symmetric traceless part $\omega_{<ij>}$ generates an adiabatic tensor mode. Using again (5.12)-(5.14) we see that

$$\text{Tensor adiabatic mode:} \quad \gamma_{ij} = -2\omega_{<ij>} \,. \tag{5.47}$$

This derivation proves that no matter the ingredient of the universe and the expansion history, there are always a constant scalar and a constant tensor modes. This can be derived in several different ways [31], and has been known for a long time and [32]. If the system has a single active scalar degree of freedom, as in single field inflation, then (5.46) must be the solution on large scales. This is what makes early universe cosmology so interesting if you care about high energy physics! Primordial perturbations, produced in the first fraction of a second after the big bang, at energy probably much larger than anything we will be ever able to reproduce on earth are conserved during most of the history of the universe. So, we can measure them at late times in cosmological observables such as the CMB or Large Scale Structures. Remarkably, it is precisely the scalar adiabatic mode that generates all cosmological perturbations we have ever measured in our universe. Anything deviation from this scalar adiabatic mode is constrained to be less than about a percent [33]. The tensor adiabatic mode on the other hand has not yet measured, but there is large ongoing experimental effort to detect it in the odd-parity polarization of the CMB.

Summarizing, in single fluid cosmologies, we have found that residual large diffs lead to non-linear symmetries of cosmological perturbations. Since these are diff they can be computed to linear order in ϵ and in perturbation but allowing for terms order $\epsilon h_{\mu\nu}$ simply from

$$\Delta h_{\mu\nu}(x) \equiv -2\nabla_{(\mu}\epsilon_{\nu)} = -2\bar{\nabla}_{(\mu}\epsilon_{\nu)} - 2\delta\Gamma^\alpha_{\mu\nu}\epsilon_\alpha \,. \tag{5.48}$$

The result is (Exercise)

$$\Delta\mathcal{R} = H\epsilon_0 - \frac{1}{3}\partial_i\epsilon^i + \frac{1}{2}\vec{\partial}\epsilon_0 \cdot \vec{\partial}\psi - \epsilon^\mu\partial_\mu\mathcal{R} \,, \tag{5.49}$$

$$\Delta\gamma_{ij} = -2\partial_{<i}\epsilon_{j>} - \epsilon^\mu\partial_\mu\gamma_{ij} \,, \tag{5.50}$$

where the first two terms in $\mathcal{R}$ and the first in γ_{ij} are the same non-linear shifts as in (5.26), indicating that symmetry is non-linearly realized. More explicitly and for further reference, the resulting symmetries in real and Fourier space are (Exercise)

$$\gamma_{ij}(t,\mathbf{x}) \to \gamma_{ij}(t,\mathbf{x}) - 2\omega_{<ij>} - \omega_{<lm>}x^m\partial_l\gamma_{ij}(t,\mathbf{x})\,, \tag{5.51}$$

$$\gamma_{ij}(t,\mathbf{k}) \to \gamma_{ij}(t,\mathbf{k}) - 2\omega_{<ij>}(2\pi)^3\delta^3(\mathbf{k}) - \ldots\,, \tag{5.52}$$

$$\mathcal{R}(t,x) \to \mathcal{R}(t,\mathbf{x}) - \frac{\omega_{ii}}{3} - \frac{\omega_{ii}}{3}x^i\partial_i\mathcal{R}(t,\mathbf{x})\,, \tag{5.53}$$

$$\mathcal{R}(t,\mathbf{k}) \to \mathcal{R}(t,\mathbf{k}) - \frac{\omega_{ii}}{3}(2\pi)^3\delta^3(\mathbf{k}) + \frac{\omega_{ii}}{3}\left(3+\mathbf{k}\cdot\partial_{\mathbf{k}}\right)\mathcal{R}(t,\mathbf{k})\,. \tag{5.54}$$

Infinitely many other symmetry can be found, that are analogous to this one. See e.g. [6, 7, 9, 29, 34] for more details.

5.3 Shift symmetric cosmologies

In the previous example, the only symmetry at play were diffs. But adiabatic modes can be generalized to account for physical perturbations that are locally indistinguishable from a change of coordinates *plus* a global symmetry transformation [10]. This is relevant since in many models additional symmetries besides diffs play an important role. For example, here I will discuss cosmologies with a single scalar field ϕ that enjoys a shift symmetry $\phi \to \phi+$const. This symmetry is present in almost all attempts to embed inflation in a UV-finite theory of gravity such as string theory. The reason is, unless an approximate symmetry is at play, one expects dimension five and six Planck suppressed corrections in the low-energy effective action, from the coupling to Planck scale modes such as higher string oscillations. These corrections typically lead to large corrections to the η slow-roll parameter (the so called η problem), which shorten the duration of inflation to just a few efoldings.

The following is based on [10]. To next to leading order in derivatives, a scalar shift symmetry theory must take the form

$$\mathcal{L} = -\frac{M_{\rm Pl}^2}{2}R + P(X) + G(X)\Box\phi\,, \quad \text{with} \quad X \equiv -\frac{1}{2}(\nabla\Phi)^2\,, \tag{5.55}$$

where P and G are arbitrary functions of the canonical kinetic term X. To find adiabatic modes related to this new symmetry, we will follow the same general procedure as in the previous section:

1. First, we choose a convenient gauge that fixes small gauge transformations. Let us again choose comoving gauge, where metric takes the form (5.25) and instead of $\delta u = 0$ as in (5.23) we will impose $\phi(t,\mathbf{x}) = \bar\phi(t)$ or in other words $\delta\phi = 0$.

2. Then, we identify all residual symmetry transformation that preserves the gauge choice. In addition to the adiabatic mode we found in the previous section, now we can also consider the diff $\epsilon^0 = c/\dot{\bar\phi}$ for any constant c. This does change our gauge condition $\delta\phi = 0$ into

$$\delta\phi \to \delta\phi + \epsilon_0\dot{\bar\phi} = \delta\phi - c\,. \tag{5.56}$$

   But this change can be re-absorbed in a global shift of the field $\phi \to \phi + c$.

3. Now, we mpose that the resulting zero-momentum modes solves all equations of motion even *finite* momentum, $q \neq 0$. The last step leads the sought-after adiabatic modes. Inspection of the equations of motion that become trivial at $\mathbf{k} = 0$, (5.33)-(5.38), shows that this mode does not extend to finite momentum. But we can "improve" this mode to make it extendable to finite momentum. I skip this calculation here because it is conceptually the same as in the previous section. Details are given in [10]. The result is that the right diff is

$$\epsilon^{\mu} = \{c/\dot{\bar{\Phi}}, c\,\lambda(t)x^i\} \quad \text{with} \quad \lambda(t) = C_1 - \int^t \mathrm{d}t' \left(\frac{\dot{H}}{\dot{\bar{\Phi}}} + (\Theta - H)\frac{\ddot{\bar{\Phi}}}{\dot{\bar{\Phi}}^2} \right), \tag{5.57}$$

where Θ depends on the theory as $\Theta = H + \dot{\bar{\Phi}}^3 \partial_X G / 2M_{\mathrm{Pl}}^2$.

4. Finally, we use again the general formula (5.50) to find a new symmetry

$$\mathcal{R}_{\mathbf{k}} \to \mathcal{R}_{\mathbf{k}} - c\,\lambda(t)\left((2\pi)^3\delta^3(\mathbf{k}) - (3 + \mathbf{k}\cdot\partial_{\mathbf{k}})\,\mathcal{R}_{\mathbf{k}}\right) - \frac{c}{\dot{\bar{\Phi}}}\left(H(2\pi)^3\delta^3(\mathbf{k}) + \dot{\mathcal{R}}_{\mathbf{k}}\right). \tag{5.58}$$

This differs from the previous symmetry we found both because it involves a time-diff as well and so $\dot{\mathcal{R}}$ appears, and because now λ is not constant. For tensors there are no new symmetries.

Let me stress that these symmetries apply to *any* FLRW solution, not just quasi de Sitter as relevant for inflation. A complementary perspective on shift-symmetric cosmologies is provided by the Effective Field Theory approach of [35, 36] specified to this case. This is beyond the scope of this review, but see [37] for a full treatment.

6 Background-wave method

The background-wave method is the most intuitive and physically transparent was to derive the soft theorems associated to the non-linearly realized symmetries we derived in the previous section. The main idea is to separate perturbations into long, $\mathcal{O}_L$ and short $\mathcal{O}_S$ wavelengths according to

$$\mathcal{O}_L(\mathbf{q}) \equiv \Theta(\bar{k} - q)\mathcal{O}(\mathbf{q}) \qquad \mathcal{O}_S(\mathbf{q}) \equiv \Theta(q - \bar{k})\mathcal{O}(q) = \mathcal{O}(\mathbf{q}) - \mathcal{O}_L(\mathbf{q}), \tag{6.1}$$

for some reference comoving scale $\bar{k}$ and for Θ the Heaviside theta. Then one manipulates the path integral representation of the correlator to separate the measure into long and short wavelength modes

$$\langle \mathcal{O}_L(\mathbf{q})\mathcal{O}(k_1)\ldots\mathcal{O}(\mathbf{k}_n)\rangle = \int [D\mathcal{O}]\mathcal{O}(\mathbf{q})\mathcal{O}(k_1)\ldots\mathcal{O}(\mathbf{k}_n)e^{iS} \tag{6.2}$$

$$= \int [D\mathcal{O}_L D\mathcal{O}_S]\mathcal{O}(\mathbf{q})\mathcal{O}(k_1)\ldots\mathcal{O}(\mathbf{k}_n)e^{iS} \tag{6.3}$$

$$= \int [D\mathcal{O}]\mathcal{O}(\mathbf{q})\mathcal{O}(k_1)\ldots\mathcal{O}(\mathbf{k}_n)e^{iS} \tag{6.4}$$

$$\langle \mathcal{O}(\mathbf{q})\langle \mathcal{O}(k_1)\ldots\mathcal{O}(\mathbf{k}_n)\rangle_{\mathcal{O}(\mathbf{q}')}\rangle, \tag{6.5}$$

where these objects are most easily defined in the formalism

$$\langle \mathcal{O}(\mathbf{q})\mathcal{O}(k_1)\dots\mathcal{O}(\mathbf{k}_n)\rangle \tag{6.6}$$

One re-absorbs the soft classical perturbation (a.k.a. a "wave") in the classical background that the remaining fields in the correlator live on. In formulae

7 Ward Takahashi identities

Let us now discuss a standard method to derive soft-theorems from non-linearly realized symmetries: Ward-Takahashi (WT) identities. These identities are probably familiar to many reader from the renormalization of QED. Gauge transformations in fact act non-linearly on the four-vector potential $A_\mu \to A_\mu + \partial_\mu \alpha$, for some function of spacetime $\alpha(x)$. The WT identities are the used to proved non-perturbatively that the gauge invariance of the theory is preserved by renormalization. Let us review the main idea. The starting point is indeed just the identity

$$i\langle[Q,\mathcal{O}]\rangle = \langle\Delta\mathcal{O}\rangle\,. \tag{7.1}$$

Were $\mathcal{O}$ denotes collectively the product of n operators and its variation is

$$\mathcal{O} = \prod_{a=1}^{n}\mathcal{O}(x_a) \quad \Rightarrow \quad \Delta\mathcal{O} = \sum_{a=1}^{n}\mathcal{O}(x_1)\dots\Delta\mathcal{O}(x_a)\dots\mathcal{O}(x_n)\,. \tag{7.2}$$

If we are interested in late-time cosmological observations, which probe only primordial correlators that are the equal-time products of fields but not of their conjugate momenta, which typically decay with time. Then $\mathcal{O}$ is Hermitian and we can use

$$i\langle[Q,\mathcal{O}]\rangle = 2\mathrm{Im}\langle\mathcal{O}Q\rangle\,. \tag{7.3}$$

The idea is then to compute the left- and right-hand sides of (7.1) separately. I will shortly present an explicit example for single field slow-roll inflation. One finds that the left-hand side of (7.1) depends on the non-linear part of the transformation. The reason is as follows. Using (2.7) and (2.8) to move from the Heisenberg to the interaction picture (see App. B of [6] for a subtlety in this argument)

$$Q_H\,|\Omega\rangle_H = U_I^\dagger Q_I U_I U_{int}^\dagger\,|\Omega\rangle_I = U_I^\dagger Q_I\,|\Omega\rangle\,. \tag{7.4}$$

Here Q_I is the charge written in terms of free fields. The free theory is invariant only under the non-linear shift in the symmetry transformation and so Q_I must contain only that term. On the other hand, the right-hand side of (7.1) depends only on the linear part. The reason is that we are interested in computing *connected* diagrams, namely diagrams that are proportional to an overall delta function. Instead the non-linear shift only contributes to disconnected ones, that are proportional to the product of two or more delta functions. Schematically

$$\Delta\mathcal{O}(\mathbf{x}) = C + \Delta_{lin}\mathcal{O}(\mathbf{x}) \quad \Rightarrow \quad \Delta\mathcal{O}(\mathbf{k}) = C(2\pi)^3\delta_D^3(\mathbf{k}) + \Delta_{lin}\mathcal{O}(\mathbf{k})\,, \tag{7.5}$$

and so

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

where we see that the constant part only contributes to disconnected diagrams.

Operator Product Expansion (OPE) The OPE states that, in the limit $\mathbf{x} \to \mathbf{y}$, the product of two operators $A(\mathbf{x})$ and $B(\mathbf{y})$ can be expanded as an infinity sum of operators $C(\mathbf{y})$ evaluated at a single point (see e.g. Ch. 20 of [43])

$$A(\mathbf{x}) B(\mathbf{y}) = \sum_C f_C^{AB} (\mathbf{x} - \mathbf{y}) C(\mathbf{y}) , \tag{7.6}$$

where f_C^{AB} are numbers that depend on the specific operators and only on the distance $\mathbf{x} - \mathbf{y}$. The word "operator" in OPE reminds us that this relation is valid inside any correlator, not just on some specific states. By naive dimensional analysis, one expects f_C^{AB} to scale as $|\mathbf{x} - \mathbf{y}|^{d_C - d_A - d_B}$, where $d_{A,B,C} > 0$ are the mass dimensions of A, B and C. Then, the strongest divergence is obtained for the smallest d_C and so a truncation of the infinite sum to the first lowest dimensional operators has a chance to be accurate. In momentum space we will use the equivalent relation (Exercise)

$$\lim_{q \ll k} A(-\mathbf{k} - \mathbf{q}/2) B(\mathbf{k} - \mathbf{q}/2) = \sum_C f_C^{AB}(\mathbf{k}) C(\mathbf{q}) , \tag{7.7}$$

where $\mathbf{k} \sim |\mathbf{x} - \mathbf{y}|^{-1}$ is the momentum associated with the inverse distance between the two points, which is becoming large, and $\mathbf{q} \sim |\mathbf{x} + \mathbf{y}|^{-1}$ is associated with the inverse average, which is held approximately constant in the OPE limit. The main input needed for using this method is a basis of operators C that are build out of products of fundamental fields and their derivatives and a way to rank them in order of increasing mass dimension. For perturbative theories, the naive or classical mass dimensions of operators, as derived from dimensional analysis of the free part of the theory (e.g. a canonical scalar or vector has dimension of mass, a canonical fermion dimension of mass$^{3/2}$ and so on) usually provides sufficient guidance to achieve this.

7.1 Slow-roll inflation: Ward-Takahashi identities

As an illustration, let us derive the WT identities for the symmetries in (5.52) and (5.54). This presentation parallels in spirit that of [6], but instead of the Schrödinger picture of "wave functionals" I use the more standard interaction picture. We start by defining the charge

$$Q = \frac{1}{2} \int d^3x \{\Pi, \Delta \mathcal{R}\} + \{\Pi_\gamma^{ij}, \Delta \gamma_{ij}\} , \tag{7.8}$$

where Π and Π_γ are the conjugate momenta of $\mathcal{R}$ and γ respectively

$$[\mathcal{R}(t, \mathbf{x}), \Pi(t, \mathbf{y})] = i \delta_D^3(\mathbf{x} - \mathbf{y}) \qquad [\gamma_{ij}(t, \mathbf{x}), \Pi_\gamma^{ij}(t, \mathbf{y})] = i \delta_D^3(\mathbf{x} - \mathbf{y}) , \tag{7.9}$$

and the parenthesis indicated the anti-commutator and are used to make Q hermitian. Also, we recall the explicit form of the symmetry transformation was given in (5.54). Since those transformations are valid for arbitrary ω_{ij}, which eventually cancels in the final soft theorem, in the following I'll set $-\omega_{ii}/3 = 1$ and $-2\omega_{[ij]} = 1$ to simplify the notation. So our non-linearly realized symmetries are

$$\Delta\mathcal{R}(\mathbf{k}) = (2\pi)^3\delta^3(\mathbf{k}) - (3 + \mathbf{k}\cdot\partial_{\mathbf{k}})\,\mathcal{R}(\mathbf{k})\,, \tag{7.10}$$

$$\Delta\gamma_{ij}(\mathbf{k}) = (2\pi)^3\delta^3(\mathbf{k}) - ..,\tag{7.11}$$

For example for the part of $\Delta\mathcal{R}$ that is linear in $\mathcal{R}$ in real space

$$\{\Pi,\mathcal{R}\}^\dagger = \{\mathcal{R}^\dagger,\Pi^\dagger\} = \{\mathcal{R}^\dagger,\Pi^\dagger\} = \{\mathcal{R},\Pi\} = \{\Pi,\mathcal{R}\} \quad \Rightarrow \quad Q^\dagger = Q\,. \tag{7.12}$$

where I used the symmetry of the anti-commutator in the last step. The expression for Q is chosen so that indeed it generated the transformation upon commutation as in (7.1). I will focus now on the derivation for $\mathcal{R}$ and leave the almost identical calculation for γ_{ij} as an exercise.

The left-hand side As argued in (7.3) to compute the left hand side of the WT identity (7.1) we need the imaginary part of $\langle\mathcal{O}Q\rangle$ and we can compute $Q\,|\Omega\rangle$ in the interaction picture, where the free $\mathcal{R}$ and Π fields are given by

$$\mathcal{R}(\mathbf{x}) = \int\frac{d^3k}{(2\pi)^3}\left[a_{\mathbf{k}}f_k(t)e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}}f^*_k(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right]\,, \tag{7.13}$$

$$\Pi(\mathbf{x}) = \int\frac{d^3k}{(2\pi)^3}\left[a_{\mathbf{k}}g_k(t)e^{i\mathbf{k}\cdot\mathbf{x}} + a^\dagger_{\mathbf{k}}g^*_k(t)e^{-i\mathbf{k}\cdot\mathbf{x}}\right]\,, \tag{7.14}$$

where $f_k(t)$ and $g_k(t)$ are the solutions of the classical linearized Hamilton equations of motion and depend only on $|\mathbf{k}|$ by virtue of isotropy. 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 (7.9) fixes the so-called wronskian

$$f_k g^*_k - f^*_k g_k = i\,. \tag{7.15}$$

We can then massage the equation as

$$Q_I\,|0\rangle = \int d^3x\Pi(x)\,|0\rangle = g^*_0(t)a^\dagger_{\mathbf{0}}\,|0\rangle \tag{7.16}$$

$$= \frac{g^*_0(t)}{f^*_0(t)}f^*_0(t)a^\dagger_{\mathbf{0}}\,|0\rangle = \frac{g^*_0(t)}{f^*_0(t)}\mathcal{R}(\mathbf{0})\,|0\rangle\,, \tag{7.17}$$

where $\mathcal{R}(\mathbf{0}) = \mathcal{R}(\mathbf{k}=\mathbf{0})$ is the Fourier space field and as promised I only used the non-linear part of $\Delta\mathcal{R}$. I can then re-write this in the Heisenberg picture by repeating (7.4) backwards. So we need to compute

$$i\langle[Q,\mathcal{O}]\rangle = 2\mathrm{Im}\left[\langle\mathcal{O}\mathcal{R}(\mathbf{0})\rangle\frac{g^*_0(t)}{f^*_0(t)}\right]\,. \tag{7.18}$$

For concreteness and because it is relevant for observations, let us assume that

$$\mathcal{O} = \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\,, \tag{7.19}$$

where the time argument is $\tau \to 0$ so that $\mathcal{R}$ is approximately constant and I can keep τ implicit. By hermiticity one finds that $\langle \mathcal{O}\mathcal{R}\rangle$ is real

$$\begin{aligned}\langle \mathcal{O}\mathcal{R}(\mathbf{0})\rangle^* &= \langle \mathcal{R}^\dagger(\mathbf{0})\mathcal{O}^\dagger\rangle = \langle \mathcal{R}(\mathbf{0})\mathcal{R}(-\mathbf{k}')\mathcal{R}(-\mathbf{k})\rangle \\ &= \langle \mathcal{R}(\mathbf{0})\mathcal{R}(\mathbf{k}')\mathcal{R}(\mathbf{k})\rangle = \langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\mathcal{R}(\mathbf{0})\rangle = \langle \mathcal{O}\mathcal{R}(\mathbf{0})\rangle\,,\end{aligned} \tag{7.20}$$

where I used that $\mathcal{R}^\dagger(\mathbf{k}) = \mathcal{R}^\dagger(-\mathbf{k})$ which is a consequence of $\mathcal{R}$ being hermitian in real space and that all equal time $\mathcal{R}$ commute with each other. For the other factor in (7.18) we can use the Wronskian condition

$$\mathrm{Im}\,\frac{g_0^*(t)}{f_0^*(t)} = \frac{\mathrm{Im}\,[g_0^*(t)f_0(t)]}{|f_0(t)|^2} = -\frac{i}{2}\frac{[g_0^*(t)f_0(t) - g_0(t)f_0^*(t)]}{|f_0(t)|^2} = \frac{1}{2|f_0(t)|^2}\,, \tag{7.21}$$

where we recognize the $\mathcal{R}$ power spectrum at zero momentum

$$\langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\rangle = (2\pi)^3\delta_D^3\left(\mathbf{k}+\mathbf{k}'\right)|f_k(t)|^2\,. \tag{7.22}$$

Our calculation of the left-hand side is complete

$$i\langle[Q, \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')]\rangle = \frac{1}{P_{\mathcal{R}}(0)}\langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\mathcal{R}(\mathbf{0})\rangle\,. \tag{7.23}$$

The right-hand side On the right-hand side of the (7.1), I can just use the linear part of the transformation in (7.10) and find

$$\langle \Delta\mathcal{O}\rangle = -\left(3 + \mathbf{k}\cdot\partial_{\mathbf{k}} + 3 + \mathbf{k}'\cdot\partial_{\mathbf{k}'}\right)\langle \mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\rangle\,. \tag{7.24}$$

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

$$\langle \Delta\mathcal{O}\rangle' = -\left(3 + k\partial_k\right)P_{\mathcal{R}}(k) = (1-n_s)P_{\mathcal{R}}(k)\,. \tag{7.25}$$

We conclude with the WT identity in its final form

$$\lim_{\mathbf{q}\to 0}\langle \mathcal{R}(\mathbf{q})\mathcal{R}(\mathbf{k})\mathcal{R}(\mathbf{k}')\rangle' = (1-n_s)P_{\mathcal{R}}(k)P_{\mathcal{R}}(q)\,. \tag{7.26}$$

A few comments are in order:

- 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. In comparing with observations several non-linear effects due to the late evolution need to be computed and affect the squeezed limit, as e.g. in [44–49]. The exception is the cross correlation of CMB temperature and CMB spectral distortions [50], which has no such squeezed late time non-Gaussianity [51].

- 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 [52]

$$\lim_{\mathbf{q}_p \to 0} \langle \mathcal{R}(\mathbf{q}_p)\mathcal{R}(\mathbf{k}_p)\mathcal{R}(\mathbf{k}'_p)\rangle' = \mathcal{O}(q^2)\,. \tag{7.27}$$

  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 [53–55]. The $\mathcal{O}(q^2)$ term is model dependent but has a lower bound of order η [56, 57].

- Many other soft theorems exist, also with soft tensor and vectors [1, 6, 9, 11].

- There is a loophole in the derivation, which I will discuss next.

7.2 Ultra-slow-roll inflation: OPE derivation

It was pointed out in [58, 59] that there are models of single-field inflation in which the soft theorem (7.26) is violated. The violation is actually small in canonical models, but can be large in presence of a small speed of sound [60] The loophole is that in these models $\mathcal{R}$ grows on superHubble scales and so it does not asymptote the adiabatic solution (see e.g. [11, 61]). In the general case this is the end of the story, but in the simplest model, namely Ultra-Slow-Roll inflation [62], there is an additional shift symmetry in the problem. To leading order in derivatives, the Lagrangian that of (5.55). In Sec 5.3 we saw that this new symmetry generates new adiabatic modes. Here we will see that the associated symmetries lead to a new, corrected soft theorem, which is satisfied by the explicit calculation. This discussion is based on [10]. Instead of using the WT identities as in the previous section, I will instead use the OPE.
If we are only interested in scalar modes $\mathcal{R}$, the leading order terms in the OPE are

$$\mathcal{R}_{\mathbf{k}-\frac{1}{2}\mathbf{q}}\mathcal{R}_{-\mathbf{k}-\frac{1}{2}\mathbf{q}} \xrightarrow{\mathbf{q}\to 0} h(k)(2\pi)^3\delta^3(\mathbf{q}) + f(k)\mathcal{R}_{-\mathbf{q}} + g(k)\dot{\mathcal{R}}_{-\mathbf{q}} \tag{7.28}$$

$$+ f^i(\mathbf{k})\partial_i\mathcal{R}_{-\mathbf{q}} + g^i(\mathbf{k})\partial_i\dot{\mathcal{R}}_{-\mathbf{q}} + \mathcal{O}(q^2\mathcal{R}, \mathcal{R}^2)\,. \tag{7.29}$$

Notice that the $\mathbf{q}$ and $\mathbf{k}$ dependence has been completely factored between the operators $\{\mathcal{R}, \dot{\mathcal{R}}, \dots\}$ and the coefficients $\{f, g, \dots\}$. The normalization of the first, constant term is easily determined by taking the expectation value on each side and using $\langle\mathcal{R}\rangle = 0$, then

$$h(k) = P_{\mathcal{R}}(k)\,. \tag{7.30}$$

The coefficients f_i and g_i vanish. To see this notice that by rotational invariance one must have

$$f_i(k) = F(k)k^i\,, \tag{7.31}$$

which is an odd function under $\mathbf{k} \to -\mathbf{k}$. But the left-hand side of (7.29) is even under this parity transformation (recall that $\mathcal{R}$ are at equal time and so commute) and so $f^i = 0$.

The coefficients f and g are fixed by the symmetries we derived previously

$$\mathcal{R}(\mathbf{k}) \to \mathcal{R}(\mathbf{k}) - \lambda(2\pi)^3\delta^3(\mathbf{k}) + \lambda\,(3 + \mathbf{k}\cdot\partial_{\mathbf{k}})\,\mathcal{R}(\mathbf{k})\,, \tag{7.32}$$

$$\mathcal{R}_{\mathbf{k}} \to \mathcal{R}_{\mathbf{k}} - \lambda(t)\left((2\pi)^3\delta^3(\mathbf{k}) - (3 + \mathbf{k}\cdot\partial_{\mathbf{k}})\,\mathcal{R}_{\mathbf{k}}\right) - \frac{1}{\dot{\bar{\Phi}}}\left(H(2\pi)^3\delta^3(\mathbf{k}) + \dot{\mathcal{R}}_{\mathbf{k}}\right)\,. \tag{7.33}$$

Let us call $Q_{1,2}$ the respective charges. By taking the commutator of each side of the OPE (7.29) with Q_1 and then the expectation value we find

$$\langle\left[iQ_1, \mathcal{R}_{-\mathbf{k}-\frac{1}{2}\mathbf{q}}\mathcal{R}_{\mathbf{k}-\frac{1}{2}\mathbf{q}}\right]\rangle = (1-n_s)P(k)(2\pi)^3\delta^3(\mathbf{q})\,, \tag{7.34}$$

$$\langle\left[iQ_1, f(k)\mathcal{R}_{-\mathbf{q}} + g(k)\dot{\mathcal{R}}_{-\mathbf{q}}\right]\rangle = \langle f(k)\Delta\mathcal{R}_{-\mathbf{q}} + g(k)\partial_t\Delta\mathcal{R}_{-\mathbf{q}}\rangle = f(k)(2\pi)^3\delta^3(\mathbf{q})\,, \tag{7.35}$$

where only the linear part, $\mathcal{O}(\mathcal{R}^1)$ contributed in the first equation and only the non-linear, constant part $\mathcal{O}(\mathcal{R}^0)$ in the second. Since these two expression need to be equal in the limit $\mathbf{q}\to 0$ in virtue of the OPE relation, we conclude that

$$f(k) = (1-n_s)P(k)\;. \tag{7.36}$$

Plugging this result back into the OPE and performing the same procedure again using Q_2 instead of Q_1 fixes g to

$$g(k) = \frac{1}{\Theta}\frac{\dot{\bar{\Phi}}}{\ddot{\bar{\Phi}}}\left[(1-n_s)P(k)H - \dot{P}(k)\right]\;. \tag{7.37}$$

Now that we know all leading order coefficients appearing in the OPE, all we have to do is to multiply the OPE equation (7.29) by $\mathcal{R}_{\mathbf{q}}$ and take the expectation value. The result is

$$\lim_{\mathbf{q}\to 0}\langle\mathcal{R}_{\mathbf{q}}\mathcal{R}_{\mathbf{k}-\frac{1}{2}\mathbf{q}}\mathcal{R}_{-\mathbf{k}-\frac{1}{2}\mathbf{q}}\rangle' = -\frac{\dot{\bar{\Phi}}\dot{P}_{\mathcal{R}}(q)}{2\ddot{\bar{\Phi}}\Theta}\left[(n_s-1)HP_{\mathcal{R}}(k) + \dot{P}_{\mathcal{R}}(k)\right] + (1-n_s)P(k)P(q)\;, \tag{7.38}$$

where I used that

$$\langle\mathcal{R}_{\mathbf{q}}\mathcal{R}_{\mathbf{k}-\frac{1}{2}\mathbf{q}}\mathcal{R}_{-\mathbf{k}-\frac{1}{2}\mathbf{q}}\rangle' = \frac{1}{2}\langle\{\mathcal{R}_{\mathbf{q}}, \mathcal{R}_{\mathbf{k}-\frac{1}{2}\mathbf{q}}\mathcal{R}_{-\mathbf{k}-\frac{1}{2}\mathbf{q}}\}\rangle\,. \tag{7.39}$$

A few comments are in order:

- The new soft theorem (7.38) reduces to Maldacena's consistency relation (7.26) in slow roll inflation because then $\dot{P}_{\mathcal{R}}(q) \simeq q^2\tau^2 P_{\mathcal{R}}(q)$ and so the first term is subleading as $\mathbf{q}\to 0$ and it is of the same order as terms we neglected in the OPE. This is reassuring.

- In USR inflation instead the first term is actually leading and the second subleading. To see this, let us consider the simplest model

$$\mathcal{L} = -\frac{M_{\rm Pl}^2}{2}R - \frac{1}{2}\partial_\mu\phi\partial^\mu\phi\,. \tag{7.40}$$

i.e. (5.55) with P and G the identity function. From the background equations of motion

$$\ddot{\bar{\phi}} + 3H\dot{\bar{\phi}} = 0\,, \mathcal{R} \tag{7.41}$$

we see that $\dot{\bar{\phi}} \propto a^{-3}$ and so

$$\epsilon = -\frac{\dot{H}}{H^2} = \frac{1}{2}\frac{\dot{\bar{\phi}}^2}{M_{\rm Pl}^2 H^2} \propto a^{-6}\,, \tag{7.42}$$

which makes ϵ negligibly small after few efoldings. The power spectrum and bispectrum can be computed in the usual in-in formalism and one finds [58]

$$\langle \mathcal{R}_{\mathbf{k}} \mathcal{R}_{-\mathbf{k}} \rangle' = \frac{H^2}{4M_p^2 \epsilon k^3}\,, \quad \langle \mathcal{R}_{\mathbf{k}_1} \mathcal{R}_{\mathbf{k}_2} \mathcal{R}_{\mathbf{k}_3} \rangle = \frac{3H^4}{16 M_p^4 \epsilon^2} \frac{\sum_i k_i^3}{\prod_i k_i^3} = 3\sum_{j>i} P(k_i) P(k_j)\;. \tag{7.43}$$

The power spectrum is just the usual expression except that now ϵ is growing fast with time $P_{\mathcal{R}} \propto a^6$ (equivalently $\mathcal{R} \propto a^3$). The spectral tilt can be computed[16] to be

$$n_s - 1 = -2\epsilon - 6 - \eta \simeq 0. \tag{7.44}$$

Plugging these explicit results into our soft theorem (7.39) we see that it is perfectly satisfied.

- The soft theorem (7.39) is not a consistency relation in the sense that the right hand side depends on $\dot{P}$ during inflation, which is a late time observable. So this soft theorem, unlike Maldacena's result for slow-roll inflation, unfortunately cannot be tested with observations.

- The result is still useful. For example is can be straightforwardly generalized to the n-point correlation function (which have not been computed explicitly)

$$\lim_{\mathbf{q}\to 0} \langle \mathcal{R}_{\mathbf{q}}, \prod_{a=1}^{n} \mathcal{R}_{\mathbf{k}_a - \mathbf{q}/n} \rangle' = -\frac{\dot{\bar{\Phi}}\dot{P}(q)}{2\ddot{\bar{\Phi}}\Theta}\left[H\mathcal{D}^{(n)} B_n(\mathbf{k}_a, t) + \dot{B}_n(\mathbf{k}_a, t)\right] \tag{7.45}$$

$$- P(q)\mathcal{D}^{(n)} B_n(\mathbf{k}_a, t) + \mathcal{O}(q)\,, \tag{7.46}$$

where

$$\mathcal{D}^{(n)} \equiv \left[3(n-1) + \sum_{a=1}^{n} \mathbf{k}_a \cdot \partial_{\mathbf{k}_a}\right]\,, \tag{7.47}$$

$$B_n(\mathbf{k}_a, t) \equiv \langle \prod_{a=1}^{n} \mathcal{R}_{\mathbf{k}_a, t} \rangle\,. \tag{7.48}$$

[16] To get this result you need to evolve the power spectrum until some fixed, k-independent time τ_* when USR stops and a phase of slow-roll inflation starts. Compute n_s at horizon exit gives the wrong result because $P_{\mathcal{R}}$ continues to evolve after that.

8 Conclusions

References

[1] J. M. Maldacena, "Non-Gaussian features of primordial fluctuations in single field inflationary models," *JHEP* **05** (2003) 013, arXiv:astro-ph/0210603 [astro-ph].

[2] V. Assassi, D. Baumann, and D. Green, "On Soft Limits of Inflationary Correlation Functions," *JCAP* **1211** (2012) 047, arXiv:1204.4207 [hep-th].

[3] V. Assassi, D. Baumann, and D. Green, "Symmetries and Loops in Inflation," *JHEP* **02** (2013) 151, arXiv:1210.7792 [hep-th].

[4] L. Senatore and M. Zaldarriaga, "A Note on the Consistency Condition of Primordial Fluctuations," *JCAP* **1208** (2012) 001, arXiv:1203.6884 [astro-ph.CO].

[5] A. Kehagias and A. Riotto, "Operator Product Expansion of Inflationary Correlators and Conformal Symmetry of de Sitter," *Nucl. Phys.* **B864** (2012) 492–529, arXiv:1205.1523 [hep-th].

[6] K. Hinterbichler, L. Hui, and J. Khoury, "An Infinite Set of Ward Identities for Adiabatic Modes in Cosmology," *JCAP* **1401** (2014) 039, arXiv:1304.5527 [hep-th].

[7] M. Mirbabayi and M. Zaldarriaga, "Double Soft Limits of Cosmological Correlations," *JCAP* **1503** no. 03, (2015) 025, arXiv:1409.6317 [hep-th].

[8] A. Kehagias, A. M. Dizgah, J. Noreña, H. Perrier, and A. Riotto, "A Consistency Relation for the Observed Galaxy Bispectrum and the Local non-Gaussianity from Relativistic Corrections," *JCAP* **1508** no. 08, (2015) 018, arXiv:1503.04467 [astro-ph.CO].

[9] E. Pajer and S. Jazayeri, "Systematics of Adiabatic Modes: Flat Universes," *JCAP* **1803** no. 03, (2018) 013, arXiv:1710.02177 [astro-ph.CO].

[10] B. Finelli, G. Goon, E. Pajer, and L. Santoni, "Soft Theorems For Shift-Symmetric Cosmologies," *Phys. Rev.* **D97** no. 6, (2018) 063531, arXiv:1711.03737 [hep-th].

[11] L. Hui, A. Joyce, and S. S. C. Wong, "Inflationary soft theorems revisited: a generalized consistency relation," arXiv:1811.05951 [hep-th].

[12] S. Weinberg, *The Quantum theory of fields. Vol. 1: Foundations*. Cambridge University Press, 2005.

[13] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014. http://www.cambridge.org/us/academic/subjects/physics/theoretical-physics-and-mathematical-physics/quantum-field-theory-and-standard-model.

[14] D. Marolf, I. A. Morrison, and M. Srednicki, "Perturbative S-matrix for massive scalar fields in global de Sitter space," *Class. Quant. Grav.* **30** (2013) 155023, arXiv:1209.6039 [hep-th].

[15] X. Chen, "Primordial Non-Gaussianities from Inflation Models," *Adv. Astron.* **2010** (2010) 638979, arXiv:1002.1416 [astro-ph.CO].

[16] T. Brauner, "Spontaneous Symmetry Breaking and Nambu-Goldstone Bosons in Quantum Many-Body Systems," *Symmetry* **2** (2010) 609–657, arXiv:1001.5212 [hep-th].

[17] E. Fabri and L. E. Picasso, "Self-conjugate multiplets, tcp and locality," *Nuovo Cim.* **A60** (1969) 481–493.

[18] S. R. Coleman and J. Mandula, "All Possible Symmetries of the S Matrix," *Phys. Rev.* **159** (1967) 1251–1256.

[19] K. Schalm, G. Shiu, and T. van der Aalst, "Consistency condition for inflation from (broken) conformal symmetry," *JCAP* **1303** (2013) 005, arXiv:1211.2157 [hep-th].

[20] P. Creminelli, "Conformal invariance of scalar perturbations in inflation," *Phys. Rev.* **D85** (2012) 041302, arXiv:1108.0874 [hep-th].

[21] N. Arkani-Hamed and J. Maldacena, "Cosmological Collider Physics," arXiv:1503.08043 [hep-th].

[22] E. Pajer, G. L. Pimentel, and J. V. S. Van Wijck, "The Conformal Limit of Inflation in the Era of CMB Polarimetry," *JCAP* **1706** no. 06, (2017) 009, arXiv:1609.06993 [hep-th].

[23] S. Rychkov, *EPFL Lectures on Conformal Field Theory in D¿= 3 Dimensions*. SpringerBriefs in Physics. 2016. arXiv:1601.05000 [hep-th].

[24] D. Simmons-Duffin, "The Conformal Bootstrap," in *Proceedings, Theoretical Advanced Study Institute in Elementary Particle Physics: New Frontiers in Fields and Strings (TASI 2015): Boulder, CO, USA, June 1-26, 2015*, pp. 1–74. 2017. arXiv:1602.07982 [hep-th].

[25] H. Osborn, *Lectures on Conformal Field Theories in more than two dimensions*. 2010. http://www.damtp.cam.ac.uk/user/ho/CFTNotes.pdf.

[26] A. Bzowski, P. McFadden, and K. Skenderis, "Holographic predictions for cosmological 3-point functions," *JHEP* **03** (2012) 091, arXiv:1112.1967 [hep-th].

[27] A. Bzowski, P. McFadden, and K. Skenderis, "Implications of conformal invariance in momentum space," *JHEP* **03** (2014) 111, arXiv:1304.7760 [hep-th].

[28] S. Weinberg, "Adiabatic modes in cosmology," *Phys. Rev.* **D67** (2003) 123504, arXiv:astro-ph/0302326 [astro-ph].

[29] K. Hinterbichler, L. Hui, and J. Khoury, "Conformal Symmetries of Adiabatic Modes in Cosmology," *JCAP* **1208** (2012) 017, arXiv:1203.6351 [hep-th].

[30] S. Weinberg, *Cosmology*. 2008. http://www.oup.com/uk/catalogue/?ci=9780198526827.

[31] D. H. Lyth, K. A. Malik, and M. Sasaki, "A General proof of the conservation of the curvature perturbation," *JCAP* **0505** (2005) 004, arXiv:astro-ph/0411220 [astro-ph].

[32] D. S. Salopek and J. R. Bond, "Nonlinear evolution of long wavelength metric fluctuations in inflationary models," *Phys. Rev.* **D42** (1990) 3936–3962.

[33] **Planck** Collaboration, P. A. R. Ade *et al.*, "Planck 2015 results. XIII. Cosmological parameters," arXiv:1502.01589 [astro-ph.CO].

[34] P. Creminelli, J. Norena, and M. Simonovic, "Conformal consistency relations for single-field inflation," *JCAP* **1207** (2012) 052, arXiv:1203.4595 [hep-th].

[35] P. Creminelli, M. A. Luty, A. Nicolis, and L. Senatore, "Starting the Universe: Stable Violation of the Null Energy Condition and Non-standard Cosmologies," *JHEP* **12** (2006) 080, arXiv:hep-th/0606090 [hep-th].

[36] C. Cheung, P. Creminelli, A. L. Fitzpatrick, J. Kaplan, and L. Senatore, "The Effective Field Theory of Inflation," *JHEP* **03** (2008) 014, arXiv:0709.0293 [hep-th].

[37] B. Finelli, G. Goon, E. Pajer, and L. Santoni, "The Effective Theory of Shift-Symmetric Cosmologies," arXiv:1802.01580 [hep-th].

[38] S. Endlich, A. Nicolis, and J. Wang, "Solid Inflation," *JCAP* **1310** (2013) 011, arXiv:1210.0569 [hep-th].

[39] A. Gruzinov, "Elastic inflation," *Phys. Rev.* **D70** (2004) 063518, arXiv:astro-ph/0404548 [astro-ph].

[40] G. L. Pimentel, "Inflationary Consistency Conditions from a Wavefunctional Perspective," *JHEP* **02** (2014) 124, arXiv:1309.1793 [hep-th].

[41] L. Berezhiani and J. Khoury, "Slavnov-Taylor Identities for Primordial Perturbations," *JCAP* **1402** (2014) 003, arXiv:1309.4461 [hep-th].

[42] L. Berezhiani, J. Khoury, and J. Wang, "Non-Trivial Checks of Novel Consistency Relations," *JCAP* **1406** (2014) 056, arXiv:1401.7991 [hep-th].

[43] S. Weinberg, *The quantum theory of fields. Vol. 2: Modern applications.* Cambridge University Press, 2013.

[44] P. Creminelli and M. Zaldarriaga, "CMB 3-point functions generated by non-linearities at recombination," *Phys. Rev.* **D70** (2004) 083532, arXiv:astro-ph/0405428 [astro-ph].

[45] N. Bartolo, S. Matarrese, and A. Riotto, "CMB Anisotropies at Second Order I," *JCAP* **0606** (2006) 024, arXiv:astro-ph/0604416 [astro-ph].

[46] N. Bartolo, S. Matarrese, and A. Riotto, "CMB Anisotropies at Second-Order. 2. Analytical Approach," *JCAP* **0701** (2007) 019, arXiv:astro-ph/0610110 [astro-ph].

[47] L. Boubekeur, P. Creminelli, G. D'Amico, J. Norena, and F. Vernizzi, "Sachs-Wolfe at second order: the CMB bispectrum on large angular scales," *JCAP* **0908** (2009) 029, arXiv:0906.0980 [astro-ph.CO].

[48] P. Creminelli, C. Pitrou, and F. Vernizzi, "The CMB bispectrum in the squeezed limit," *JCAP* **1111** (2011) 025, arXiv:1109.1822 [astro-ph.CO].

[49] N. Bartolo, S. Matarrese, and A. Riotto, "Non-Gaussianity in the Cosmic Microwave Background Anisotropies at Recombination in the Squeezed limit," *JCAP* **1202** (2012) 017, arXiv:1109.2043 [astro-ph.CO].

[50] E. Pajer and M. Zaldarriaga, "A New Window on Primordial non-Gaussianity," *Phys. Rev. Lett.* **109** (2012) 021302, arXiv:1201.5375 [astro-ph.CO].

[51] G. Cabass, E. Pajer, and D. van der Woude, "Spectral distortion anisotropies from single-field inflation," *JCAP* **1808** no. 08, (2018) 050, arXiv:1805.08775 [astro-ph.CO].

[52] E. Pajer, F. Schmidt, and M. Zaldarriaga, "The Observed Squeezed Limit of Cosmological Three-Point Functions," *Phys. Rev.* **D88** no. 8, (2013) 083502, arXiv:1305.0824 [astro-ph.CO].

[53] T. Baldauf, U. Seljak, L. Senatore, and M. Zaldarriaga, "Galaxy Bias and non-Linear Structure Formation in General Relativity," *JCAP* **1110** (2011) 031, arXiv:1106.5507 [astro-ph.CO].

[54] L. Dai, E. Pajer, and F. Schmidt, "Conformal Fermi Coordinates," *JCAP* **1511** no. 11, (2015) 043, arXiv:1502.02011 [gr-qc].

[55] L. Dai, E. Pajer, and F. Schmidt, "On Separate Universes," *JCAP* **1510** no. 10, (2015) 059, arXiv:1504.00351 [astro-ph.CO].

[56] P. Creminelli, A. Perko, L. Senatore, M. Simonovi?, and G. Trevisan, "The Physical Squeezed Limit: Consistency Relations at Order q^2," *JCAP* **1311** (2013) 015, arXiv:1307.0503 [astro-ph.CO].

[57] G. Cabass, E. Pajer, and F. Schmidt, "How Gaussian can our Universe be?," *JCAP* **1701** no. 01, (2017) 003, arXiv:1612.00033 [hep-th].

[58] M. H. Namjoo, H. Firouzjahi, and M. Sasaki, "Violation of non-Gaussianity consistency relation in a single field inflationary model," *EPL* **101** no. 3, (2013) 39001, arXiv:1210.3692 [astro-ph.CO].

[59] X. Chen, H. Firouzjahi, M. H. Namjoo, and M. Sasaki, "A Single Field Inflation Model with Large Local Non-Gaussianity," *EPL* **102** no. 5, (2013) 59001, arXiv:1301.5699 [hep-th].

[60] Y.-F. Cai, X. Chen, M. H. Namjoo, M. Sasaki, D.-G. Wang, and Z. Wang, "Revisiting non-Gaussianity from non-attractor inflation models," *JCAP* **1805** no. 05, (2018) 012, arXiv:1712.09998 [astro-ph.CO].

[61] M. Akhshik, H. Firouzjahi, and S. Jazayeri, "Cosmological Perturbations and the Weinberg Theorem," *JCAP* **1512** no. 12, (2015) 027, arXiv:1508.03293 [hep-th].

[62] W. H. Kinney, "Horizon crossing and inflation with large eta," *Phys. Rev.* **D72** (2005) 023515, arXiv:gr-qc/0503017 [gr-qc].